Stochastic volatility in mean models with heavy-tailed distributions: A maximum likelihood approach using structured hidden Markov models

Carlos A. Abanto-Valle†, Roland Langrock‡, Ming-Hui Chen* and Michel V. Cardoso†

† Department of Statistics, Federal University of Rio de Janeiro, Caixa Postal 68530, CEP: 21945-970, Rio de Janeiro, Brazil

‡School of Mathematics and Statistics, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, KY169LZ, UK

*Department of Statistics, University of Connecticut, U-4120, Storrs, CT 06269, USA

Abstract

In this article, we introduce a likelihood-based estimation method for the stochastic volatility in mean (SVM) model with scale mixtures of normal (SMN) distributions (Abanto-Valle et al., 2012). Our estimation method is based on the fact that the powerful hidden Markov model (HMM) machinery can be applied in order to evaluate an arbitrarily accurate approximation of the likelihood of an SVM model with SMN distributions. The method is based on the proposal of Langrock et al. (2012) and makes explicit the useful link between HMMs and SVM models with SMN distributions. Likelihood-based estimation of the parameters of stochastic volatility models in general, and SVM models with SMN distributions in particular, is usually regarded as challenging as the likelihood is a high-dimensional multiple integral. However, the HMM approximation, which is very easy to implement, makes numerical maximum of the likelihood feasible and leads to simple formulae for forecast distributions, for computing appropriately defined residuals, and for decoding, i.e. estimating the volatility of the process.

Keywords: feedback effect, non-Gaussian and nonlinear state-space models, scale mixture of normal distributions, Value-at-Risk.

1 Introduction

Over the last two decades, stochastic volatility models have proven to be useful for modeling time-varying variances, mainly in financial applications where policy makers or stockholders are constantly facing decision problems that usually depend on measures of volatility and risk. An attractive feature of the stochastic volatility model is its close relationship to financial economic theories (Melino and Turnbull, 1990) and its ability to capture the main empirical properties often observed in daily series of financial returns (Carnero et al., 2004).

Many empirical studies have shown strong evidence of heavy-tailed conditional mean errors in financial time series; see for example Mandelbrot (1963) and Fama (1965). In the stochastic volatility literature, Liesenfeld and Jung (2000), Chib et al. (2002), Jacquier et al. (2004) and Abanto-Valle et al. (2010), amongst others, have provided consistent evidence that leptokurtic distributions, such as the Student's t, the GED or the SMN distributions, are more adequate to capture this empirical regularity by relaxing the normality assumption in the distribution of the returns.

Furthermore, evidence has been provided that unexpected returns and innovations to the volatility process are negatively correlated (Harvey and Shephard, 1996; Yu, 2005). The two popular theories associated with the negative return-volatility relation are the leverage hypothesis and the volatility feedback hypothesis. Conceptually, the idea of a reward to a risk-averse investor for holding a risky asset appears theoretically sound and intuitively appealing. The theory generally predicts a positive relation between expected stock returns and volatility if investors are risk averse. In other words, a general agreement prevails that the more uncertain the investment, the higher the return expected by the investor. However, empirical studies that attempt to test this important relation yield mixed results. French et al. (1987) found a positive and significant relationship and Theodossiou and Lee (1995) reported a positive but insignificant relationship between stock market volatility and stock returns. Consistent with the asymmetric volatility argument, Nelson (1991) and, more recently, Brandt and Kang (2004), Loudon (2006) and Abanto-Valle et al. (2012) reported evidence of a negative and often significant relationship between volatility and returns. Overall, there appears to be stronger evidence of a negative relationship between unexpected returns and innovations to the volatility process, which French et al. (1987) interpreted as indirect evidence of a positive correlation between the expected risk premium and *ex ante* volatility. This theory, known as feedback volatility, states that bad (good) news lead to decreased (increases) stock prices and increased volatility, therefore determining a further decrease of the price. An alternative explanation for asymmetric volatility where causality runs in the opposite direction is the leverage effect put forward by Black (1976), who asserted that a negative (positive) return shock leads to an increase (decrease) in the company's financial leverage ratio, which has an upward (downward) effect on the volatility of its stock returns. However, French et al. (1987) and Schwert (1989) argued that leverage alone cannot account for the magnitude of the negative relationship. For example, Campbell and Hentschel (1992) found evidence of both volatility feedback and leverage effects, whereas Bekaert and Wu (2000) presented results suggesting that the volatility feedback effect dominates the leverage effect empirically.

Frequently, the volatility of daily stock returns has been estimated with stochastic volatility models, but the results have relied on a extensive pre-modeling of these series to avoid the problem of simultaneous estimation of the mean and variance. Koopman and Uspensky (2002) introduced the stochastic volatility in mean (SVM) model to deal with this problem and the unobserved volatility is incorporated as an explanatory variable in the mean equation of the returns under the normality assumption of the innovations. Abanto-Valle et al. (2012) proposed to enhance the robustness of the specification of the innovation return in SVM models by introducing SMN distributions, referring to this generalization as the SVM-SMN class of models. This rich class contains as proper elements the SVM with normal (SVM-N), Student-t (SVM-T), slash (SVM-S) and the generalised Student-t (SVM-GT) distributions. Abanto-Valle et al. (2012) proposed an efficient Markov Chain Monte Carlo (MCMC) procedure for Bayesian estimation of SVM-SMN models. However, the resulting MCMC algorithm has some undesirable features. The procedure is quite involved, requiring a large amount of computer intensive simulations. In addition, the computational cost increases rapidly with the sample size.

In this paper, we apply an alternative estimation method, using an approximation of the likelihood function via hidden Markov models (HMMs). The key idea, the use of iterated numerical integration, was introduced by Kitagawa (1987). In the context of stochastic volatility models, it was applied by Fridman and Harris (1998), by Bartolucci and De Luca (2001; 2003), and by Clements et al. (2006), although none of these papers makes explicit the link between stochastic volatility models and HMMs. The method involves an approximation to the SVM likelihood that can be made arbitrarily accurate, and that is competitive in terms of computational effort, due to the powerful HMM forward algorithm becoming applicable. Further advantages of the HMM formulation of SVM-SMN models are that simple explicit formulae exist for the residuals and the forecast distributions, and that estimates of the latent log-volatility process can be obtained by using the Viterbi algorithm.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction to SMN distributions. Section 3 outlines the general class of the SVM–SMN models as well the maximum likelihood based estimation procedure using HMMs methods. In Section 4 we perform a simulation study in order to verify frequentist properties of the likelihood estimators. Section 5 is devoted to the application and model comparison among particular members of the SVM-SMN models using international market indexes. Finally, some concluding remarks and suggestions for future developments are given in Section 6.

