Working paper

2021-01

Statistics and Econometrics

ISSN 2387-0303

Integrated nested Laplace approximations for threshold stochastic volatility models

P. de Zea Bermudez, J.M. Marín, Hårvard Rue, Helena Veiga

Serie disponible en

<http://hdl.handle.net/10016/12>

Creative Commons Reconocimiento-
NoComercial- SinObraDerivada 3.0 España
([CC BY-NC-ND 3.0 ES](http://creativecommons.org/licenses/by-nc-nd/3.0/es/))

Integrated nested Laplace approximations for threshold stochastic volatility models*

P. de Zea Bermudez[†] J. M. Marín[‡] Håvard Rue[§] Helena Veiga[¶]

December 30, 2020

ABSTRACT

The aim of the paper is to implement the integrated nested Laplace (INLA) approximations, known to be very fast and efficient, for a threshold stochastic volatility model. INLA replaces MCMC simulations with accurate deterministic approximations. We use proper although not very informative priors and 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, they provide more accurate estimates of the model's parameters, but as sample size increases both type of priors lead to reliable estimates of the parameters. We also validate the estimation method in-sample and out-of-sample by applying it to six series of returns including stock market, commodity and cryptocurrency returns and by forecasting their one-day-ahead volatilities, respectively. Our empirical results support that the TSV model does a good job in forecasting the one-day-ahead volatility of stock market and gold returns but faces difficulties when the volatility of returns is extreme, which occurs in the case of cryptocurrencies.

JEL-Classifications: C13; C32; C52; C58

Keywords: INLA; PC priors; Threshold Stochastic Volatility Model

1. Introduction

Stochastic volatility models are widespread proposed in the literature to cope with the main empirical features of financial time series. Very recent examples are the models proposed by Asai

*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-I00, 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.

[†]Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal and CEAUL - Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal. e-mail: pcbermudez@fc.ul.pt.

[‡]Department of Statistics Universidad Carlos III de Madrid, Spain. Email: jmmarin@est-econ.uc3m.es.

[§]King Abdullah University of Science and Technology. Email: Haavard.Rue@kaust.edu.sa.

[¶]Universidad Carlos III de Madrid (Department of Statistics and Instituto Flores de Lemus), Spain and BRU-IUL, Instituto Universitário de Lisboa, Portugal. Email: mhveiga@est-econ.uc3m.es. *Corresponding author.*

et al. (2017) and Mao et al. (2017, 2020) that accommodate a general asymmetric function for volatility. 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 depending on the sign of the lagged returns (Breidt, 1996; So et al., 2002).¹

This paper focuses on the second type of volatility asymmetry and proposes to use Integrated Nested Laplace Approximation to estimate the parameters and the latent volatility of threshold stochastic (TSV) models. Stochastic volatility (SV) models are very flexible to accommodate the empirical features of the data but their estimation is not easy due to the intractable form of the likelihood function. Over the last twenty years many methods have been proposed, 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) and MCMC procedures (Jacquier et al., 1994; Shephard and Pitt, 1997; Kim et al., 1998). Very recently, de Zea Bermudez et al. (2020) proposed 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 calculation 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 back are possible now. Namely, Bayesian inference could not really be performed from an application 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 intensive. Moreover, the assessment of chains convergence can sometimes be a challenge. INLA was conceived with the purpose of providing 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 Markov random fields (GRMF). In fact, many interesting and useful models may be rewritten in such a way that they possess a latent (unobserved) structure which can be represented by a GRMF. INLA is used in a very wide range of Bayesian applications, varying from time series and hierarchical models to space and space-time modelling (Blangiardo and Cameletti, 2015; Krainski et al., 2019; Gómez-Rubio, 2020). Recently, Gómez-Rubio and Rue (2017) 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` and `stochvol.t` that use Gaussian and Student- t distributions, respectively. INLA functions have been specifically adapted by Håvard Rue for dealing with TSV models.

We evaluate the performance of the Bayesian approach 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 using PC priors that convey a sound

¹For 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.

information about the TSV model parameters; the model persistence, in accordance with the financial literature, is one of such parameters. In a second phase, the estimation procedure using PC priors is evaluated in a slightly less informative scenario. Finally, traditional vague priors are chosen with the intent of “letting the data speak”.

