Integrated nested Laplace approximations for threshold stochastic volatility models

P. de Zea Bermudez

Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa and CEAUL - Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal.

J. Miguel Marín

Department of Statistics Universidad Carlos III de Madrid, Spain.

Håvard Rue

King Abdullah University of Science and Technology. South Arabia.

Helena Veiga*

Department of Statistics and Instituto Flores de Lemus, Universidad Carlos III de Madrid, Spain and BRU-IUL, Instituto Universitário de Lisboa, Portugal.

Abstract

The aim is to implement the integrated nested Laplace approximations (INLA), known to be very fast and efficient, for estimating the parameters of the threshold stochastic volatility (TSV) model. INLA replaces Markov chain Monte Carlo (MCMC) simulations with accurate deterministic approximations. Weakly informative proper priors are used, as well as Penalizing Complexity (PC) priors. The simulation results favor the use of PC priors, specially when the sample size varies from small to moderate. For these sample sizes, PC priors provide more accurate estimates of the model parameters. However, as sample size increases, both types of priors lead to similar estimates of the parameters. The estimation method is applied to six series of returns, including stock market, commodity and cryptocurrency returns, and its performance is assessed, by means of in-sample and out-of-sample approaches; the forecasting of one-day-ahead volatilities is also carried out. The empirical results support that the TSV is the model that generally fits the best to the series of returns and most of the times ranks the first in terms of forecasting one-day-ahead volatility, when compared to the symmetric stochastic volatility model.

Keywords: INLA, PC priors, Threshold Stochastic Volatility Model.

JEL-Classifications: C13; C32; C52; C58

*Corresponding author

Email address: mhveiga@est-econ.uc3m.es. (Helena Veiga)

1. Introduction

Stochastic volatility models have been widely proposed in the literature to cope with the main empirical features of financial time series. Very recent examples are the models suggested by Asai et al. (2017) and Mao et al. (2017, 2020) that accommodate a general asymmetric function for volatility; see also Turkman et al. (2014). Asymmetric effects have been traditionally dealt with either by considering a negative correlation between returns and future volatility (Harvey and Shephard, 1996) or by allowing the parameters of the log-volatility equation to differ according to the sign of the lagged returns (Breidt, 1996; So et al., 2002).

For more examples on the threshold stochastic volatility; see Asai and McAleer (2004, 2005, 2011), Chen et al. (2008, 2013), Elliott et al. (2011), Ghosh et al. (2015), Wu and Zhou (2015) and Wirjanto et al. (2016) among others.

This paper focuses on the second type of volatility asymmetry and proposes to use INLA to estimate the parameters and the latent volatility of TSV models. Stochastic volatility (SV) models are very flexible and easily accommodate the empirical features of the data. However, parameter estimation is not easy due to the intractable form of the likelihood function. Over the last twenty years many methods have been developed, such as the generalized method of moments (Melino and Turnbull, 1990; Sørensen, 2000), quasi-maximum likelihood (Harvey et al., 1994), simulated maximum likelihood (Danielsson, 1994; Sandmann and Koopman, 1998), approximate maximum likelihood (Fridman and Harris, 1998), maximum likelihood that relies on a nested structure of filters (Karamé, 2018) and MCMC procedures (Shephard and Pitt, 1997; Kim et al., 1998; Omori et al., 2007; Omori and Watanabe, 2008; Sakaria and Griffin, 2017). Note that the method proposed by Shephard and Pitt (1997) omits a term in the posterior density of a block of state errors, given the parameters and the other blocks, that is later corrected in Watanabe and Omori (2004). Recently, de Zea Bermudez et al. (2020) propose using data cloning to estimate the parameters of several SV models. Data cloning enables the estimation of the model parameters by maximum likelihood and the computation of the corresponding asymptotic variances through MCMC; see Lele et al. (2007, 2010).

The importance of MCMC algorithms in Statistics is nowadays paramount. Solutions to problems which were computationally unsolvable a few years ago are possible now. For example, Bayesian inference could not really be performed from an applied point of view, apart from the implementation of a very reduced number of academic-type models. However, the incredible benefits of using MCMC methods, such as the celebrated Metropolis-Hastings algorithm and the Gibbs sampler, come along with a very significant drawback. It is generally known that computer codes that implement MCMC algorithms can really be time consuming. Moreover, the assessment of the chains convergence can sometimes be a challenge. INLA, Rue et al. (2009), was conceived to provide Bayesian Statistics users with a set of tools which would enable to reduce the computation time while not affecting the accuracy of the results. Rue et al. (2009) establish the fundamental theory implemented in INLA. It lies on the fact that a wide range of models can be expressed as latent Gaussian models (LGM); see Rue and Held (2005). In fact, these models may be rewritten in such a way, that they possess a latent (unobserved) structure which can be represented by a Gaussian Markov Random Field (GMRF). INLA is used in a very extensive range of Bayesian applications, varying from time series and hierarchical models to space and space-time modeling (Blangiardo and Cameletti, 2015; Krainski et al., 2019; Gómez-Rubio, 2020; Amaral Turkman et al., 2021). Recently, Gómez-Rubio and Rue (2018) propose

embedding MCMC algorithms within INLA in order to extend the set of models that can be implemented. SV models can be fitted in INLA by means of the functions `stochvol`, `stochvol.t` and `stochvolsn` that use Gaussian, Student-t and skew-normal distributions, respectively. INLA functions have been specifically adapted by Håvard Rue for dealing with TSV and symmetric SV models. However the implementation of the asymmetric SV model by Harvey and Shephard (1996) is still not possible. It is well known that regarding modeling real series of returns like stock market returns, asymmetric volatility models show a better performance than standard symmetric models. The fact that INLA still cannot deal with asymmetric volatility models, as the one proposed by Harvey and Shephard (1996), can be a drawback for tackling some real situations in financial time series modeling.

The Bayesian approach is evaluated using PC priors, proposed by Simpson et al. (2017), and the traditional families of priors, e.g., the gamma for a scale parameter. Initially, the performance of the estimation approach is assessed through PC priors that convey sound information about the TSV model parameters; the persistence of the model in accordance with the financial literature is one of such parameters. In a second part, the estimation procedure using PC priors is evaluated in a slightly less informative scenario. Finally, traditional vague priors are considered with the intent of "letting the data speak".

The accomplishment of INLA for estimating the parameters of TSV models is assessed by means of a simulation study. Samples of returns are used with sizes ranging from 500 to 4000 and two distributions: the Gaussian and the Student-t. The values of the parameters in the simulation study reflects real financial data features, such as a very large and positive persistence. The results of the study show that INLA estimates of the TSV parameters are quite reliable in finite samples, except the degrees of freedom of the Student-t distribution. As expected, the results are slightly better when more informative PC priors are considered. The differences between the traditional non-informative priors and the more informative PC priors are mostly observed for small and moderate sample sizes. The estimate of the degrees of freedom, which is known to be a difficult parameter to estimate, is particularly more accurate for small sample sizes when PC priors are considered.

The performance of the INLA in estimating the parameters of the symmetric SV model with errors following a skew-normal distribution is also analyzed. The simulation results show that INLA estimates obtained for the symmetric SV parameters are quite reliable in finite samples. This model is used as a benchmark in the empirical application in order to evaluate the potential improvement of the TSV model for the volatility forecasting. Furthermore, the finite sample properties of the INLA estimator are compared to the MCMC estimator of the TSV parameters. INLA proves to be quite reliable and requires far less computation time when compared to MCMC.

The implementation of TSV models in INLA is demonstrated using three sets of daily time series. The first set consists of two daily times series of international stock market indexes, the S&P 500 and the Nikkei 225; the second set is composed of two time series of futures prices of two important commodities, gold and oil; the third set incorporates data of two well known cryptocurrencies, the Bitcoin and the Ethereum. These applications involve the volatility estimation in-sample and the one-day-ahead volatility forecasting. The empirical results indicate that the threshold volatility asymmetry seems to be relevant to the estimation and forecasting of the volatility of financial time series returns.

The use of INLA for estimating threshold stochastic volatility models is the major contribu-

tion of this paper. The performance of the estimation method is assessed by using informative PC priors, less informative PC priors and "vague" standard priors. The use of PC priors produces reliable estimates of the model parameters for small and moderate sample sizes. From an application point of view, it is possible to conclude that the one-day-ahead volatilities are well predicted by the model.

The paper is organized as follows: Section 2 describes the TSV model. The first part of Section 3 reviews basic concepts underlying INLA, as well as how to express TSV models in terms of a LGM. In the second part of Section 3, PC priors are presented and briefly discussed. The simulation study is presented in Section 4 using different kinds of prior distributions for the parameters and hyperparameters. Finally, several real data applications are presented in Section 5, and Section 6 concludes the paper.

2. Model description

Let y_t be the return at time t, σ_t^2 its volatility, $h_t \equiv \log \sigma_t^2$ and ϵ_t be an independent and identically distributed (IID) sequence with mean zero and variance one. The TSV model is given as

$$y_t = \sigma_t \epsilon_t = \exp(h_t/2)\epsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

where the latent (unobserved) variable at time t is the log-volatility given by:

$$h_t = \log \sigma_t^2 = \mu + \phi h_{t-1} + \delta I(\epsilon_{t-1} < 0) + \eta_t, \tag{2}$$

where $I(.)$ is an indicator function that takes the value one when the standardized return is negative in $t-1$ and zero otherwise. The log-volatility disturbance η_t is a Gaussian white noise with mean zero and variance σ_η^2, and it is independent of ϵ_t for all leads and lags. In situations where the financial returns are known to be heavy-tailed, ϵ_t may be modelled as a Student-t. The parameter ϕ, $-1 < \phi < 1$, controls the persistence of the volatility; values of ϕ as large as or even larger than 0.90 are quite common in real applications. Finally, the asymmetric response of the volatility to shocks is often captured by $\delta > 0$. Yet, $\delta = 0$ reduces the model to the symmetric SV of Taylor (1982) known in the literature as the ARSV model.

3. Bayesian approach using INLA

3.1. INLA basics

Traditionally, Bayesian inference is carried out by means of some MCMC algorithm, such as the Gibbs Sampler or the Metropolis-Hastings algorithm. However, it is understood that MCMC algorithms can be very inefficient since they can take very long time to run, specially as model complexity increases. Moreover, frequently there are difficult issues to address in order to assess chain convergence. INLA has been proposed by Rue et al. (2009). Contrary to MCMC algorithms, which are simulation-based methods, INLA uses numerical approximations and can be applied whenever the latent structure of the model can be represented as a GMRF. GMRF theory is extensively addressed in Rue and Held (2005) and provides very attractive computational properties. In this subsection a very general explanation of the fundamental aspects underlying the software INLA is given. Some further details can be found in the Appendix.

The Laplace method, one of the building blocks of INLA, and in its simplest formulation, enables to approximate a well-behaved and unimodal function $f(.)$ by a normal distribution, with mean x^* and variance $\sigma^{2^*} = -1/\frac{\partial^2 \log f(x)}{\partial x^2}|_{x=x^*}$, where x^* is the mode of $f(.)$ (see e.g. Blangiardo and Cameletti, 2015, for details).