2 Scale mixture of normal distributions

A random variable Y belongs to the SMN family if it can be expressed as

$$Y \;=\; \mu + \kappa(\lambda)^{1/2}X, \tag{1}$$

where μ is a location parameter, $X \sim \mathcal{N}(0, \sigma^2)$, λ is a positive mixing random variable with cumulative distribution function (*cdf*) $H(. \mid \boldsymbol{\nu})$ and probability density function *pdf* $h(.|\boldsymbol{\nu})$, $\boldsymbol{\nu}$ is a scalar or parameter vector indexing the distribution of λ and $\kappa(.)$ is a positive weight function. As in Lange and Sinsheimer (1993) and Choy et al. (2008), we restrict our attention to the case in that $\kappa(\lambda) = 1/\lambda$. Given λ, we have $Y|\lambda \sim \mathcal{N}(\mu, \lambda^{-1}\sigma^2)$, and the *pdf* of Y is given by

$$f_{SMN}(y|\mu, \sigma^2, \nu) \;=\; \int_{-\infty}^{\infty} \phi(y|\mu, \lambda^{-1}\sigma^2)dH(\lambda|\boldsymbol{\nu}), \tag{2}$$

where $\phi(. \mid \mu, \sigma^2)$ denotes the density of the univariate $\mathcal{N}(\mu, \sigma^2)$ distribution. From equation (2), we have that the *cdf* of the SMN distributions is given by

$$\begin{aligned} F_{SMN}(y|\mu,\sigma^2,\nu) &= \int_{-\infty}^{y}\int_{-\infty}^{\infty}\phi(u|\mu,\lambda^{-1}\sigma^2)dH(\lambda|\boldsymbol{\nu})du \\ &= \int_{-\infty}^{\infty}\Phi\left(\frac{\lambda^{1/2}[y-\mu]}{\sigma}\right)dH(\lambda|\boldsymbol{\nu}), \end{aligned} \tag{3}$$

where $\Phi(.)$ is the *cdf* of the standard normal distribution. The notation $Y \sim \mathcal{SMN}(\mu, \sigma^2, \boldsymbol{\nu}; H)$ will be used when Y has *pdf* (2) and *cdf* (3). As was mentioned above, the SMN family constitutes a class of thick-tailed distributions including the normal, the Student-t, the Slash and Generalized Student-t distributions, which are obtained respectively by choosing the mixing variables as: $\lambda = 1$, $\lambda \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$, $\lambda \sim \mathcal{B}e(\nu, 1)$ and $\lambda \sim \mathcal{G}(\frac{\nu_1}{2}, \frac{\nu_2}{2})$, where $\mathcal{G}(.,.)$, $\mathcal{B}e(.,.)$ denote the gamma and beta distributions respectively. In the Generalized Student-t, we set $\nu_1 = \nu$ and $\nu_2 = 1$, respectively, in order to avoid identifiability problems.

3 The heavy-tailed stochastic volatility in mean model

The SVM model with heavy-tails is defined by

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 \mathrm{e}^{h_t} + e^{\frac{h_t}{2}}\epsilon_t, \tag{4a}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma_\eta \eta_t, \tag{4b}$$

where y_t and h_t are, respectively, the compounded return and the log-volatility at time t. We assume that $|\phi| < 1$, i.e. that the log-volatility process is stationary and that the initial value $h_1 \sim \mathcal{N}(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$. The innovations ϵ_t and η_t are assumed to be mutually independent, $\epsilon_t \sim \mathcal{SMN}(0, 1, \nu; H)$ and $\eta_t \sim \mathcal{N}(0, 1)$, respectively. The aim of the SVM-SMN class of models is to simultaneously estimate the *ex-ante* relation between returns and volatility and the volatility feedback effect in the presence of outliers. This class of models includes SVM models with normal distribution (SVM-N) (Koopman and Uspensky, 2002), the Student-t (SVM-T), with slash (SVM-S) and generalized Student-t (SVM-GT) distributions as special cases.

3.1 Model fitting strategy

3.1.1 Likelihood evaluation by iterated numerical integration

To formulate the likelihood, we will require the conditional *pdfs* of the random variables y_t, given h_t and y_{t-1} ($t = 1, \ldots, T$), and of the random variables h_t, given h_{t-1} ($t = 2, \ldots, T$). We denote these by $p(y_t \mid y_{t-1}, h_t)$ and $p(h_t \mid h_{t-1})$, respectively. For any member of the class of SMN distributions, the likelihood of the model defined by equations (4a) and (4b) can then be derived as

$$\begin{aligned}
\mathcal{L} &= \int \ldots \int p(y_1, \ldots, y_T, h_1, \ldots, h_T \mid y_0) dh_T \ldots dh_1 \\
&= \int \ldots \int p(y_1, \ldots, y_T \mid y_0, h_1, \ldots, h_T) p(h_1, \ldots, h_T) dh_T \ldots dh_1 \\
&= \int \ldots \int p(h_1) p(y_1 \mid y_0, h_1) \prod_{t=2}^{T} p(y_t \mid y_{t-1}, h_t) p(h_t \mid h_{t-1}) dh_T \ldots dh_1,
\end{aligned}$$

exploiting the dependence structure of state-space models (SSMs) – of which stochastic volatility in mean models are a special type – in the last step. Hence, the likelihood is a high-order multiple integral that cannot be evaluated directly. Via numerical integration, using a simple rectangular rule based on m equidistant intervals, $B_i = (b_{i-1}, b_i)$, $i = 1, \ldots, m$, with midpoints b_i^* and of length b, the likelihood can be approximated as follows:

$$\begin{aligned}
\mathcal{L} \approx b^T \sum_{i_1=1}^{m} \ldots \sum_{i_T=1}^{m} & p(h_1 = b_{i_1}^*) p(y_1 \mid y_0, h_1 = b_{i_1}^*) \\
& \times \prod_{t=2}^{T} p(h_t = b_{i_t}^* \mid h_{t-1} = b_{i_{t-1}}^*) p(y_t \mid y_{t-1}, h_t = b_{i_t}^*) = \mathcal{L}_{\text{approx}} \, . \qquad (5)
\end{aligned}$$

This approximation can be made arbitrarily accurate by increasing m, provided that the interval (b_0, b_m) covers the essential range of the log-volatility process. We note that this simple midpoint quadrature is by no means the only way in which the integral can be approximated (cf. Langrock et al., 2012).

3.1.2 Fast evaluation of the approximate likelihood using HMM techniques

In the form given in (5), the approximate likelihood can be evaluated directly, but the evaluation will usually be computationally intractable as it involves m^T summands. However, instead of the brute force summation in (5), an efficient recursive scheme can be used to evaluate the approximate likelihood. To see this, note that the numerical integration essentially corresponds to a discretization of the state space, i.e. of the support of the log-volatility process h_t. Therefore, the approximate likelihood given in (5) can be evaluated using the well-developed and powerful HMM machinery, which have exactly the same dependence strucutre as SSMs, but a finite and hence discrete state space (cf. Langrock, 2011; Langrock et al., 2012). A key property of HMM, which we exploit here, is that the likelihood can be evaluated efficiently using the so-called forward algorithm, a recursive scheme which iteratively traverses forward along the time series, updating the likelihood in each step (Zucchini and MacDonald, 2009). For an HMM, applying the forward algorithm results in a convenient, closed-form matrix product expression for the likelihood, and this is exactly what is obtained also for the SVM-SMN class of models:

$$\mathcal{L}_{\text{approx}} = \boldsymbol{\delta}\mathbf{P}(y_1)\boldsymbol{\Gamma}\mathbf{P}(y_2)\boldsymbol{\Gamma}\mathbf{P}(y_3)\cdots\boldsymbol{\Gamma}\mathbf{P}(y_{T-1})\boldsymbol{\Gamma}\mathbf{P}(y_T)\mathbf{1}^\top . \tag{6}$$

