Bayesian Analysis of a Threshold Stochastic Volatility Model

TONY S. WIRJANTO,1,2* ADAM W. KOLKIEWICZ1 AND ZHONGXIAN MEN1

1 Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
2 School of Accounting and Finance, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

ABSTRACT
This paper proposes a parsimonious threshold stochastic volatility (SV) model for financial asset returns. Instead of imposing a threshold value on the dynamics of the latent volatility process of the SV model, we assume that the innovation of the mean equation follows a threshold distribution in which the mean innovation switches between two regimes. In our model, the threshold is treated as an unknown parameter. We show that the proposed threshold SV model can not only capture the time-varying volatility of returns, but can also accommodate the asymmetric shape of conditional distribution of the returns. Parameter estimation is carried out by using Markov chain Monte Carlo methods. For model selection and volatility forecast, an auxiliary particle filter technique is employed to approximate the filter and prediction distributions of the returns. Several experiments are conducted to assess the robustness of the proposed model and estimation methods. In the empirical study, we apply our threshold SV model to three return time series. The empirical analysis results show that the threshold parameter has a non-zero value and the mean innovations belong to two separately distinct regimes. We also find that the model with an unknown threshold parameter value consistently outperforms the model with a known threshold parameter value. Copyright © 2016 John Wiley & Sons, Ltd.

KEY WORDS threshold stochastic volatility; Bayesian inference; Markov chain Monte Carlo; deviance information criteria (DIC)

INTRODUCTION
Volatilities of financial asset returns are observed to be time varying, clustered and highly persistent over time. Properly capturing these features is essential for volatility prediction. In the literature, statistical models such as the autoregressive conditional heteroskedasticity (ARCH) model introduced by Engle (1982), the generalized autoregressive conditional heteroskedasticity (GARCH) model, introduced by Bollerslev (1986), and the stochastic volatility (SV) models, proposed by Taylor (1986), have been used to characterize this volatility behavior. The main difference between the ARCH, GARCH and SV models is that in the former the conditional variance is a deterministic function of the past volatilities and the squared observed returns, while in the latter the logarithm of the conditional volatility of the asset returns is assumed to follow a latent stationary first-order autoregressive (AR(1)) process with a univariate normal innovation. In both of these volatility models the innovation of the asset returns is usually assumed to follow either a standard normal or Student’s t-distribution. Empirically, SV models are preferred over the ARCH/GARCH-type models in modeling and predicting volatilities of financial asset returns. This is partially due to the fact that the SV models tend to exhibit lower persistence in volatility of returns than the ARCH/GARCH models. This paper builds on the SV model.

Researchers working on the SV models typically focus on two particular aspects of the models. The first one is the characterization of the properties of conditional volatility such as persistence, clustering and asymmetry. The other one is the description of the shape of the conditional distribution of asset returns. In other words, the specification of the error process for the mean equation has been a subject of much research in the literature. This is because it is seen to affect the marginal distribution of the returns and, as a result, critically affect the empirical performance of the model. The innovation of the mean equation is initially assumed to have a standard normal distribution. This assumption is subsequently found to be restrictive in terms of its ability to characterize the marginal distribution of financial returns. Instead it has been argued that the residuals of the SV model should be heavy-tailed and left-skewed to be consistent with the stylized facts of financial returns. To characterize the heavy tail property of the marginal distribution of returns, Chib et al. (2002), Jacquier et al. (2004), Shimada and Tsukuda (2005) and Zhang and King (2008), among others, extend the SV model to allow the mean innovation process of the model to have a Student’s t-distribution. Durham (2007) applies a mixture-of-normal distribution to the innovation process in the mean equation. In a more general setting, Xu and Wirjanto (2008) use a mixture-of-bivariate normal distributions to approximate

*Correspondence to: Tony S. Wirjanto, Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1. E-mail: twirjant@uwaterloo.ca

Copyright © 2016 John Wiley & Sons, Ltd
the joint distribution of the mean innovation and that of the latent process, and the authors find that the mixture-of-bivariate normal distributions can describe various shapes of the joint distribution of the innovations of the two equations that define the model.

In addition, researchers working in this area also have come to the realization recently that the volatility of asset returns can exhibit very different types of behavior over different segments of a given time period. Motivated by this, So et al. (2002) propose a threshold SV (TSV) model in which the latent volatility follows a threshold AR(1) process, where the threshold parameter value is fixed at zero. However the assumption of a standard normal distribution for the mean innovation is maintained in the model. Subsequently, Chen et al. (2008) extended this model by endowing a Student’s $t$-distribution to the mean innovation to accommodate the heavy tail property of the empirical distribution of asset returns. In addition, the authors treat the threshold parameter as unknown and estimated from the data. For the two threshold SV models, the logarithms of conditional volatilities and the parameters in different states are estimated by complex MCMC procedures. Based on the TSV models considered by So et al. (2002) and Chen et al. (2008), a TSV model is recently proposed in Xu (2012), in which the innovation of the mean equation is allowed to be correlated with both of the two latent innovations of the threshold AR(1) process. However, this generality is introduced at the cost of having to make volatility observable through the construction of realized volatility, so that it is tractable to estimate the resulting model by a maximum likelihood estimator (MLE).

As pointed out in a review paper by Wirjanto and Xu (2012), both the TSV models and the mixture SV models have been successfully used in analyzing financial data. This paper presents a new parsimonious TSV model where the mean error process follows a threshold univariate normal distribution, while the latent variables follow an AR(1) process endowed with univariate normal innovations. Specifically, the error distribution of the mean equation is assumed to have zero mean, and it switches between two regimes determined by a threshold level. In our specification of the model, there is no need to introduce a location parameter in the latent AR(1) process, since the normal distributions in the two regimes are allowed to have non-unit variances. In this setting, the latent AR(1) process partly describes the conditional volatility of returns. In addition, we treat the threshold as an unknown parameter, which is to be estimated from data. This simple and parsimonious TSV model represents our first contribution to the literature.

Our second contribution lies in the development of an efficient MCMC algorithm to fit the proposed model. We develop a version of the Metropolis–Hastings (MH) method proposed in Men et al. (2012), where the latent states are augmented as parameters and sampled one at a time via the MH sampler with a univariate normal distribution as the proposal. Due to the fact that the proposal distribution is a part of the target distribution, the proposed simulation method is straightforward to implement and exhibits good convergence properties.