The second basic ingredient of INLA is the definition of a latent Gaussian model. The explanation to be presented is fundamentally based on Blangiardo and Cameletti (2015). Let $(Y_1, Y_2, \ldots, Y_n)$ be an n-dimensional random vector of size n and $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ the corresponding observed data. Let us assume that Y_i conditional on ϕ_i, $i = 1, 2, \ldots, n$ are independent. Likewise in a generalized linear model framework, the ϕ_i are linked to the values of some covariates, $x_1, x_2, ..., x_M$, which (allegedly) affect the observations y_i, $i = 1, 2, \ldots, n$. Moreover, additional covariates, $z_1, z_2, ..., z_L$ that affect the outcome variables by means of functions, $f_l(.)$, might also be considered. These functions, $f_l(.)$, account for other kinds of random effects. All that can be done through a link function $g(.)$ such that

$$g(\phi_i) = h_i = \beta_0 + \sum_{j=1}^{M} \beta_j x_{ji} + \sum_{l=1}^{L} f_l(z_{li}). \tag{3}$$

Let us assume that $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2 \ldots . \beta_M, f_1(.), f_2(.), \ldots, f_L(.))$ is a GMRF, i.e., $\boldsymbol{\theta} \mid \boldsymbol{\Psi}$ is a multivariate normal distribution with mean vector zero and precision matrix (inverse of the covariance matrix), $Q(\Psi)$ with a Markov structure, depending on a vector of hyperparameters $\boldsymbol{\Psi} = (\Psi_1, \Psi_2, \ldots, \Psi_K)$. Given the data $\mathbf{y}$, the likelihood of $(\boldsymbol{\theta}, \boldsymbol{\Psi})$ is given by

$$L(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y}) = \prod_{i=1}^{n} p(y_i \mid \theta_i, \boldsymbol{\Psi}).$$

If $p(\boldsymbol{\Psi})$ is the prior distribution of $\boldsymbol{\Psi}$, then the joint posterior distribution of the parameters and hyperparameters is given by

$$p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y}) \propto p(\boldsymbol{\Psi}) p(\boldsymbol{\theta} \mid \boldsymbol{\Psi}) L(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y}) = p(\boldsymbol{\Psi}) p(\boldsymbol{\theta} \mid \boldsymbol{\Psi}) \prod_{i=1}^{n} p(y_i \mid \theta_i, \boldsymbol{\Psi}),$$

which is equivalent to

$$p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y}) \propto p(\boldsymbol{\Psi}) |Q(\boldsymbol{\Psi})|^{1/2} \exp\left(-\frac{1}{2} \boldsymbol{\theta}^T Q(\boldsymbol{\Psi}) \boldsymbol{\theta}) + \sum_{i=1}^{n} \log(p(y_i \mid \theta_i, \boldsymbol{\Psi})) \right),$$

where $|Q(\boldsymbol{\Psi})|$ refers to the determinant of the matrix $Q(\boldsymbol{\Psi})$ and $\boldsymbol{\theta}^T$ stands for the transpose operation on the of the vector $\boldsymbol{\theta}$.

The interest focuses on calculating the marginal posterior distribution of θ_i, for each i, which is given by:

$$p(\theta_i \mid \mathbf{y}) \;\; = \;\; \int p(\theta_i, \boldsymbol{\Psi} \mid \mathbf{y}) d\boldsymbol{\Psi} = \int p(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y}) p(\boldsymbol{\Psi} \mid \mathbf{y}) d\boldsymbol{\Psi}, \tag{4}$$

and also on the marginal posterior distribution of each hyperparameter Ψ_k,

$$p(\Psi_k \mid \mathbf{y}) = \int p(\boldsymbol{\Psi} \mid \mathbf{y}) d\boldsymbol{\Psi}_{-k}, \tag{5}$$

where, as usual, $\boldsymbol{\Psi}_{-k}$ represents all the components of the vector $\boldsymbol{\Psi}$ but the k^{th}.

Looking at (4) and (5), it can clearly be seen that, in order to obtain the posterior marginals of each θ_i and each Ψ_k, the distributions $p(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y})$ and $p(\boldsymbol{\Psi}_k \mid \mathbf{y})$ need to be derived. Basically, what INLA does is to approximate $p(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y})$ by Laplace approximation and $p(\boldsymbol{\Psi}_k \mid \mathbf{y})$ by numerical integration. Some further details can be found in the Appendix.

3.2. Volatility models with INLA

INLA has been scarcely used for estimating the parameters of SV models (see Martino et al., 2011; Ehlers and Zevallos, 2015, for some examples on the autoregressive SV model). The TSV model given in (1) and (2) with $\epsilon_t N(0,1)$ can be expressed as follows:

$$y_t \mid h_t \sim N(0, e^{h_t}),\ t = 1, 2, \ldots$$

with $h_t = \mu + \phi h_{t-1} + \delta I(\epsilon_{t-1} < 0) + \eta_t$. Therefore,

$$h_t \mid h_{t-1}, \mu, \phi, \delta, \tau_\eta \sim N(\mu + \phi h_{t-1} + \delta I(\epsilon_{t-1} < 0), 1/\tau_\eta),$$

where $\tau_\eta = 1/\sigma_\eta^2$.

For T returns, let $\boldsymbol{\theta} = (h_1, h_2, \ldots, h_T)$ be the latent field and $\boldsymbol{\Psi} = (\mu, \phi, \delta, \tau_\eta)$. The distribution of $\boldsymbol{\theta} \mid \boldsymbol{\Psi}$ is $N(\mathbf{0}, \boldsymbol{Q}^{-1}(\boldsymbol{\Psi}))$, where $\boldsymbol{Q}$ is the precision matrix.

Martino et al. (2011) and Ehlers and Zevallos (2015) consider the following prior distributions for μ, ϕ and η_t:

$$\mu \sim N(a, b^{-1}), \quad \text{logit}\left(\frac{\phi+1}{2}\right) \sim N(c, d^{-1}) \quad \text{and} \quad \log(\tau_\eta) \sim \text{log-gamma}(e, f), \tag{6}$$

where a, $c \in \mathbb{R}$ are location parameters, b, $d \in \mathbb{R}^+$ are precision parameters and e and f are real positive numbers. The authors chose these hyperparameters to reflect absence of prior information about the model parameters. Specifically, both Martino et al. (2011) and Ehlers and Zevallos (2015) select a very low value for the precision parameter of the normal distribution so that the resulting distribution resembles an uniform distribution in the interval $(0, 1)$. In the present work, a different prior for ϕ is considered. The INLA function `betacorrelation` is used, which is a $Beta(a, b)$ with $a > 0$ and $b > 0$, with support in $(-1, 1)$. The probability density function of such a $Beta(a, b)$ is given by

$$p(\phi \mid a, b) = 0.5 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \phi^{a-1}(1-\phi)^{b-1},\ -1 < \phi < 1.$$

When $a = b = 1$, the $Beta(a, b)$ defined above coincides with the uniform distribution in the interval $(-1, 1)$. Besides the prior distributions indicated in (6), PC Priors, proposed by Simpson

et al. (2017), are also considered. A brief review of the motivation, definition and interest of these priors is presented in subsection 3.3.

3.3. PC Priors

PC priors are specially suited for providing information about hyperparameters. They are always proper, which guarantees that the posterior distributions are also proper. They can cope with different degrees of information. They can be vague, mildly informative or very informative. The main motivation for devising PC priors is to provide the model with information without having to resort to an usual (and standard) parametric family, for instance the gamma family for a scale parameter. They just require the specification of two quantities: a threshold u and a probability α. The pair (u, α) is such that

$$P(\xi > u) = \alpha,$$

being ξ the parameter of interest and $0 < \alpha < 1$. Very frequently, the interest does not lie in the parameter ξ but in some appropriate function of the parameter, $t(\xi)$. These priors enable the INLA software to be developed in a more unified framework for Bayesian model implementation. In a nutshell, a user of Bayesian statistics does not need to know thorough details about the theory underlying PC priors; he/she just needs to express, in terms of probability statements, how likely certain are events located in the tail of the distribution.

The degree of information that the expert wants to convey to the distribution is dictated by the choice of (u, α). PC priors are based on four principles: the existence of a base model (a more flexible model than the one being considered); a constant rate penalization; a measure of complexity based on Kullback-Leibler divergence and a user-defined scaling in terms of (u, α) (see Simpson et al., 2017). Fundamentally, PC priors highly regard simplicity vs. complexity, in the sense that a simpler model is to be preferred than a more complex one.

Several PC priors for ϕ and τ are presented in Figure 1. For making these plots the INLA functions `inla.pc.dcor1` and `inla.pc.dprec` are used for ϕ and τ , respectively. It can be seen that the priors are quite informative. Different degrees of information will result from setting higher (lower) values of u and α.

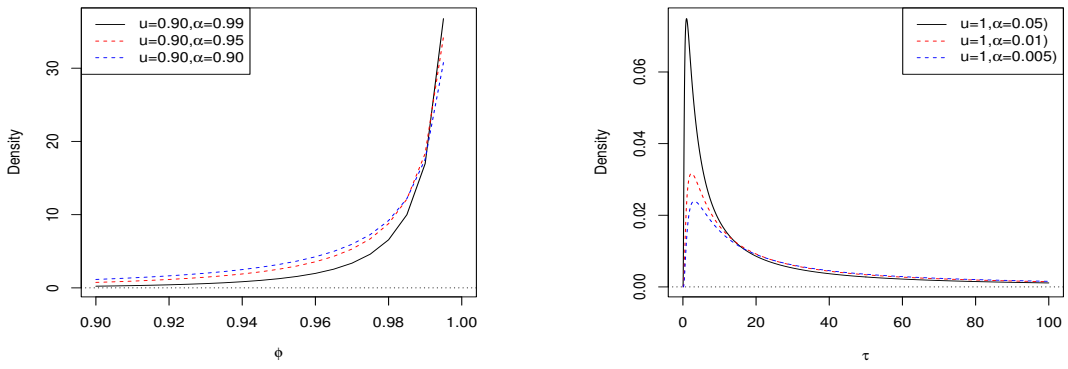

Figure 1. PC priors for persistence and precision parameters.

4. Simulation study

The performance of the INLA estimator of TSV parameters in finite samples is analyzed by Monte Carlo simulation. For this purpose, five hundred samples from the following TSV model are simulated:

$$\begin{aligned} y_t &= \exp(h_t/2)\epsilon_t, \\ h_t &= \mu + \phi h_{t-1} + \delta I(\epsilon_{t-1} < 0) + \eta_t. \end{aligned}$$

Sample sizes of $n = 500$, 1000, 2000 and 4000 have been considered.

The samples are simulated considering $\mu = 0$, $\phi = 0.98$, $\delta = 0.07$ and $\sigma_\eta^2 = 0.05$ (see Mao et al., 2017). Standardized returns can follow either a Gaussian or a Student-t distribution.

The first set of PC priors, the most informative one, for ϕ, σ_η and ν (the degrees of freedom of the Student-t model) satisfy the following probabilities:

$$P(\phi > 0.90) = 0.95; \quad P(\sigma_\eta > 1) = 0.01; \quad P(\nu > 15) = 0.5.$$

The priors for δ and for μ are $N(0, 1)$. The values of u and α considered for the prior distribution of ν are the default choices of the INLA `stochvol` function. The precision of the log volatility process is given by

$$\tau_h = \sigma_h^{-2} = \frac{1 - \phi^2}{\sigma_\eta^2 + 0.25\,\delta^2}.$$

The PC priors for ϕ and τ_h can be seen in Figure 1, whereas Figure 2 contains the plots of the PC prior for ν. This plot was obtained using INLA's function `inla.pc.ddof`.