Here, the $m \times m$-matrix $\boldsymbol{\Gamma} = \left(\gamma_{ij}\right)$ is the analogue to the transition probability matrix in case of an HMM, defined by $\gamma_{ij} = p(h_t = b_j^* \mid h_{t-1} = b_i^*) \cdot b$, which is an approximation of the probability of the log-volatility process changing from some value in the interval B_i to some value in the interval B_j; this conditional probability is determined by Eq. (4b). The vector $\boldsymbol{\delta}$ is the analogue to the Markov chain initial distribution in case of an HMM, here defined such that δ_i is the density of the $\mathcal{N}(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$-distribution — the stationary distribution of the log-volatility process — multiplied by b. Furthermore, $\mathbf{P}(y_t)$ is an $m \times m$ diagonal matrix with ith diagonal entry $p(y_t \mid y_{t-1}, h_t = b_i^*)$, hence the analogue to the matrix comprising the state-dependent probabilities in case of an HMM; this conditional probability is determined by Eq. (4a). Finally, $\mathbf{1}^\top$ is a column vector of ones. Using the matrix product expression given in (6), the computational effort required to evaluate the approximate likelihood is linear in the number of observations, T, and quadratic in the number of intervals used in the discretization, m. In practice, this means that the likelihood can typically be calculated in a fraction of a second, even for T in the thousands and say $m = 100$, a value which renders the approximation virtually exact (see the simulation experiments below). Furthermore, $\mathcal{L}_{\text{approx}} \to \mathcal{L}$ as $b_m, m \to \infty$ and $b_0 \to -\infty$. It should perhaps be noted here that, although we are using the HMM forward algorithm to evaluate the (approximate) likelihood, the specifications of

$\boldsymbol{\delta}$, $\boldsymbol{\Gamma}$ and $\mathbf{P}(x_t)$ given above do not define exactly an HMM, since in general the row sums of $\boldsymbol{\Gamma}$ will only approximately equal one, and the components of the vector $\boldsymbol{\delta}$ will only approximately sum to one. If desired, this can be remedied by scaling each row of $\boldsymbol{\Gamma}$ and the vector $\boldsymbol{\delta}$ to total 1.

3.1.3 Forecasts and model checking

The HMM forward algorithm can also be used to obtain forecast distributions for SVM models. For example, it is straightforward to find the cumulative distribution function of the one-step-ahead forecast distribution on day $t-1$, i.e., the conditional distribution of the return on day t, given all previous observations. This is given by

$$F(y_t \mid y_{t-1}, y_{t-2}, \ldots, y_0) \approx \sum_{i=1}^{m} \zeta_i F(y_t \mid y_{t-1}, h_t = b_i^*), \tag{7}$$

where ζ_i i is the ith entry of the vector $\boldsymbol{\alpha}_{t-1}/(\boldsymbol{\alpha}_{t-1}\mathbf{1}^\top)$, obtained from the "forward probabilities"

$$\boldsymbol{\alpha}_{t-1} = \boldsymbol{\delta}\mathbf{P}(y_1)\boldsymbol{\Gamma}\mathbf{P}(y_2)\boldsymbol{\Gamma}\mathbf{P}(y_3)\cdots\boldsymbol{\Gamma}\mathbf{P}(y_{t-1}),$$

with $\boldsymbol{\delta}$, $\mathbf{P}(y_k)$ and $\boldsymbol{\Gamma}$ defined as above. The corresponding expression for longer forecast horizons is similar (see Chapter 5 of Zucchini and MacDonald, 2009, for details). The approximation in 7 becomes virtually exact for values of m about 100. A closed-form expression for obtaining state predictions, i.e., volatility predictions, is also available. Furthermore, the forecast distribution given in Eq. (7) can be used in order to perform model checking via residuals (Kim et al., 1998). The one-step-ahead forecast pseudo-residual (or quantile residual) is given by

$$r_t = \Phi^{-1}(F(y_t \mid y_{t-1}, y_{t-2}, \ldots, y_0)), \tag{8}$$

for $t = 1, \ldots, T$, where where $\Phi()$ denotes the cumulative distribution function of the standard normal. For a correctly specified model, the r_t follow a standard normal distribution (Rosenblatt, 1952; Smith, 1985; Kim et al., 1998; Gerlach et al., 1999; Liesenfeld and Richard, 2003, see, e.g.,). Thus, forecast pseudo-residuals can be used to identify extreme values, and the suitability of the model can be checked by using, for example, qq-plots or formal tests for normality.

3.1.4 Decoding

Again building on standard HMM machinery, estimates of the underlying log-volatility values can easily be obtained using the Viterbi algorithm, which is an efficient dynamic programming algorithm for computing the most likely Markov chain state sequence to have given rise to observations stemming from an HMM (see Langrock et al., 2012; Zucchini and MacDonald, 2009, for details)

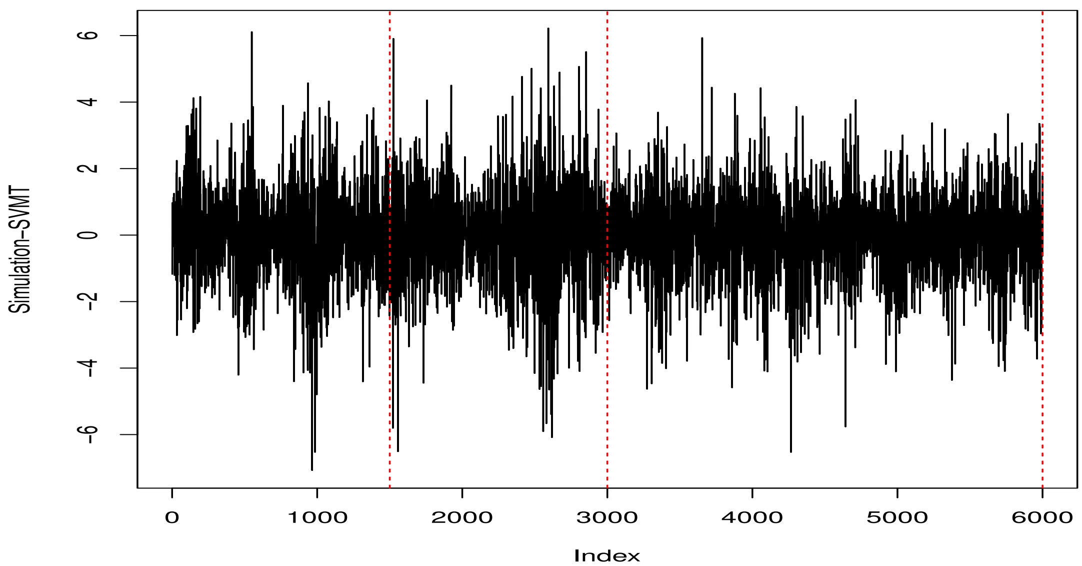

Figure 1: Simulated data set from the SVM-T with $\boldsymbol{\beta} = (0.2, 0.07, -0.18)^\top$, $\mu = 0.1$, $\phi = 0.98$, $\sigma_\eta = 0.1$ and $\nu = 10$ and $y_0 = 0.2$.

