Working paper

2019-03

Statistics and Econometrics

ISSN 2387-0303

Data cloning estimation for asymmetric stochastic volatility models

Patrícia de Zea Bermudez, J. Miguel Marín, Helena Veiga

uc3m | Universidad Carlos III de Madrid

Data cloning estimation for asymmetric stochastic volatility models

Patrícia de Zea Bermudez

Departamento de Estatística e Investigação Operacional

Faculdade de Ciências da Universidade de Lisboa

and CEAUL

J. Miguel Marín

Department of Statistics and Instituto Flores de Lemus

Universidad Carlos III de Madrid

Helena Veiga*

Department of Statistics and Instituto Flores de Lemus

Universidad Carlos III de Madrid

BRU-IUL, Instituto Universitário de Lisboa

March 19, 2019

*Email: mhveiga@est-econ.uc3m.es. Corresponding author.

Abstract

The paper proposes the use of data cloning (DC) to the estimation of general asymmetric stochastic volatility (ASV) models with flexible distributions for the standardized returns. These models are able to capture the asymmetric volatility, the leptokurtosis and the skewness of the distribution of returns. Data cloning is a general technique to compute maximum likelihood estimators, along with their asymptotic variances, by means of a Markov chain Monte Carlo (MCMC) methodology. The main aim of this paper is to illustrate how easily general ASV models can be estimated and consequently studied via data cloning. Changes of specifications, priors and sampling error distributions are done with minor modifications of the code. Using an intensive simulation study, the finite sample properties of the estimators of the parameters are evaluated and compared to those of a benchmark estimator that is also user-friendly. The results show that the proposed estimator is computationally efficient and robust, and can be an effective alternative to the exiting estimation methods applied to ASV models. Finally, we use data cloning to estimate the parameters of general ASV models and forecast the one-step-ahead volatility of S&P 500 and FTSE-100 daily returns.

Keywords: Asymmetric Volatility; Data Cloning; Non-Gaussian Nonlinear Time Series Models; Skewed and Heavy-Tailed distributions

1 Introduction

Complex parametric stochastic volatility models have been extensively proposed in the literature to cope with the main empirical facts of financial time series. Very recent examples are the models proposed by Mao et al. (2015, 2017) and Asai et al. (2017) that accommodate a general asymmetric function for volatility. Asymmetric effects have been traditionally modeled, 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] Mao et al. (2015) combine these two sources of asymmetry and propose a quite flexible model for financial returns.

On the other hand, heavy-tails in the distribution of returns are currently accepted as an empirical fact that should be included in the volatility models. Although SV models are able to generate more kurtosis than GARCH-type models, there is often higher kurtosis in the returns than models imply. One way to cope with this is to assume a heavy-tailed distribution for the standardized returns such as the Student-t distribution, which has been shown to fit better the returns than other fat-tailed distributions (see, e.g, Liesenfeld and Jung, 2000; Asai, 2008). Others show that the tail of a Student-t distribution is not heavy enough for modeling financial returns and besides asymmetric volatility, asymmetrically heavy-tailed distributions should also be considered for improving models' fit (Nakajima and Omori, 2012; Abanto-Valle et al., 2015).

Unfortunately, the extra flexibility of SV models makes parameter and volatility estimation difficult due to the intractable form of the likelihood function. During the last decades several methods have been proposed for the estimation of SV models, including generalized method of moments (Melino and Turnbull, 1990; Sørensen, 2000), quasi-maximum likelihood (Harvey et al., 1994), simulated maximum likelihood (Danielsson, 1994; Sandmann and

[1] For more examples on the negative correlation between returns and future volatility see Yu (2005), and 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.

Koopman, 1998), approximate maximum likelihood (Fridman and Harris, 1998) and MCMC procedures, among others (see, e.g., Jacquier et al., 1994; Shephard and Pitt, 1997; Kim et al., 1998). Amidst all these estimation methods, one of the best alternatives is to use an MCMC algorithm, although MCMC is very computationally demanding and requires hard code, which makes its implementation difficult for practitioners and researchers. A very recent paper by Barra et al. (2017), based on importance sampling weighted expectation maximization, argues that this approach is computationally efficient and can be regarded as an effective alternative to MCMC methods. However, it requires, again, heavy programming and there is not free available software that implements the method.

In this paper we propose the use of data cloning for ASV models. Data cloning is a general technique to obtain maximum likelihood estimators, along with their asymptotic variances, by means of a MCMC procedure (Lele et al., 2007, 2010). Its main strength lies in the easiness with which any changes in the model specification can be made. Furthermore, the method is not very influenced by the selection of the prior (proper) distributions of the parameters, assuming an appropriate number of clones (Lele et al., 2007, 2010).

For implementing data cloning, we use the free software `dclone` in a R framework and for the Bayesian approach we use JAGS (Plummer, 2003). However, the major weakness of the Bayesian MCMC implementation is the slow convergence and inefficiency in terms of simulation, given that it is based on a single move Gibbs sampling algorithm in the same way as WinBUGS (see Meyer and Yu, 2000; Yu, 2005, 2012, for WinBUGS implementation of SV models). Data cloning improves the accuracy of the parameter estimates of the models considered in the paper in comparison to that of the standard Bayesian approach, making it an effective alternative estimation method for flexible and complex SV models.

The finite sample properties of the proposed estimators are studied by means of an intensive Monte Carlo study and their performance is compared to that of a standard Bayesian approach. The simulation results show that the parameters are estimated more accurately with data cloning than with a standard Bayesian approach, and the differences between the estimates are quite significant for small samples. As the sample size increases the standard Bayesian approach improves its performance, although it often remains less accurate than data cloning when more complex distributions are used for the standardized returns.

Finally, the estimation performance of the proposed estimator for parameters and volatilities is illustrated by fitting three asymmetric stochastic volatility models to daily S&P 500 and FTSE-100 returns. The results allow to identify which type of asymmetric volatility best characterizes the data. We also provide one-day-ahead volatility forecasts and evaluate them against estimates of realized volatility obtained from high frequency data.

The remainder of the paper is organized as follows. Section 2 describes the models studied in the paper and the different sources of asymmetric volatility. Section 3 suggests data cloning for the estimation of flexible and complex asymmetric stochastic volatility models. Section 4 investigates the finite sample performance of the proposed estimator, and compares it to that of other Bayesian user-friendly procedure. An empirical application using two time series of financial returns is presented in Section 5 and finally Section 6 provides some concluding remarks.

