On the Estimation of Asymmetric Long Memory Stochastic Volatility Models [*]

Omar Abbara [†] Mauricio Zevallos [‡]

March 20, 2023

Abstract

The Asymmetric Long Memory Stochastic Volatility (A-LMSV) model has two attractive features for modeling financial returns: i) the autocorrelation function of the log-variance presents hyperbolic decay, and ii) the two driven random noises that define the model have nonzero correlation. In this work we present a maximum likelihood method for estimating both the parameters and the unobserved components, together with a method for value-at-risk (VaR) forecasting. Our method takes advantage of a state space representation of the model which is written as a dynamic linear model with Markov switching. Then, the likelihood is readily calculated by the Kalman filter. The proposed method is assessed by Monte Carlo experiments and real-life illustrations.

Keywords: value-at-risk, leverage effect, long memory
JEL classifications: C22, C58, G15.

1 Introduction

One of the best known model in financial econometrics is the *stochastic volatility* (SV) model, first proposed by Taylor (1982, 1986). This class of models is driven by two independent disturbances (random noises): the first impacts the return equation and

[*]We thank the computational support of the Center for Applied Research on Econometrics, Finance and Statistics (CAREFS), Brazil. The second author acknowledges the financial support of the São Paulo State Research Foundation (FAPESP), grant 2018/04654-9.

[†]Canvas Capital S.A. Email: muhieddine@gmail.com

[‡]Department of Statistics, University of Campinas, Brazil. Email: amadeus@unicamp.br

the other is an innovation in the log-variance process. For modeling financial returns, SV models are useful alternatives to GARCH models in which volatility only depends on squared past returns and one innovation process, see Broto and Ruiz (2004) and Shephard and Andersen (2009).

A distinctive *stylized fact* for volatility modeling is the presence of asymmetric behavior of the volatility after a positive/negative return. This feature, called the *leverage effect*, was first recognized by Black (1976) and was introduced in the GARCH literature by Nelson (1991) through the EGARCH models. The leverage effect can also be included in SV models if the two disturbances are correlated, resulting the so called *Asymmetric Stochastic Volatility* (A-SV) model. Omori et al. (2007) is an important reference for the estimation of A-SV models. The authors proposed a Bayesian estimator on a model where the joint distribution of the random noises is approximated. Abbara and Zevallos (2023b) combined the approximation method of Omori et al. (2007) and the the SV estimation procedure of Shumway and Stoffer (2006, chapter 6) to propose a maximum likelihood estimation method for A-SV models. In addition, the latter work includes a procedure for value-at-risk forecasting.

Another important empirical feature is the presence of long memory in volatility. This property is characterized by the hyperbolic decay in the autocorrelation function of (log) squared returns and has been usually incorporated in SV models through *autoregressive fractionally integrated moving average* (ARFIMA) processes for the log-variances. In that case, one obtains the so called *Long Memory Stochastic Volatility* (LMSV) model. In the literature, several methods have been proposed to estimate LMSV models. For instance, in the frequency domain we can cite Deo and Hurvich (2001) and in time domain Chan and Petris (2000). In addition, a fast maximum likelihood estimation method was proposed by Abbara and Zevallos (2023b), along with a procedure for VaR forecasting.

The class of SV models which presents both long memory and the leverage effect is called the *Asymmetric Long Memory Stochastic Volatility* (A-LMSV) model. Compared to other SV specifications, there are only a few works related to the estimation and forecasting of A-LMSV models, see for example Ruiz and Veiga (2008), who compares the properties of A-LMSV and FIEGARCH models. Therefore, our main objective is to fill this gap by proposing a fast procedure, based on the Kalman filter, to estimate A-LMSV models by maximum likelihood. This proposal extends the contributions of Abbara and Zevallos (2023a,b) and also provides a method for volatility forecasting.

There are some related works outside the framework discussed in the paper that deal with both long memory and leverage effect. For instance, the FIEGARCH proposed by Bollerslev and Mikkelsen (1996), and the works of Asai et al. (2012, 2017), where a new method for simultaneously modeling daily returns and realized volatility is presented/ see also Bilayi-Biakana et al. (2019) and Phillip et al. (2020).

The paper is organized in 5 sections, including this introduction. Section 2 presents the method while section 3 assess the estimation proposal through Monte Carlo experiments. Section 4 presents an illustration with real-life time series and Section 5 concludes. The derivation of the Kalman filter algorithm is summarized in the Appendix.

2 Methods

Let $\{r_t\}_{t=1}^{T}$ be a time series of returns. The LMSV model is given by the equations:

$$\begin{aligned} r_t &= \beta e^{h_t/2}\varepsilon_t, & (1)\\ \phi(B)(1-B)^d h_t &= \theta(B)\omega_t & (2)\end{aligned}$$

where B is the backshift operator, $\phi(B) = 1-\phi_1 B - \ldots - \phi_p B^p$ and $\theta(B) = 1+\theta_1 B + \ldots + \theta_q B^q$ are the autoregressive and moving average polynomials, respectively, and $(1-B)^d = \sum_{j=0}^{\infty} b_j(d)B^j$ with

$$b_j(d) = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} \quad j = 0, 1, \ldots. \tag{3}$$

In addition, the disturbances ε_t and ω_s are independent for all t and s. We do not assume any specific distribution for ε_t provided that $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = 1$, and we consider that ω_t has a normal distribution with zero mean and variance σ_ω^2.

When disturbances ε_t and ω_{t+1} are correlated then the model exhibits a leverage effect and it is called an A-LMSV model. The correlation is included between ε_t and ω_{t+1}, instead of ε_t and ω_t, to ensure that the returns process is a martingale difference, see Yu (2005) and Ruiz and Veiga (2008). Specifically, we consider that the vector $(\varepsilon_t, \omega_{t+1})$ has a covariance matrix given by:

$$\Sigma = \begin{bmatrix} 1 & \rho\sigma_\omega \\ \rho\sigma_\omega & \sigma_\omega^2 \end{bmatrix}, \tag{4}$$

and we do not assume any specific joint distribution for the vector $(\varepsilon_t, \omega_{t+1})$, but we maintain the assumption that $\omega_t \sim N(0, \sigma^2_\omega)$.

Taking the log of the squares of equation (1) yields:

$$y_t = \alpha + h_t + \eta_t, \tag{5}$$

where $y_t = \ln(r_t^2)$, $\alpha = \ln(\beta^2)$ and $\eta = \ln \varepsilon_t^2$. The goal is to write the model as a dynamic linear switching model with mixtures. To do that, we need to tackle two problems. First, we need to specify a joint distribution of (η_t, ω_{t+1}). Here, we follow an approach very similar to that Abbara and Zevallos (2023b) for A-SV models. Second, we want to obtain a finite linear state space representation of the ARFIMA process for h_t. In this paper, we use a autoregressive approximation[1]as discussed for LMSV models in Abbara and Zevallos (2023a).

All these issues are discussed briefly in the next subsections.

2.1 Approximation of the joint distribution of (ω_{t+1}, η_t).

The density $f(\omega_{t+1}, \eta_t)$ is decomposed as:

$$f(\omega_{t+1}, \eta_t) = f(\eta_t) f(\omega_{t+1}|\eta_t),$$

where $f(\eta_t)$ is approximated by a mixture of m normals as in Shumway and Stoffer (2006, chap.6). Thus:

$$\eta_t = \sum_{j=1}^{m} I_{jt} V_{jt}, \tag{6}$$

where $\{I_{jt}\}$ is an independent Bernoulli sequence with $P(I_{jt} = 1) = \pi_j$, and $V_{jt} \sim N(\mu_j, \sigma_j^2)$ with $\mu_1 = 0$. We consider $\pi_j = 1/m$, although it is possible to choose other values.