4 Simulation Study

In order to assess the performance of the methodology described in Section 3.1, we conducted some simulation experiments. All the calculations were performed using stand-alone code developed by the authors using the Rcpp interface inside R. First, we simulated a data set comprising $T = 6000$ observations from the SVM-T model using $\boldsymbol{\beta} = (0.2, 0.07, -0.18)^\top$, $\mu = 0.1$, $\phi = 0.98$, $\sigma_\eta = 0.1$, $\nu = 10$ and $y_0 = 0.2$, which correspond to typical values found in daily series of returns. Figure 1 shows the resulting data set. In order to investigate the influence the sample size on the accuracy and computing time, we fitted the SVM-T model using $m = 50, 100, 150, 200$, $b_m = -b_0 = 4$ and $T = 1500, 3000, 6000$, respectively. Table 1 reports the results. We observe that the parameter estimates obtained by the HMM method stabilize for values of m around 100, for all the sample sizes considered here. We also investigated the influence of the choice of b_0 and b_m, finding that the estimator performance was not affected much unless these were chosen either much too small (thus not covering the essential support of the log-volatility process, which in the given setting would be the case for example for $b_m = -b_0 = 2$) or much too large (thus leading to a partition of the support into unnecessarily wide intervals in the numerical integration, and hence a poor approximation of the likelihood, for example with $m = 50$ and $b_m = -b_0 = 15$). In practice, it can easily be checked *post-hoc* if the chosen range, specified by b_0 and b_m, is adequate, by investigating the stationary distribution of the fitted log-volatility process. A similar exercise was performed for the SVM-GT and SVM-S models. These results are shown in the supplementary material. Another important fact to be mentioned is the computing time to get the maximum likelihood estimators of the parameters. For example, for the SVM-T with $m = 200$ and 6000 observations our approach using the HMM machinery takes about 17 minutes, instead of almost 4 hours to realize 50000 iterations in order to achieve convergence of an MCMC procedure.

Next, we conducted a second simulation experiment with the objective to study properties of the maximum likelihood estimators of the SVM model's parameters. We generated 300 datasets of size $T = 2500$ from the SVM-T model, specifying $\boldsymbol{\beta} = (0.20, 0.07, -0.18)^\top$, $\phi = 0.98$, $\sigma_\eta = 0.1$, $\mu = 0.10$ and $\nu = 10$. For each generated data set, we fitted the SVM-T model using $m = 50, 100, 150, 200$ and $b_m = -b_0 = 4$, respectively. Table 2 reports the sample mean, the mean relative bias (MRB),

Table 1: SVM-T model, simulated data set: maximum likelihood estimates of the parameters and computing times for the HMM method ($b_m = -b_0 = 4$. True values of the parameters: $\boldsymbol{\beta} = (0.2, 0.07, -0.18)^\top$, $\mu = 0.1$, $\phi = 0.98$, $\sigma_\eta = 0.1$ and $\nu = 10$).

size	m	$\mathcal{L}$	ϕ	σ_η	μ	β_0	β_1	β_2	ν	time
1500	50	2427.46	0.9874	0.0878	0.1716	0.2004	0.0535	-0.1299	10.1613	43.50
			(0.8649,1.1098)	(0.0374,0.1381)	(0.0663,0.2768)	(-0.1696,0.5705)	(0.0441,0.0628)	(-0.1568,-0.103)	(4.7053,15.6173)	
	100	2427.46	0.9878	0.0874	0.161	0.1984	0.0551	-0.1287	9.7333	111.75
			(0.8662,1.1094)	(0.0371,0.1377)	(0.0559,0.2661)	(-0.1719,0.5688)	(0.0446,0.0656)	(-0.1641,-0.0934)	(4.7834,14.6831)	
	150	2427.46	0.9878	0.0874	0.161	0.1984	0.0551	-0.1287	9.7333	192.16
			(0.8662,1.1094)	(0.0371,0.1377)	(0.0559,0.2661)	(-0.1719,0.5688)	(0.0446,0.0656)	(-0.1641,-0.0934)	(4.7834,14.6831)	
	200	2427.46	0.9878	0.0874	0.161	0.1984	0.0551	-0.1287	9.7333	266.80
			(0.8662,1.1094)	(0.0371,0.1377)	(0.0559,0.2661)	(-0.1719,0.5688)	(0.0446,0.0656)	(-0.1641,-0.0934)	(4.7834,14.6831)	
3000	50	4966.28	0.9898	0.085	0.1704	0.1595	0.0809	-0.1154	9.517	96.22
			(0.906,1.0736)	(0.0498,0.1202)	(0.1025,0.2382)	(-0.1585,0.4777)	(0.0746,0.0871)	(-0.1323,-0.0985)	(6.2158,12.8181)	
	100	4966.28	0.9908	0.0805	0.1794	0.1598	0.0819	-0.1153	9.5824	220.43
			(0.9068,1.0748)	(0.0452,0.1157)	(0.1115,0.2472)	(-0.1593,0.479)	(0.0755,0.0884)	(-0.1376,-0.093)	(6.2344,12.9303)	
	150	4966.28	0.9908	0.0805	0.1794	0.1598	0.0819	-0.1153	9.5824	342.55
			(0.9068,1.0748)	(0.0452,0.1157)	(0.1115,0.2472)	(-0.1593,0.479)	(0.0755,0.0884)	(-0.1376,-0.093)	(6.2344,12.9303)	
	200	4966.28	0.9908	0.0805	0.1794	0.1598	0.0819	-0.1153	9.5824	482.76
			(0.9068,1.0748)	(0.0452,0.1157)	(0.1115,0.2472)	(-0.1593,0.479)	(0.0755,0.0884)	(-0.1376,-0.093)	(6.2344,12.9303)	
6000	50	9562.37	0.9868	0.0923	0.0784	0.1645	0.0638	-0.1391	9.2454	192.32
			(0.9261,1.0475)	(0.0674,0.1172)	(0.0225,0.1343)	(-0.0248,0.3539)	(0.058,0.0696)	(-0.156,-0.1223)	(7.0077,11.4832)	
	100	9562.35	0.9865	0.0921	0.0799	0.1647	0.0636	-0.1407	9.1819	376.17
			(0.9257,1.0473)	(0.0672,0.117)	(0.0237,0.1361)	(-0.0179,0.3474)	(0.0571,0.07)	(-0.1614,-0.1201)	(6.9742,11.3895)	
	150	9562.35	0.9865	0.0921	0.0799	0.1647	0.0636	-0.1407	9.1819	621.90
			(0.9257,1.0473)	(0.0672,0.117)	(0.0237,0.1361)	(-0.0179,0.3474)	(0.0571,0.07)	(-0.1614,-0.1201)	(6.9742,11.3895)	
	200	9562.35	0.9865	0.0921	0.0799	0.1647	0.0636	-0.1407	9.1819	989.28
			(0.9257,1.0473)	(0.0672,0.117)	(0.0237,0.1361)	(-0.0179,0.3474)	(0.0571,0.07)	(-0.1614,-0.1201)	(6.9742,11.3895)	

the mean relative absolute deviation (MRAD) and the mean square error (MSE) of the parameter estimates. A similar simulation study for the SVM-GT and SVM-S is available in the supplementary material.