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 general asymmetric autoregressive stochastic volatility (GA-SV) family of Mao et al. (2015) is given by

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

$$h_t - \mu = \phi(h_{t-1} - \mu) + m(\epsilon_{t-1}; \theta) + \eta_{t-1}, \tag{2}$$

where the log-volatility disturbance η_t is a Gaussian white noise with variance σ_η^2, $m(\epsilon_t;\theta)$ is any real parametric function of ϵ_t independent of η_t for all leads and lags and θ is the vector of parameters of this function.[2] The scale parameter μ is related with the marginal variance of the returns, while ϕ controls the rate of decay (towards zero) of the autocorrelations of power-transformed absolute returns, hence, the persistence of volatility shocks. Note that, in equations (1) and (2), the standardized return at time $t-1$ is correlated with the volatility at time t. Furthermore, if $m(\cdot;\theta)$ is not an even function, then positive and negative past standardized returns of the same magnitude have different effects on volatility (see Mao et al., 2015, for more details).

In this paper, we consider the following specifications of m that correspond to the mostly well-known forms of asymmetric volatility , and that are nested in the GA-SV family:

AARSV: $m(\epsilon_{t-1};\theta) = \gamma_1\epsilon_{t-1}$,

TSV: $m(\epsilon_{t-1};\theta) = \delta I(\epsilon_{t-1} < 0)$,

TGA-SV: $m(\epsilon_{t-1};\theta) = \delta I(\epsilon_{t-1} < 0) + \gamma_1\epsilon_{t-1} + \gamma_2(|\epsilon_{t-1}| - E(|\epsilon_{t-1}|))$.

The AARSV is the well-known asymmetric autoregressive SV model of Taylor (1994) and Harvey and Shephard (1996). TSV stands for a threshold SV model where only a constant δ is added (or not) to the model depending on the sign of past returns. $I(\cdot)$ is an indicator function that takes the value one when the argument is true and zero otherwise. This model is based on a restricted version of the threshold SV proposed by So et al. (2002), and TGA-SV corresponds to the threshold generalized asymmetric stochastic volatility model proposed by Mao et al. (2015), which includes also an EGARCH part as in Demos (2002) and Asai and McAleer (2011).

[2]The Gaussianity of η_t when $m(\epsilon_t;\theta) = 0$ is justified in Andersen, Bollerslev, Diebold, and Ebens (2001) and Andersen, Bollerslev, Diebold, and Labys (2001, 2003).

3 Data cloning estimation

The data cloning method is a computational technique that allows to obtain maximum likelihood (ML) estimators of parameters and their asymptotic variances, through a MCMC method (see, e.g, Lele et al., 2007, 2010). It is based on the intuitive idea of running an experiment several times and obtaining always the same observations. In this way, given some observed data $\mathbf{y} = (y_1, \ldots, y_n)$, a new dataset is built by means of making K clones of the original observations: $\mathbf{y}^{(K)} = (\mathbf{y}, \ldots, \mathbf{y})$. The clones are assumed to be independent of each other, and K is supposed to be large enough. Accordingly, the likelihood of $\mathbf{y}^{(K)}$ is equal to the K-th power of the likelihood of the original data $[L(\boldsymbol{\theta}|\mathbf{y})]^K$, where $\boldsymbol{\theta}$ is the vector of parameters.

In general, with data cloning, a MCMC procedure is applied by taking the cloned data $\mathbf{y}^{(K)}$ and multiplying the corresponding likelihood $[L(\boldsymbol{\theta}|\mathbf{y})]^K$ by a given proper prior distribution $\pi(\boldsymbol{\theta})$ of $\boldsymbol{\theta}$. Then, samples from the posterior distributions $\pi^{(K)}(\boldsymbol{\theta}|\mathbf{y})$ are generated. The selection or elicitation of a proper prior distribution is not very relevant and any distribution may be chosen. The mean of the posterior distribution of $\boldsymbol{\theta}$ approximates the ML estimator ($\hat{\boldsymbol{\theta}}$), when the number of clones K increases. Nevertheless, as Lele et al. (2010) point out, when the prior distribution is more informative, the convergence of the mean of the posterior distribution to the ML estimator is faster and a smaller number of clones K is needed. For K large enough, $\pi^{(K)}(\boldsymbol{\theta}|\mathbf{y})$ converges to a multivariate normal distribution whose mean is equal to the ML estimator of $\boldsymbol{\theta}$, and the covariance matrix equals $1/K$ times the inverse of the Fisher information matrix of the ML estimator of $\boldsymbol{\theta}$ (see the appendix of Lele et al., 2007, for details).

As a summary, the data cloning algorithm consists of 3 steps:

Step 1: Generate a K-cloned data set $\mathbf{y}^{(K)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$ in such a way that the observed data vector is repeated K times.

Step 2: Generate random deviates from the posterior distribution by means of a MCMC algorithm, with any proper prior distribution $\pi(\boldsymbol{\theta})$ and the cloned data vector $\mathbf{y}^{(K)} = (\mathbf{y}, \mathbf{y}, \ldots, \mathbf{y})$. The K copies of $\mathbf{y}$ are assumed to be independent of each other.

Step 3: Compute sample means and variances of the MCMC chains generated from the posterior distribution of the vector of parameters $\boldsymbol{\theta}$. The means of the posterior distributions approximate the ML estimates of parameters, and K times the variances of the posterior distributions coincide with the estimated asymptotic variances of the ML estimators of the parameters.

One of the main advantages of using data cloning alternative is its easy and friendly implementation in R, by means of the packages `dclone` (Solymos, 2010) and `dcmle` (Solymos, 2016)). The syntax adopted in `dclone` is similar to BUGS language.[3]

First, it is necessary to define the prior distributions of the models' parameters. A proper set of prior distributions can be assumed. We have considered weakly informative prior distributions based on the literature on the estimation of SV models:

$$\mu \sim \mathrm{N}(0,\ 10^3) \qquad \frac{1+\phi}{2} \sim \mathrm{Beta}(1,\ 1) \qquad \sigma_\eta^{-2} \sim \mathrm{Gamma}(10^{-3},\ 10^{-3})$$

$$\gamma_1 \sim \mathrm{N}(0,\ 10^3) \qquad \gamma_2 \sim \mathrm{N}(0.05,\ 10^3) \qquad \delta \sim \mathrm{N}(0.05,\ 10^3)$$

$$\nu \sim \mathrm{Gamma}(2,\ 0.1) \qquad \lambda \sim \mathrm{Gamma}(10^{-2},\ 10^{-2})$$

where ν and λ are the degrees of freedom and the asymmetry parameter of the skew-Student-t distribution, respectively. In the simulations and the empirical application the standardized returns are assumed to follow, for instance, a skew-Student-t distribution.