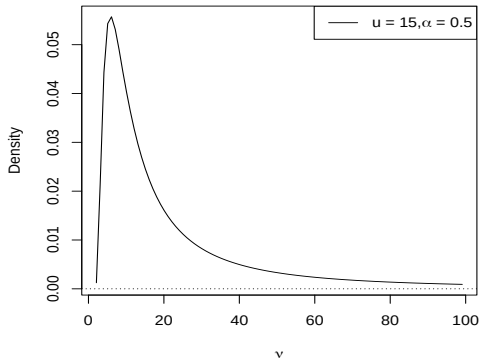

Figure 2. PC priors for the degrees of freedom.

The PC prior for ϕ reflects a very strong prior belief that the persistence of the TSV process is very high. Only a 5% change is left for values below or equal to 0.90. This probability is based on the prior knowledge that values of ϕ positive and close to 1 are quite common in financial applications (see, for example Asai and McAleer, 2005; Chen et al., 2008, 2013; Fan and Wang, 2013; Mao et al., 2017, among much others). The PC prior for ν reflects a strong belief that the Student-t distribution considered for the process ϵ_t is heavier than a Gaussian distribution.

Finally, mild to large variances are slightly more probable than high precisions. Thus, in the overall, the three PC priors considered vary from moderately informative to very informative.

The second set of PC priors, which are less informative, is defined as:

$$P(\phi > 0.90) = 0.70; \quad P(\sigma_\eta > 1) = 0.005; \quad P(\nu > 15) = 0.5.$$

As previously, the priors for δ and for μ are $N(0, 1)$. It should be mentioned that the Q matrix referred to in Subsection 3.2 is always given by $Q = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$ except for the second set of PC priors and $n = 500$. In that case, the variance of μ and δ is decreased and thus $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

The estimates of the parameters are the means of the posterior distributions. The standard deviation and the bounds of the credible intervals are summary statistics of the posterior marginal distribution of the parameters.

Table 1 reports the simulation results for Gaussian and Student-t standardized returns. The estimates obtained by INLA are very close to the true values for all the parameters, except for the degrees of freedom, even for sample sizes as small as $n = 500$. In regard to ν, the estimates improve as the sample size increases both in terms of bias and standard deviation even though really large values of n are required to obtain accurate results. Estimating the degrees of freedom is generally not an easy task and difficulties are reported by several authors, such as de Zea Bermudez et al. (2020). To highlight the speed associated to run the TSV model in INLA, the time required to complete one iteration when using standardized Student-t and Gaussian returns are presented in Table 4. The results were obtained using a Mac Pro processor 2.6 GHz, 6 core intel core i7.

Table 3 presents the simulation results using the priors considered by Martino et al. (2011) and Ehlers and Zevallos (2015) for μ and τ_h. The families of distributions chosen by these authors are indicated in (6). Regarding the prior distribution for ϕ, the uniform distribution in the interval $(-1, 1)$ is implemented here. Vague priors for the logarithm of the precision parameter and for μ have been selected. The prior for δ remains the same, i.e., N(0,1).

Table 1

Posterior means, standard deviations (sd) and 95% Credible Intervals (CI), obtained by INLA, using PC priors (n - sample size).

Parameters TSV model	$\boldsymbol{\mu}$	$\boldsymbol{\phi}$	$\boldsymbol{\tau_h}$	$\boldsymbol{\delta}$	$\boldsymbol{\nu}$
True values	**0**	**0.98**	**0.773**	**0.07**	**7**
			Normal standardized returns		
$n = 500$					
mean	0.029	0.965	1.308	0.069	
sd	0.040	0.015	0.448	0.077	
CI	(-0.049 , 0.108)	(0.928 , 0.987)	(0.63 , 2.365)	(-0.082 , 0.220)	
$n = 1000$					
mean	0.011	0.974	1.020	0.073	
sd	0.027	0.009	0.263	0.053	
CI	(-0.043 , 0.064)	(0.954 , 0.988)	(0.583 , 1.609)	(-0.031 , 0.176)	
$n = 2000$					
mean	0.005	0.977	0.897	0.072	
sd	0.019	0.006	0.179	0.037	
CI	(-0.033 , 0.042)	(0.965 , 0.987)	(0.584 , 1.284)	(0.000 , 0.145)	
$n = 4000$					
mean	0.003	0.979	0.847	0.070	
sd	0.013	0.004	0.127	0.026	
CI	(-0.024 , 0.029)	(0.971 , 0.986)	(0.616 , 1.115)	(0.019 , 0.120)	
			Student-t standardized returns		
$n = 500$					
mean	0.035	0.962	1.499	0.079	12.788
sd	0.042	0.017	0.589	0.081	9.430
CI	(-0.048 , 0.118)	(0.921 , 0.985)	(0.664 , 2.926)	(-0.080 , 0.238)	(4.845 , 36.645)
$n = 1000$					
mean	0.023	0.971	1.112	0.07	9.446
sd	0.029	0.01	0.296	0.057	3.622
CI	(-0.034 , 0.081)	(0.949 , 0.986)	(0.631 , 1.788)	(-0.041 , 0.181)	(5.141 , 18.678)
$n = 2000$					
mean	0.013	0.976	0.959	0.074	7.937
sd	0.020	0.006	0.194	0.039	1.664
CI	(-0.027 , 0.053)	(0.962 , 0.986)	(0.625 , 1.384)	(-0.004 , 0.151)	(5.435 , 11.902)
$n = 4000$					
mean	0.010	0.978	0.879	0.071	7.393
sd	0.014	0.004	0.134	0.028	1.007
CI	(-0.018 , 0.038)	(0.970 , 0.986)	(0.639 , 1.163)	(0.017 , 0.125)	(5.721 , 9.664)

The results presented in Tables 1–3 show that the benefits of using PC priors are mostly seen when the sample sizes vary from small to moderate. This conclusion is not surprising due to the fact that the PC priors considered in the simulation study are more informative than the priors used by Martino et al. (2011) and Ehlers and Zevallos (2015). This circumstance

Table 2

Posterior means, standard deviations (sd) and 95% Credible Intervals (CI), obtained by INLA, using light PC priors (n - sample size).

Parameters TSV model	$\boldsymbol{\mu}$	$\boldsymbol{\phi}$	$\boldsymbol{\tau_h}$	$\boldsymbol{\delta}$	$\boldsymbol{\nu}$
True values	**0**	**0.98**	**0.773**	**0.07**	**7**
			Normal standardized returns		
$n = 500$					
mean	0.038	0.959	1.406	0.070	
sd	0.041	0.018	0.458	0.078	
CI	(-0.041, 0.118)	(0.916 , 0.985)	(0.711 , 2.486)	(-0.083 , 0.222)	
$n = 1000$					
mean	0.014	0.972	1.072	0.073	
sd	0.027	0.009	0.264	0.053	
CI	(-0.04 , 0.068)	(0.952 , 0.987)	(0.632 , 1.662)	(-0.031 , 0.177)	
$n = 2000$					
mean	0.006	0.976	0.925	0.073	
sd	0.019	0.006	0.178	0.037	
CI	(-0.032 , 0.044)	(0.964 , 0.986)	(0.612, 1.31)	(0.000 , 0.145)	
$n = 4000$					
mean	0.003	0.979	0.862	0.070	
sd	0.013	0.004	0.127	0.026	
CI	(-0.023 , 0.03)	(0.970 , 0.986)	(0.632 , 1.129)	(0.019 , 0.120)	
			Student-t standardized returns		
$n = 500$					
mean	0.054	0.953	3.573	0.077	14.061
sd	0.044	0.022	33.856	0.085	11.123
CI	(-0.032 , 0.141)	(0.898 , 0.983)	(0.667 , 16.300)	(-0.089 , 0.243)	(5.023 , 42.051)
$n = 1000$					
mean	0.027	0.969	1.157	0.07	9.556
sd	0.029	0.010	0.299	0.057	3.742
CI	(-0.031 , 0.085)	(0.945 , 0.985)	(0.672 , 1.838)	(-0.042 , 0.182)	(5.159 , 19.104)
$n = 2000$					
mean	0.015	0.975	0.987	0.074	7.959
sd	0.020	0.006	0.194	0.039	1.682
CI	(-0.025 , 0.055)	(0.961 , 0.985)	(0.653 , 1.411)	(-0.003 , 0.151)	(5.439 , 11.974)
$n = 4000$					
mean	0.011	0.978	0.894	0.071	7.398
sd	0.014	0.004	0.133	0.028	1.008
CI	(0.017 , 0.039)	(0.969 , 0.985)	(0.654 , 1.177)	(0.017 , 0.126)	(5.724 , 9.671)

is obviously more important when the available data are scarce and consequently the posterior distributions of the parameters rely more on the information dictated by the priors. In these cases the estimates of the parameters are more accurate than those obtained using vague informative priors belonging to standard families of distributions. For large sample sizes the estimates of

Table 3

Posterior means, standard deviations (sd) and 95% Credible Intervals (CI), obtained by INLA, using standard priors (n - sample size).

Parameters	$\boldsymbol{\mu}$	$\boldsymbol{\phi}$	$\boldsymbol{\tau_h}$	$\boldsymbol{\delta}$	$\boldsymbol{\nu}$
TSV model					
True values	**0**	**0.98**	**0.773**	**0.07**	**7**
			Normal standardized returns		
$n = 500$					
mean	0.045	0.956	1.177	0.070	
sd	0.043	0.020	0.407	0.082	
CI	(-0.039 , 0.129)	(0.907 , 0.985)	(0.559 , 2.133)	(-0.09 , 0.231)	
$n = 1000$					
mean	0.013	0.973	0.951	0.073	
sd	0.028	0.010	0.254	0.054	
CI	(-0.042 , 0.069)	(0.951 , 0.988)	(0.531 , 1.519)	(-0.034 , 0.179)	
$n = 2000$					
mean	0.005	0.977	0.855	0.073	
sd	0.019	0.006	0.178	0.037	
CI	(-0.033 , 0.043)	(0.964 , 0.987)	(0.543 , 1.237)	(-0.001 , 0.146)	
$n = 4000$					
mean	0.003	0.979	0.825	0.070	
sd	0.013	0.004	0.127	0.026	
CI	(-0.024 , 0.029)	(0.971 , 0.986)	(0.594 , 1.092)	(0.019 , 0.121)	
			Student-t standardized returns		
$n = 500$					
mean	0.195	0.898	1.401	0.078	16.049
sd	0.051	0.030	14.403	0.090	14.391
CI	(0.096 , 0.294)	(0.825 , 0.94)	(0.514 , 3.559)	(-0.099 , 0.254)	(5.144 , 51.835)
$n = 1000$					
mean	0.028	0.968	1.015	0.070	10.154
sd	0.030	0.011	4.304	0.059	4.302
CI	(-0.032 , 0.088)	(0.942 , 0.985)	(0.566 , 1.647)	(-0.045 , 0.186)	(5.280 , 21.169)
$n = 2000$					
mean	0.014	0.976	0.910	0.074	8.073
sd	0.021	0.006	1.758	0.040	1.757
CI	(-0.027 , 0.056)	(0.961 , 0.986)	(0.584 , 1.325)	(-0.006 , 0.153)	(5.457 , 12.28)
$n = 4000$					
mean	0.010	0.978	0.853	0.071	7.456
sd	0.014	0.004	1.029	0.028	1.029
CI	(-0.018 , 0.039)	(0.969 , 0.986)	(0.614 , 1.135)	(0.017 , 0.126)	(5.753 , 9.782)

Table 4
Time (sec) to process a sample - Student-t and Gaussian models

Time (sec)		Sample size
Student-t model	Gaussian model	n
2.1	2.3	500
4.1	4.5	1000
7.2	5.2	2000
10.1	10.4	4000

the parameters obtained with the two sets of priors are very similar, except for the degrees of freedom of the Student-t distribution. One of the main advantages of using PC priors seems to be the estimation of the degrees of freedom of the Student-t distribution.