The performance of INLA for estimating the parameters of TSV models is assessed by means of a simulation study. We use samples of returns with sizes ranging from 500 to 4000 and two distributions: the Gaussian and the Student- t . The value of the parameters in the simulation study reflect real financial data features, such as a very large and positive persistence parameter. 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 it is 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 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; finally, the third set contains 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 support that the TSV model does a good job in forecasting the one-day-ahead volatility of stock market and gold returns but faces more difficulties when the volatility of returns is extreme, which occurs in the case of cryptocurrencies.

This paper makes several contributions to the literature. First, we propose the use of INLA to estimate threshold stochastic volatility models. Second, we compare the performance of the estimation method by using informative PC priors, less informative PC and “vague” standard priors. We observe that the use of PC priors leads to reliable estimates of parameters in small and moderate sample sizes. Finally, we conclude that the model predicts the one-day-ahead volatilities quite well for stock market and gold returns, but it encounters difficulties when the series of returns are extremely volatile. This is the case of cryptocurrencies. The solution could be the inclusion of jumps into the model. An extension that we will leave for future research.

The paper is organized as follows. Section 2 describes the TSV model. In the first part of Section 3 the basic concepts underlying INLA are reviewed, as well as how to express TSV models in terms of a GRMF. In the second part of Section 3, PC priors are presented and briefly discussed. The simulation study is given 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

$$y_t = \sigma_t \epsilon_t = \exp(h_t/2) \epsilon_t, \quad t = 1, \dots, T, \quad (1)$$

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

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

where $I(\cdot)$ is an indicator function that takes the value of 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. Given that financial returns are known to be heavy-tailed distributed, we allow ϵ_t to follow a Student- t distribution. Finally, the parameter ϕ controls the persistence of the volatility; values of ϕ as large as or even larger than 0.90 are quite common in real applications.

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 well known that MCMC algorithms can be very inefficient in the sense that they can take very long 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). It is extremely fast and gives very precise results. Contrarily to MCMC algorithms, which are simulation-based methods, INLA uses numerical approximations. INLA can be applied whenever the latent structure of the model can be represented as a Gaussian random field. Due to the fact that a Markov property is frequently a reasonable assumption, we end up with GRMF. GRMF theory is extensively addressed in Rue and Held (2005) and provides very attractive computational properties. In this subsection we will fill in the reader with a very general explanation of the fundamental aspects underlying the software INLA. Some further details can be found in the Appendix.

The Laplace method, which is one of the building blocks of INLA, basically enables to approximate a well-behaved and unimodal function $f(\cdot)$ 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 $f(\cdot)$ (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 $(Y_1, Y_2, \dots, Y_n)$ be a random sample of size n and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ the corresponding observed data. Let us assume that the distribution of Y_i , $i = 1, 2, \dots, n$ is characterized by some parameter ϕ_i . Likewise in a generalized linear model framework, we might be interested to link ϕ_i to the values of some M covariates, $x_1, x_2, \dots, x_M$, which (allegedly) affect the observations y_i , $i = 1, 2, \dots, n$, in a linear manner. Moreover, we might also want to consider additional covariates, $z_1, z_2, \dots, z_L$ that affect the outcome variables by means of functions, $f_l(\cdot)$, that account for other kinds of effects. All that can be done through a link function $g(\cdot)$ such that

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

Let $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_M, f_1(\cdot), f_2(\cdot), \dots, f_L(\cdot))$ represent the latent field components and $\boldsymbol{\Psi} = (\Psi_1, \Psi_2, \dots, \Psi_K)$ be the vector of the model hyperparameters. The likelihood of $(\boldsymbol{\theta}, \boldsymbol{\Psi})$ is given by

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

Provided that the prior distribution of $\boldsymbol{\theta} \mid \boldsymbol{\Psi}$ is a multivariate normal with mean vector $\mathbf{0}$ and precision matrix $Q(\boldsymbol{\Psi})$, and 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 $|\mathbf{A}|$ indicates the determinant of the matrix $\mathbf{A}$, and $\mathbf{A}^T$ stands for the transpose operation on the matrix $\mathbf{A}$.