Second, we have determined the optimal number of clones to be used. The library `dclone` includes several diagnosis measures, such as the function `dcdiag` that calculates some statistics that assist the user in that choice. One of these statistics is the maximum eigenvalue of the posterior covariance matrix provided by Lele et al. (2010), which gives

[3]See code example in Appendix A.

us information about the degeneration of the posterior distribution; accordingly, if it is close to zero, prior distributions have small influence on the results. In our simulations and application, we use 15 clones. We have also experimented with a larger number of clones but the improvement in the estimates of the parameters is irrelevant. Other measures related with the selection of the optimal number of clones are: the mean squared error (MSE) and the R^2 statistic. They are both based on a χ^2 approximation, and they should converge to zero when the number of clones increases. As before, the optimal number of clones selected by these measures is 15.

3.1 Example of model implementation in data cloning

We use the wrapper included in the library `dcmle` in order to implement the SV models (see Solymos, 2016). Furthermore, we consider a `dcFit` object that includes the parameters of the model, the data, the name of the underlying instructions of a standard Bayesian program, in this paper JAGS that is proposed by Plummer (2003), and the variable that controls the number of clones. The corresponding code is:

```
library(dcmle)

DCloneObject = NULL

dataInitial = list(y=y, t_max=t_max, htmean0=htmean0, epsilon0=epsilon0)
parameters = c("mu", "gamma1", "phi", "sigeta2", "nu", "httot", "skew")

DCloneObject = new("dcFit")

DCloneObject@data = list(
y = dcdim(y),
t_max=t_max, htmean0=htmean0, epsilon0=epsilon0,
```

```
Ncl=1)


DCloneObject@model = "Instructions.txt"
DCloneObject@multiply = "Ncl"
DCloneObject@unchanged = c("t_max", "htmean0", "epsilon0")
DCloneObject@params = parameters



Ncluster = makePSOCKcluster(3)
m1 = dcmle(DCloneObject, n.clones=15, cl=Ncluster)
stopCluster(elcluster)
```

The Appendix A provides an example that runs data cloning for a particular ASV model.

4 Simulation study

In this section we conduct simulation experiments in order to assess the performance of the approximated ML estimators of the parameters of different asymmetric stochastic volatility models via data cloning, hereafter called DC estimator. We consider three possible designs for the Monte Carlo experiments depending on the error distribution of the standardized returns. They are either assumed to follow a $N(0,1)$, a Student-t with v degrees of freedom or a skew-Student-t distribution with parameters v and λ, where v corresponds to the degrees of freedom and λ to the skewness parameter (see Fernández and Steel, 1998, for details on the skew-Student-t distribution). Note that $\lambda = 1$ corresponds to a symmetric distribution, while $\lambda < 1$ implies negative skewness. The number of replicates is 200 and the full set of parameter values is $(\mu, \phi, \delta, \gamma_1, \gamma_2, \sigma^2_\eta, \nu, \lambda) = (0,\ 0.98,\ 0.07,\ -0.08,\ 0.1,\ 0.05,\ 5,\ 0.9)$. Finally, we consider four sample sizes $T = 500,\ 1000,\ 2000$ and 4000 and three asymmetric stochastic volatility models; see Section 2.

4.1 Finite sample properties of the data cloning and benchmark estimators

Table 1 reports the Monte Carlo averages and standard deviations of DC and the Bayesian parameter estimates when the errors are assumed to be Gaussian. The two estimators provide quite similar parameter estimates for large T and for the AARSV model, although the log-volatility intercept is, for the majority of the cases, estimated with more precision by DC. The most relevant differences between DC and Bayesian methodology are observed for small and moderate sample sizes. DC estimates of the parameters are often closer to the true values of the parameters and have smaller standard errors. Regarding the TSV and TGA-SV models, DC also provides more accurate estimates and more precision in the estimation of the parameters for small and moderate sample sizes. Even for large sample sizes some parameters are estimated with more precision when using the DC estimator. It seems that there is a correlation between the complexity of the model and the performance of the estimators, that is, the higher the complexity of the model, the larger the differences between the Bayesian and DC estimators are, in favor of the DC estimator.

Table 2 shows the results of the Monte Carlo simulation when the standardized returns follow a Student-t with five degrees of freedom. The degrees of freedom are jointly estimated with the rest of the parameters in all asymmetric stochastic volatility models. For this design, we observe an important improvement in the accuracy of the estimates provided by DC in comparison to that of the Bayesian approach. Regarding the AARSV and TSV models, we observe that Bayesian estimates σ_η^2 and v (the degrees of freedom of the Student-t distribution) very inaccurately and imprecisely when considering small and moderate sample sizes. Even for $T = 4000$ these estimates are much more accurate with DC than with the standard Bayesian approach. Furthermore, the precision in the estimation of the parameters is often higher with DC. Looking at the TGA-SV results, we observe that the Bayesian method has an enormous difficulty in estimating γ_2 and v. Once more, DC not only provides more accurate estimates but also more precision in the estimation of parameters.

Finally, Table 3 provides Monte Carlo averages and standard deviations of the parameter estimates when the standardized returns are skew-Student-t. The simulation results for the AARSV and TSV models show that for small and moderate sample sizes, the parameters γ_1, σ_η^2 and v are better estimated with DC. As the sample size increases, the Bayesian estimator improves in precision and accuracy, but it underperforms in comparison to the DC estimator, in particular, it is not able to estimate well σ_η^2 and the degrees of freedom of the skew-Student-t distribution. Regarding the TGA-SV, the estimation precision of DC estimator is higher than that of Bayesian, but since the model is quite complex, some parameters are not estimated with the expected accuracy. In this case a larger sample size is required.

5 Empirical application

We estimate the three asymmetric SV addressed in Section 2 using daily returns for two major stock market indexes, namely: S&P 500 and FTSE-100. The S&P 500 returns span the period between July 14, 2000 and May 21, 2018, while FTSE-100 returns range from August 10, 2000 till June 18, 2018. The total number of observations for both series is 4500.

Figure 1 shows the time series of the returns. The samples cover the period of the global financial crisis, which corresponds to a period of high volatility.

Table 4 presents the descriptive statistics of the returns. The empirical distribution of the returns is lepkurtic and skewed to the left, which supports the use of the skew-Student-t distribution to model the standardized returns.