The skew-normal distribution for the standardized returns has also be examined. Instead of the TSV model the ARSV model has been considered, a decision reached for two reasons. First, the TSV, an asymmetric volatility model, is compared to the basic ARSV that does not take into account the asymmetric response of the volatility to positive and negative shocks of the same magnitude. It must be kept in mind that ARSV is obtained from the TSV model by making $\delta = 0$. Second, the implementation of the TSV-skew-normal is still not possible in INLA. Table 5 reports the results obtained. The INLA estimates of the ARSV-skew-normal parameters are accurate in finite samples, in particular when the sample size increases.

Finally, in the Appendix B simulation results are reported using MCMC to obtain estimates of the parameters of the TSV model with normal and Student-t standardized returns; see Table A. The main conclusion is that, whenever is possible to implement, INLA provides very reliable estimates of the TSV model in finite samples at an incomparable speed.

Table 5

Posterior means, standard deviations (sd) and 95% Credible Intervals (CI), obtained by INLA, using PC priors (n - sample size).

Parameters ARSV model	$\boldsymbol{\mu}$	$\boldsymbol{\phi}$	$\boldsymbol{\tau_h}$	$\boldsymbol{\alpha}$
True values	**0**	**0.98**	**0.773**	**-0.20**
	Skew-normal standardized returns			
$n = 500$				
mean	0.002	0.962	1.665	-0.146
sd	0.095	0.015	0.504	0.114
CI	(-0.184 , 0.188)	(0.927 , 0.984)	(0.892 , 8.011)	(-0.370 , 0.079)
$n = 1000$				
mean	0.008	0.971	1.229	-0.159
sd	0.093	0.008	0.279	0.088
CI	(-0.175 , 0.190)	(0.953 , 0.985)	(0.762 , 1.852)	(-0.331 , 0.013)
$n = 2000$				
mean	-0.006	0.975	1.015	-0.179
sd	0.089	0.006	0.180	0.065
CI	(-0.180 , 0.168)	(0.963 , 0.985)	(0.697 , 1.402)	(-0.304 , -0.050)
$n = 4000$				
mean	-0.008	0.977	0.911	-0.189
sd	0.084	0.004	0.125	0.046
CI	(-0.173 , 0.156)	(0.969 , 0.984)	(0.683 , 1.173)	(-0.280 , -0.097)

5. Empirical application

The symmetric ARSV and TSV models are estimated in six different daily series: two are international stock market indexes (S&P 500 and Nikkei 225), two are commodity futures prices series (gold and oil), and the last two correspond to quotations of two well known cryptocurrencies (Bitcoin and Ethereum). S&P 500 and Nikkei 225 returns cover the period between January 5, 2000 and July 22, 2020, with a total number of observations of 5170 and 5036, respectively. Gold and oil futures returns range from February 29, 2000 (oil, March 22, 2000) until July 22, 2020 with a total number of observations of 5111 and 5097, respectively. Finally, Bitcoin and Ethereum returns range from September 17, 2014 (Ethereum, August 7, 2015) to July 23, 2020, with a total number of observations of 2136 and 1812, respectively.

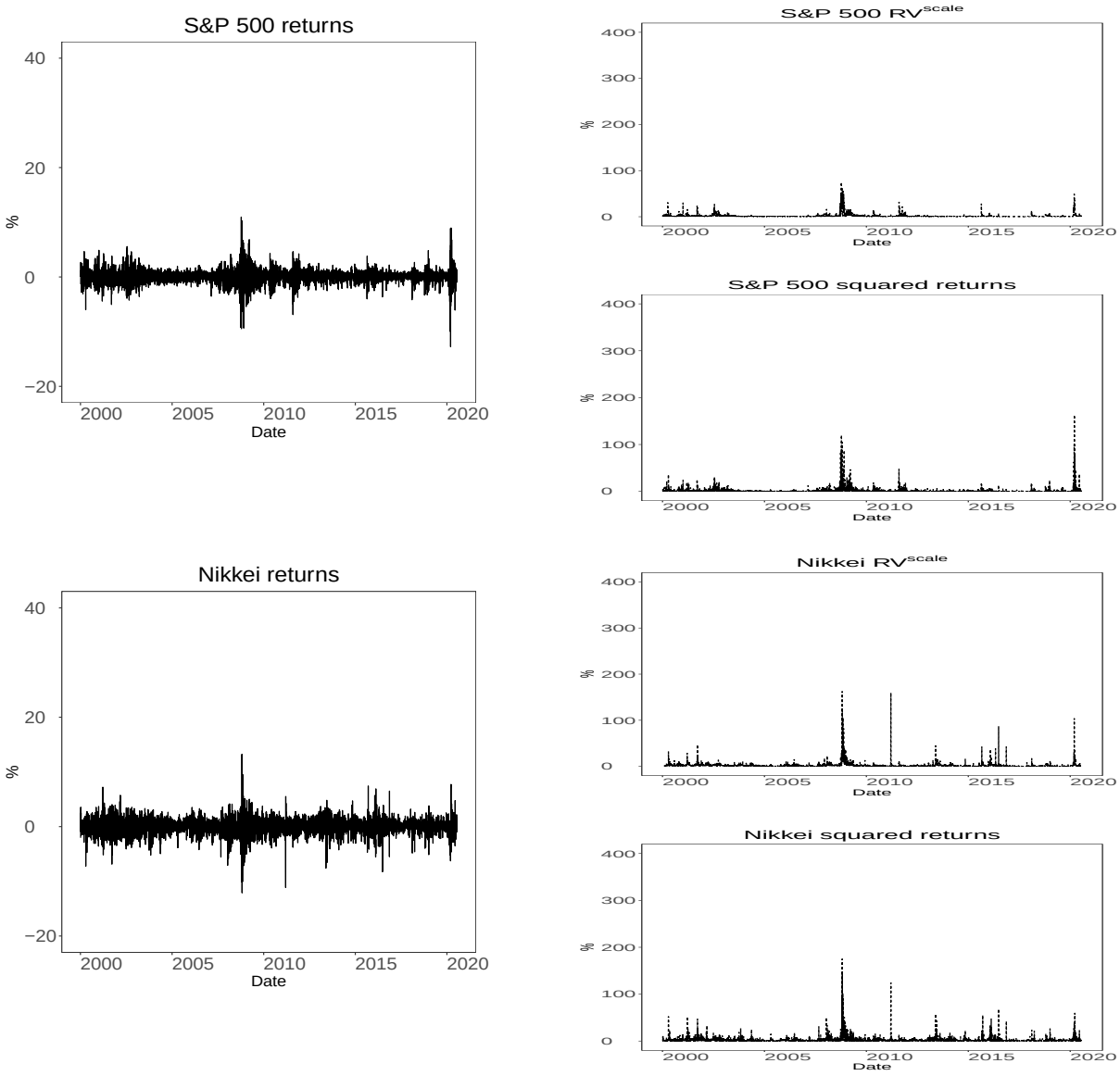

Figure 3. Series of returns, RV^{scale}, RV^{on} and squared returns in percentage.

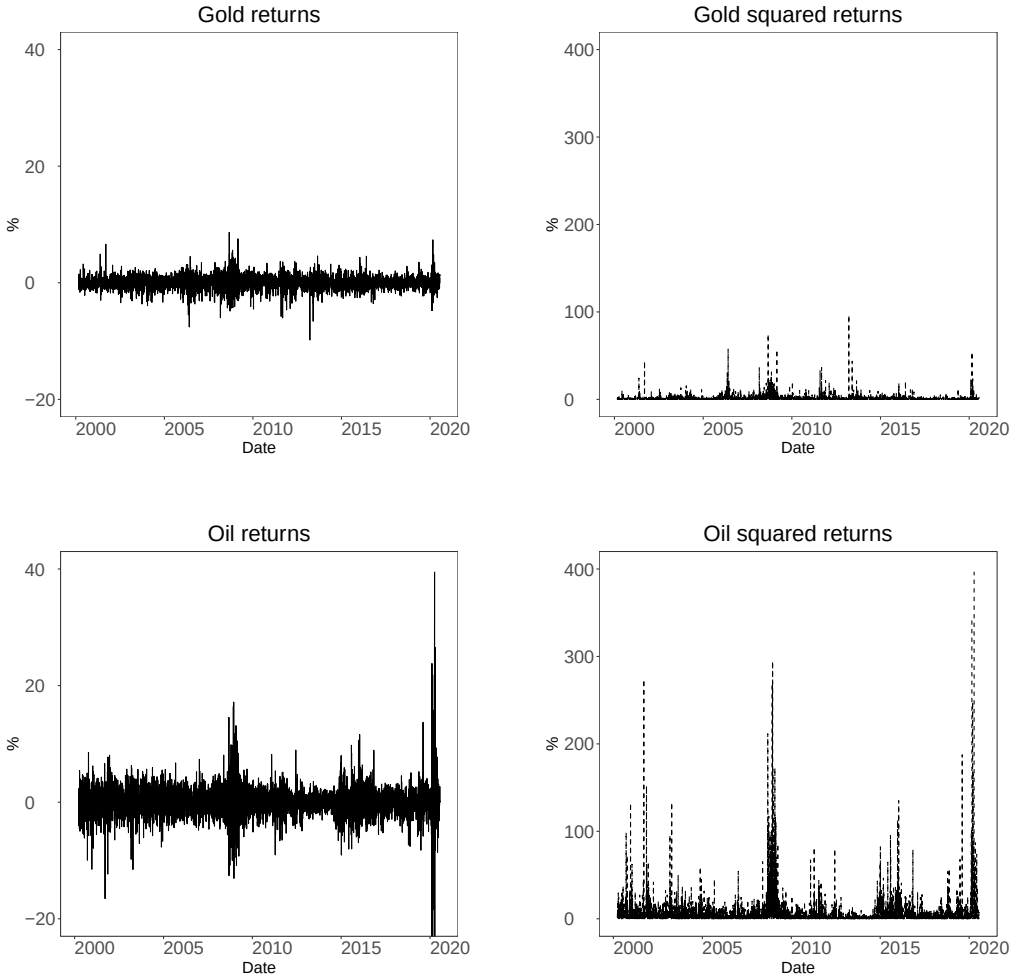

Figure 3. Series of returns and squared returns in percentage (cont.).

Figure 3 shows the time series of returns, the squared returns and the realized volatility for the S&P 500 and Nikkei 225. Squared returns are often used as a proxy of volatility in the literature, although there is evidence that good volatility forecasts might have poor predictive power for daily squared returns, see Andersen and Bollerslev (1998) and Hansen and Lunde (2006) for details on this issue. Notice that stock market returns show a period of high volatility that corresponds to the last global financial crisis. In particular, S&P 500 returns show a period of high volatility at the end of the sample that corresponds to the effects of the Coronavirus pandemic. The behavior of commodity returns is similar to that of stock market returns, specially oil returns. Notice also a huge increase in oil volatility at the end of the sample corresponding to oil negative prices due to a decrease in demand. Finally, cryptocurrency returns are the most volatile of all series of returns.