On the other hand, from (6) and following the same strategy of Abbara and Zevallos (2023b), an approximation of $f(\omega_{t+1}|\eta_t)$ is given by:

$$f(\omega_{t+1}|\eta_t, I_{jt} = 1) = f(\omega_{t+1}|V_{jt}) \approx N(d_t \rho \sigma_\omega \exp(\mu_j/2)(a_j + b_j(V_{jt} - \mu_j)), \sigma^2_\omega(1 - \rho^2)). \tag{7}$$

[1]There is an alternative for autoregressive approximation, originally proposed by Chan and Petris (2000), where the LMSV model is constructed based on the MA representation of the first difference of y_t. We will study this representation in the next steps of our work.

As a result, the joint distribution $f(\omega_{t+1}, \eta_t)$ satisfies:

$$\begin{bmatrix} \eta_t \\ \omega_{t+1} \end{bmatrix} |d_t, I_{jt} = 1 \sim \begin{bmatrix} \mu_j + \sigma_j z_t \\ d_t \rho \sigma_\omega (a_j + b_j \sigma_j z_t) \exp(\mu_j/2) + \sigma_\omega \sqrt{1-\rho^2} z_t^* \end{bmatrix}, \tag{8}$$

with $(z_t, z_t^*) \sim N(0, I_2)$, and from equation (8) we obtain:

$$A_{jt} = E(\omega_{t+1}|d_t, I_{jt} = 1) = d_t \rho \sigma_\omega a_j \exp(\mu_j/2), \tag{9}$$

$$B_{jt} = \mathrm{Var}(\omega_{t+1}|d_t, I_{jt} = 1) = \rho^2 \sigma_\omega^2 b_j^2 \sigma_j^2 \exp(\mu_j) + \sigma_\omega^2 (1-\rho^2), \tag{10}$$

where d_t is calculated as $d_t = I(r_t \geq 0) - I(r_t < 0)$ for $t = 1, \ldots, n$.

2.2 State-space model and the Kalman filter algorithm

As stated in equation (2), h_t follows an ARFIMA(p, d, q) process, and it can be expressed as an infinite AR expansion given by $g(B) = \phi(B)(1-B)^d = \sum_{j=0}^{\infty} g_j B^j h_t$, see Palma (2007) for more details. Thus, as in Abbara and Zevallos (2023a), we truncate the polynomial $g(B)$ up to the power K:

$$\sum_{j=0}^{K} g_j B^j h_t \approx \theta(B) \omega_t \tag{11}$$

and then we obtain an ARMA(K, q) model which has the following state-space representation for the ARFIMA(1,d,1) case:

$$y_t = \Theta \mathbf{X}_t + \eta_t, \tag{12}$$

$$\mathbf{X}_t = \Phi \mathbf{X}_{t-1} + \mathbf{H} \omega_t, \tag{13}$$

where

$$\Theta = \begin{bmatrix} \alpha & 0 & 0 & \ldots & \theta & 1 \end{bmatrix}, \tag{14}$$

$$\mathbf{X}_t = \begin{bmatrix} 1 & X_{t-K+1} & X_{t-K+2} & \ldots & X_t \end{bmatrix}', \tag{15}$$

$$\Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & g_K & g_{K-1} & g_{K-2} & g_{K-3} & \ldots & g_1 \end{bmatrix}, \tag{16}$$

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & \ldots & 1 \end{bmatrix}', \tag{17}$$

$$g_j = \phi b_{j-1}(d) - b_j(d), \quad j = 1, \ldots, K. \tag{18}$$

Equations (12) and (13) with the innovations given by the joint distribution (8) form a conditionally linear Gaussian state-space model. Therefore we can easily derive the Kalman filter algorithm.

Let $\mathbf{X}_{t|t-1} = \mathrm{E}(\mathbf{X}_t|y_{1:t-1})$, $\mathbf{X}_{t|t} = \mathrm{E}(\mathbf{X}_t|y_{1:t})$ the predicted and filtered values of the state vector $\mathbf{X}_t$, and consider $\mathbf{P}_{t|t-1} = \mathrm{Var}(\mathbf{X}_t - \mathbf{X}_{t|t-1}|y_{1:t-1})$, $\mathbf{P}_{t|t} = \mathrm{Var}(\mathbf{X}_t - \mathbf{X}_{t|t}|y_{1:t})$ and $\pi_{jt} = P(I_{jt} = 1|y_{1:t})$. Following similar steps as Abbara and Zevallos (2023a), we have (see Appendix for details):

$$
\begin{aligned}
\mathbf{X}_{t|t-1} &= \Phi\mathbf{X}_{t-1|t-1} + \mathbf{H}\Big(\sum_{j=1}^{m}\pi_{jt-1}A_{jt-1}\Big), &(19)\\
\epsilon_{jt} &= y_t - \mu_j - \Theta\mathbf{X}_{t|t-1}, &(20)\\
\mathbf{\Sigma_{jt}} &= \mathbf{\Theta P_{t|t-1}\Theta'} + \sigma_{\mathbf{j}}^{\mathbf{2}} &(21)\\
\mathbf{k}_{jt} &= \frac{\mathbf{P}_{t|t-1}\Theta'}{\Sigma_{jt}} &(22)\\
\mathbf{X}_{t|t} &= \mathbf{X}_{t|t-1} + \sum_{j=1}^{m}\pi_{jt}\mathbf{k}_{jt}\epsilon_{jt}, &(23)\\
\mathbf{P}_{t|t-1} &= \Phi\mathbf{P}_{t-1|t-1}\Phi' + \Big(\sum_{j=1}^{m}\pi_{jt-1}B_j\Big)\mathbf{HH}', &(24)\\
\mathbf{P}_{t|t} &= \sum_{j=1}^{m}\pi_{jt}(\mathbf{I}_{K+1} - \mathbf{k}_{jt}\Theta)\mathbf{P}_{t|t-1}, &(25)
\end{aligned}
$$

for $t = 1, \ldots, T$. The probabilities π_{jt} are calculated as follows:

$$\pi_{jt} = \frac{\pi_j f_j(t|t-1)}{\sum_{k=1}^{m}\pi_k f_k(t|t-1)}, \tag{26}$$

where $f_j(t|t-1)$ is approximated by $N(\alpha + h_{t|t-1} + \mu_j, \Sigma_{jt})$. Then, the log-likelihood is given by:

$$l(\Lambda) = \sum_{t=1}^{n}\ln\left\{\sum_{j=1}^{m}\pi_j f_j(t|t-1)\right\}, \tag{27}$$

where $\Lambda = (\lambda, \tau)$, $\lambda = (\alpha, d, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, \sigma_\omega, \rho)$ and $\tau = (\mu_2, \ldots, \mu_m, \sigma_1, \ldots, \sigma_m)$ are the set of parameters of the model.

Thus, let $r_1, \ldots, r_T$ be a sample of returns; we can summarize the steps to estimate the parameters as follows:

1. Calculate $y_t = \ln(r_t^2)$ for $t = 1, \ldots, T$.

2. Choose appropriate values of K and m.

3. Set initial values for $\mathbf{X}_{t|t-1}, \mathbf{X}_{t|t}, \mathbf{P}_{t|t-1}$, $\mathbf{P}_{t|t-1}$ and π_{jt}.

4. Compute the log-likelihood as in (27) through the recursive calculation of the quantities given in equations (20) to (25) for $t \in \{1, ..., T\}$ and $j \in \{1, ..., m\}$. Then, maximize the log-likelihood by a numerical routine and obtain an estimate of Λ.