As a third contribution to the literature, we provide a formal test to assess the goodness-of-fit of the model and to conduct volatility forecasts based on the fitted model. However, both the diagnostic test and the volatility forecasts require the computation of filter and prediction distributions of the latent states. To this end, we employ an auxiliary particle filter (APF) technique introduced by Pitt and Shephard (1999), which has proven to be an efficient tool for calculating the likelihood of the SV models recursively.

The remainder of the paper is organized as follows. The next section introduces our proposed TSV model and presents an efficient MCMC algorithm for the estimation of its parameters. The latent states are augmented as parameters and then estimated by a sample mean based on simulated values from the corresponding full conditional. The third section describes the procedure used for model selection and the assessment of the goodness-of-fit of the model. The fourth section conducts simulation studies to examine the performance of the proposed model and its MCMC algorithm. The fifth section provides empirical applications of the proposed model to three return datasets, and the sixth section concludes the paper.

MODEL AND ESTIMATION

The TSV model

A discrete-time SV model is proposed by Taylor (1986). In its canonical form, the observed time series of asset returns is modeled by two independent processes, which describe the process of generating the returns and the process generating the corresponding future volatilities of returns. Let $y_t$ denote the demeaned observed return of a financial asset at time $t$, $t \leq T$, where $T$ is a positive integer denoting the sample size. The process of $y_t$ can be characterized as a product of an exponential function of a random variable $h_t$ and another random variable $\epsilon_t$. That is, the model structure can be described as

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

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma u_{t+1}, \quad t = 1, \ldots, T - 1$$

$$h_1 \sim N(\mu, \sigma^2/(1 - \phi^2))$$

Copyright © 2016 John Wiley & Sons, Ltd

where $\epsilon_t$ is assumed to follow a standard normal distribution. It is also assumed that $\epsilon_t$ is independent of the innovation of the latent process $u_{t+1}$. Following convention in the literature, we set $u_t \sim N(0, 1)$, where $N(a, b)$ is a univariate normal distribution with mean $a$ and variance $b$. For the latent first-order AR(1) process (2) to be weakly stationary, the persistence parameter $\phi$ is assumed to satisfy the restriction that $|\phi| < 1$. An extended version of the SV model defined in equations (1)–(3) is given by Harvey and Shephard (1996), where a correlation between $\epsilon_t$ and $u_{t+1}$ is permitted. In practice, this correlation between returns and future volatilities is usually found to have a negative sign and interpreted as the leverage effect.

As argued in So et al. (2002) and Chen et al. (2008), the conditional volatilities, $h_t$, exhibit different dynamic patterns in response to the arrival of different types of market news. That is, volatility changes over time according to the previously observed return. In particular, So et al. (2002) observe this asymmetry following the arrival of bad news and good news in the market, where typically volatility tends to be higher with the arrival of bad news. Based on this consideration, So et al. (2002) and Chen et al. (2008) propose TSV models. As stated earlier, the main feature of the TSV models, as a departure from the standard SV models, is that the latent states follow a threshold AR(1) process given by

\[ \begin{align*}
    h_{t+1} - \mu_0 &= \phi_0(h_t - \mu_0) + \sigma_0 u_{t+1}, \quad y_{t-1} < r \\
    h_{t+1} - \mu_1 &= \phi_1(h_t - \mu_1) + \sigma_1 u_{t+1}, \quad y_{t-1} \geq r
\end{align*} \tag{4} \]

The specification of the threshold process (4) stems from the threshold framework originally proposed by Tong and Lim (1980). In the TSV model of So et al. (2002), the authors treat the threshold parameter, $r$, as known and pre-set it at a zero value, whereas in Chen et al. (2008) $r$ is assumed to be an unknown parameter, which changes over different periods of time. In both approaches, the threshold parameter value, $r$, plays an important part in determining the distributions of future volatilities. For the two types of TSV models, cumbersome MCMC algorithms are developed for parameter estimation.

This paper proposes a new, and arguably more parsimonious, TSV model. Instead of endowing a threshold value to the latent process as in So et al. (2002) and Chen et al. (2008), we endow it to the error process of the mean equation. That is, we assume that the latent volatility states, $h_t$, follow an AR(1) process, but let $\epsilon_t$ follow a threshold normal distribution triggered by the observed return time series. In particular, our TSV model is specified as

\[ y_t = \exp(h_t/2)\epsilon_t \tag{5} \]

\[ h_{t+1} = \phi h_t + \sigma u_{t+1} \tag{6} \]

\[ h_1 \sim N(0, \sigma^2/(1 - \phi^2)) \tag{7} \]

where

\[ \begin{align*}
    \epsilon_t &\sim N(0, \lambda_1^2) \quad \text{if } y_{t-1} < r \\
    \epsilon_t &\sim N(0, \lambda_2^2) \quad \text{if } y_{t-1} \geq r
\end{align*} \tag{8} \]

As in the previous specification of the TSV model, $r$ is the threshold parameter, which represents the level of the time series of $y_t$. Since the two normal distributions in equation (8) are allowed to have non-unit variances, there are no location parameters in Eqs. 6 and (7). Under our specification, the innovations of $\epsilon_t$ shift between the two regimes depending on the previous returns. More specifically, the variances of the two normal distributions switch between the two regimes according to the threshold level of previous price change. Compared with the existing TSV models, we let the threshold parameter, $r$, determine the marginal distribution of future returns. Our model shares some similarities with the mixture SV model in Durham (2007), where the proportion, indicating which mixture component is used to generate the innovation of the mean equation, is estimated. In contrast, in our model the two regime distributions are determined by an unknown threshold parameter value to be estimated from the data.

A review of the estimation of the SV models is given in Broto and Ruiz (2004) and Chib et al. (2009). As mentioned in So et al. (2002), Chen et al. (2008) and, more recently, Xu (2012), the estimation of the TSV model’s parameters is even more challenging than the estimation of the SV models. Although our formulations of the TSV model and MCMC algorithm lend themselves to high tractability in terms of parameter estimation compared to other existing approaches, it is still computationally quite intensive. So, to keep the estimation burden at a reasonable level, in this paper we do not assume that $\epsilon_t$ and $u_{t+1}$ are correlated. However, our MCMC estimation method can in principle be extended to the case with a non-zero correlation. Also, a heavy tailed distribution, such as a Student’s $t$-distribution, can be readily imposed on the innovation of $\epsilon_t$. We leave these extensions as avenues for future research.