Table 6 reports the data descriptive statistics and the results of Jarque-Bera's test. The empirical distributions of returns are in general lepkurtic and skewed to the left. The exception is the distribution of Ethereum returns that is skewed to the right. Normality is rejected for all series.

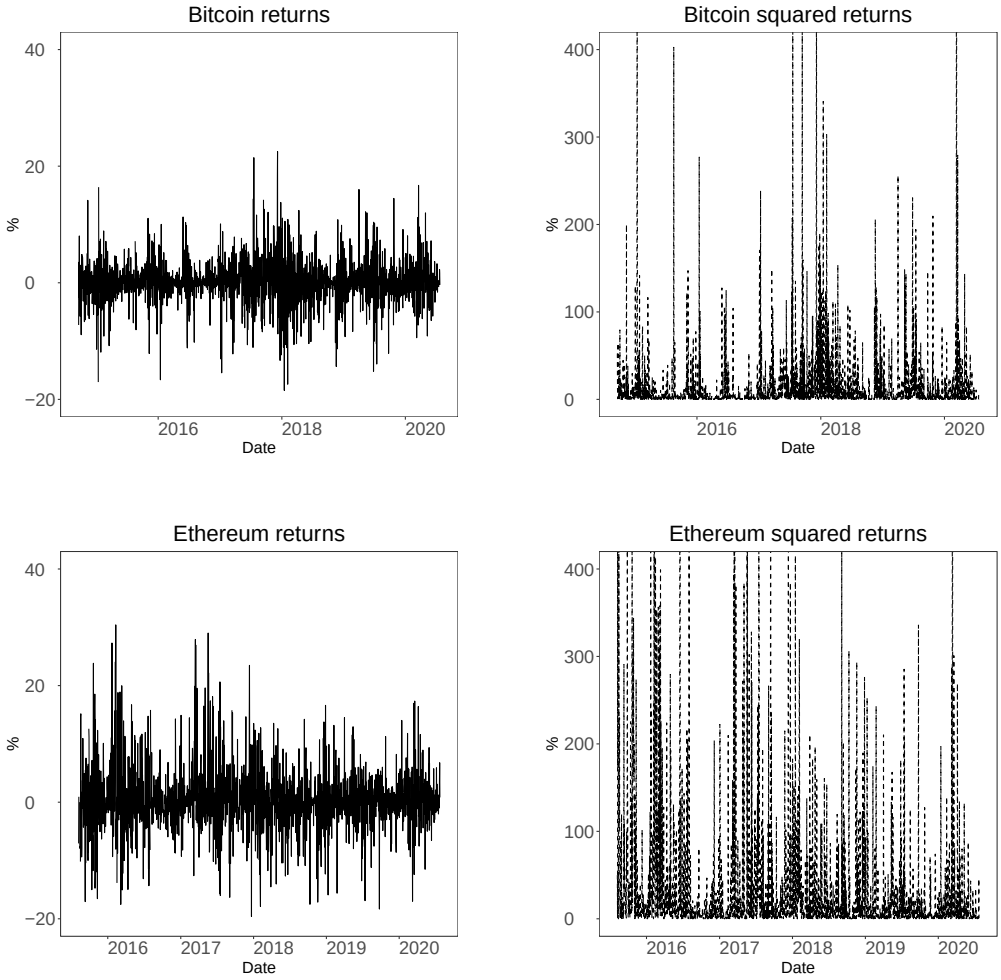

Figure 3. Series of returns and squared returns in percentage (cont.).

Table 6
Returns' summary statistics

Series	Mean	SD	Kurtosis	Skewness	Jarque-Bera	p-value
SP500	0.016	1.258	11.076	-0.383	26493	0.000
Nikkei	0.004	1.502	6.334	-0.385	8520.9	0.000
Gold	0.036	1.110	5.796	-0.194	7167.4	0.000
Oil	-0.0876	6.860	3132.4	-50.290	2.082e+09	0.000
Bitcoin	0.146	3.941	13.416	-0.953	16252	0.000
Ethereum	0.324	6.303	8.091	0.092	4910.8	0.000

5.1. Parameter estimation results

The parameters of the TSV and ARSV models are estimated using the previous PC priors and assume three possible distributions for the standardized returns.

Table 7 reports the results of the estimation of the parameters of the TSV and ARSV models. Regarding the TSV model, it can be seen that stock market returns present high volatility persistence (ϕ around 0.97), volatility asymmetry (negative returns increase the volatility the day after), and high degrees of freedom of the Student-t distribution. On the other hand, commodity futures returns seem to require fat-tailed distributions, present a high volatility persistence (larger than stock market returns) and volatility asymmetry. Yet negative returns of oil futures increase volatility in the following day while negative returns of gold futures decrease it. The oil volatility asymmetry is similar to that of stock market returns. In fact, due to the opening of commodity markets to financial investors, oil markets have become increasingly driven by flows of financial investors and less by their fundamentals, strengthening the links between oil and stock markets (see Silvennoinen and Thorp, 2013; Büyüksahin and Robe, 2014; Martín-Barragán et al., 2015). Furthermore, a different kind of volatility asymmetry is observed for gold futures returns. Positive gold returns today will increase tomorrow's volatility.

This phenomenon is known as "inverse leverage effect" or inverted asymmetric reaction to positive and negative shocks. Commodity prices are often influenced by macroeconomic fundamentals such as inflation, interest rates and industrial production (Hammoudeh and Yuan, 2008). Increases in commodity prices anticipate rises in inflation, leading to monetary policy tightening and augmentation of interest rates. Therefore, "*...investors interpret positive gold price changes as a signal for future adverse conditions and uncertainty in other asset markets. This introduces uncertainty in the gold market and thus higher volatility.*" (Baur, 2012). Moreover, high volatility persistence in the cryptocurrency returns that follow fat-tailed distributions is observed. Furthermore, Ethereum returns do not present volatility asymmetry while Bitcoin returns show an inverted asymmetric reaction; see Figure 3 and Baur and Dimpfl (2018) for similar results regarding Bitcoin returns.

Finally, the estimation of the ARSV model parameters, not including the volatility asymmetry, leads in general to an increase of the parameter estimate, that corresponds to the volatility persistence, and to a decrease of the degrees of freedom of the Student-t distribution. This suggests that the distribution of the standardized returns is more fat-tailed. Looking at the logarithm of the marginal likelihood, the results suggest that the TSV-Student-t outperforms the other models for the majority of the series of returns, except for the gold returns. The best in-sample model is the ARSV-Student-t model. Yet the difference between the TSV-Student-t and ARSV-t is small.

Table 7

Posterior means and standard deviations (sd) of TSV parameters using INLA. CI stands for 95% credible intervals and lml represents the log marginal likelihood. Standardized returns follow a $N(0,1)$ distribution.

	Normal			
Series	μ	ϕ	τ_h	δ
SP500				
mean	-0.159	0.973	1.651	0.325
sd	0.01	0.003	0.191	0.022
CI	(-0.18 , -0.139)	(0.966 , 0.979)	(1.298 , 2.047)	(0.282 , 0.367)
lml (1)	-6960.698			
lml (2)	-6958.572			
Nikkei				
mean	-0.096	0.965	1.916	0.222
sd	0.011	0.005	0.220	0.022
CI	(-0.118 , -0.074)	(0.955 , 0.975)	(1.515 , 2.379)	(0.179 , 0.266)
lml (1)	-8435.924			
lml (2)	-8433.798			
Gold				
mean	0.003	0.975	1.952	-0.015
sd	0.01	0.006	0.275	0.02
CI	(-0.016 , 0.022)	(0.963 , 0.985)	(1.465 , 2.545)	(-0.054 , 0.025)
lml (1)	-7242.944			
lml (2)	-7240.818			
Oil				
mean	-0.031	0.975	1.357	0.136
sd	0.01	0.004	0.163	0.021
CI	(-0.051 , -0.010)	(0.966 , 0.983)	(1.058 , 1.697)	(0.094 , 0.177)
lml (1)	-11195.95			
lml (2)	-11193.82			
Bitcoin				
mean	0.204	0.902	0.523	-0.037
sd	0.031	0.017	0.06	0.062
CI	(0.142 , 0.265)	(0.867 , 0.933)	(0.414 , 0.65)	(-0.158 , 0.084)
lml (1)	-5433.214			
lml (2)	-5431.089			
Ethereum				
mean	0.309	0.885	0.557	0.054
sd	0.038	0.023	0.065	0.069
CI	(0.234 , 0.384)	(0.836 , 0.925)	(0.440 , 0.695)	(-0.082 , 0.189)
lml (1)	-5551.211			
lml (2)	-5549.086			

lml (1) - log marginal likelihood (integration)
lml (2) - log marginal likelihood (Gaussian)

Table 7
(continued)

Posterior means and standard deviations (sd) of TSV parameters using INLA. CI stands for 95% credible intervals and lml represents the log marginal likelihood. Standardized returns follow a Student-t distribution.

			Student-t		
Series	μ	ϕ	τ_h	δ	ν
SP500					
mean	-0.156	0.975	1.696	0.320	24.295
sd	0.01	0.003	0.189	0.022	8.446
CI	(-0.176, -0.136)	(0.968, 0.980)	(1.338, 2.081)	(0.277, 0.362)	(13.466, 45.842)
lml (1)	-6957.335				
lml (2)	-6954.290				
Nikkei					
mean	-0.089	0.970	2.045	0.206	24.905
sd	0.011	0.005	0.261	0.021	9.659
CI	(-0.11, -0.068)	(0.96, 0.979)	(1.583, 2.606)	(0.164, 0.248)	(12.902, 49.641)
lml (1)	-8434.467				
lml (2)	-8431.423				
Gold					
mean	0.015	0.993	2.428	-0.033	5.628
sd	0.007	0.002	0.576	0.014	0.465
CI	(0.002, 0.029)	(0.989, 0.997)	(1.455, 3.701)	(-0.06, -0.005)	(4.808, 6.634)
lml (1)	-7143.408				
lml (2)	-7140.363				
Oil					
mean	-0.040	0.988	1.945	0.122	9.324
sd	0.008	0.003	0.327	0.017	1.126
CI	(-0.056, -0.024)	(0.982, 0.992)	(1.362, 2.64)	(0.089, 0.155)	(7.401, 11.818)
lml (1)	-11162.300				
lml (2)	-11159.260				
Bitcoin					
mean	0.096	0.977	0.633	-0.095	3.490
sd	0.021	0.007	0.126	0.043	0.320
CI	(0.056, 0.137)	(0.962, 0.988)	(0.416, 0.907)	(-0.180, -0.010)	(2.948, 4.199)
lml (1)	-5383.094				
lml (2)	-5380.050				
Ethereum					
mean	0.173	0.943	0.745	0.005	4.380
sd	0.029	0.014	0.143	0.056	0.740
CI	(0.115, 0.231)	(0.910, 0.966)	(0.471, 1.021)	(-0.105, 0.113)	(3.392, 6.236)
lml (1)	-5533.490				
lml (2)	-5530.446				

lml (1) - log marginal likelihood (integration)
lml (2) - log marginal likelihood (Gaussian)

Table 7
(continued)

Posterior means and standard deviations (sd) of ARSV parameters using INLA. CI stands for 95% credible intervals and lml represents the log marginal likelihood. Standardized returns follow a $N(0,1)$ distribution.