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

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

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

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

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 $p(\boldsymbol{\Psi} \mid \mathbf{y})$ need to be derived. Basically, what INLA does is to approximate these two distributions by Laplace approximation. 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) can be expressed as follows:

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

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

$$h_t \mid h_{t-1}, h_{t-2}, \dots, h_1, \mu, \phi, \delta, \tau_\eta \sim N(\mu + \phi h_{t-1} + \delta I(\epsilon_{t-1} < 0), 1/\tau_\eta).$$

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

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

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

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 choose these hyperparameters to reflect absence of prior information about the model parameters. Specifically, both Martino et al. (2011) and Ehlers and Zevallos (2015) choose 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, we use a different prior for ϕ which in fact boils down to be the same. We use the INLA function `betacorrelation`, which is a $Beta(a, b)$ with $a > 0$ and $b > 0$, and has support in $(-1, 1)$.² 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 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 events located in the tail of the distribution are.

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.

²The probability density function of such a $Beta(a, b)$ distribution is given by $p(\rho \mid a, b) = 0.5 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \rho^{a-1} (1 - \rho)^{b-1}$, $-1 < \rho < 1$.

PC priors are especially easy to derive for some model's parameters, such as for the precision and for the persistence. Simpson et al. (2017) show that the PC prior for $\tau = \sigma^{-2}$ is a type-2 Gumbel distribution with pdf given by

$$\pi(\tau) = \frac{\lambda}{2} \tau^{-3/2} \exp(-\lambda \tau^{-1/2}),$$

for $\tau > 0$, λ positive given by $\lambda = -\ln(\alpha)/u$ and where $\tau = \infty$ is the base-model. For ϕ , the PC prior is

$$\pi(\phi) = \frac{\lambda \exp(-\lambda \sqrt{1-\phi})}{1 - \exp(-\sqrt{2}\lambda)} \frac{1}{2\sqrt{1-\phi}},$$

where $\phi = 1$ is the base-model. The parameter λ is defined so that $P(\phi > u) = \alpha$, $-1 < u < 1$ and $\sqrt{\frac{1-u}{2}} < \alpha < 1$. Several PC priors for ϕ and τ are presented in Figure 1. For making these plots the INLA functions `inla.pc.dprec` and `inla.pc.dcor1` are used for τ and ϕ , respectively (see e.g. Franco-Villoria et al., 2018). 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 precision and persistence parameters.

4. Simulation study

We run Monte Carlo simulations to analyze the performance of the INLA estimators of TSV parameters in finite samples. For this purpose, we simulate five hundred samples from the following TSV model:

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

with sample sizes of $n = 500, 1000, 2000$ and 4000 observations. 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 the threshold, δ and for μ are $N(0, 1)$. The values of u and α chosen 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 previous knowledge that positive and close to 1 values of ϕ 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 most likely 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 to very informative.

The second set of PC priors, which are less informative, are 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 $N(0, 1)$.³

³The Q matrix (referred to in Subsection 3.2) is always given as $Q = \begin{bmatrix} 0.1 & 0 \\ 0 & 0.1 \end{bmatrix}$ except for the second set of PC priors and $n = 500$. In that case, the variance of μ and δ is decreased and thus $Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Table 2 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 what regards 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 we present in Table 1 the time required to complete one iteration when using standardized Student- t and Gaussian returns. The results were obtained using a Mac Pro processor 2.6 GHz, 6 core intel core i7.

Table 1
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

Table 4 presents the simulation results using the priors considered by Martino et al. (2011) and Ehlers and Zevallos (2015) for ϕ , μ and τ_h . The families of distributions chosen by these authors are indicated in (6). The hyperparameters are selected in such a way so to obtain a prior for ϕ approximately uniform in the interval (0,1), and vague priors for the logarithm of the precision parameter and for μ . The prior for δ remains the same, i.e., $N(0,1)$.

A careful examination of the results presented in Tables 2–4 shows that the benefits of using the PC priors, are mostly seen when the samples sizes vary from small to moderate. This conclusion is not surprising due to 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 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 vaguely informative priors belonging to standard families of distributions. For large sample sizes, the parameters' estimates 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 advantage of using PC priors seems to lie on the estimation of the degrees of freedom of the Student- t distribution.