2.3 VaR forecasting

From the state-space representation, a one-step-ahead forecast for the volatility is given by:

$$\sigma_{T+1|T} = \beta \exp(0.5\Theta X_{T+1|T}). \tag{28}$$

It can be used for value-at-risk (VaR) prediction through the evaluation of the empirical distribution of ε_t. Thus, for $t \in \{1, ..., T\}$ the procedure consists of the following steps:

1. Calculate $\hat{\sigma}_{t|t-1} = \hat{\beta} \exp(0.5\hat{\Theta}\hat{X}_{t|t-1})$

2. Calculate $e_t = r_t / \hat{\sigma}_{t|t-1}$

3. Calculate the γ-quantile of e_t, called q_γ,

4. The one-step $100(1-\gamma)\%$-VaR forecast is equal to $q_\gamma \hat{\sigma}_{T|T-1}$.

This proposal has a small difference compared to that of Abbara and Zevallos (2023a). There, we used the filtered volatility ($\sigma_{t|t}$), instead of the predicted ones, for the calculation of e_t. We tested both volatilities and found better results with the predicted volatilities $\sigma_{t|t-1}$.

2.4 Implementation

All routines were implemented in R (R Core Team, 2020) using the packages `Rcpp` and `RcppArmadillo`. The likelihood optimization routine was performed using the `nlminb` function in R. We used unconstrained optimization with transformed parameters to account for the restrictions: $\phi \in (-1, 1)$, $\theta \in (-1, 1)$, $\rho \in (-1, 1)$ and $d \in (0, 1)$, and the variances of the model are in interval $(0, \infty)$. We emphasize that the parameter d

can be outside the stationary interval, extending the stationary case studied by Ruiz and Veiga (2008).

We consider $\pi_j = 1/m$, $m \in \{2, 3\}$ and $K = 75$. In addition, the initial values of the Kalman filter quantities are: $\mathbf{P}_{1|0} = \mathbf{I}_{K+1}\sigma_\omega^2$, $\mathbf{X}_{1|0} = \mathbf{0}_{K+1}$ and $\pi_{j1|0} = 1/m$. Moreover, the initial values of the parameters are $d_0 = 0.75$, $\phi_0 = \theta_0 = 0$, $\sigma_{\omega,0} = 0.3$, while the initial values of α_0, ρ_0, $\mu_{j,0}$ (for $j > 1$) and $\sigma_{j,0}$ are obtained estimating the A-SV model. These choices permit fast estimation of the A-LMSV model but we stress that other initial values do not affect the convergence of the optimization procedure.

3 Monte Carlo Experiments

Here we evaluate the performance of our proposal in terms of point estimation by a Monte Carlo experiment. The true values of the data generating process were chosen based on the real-life results presented in Section 4, and we generated 1,000 samples with length $T = 5,000$ observations. The performance of the estimates was assessed by the bias, the root mean square error (RMSE) and the standard deviation. Like our previous works, we considered $m \in \{2, 3\}$ to measure the sensitivity of the estimates to the approximation of η_t.

As in Abbara and Zevallos (2023b) we considered two distributions for ε_t: i) the standard Gaussian, and ii) the t-Student distribution with $v = 5$ degrees of freedom, which is standardized to have unit variance. We also generated samples for ω_t after generating ε_t considering $\omega_{t+1}|\varepsilon_t \sim N(\rho\sigma_\omega\varepsilon_t, (1-\rho^2)\sigma_\omega^2)$.

The results are presented in Table 1. One can observe very good results for $\hat{d}$, $\hat{\sigma}_\omega$ and $\hat{\alpha}$, because of the small bias values. Moreover, the bias and RMSE are higher for $\hat{\sigma}_\omega$ compared with the other two parameters, but they are at most 14% of the true parameter value. However, compared to those estimates, we did not obtain good results for $\hat{\rho}$, mainly because the higher bias and RMSE values. According to its standard deviation (which is small), the RMSE values of $\hat{\rho}$ are large because of the bias.

[Table 1 around here]

Next, we compared the results obtained from different values of m. First, for both distributions, the performance of $\hat{\rho}$ and $\hat{\sigma}_\omega$ improved when using $m = 3$, because of the decrease in the bias and RMSE values. Second, the results for $\hat{d}$ were very good and similar across the values of m, for both distributions. Third, we found mixed results for $\hat{\alpha}$. Its performance improves for $m = 3$ when $\varepsilon_t \sim N(0, 1)$ but not when $\varepsilon_t \sim t_5$.

4 Illustrations

In this section, the proposed method is illustrated through the analysis of three time series returns of major country stock market indexes: the Brazilian (IBOV), the Spanish (IBEX) and the Japanese (Nikkei) indexes. All time series start on January 2, 1996, and end on December 30, 2021 for the IBOV and Nikkei, andon May 15, 2021 for the IBEX. The sample sizes are: 6,409 for IBOV, 6,415 for IBEX and 6,371 for Nikkei. The data were collected from Bloomberg.

4.1 Real-life estimates of the A-LMSV model

First, we present the results of the first 5,000 returns of each series. Figure 1 presents the time series plots and the empirical autocorrelation function (ACF) of the log-squared returns. The hyperbolic decay of the ACF can be seen, so a model with long memory is a proper choice. Since several works in the literature account for the existence of a leverage effect in equity returns, see for example Omori et al. (2007), an interesting candidate is the A-LMSV model. Estimates of the A-LMSV model are presented in Table 2 for $m \in \{2, 3\}$.

[Figure 1 and Table 2 around here]

Table 2 shows that most $\hat{d}$ values are between 0.5 and 0.7, outside the stationary range. The only exceptions occurs for the Nikkei using the ARFIMA$(1, d, 0)$ model, for all values of m, and for the ARFIMA$(1, d, 1)$ model with $m = 3$. Furthermore, $\hat{\rho}$ values are negative for all series, indicating the presence of a leverage effect. Also, the values of $\hat{\rho}$ are, in general, higher for IBEX compared with the other series.

Another important result from Table 2 is the instability of θ and ϕ estimates. The significance of $\hat{\phi}$ and $\hat{\theta}$ for ARFIMA$(1, d, 0)$, $(0, d, 1)$ and $(1, d, 1)$ differs substantially, depending on the specification of the model. For instance, in the case of the IBEX, these parameters are not significant for the $(1, d, 0)$ and $(0, d, 1)$ fits, while they are significant for the $(1, d, 1)$ fit, although it seems that the polynomials $\phi(B)$ and $\theta(B)$

have a common root, so the appropriate model is $(0, d, 0)$. For the Nikkei time series, $\hat{\phi}$ is significant for the $(1, d, 0)$ fit but $\hat{d}$ is not, while $\hat{\phi}$ are not significant for $(1, d, 1)$. Finally, for the IBOV case, all parameter estimates are significant for all models, except $\hat{\phi}$ for the $(1, d, 1)$ model, for all values of m.

In Figure 2 we present the evolution of the predicted volatility for all series using the ARFIMA$(0, d, 0)$ specification. It can be seen that the volatility reproduces the variability of the returns. In particular we observe a increase in volatility around 1998 (period of the Russian crisis) for IBEX and IBOV and a sharp increase in volatility for all series in 2008, the period of the global financial crisis.

[Figure 2 around here]

4.2 Backtesting results

Here we assess the one-step-ahead VaR forecast performance for the three series. The exercise consists of a rolling window procedure with time length of 5,000 observations where for each iteration, the parameters and VaR forecasts are calculated[2].