	Normal		
Series	μ	ϕ	τ_h
SP500			
mean	-0.073	0.980	0.936
sd	0.084	0.003	0.125
CI	(-0.236 , 0.093)	(0.974 , 0.986)	(0.709 , 1.198)
lml (1)	-7051.242		
lml (2)	-7051.591		
Nikkei			
mean	0.195	0.976	1.495
sd	0.078	0.004	0.208
CI	(0.036 , 0.344)	(0.967 , 0.984)	(1.118 , 1.933)
lml (1)	-8475.755		
lml (2)	-8476.100		
Gold			
mean	-0.058	0.974	2.025
sd	0.068	0.006	0.278
CI	(-0.189 , 0.077)	(0.962 , 0.984)	(1.531 , 2.621)
lml (1)	-7232.213		
lml (2)	-7232.551		
Oil			
mean	0.142	0.992	0.509
sd	0.106	0.002	0.091
CI	(-0.065 , 0.35)	(0.988 , 0.995)	(0.349 , 0.706)
lml (1)	-11227.440		
lml (2)	-11227.880		
Bitcoin			
mean	0.149	0.971	0.256
sd	0.105	0.006	0.038
CI	(-0.057 , 0.355)	(0.957 , 0.981)	(0.189 , 0.337)
lml (1)	-5458.725		
lml (2)	-5459.073		
Ethereum			
mean	0.076	0.985	0.165
sd	0.101	0.003	0.027
CI	(-0.123 , 0.275)	(0.978 , 0.990)	(0.118 , 0.223)
lml (1)	-5580.416		
lml (2)	-5580.711		

lml (1) - log marginal likelihood (integration)
lml (2) - log marginal likelihood (Gaussian)

Table 7
(continued)

Posterior means and standard deviations (sd) of ARSV parameters using INLA. CI stands for 95% credible intervals and lml represents the log marginal likelihood. Standardized returns follow a Student-t distribution.

	Student-t			
Series	μ	ϕ	τ_h	ν
SP500				
mean	-0.060	0.984	0.978	16.664
sd	0.085	0.003	0.139	4.501
CI	(-0.225 , 0.108)	(0.977 , 0.989)	(0.723 , 1.268)	(10.254 , 27.694)
lml (1)	-7045.550			
lml (2)	-7045.477			
Nikkei				
mean	0.180	0.982	1.593	15.613
sd	0.084	0.004	0.257	3.67
CI	(0.009 , 0.341)	(0.973 , 0.989)	(1.144 , 2.150)	(9.991 , 24.282)
lml (1)	-8469.487			
lml (2)	-8469.578			
Gold				
mean	-0.009	0.992	2.660	5.639
sd	0.081	0.002	0.541	0.463
CI	(-0.168 , 0.151)	(0.988 , 0.996)	(1.733 , 3.844)	(4.813 , 6.633)
lml (1)	-7133.154			
lml (2)	-7133.138			
Oil				
mean	0.095	0.996	0.589	9.336
sd	0.103	0.001	0.127	1.121
CI	(-0.108 , 0.296)	(0.994 , 0.998)	(0.372 , 0.867)	(7.412 , 11.808)
lml (1)	-11189.630			
lml (2)	-11189.600			
Bitcoin				
mean	0.067	0.992	0.294	3.484
sd	0.101	0.002	0.059	0.297
CI	(-0.132 , 0.266)	(0.987 , 0.995)	(0.193 , 0.422)	(2.978 , 4.145)
lml (1)	-5392.265			
lml (2)	-5392.269			
Ethereum				
mean	0.049	0.994	0.215	3.960
sd	0.101	0.002	0.045	0.450
CI	(-0.149 , 0.246)	(0.99 , 0.996)	(0.14 , 0.315)	(3.22 , 4.981)
lml (1)	-5544.046			
lml (2)	-5544.056			

lml (1) - log marginal likelihood (integration)
lml (2) - log marginal likelihood (Gaussian)

Table 7
(continued)

Posterior means and standard deviations (sd) of ARSV parameters using INLA. CI stands for 95% credible intervals and lml represents the log marginal likelihood. Standardized returns follow a skew-normal distribution with parameter α.

Series	μ	ϕ	τ_h	α
SP500				
mean	-0.071	0.981	0.960	-0.140
sd	0.083	0.003	0.130	0.041
CI	(-0.232 , 0.094)	(0.975 , 0.987)	(0.720 , 1.229)	(-0.220 , -0.060)
lml (1)	-7046.613			
lml (2)	-7046.700			
Nikkei				
mean	0.201	0.977	1.525	-0.107
sd	0.078	0.004	0.203	0.039
CI	(0.041 , 0.350)	(0.968 , 0.985)	(1.154 , 1.947)	(-0.183 , -0.030)
lml (1)	-8473.292			
lml (2)	-8473.434			
Gold				
mean	-0.054	0.975	2.060	-0.144
sd	0.067	0.005	0.284	0.035
CI	(-0.184 , 0.080)	(0.963 , 0.984)	(1.551 , 2.667)	(-0.213 , -0.075)
lml (1)	-7226.204			
lml (2)	-7226.199			
Oil				
mean	0.013	0.993	0.465	-0.171
sd	0.032	0.001	0.080	0.037
CI	(-0.049 , 0.075)	(0.990 , 0.996)	(0.321 , 0.632)	(-0.244 , -0.099)
lml (1)	-11218.940			
lml (2)	-11219.000			
Bitcoin				
mean	0.150	0.971	0.256	-0.194
sd	0.105	0.006	0.037	0.065
CI	(-0.055 , 0.355)	(0.958 , 0.981)	(0.191 , 0.334)	(-0.317 , -0.064)
lml (1)	-5456.381			
lml (2)	-5456.264			
Ethereum				
mean	0.079	0.985	0.165	0.116
sd	0.101	0.003	0.026	0.082
CI	(-0.121 , 0.278)	(0.978 , 0.990)	(0.118 , 0.221)	(-0.050 , 0.270)
lml (1)	-5580.665			
lml (2)	-5580.646			

lml (1) - log marginal likelihood (integration)
lml (2) - log marginal likelihood (Gaussian)

5.2. Volatility results

Figure 4 shows the posterior mean of the daily estimated volatility ($\hat{\sigma}_t$) in percentage (solid black line), together with the one-day-ahead volatility forecast (dashed red line) for the six series of returns. The estimated volatility is obtained from the model with the best in-sample goodness-of-fit, and the one-day-ahead forecast volatility is obtained from the model with the smallest value of the loss function proposed by Patton (2011) when using squared returns as volatility proxy, see Tables 7 and 8, respectively. The volatility estimates range from the beginning of the series until December 31, 2018, while the forecasts cover the period between the beginning of January, 2019 and the 22nd or 23rd of July 2020. A rolling window forecasting scheme is used with a fixed number of observations. The first window includes all the observations until December 31, 2018. The second window contains the second observation of the first window and one extra observation at the end, assuring that the number of observations is the same than that of the first window.

The in-sample period includes the last global financial crisis while the out-of-sample period incorporates the Coronavirus pandemic. During the last global financial crisis, stock markets were confronted with unprecedented declines that lead to a decrease of wealth, with costly social and economic consequences. All this uncertainty resulted in an increase of the estimated volatility around 2008; see Figure 4. On the other hand, the one-day-ahead volatility forecasts show huge picks around March 17, 24 and 25, 2020 for the S&P 500, Nikkei and gold returns, respectively, on April 22, 2020 for the oil returns and on March 13, 2020 for the cryptocurrency returns. On March 23, 2020, the S&P 500 index lost around 35% of its February 19, 2020 value. A decrease comparable in magnitude to the declines occurred during the last global financial crisis, and to those corresponding to the Black Monday back in 1987 (Lyócsa et al., 2020). Ashraf (2020) also finds evidence of a negative response of stock markets to the rise of COVID-19 cases. Furthermore, these declines spread to other stock markets causing a wave of contagion during the period of confinement (see Okorie and Lin, 2020). The evidence is unclear regarding gold and the cryptocurriencies. While gold acted as a safe commodity (Ji et al., 2020) during the first wave of the pandemic the evidence for the Bitcoin is controversial. Goodell and Goutte (2020) provides evidence that Bitcoin acted as a safe haven, but Conlon et al. (2020) and Corbet et al. (2020) provide an opposite evidence.

The measure of realized volatility that has been used is the realized kernel proposed by Barndorff-Nielsen et al. (2008, 2009) that minimizes the bias produced by the existence of market microstructure noise. This measure has been converted to an unbiased estimator for the close-to-close (whole day's) volatility by applying the correction proposed by Hansen and Lunde (2005), the scale realized volatility denoted here RV_t^{scale}. Yet, it should be mentioned that the RV^{scale} is only used in the case of the S&P 500 and Nikkei returns. The realized kernel data has been downloaded from the *Oxford Man Institute's Realized Library*; see `https://realized.oxford-man.ox.ac.uk/data/download`.

The mean squared error (MSE) and the QLIKE loss function proposed by Patton (2011) are two loss functions chosen for comparing the predictive ability of models; see, for instance, Takahashi et al. (2016) for the use of the latter loss function. The MSE is one of the most widely used loss functions in empirical applications. It satisfies the necessary condition for robustness since generates an optimal forecast equal to the conditional variance, and it also fulfils the sufficient condition of Hansen and Lunde (2006). The QLIKE loss function is also a robust loss function. However, the MSE is more sensitive to outliers while the QLIKE corrects

some of this sensitiveness by being more robust to outliers coming from the right side of the distribution. Patton (2011) states that the good properties of the loss functions are based on the fact that the proxies of volatility are conditionally unbiased estimators of the daily conditional variance, implying that the proxies of the volatility are conditionally unbiased estimators of the true volatility. Both the squared returns and the realized volatility are conditionally unbiased estimators of the daily conditional variance, but the RV is more efficient than the squared returns. The more efficient is the proxy of volatility used, less distortions causes on the loss functions; see Patton (2011) for details on these issues.

Models' statistical predictive accuracy is tested with the model confident set procedure (MCS) proposed by Hansen et al. (2011). The procedure consists of a sequence of statistical tests to construct the "Superior Set of Models" (SSM) at a given confidence level, $100(1-\alpha)\%$. Models in the SSM have statistically the same predictive accuracy, measured by a loss function at significant level α. Table 8 displays the 80% SSM obtained from comparing all the models (see Granger, 1996, that considers confidence levels of 50% and 80% adequate *to provide "warning" signals that the model is breaking down*).

Table 8 shows that according to the MSE loss function, and independently of the volatility proxy, all the models are in the SSM and therefore have statistically the same predictive ability. The conclusions are different for the majority of financial returns when the QLIKE loss function is used, and the volatility proxy is the squared returns. For the S&P 500, Nikkei 225, Bitcoin and Ethereum returns the TSV is the only model in the SSM. It should be kept in mind that the QLIKE is more robust regarding extreme values of the squared returns, and therefore, more adequate to evaluate the forecasting performance of models. Regarding the distribution of the standardized returns, TSV models with normal or Student-t standardized returns have statistically the same predictive ability except for the S&P 500 returns. In this case, the best model for forecasting the one-day-ahead volatility is the TSV-$N(0,1)$.