**Bayesian analysis**

It is important to stress that the main advantage of our proposed TSV model is that the estimation part is highly tractable. So et al. (2002) and Chen et al. (2008) derive cumbersome MCMC methods for their TSV models, where
the latent states are sampled by using the method in Kim et al. (1998) and complemented by use of the Kalman filter. In this paper, we propose a simple MCMC algorithm to fit our proposed TSV model. As usually the case with a Bayesian estimation approach, we treat all of the latent states as parameters in the estimation procedure. In the MCMC approach, the state variables are sampled through the corresponding posterior distributions, which are known as the full conditionals. As discussed in Jacquier et al. (2004), the quality of the MCMC estimation methods for the SV models largely depends on the simulation efficiency of the latent states. In other words, the sampling of the log volatilities is critical to the performance of the proposed procedure as a whole. Our experience shows that this argument also extends to the TSV model. In our developed MCMC algorithm, the latent states are sampled one at a time through an MH method with a univariate normal distribution as the proposal. We find that for each individual state, the MH algorithm can result in a moderately high acceptance rate. As in any Bayesian estimation approach, the prior distributions of the parameters contained in the TSV model have to be specified in advance. Since we do not have much information on these priors, we use non-informative priors for all of the parameters, except for the persistence parameter \( \phi \), where we apply a uniform distribution over the interval \((-1, 1)\) to ensure the weak stationarity of the AR(1) process (6).

Now, we are ready to obtain the full conditionals for each parameter and individual latent state. Let \( y \) be the collection of all observations and \( h \) be the vector of all of the latent states. Define \( \theta = (\phi, \sigma, \lambda_1, \lambda_2, r) \), which is a vector of parameters of the TSV model. Instead of sampling \( \phi \) directly, we sample \( \sigma^2 \), since the latter has an inverse Gamma posterior distribution. The MCMC algorithm for the TSV model is outlined in Table 1, followed by a detailed description of each step. Some of the full conditionals can be found in the Appendix.

**Step 1. Sample \( h \).** It is easy to see that the conditional distribution of \( y_t \) depends on the previous observation \( y_{t-1} \) and the threshold parameter, \( r. \) If \( y_{t-1} < r \), the full conditional of \( h_t \) is given by

\[
f(h_t | y, y_{t-1} < r, \theta) \propto \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2\lambda_1^2} \right\} \times \exp\left\{ -\frac{(h_t - \mu_t)^2}{2\sigma_t^2} \right\}
\]

where

\[
\mu_t = -\frac{\sigma_t^2}{2} + \frac{\phi(h_{t-1} + h_{t+1})}{1 + \phi^2}, \quad \sigma_t^2 = \frac{\sigma^2}{1 + \phi^2}
\]

Thus the full conditional of \( h_t \) can be sampled by the MH method with the univariate normal distribution \( \mathcal{N}(\mu_t, \sigma_t^2) \) as the proposal. The full conditional of \( h_t \), when \( y_{t-1} \geq r \), can be obtained in a similar way.

**Step 2. Sample the threshold parameter, \( r \).** Given that all other parameters and latent states in the model have been sampled, the full conditional of \( r \) can be expressed as follows:

\[
f(r | y, \theta_{-r}, h) \propto \prod_{t=2, y_{t-1} < r}^T \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2\lambda_1^2} \right\} \prod_{t=2, y_{t-1} \geq r}^T \exp\left\{ -\frac{y_t^2 \exp(-h_t)}{2\lambda_2^2} \right\}
\]

where \( \theta_{-r} \) defines a vector of parameters excluding \( r \). The full conditional of \( r \) is not in the form of any standard distribution. Therefore, we use a random-walk MH algorithm with the target density given in Eq. 10 to simulate \( r \), where the proposal distribution is a univariate normal distribution with the mean value given by the previous iteration, and the variance is adjusted so that the acceptance rate can be tuned. A similar method is proposed in Chen et al. (2008). In this paper, the variance is adjusted to obtain an acceptance rate between 20% and 50%. Such rates are commonly accepted in the Bayesian literature. In order to ensure that there is enough data in each of the two regimes, we require that \( r \) has a uniform prior distribution over the interval between the first and the third quartiles of the observed data. This prior distribution assumption for the threshold value is also used in Chen et al. (2008).

**Step 3. Sample \( \phi \) and \( \sigma^2 \).** The full conditional of \( \phi \) is a univariate normal distribution truncated in the interval \((-1, 1)\), which can be simulated by using statistical software. The full conditional of \( \sigma^2 \) is an inverse Gamma distribution, from which the sampling can be done relatively easily. The derivation of the full conditionals of \( \phi \) and \( \sigma^2 \) is not given in this paper, as it can be found in other published papers on stochastic volatility (e.g. in Kim et al., 1998).
Step 4. Sample $\lambda_1$ and $\lambda_2$. It is obvious that the full conditionals of these two parameters depend on $y$, $h$ and $r$. In our implementation, instead of sampling $\lambda_1$ and $\lambda_2$ directly, we sample $\lambda_1^2$ and $\lambda_2^2$. Since the full conditionals of $\lambda_1^2$ and $\lambda_2^2$ are similar, below we only present the full conditional of $\lambda_1^2$:

$$f (\lambda_1^2 | y, \theta_{-\lambda_1}, h) \propto \frac{1}{(\lambda_1^2)^{T_1 + 1/2}} \exp \left\{ -\frac{1}{2} \sum_{t=2}^{T} \frac{y_t^2 \exp(-h_t)}{\lambda_1^2} \right\}$$

(11)

In the last line we use $T_1 = \sum_{t=2}^{T} I(y_{t-1} < r)$, where $I(.)$ is an indicator function. The full conditional of $\lambda_1^2$ has an inverse Gamma distribution:

$$\lambda_1 \sim IG \left( \frac{T_1}{2} - 1, \beta \right), \quad \beta = \frac{1}{2} \sum_{t=2}^{T} \left[ y_t^2 \exp(-h_t) \right]$$

(12)

Now, we can sample $\lambda_1^2$ from equation (12) and then obtain $\lambda_1$. The sampling of $\lambda_2$ can be done in the same fashion.

In the implementation of the four MCMC steps described above, we adopt an exact sequential order. Using a randomized order could in principle shorten the length of the dependence in the simulated histories and thus save computing time. As we use the C++ programming language to perform the estimation process, the running time is not a potential concern. On an Asus laptop computer with two 2.40 Hz Intel Cores, our estimation process took only about 40 minutes for 100,000 iterations. Our sequential algorithm is arguably consistent with the literature for Bayesian inference of the ARCH/GARCH, SV, stochastic conditional duration (SCD) and the threshold SCD models. As an attempt to remove the possible serial correlation in the posterior sample, we retain one draw for every five draws during the MCMC iterations. The C++ codes and data are available from the authors upon request.