We compare two models: the A-LMSV described earlier and the LMSV model (both using $K = 75$). VaR forecasts for the LMSV model were obtained in the same way as discussed in subsection 2.3 (that is, using the predicted volatility instead of the filtered volatility, to calculate e_t). Our goal is to assess whether the inclusion of leverage effect in the model improves the performance of the estimated VaRs.

For both models we consider that h_t follows an ARFIMA$(0, d, 0)$ process. We choose this specification because, as discussed in the previous section, when estimating ARFIMA$(1, d, 0)$ and ARFIMA$(0, d, 1)$ models, some parameter estimates are unstable. Additionally, as in Abbara and Zevallos (2023a,b), we use $m \in \{2, 3\}$ and compute VaRs for long and short positions. The latter is valuable because the A-LMSV posits a different behavior for volatility after a negative shock compared to a positive one, so we expected a different behavior of buyers and sellers (of the asset).

The quality of the VaR forecasts was assessed according to Kupiec (1995), Christoffersen (1998) and Christoffersen and Pelletier (2004). These results are presented in

[2] At each iteration, we considered as initial parameter values the parameter estimates of the previous iteration, obviously except for the first iteration. This choice improved the time of convergence for most of the estimates. However, there were no substantial differences when choosing the initial values as described in subsection 2.4

Table 3.

[Table 3 around here]

Overall, we obtained very good results using both methods, with the exception of 99%-VaR for a long position in the IBOV and in Nikkei. For most of the VaRs the proportion of violations are close to the nominal level as assessed by the Kupiec test, and most of the violations are independent.

There are some situations where the A-LMSV presents better results compared with the LMSV. For instance, for a short position in the IBOV, we find higher p-values of the Kupiec test for 97.5% and 99%-VaR, with both values of m. Besides this, we also find higher p-values of the Kupiec test for both long and short positions in the IBEX 95%-VaR with $m = 3$.

Figure 3 plots the 99%-VaR forecasts for the A-LMSV using $m = 3$ for all series. Here, we observe an increase in the VaRs in the beginning of 2020, the start of the Covid-19 pandemic. We do not observe any loss much larger than its respective VaR forecast most of the time, and for IBOV we observe many days with returns below the predicted VaRs during the pandemic crisis, in the months of February and March 2020.

[Figure 3 around here]

Finally, here we compare the VaR forecasts of the LMSV and A-LMSV models using the T_{max} statistic, which is part of the Model Confidence Set approach of Hansen et al. (2011). The statistic of the test was calculated using the asymmetric loss function, and the p-values were obtained by a block-bootstrap procedure. The comparison was performed using the `MCS` package in R/ see Bernardi and Catania (2016) for more details. In Table 4 we present the p-values of the test for each model, and the model with the highest p-value presents better results. We can see from Table 4 that the A-LMSV has the highest p-value for most VaRs, for long or short positions. The exceptions occur for 97.5%-VaR for the IBEX (both long and short positions), 99%-VaR for a short position in the IBEX and for 95%-VaR for short position in the IBOV (with $m = 3$). Therefore, the inclusion of the leverage effect in the model improves VaR forecasting in most situations, although both the LMSV and A-LMSV models present a close proportion of violations.

[Table 4 around here]

5 Conclusions

This paper presents a new maximum likelihood method for parameter estimation and value-at-risk prediction of A-LMSV models. The likelihood is readily obtained from a Kalman filter algorithm derived from a dynamic linear model with mixtures. To obtain the model: i) an approximation of the joint distribution of (η_t, ω_{t+1}) is proposed, and ii) autoregressive truncation of the long memory process of h_t is considered. The derivation of the algorithm follows similar steps as those presented by Abbara and Zevallos (2023a,b). In terms of parameter estimation, Monte Carlo experiments indicated sound results for most parameters but we observed some problems when estimating ρ. However, we obtained good results in the backtesting exercise for VaR forecasts and we found that the inclusion of the leverage effect in the model improved the VaR forecast performance in comparison with the LMSV and A-LMSV models.

Appendix

Here, we present the derivation of the Kalman filter expressions presented in Section 2.2.

The predicted vector $\mathbf{X}_{t|t-1}$ is obtained as follows:

$$\begin{aligned}
\mathbf{X}_{t|t-1} &= \mathrm{E}(\mathbf{X}_t|y_{1:t-1}) \\
&= \mathrm{E}(\Phi\mathbf{X}_{t-1} + \mathbf{H}\omega_t|y_{1:t-1}) \\
&= \Phi\mathrm{E}(\mathbf{X}_{t-1}|y_{1:t-1}) + \mathbf{H}\mathrm{E}(\omega_t|y_{1:t-1}) \\
&= \Phi\mathbf{X}_{t-1|t-1} + \mathbf{H}\mathrm{E}\Big(\mathrm{E}(\omega_t|y_{1:t-1}, I_{jt-1}=1)|y_{1:t-1}\Big) \\
&= \Phi\mathbf{X}_{t-1|t-1} + \mathbf{H}\mathrm{E}\Big(\sum_{j=1}^{m}\mathrm{E}(\omega_t|y_{1:t-1}, I_{jt-1}=1)I_{jt}|y_{1:t-1}\Big) \\
&= \Phi\mathbf{X}_{t-1|t-1} + \mathbf{H}\sum_{j=1}^{m}A_{jt-1}\mathrm{E}(I_{jt-1}|y_{1:t-1}) \\
&= \Phi\mathbf{X}_{t-1|t-1} + \mathbf{H}\sum_{j=1}^{m}A_{jt-1}\pi_{jt-1} \qquad (29)
\end{aligned}$$

We define the innovation ϵ_{jt} as:

$$\begin{aligned}
\epsilon_{jt} &= y_t - \mathrm{E}(y_t|y_{1:t-1}, I_{jt}=1) \\
&= y_t - \mathrm{E}[\Theta\mathbf{X}_t|y_{1:t-1}, I_{jt}=1] - \mathrm{E}(V_{jt}|y_{1:t-1}) \\
&= y_t - \mathrm{E}(V_{jt}) - \Theta\mathbf{X}_{t|t-1} \\
&= y_t - \mu_j - \Theta\mathbf{X}_{t|t-1} \qquad (30)
\end{aligned}$$

For now, our objective is to find the joint distribution between $\mathbf{X}_{t|t-1}$ and ϵ_{jt}. Thus, let $\mathbf{P}_{t|t-1} = \mathrm{Var}(\mathbf{X}_t - \mathbf{X}_{t|t-1}|y_{1:t-1})$, so the covariance between both quantities is equal to:

$$\begin{aligned}
\mathrm{cov}(\mathbf{X}_t, \epsilon_{jt}|y_{1:t-1}) &= \mathrm{cov}(\mathbf{X}_t, y_t - \mu_j - \Theta\mathbf{X}_{t|t-1}|y_{1:t-1}) \\
&= \mathrm{cov}(\mathbf{X}_t - \mathbf{X}_{t|t-1}, \Theta\mathbf{X}_t - \Theta\mathbf{X}_{t|t-1} + V_{jt}|y_{1:t-1}) \\
&= \mathrm{cov}(\mathbf{X}_t - \mathbf{X}_{t|t-1}, \Theta(\mathbf{X}_t - \mathbf{X}_{t|t-1})|y_{1:t-1}) \\
&= \mathrm{Var}(\mathbf{X}_t - \mathbf{X}_{t|t-1}|y_{1:t-1})\Theta' \\
&= \mathbf{P}_{t|t-1}\Theta'
\end{aligned}$$

since $\mathrm{cov}(\omega_t, V_{jt}) = 0$. Finally, the joint distribution is equal to:

$$\begin{bmatrix} \mathbf{X}_t \\ \epsilon_{jt} \end{bmatrix} |y_{1:t-1}, I_{jt} = 1 \sim N\left(\begin{bmatrix} \mathbf{X}_{t|t-1} \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{P}_{t|t-1} & \mathbf{P}_{t|t-1}\Theta' \\ \Theta\mathbf{P}_{t|t-1} & \Sigma_{jt} \end{bmatrix} \right), \qquad (31)$$