The highest (yet still small) mean relative bias was found for the parameter ν giving the degrees of freedom of the conditional Student-t distribution. The results obtained when using $m = 50$ are similar to those using higher values of m and hence finer approximations of the likelihood. Overall, it can be concluded that the use of HMM machinery to numerically maximize the approximate likelihood function of SVM models leads to very good estimator performance, yet involves only a modest computational effort.

5 Empirical Application

5.1 The Data

In this section, we analyze the indexes from the São Paulo Stock, Mercantile & Futures Exchange, Tokyo Stock Exchange and the New York Stock Exchange. The considered indexes are the IBOVESPA (IBVSP), Nikkei 225 (NIKKEI) and the S&P 500 (SP500) respectively. The period of analysis is from January 5, 1998, until June 30, 2011. All the datasets were downloaded from http://finance.yahoo.com. Stock returns are computed as $y_t = 100 \times (\log P_t - \log P_{t-1})$, where P_t is the (adjusted) closing price on day t. Table 3 reports a summary of descriptive statistics for the series returns. The IBVSP returns show positive skewness and the NIKKEI and SP500 returns negative skewness (NIKKEI and SP500). All the series show kurtosis greater than three, confirming a well-known stylized fact for return series, namely the departure from normality. We analyze the data with the aim of providing robust inference by using the SMN class of distributions. In our analysis, we compare the SVM-N, SVM-T, SVM-GT and SVM-S models for each one of the series described above.

In order to obtain the maximum likelihood estimates (MLEs) of the parameters in the SVM models, we apply the HMM machinery as introduced in Section 3.1. To ensure a good approximation of the estimators, we use $b_m = -b_0 = 4$ and $m = 200$. As before, all the calculations were performed using stand-alone code developed by the authors using the Rcpp interface inside

Table 2: SVM-T model: Simulaton study results based on 300 replicates using the the HMM method ($b_{max} = -b_{min} = 4$).

Parameter	True value	mean	MRB	MARB	MSE
$m = 50$					
ϕ	0.98	0.9807	0.0008	0.0056	0.0001
σ_η	0.10	0.0990	-0.0010	0.1392	0.0003
μ	0.10	0.1068	0.0685	0.8932	0.0130
β_1	0.20	0.1994	-0.0028	0.2607	0.0040
β_2	0.07	0.0693	-0.0097	0.2430	0.0050
β_3	-0.18	-0.1787	-0.0071	0.2612	0.0040
ν	10.0	10.4946	0.0495	0.1612	1.6120
$m = 100$					
ϕ	0.98	0.9812	0.0013	0.0061	0.0001
σ_η	0.10	0.0965	-0.0353	0.1618	0.0004
μ	0.10	0.1038	0.0386	0.9051	0.0130
β_1	0.20	0.1971	-0.0141	0.2553	0.0040
β_2	0.07	0.0695	-0.0070	0.2477	0.0050
β_3	-0.18	-0.1767	-0.0183	0.2597	0.0040
ν	10.0	10.5322	0.0532	0.1661	1.6610
$m = 150$					
ϕ	0.98	0.9812	0.0013	0.0063	0.0001
σ_η	0.10	0.0965	-0.0342	0.1628	0.0004
μ	0.10	0.1039	0.0392	0.9059	0.0130
β_1	0.20	0.1962	-0.0188	0.2530	0.0040
β_2	0.07	0.0693	-0.0090	0.2471	0.0050
β_3	-0.18	-0.1767	-0.0230	0.2573	0.0040
ν	10.0	10.5322	0.0514	0.1621	1.6200
$m = 200$					
ϕ	0.98	0.9812	0.0013	0.0063	0.0001
σ_η	0.10	0.0965	-0.0343	0.1629	0.0004
μ	0.10	0.1039	0.0394	0.9058	0.0130
β_1	0.20	0.1963	-0.0184	0.2535	0.0040
β_2	0.07	0.0694	-0.0090	0.2473	0.0050
β_3	-0.18	-0.1759	-0.0227	0.2577	0.0040
ν	10.0	10.5141	0.0514	0.1620	1.6210

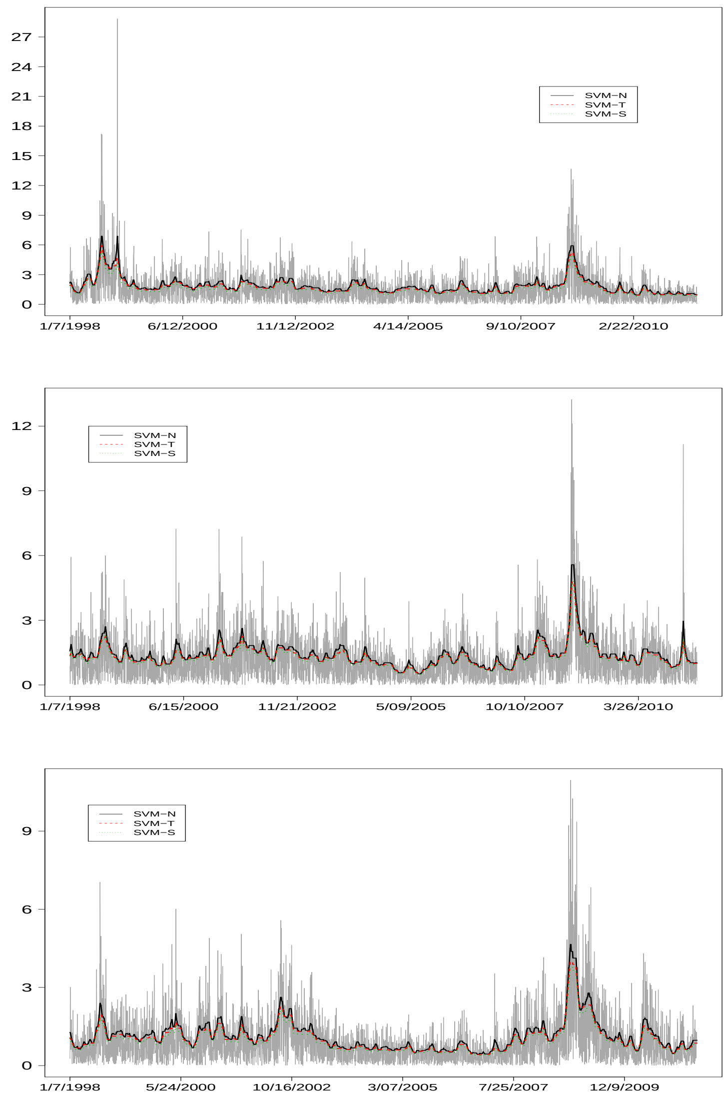

Figure 2: Decoded $e^{\frac{h_t}{2}}$ using the viterbi algorithm: top (IBVSP), center (NIKKEI) and bottom (SP500). The solid line (SVM-N), dotted red line (SVM-T) and dotted green line (SVM-S). The grey line indicates the absolute returns.