Regarding the realized volatility proxy, note that the TSV-$N(0,1)$ is the only model in the SSM for the S&P 500. The QLIKE loss function for the Nikkei 225 returns shows that the majority of models have the same predictive ability and that threshold stochastic volatility models are always ranked first.

Finally, both the in-sample and forecasting evaluations favor the TSV model. Threshold volatility asymmetry seems to be relevant for the estimation and forecasting of the volatility of financial time series returns, independently if the series are equity, commodities or cryptocurrencies.

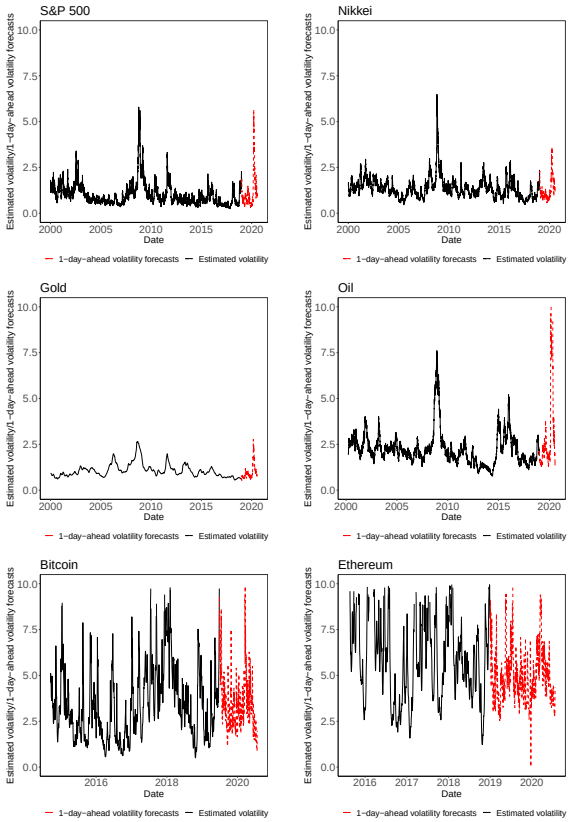

Figure 4. TSV estimated volatilities (solid line) and one-day-ahead volatility forecasts (dashed line).

Table 8: MCS procedure

Results obtained from the MCS with two different loss functions and two proxies of volatility, the mean squared error (MSE) and the QLIKE by Patton (2011), and 5000 bootstrap samples. - indicates the models that are excluded from the Superior Set of Models at the 80% confidence level. The numbers in the table correspond to the rankings inside of the SSM.

	MSE						**QLIKE**					
Models	SP500	Nikkei	Gold	Oil	Bitcoin	Ethereum	SP500	Nikkei	Gold	Oil	Bitcoin	Ethereum
	proxy of volatlity = squared returns											
TSV-$N(0,1)$	3	1	1	3	4	2	1	2	2	1	1	2
TSV-Student-t	2	2	4	2	1	1	-	1	4	2	2	1
ARSV-$N(0,1)$	1	4	2	1	2	3	-	-	3	3	-	-
ARSV-Student-t	5	5	5	4	5	4	-	-	5	4	-	-
ARSV-skew-normal	4	3	3	5	3	5	-	-	1	5	-	-
p-value	0.224	0.301	0.237	0.261	0.543	0.623	0.030	0.716	0.242	0.427	0.202	0.573
	proxy of volatlity = RV^{scale}											
TSV-$N(0,1)$	1	1					1	1				
TSV-Student-t	4	2					-	2				
ARSV-$N(0,1)$	3	3					-	3				
ARSV-Student-t	5	5					-	4				
ARSV-skew-normal	2	4					-	5				
p-value	0.232	0.344					0.000	0.330				

The p-value corresponds to the highest p-value of the model that has been excluded from the SSM.

6. Conclusion

The Bayesian approach is used in this paper by means of INLA, which enables to approximate the posterior marginal distributions of the parameters of a threshold stochastic volatility model using numerical approximations. The benefits of using PC priors for the accuracy of the parameter estimates are evaluated in finite samples and, finally, estimates of volatility and volatility forecasts for six series of returns are also provided. These series correspond to stock market, commodity and cryptocurrency returns.

The simulation results show that the PC priors enable us to obtain quite reliable estimates of the parameters in finite samples in comparison to those which result from using "standard" priors. This is particularly evident when the sample sizes are small or moderate. As the sample size increases, the estimates of the parameters become more robust to the choice of the priors.

On the other hand, our empirical results show that: First, the series of returns considered present high volatility persistence and volatility asymmetry, except the Ethereum returns that do not report a statistically significant volatility asymmetry. Second, the sign of the volatility asymmetry is not similar for all series. Third, the threshold stochastic volatility model is the one that shows the best in-sample fit for the majority of the series, and finally the out-of-sample results and, in particular the results regarding the QLIKE loss function, suggest that the threshold stochastic volatility model is relevant to the estimation and forecasting of the volatility of financial time series returns, independently if the series are equity, commodities or cryptocurrencies.

Acknowledgements

We thank two anonymous referees for their careful and constructive comments that helped to improve the paper, and Professor Maria Antónia Amaral Turkman, Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal, for the careful reading of the manuscript and for all the very useful comments and suggestions. We also thank Professor John McDermott for correcting the English. The first author was partially financed by national funds through FCT - Fundação para a Ciência e a Tecnologia under the projects PTDC/MAT-STA/28649/2017 and UIDB/00006/2020. The fourth author acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities, research project PGC2018-096977-B-l00, from the Agencia Estatal de Investigación PID2019-108079GB-C21/AIE/10.13039/501100011033 and from Fundação para a Ciência e a Tecnologia, grant UIDB/00315/2020.

Appendix

A. INLA

In general, given that $p(x \mid z) = p(x, z)/p(z)$ then $p(z) = p(x, z)/p(x \mid z)$. Therefore, if z is now conditioned on some w then $p(z \mid w) = p(x, z \mid w)/p(x \mid z, w)$. Here, the posterior marginal

in (5) can be expressed as:

$$p(\boldsymbol{\Psi} \mid y) = \frac{p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y})}{p(\boldsymbol{\theta} \mid \boldsymbol{\Psi}, \mathbf{y})},$$

and by applying the Bayes' theorem

$$p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\Psi}) p(\boldsymbol{\theta}, \boldsymbol{\Psi})}{p(y)}.$$

Given that

$$p(\boldsymbol{\theta}, \boldsymbol{\Psi}) = p(\theta \mid \boldsymbol{\Psi}) p(\boldsymbol{\Psi}),$$

then

$$p(\boldsymbol{\Psi} \mid \mathbf{y}) \propto \frac{p(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\Psi}) p(\theta \mid \boldsymbol{\Psi}) p(\boldsymbol{\Psi})}{p(\boldsymbol{\theta} \mid \boldsymbol{\Psi}, \mathbf{y})}.$$

Considering that the difficult part of the previous expression is the denominator, an approximation is obtained, $\tilde{p}(\boldsymbol{\theta} \mid \boldsymbol{\Psi}, \mathbf{y})$, to $p(\boldsymbol{\theta} \mid \boldsymbol{\Psi}, \mathbf{y})$ by means of Laplace's method. This expression is to be calculated in the mode $\boldsymbol{\theta}^*$ for a certain value of $\boldsymbol{\Psi}$, that is, $\boldsymbol{\theta} = \boldsymbol{\theta}^*(\boldsymbol{\Psi})$.

The procedure used for approximating $p(\theta_i, \boldsymbol{\Psi} \mid \mathbf{y})$ lies on, primarily, expressing the vector $\boldsymbol{\theta}$ as $(\theta_i, \boldsymbol{\theta}_{-i})$ and then using, as before, the Laplace approximation:

$$p(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y}) = \frac{p((\theta_i, \boldsymbol{\theta}_{-i}) \mid \boldsymbol{\Psi}, \mathbf{y})}{p(\boldsymbol{\theta}_{-i} \mid \theta_i, \boldsymbol{\Psi}, \mathbf{y})} = \frac{p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y})}{p(\boldsymbol{\Psi} \mid \mathbf{y})} \frac{1}{p(\boldsymbol{\theta}_{-i} \mid \theta_i, \boldsymbol{\Psi}, \mathbf{y})} \propto \frac{p(\boldsymbol{\theta}, \boldsymbol{\Psi} \mid \mathbf{y})}{p(\boldsymbol{\theta}_{-i} \mid \theta_i, \boldsymbol{\Psi}, \mathbf{y})}.$$

The denominator is replaced by $\tilde{p}(\boldsymbol{\theta}_{-i} \mid \theta_i, \boldsymbol{\Psi}, \mathbf{y})$ which is obtained by the Laplace approximation computed at the mode $\theta^*_{-i}(\theta_i, \boldsymbol{\Psi})$. Then, it is obtained the $\tilde{p}(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y})$ which is the Laplace approximation of $p(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y})$. The marginal posterior distribution of θ_i is then given by:

$$\tilde{p}(\theta_i \mid \mathbf{y}) \approx \int \tilde{p}(\theta_i \mid \boldsymbol{\Psi}, \mathbf{y}) \tilde{p}(\boldsymbol{\Psi} \mid \mathbf{y}) d\boldsymbol{\Psi},$$

which is calculated as

$$\sum_j \tilde{p}(\theta_i \mid \boldsymbol{\Psi}^{(j)}, \mathbf{y}) p(\boldsymbol{\Psi}^{(j)} \mid \mathbf{y}) \Delta_j,$$

where Δ_j represents the weight which is associated to the integration point $\boldsymbol{\Psi}^{(j)}$. For further details see, *e.g.*, Blangiardo and Cameletti (2015).

B. MCMC simulations

In this appendix, the MCMC estimates of the parameters of the TSV model are presented and compared to the INLA estimates. `Nimble` (see de Valpine et al., 2017, 2021) has been applied for comparison purposes using similar standard priors to those of Table 3. `Nimble` also allows

the parallelization of the processes. For estimating the TSV model 3 chains with 10000 iterations each have been used and a burn in of 5000 has been considered.

Comparing the simulation results of using INLA (Table 3) for estimating the TSV parameters to the ones reported in Table A, notice that for normal standardized returns, INLA estimates are in general more accurate and present smaller standard errors for all sample sizes, particularly for the parameters μ, ϕ and δ. Concerning Student-t standardized returns, INLA is able to estimate all parameters of the TSV model more precisely than the MCMC estimator, except for τ_h. Therefore, the results show that INLA, whenever it is possible to implement, provides very reliable estimates of the TSV model in finite samples and requires much less computation time; see Table B.

Table A

Posterior means, standard deviations (sd) and 95% Credible Intervals (CI), obtained by MCMC, using standard priors (n - sample size).