Based on the distribution in (31), we can obtain:

$$\begin{aligned}
\mathrm{E}(\mathbf{X}_t|y_{1:t}, I_{jt}=1) &= \mathrm{E}(\mathbf{X}_t|\epsilon_{jt}, y_{1:t-1}, I_{jt}=1) \\
&= \mathbf{X}_{t|t-1} + \mathbf{k}_{jt}\epsilon_{jt} \qquad (32)
\end{aligned}$$

where $\mathbf{k}_{jt}$ is the Kalman gain given by:

$$\mathbf{k}_{jt} = \frac{\mathbf{P}_{t|t-1}\Theta'}{\Sigma_{jt}} = \frac{\mathbf{P}_{t|t-1}\Theta'}{\Theta\mathbf{P}_{t|t-1}\Theta' + \sigma_j^2}. \qquad (33)$$

Finally, $\mathbf{X}_{t|t}$ is obtained as follows:

$$\begin{aligned}
\mathbf{X}_{t|t} &= \mathrm{E}(\mathbf{X}_t|y_{1:t}) \\
&= \mathrm{E}\Big(\mathrm{E}(\mathbf{X}_t|y_{1:t}, I_{jt}=1)|y_{1:t}\Big) \\
&= \mathrm{E}\Big(\sum_{j=1}^{m}\mathrm{E}(\mathbf{X}_t|y_{1:t}, I_{jt}=1)I_{jt}|y_{1:t}\Big) \\
&= \sum_{j=1}^{m}\mathrm{E}(\mathbf{X}_t|y_{1:t}, I_{jt}=1)\mathrm{E}(I_{jt}|y_{1:t}) \\
&= \sum_{j=1}^{m}\mathrm{E}(\mathbf{X}_t|y_{1:t}, I_{jt}=1)\pi_{jt} \qquad (34)
\end{aligned}$$

By replavcing Equation (32) in (34), we obtain:

$$\begin{aligned}
\mathbf{X}_{t|t} &= \sum_{j=1}^{m}(\mathbf{X}_{t|t-1}+\mathbf{k}_{jt}\epsilon_{jt})\pi_{jt} \\
&= \mathbf{X}_{t|t-1}+\sum_{j=1}^{m}\pi_{jt}\mathbf{k}_{jt}\epsilon_{jt}.
\end{aligned}$$

Let $\Omega_t = \sum_{j=1}^{m}\pi_{jt}A_{jt}$. The expression for $\mathbf{P}_{t|t-1}$ is obtained as follows:

$$\begin{aligned}
\mathbf{P}_{t|t-1} &= \mathrm{E}\Big((\mathbf{X}_t-\mathbf{X}_{t|t-1})(\mathbf{X}_t-\mathbf{X}_{t|t-1})'|y_{1:t-1}\Big) \\
&= \mathrm{E}\Big((\Phi\mathbf{X}_{t-1}+\mathbf{H}\omega_t-\Phi\mathbf{X}_{t-1|t-1}-\mathbf{H}\Omega_t)(\Phi\mathbf{X}_{t-1}+\mathbf{H}\omega_t-\Phi\mathbf{X}_{t-1|t-1}-\mathbf{H}\Omega_t)'|y_{1:t-1}\Big) \\
&= \mathrm{E}\Big(\Phi(\mathbf{X}_{t-1}-\mathbf{X}_{t-1|t-1})(\mathbf{X}_{t-1}-\mathbf{X}_{t-1|t-1})'\Phi'|y_{1:t-1}\Big)+\mathrm{E}(\mathbf{H}\mathbf{H}'(\omega_t-\Omega_t)^2|y_{1:t-1}) \\
&= \Phi\mathbf{P}_{t-1|t-1}\Phi'+\mathbf{H}\mathbf{H}'\mathrm{E}\big((\omega_t-\Omega_t)^2|y_{1:t-1}\big) \\
&= \Phi\mathbf{P}_{t-1|t-1}\Phi'+\mathbf{H}\mathbf{H}'\mathrm{E}\Big(\sum_{j=1}^{m}E((\omega_t-\Omega_t)^2|y_{1:t-1}, I_{jt-1}=1)I_{jt-1}|y_{1:t-1}\Big) \\
&= \Phi\mathbf{P}_{t-1|t-1}\Phi'+\mathbf{H}\mathbf{H}'\sum_{j=1}^{m}\mathrm{E}(B_jI_{jt-1}|y_{1:t-1}) \\
&= \Phi\mathbf{P}_{t-1|t-1}\Phi'+\mathbf{H}\mathbf{H}'\sum_{j=1}^{m}B_j\pi_{jt-1} \qquad (35)
\end{aligned}$$

Finally we obtain the expression for $\mathbf{P}_{t|t}$. Thus,

$$\begin{aligned}
\mathbf{P}_{t|t} &= \mathrm{E}\big[(\mathbf{X}_t-\mathbf{X}_{t|t})(\mathbf{X}_t-\mathbf{X}_{t|t})'|y_{1:t}\big] = \mathrm{E}\Big[\mathrm{E}\big[(\mathbf{X}_t-\mathbf{X}_{t|t})(\mathbf{X}_t-\mathbf{X}_{t|t})'|y_{1:t}, I_{jt}=1\big]|y_{1:t}\Big], \\
&= \sum_{j=1}^{m}\mathrm{E}\big[(\mathbf{X}_t-\mathbf{X}_{t|t})(\mathbf{X}_t-\mathbf{X}_{t|t})'|y_{1:t}, I_{jt}=1\big]\pi_{jt}. \qquad (36)
\end{aligned}$$

Note that:

$$\begin{aligned}
\mathrm{E}\big[(\mathbf{X}_t - \mathbf{X}_{t|t})(\mathbf{X}_t - \mathbf{X}_{t|t})'|y_{1:t}, I_{jt} = 1\big] &= \mathrm{E}\big[(\mathbf{X}_t - \mathbf{X}_{t|t})(\mathbf{X}_t - \mathbf{X}_{t|t})'|\epsilon_{jt}, y_{1:t-1}, I_{jt} = 1\big] \\
&= \mathrm{Var}(\mathbf{X}_t|\epsilon_{jt}, y_{1:t-1}, I_{jt} = 1) \\
&= \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{k}_{jt}\Theta, \qquad (37)
\end{aligned}$$

because of the joint distribution in (31). Substituting Equation (37) in (36) yields:

$$\mathbf{P}_{t|t} = \sum_{j=1}^{m} \big[\mathbf{I}_{K+1} - \mathbf{k}_{jt}\Theta\big]\mathbf{P}_{t|t-1}\pi_{jt}. \qquad (38)$$

References

O. Abbara and M. Zevallos. Estimation and forecasting of long memory stochastic volatility models. *Studies in Nonlinear Dynamics and Econometrics*, 27(1):1–24, 2023a. doi: https://doi.org/10.1515/snde-2020-0106.