MODEL ASSESSMENT AND MODEL SELECTION

In statistical practice, model assessment is conducted to determine whether the estimated model fits the corresponding data, and model selection is carried out to choose a better model for a given dataset. Both procedures require the calculation of the model likelihood. Obtaining the likelihood function of the TSV model requires a multivariate integration to remove the latent states, which, for large datasets, must be done numerically. One such method is an auxiliary particle filter (APF), introduced by Pitt and Shepard (1999). In the context of the SV model, this approach has proven to be an effective tool to approximate recursively the filter and prediction distributions of the latent states, from which the model likelihood can be calculated by successive conditioning. Below we adapt this APT procedure to the TSV model.

The sample likelihood function of the TSV model can be obtained via successive conditioning:

$$f(y | \theta) = f(y_1 | \theta) \prod_{t=2}^{T} f(y_t | F_{t-1}, \theta)$$

(13)

where $F_t = (y_1, \ldots, y_t)$ is the information flow known at time $t$. The conditional density of $y_{t+1}$ given $\theta$ and $F_t$ has the following expression:

$$f(y_{t+1} | F_t, \theta) = \int f(y_{t+1} | h_{t+1}, \theta) dF(h_{t+1} | F_t, \theta)$$

$$= \int f(y_{t+1} | h_{t+1}, \theta) \pi(h_{t+1} | h_t, \theta) dF(h_t | F_t, \theta)$$

(14)

As mentioned earlier, since the analytic expression of the conditional density of $y_{t+1}$ is difficult to obtain, we resort to the APF to calculate Eq. 14 numerically.

To describe the APF for the TSV model, suppose that we have a sample of particles $\{h_t^{(i)}, i = 1, \ldots, N\}$ of $h_t$ from the filter distribution of $(h_t | F_t, \theta)$ with weights $\{\pi_{tt}, i = 1, \ldots, N\}$, such that $\sum_{i=1}^{N} \pi_{tt} = 1$. Then, the one-step-ahead prediction density of $h_{t+1}$ is approximated as follows:

$$f(h_{t+1} | F_t, \theta) \approx \sum_{i=1}^{N} \pi_{tt} f(h_{t+1} | h_t^{(i)}, \theta).$$

(15)
Using this approximation, the one-step ahead prediction distribution of \( h_{t+1} \) can be easily sampled, and, hence, the conditional density (14) can be evaluated numerically by

\[
f(y_{t+1}|\mathcal{F}_t, \theta) \approx \sum_{i=1}^{N} \pi_{ti} f(y_{t+1}|h_{t+1}^{(i)}, \theta)
\]

where \( h_{t+1}^{(i)} \) are particles from the approximation (15) of the prediction distribution of \( (h_{t+1}|\mathcal{F}_t, \theta) \).

It is valid to evaluate Eqs. 15 and (16) numerically, since \( h_{t+1} \) in Eq. 15 follows a mixture of normal distributions. We should note that Eq. 15 can also be used for volatility forecasting, meaning that volatilities can be predicted via the APF simultaneously. Details on APF can be found in many studies, such as in Chib et al. (2006) for SV models, and in Men et al. (2012) for stochastic conditional duration models.

For the assessment of the overall fit of the model, we apply probability integral transforms (PITs) proposed in Diebold et al. (1998) to the TSV model. Suppose that \( \{f(y_t|\mathcal{F}_{t-1})\}_{t=1}^{T} \) is a sequence of conditional densities of \( y_t \) that governs the time series of \( y_t \), and \( \{p(y_t|\mathcal{F}_{t-1})\}_{t=1}^{T} \) is the corresponding sequence of one-step-ahead density forecasts. The PIT of \( y_t \) is defined as

\[
u(t) = \int_{-\infty}^{y_t} p(z|\mathcal{F}_{t-1}) \, dz
\]

As proved in Diebold et al. (1998), under a null hypothesis that the sequence \( \{p(y_t|\mathcal{F}_{t-1})\}_{t=1}^{T} \) coincides with \( \{f(y_t|\mathcal{F}_{t-1})\}_{t=1}^{T} \), the sequence \( \{\nu(t)\}_{t=1}^{T} \) are i.i.d. and uniformly distributed on the interval (0,1). In practice, for a given dataset, the uniformity of the distribution of \( \{\nu(t)\}_{t=1}^{T} \) can be tested using a Kolmogorov–Smirnov (KS) test.

For the TSV model, the PITs can be calculated from the formulas presented below:

\[
u(t) = \frac{1}{N} \sum_{i=1}^{N} \int_{-\infty}^{y_t} \frac{1}{\sqrt{2\pi\lambda_1}} \exp\left(-\frac{z^2}{2\lambda_1}\right) \, dz
\]

\[
u(t) = \frac{1}{N} \sum_{i=1}^{N} \int_{-\infty}^{y_t} \frac{1}{\sqrt{2\pi\lambda_2}} \exp\left(-\frac{z^2}{2\lambda_2}\right) \, dz
\]

In the computation of \( \nu(t) \), \( h_{t+1}^{(i)} \) are particles from the corresponding prediction distribution of \( h_t \) with weights \( 1/N \). In this paper, whenever the APF is implemented, we find that 3000 particles are sufficient to give us satisfactory results, but, to guarantee high accuracy of the presented results, we use 10,000 particles in all our applications.

Once the likelihood for the model has been obtained, a model comparison between competing models can be conducted. This is usually done by using the Akaike information criterion (AIC) (Akaike, 1987) and Bayesian information criterion (BIC) (Schwarz, 1978). It is well known that AIC tends to select a model with a larger number of parameters, whereas BIC tends to prefer a model with a smaller number of parameters. In addition, both AIC and BIC require an exact specification of the number of model parameters for the criterion calculation.