Table 3: Summary statistics of the return indexes

Return	size	mean	SD	Minimum	Maximum	Skewness	Kurtosis
IBVSP	3338	0.0531	2.1964	-17.2082	28.8325	0.5748	16.7109
NIKKEI	3310	-0.0127	1.6010	-12.1110	13.2346	-0.3206	9.0383
SP500	3393	0.0089	1.3385	-9.4695	10.9572	-0.1494	10.2474

R package. Tables 4 and 5 show the results for the SVM-N, SVM-T, SVM-GT and SVM-S models for each one of the return indexes series.

For all three series of returns and all four models considered, we find that the MLEs of ϕ are above 0.97, indicating a high persistence of the log-volatility process. For the SVM-GT model, the MLEs and the 95% confidence intervals of ϕ tend to be higher than those obtained using the other three models, while the persistence in the log-volatility process underlying the SVM-N model is smaller than that found using the other three models. The MLE of σ_η is smaller in the SVM-GT than those of the SVM-N, SVM-T and the SVM-S models, indicating that the volatility of the SVM-GT is less variable than those of the other three models. The variability of the log-volatility process is highest in case of the SVM-N model.

For the mean process, we find that the MLEs of β_0 are positive and statistically significant, since the 95% confidence intervals do not contain zero (for all models and series considered). For the IBVSP and the NIKKEI return series, there is an indication that β_1 might not be relevant. In the SP500 case, β_1 is significant. The β_2 parameter, which measures both the *ex ante* relationship between returns and volatility and the volatility feedback effect, is estimated to be negative and is deemed statistically significant, for all models and indexes considered. This result confirms previous results in the literature and indicates that when investors expect higher persistent levels of volatility in the future they require compensation for this in the form of higher expected returns.

The magnitude of the tail fatness is measured by the shape parameter ν in the SVM-T, SVM-GT and SVM-S models. The MLEs of ν are obtained as 13.11, 12.69 and 12.03 in the SVM-T for the IBVSP, NIKKEI and SP500, respectively. The MLEs of ν in the SVM-GT are 5.59, 6.71 and 6.92 for the IBVSP, NIKKEI and SP500 , respectively. Finally, the MLEs of ν in the SVM-S are approximately 4.49, 4.33, 3.53 for the IBVSP, NIKKEI and SP500, respectively. These results

Table 4: Results obtained when fitting the SVM-N model to the three series of index returns considered (using $m = 200$ and $b_{max} = -b_{min} = -4$).

SVM-N									
-	length	llk	ϕ	σ_η	μ	β_0	β_1	β_2	time
IBVSP	3338	-6792.32	0.9756	0.1590	1.1358	0.1971	0.0030	-0.0309	158.21
			(0.9655,0.9857)	(0.1283,0.1898)	(0.911,1.3605)	(0.1075,0.2867)	(-0.032,0.038)	(-0.0568,-0.0051)	
Nikkei	3310	-5812.40	0.9773	0.1586	0.5763	0.1343	-0.0228	-0.0592	162.07
			(0.9672,0.9874)	(0.1297,0.1874)	(0.3354,0.8172)	(0.0672,0.2013)	(-0.058,0.0123)	(-0.0947,-0.0237)	
SP500	3393	-5062.53	0.9806	0.1660	-0.0636	0.1146	-0.0834	-0.0726	147.20
			(0.9728,0.9884)	(0.1353,0.1967)	(-0.3595,0.2324)	(0.0736,0.1556)	(-0.1179,-0.0488)	(-0.1086,-0.0367)	

suggest that the noise in the stock returns is better explained by heavy-tailed distributions.

In Figure 2, we plot the decoded volatility, $e^{\frac{h_t}{2}}$, obtained by applying the Viterbi algorithm for the IBVSP (top), NIKKEI (center) and SP500 (bottom) series. The solid line, dotted red line and and green line indicate the values obtained by the SVM-N, SVM-T and SVM-S, respectively. There are notable differences especially in high volatility periods. This can have a substantial impact, for instance, in the valuation of derivative instruments and several strategic or tactical asset allocation topics. We considered here only the SVM-N, SVM-T and SVM-S models, becuase they are the better models for three indexes considered here.

To compare the different models considered, we calculate the Akaike information criterion (AIC), defined as $\text{AIC} = -2\log L + 2p$, where $\log L$ is the log-likelihood of the fitted model and where p denotes the number of parameters of the model evaluated. From a suite of candidate models, we favor the model which has the smallest AIC. As reported in Table 6, the AIC selects the SVM-S as the best model for the IBVSP series, and the SVM-T for the NIKKEI and SP500 series, respectively.

We now perform an out-of-sample analysis of the forecast performance of the models covered in Table 6. The observation period is January 5, 1988 until September 29, 2014. For each return series the data are now divided into a calibration and a validation sample:

- Calibration sample (in-sample period): from January 5, 1998 until June 30, 2011.
- Validation sample (out-of-sample period): from July 1, 2011 until September 29, 2014.

As a first step, the SVM-N, SVM-S and SVM-T models was fitted to the calibration sample of each series. This was done by using the HMM method with $m = 200$, a value that is large enough to ensure that any anomalies that may occur could not be attributed to inaccuracies in the approximation of the likelihood. Then, for each one of the observations in the validation sample, the (one-step-ahead forecast) pseudo-residual was computed according to equation 8. As described in Section 3.1.3, non-normality of these residuals is an indication of mis-specification of the corresponding model.

The p-values for the Jarque–Bera tests applied to the pseudo-residuals are listed in Table 7. As an example, the qq-plots for the three return series and SVM-N, SVM-S and SVM-T models

Table 5: Results obtained when fitting the SVM-T, SVM-GT and SVM-S models to the three series of index returns considered (using $m = 200$ and $b_{max} = -b_{min} = -4$).