O. Abbara and M. Zevallos. Maximum likelihood inference for asymmetric stochastic volatility models. *Econometrics*, 11(1):1–18, 2023b. doi: https://doi.org/10.3390/econometrics11010001.

M. Asai, M. McAleer, and M. Medeiros. Asymmetry and long memory in volatility modeling. *Journal of Financial Econometrics*, 10(3):495–512, 2012.

M. Asai, C. Chang, and M. McAleer. Realized stochastic volatility with general asymmetry and long memory. *Journal of Econometrics*, 199:202–2012, 2017.

M. Bernardi and L. Catania. Comparison of value-at-risk models using the mcs package. *Computational Statistics*, 31:579–608, 2016.

C. Bilayi-Biakana, G. Ivanoff, and R. Kulik. The tail empirical process for long memory stochastic volatility models with leverage. *Electronic Journal of Statistics*, 13:3453–3484, 2019.

F. Black. Studies of stock price volatility changes. In *Proceedings of the Business and Economics Section of the American Statistical Association*, pages 177–181, 1976.

T. Bollerslev and H. O. Mikkelsen. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, 73(1):151–184, 1996.

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

N. Chan and G. Petris. Long memory stochastic volatility: a bayesian approach. *Communications in Statistics - Theory and Methods*, 29:1367–1378, 2000.

P. Christoffersen. Evaluating interval forecasts. *Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance*, 39:841–862, 1998.

P. Christoffersen and D. Pelletier. Backtesting value-at-risk: a duration-based approach. *Journal of Financial Econometrics*, 2:84–108, 2004.

R. Deo and C. Hurvich. On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory*, 18: 686–710, 2001.

P Hansen, A. Lunde, and J. Mason. The model confidence set. *Econometrica*, 79(2): 453–497, 2011.

P. Kupiec. Techniques for verifying the accuracy of risk management models. *Journal of Derivatives*, 3:73–84, 1995.

D. Nelson. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica*, 59(2):347–370, 1991.

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

W. Palma. *Long Memory Time Series: Theory and Methods*. John Wiley and Sons, New Jersey, USA, 2007.

A. Phillip, J. Chan, and S. Peiris. On generalized bivariate student-t gegenbauer long memory stochastic volatility models with leverage: Bayesian forecasting of cryptocurrencies with a focus on bitcoin. *Econometrics and Statistics*, (16):69–90, 2020. doi: https://doi.org/10.1016/j.ecosta.2018.10.003.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

E. Ruiz and H. Veiga. Modelling long-memory volatilities with leverage effect: A-lmsv versus fiegarch. *Computational Statistics and Data Analysis*, (52):2846–2862, 2008. doi: 10.1016/j.csda.2007.09.031.

N. Shephard and T. Andersen. Stochastic volatility: Origins and overview. In R. Andersen, J. Davis, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 233–254. Springer-Herlag, Berlin Heidelberg, Germany, 2009.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications*. Springer-Verlag, New York, 2006.

S.J. Taylor. Financial returns modelled by the product of two stochastic processes - a study of daily sugar prices, 1961-79. In O. D. Anderson, editor, *Time Series Analysis: Theory and Practice*, volume 1, pages 203–226. Elsevier/North-Holland, New York, 1982.

S.J. Taylor. *Modelling Financial Time Series*. Wiley, New York, 1986.

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

Table 1: Results of Monte Carlo experiments for the A-LMSV model ($d = 0.65$, $\sigma_\omega = 0.35$, $\alpha = -8$, $\rho = -0.45$). For each case, the bias, standard deviation (SD), and root mean square error (RMSE) of the estimates are presented. Results are based on 1,000 replications of time series of size $T = 5,000$.

			$\boldsymbol{m=2}$				$\boldsymbol{m=3}$			
			$\hat{d}$	$\hat{\sigma}_\omega$	$\hat{\alpha}$	$\hat{\rho}$	$\hat{d}$	$\hat{\sigma}_\omega$	$\hat{\alpha}$	$\hat{\rho}$
Case 1	$\varepsilon_t \sim N(0,1)$	Bias	0.079	-0.045	-0.163	0.187	0.083	-0.008	0.349	0.169
		SD	0.071	0.069	0.514	0.064	0.067	0.069	0.497	0.069
		RMSE	0.106	0.082	0.539	0.198	0.107	0.069	0.607	0.183
Case 2	$\varepsilon_t \sim t_5$	Bias	0.077	-0.049	-0.491	0.216	0.081	-0.020	0.034	0.184
		SD	0.080	0.082	0.532	0.066	0.076	0.082	0.583	0.076
		RMSE	0.111	0.095	0.724	0.226	0.112	0.085	0.584	0.199

Table 2: Results of the estimation of the A-LMSV model for the real-life time series using the first 5,000 observations. The standard errors and standard deviations are in parentheses.

			$m=2$						$m=3$					
			$\hat{\alpha}$	$\hat{d}$	$\hat{\sigma}_\omega$	$\hat{\rho}$	$\hat{\theta}$	$\hat{\phi}$	$\hat{\alpha}$	$\hat{d}$	$\hat{\sigma}_\omega$	$\hat{\rho}$	$\hat{\theta}$	$\hat{\phi}$
IBOV	0d0	Estimate	-8.133	0.631	0.361	-0.376			-7.628	0.639	0.407	-0.407		
		Std. error	(0.209)	(0.055)	(0.065)	(0.075)			(0.213)	(0.052)	(0.064)	(0.076)		
	0d1	Estimate	-8.223	0.573	1.074	-0.180	-0.536		-7.666	0.623	0.841	-0.268	-0.467	
		Std. error	(0.202)	(0.059)	(0.032)	(0.039)	(0.076)		(0.221)	(0.055)	(0.040)	(0.054)	(0.080)	
	1d0	Estimate	-8.170	0.609	0.603	-0.279		-0.471	-7.640	0.600	0.712	-0.292		-0.459
		Std. error	(0.208)	(0.066)	(0.132)	(0.074)		(0.110)	(0.204)	(0.062)	(0.128)	(0.071)		(0.081)
	1d1	Estimate	-8.222	0.568	1.072	-0.180	-0.497	-0.052	-7.652	0.615	0.825	-0.266	-0.286	-0.245
		Std. error	(0.201)	(0.060)	(0.032)	(0.039)	(0.109)	(0.092)	(0.218)	(0.054)	(0.040)	(0.054)	(0.155)	(0.139)
IBEX	0d0	Estimate	-8.750	0.670	0.322	-0.464			-8.298	0.671	0.377	-0.481		
		Std. error	(0.240)	(0.046)	(0.051)	(0.083)			(0.242)	(0.046)	(0.056)	(0.080)		
	0d1	Estimate	-8.737	0.653	0.257	-0.492	0.312		-8.287	0.656	0.301	-0.510	0.308	
		Std. error	(0.223)	(0.047)	(0.080)	(0.085)	(0.399)		(0.226)	(0.047)	(0.095)	(0.082)	(0.417)	
	1d0	Estimate	-8.734	0.619	0.211	-0.513		0.445	-8.286	0.633	0.268	-0.520		0.371
		Std. error	(0.202)	(0.063)	(0.076)	(0.087)		(0.267)	(0.211)	(0.060)	(0.093)	(0.082)		(0.285)
	1d1	Estimate	-8.750	0.526	0.264	-0.506	-0.729	0.879	-8.291	0.559	0.314	-0.508	-0.694	0.839
		Std. error	(0.183)	(0.114)	(0.073)	(0.090)	(0.187)	(0.107)	(0.191)	(0.097)	(0.086)	(0.085)	(0.244)	(0.145)
Nikkei	0d0	Estimate	-8.801	0.661	0.335	-0.342			-8.279	0.697	0.355	-0.363		
		Std. error	(0.239)	(0.063)	(0.065)	(0.087)			(0.273)	(0.065)	(0.067)	(0.093)		
	0d1	Estimate	-8.801	0.661	0.335	-0.342	0.000		-8.279	0.696	0.333	-0.371	0.068	
		Std. error	(0.239)	(0.064)	(0.166)	(0.108)	(0.476)		(0.272)	(0.064)	(0.176)	(0.108)	(0.545)	
	1d0	Estimate	-8.748	0.000	0.168	-0.420		0.977	-8.208	0.377	0.121	-0.441		0.906
		Std. error	(0.128)	(0.000)	(0.022)	(0.098)		(0.006)	(0.145)	(0.179)	(0.028)	(0.102)		(0.065)
	1d1	Estimate	-8.787	0.565	1.088	-0.118	-0.478	-0.086	-8.213	0.396	0.235	-0.437	-0.544	0.912
		Std. error	(0.203)	(0.064)	(0.034)	(0.040)	(0.112)	(0.090)	(0.153)	(0.151)	(0.096)	(0.100)	(0.232)	(0.049)