Parameters	$\boldsymbol{\mu}$	$\boldsymbol{\phi}$	$\boldsymbol{\tau_h}$	$\boldsymbol{\delta}$	$\boldsymbol{\nu}$
True values	**0**	**0.98**	**0.773**	**0.07**	**7**
			Normal		
$n = 500$					
mean	1.509	0.753	1.018	0.048	
sd	0.347	0.055	0.281	0.134	
CI	(0.934 , 2.258)	(0.636 , 0.845)	(0.622 , 1.697)	(-0.208 , 0.312)	
$n = 1000$					
mean	0.754	0.874	0.858	0.057	
sd	0.153	0.024	0.143	0.082	
CI	(0.487 , 1.079)	(0.824 , 0.916)	(0.607 , 1.164)	(-0.100 , 0.217)	
$n = 2000$					
mean	0.429	0.926	0.830	0.062	
sd	0.108	0.017	0.115	0.050	
CI	(0.277 , 0.689)	(0.887 , 0.950)	(0.624 , 1.068)	(-0.035 , 0.160)	
$n = 4000$					
mean	0.251	0.954	0.802	0.066	
sd	0.046	0.007	0.091	0.032	
CI	(0.175 , 0.356)	(0.939 , 0.966)	(0.633 , 0.985)	(0.005 , 0.129)	
			Student-t		
$n = 500$					
mean	0.138	0.909	1.145	0.067	15.679
sd	0.090	0.037	0.523	0.104	9.746
CI	(-0.010 , 0.337)	(0.824 , 0.964)	(0.464 , 2.498)	(-0.132 , 0.272)	(5.468 , 41.844)
$n = 1000$					
mean	0.026	0.965	0.922	0.069	11.869
sd	0.039	0.013	0.267	0.061	5.830
CI	(-0.044 , 0.107)	(0.937 , 0.986)	(0.456 , 1.502)	(-0.048 , 0.190)	(5.747 , 27.542)
$n = 2000$					
mean	0.010	0.974	0.863	0.074	8.942
sd	0.024	0.007	0.187	0.041	2.549
CI	(-0.035 , 0.059)	(0.959 , 0.986)	(0.523 , 1.252)	(-0.004 , 0.155)	(5.801 , 15.456)
$n = 4000$					
mean	-0.002	0.977	0.819	0.072	7.489
sd	0.034	0.005	0.140	0.041	1.460
CI	(-0.087, 0.039)	(0.966 , 0.985)	(0.547 , 1.093)	(-0.006 , 0.147)	(4.976 , 10.486)

Table B
Time (sec) to process a sample - Student-t and Gaussian models using Nimble

Time (sec)		Sample size
Student-t model	Gaussian model	n
30.4	29.9	500
53.9	42.4	1000
91.1	67.6	2000
170.5	124.8	4000

References

Amaral Turkman, M. A., K. F. Turkman, P. de Zea Bermudez, S. Pereia, P. Pereira, and M. de Carvalho (2021). Calibration of the bulk and extremes of spatial data. *Revstat Statistical Journal 19*, 308–324.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*(4), 885–905.

Asai, M., C.-L. Chang, and M. McAleer (2017). Realized stochastic volatility with general asymmetry and long memory. *Journal of Econometrics 199*, 202–212.

Asai, M. and M. McAleer (2004). Dynamic leverage and threshold effects in stochastic volatility models. *Unpublished manuscript, Faculty of Economics, Tokyo Metropolitan University*.

Asai, M. and M. McAleer (2005). Dynamic asymmetric leverage in stochastic volatility models. *Econometric Reviews 24*(3), 317–332.

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

Ashraf, B. N. (2020). Stock markets' reaction to COVID-19: Cases or fatalities? *Research in International Business and Finance 54*, 101249.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica 76*(6), 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal 12*(3), C1–C32.

Baur, D. G. (2012). Asymmetric volatility in the gold market. *The Journal of Alternative Investments 14*(4), 26–38.

Baur, D. G. and T. Dimpfl (2018). Asymmetric volatility in cryptocurrencies. *Economics Letters 173*, 148–151.

Blangiardo, M. and M. Cameletti (2015). *Spatial and Spatio-Temporal Bayesian Models with R-INLA* (First ed.). Wiley.

Breidt, F. J. (1996). A threshold autoregressive stochastic volatility model. In *VI Latin American Congress of Probability and Mathematical Statistics (CLAPEM), Valparaiso, Chile*. Citeseer.

Büyüksahin, B. and M. A. Robe (2014). Speculators, commodities and cross-market linkages. *Journal of International Money and Finance 42*, 38–70.

Chen, C. W. S., F. C. Liu, and M. K. P. So (2008). Heavy-tailed-distributed threshold stochastic volatility models in financial time series. *Australian & New Zealand Journal of Statistics 50*(1), 29–51.

Chen, C. W. S., F. C. Liu, and M. K. P. So (2013). Threshold variable selection of asymmetric stochastic volatility models. *Computational Statistics 28*(6), 2415–2447.

Conlon, T., S. Corbet, and R. J. McGee (2020). Are cryptocurrencies a safe haven for equity markets? An international perspective from the COVID-19 pandemic. *Research in International Business and Finance 54*, 101248.

Corbet, S., C. Larkin, and B. Lucey (2020). The contagion effects of the COVID-19 pandemic: Evidence from gold and cryptocurrencies. *Finance Research Letters 35*, 101554.

Danielsson, J. (1994). Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics 64*(1), 375–400.

de Valpine, P., C. Paciorek, D. Turek, N. Michaud, C. Anderson-Bergman, F. Obermeyer, C. Wehrhahn Cortes, A. Rodrìguez, D. Temple Lang, and S. Paganin (2021). *NIMBLE User Manual*. R package manual version 0.11.1.

de Valpine, P., D. Turek, C. Paciorek, C. Anderson-Bergman, D. Temple Lang, and R. Bodik (2017). Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics 26*, 403–413.

de Zea Bermudez, P., J. M. Marín, and H. Veiga (2020). Data cloning estimation for asymmetric stochastic volatility models. *Econometric Reviews 39*, 1057–1074.

Ehlers, R. and M. Zevallos (2015). Bayesian estimation and prediction of stochastic volatility models via INLA. *Communications in Statistics - Simulation and Computation 44*(3), 683–693.

Elliott, R. J., C. C. Liew, and T. K. Siu (2011). On filtering and estimation of a threshold stochastic volatility model. *Applied Mathematics and Computation 218*(1), 61–75.

Fan, T.-H. and Y.-F. Wang (2013). An empirical Bayesian forecast in the threshold stochastic volatility models. *Journal of Statistical Computation and Simulation 83*(3), 486–500.

Fridman, M. and L. Harris (1998). A maximum likelihood approach for non-Gaussian stochastic volatility models. *Journal of Business & Economic Statistics 16*(3), 284–291.

Ghosh, H., B. Gurung, and P. Source (2015). Kalman filter-based modelling and forecasting of stochastic volatility with threshold. *Journal of Applied Statistics 42*(3), 492–507.

Gómez-Rubio, V. (2020). *Bayesian Inference with INLA* (First ed.). CRC Press, Chapman and Hall.

Gómez-Rubio, V. and H. Rue (2018). Markov chain Monte Carlo with the integrated nested laplace approximation. *Statistics and Computing 28*, 1033–1051.

Goodell, J. W. and S. Goutte (2020). Co-movement of COVID-19 and Bitcoin: Evidence from wavelet coherence analysis. *Finance Research Letters*, 101625.

Granger, C. W. J. (1996). Can we improve the perceived quality of economic forecasts? *Journal of Applied Econometrics 11*(5), 455–473.

Hammoudeh, S. and Y. Yuan (2008). Metal volatility in presence of oil and interest rate shocks. *Energy Economics 30*(2), 606 – 620.

Hansen, P. R. and A. Lunde (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? *Journal of Applied Econometrics 20*(7), 873–889.

Hansen, P. R. and A. Lunde (2006). Consistent ranking of volatility models. *Journal of Econometrics 131*(1), 97–121.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*, 453–497.

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

Harvey, A. C. and N. Shephard (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics 14*(4), 429–34.

Ji, Q., D. Zhang, and Y. Zhao (2020). Searching for safe-haven assets during the COVID-19 pandemic. *International Review of Financial Analysis 71*, 101526.

Karamé, F. (2018). A new particle filtering approach to estimate stochastic volatility models with markov-switching. *Econometrics and statistics 8*, 204–230.

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

Krainski, E. T., V. Gómez-Rubio, H. Bakka, A. Lenzi, D. Castro-Camilo, D. Simpson, F. Lindgren, and H. Rue (2019). *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA* (First ed.). CRC Press, Chapman and Hall.

Lele, S. R., B. Dennis, and F. Lutscher (2007). Data cloning: Easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecology Letters 10*(7), 551–563.

Lele, S. R., K. Nadeem, and B. Schmuland (2010). Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association 105*(492), 1617–1625.

Lyócsa, S., E. Baumöhl, T. Výrost, and P. Molnár (2020). Fear of the coronavirus and the stock markets. *Finance Research Letters 36*, 101735.

Mao, X., V. Czellar, E. Ruiz, and H. Veiga (2020). Asymmetric stochastic volatility models: Properties and particle filter-based simulated maximum likelihood estimation. *Econometrics and Statistics 13*, 84–105.

Mao, X., E. Ruiz, and H. Veiga (2017). Threshold stochastic volatility: Properties and forecasting. *International Journal of Forecasting 33*(4).

Martino, S., K. Aas, O. Lindqvist, N. L. R., and H. Rue (2011). Estimating stochastic volatility models using integrated nested laplace approximations. *The European Journal of Finance 17*(7), 487–503.

Martín-Barragán, B., S. B. Ramos, and H. Veiga (2015). Correlations between oil and stock markets: A wavelet-based approach. *Economic Modelling 50*, 212–227.

Melino, A. and S. M. Turnbull (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics 45*(1-2), 239–265.

Okorie, D. I. and B. Lin (2020). Stock markets and the covid-19 fractal contagion effects. *Finance Research Letters*, 101640.

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

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

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*(1), 246–256.

Rue, H. and L. Held (2005). *Gaussian Markov Random Fields - Theory and Applications* (First ed.). Chapman and Hall/CRC, Boca Raton, Florida.

Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested laplace approximations (with discussion). *Journal of the Royal Statistical Society: Series B 71*(2), 319–392.

Sakaria, D. and J. E. Griffin (2017). On efficient bayesian inference for models with stochastic volatility. *Econometrics and Statistics 3*, 23–33.

Sandmann, G. and S. J. Koopman (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics 87*(2), 271–301.

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

Silvennoinen, A. and S. Thorp (2013). Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money 24*, 42–65.

Simpson, D. P., H. Rue, A. Riebler, T. Martins, and S. H. Sørbye (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science 32*, 1–28.

So, M. K. P., W. K. Li, and K. Lam (2002). A threshold stochastic volatility model. *Journal of Forecasting 21*(7), 473–500.

Sørensen, M. (2000). Prediction-based estimating functions. *The Econometrics Journal 3*(2), 123–147.

Takahashi, M., T. Watanabe, and Y. Omori (2016). Volatility and quantile forecasts by realized stochastic volatility models with generalized hyperbolic distribution. *International Journal of Forecasting 32*(2), 437–457.

Taylor, S. J. (1982). *Time series Analysis: Theory and Practice*, Volume 1, Chapter Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75, pp. 203–206. North-Holland, Amsterdam.

Turkman, K. F., M. González Scotto, and P. de Zea Bermudez (2014). *Non-Linear Time Series: Extreme Events and Integer Value Problems* (First ed.). Springer.

Watanabe, T. and Y. Omori (2004). A Multi-Move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard & Pitt (1997). *Biometrika 91*(1), 246–248.

Wirjanto, T., A. Kolkiewicz, and Z. Men (2016). Bayesian analysis of a threshold stochastic volatility model. *Journal of Forecasting 35*(5), 462–476.

Wu, X.-Y. and H.-L. Zhou (2015). A triple-threshold leverage stochastic volatility model. *Studies in Nonlinear Dynamics and Econometrics 19*, 483–500.