SVM-T										
-	length	llk	ϕ	σ_η	μ	β_0	β_1	β_2	ν	time
IBVSP	3338	-6793.81	0.9880 (0.9796,0.9964)	0.1218 (0.0923,0.1513)	0.8718 (0.4926,1.2509)	0.2096 (0.118,0.3012)	0.10 (-0.0333,0.0353)	-0.0338 (-0.0645,-0.0032)	13.1162 (8.0314,18.201)	366.29
NIKKEI	3310	-5815.45	0.9826 (0.9738,0.9913)	0.1288 (0.0997,0.1580)	0.3725 (0.1057,0.6393)	0.1094 (0.0395,0.1793)	-0.0323 (-0.0666,0.0020)	-0.0622 (-0.1058,-0.0186)	12.6989 (7.3276,18.0701)	366.51
SP500	3393	-5056.31	0.9878 (0.982,0.9936)	0.1267 (0.0997,0.1537)	-0.1257 (-0.4774,0.226)	0.1196 (0.0781,0.1612)	-0.0516 (-0.0848,-0.0183)	-0.0799 (-0.1237,-0.0361)	12.0287 (7.0504,17.007)	377.98
SVM-GT										
IBVSP	3338	-6843.49	0.9852 (0.9776,0.9929)	0.0949 (0.0681,0.1217)	2.5666 (2.298,2.8351)	0.2509 (0.1319,0.37)	0.0146 (-0.0187,0.0478)	-0.0120 (-0.0212,-0.0027)	5.5986 (4.9512,6.2459)	182.37
NIKKEI	3310	-5838.77	0.9908 (0.9838,0.9977)	0.0940 (0.0686,0.1193)	2.2594 (1.854,2.6648)	0.1651 (0.0869,0.2433)	-0.0365 (-0.0698,-0.0031)	-0.0150 (-0.024,-0.0061)	6.7067 (5.5756,7.8378)	206.77
SP500	3393	-5073.02	0.9905 (0.9853,0.9957)	0.1124 (0.0869,0.1379)	1.7885 (1.3388,2.2382)	0.1180 (0.0749,0.1611)	-0.0573 (-0.0896,-0.0251)	-0.0125 (-0.0204,-0.0047)	6.9203 (5.7765,8.0642)	221.63
SVM-S										
IBVSP	3338	-6788.26	0.9843 (0.9754,0.9933)	0.1368 (0.1057,0.1678)	0.9315 (0.6112,1.2518)	0.2061 (0.1153,0.2969)	0.0101 (-0.0246,0.0449)	-0.0455 (-0.0802,-0.0108)	4.4962 (2.3816,6.6107)	1070.63
NIKKEI	3310	-5811.25	0.9808 (0.9710,0.9906)	0.1430 (0.1111,0.1749)	0.3284 (0.0376,0.6193)	0.1344 (0.066,0.2028)	-0.0273 (-0.0624,0.0077)	-0.0765 (-0.1247,-0.0282)	4.3333 (2.0783,6.5884)	922.22
SP500	3393	-5056.46	0.9881 (0.9819,0.9942)	0.1331 (0.1057,0.1606)	-0.3523 (-0.7472,0.0426)	0.1019 (0.0609,0.1429)	-0.0605 (-0.0943,-0.0267)	-0.0741 (-0.1249,-0.0234)	3.4195 (2.2619,4.5772)	1080.35

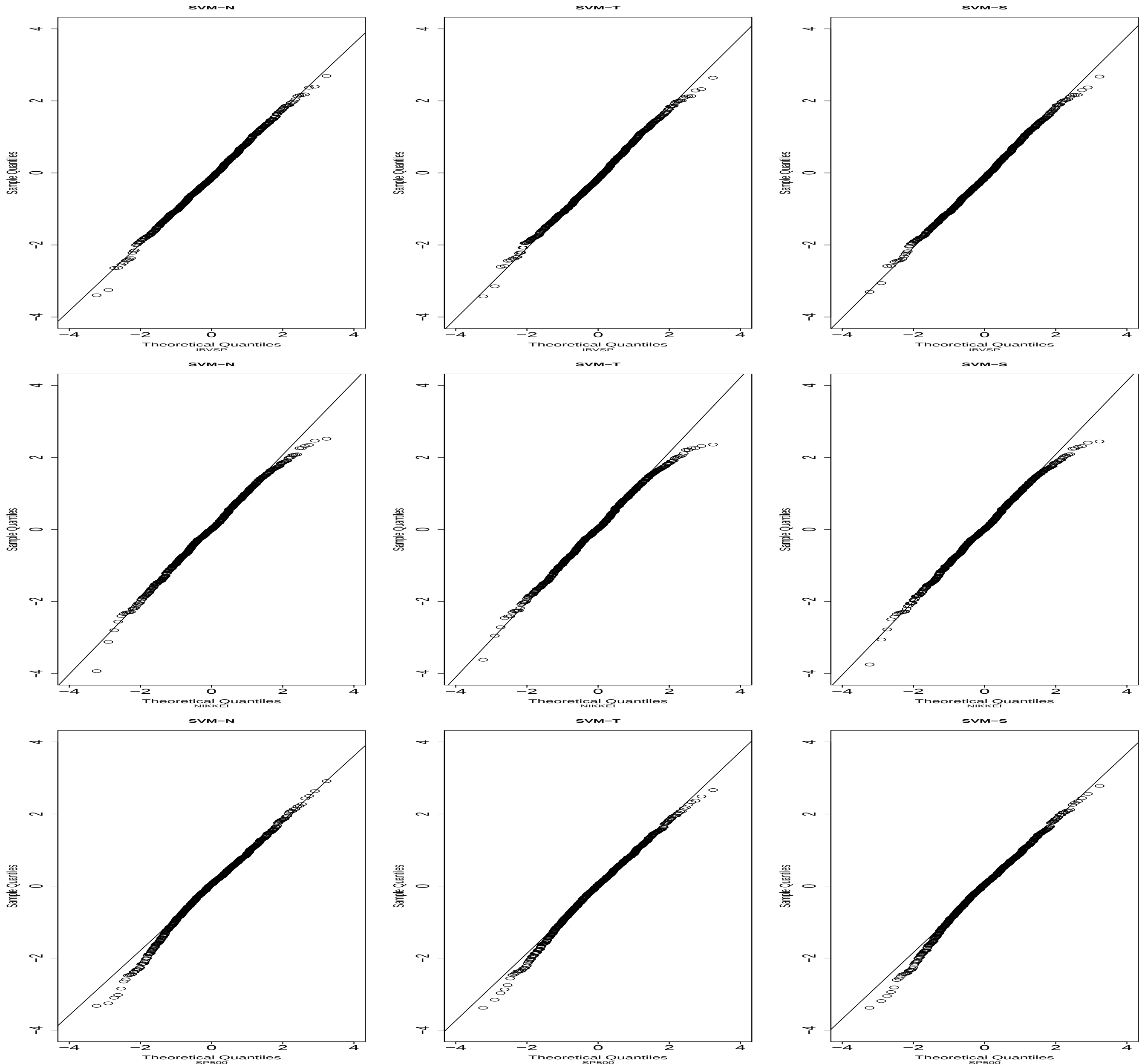

Figure 3: qq-plot of the forecast pseudo-residuals for SVM-N (left), SVM-T (middle) and SVM-S (rigth) and the three returns indexes IBVSP(top), NIKKEI(middle) and SP500 (bottom).

Table 6: Model comparison via AIC; for each series considered, the minimum AIC is highlighted in bold.

		AIC		
Return	SVM-N	SVM-T	SVM-GT	SVM-S
IBVSP	13596.64	13601.62	13700.98	**13590.52**
NIKKEI	11636.80	**11636.40**	11691.54	11636.50
SP500	10137.06	**10126.62**	10160.04	10126.92

Table 7: p-values of Jarque–Bera tests applied to one-step-ahead ahead forecast pseudo-residuals

	SVM-N	SVM-S	SVM-T
IBVSP	0.97	0.51	0.49
NIKKEI	0.11	0.12	0.07
SP500	0.0002	0.0006	0.004

are given in Figure 3. For IBVSP return index the qq-plot indicate a lack of fit in the left tail (SVM-N) and rigth tail (SVM-T and SVM-S). The JB test accept the hypothesis of normality of the residuals of the three models at the 5% and 10% level. For the NIKKEI return index the qq-plot reveals a poor fit in left tail (SVM-N)evels and a similar fit in the right tail. The JB test accept the normalty assumption of the pseudo-residuals at the 10% level for the SVM-N and SVM-S models and reject the SVM-T and at the 5% level the JB test accept normality of the three models. Finally, considering the SP500 returns the qq-plot identifies a poor fit in the left tail for the three models, it is more evidently in the normal case. The JB test confirms these findings and reject the normality assumption of the pseudo-residuals. This apparently mis-specification could be caused by the presence of correlation between the perturbation terms defined by equations (4a) and (4b).