Table 3: Backtesting results using LMSV and A-LMSV models. In both models h_t follows an ARFIMA$(0, d, 0)$ process.

Asset	Model	Stat/test	$m = 2$						$m = 3$					
			Long position			Short position			Long position			Short position		
			1%	2.5%	5%	5%	2.5%	1%	1%	2.5%	5%	5%	2.5%	1%
IBOV	LMSV	Proportion	0.014	0.028	0.049	0.047	0.021	0.011	0.015	0.028	0.048	0.048	0.020	0.008
		Kupiec	0.144	0.451	0.908	0.542	0.339	0.830	0.090	0.451	0.718	0.718	0.187	0.548
		Christoffersen	0.194	0.237	0.417	0.720	0.234	0.343	0.144	0.237	0.331	0.926	0.360	0.754
		Christoffersen-Pelletier	0.072	0.256	0.306	0.745	0.748	0.050	0.104	0.242	0.352	0.353	0.976	0.391
	A-LMSV	Proportion	0.016	0.028	0.049	0.050	0.027	0.010	0.016	0.029	0.047	0.048	0.023	0.009
		Kupiec	0.050	0.425	0.859	0.956	0.640	0.981	0.029	0.337	0.583	0.671	0.576	0.767
		Christoffersen	0.002	0.067	0.183	0.958	0.608	0.869	0.002	0.067	0.257	0.908	0.815	0.848
		Christoffersen-Pelletier	0.028	0.142	0.262	0.994	0.255	0.179	0.023	0.133	0.500	0.515	0.873	0.299
IBEX	LMSV	Proportion	0.011	0.028	0.053	0.048	0.025	0.009	0.011	0.027	0.056	0.046	0.027	0.010
		Kupiec	0.586	0.476	0.568	0.728	0.936	0.838	0.586	0.690	0.257	0.412	0.690	0.960
		Christoffersen	0.701	0.289	0.234	0.878	0.371	0.848	0.701	0.683	0.223	0.705	0.293	0.848
		Christoffersen-Pelletier	0.061	0.207	0.764	0.001	0.915	0.181	0.061	0.227	0.994	0.002	0.030	0.479
	A-LMSV	Proportion	0.011	0.025	0.054	0.047	0.023	0.012	0.012	0.028	0.053	0.052	0.023	0.015
		Kupiec	0.628	0.915	0.452	0.645	0.683	0.460	0.460	0.440	0.608	0.694	0.683	0.088
		Christoffersen	0.350	0.045	0.003	0.894	0.890	0.619	0.332	0.067	0.021	0.507	0.890	0.170
		Christoffersen-Pelletier	0.003	0.023	0.779	0.041	0.826	0.755	0.062	0.199	0.852	0.002	0.867	0.548
Nikkei	LMSV	Proportion	0.015	0.029	0.053	0.056	0.034	0.018	0.015	0.028	0.050	0.050	0.030	0.016
		Kupiec	0.066	0.334	0.585	0.304	0.037	0.012	0.110	0.526	0.956	0.946	0.259	0.039
		Christoffersen	0.115	0.212	0.494	0.558	0.108	0.031	0.161	0.571	0.425	0.974	0.516	0.078
		Christoffersen-Pelletier	0.439	0.804	0.288	0.131	0.375	0.850	0.423	0.684	0.486	0.051	0.253	0.541
	A-LMSV	Proportion	0.015	0.029	0.055	0.055	0.033	0.018	0.014	0.028	0.051	0.051	0.030	0.018
		Kupiec	0.066	0.334	0.364	0.364	0.076	0.012	0.175	0.424	0.858	0.858	0.259	0.012
		Christoffersen	0.115	0.062	0.453	0.533	0.190	0.031	0.305	0.533	0.454	0.932	0.516	0.031
		Christoffersen-Pelletier	0.971	0.341	0.629	0.017	0.098	0.702	0.522	0.998	0.826	0.058	0.212	0.285

Table 4: Comparison of VaR forecasts. This table shows the p-values of the T_{max} statistic proposed by Hansen et al. (2011) for the model confidence set (MCS). Higher p-values indicate a better model, and for all tests we applied the asymmetric loss function.

	IBOV		IBEX		Nikkei	
	$m = 2$	$m = 3$	$m = 2$	$m = 3$	$m = 2$	$m = 3$
99%-VaR, long position						
LMSV	0.039	0.502	0.933	0.939	0.039	0.053
A-LMSV	1.000	1.000	1.000	1.000	1.000	1.000
97.5%-VaR, long position						
LMSV	0.232	0.588	1.000	1.000	0.087	0.763
A-LMSV	1.000	1.000	0.738	0.700	1.000	1.000
95%-VaR, long position						
LMSV	0.007	0.012	0.571	0.742	0.216	0.583
A-LMSV	1.000	1.000	1.000	1.000	1.000	1.000
95%-VaR, short position						
LMSV	0.536	1.000	0.677	0.624	0.007	0.116
A-LMSV	1.000	0.446	1.000	1.000	1.000	1.000
97.5%-VaR, short position						
LMSV	0.231	0.731	1.000	1.000	0.037	0.217
A-LMSV	1.000	1.000	0.675	0.388	1.000	1.000
99%-VaR, short position						
LMSV	0.219	0.974	1.000	1.000	0.185	0.381
A-LMSV	1.000	1.000	0.243	0.239	1.000	1.000