5.1 Estimation results

We estimate the parameters and volatilities of the three asymmetric stochastic volatility models using a standard Bayesian approach and DC. Given the descriptive statistics of the data, we assume that the standardized returns follow a skew-Student-t distribution.[4]

Table 5 and Figure 2 show the estimation results: parameters and volatility. For precision, we estimate the constant of the log-volatility process as $\mu(1-\phi)$. Regarding the S&P 500 returns, we observe that DC generates often slightly smaller standard deviations of the parameters than the standard Bayesian approach. The estimates of the parameters are quite similar for the AARSV and the TSV models. This is expected given that the number of observations involved in the estimation is quite large. The largest difference is obtained for the estimation of the degrees of freedom of the skew-Student-t distribution. DC provides smaller parameter and standard deviation estimates. Regarding the TGA-SV model, the largest differences are in the estimation of the parameters δ and ν (again the degrees of freedom of the skew-Student-t distribution). DC provides smaller estimated standard deviations for these two parameters, which might indicate a more reliable estimation. In terms of goodness-of-fit, we observe that the Deviance Information Criterion (DIC) is slightly larger for the TGA-SV model. On the other hand, AARSV and TSV are identical in terms of goodness-of fit. It is also the TGA-SV model that provides the highest estimates of the volatility, specially in periods of high volatility.

Looking at the estimation results of the FTSE-100 returns, the largest differences are observed in the estimation of the parameters: δ, γ_2 and ν. DC estimates γ_2 with more precision than the standard Bayesian approach, but provides a smaller precision for the parameter of the degrees of freedom of the skew-Student-t distribution. All the models are similar in terms of goodness-of-fit. Finally, regarding the estimated volatility, we observe that in the peak of volatility of the last global financial crisis, the TSV model estimates more volatility than the other models, and this difference is slightly larger when using DC (see Figure 2).

[4] See Fernández and Steel (1998) for the specification of the skew-Student-t distribution

5.2 Forecasting results

We use a fixed rolling window scheme with 4400 observations. In total, for all models, we obtain 100 out-of-sample one-step-ahead forecasts.

The models' forecasting performance is evaluated using the model confident set (MCS) procedure proposed by Hansen et al. (2011) and programmed in R by Catania and Bernardi (2015). Hansen et al. (2011) procedure consists of a sequence of statistic tests to construct a set of models called "Superior Set Model" (SSM). Models in the SSM have statistically the same predictive ability and are ranked by the value of the loss function. The test statistic is calculated for an arbitrary loss function and evaluates point forecasts.

Table 6 displays the SSM of the forecasting models under study. The loss functions are the squared error (SE) and the Qlike of Patton (2011), which is robust to the presence of noise in the volatility proxy.[5] The results depend mainly on the loss function. The TGA-SV model is always ranked first for the S&P 500 returns. None of the models are excluded from the SSM, which implies they have a similar predictive ability. Nevertheless, depending on the loss function the models are ranked differently when the estimation method is the standard Bayesian approach. Regarding the FTSE-100 returns, the results are more stable. All the models belong to the SSM regardless the loss function and confidence level chosen, and the TSV is always ranked first followed by the TGA-SV model.

6 Conclusion

In this paper we propose the use of data cloning to the estimation of general asymmetric stochastic volatility models that are able to capture the asymmetric volatility, the leptokurtosis and skewness of the distribution of standardized returns. Data cloning is used together with a standard Bayesian approach, which improves the accuracy of the parameter estimates

[5]The SE is defined as $(RV_{t+1} - \widehat{\sigma_{t+1}})^2$ while the Qlike is defined as $Qlike \equiv \frac{RV_{t+1}}{\widehat{\sigma_{t+1}}} - \left(\frac{RV_{t+1}}{\widehat{\sigma_{t+1}}}\right) - 1$. RV_{t+1} corresponds to the realized volatility obtained with intradaily returns (5-minutes) calculated in https://realized.oxford-man.ox.ac.uk/data.

in comparison to the accuracy of the parameter estimates obtained from an alternative Bayesian method, making it an effective alternative estimation method for flexible and complex SV models. Furthermore and most important, it does not imply hard programming and the R packages are freely available.

The models that are estimated in this paper incorporate different sources of asymmetric volatility: the first models the asymmetric volatility through the correlation between the standardized returns and log-volatility, the second allows the constant of the log-volatility equation to change according to the sign of the returns, and the last model nests the first two and also includes a magnitude effect similar to that of the EGARCH model.

We analyze the finite sample properties of the data cloning and Bayesian estimators and show that data cloning leads to accurate estimates of the parameters specially for small and moderate samples sizes. Additionally, and even for large sample sizes, it provides more accurate estimates than those of a standard Bayesian approach when the models are quite complex.

Finally, we use data cloning to estimate the parameters of three asymmetric SV models and forecast the one-step-ahead volatility of S&P 500 and FTSE-100 daily returns. Data cloning often provides estimates of parameters that have less variance, when compared to those estimated with a standard Bayesian approach. Although for the FTSE-100 volatility forecasting the results provided by the two estimation methods are similar, when we consider the volatility forecasting of the S&P 500 returns, the estimation results are slightly different depending on the loss function used. Nevertheless, for this series of returns the TGA-SV model is always ranked first.

Acknowledgements

The first author was financially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the projects UID/MAT/00006/2013 and UID/MAT/00006/2019. The third author acknowledges financial support from Spanish Ministry of Economy and Competitiveness, research projects ECO2015-70331-C2-2-R and ECO2015-65701-P and from Fundação para a Ciência e a Tecnologia, grant UID/GES/00315/2013.

Figures and Tables

Table 1: Simulation results

Monte Carlo results of a standard Bayesian approach and DC estimators of ASV model's parameters with Gaussian errors. Monte Carlo averages of estimates and standard deviations are reported (in parenthesis). The first column of each parameter corresponds to the estimates obtained with a **standard Bayesian approach**, while the second column corresponds to the estimates obtained with **DC**.