6 Discussion

In this article, we presented an implementation of a maximum likelihood-based estimation approach for the SVM model (Koopman and Uspensky, 2002). The SVM model allows to investigate the dynamic relationship between returns and their time-varying volatility. The Gaussian assumption of the mean innovation was replaced by univariate thick-tailed processes, known as scale mixtures of normal distributions. We studied three specific sub-classes, viz. the Student-t, slash and the generalized Student-t distributions, and compared parameter estimates and model fit with the default normal model. We illustrated our methods through an empirical application of the IBVSP, NIKKEI and SP500 index returns. The AIC was used to asses in-sample fit. According to the AIC, the SVM-S model showed the best fit for the IBVSP series and the SVM-T model for the NIKKEI and SP500 series, respectively. For all indexes and models considered, the β_2 estimate, which measures both the *ex ante* relationship between returns and volatility and the volatility feedback effect, was found to be negative. The results are in line with those of French et al. (1987), who found a similar relationship between unexpected volatility dynamics and returns, and confirm the hypothesis that investors require higher expected returns when unanticipated increases in future volatility are highly persistent. This is consistent with our findings of higher values of ϕ combined with larger negative values for the in-mean parameter.

Our SVM-SMN models showed considerable flexibility to accommodate outliers, however their robustness aspects could be seriously affected by the presence of skewness and heavy-tailedness simultaneously. To deal with this problem, the scale mixtures of skew-normal distributions can be used, or alternatively, the conditional distribution of the retunrs could be modeled nonparametrically (Langrock et al., 2014). A deeper investigation of such modifications is beyond the scope of the present paper, but provides stimulating topics for further research.

Acknowledgements

The research of Carlos A. Abanto-Valle was partially supported by the CNPq-Brazil and Fundo de Amparo à Pesquisa do Estado de Rio de Janeiro (FAPERJ). The research of Michel V. Cardoso was supported by the Comissão de Aperfeiçoamento de Pessoal de Nível Superior de Pessoal (CAPES).

References

Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V., and Enriquez, I. (2010), "Robust Bayesian analysis of heavy-tailed stochastic volatility modelsusing scale mixtures of normal distributions," *Computational Statistics and Data Analysis*, 54, 2883–2898.

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

Bartolucci, F., and De Luca, G. (2001), "Maximum likelihood estimation of a latent variable time-series model," *Applied Stochastic Models in Business and Industry*, 17, 5–17.

Bartolucci, F., and De Luca, G. (2003), "Likelihood-based inference for asymmetric stochastic volatility models," *Computational Statictics and Data Analysis*, 42, 445–449.

Bekaert, G., and Wu, G. (2000), "Asymmetric Volatility and Risk in Equity Markets," *Review of Financial Studies*, 13, 1–42.

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

Brandt, M. W., and Kang, Q. (2004), "Generalized autoregressive conditional heteroskedasticy," *Journal of Financial Economics*, 72, 217–257.

Campbell, J. Y., and Hentschel, L. (1992), "No news is good News, an asymmetric model of changing volatility in stock returns," *Journal of Financial Economics*, 31, 281–318.

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

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., Wan, W. Y., and Chan, C. (2008), "Bayesian Student-t Stochastic Volatility Models via Scale Mixtures," *Advances in Econometrics*, 23, 595–618.

Clements, A. E., Hurn, S., and White, S. I. (2006), "Mixture distribution-based forecasting using stochastic volatility model," *Applied Stochastic Models in Business and Industry*, 22, 547–557.

Fama, E. (1965), "Portfolio analysis in a stable paretian market," *Management Science*, 11, 404–419.

French, K. R., Schert, W. G., and Stambugh, R. F. (1987), "Expected Stock Return and Volatility," *Journal of Financial Economics*, 19, 3–29.

Fridman, M., and Harris, L. (1998), "A maximum likelihood approach for non-Gaussian stochastic volatility models," *Journal of Business and Economic Statistics*, 16, 284–291.

Gerlach, R., Carter, C., and Kohn, R. (1999), "Diagnostics for time series analysis," *Journal of Time Series Analysis*, 20, 309–330.

Harvey, A. C., and Shephard, N. (1996), "The estimation of an asymmetric stochastic volatility model for asset returns," *Journal of Business and Economic Statistics*, 14, 429–434.

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.

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

Kitagawa, G. (1987), "Non-Gaussian state-space modeling of nonstationary time series (with discussion)," *Journal of the American Statistical Association*, 82, 1032–1041.

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

Lange, K. L., and Sinsheimer, J. S. (1993), "Normal/independent distributions and their applications in robust regression," *J. Comput. Graph. Stat*, 2, 175–198.

Langrock, R. (2011), "Some applications of nonlinear and non-Gaussian state-space modelling by means of hidden Markov models," *Journal of Applied Statistics*, 38, 2955–2970.

Langrock, R., MacDonald, I. L., and Zucchini, W. (2012), "Some nonstandard stochastic volatility models and their estimation usingstructured hidden Markov models," *Journal of Empirical Finance*, 19, 147–161.

Langrock, R., Michelot, T., Sohn, A., and Kneib, T. (2014), "Semiparametric stochastic volatility modelling using penalized splines," *ArXiv e-prints*, .
URL: *http://arxiv.org/abs/1308.5836v3*

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

Liesenfeld, R., and Richard, J.-F. (2003), "Univariate and multivariate stochastic volatility models: estimation and diagnostics," *Journal of Empirical Finance*, 10, 505–531.

Loudon, G. F. (2006), "Is the risk-return relation positive?. Further evidence froma stochastic volatility in mean approach," *Applied Financial Economics*, 16, 981–992.

Mandelbrot, B. (1963), "The variation of certain speculative prices," *Journal of Business*, 36, 314–419.

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

Nelson, D. B. (1991), "Leverage, Heavy-Tails and Correlated Jumps in Stochastic Volatility Models," *Econometrica*, 59, 347–370.

Rosenblatt, M. (1952), "Remarks on a multivariate transformation," *Annals of Mathematical Statistics Volume 23*, 23, 470–472.

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

Smith, J. Q. (1985), "Diagnostic checks of non-standard time series models," *Journal of Forecasting*, 4, 283–291.

Theodossiou, P., and Lee, U. (1995), "Relationship between Volatility and Expected Return Across International Stock Markets," *Journal of Business Finance and Accounting*, 22, 289–300.

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

Zucchini, W., and MacDonald, I. L. (2009), *Hidden Markov Models for Time Series: An Introduction Using R*, London: Chapman & Hall.