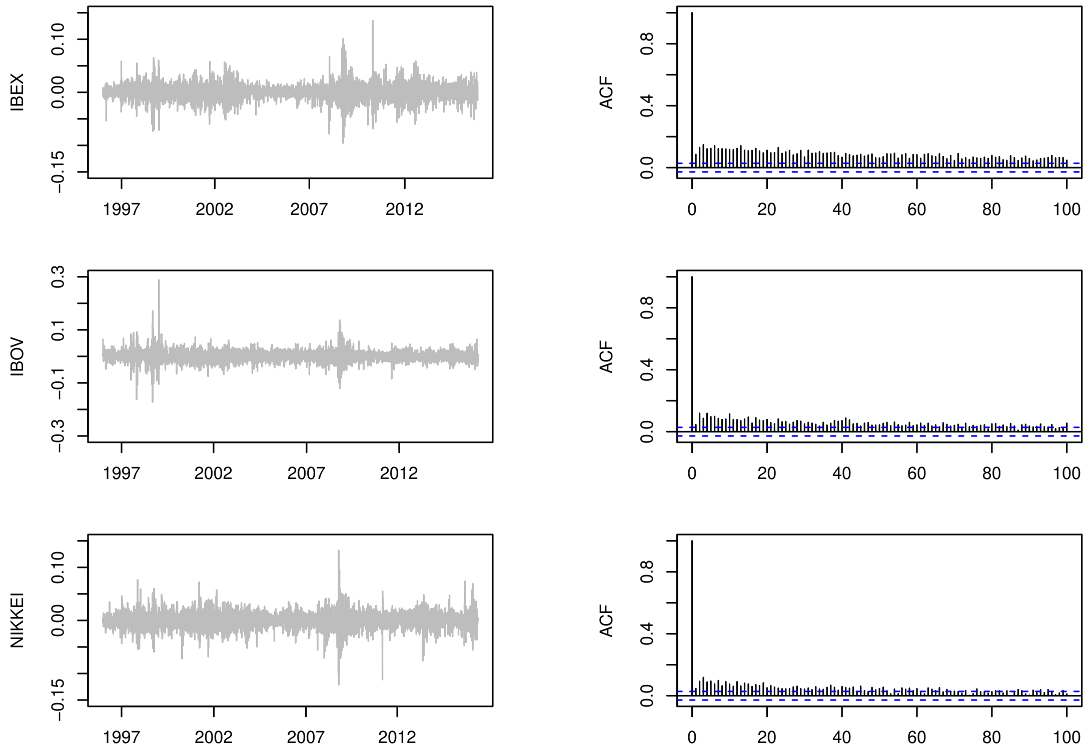

Figure 1: Plots of the first 5,000 returns (first column) and the empirical ACF function of the log-squared returns.

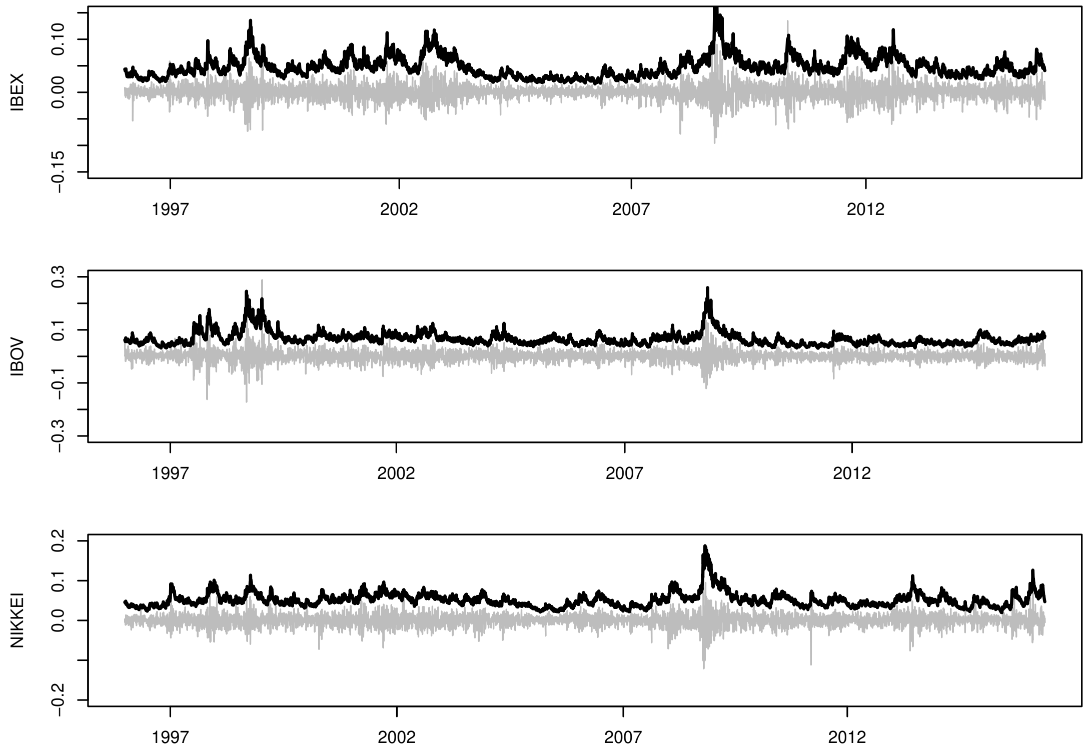

Figure 2: Plots of the first 5,000 returns and the predicted volatility (times 3) obtained from the A-LMSV model. The volatilities were estimated by ARFIMA$(0, d, 0)$ with $m = 3$, with the parameter estimates were presented in Table 2

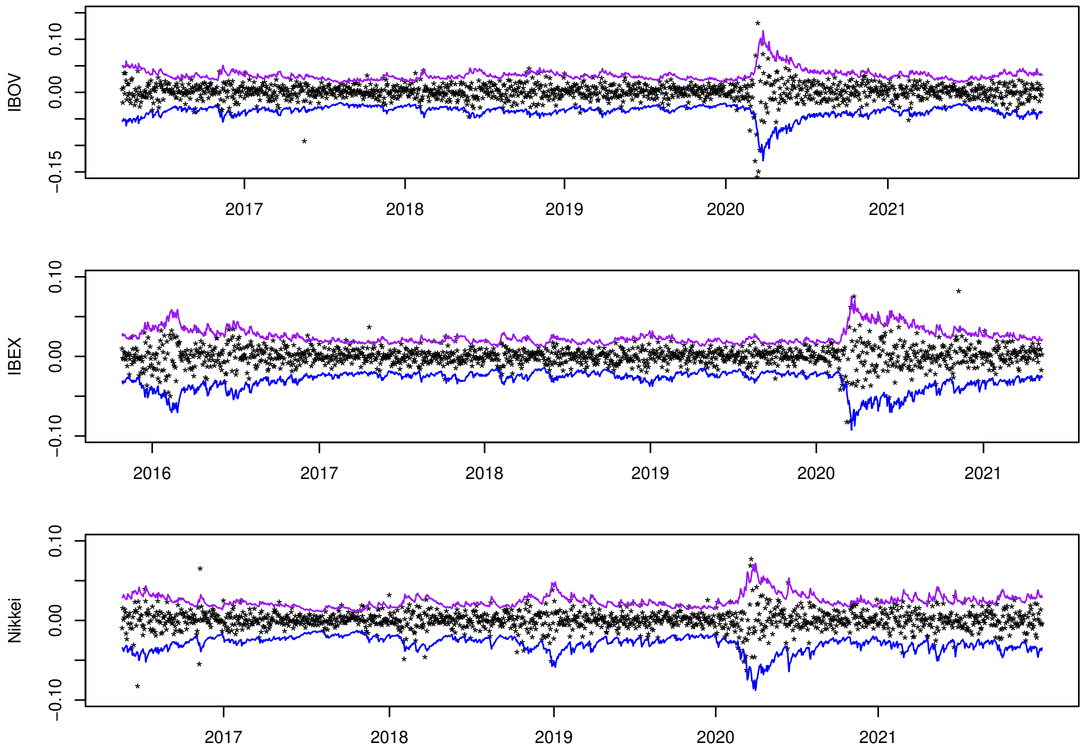

Figure 3: Plots of backtesting results for 99%-VaR, when h_t follows an ARFIMA$(0, d, 0)$ process and $m = 3$. The blue line represents the 99%-VaR for a long position while the purple line denotes a short position (with change of signal).