Parameters	μ		ϕ		δ		γ_1		γ_2		σ^2_η	
True	**0**		**0.98**		**0.07**		**-0.08**		**0.10**		**0.05**	
AARSV												
$T = 500$												
Mean	-0.105	-0.043	0.969	0.968	–	–	-0.082	-0.083	–	–	0.068	0.060
Std.dev.	(1.124)	(0.624)	(0.025)	(0.017)	–	–	(0.042)	(0.041)	–	–	(0.029)	(0.027)
$T = 1000$												
Mean	-0.090	-0.021	0.975	0.974	–	–	-0.084	-0.085	–	–	0.061	0.058
Std.dev.	(0.501)	(0.377)	(0.011)	(0.010)	–	–	(0.028)	(0.027)	–	–	(0.018)	(0.017)
$T = 2000$												
Mean	0.023	0.024	0.977	0.977	–	–	-0.081	-0.081	–	–	0.054	0.053
Std.dev.	(0.265)	(0.265)	(0.018)	(0.010)	–	–	(0.006)	(0.006)	–	–	(0.017)	(0.010)
$T = 4000$												
Mean	0.000	-0.015	0.978	0.978	–	–	-0.081	-0.078	–	–	0.052	0.051
Std.dev.	(0.178)	(0.198)	(0.004)	(0.003)	–	–	(0.014)	(0.014)	–	–	(0.007)	(0.006)
TSV												
$T = 500$												
Mean	-0.376	0.402	0.945	0.947	0.100	0.101	–	–	–	–	0.110	0.097
Std.dev.	(2.507)	(1.621)	(0.044)	(0.037)	(0.087)	(0.088)	–	–	–	–	(0.064)	(0.058)
$T = 1000$												
Mean	-0.238	0.240	0.967	0.967	0.084	0.084	–	–	–	–	0.073	0.069
Std.dev.	(1.981)	(1.269)	(0.016)	(0.015)	(0.057)	(0.057)	–	–	–	–	(0.028)	(0.028)
$T = 2000$												
Mean	-0.027	0.170	0.974	0.974	0.079	0.080	–	–	–	–	0.060	0.059
Std.dev.	(1.163)	(0.979)	(0.008)	(0.008)	(0.040)	(0.040)	–	–	–	–	(0.016)	(0.016)
$T = 4000$												
Mean	-0.029	0.055	0.978	0.977	0.075	0.075	–	–	–	–	0.055	0.055
Std.dev.	(0.771)	(0.732)	(0.005)	(0.005)	(0.027)	(0.027)	–	–	–	–	(0.010)	(0.009)
TGA-SV												
$T = 500$												
Mean	-0.383	0.163	0.954	0.958	0.109	0.110	-0.066	-0.067	0.122	0.113	0.079	0.069
Std.dev.	(3.723)	(3.623)	(0.039)	(0.033)	(0.159)	(0.161)	(0.082)	(0.083)	(0.147)	(0.136)	(0.055)	(0.048)
$T = 1000$												
Mean	-0.635	-0.267	0.970	0.970	0.107	0.108	-0.066	-0.066	0.121	0.117	0.066	0.063
Std.dev.	(2.912)	(2.517)	(0.015)	(0.014)	(0.105)	(0.106)	(0.056)	(0.056)	(0.089)	(0.083)	(0.032)	(0.030)
$T = 2000$												
Mean	0.029	0.148	0.975	0.975	0.075	0.075	-0.079	-0.080	0.122	0.116	0.056	0.055
Std.dev.	(1.702)	(1.563)	(0.008)	(0.007)	(0.065)	(0.066)	(0.031)	(0.031)	(0.059)	(0.056)	(0.016)	(0.015)
$T = 4000$												
Mean	-0.049	-0.001	0.978	0.978	0.074	0.074	-0.079	-0.079	0.107	0.103	0.053	0.052
Std.dev.	(1.241)	(1.203)	(0.004)	(0.004)	(0.047)	(0.048)	(0.024)	(0.024)	(0.040)	(0.038)	(0.008)	(0.008)

Table 2: Simulation results

Monte Carlo results of a standard Bayesian approach and DC estimators of ASV model's parameters with Student-t standardized returns. Monte Carlo averages of estimates and standard deviations are reported (in parenthesis). The first column of each parameter corresponds to the estimates obtained with a **standard Bayesian approach**, while the second column corresponds to the estimates obtained with **DC**.

Parameters	μ		ϕ		δ		γ_1		γ_2		σ_η^2		ν	
True	**0**		**0.98**		**0.07**		**-0.08**		**0.10**		**0.05**		**5.00**	
AARSV														
$T = 500$														
Mean	0.029	-0.006	0.961	0.965	–	–	-0.092	-0.088	–	–	0.093	0.063	9.500	6.100
Std.dev.	(0.974)	(0.714)	(0.036)	(0.039)	–	–	(0.040)	(0.037)	–	–	(0.054)	(0.036)	(4.785)	(3.657)
$T = 1000$														
Mean	0.001	-0.005	0.974	0.974	–	–	-0.083	-0.082	–	–	0.064	0.055	6.484	5.422
Std.dev.	(0.484)	(0.369)	(0.014)	(0.011)	–	–	(0.027)	(0.026)	–	–	(0.028)	(0.022)	(2.185)	(1.148)
$T = 2000$														
Mean	0.023	0.005	0.977	0.977	–	–	-0.082	-0.081	–	–	0.057	0.053	5.769	5.294
Std.dev.	(0.266)	(0.264)	(0.007)	(0.007)	–	–	(0.016)	(0.016)	–	–	(0.015)	(0.013)	(1.151)	(0.822)
$T = 4000$														
Mean	-0.008	-0.053	0.979	0.978	–	–	-0.082	-0.078	–	–	0.053	0.056	5.296	5.074
Std.dev.	(0.190)	(0.199)	(0.004)	(0.004)	–	–	(0.012)	(0.011)	–	–	(0.009)	(0.008)	(0.559)	(0.443)
TSV														
$T = 500$														
Mean	0.239	0.593	0.933	0.950	0.071	0.073	–	–	–	–	0.145	0.095	10.667	6.857
Std.dev.	(2.655)	(2.025)	(0.060)	(0.039)	(0.100)	(0.097)	–	–	–	–	(0.093)	(0.063)	(5.463)	(5.747)
$T = 1000$														
Mean	-0.147	0.158	0.960	0.964	0.092	0.093	–	–	–	–	0.089	0.075	6.495	5.285
Std.dev.	(1.857)	(1.357)	(0.026)	(0.018)	(0.060)	(0.059)	–	–	–	–	(0.043)	(0.032)	(2.722)	(1.408)
$T = 2000$														
Mean	-0.013	0.405	0.974	0.976	0.075	0.062	–	–	–	–	0.062	0.058	5.466	5.069
Std.dev.	(1.298)	(0.888)	(0.009)	(0.005)	(0.043)	(0.041)	–	–	–	–	(0.017)	(0.011)	(0.825)	(0.717)
$T = 4000$														
Mean	0.010	0.035	0.977	0.978	0.073	0.074	–	–	–	–	0.056	0.054	5.248	5.069
Std.dev.	(0.860)	(0.660)	(0.005)	(0.004)	(0.029)	(0.030)	–	–	–	–	(0.010)	(0.008)	(0.554)	(0.461)
TGA-SV														
$T = 500$														
Mean	-0.790	-1.157	0.953	0.955	0.106	0.101	-0.079	-0.076	-0.015	0.071	0.089	0.052	13.511	8.975
Std.dev.	(4.411)	(5.884)	(0.069)	(0.129)	(0.163)	(0.154)	(0.069)	(0.063)	(0.187)	(0.160)	(0.063)	(0.059)	(6.657)	(10.436)
$T = 1000$														
Mean	-1.090	-1.129	0.972	0.975	0.092	0.076	-0.077	-0.078	0.028	0.100	0.054	0.040	9.651	7.079
Std.dev.	(3.023)	(2.763)	(0.012)	(0.011)	(0.100)	(0.088)	(0.046)	(0.038)	(0.159)	(0.130)	(0.031)	(0.030)	(5.866)	(6.162)
$T = 2000$														
Mean	-0.447	-0.678	0.976	0.976	0.072	0.082	-0.083	-0.077	0.037	0.081	0.050	0.049	8.556	5.828
Std.dev.	(1.298)	(1.494)	(0.007)	(0.006)	(0.069)	(0.045)	(0.029)	(0.024)	(0.136)	(0.098)	(0.022)	(0.016)	(5.182)	(2.908)
$T = 4000$														
Mean	-0.330	-0.593	0.979	0.978	0.077	0.070	-0.080	-0.083	0.071	0.081	0.048	0.054	6.254	5.548
Std.dev.	(1.298)	(1.857)	(0.004)	(0.004)	(0.048)	(0.061)	(0.019)	(0.019)	(0.094)	(0.076)	(0.017)	(0.015)	(2.600)	(1.482)