5. Empirical application

The TSV model presented in Section 2 is estimated on 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

Table 2

Finite sample properties of the INLA estimator of the parameters of the TSV model with normal and Student- t standardized returns using PC priors.

Parameters	μ	ϕ	τ_h	δ	ν	n
True values	0	0.98	0.773	0.07	7	
	Normal					
mean	0.029	0.965	1.317	0.069		500
sd	0.040	0.016	0.447	0.077		500
mean	0.011	0.974	1.026	0.072		1000
sd	0.027	0.009	0.262	0.053		1000
mean	0.005	0.977	0.903	0.072		2000
sd	0.019	0.006	0.177	0.037		2000
mean	0.003	0.979	0.850	0.070		4000
sd	0.013	0.004	0.126	0.026		4000
	Student-t					
mean	0.035	0.962	1.504	0.079	12.641	500
sd	0.042	0.017	0.592	0.081	9.247	500
mean	0.023	0.971	1.112	0.070	9.446	1000
sd	0.029	0.010	0.296	0.057	3.622	1000
mean	0.013	0.976	0.956	0.074	7.937	2000
sd	0.020	0.006	0.194	0.039	1.664	2000
mean	0.010	0.978	0.879	0.071	7.937	4000
sd	0.014	0.004	0.134	0.028	1.007	4000

Note: Posterior means and standard deviations of the parameters; n represents the sample size.

(Bitcoin and Ethereum). S&P 500 and Nikkei 225 returns span 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 shows the time series of returns and absolute returns. Absolute returns are often used as a proxy of volatility in the literature. We observe that stock market returns show a period of high volatility that corresponds to the last global financial crisis, and 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 commodities returns is similar to that of stock markets returns, specially oil returns. We also observe a huge increase in oil volatility at

Table 3

Finite sample properties of the INLA estimator of the parameters of the TSV model with normal and Student- t standardized returns using mild informative PC priors.

Parameters	μ	ϕ	τ_h	δ	ν	n
True values	0	0.98	0.773	0.07	7	
	Normal					
mean	0.038	0.959	1.412	0.070		500
sd	0.041	0.018	0.456	0.078		500
mean	0.014	0.972	1.077	0.073		1000
sd	0.027	0.009	0.263	0.053		1000
mean	0.006	0.976	0.925	0.073		2000
sd	0.019	0.006	0.178	0.037		2000
mean	0.003	0.979	0.862	0.070		4000
sd	0.013	0.004	0.127	0.026		4000
	Student-t					
mean	0.094	0.940	1.841	0.077	14.060	500
sd	0.045	0.023	2.736	0.085	11.137	500
mean	0.027	0.969	1.157	0.070	9.556	1000
sd	0.029	0.010	0.299	0.057	3.742	1000
mean	0.015	0.975	0.987	0.074	7.959	2000
sd	0.020	0.006	0.194	0.039	1.682	2000
mean	0.011	0.978	0.894	0.071	7.398	4000
sd	0.014	0.004	0.133	0.028	1.008	4000

Note: Posterior means and standard deviations of the parameters; n represents the sample size.

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 5 reports the data descriptive statistics. The empirical distributions of returns are in general leptokurtic and skewed to the left, except the distribution of Ethereum returns that is skewed to the right. Normality is rejected for all series.

Table 4

Finite sample properties of the INLA estimator of the parameters of the TSV model with normal and Student- t standardized returns using “standard” priors.

Parameters	μ	ϕ	τ_h	δ	ν	n
True values	0	0.98	0.773	0.07	7	
Normal						
mean	0.058	0.948	1.317	0.070		500
sd	0.043	0.022	0.572	0.082		500
mean	0.013	0.971	0.949	0.073		1000
sd	0.028	0.010	0.254	0.054		1000
mean	0.005	0.977	0.861	0.073		2000
sd	0.019	0.006	0.177	0.037		2000
mean	0.003	0.979	0.828	0.070		4000
sd	0.013	0.004	0.127	0.026		4000
Student-t						
mean	0.035	0.896	1.399	0.077	16.006	500
sd	0.050	0.030	0.925	0.090	14.362	500
mean	0.028	0.968	1.016	0.070	10.140	1000
sd	0.030	0.011	0.278	0.059	4.303	1000
mean	0.014	0.976	0.910	0.074	8.072	2000
sd	0.021	0.006	0.189	0.040	1.757	2000
mean	0.010	0.978	0.853	0.071	7.456	4000
sd	0.014	0.004	0.133	0.028	1.029	4000

Note: Posterior means and standard deviations of the parameters. n corresponds to the sample size.