In the Bayesian inference of our TSV model, all of the latent states are estimated as parameters. Since the states are highly correlated, it is not reasonable to treat each state as an independent parameter. Hence we argue that AIC and BIC are not appropriate measures for model selection for SV and TSV models. Devising a generic model comparison criterion that works both as a measure of fit and as a measure of model complexity is a difficult problem. To overcome this obstacle, Spiegelhalter et al. (1998) propose a deviance information criterion (DIC), which is a more appropriate measure for assessing model fit for SV and TSV models. In this paper, following the literature for SV models, the latent states in \( z \) are treated as parameters and we use a modified measure of the DIC. Define

\[
D(\theta, z) = -2 \log f(y|\theta, z)
\]

as the deviance of the model, where \( f(y|\theta, z) \) is the conditional density of the model. The deviance defined in Eq. 19 is consistent with our MCMC estimation algorithm since we augment \( z \) as a vector of parameters. Then, the DIC for our TSV model is calculated as

\[
DIC = \hat{D} + P_D
\]

The first term \( \hat{D} \) is a Bayesian measure of model fit, defined as the posterior mean of the deviance

\[
\hat{D}(\theta, z) = E_{[\theta, x_N]}[D(\theta, z)]
\]

The larger the value of \( \hat{D} \), the worse is the model fit. The second term is defined as

\[
P_D = \hat{D} - D(\hat{\theta}, \hat{z})
\]

\[
= E_{[\theta, x_N]}[D(\theta, z)] - D(E[\theta, z|y])
\]

where \( D(\hat{\theta}, \hat{z}) \) is the deviance at the posterior mean of \( (\theta, z) \). It measures the degree of complexity of the model.
In other words, \( P_D \) represents a trade-off between the posterior mean of the deviance of the model and the deviance under the posterior mean of \((\theta, z)\). The larger \( P_D \) is, the easier it is for the model to fit the data. The term \( P_D \) is usually called the effective number of parameters.

Some comments on the use of DIC are in order. First, one of the merits of the DIC is that it does not depend on the number of parameters directly. This can be seen in the calculation of \( P_D \). Comparing with the penalty functions in the AIC and BIC, \( P_D \) can be viewed the penalty for the inclusion of model parameters. Second, as the DIC in Spiegelhalter et al. (2002) mostly deals with generalized linear models, to apply this criterion to more complicated models such as the SV and TSV models a modified version of DIC has to be considered. As discussed in Celeux et al. (2006), there are eight different ways for the calculation of DIC for missing data models depending on the ‘focus’, which is the interest of parameters. In their paper, the authors point out that the ‘focus’ is also applicable for hierarchical models. Because the SV models, including our proposed TSV model, are hierarchical, our calculation of DIC in Eq. 19 is suitable. It is also argued in Celeux et al. (2006) that \( P_D \) may have negative values for some ‘focus’ of interest. Lastly, Berg et al. (2004) use this DIC criterion for model comparison of univariate SV models, while Men et al. (2012) use DIC for model selection of stochastic conditional duration models. In Berg et al. (2004), the software WinBUGS is used for the fit of various SV models. The DIC was integrated in WinBUGS and calculated in the same way that we used in this paper.

**SIMULATION STUDIES**

We assess the proposed TSV model and the MCMC algorithm via simulation studies. There are two cases of the TSV model to be considered. One is to assume that the threshold parameter value is a known constant fixed at zero. In other words, we let the distribution of the mean innovation change with the signs of the previously observed returns. That is, the bad news and good news from the market will affect the distribution of the next observed asset return asymmetrically. In the other case, we consider the threshold parameter value as unknown to be estimated by the proposed MCMC algorithm from the data. In each of the two cases, we simulate 5000 observations of returns artificially, using the given parameters. The proposed estimation method is applied to fit the TSV model to the return time series formed by the first 4000 simulated returns. The remaining 1000 observations are used for the comparison of out-of-sample one-step-ahead predicted volatilities. The derived MCMC algorithm is iterated 100,000 times. After discarding the first 50,000 samples as the burn-in, we select one value from each five-observation gap window. The purpose for doing this is to remove the potential serial correlation in the posterior sample. We then obtain a sample with 10,000 draws for each of the posterior distributions. The parameters and latent states are estimated by sample means. Table II provides a comparison between the true and estimated parameters, based on the simulated asset return data with standard errors and Bayesian highest probability density (HPD) intervals. It is evident that, with relatively small standard deviations, the estimated parameters are close to their true values. Figure 1 depicts the paths and histograms of the simulated time series from the full conditionals of parameters.

To assess whether the fitted model agrees with the simulated return data, we test the PITs carried out through the APF procedure. In the implementation of the APF algorithm, there are 10,000 particles used to approximate the filter and the one-step-ahead prediction distributions of the states when the likelihood of the model and the volatility forecasting are performed recursively. Figure 2 presents Goodness-of-fit test via a scatter plot and histogram of the PITs from the fitted TSV model of the second case. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which was detailed in Diebold et al. (1998). In Figure 3, we compare the empirical cumulative

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True</th>
<th>Est.</th>
<th>SD</th>
<th>HPD CI (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: TSV model with ( r = 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.94</td>
<td>0.9326</td>
<td>0.0218</td>
<td>(0.8895, 0.9710)</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.12</td>
<td>0.1378</td>
<td>0.0276</td>
<td>(0.0894, 0.1916)</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>0.5</td>
<td>0.5124</td>
<td>0.0119</td>
<td>(0.4894, 0.5363)</td>
</tr>
<tr>
<td>( \lambda_2 )</td>
<td>1.2</td>
<td>1.2170</td>
<td>0.0283</td>
<td>(1.1627, 1.2736)</td>
</tr>
<tr>
<td>Panel B: TSV model with ( r ) flexible</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \phi )</td>
<td>0.94</td>
<td>0.9392</td>
<td>0.0174</td>
<td>(0.9056, 0.9725)</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>0.12</td>
<td>0.1281</td>
<td>0.0228</td>
<td>(0.0767, 0.1685)</td>
</tr>
<tr>
<td>( r )</td>
<td>0.02</td>
<td>0.0211</td>
<td>0.0015</td>
<td>(0.0172, 0.0226)</td>
</tr>
<tr>
<td>( \lambda_1 )</td>
<td>0.5</td>
<td>0.5121</td>
<td>0.0133</td>
<td>(0.4870, 0.5389)</td>
</tr>
<tr>
<td>( \lambda_2 )</td>
<td>1.2</td>
<td>1.2130</td>
<td>0.0315</td>
<td>(1.1514, 1.2750)</td>
</tr>
</tbody>
</table>
Figure 1. Time series and histograms of samples from the full conditionals of parameters in the TSV model based on the simulated return data

Figure 2. Goodness-of-fit test via a scatter plot (top) and a histogram (bottom) of the PITs produced by the fitted TSV model based on simulated return data. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which was detailed in Diebold et al. (1998). The KS test statistic is 0.0059 with the corresponding p-value 0.999; so we cannot reject the null hypothesis that the fitted flexible TSV model agrees with the generated return data. This suggests that our proposed TSV model and the estimation MCMC algorithm work well. An alternative way to assess the fitted model is to compare the time series of the artificially simulated absolute returns with the estimated (or filtered) volatilities, and the one-step-ahead in-sample and out-of-sample forecast volatilities. It is observed that the filtered and out-of-sample forecast volatilities have similar trends of the dynamics of the absolute artificial asset returns, except that the magnitude of the artificial asset returns is higher. Figure 4 provides this comparison, where the time series before and after the virtual dotted line are the in-sample and out-of-sample forecast volatilities using the fitted TSV model based on the first 4000 generated returns.