Table 3: Simulation results

Monte Carlo results of a standard Bayesian approach and DC estimators of ASV model's parameters with Skew-Student-t standardized returns. Monte Carlo averages of estimates and standard deviations are reported (in parenthesis). The first column of each parameter corresponds to the estimates obtained with **standard Bayesian approach**, while the second column corresponds to the estimates obtained with **DC**.

Parameters	μ		ϕ		δ		γ_1		γ_2		σ_η^2		ν		λ	
True	**0**		**0.98**		**0.07**		**-0.08**		**0.10**		**0.05**		**5.00**		**0.90**	
								AARSV								
$T = 500$																
Mean	-0.671	-0.036	0.951	0.970	–	–	-0.087	-0.081	–	–	0.108	0.061	9.794	6.518	0.902	1.005
Std.dev.	(1.840)	(0.997)	(0.017)	(0.062)	–	–	(0.034)	(0.046)	–	–	(0.034)	(0.087)	(5.296)	(5.069)	(0.030)	(0.034)
$T = 1000$																
Mean	-0.229	0.141	0.973	0.976	–	–	-0.084	-0.083	–	–	0.066	0.052	6.255	4.946	0.902	0.899
Std.dev.	(1.036)	(0.488)	(0.013)	(0.008)	–	–	(0.025)	(0.024)	–	–	(0.026)	(0.013)	(2.331)	(0.960)	(0.020)	(0.019)
$T = 2000$																
Mean	-0.025	0.175	0.977	0.977	–	–	-0.084	-0.082	–	–	0.056	0.056	5.536	5.326	0.899	0.901
Std.dev.	(0.436)	(0.368)	(0.007)	(0.004)	–	–	(0.019)	(0.013)	–	–	(0.014)	(0.010)	(0.864)	(0.929)	(0.016)	(0.015)
$T = 4000$																
Mean	0.028	0.056	0.978	0.979	–	–	-0.081	-0.079	–	–	0.054	0.051	5.253	5.024	0.900	0.903
Std.dev.	(0.246)	(0.214)	(0.005)	(0.004)	–	–	(0.012)	(0.012)	–	–	(0.008)	(0.008)	(0.555)	(0.424)	(0.009)	(0.008)
								TSV								
$T = 500$																
Mean	0.486	0.955	0.918	0.937	0.075	0.074	–	–	–	–	0.164	0.126	9.896	6.095	0.903	0.907
Std.dev.	(2.869)	(1.567)	(0.117)	(0.067)	(0.099)	(0.091)	–	–	–	–	(0.120)	(0.115)	(4.810)	(3.516)	(0.028)	(0.021)
$T = 1000$																
Mean	0.075	0.202	0.963	0.966	0.083	0.091	–	–	–	–	0.085	0.063	6.270	4.927	0.897	0.897
Std.dev.	(2.085)	(1.427)	(0.022)	(0.019)	(0.066)	(0.058)	–	–	–	–	(0.037)	(0.023)	(2.383)	(0.992)	(0.021)	(0.021)
$T = 2000$																
Mean	0.063	-0.255	0.973	0.975	0.076	0.087	–	–	–	–	0.065	0.063	5.390	4.892	0.898	0.894
Std.dev.	(1.327)	(1.003)	(0.009)	(0.009)	(0.042)	(0.034)	–	–	–	–	(0.018)	(0.021)	(0.900)	(0.614)	(0.013)	(0.014)
$T = 4000$																
Mean	0.102	0.290	0.977	0.978	0.072	0.066	–	–	–	–	0.057	0.053	5.160	4.985	0.899	0.899
Std.dev.	(0.908)	(0.850)	(0.005)	(0.004)	(0.031)	(0.037)	–	–	–	–	(0.010)	(0.007)	(0.504)	(0.430)	(0.010)	(0.011)
								TGA-SV								
$T = 500$																
Mean	-0.190	-0.052	0.954	0.963	0.090	0.114	-0.091	-0.069	0.031	0.104	0.088	0.039	11.576	7.989	0.899	0.891
Std.dev.	(3.830)	(3.448)	(0.030)	(0.022)	(0.174)	(0.169)	(0.081)	(0.073)	(0.187)	(0.175)	(0.066)	(0.038)	(6.611)	(9.536)	(0.029)	(0.016)
$T = 1000$																
Mean	-1.066	-1.053	0.970	0.974	0.098	0.098	-0.082	-0.088	0.051	0.069	0.057	0.041	9.068	6.019	0.899	0.903
Std.dev.	(3.251)	(2.936)	(0.015)	(0.010)	(0.098)	(0.085)	(0.047)	(0.049)	(0.155)	(0.178)	(0.037)	(0.022)	(5.731)	(4.785)	(0.018)	(0.016)
$T = 2000$																
Mean	-0.471	-0.421	0.977	0.977	0.076	0.095	-0.082	-0.063	0.080	0.151	0.048	0.036	7.105	4.733	0.900	0.900
Std.dev.	(2.022)	(1.530)	(0.007)	(0.004)	(0.065)	(0.063)	(0.036)	(0.023)	(0.128)	(0.059)	(0.023)	(0.021)	(4.506)	(0.791)	(0.013)	(0.014)
$T = 4000$																
Mean	-0.338	-0.363	0.978	0.979	0.079	0.083	-0.077	-0.074	0.096	0.116	0.047	0.044	5.696	4.793	0.900	0.904
Std.dev.	(1.420)	(1.362)	(0.004)	(0.005)	(0.048)	(0.044)	(0.024)	(0.020)	(0.086)	(0.047)	(0.020)	(0.018)	(2.383)	(0.566)	(0.010)	(0.008)