Table 5
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

Figure 3. Series of returns and absolute returns in percentage.

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

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

5.1. Parameter estimation results

The parameters of the TSV model are estimated using the previous PC priors. Given the features of the data, we assume that the standardized returns follow a Student- t distribution.

Table 6

Posterior means and standard deviations (in parenthesis) of TSV parameters using INLA.

Series	μ	ϕ	τ_h	δ	ν
SP500	-0.154 (0.010)	0.975 (0.003)	1.757 (0.209)	0.319 (0.021)	22.167 (7.204)
Nikkei	-0.087 (0.011)	0.970 (0.005)	2.053 (0.258)	0.206 (0.021)	24.393 (9.208)
Gold	0.015 (0.007)	0.993 (0.002)	2.431 (0.557)	-0.033 (0.014)	5.642 (0.465)
Oil	-0.040 (0.008)	0.988 (0.003)	1.923 (0.334)	0.121 (0.017)	9.377 (1.131)
Bitcoin	0.098 (0.021)	0.976 (0.007)	0.647 (0.128)	-0.095 (0.043)	3.481 (0.319)
Ethereum	0.171 (0.030)	0.940 (0.018)	0.765 (0.130)	0.003 (0.056)	4.588 (0.778)

Table 6 reports the parameter’s estimation results. 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 tomorrow 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 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). Regarding gold futures returns, we observe that volatility asymmetry is different. Positive gold returns today will increase volatility tomorrow.

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 increases in 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). We also observe high volatility persistence in the cryptocurrencies returns that follow fat tailed distributions. Furthermore, Ethereum returns do not present volatility asymmetry while Bitcoin returns show an

inverted asymmetric reaction; see Figure 3.

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 volatility estimates range from the beginning of the series until December 31, 2018, while the forecasts span the period between the beginning of January, 2019 and the 22nd or 23rd of July 2020. We use a rolling window forecasting scheme with a fix number of observations. The first window includes all observations until December 31, 2018, and then we roll this window one day keeping the same number of observations inside the window.

The in-sample period includes the last global financial crisis while the out-of-sample period includes the Coronavirus pandemic. During the last global financial crisis, stock markets witnessed unprecedented declines that lead to a decrease of wealth, and to its costly social and economic consequences. All this uncertainty is reflected 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 16, 24 and 25, 2020 for the S&P 500, Nikkei and Gold returns, respectively, on April 21, 2020 for the Oil returns and on March 12, 2020 for the cryptocurrency returns. On March 23, 2020, the S&P 500 index lost around 35% of its February 19, 2020 value. This decline is comparable in magnitude to the declines occurred on 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 an increase of COVID-19 cases. Furthermore, these declines were spread to other stock markets causing a wave of contagion during the period of confinement (see Okorie and Lin, 2020). Regarding Gold and the cryptocurrencies, the evidence is not clear, 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), Conlon et al. (2020) and Corbet et al. (2020) provide an opposite evidence.

Volatility is not observable and therefore, we use as a proxy the absolute returns (see Ghysels et al., 2006). Figure 5 plots the box-plots of the differences between the one-day-ahead volatility forecasts and the absolute returns. If the TSV model provides accurate volatility forecasts, the median of these differences should be as close as possible to zero, and their distributions symmetric. Although the distributions of the differences between the volatility forecasts and the absolute returns do not fully satisfy the previous conditions, we plot in green those that are closer to these conditions and in red those that are more distant. All in all, it seems that for the cryptocurrency returns the TSV model fails to provide accurate one-day-ahead volatility forecasts, while for the stock market and gold returns, it does a better job in forecasting the volatility. Note that the volatility of cryptocurrency returns is more extreme, which may require to extend the TSV model in order to include jumps.