Copyright © 2016 John Wiley & Sons, Ltd

In order to check the stability and robustness of our proposed MCMC estimation methods, we carried out two sets of simulation studies. First, we simulated artificial asset returns sets, with each of them having a sample size of 2500 from the TSV model, with various parameter values. The TSV model was then fitted respectively to the simulated datasets. We performed 500 replications for this exercise. In order to further account for the uncertainty originated from the parameter initialization, we let the estimation process randomly choose the initialization values in each replication. Because a burn-in of sampled values was dropped off before making the estimation, we believe that the initialization uncertainty of the estimation process is negligible. The true parameters are compared to the estimated parameters in Table III, with standard deviations given in parentheses. It can be seen that the estimated parameters are very close to the true values used for the asset return simulation. Second, in Table IV, we simulated artificial asset returns sets with a sample size of 2000, 2500 and 3000, respectively, from the TSV model with fixed parameter values. The TSV model was then fitted respectively to these datasets. Again, we carried out 500 replications for this exercise and provide the estimated parameters with standard deviations reported in the table. Tables III and IV show that our estimation methods are quite stable and robust in terms of the parameter values and the sample sizes.

Simulation studies illustrate that our proposed TSV model with known and unknown threshold parameters work well in terms of capturing the asymmetry of the simulated return time series. The developed MCMC estimation procedure can recover the true parameters used to artificially generate the return data more precisely. As the latent states are augmented as parameters and estimated during the Bayesian inference process, this provides us with a straightforward way of conducting volatility forecasts.

APPLICATIONS

In this section, we apply the proposed TSV model and the developed MCMC algorithm to three datasets of asset returns. The first dataset includes the returns of CAC 40 index from 3 February 1990 to 4 January 2013, containing
Table III. Monte Carlo simulation with the same sample size of simulated artificial asset returns

<table>
<thead>
<tr>
<th>Parameter</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi$</td>
<td>$\sigma$</td>
<td>$\lambda_1$</td>
<td>$\lambda_2$</td>
<td>$r$</td>
</tr>
<tr>
<td>True</td>
<td>0.94</td>
<td>0.18</td>
<td>0.5</td>
<td>1.2</td>
<td>0.0</td>
</tr>
<tr>
<td>Est.</td>
<td>0.920</td>
<td>0.208</td>
<td>0.502</td>
<td>1.203</td>
<td>0.002</td>
</tr>
<tr>
<td>SD</td>
<td>(0.024)</td>
<td>(0.035)</td>
<td>(0.019)</td>
<td>(0.046)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>True</td>
<td>0.97</td>
<td>0.13</td>
<td>0.5</td>
<td>0.7</td>
<td>0.04</td>
</tr>
<tr>
<td>Est.</td>
<td>0.961</td>
<td>0.148</td>
<td>0.505</td>
<td>0.708</td>
<td>0.042</td>
</tr>
<tr>
<td>SD</td>
<td>(0.012)</td>
<td>(0.023)</td>
<td>(0.025)</td>
<td>(0.037)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>True</td>
<td>0.96</td>
<td>0.20</td>
<td>0.6</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Est.</td>
<td>0.952</td>
<td>0.215</td>
<td>0.606</td>
<td>0.807</td>
<td>0.301</td>
</tr>
<tr>
<td>SD</td>
<td>(0.013)</td>
<td>(0.025)</td>
<td>(0.033)</td>
<td>(0.045)</td>
<td>(0.027)</td>
</tr>
</tbody>
</table>

Note: The table reports the results of the simulation experiments. For each TSV model, we vary the values of parameters and fix the sample size at 2500 of the simulated artificial asset returns. The TSV model is then fitted to the datasets by the proposed MCMC estimation method, and the average estimated parameter values are presented. The standard deviations are reported in parentheses.

Table IV. Monte Carlo simulation with different sample sizes of simulated artificial asset returns

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameter</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\phi$</td>
<td>$\sigma$</td>
<td>$\lambda_1$</td>
<td>$\lambda_2$</td>
<td>$r$</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>True</td>
<td>0.94</td>
<td>0.18</td>
<td>0.5</td>
<td>1.2</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>0.920</td>
<td>0.207</td>
<td>0.502</td>
<td>1.205</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(0.031)</td>
<td>(0.035)</td>
<td>(0.020)</td>
<td>(0.050)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>2500</td>
<td>True</td>
<td>0.96</td>
<td>0.20</td>
<td>0.6</td>
<td>0.8</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>0.921</td>
<td>0.206</td>
<td>0.499</td>
<td>1.203</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(0.024)</td>
<td>(0.036)</td>
<td>(0.018)</td>
<td>(0.048)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>3000</td>
<td>True</td>
<td>0.96</td>
<td>0.196</td>
<td>0.503</td>
<td>1.208</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>Est.</td>
<td>0.927</td>
<td>0.196</td>
<td>0.503</td>
<td>1.208</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>(0.020)</td>
<td>(0.027)</td>
<td>(0.019)</td>
<td>(0.044)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Note: The table reports the results of the simulation experiments. For the TSV model, we fix the values of parameters and simulate artificial asset returns data sets with sample sizes of 2000, 2500 and 3000, respectively. The TSV model is then fitted to the datasets by the proposed MCMC estimation method, and the average estimated parameter values are presented. The standard deviations are reported in parentheses.

Figure 5. Goodness-of-fit test via the scatter plot (top) and the histogram (bottom) of the PITs produced by the fitted flexible threshold SV model to the return data of the CAC 40 index. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which was detailed in Diebold et al. (1998)
Figure 6. Comparison between the empirical CDF of the PITS and the theoretical CDF of a uniform distribution over the interval (0,1) based on the fitted flexible threshold SV model to the return data of the CAC 40 index. The two horizontal lines in the histogram plot are the 95% confidence intervals of the uniformity, constructed under the normal approximation of a binomial distribution, the calculation of which was detailed in Diebold et al. (1998).