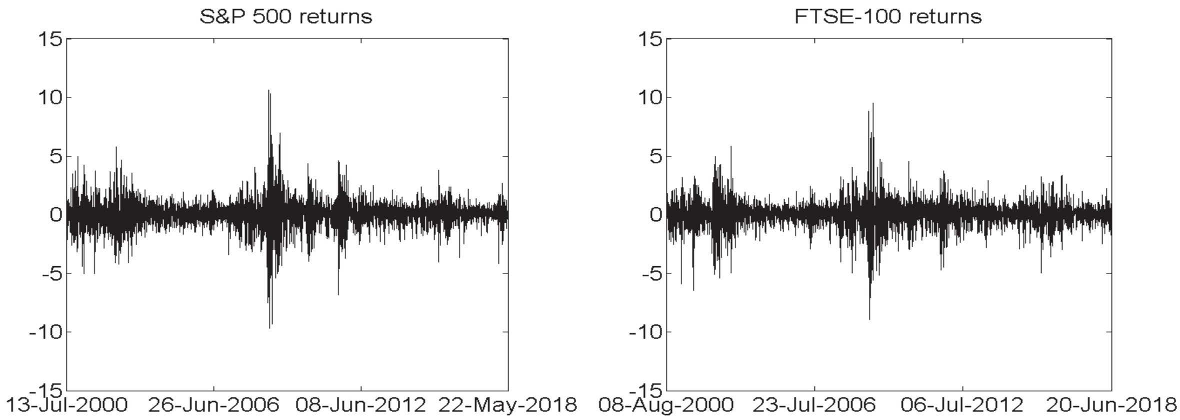


Figure 1: Daily returns in percentage

Table 4: Descriptive statistics of returns

Data	Mean	Std. Dev.	Skewness	Kurtosis
S&P 500	0.0135	1.1912	-0.2013	11.5699
FTSE-100	0.0039	1.1517	-0.1691	9.7571

Table 5: Estimates obtained using DC and a standard Bayesian approach

		Bayesian approach			DC		
Series	Param	AARSV	TSV	TGA-SV	AARSV	TSV	TGA-SV
S&P 500							
	$\mu(1-\phi)$	-0.010	-0.143	-0.030	-0.011	-0.141	-0.031
		(0.002)	(0.012)	(0.017)	(0.003)	(0.013)	(0.016)
	ϕ	0.976	0.979	0.976	0.977	0.980	0.978
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	δ		0.285	0.010		0.280	0.014
			(0.024)	(0.032)		(0.026)	(0.030)
	γ_1	-0.173		-0.179	-0.169		-0.174
		(0.012)		(0.018)	(0.011)		(0.019)
	γ_2			-0.104			-0.106
				(0.031)			(0.032)
	σ_η^2	0.018	0.025	0.019	0.016	0.024	0.017
		(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
	υ	15.338	19.780	29.679	12.765	15.830	24.521
		(4.279)	(7.626)	(12.540)	(2.429)	(4.323)	(11.764)
	λ	0.982	0.997	0.983	0.983	0.997	0.983
		(0.009)	(0.010)	(0.010)	(0.009)	(0.010)	(0.010)
	DIC	89.929	89.929	89.930			
FTSE-100							
	$\mu(1-\phi)$	-0.006	-0.130	-0.019	-0.006	-0.128	-0.019
		(0.002)	(0.011)	(0.014)	(0.002)	(0.010)	(0.018)
	ϕ	0.983	0.981	0.982	0.983	0.981	0.983
		(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	δ		0.257	0.035		0.252	0.028
			(0.021)	(0.030)		(0.020)	(0.035)
	γ_1	-0.143		-0.129	-0.145		-0.134
		(0.009)		(0.016)	(0.010)		(0.015)
	γ_2			0.031			0.002
				(0.021)			(0.032)
	σ_η^2	0.010	0.017	0.009	0.010	0.018	0.010
		(0.002)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)
	υ	37.075	41.067	31.412	42.751	60.400	42.478
		(13.712)	(16.118)	(10.759)	(27.071)	(49.903)	(32.139)
	λ	0.976	0.990	0.976	0.975	0.990	0.976
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
	DIC	89.929	89.929	89.929			

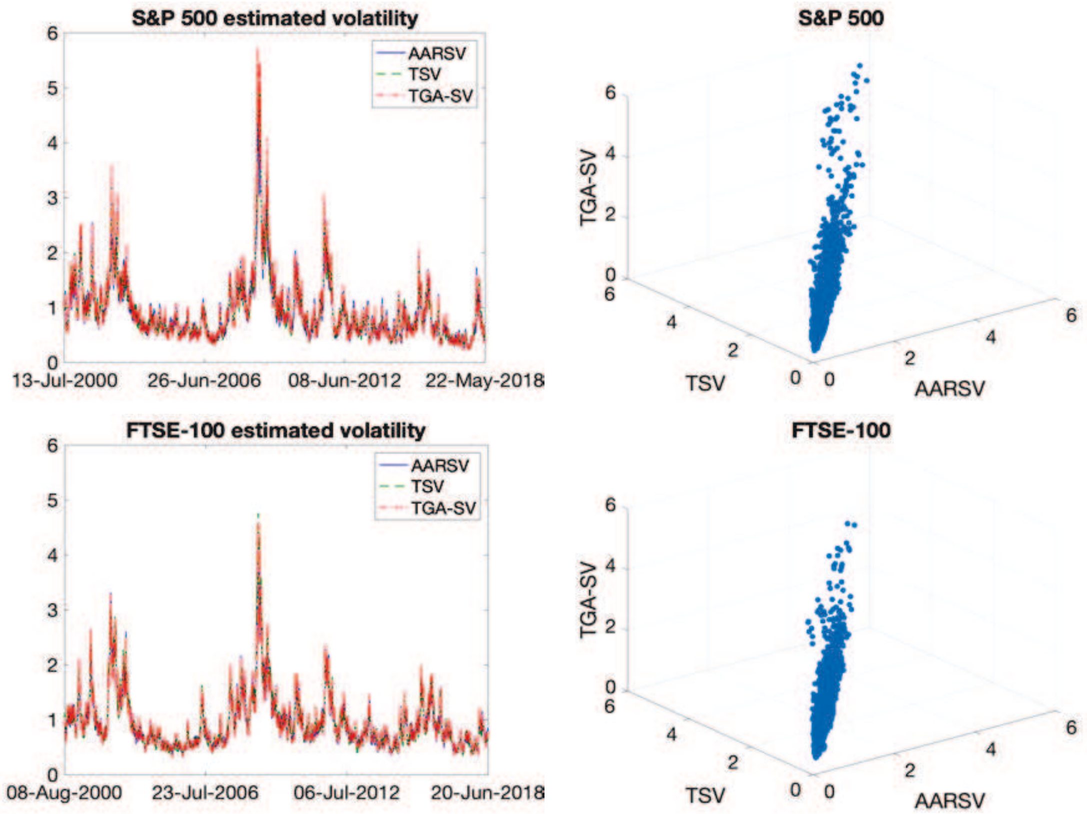
Note: Standard errors are in parentheses. The values of the DIC are divided by 10^6.