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

Figure 5. Distributions of the differences between the volatility forecasts and absolute returns.

6. Conclusion

We use the Bayesian approach using the numerical methods known as Integrated Laplace approximations to obtain the approximations to the posterior marginals of the parameters of a threshold stochastic volatility model. We test the benefits of using PC priors for the accuracy of the parameter estimates in finite samples and, finally, we also provide estimates of volatility and volatility forecasts for six series of returns, which correspond to stock market, commodity and cryptocurrency returns.

The simulation results show that the PC priors help in obtaining quite reliable estimates of the parameters in finite sample in comparison to those in which “standard” priors are used. In particular if the sample size is 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 all series of returns present high volatility persistence and volatility asymmetry, except the Ethereum returns that do not report a statistically significant volatility asymmetry. The sign of the volatility asymmetry is not similar for all series. Finally, the out-of-sample results and, in particular, the distribution of the differences between the one-day-ahead volatility forecasts and the absolute returns, suggest that the threshold stochastic volatility provides accurate forecasts of the volatility for stock market and gold returns. Nevertheless, it does a poor job in forecasting the volatility of the cryptocurrency returns. This is expected since these series are extreme volatile. The model might be extended in order to incorporate jumps, but this extension is left for future research.

Appendix

Let us consider the posterior marginal given in (5). We get⁴:

$$p(\Psi | y) = \frac{p(\theta, \Psi | y)}{p(\theta | \Psi, y)},$$

and by applying the Bayes’ theorem we obtain

$$p(\theta, \Psi | y) = \frac{p(y | \theta, \Psi)p(\theta, \Psi)}{p(y)}.$$

Given that

$$p(\theta, \Psi) = p(\theta | \Psi)p(\Psi),$$

we get

$$p(\Psi | y) \propto \frac{p(y | \theta, \Psi)p(\theta | \Psi)p(\Psi)}{p(\theta | \Psi, y)}.$$

Considering that the difficult part of the previous expression is the denominator, we can then obtain an approximation, $\tilde{p}(\theta | \Psi, y)$, to $p(\theta | \Psi, y)$ by means of Laplace’s method. This expression is to be calculated in the mode θ^* for a certain value of Ψ , that is, $\theta = \theta^*(\Psi)$.

⁴Given that $p(x | z) = p(x, z)/p(z)$ then $p(z) = p(x, z)/p(x | z)$. Therefore, if z is now conditioned on some w then $p(z | w) = p(x, z | w)/p(x | z, w)$.

The procedure used for approximating $p(\theta_i, \Psi \mid \mathbf{y})$ lies on, primarily, expressing the vector θ as (θ_i, θ_{-i}) and then using, as before, the Laplace approximation. We get:

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

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

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

which is calculated as

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

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

References

- 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.
- Baur, D. G. (2012). Asymmetric volatility in the gold market. *The Journal of Alternative Investments* 14(4), 26–38.
- 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)*, Valparaíso, 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 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. Zavallos (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.
- Franco-Villoria, M., M. Ventrucchi, and H. Rue (2018). Bayesian varying coefficient models using PC priors. *arXiv:1806.02084*, 1–21.
- 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.
- Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics* 131(1), 59–95.
- Goodell, J. W. and S. Goutte (2020). Co-movement of COVID-19 and Bitcoin: Evidence from wavelet coherence analysis. *Finance Research Letters*, 101625.
- Gómez-Rubio, V. (2020). *Bayesian Inference with INLA* (First ed.). CRC Press, Chapman and Hall.
- Gómez-Rubio, V. and H. Rue (2017). Markov chain Monte Carlo with the integrated nested laplace approximation. *arXiv:1701.07844 [stat.CO]*, 1–25.

- Hammoudeh, S. and Y. Yuan (2008). Metal volatility in presence of oil and interest rate shocks. *Energy Economics* 30(2), 606 – 620.
- 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.
- Jacquier, E., N. G. Polson, and P. E. Rossi (1994). Bayesian analysis of stochastic volatility models: Reply. *Journal of Business & Economic Statistics* 12(4), 413–417.
- 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.
- 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.
- 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.
- 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.
- 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.