Figure 7. Comparison between true, estimated and one-step-ahead forecast volatilities based to the flexible threshold SV model fitted to the return data of the CAC 40 index.

5784 observations of returns, in which the first 5000 observations are fitted by the TSV model, while the remaining 784 observations are used for volatility forecasts. The second dataset, which is used in the previous studies, consists of 945 observations on the daily pound/dollar exchange rate from 1 October 1981 to 28 June 1985, called EXC hereafter. This dataset has been analyzed in Harvey et al. (1994), Shephard and Pitt (1997), Meyer and Yu (2000), Skaug and Yu (2008) and Yu (2011), respectively. Since there are not many observations in this dataset, we fit all of the observations by the proposed threshold SV model and compare only the absolute observed returns with in-sample forecast volatilities and the estimated volatilities from the MCMC estimation process. The last dataset includes the daily returns of the Australian All Ordinaries stock index, called AUX in short hereafter. The dataset contains 1508 observations from 2 January 2000 to 30 December 2005, excluding weekends and holidays. For comparison purposes, the first 1200 observations are fitted by the proposed threshold SV model, and the remaining 308 observations are used for comparison with the forecast volatilities.

We perform 100,000 iterations of the algorithm. To allow for a burn-in period to eliminate the initial value influences that may affect the final estimates of parameters, we discard the first 50,000 sampled points. We selected one

---

1 We thank Professor Xibin Zhang for kindly providing us with these data, which were analyzed in Zhang and King (2008).
sample point after each five sampled points from the reserved 50,000 sampled points, and then estimated the parameters and latent states by sample means. Tables V, VI and VII report the estimated parameters of the TSV models with known and unknown threshold parameter values for the three datasets. Since the analysis results are similar for all of the three datasets, here we only provide the assessment for the fitted flexible TSV model (i.e. the TSV model with an unknown threshold parameter value) for the CAC 40 index data. We provide scatter plots and histograms of the PITs from the fitted model in Figure 5. The comparison between the empirical CDF of the PITs with the theoretical CDF of a uniform distribution over the interval (0, 1) is given in Figure 6. The $p$-value of the KS test is 0.0573. Thus we cannot reject the null hypothesis that the fitted TSV model agrees with the empirical return data at the 5% significance level. Figure 7 provides additional checks on the fitted model by comparing favourably estimated and one-step-ahead forecast volatilities based to the fitted flexible threshold SV model on the return data of the CAC 40 index.

For model selection, after fitting a TSV model by our MCMC estimation method to the dataset, we calculate the three components of the DIC measure: $D$, $P_D$ and DIC, respectively. These values are listed in Table VIII under panel A and B, respectively. As discussed earlier, the DIC criterion does not depend on the number of parameters of the

### Table V. True and estimated parameters of the TSV model based on the CAC 40 index return data

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>SD</th>
<th>HPD CI (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: $TSV$ model with $r = 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9862</td>
<td>0.0030</td>
<td>(0.9800, 0.9919)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1304</td>
<td>0.0105</td>
<td>(0.1106, 0.1516)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.0118</td>
<td>0.0006</td>
<td>(0.0108, 0.0131)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0111</td>
<td>0.0006</td>
<td>(0.0101, 0.0122)</td>
</tr>
<tr>
<td>Panel B: $TSV$ model with $r$ flexible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9858</td>
<td>0.0032</td>
<td>(0.9791, 0.9916)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1348</td>
<td>0.0116</td>
<td>(0.1120, 0.1572)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.0122</td>
<td>0.0008</td>
<td>(0.0109, 0.0136)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.0111</td>
<td>0.0007</td>
<td>(0.0099, 0.0125)</td>
</tr>
</tbody>
</table>

### Table VI. True and estimated parameters of the TSV model based on the on the return data of daily exchange rates for UK sterling/US dollar from 1 October 1981 to 28 June 1985

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>SD</th>
<th>HPD CI (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: $TSV$ model with $r = 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9735</td>
<td>0.0132</td>
<td>(0.9471, 0.9961)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1840</td>
<td>0.0373</td>
<td>(0.1165, 0.2602)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.6722</td>
<td>0.0926</td>
<td>(0.5146, 0.8571)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.7079</td>
<td>0.0978</td>
<td>(0.5408, 0.9042)</td>
</tr>
<tr>
<td>Panel B: $TSV$ model with $r$ flexible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9726</td>
<td>0.0142</td>
<td>(0.9464, 0.9956)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1872</td>
<td>0.0418</td>
<td>(0.1226, 0.2653)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.6794</td>
<td>0.0897</td>
<td>(0.5327, 0.8830)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.7178</td>
<td>0.0932</td>
<td>(0.5661, 0.9259)</td>
</tr>
</tbody>
</table>

### Table VII. True and estimated parameters of the TSV model based on the Australian All Ordinaries stock index returns

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>SD</th>
<th>HPD CI (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: $TSV$ model with $r = 0$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9654</td>
<td>0.0162</td>
<td>(0.9344, 0.9952)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2044</td>
<td>0.0376</td>
<td>(0.1351, 0.2796)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.8330</td>
<td>0.1153</td>
<td>(0.6571, 1.0913)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.7719</td>
<td>0.1070</td>
<td>(0.6180, 1.1086)</td>
</tr>
<tr>
<td>Panel B: $TSV$ model with $r$ flexible</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.9602</td>
<td>0.0159</td>
<td>(0.9289, 0.9878)</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2169</td>
<td>0.0382</td>
<td>(0.1437, 0.2929)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>0.7814</td>
<td>0.0692</td>
<td>(0.6535, 0.9286)</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.6986</td>
<td>0.0688</td>
<td>(0.5719, 0.8444)</td>
</tr>
<tr>
<td>$r$</td>
<td>0.2653</td>
<td>0.2311</td>
<td>(0.2573, 0.4780)</td>
</tr>
</tbody>
</table>
model directly; thus it is appropriate for hidden Markov models such as the TSV model proposed in this paper. For comparison purposes, all three datasets are also fitted to the standard SV model proposed by Taylor (1986), and their DIC values are reported in Table VIII under panel C. It is evident that for each dataset the DIC attains the smallest values for the threshold SV model with an unknown threshold parameter value. This leads us to the conclusion that the flexible TSV model (i.e. the TSV model with an unknown threshold parameter value) is preferred to the TSV model with a threshold value fixed at zero. With a non-zero threshold parameter value, there are 1555 out of 5000 (31.1%), 444 out of 945 (46.98%) and 646 out of 1200 (53.83%) observations in regime 1, which is defined by $y_t < r$, for the CAC 40, EXC and AUX, respectively. In contrast, with the threshold parameter value fixed at zero, there are 2531 out of 5000 (50.62%), 469 out of 945 (50.37%) and 782 out of 1200 (65.17%) observations in regime 1 for the CAC 40, EXC and AUX return data, respectively. Clearly, under the flexible TSV model, the number of observations in regime 1 is less than that under the fixed TSV model (i.e. the TSV model with a known threshold parameter value) for the CAC 40 and EXC data, while for the AUX data there are more observations in regime 1 under the flexible TSV model. As expected, evidence in Table VIII shows that the DIC has the largest values for the SV model in its canonical form.

### CONCLUDING REMARKS

In this paper, we propose a parsimonious threshold SV model with high tractability in terms of parameter estimation of the model. Instead of letting the logarithm of the conditional volatility follow a threshold autoregressive process, as is done in the literature, we assume that the innovation in the observation equation is endowed with a threshold univariate normal distribution with mean zero and the variances depending on the regime. In the proposed model, the threshold is treated as an unknown parameter to be estimated via a Bayesian estimation procedure. Efficient MCMC algorithms are developed for estimation of the parameters of this model, in which the latent states are estimated as a by-product. Using the auxiliary particle filter, the one-step-ahead in-sample and out-of-sample volatility forecasting are carried out in a straightforward way as well. Simulation studies and applications to a number of financial return data indicate that the proposed threshold SV model and the MCMC algorithm work well in terms of parameter estimation and volatility forecasting.

### REFERENCES

The full conditional of $h_t$, given that $y_{t-1} < r$, is given by

$$f(h_t | y_t, y_{t-1} < r, \theta) \propto f(y_t | h_t, y_{t-1} < r, \theta) \cdot f(h_t | h_{t-1}, y_{t-1}, \theta) \cdot f(h_t | y_t, h_{t+1}, \theta)$$

$$\propto \exp\left\{-\frac{h_t}{2}\right\} \exp\left\{-\frac{y_t^2 \exp(-h_t)}{2\lambda_t^2}\right\}$$

$$\times \exp\left\{-\frac{(h_t - \phi h_{t-1})^2}{2\sigma^2}\right\} \exp\left\{-\frac{(h_{t+1} - \phi h_t)^2}{2\sigma^2}\right\}$$

$$\propto \exp\left\{-\frac{y_t^2 \exp(-h_t)}{2\lambda_t^2}\right\} \times \exp\left\{-\frac{(h_t - \mu_t)^2}{2\sigma_t^2}\right\}$$

where

$$\mu_t = \frac{\sigma_t^2}{2} + \phi(h_{t-1} + h_{t+1}), \quad \sigma_t^2 = \frac{\sigma^2}{1 + \phi^2}$$
Thus the full conditional of $h_t$ can be sampled by the MH method with the univariate normal distribution $\mathcal{N}(\mu_t, \sigma_t^2)$ as the proposal. The full conditional of $h_t$, when $y_{t-1} \geq r$, can be obtained in a similar way.

- The full conditional of $r$, given that all other parameters and latent states in the model have been sampled, can be expressed as follows:

$$f(r|y, \theta_{-r}, h) \propto \prod_{t=2, y_{t-1} < r}^{T} f(y_t|h_t, \theta) \prod_{t=2, y_{t-1} \geq r}^{T} f(y_t|h_t, \theta)$$

$$\propto \prod_{t=2, y_{t-1} < r}^{T} \exp\left\{-\frac{y_t^2 \exp(-h_t)}{2\lambda_i^2}\right\} \prod_{t=2, y_{t-1} \geq r}^{T} \exp\left\{-\frac{y_t^2 \exp(-h_t)}{2\lambda_i^2}\right\}$$

where $\theta_{-r}$ defines a vector of parameters excluding $r$.

- The full conditional of $\lambda_i^2$:

$$f\left(\lambda_i^2|y, \theta_{-\lambda}, h\right) \propto \prod_{t=2, y_{t-1} < r}^{T} f(y_t|h_t, \theta)$$

$$\propto \prod_{t=2, y_{t-1} < r}^{T} \frac{1}{\lambda_i} \exp\left\{-\frac{y_t^2 \exp(-h_t)}{2\lambda_i^2}\right\}$$

$$\propto \frac{1}{(\lambda_i^2)\frac{T_i}{2}} \exp\left\{-\frac{1}{2} \sum_{t=2, y_{t-1} < r}^{T} \frac{y_t^2 \exp(-h_t)}{\lambda_i^2}\right\}$$

where $T_i = \sum_{t=2}^{T} I(y_{t-1} < r)$, and $I(.)$ is an indicator function. The full conditional of $\lambda_i^2$ has an inverse Gamma distribution:

$$\lambda_i \sim \text{IG} \left(\frac{T_i}{2} - 1, \beta\right), \quad \beta = \frac{1}{2} \sum_{t=2, y_{t-1} < r}^{T} \left[\frac{y_t^2 \exp(-h_t)}{\lambda_i^2}\right]$$

The full conditional of $\lambda_2$ can be obtained similarly.

**Authors’ biographies:**

**Dr. Tony S. Wirjanto** is professor of financial econometrics at the department of Statistics & Actuarial Science and school of Accounting & Finance of University of Waterloo. His research interests include computational finance and financial econometrics.

**Dr. Adam W. Kolkiewicz** Kolkiewicz is a professor in statistics and financial mathematics in the Department of Statistics & Actuarial Science. In statistics, he has focused on time series analysis, robust methods of estimation, and asymptotic methods of inference. In financial risks, he has worked on valuation and hedging problems in the context of incomplete markets. He has also done research on numerical methods in finance, focusing on techniques for high-dimensional problems that use the dynamic programming approach.

**Dr. Zhongxian Men** is a Data Modeler at Title Source, Detroit, MI. His research interest lies in solving problems related to appraisal escalation process using machine learning techniques.

**Authors’ addresses:**

**Tony S. Wirjanto**, Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1; School of Accounting and Finance, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1.

**Adam W. Kolkiewicz** Title Source, 662 Woodward Ave Detroit, MI 48226 United States Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1.