Table 6: Model confidence set results

The table reports the rankings of volatility forecasters with different loss functions. – means that the model does not belong to the SSM. The statistical tests are done at 95% and 80% confidence levels. We use 5000 bootstrap samples.

Series	Models	Loss functions							
		Bayesian approach				DC			
		95%		**80%**		**95%**		**80%**	
		SE	QLIKE	SE	QLIKE	SE	QLIKE	SE	QLIKE
S&P 500	AARSV	2	3	2	3	3	3	3	3
	TSV	3	2	3	2	2	2	2	2
	TGA-SV	1	1	1	1	1	1	1	1
	p-value	0.888	0.652	0.880	0.652	0.728	0.356	0.728	0.348
FTSE-100	AARSV	3	3	3	3	3	3	3	3
	TSV	1	1	1	1	1	1	1	1
	TGA-SV	2	2	2	2	2	2	2	2
	p-value	0.568	0.536	0.568	0.547	0.425	0.462	0.439	0.458

Bayesian



DC

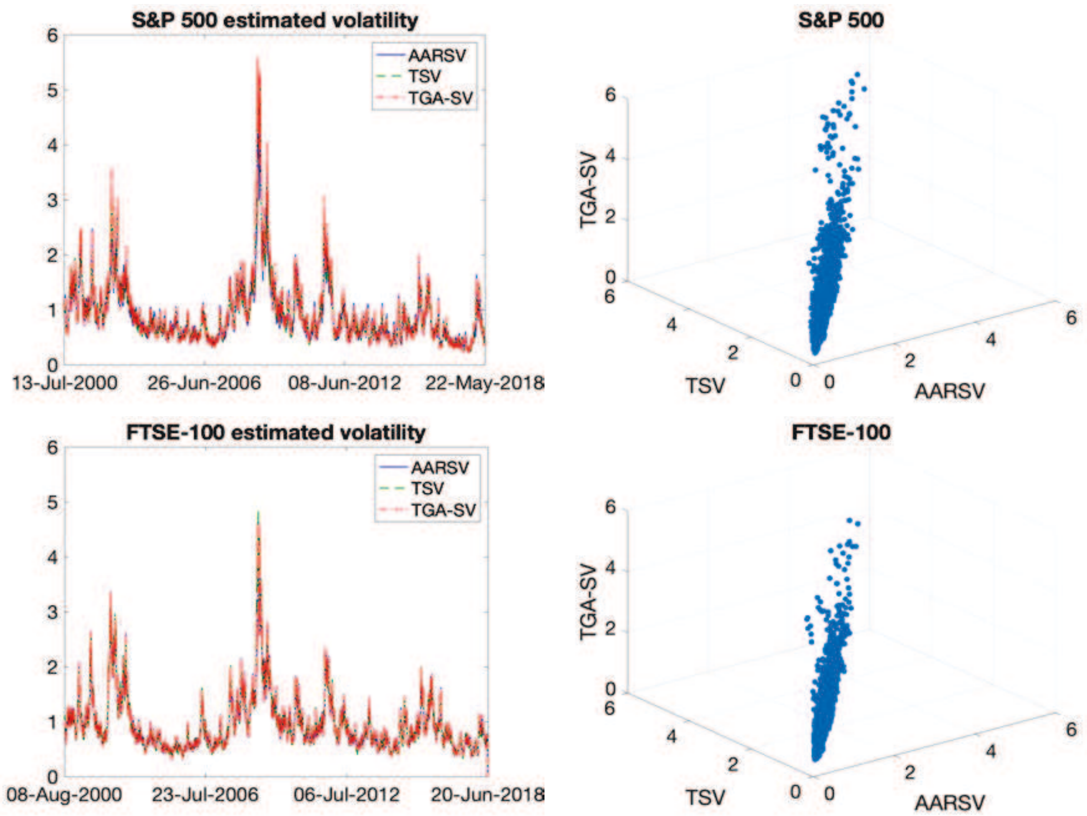


Figure 2: Estimated volatilities and volatility scatter plots

Appendix A

```
data{
for (indexclon in 1:Ncl){
for (i in 1:t_max) {
zero[i,indexclon] <- 0
}
}
}

model{

# Prior distributions
mu ~ dnorm(0, 0.001)
phistar ~ dbeta(1, 1)
tau_eta ~ dgamma(0.001, 0.001)
gamma1 ~ dnorm(0, 0.001)
nu ~ dgamma(2,0.1)
skew ~ dgamma(0.01,0.01)

phi <- 2*phistar - 1
sigeta2 <- 1/tau_eta

# Cloning process
for (indexclon in 1:Ncl){

# Initial values
htmean[1,indexclon] <- mu + phi*(htmean0-mu) + gamma1*epsilon0
ht[1,indexclon] ~ dnorm(htmean[1,indexclon], tau_eta)
```

```
ysigma2[1,indexclon] <- 1/exp(ht[1,indexclon])

# zero trick to define a skew-t
z[1,indexclon] <- ifelse(y[1,indexclon] >= 0,
y[1,indexclon]/skew, y[1,indexclon]*skew)
loglik[1,indexclon] <- logdensity.t(z[1,indexclon], 0, ysigma2[1,indexclon], nu) +
log(2) - log(skew + 1/skew)
lambda[1,indexclon] <- 10000 - loglik[1,indexclon]
zero[1,indexclon] ~ dpois(lambda[1,indexclon])

# Full iterations
for (t in 2:t_max) {
htmean[t,indexclon] <- mu + phi*(ht[t-1,indexclon]-mu) +
(gamma1*y[t-1,indexclon]/exp(ht[t-1,indexclon]/2))
ht[t,indexclon] ~ dnorm(htmean[t,indexclon], tau_eta)
ysigma2[t,indexclon] <- 1/exp(ht[t,indexclon])

# zero trick to create a skew-t
z[t,indexclon] <- ifelse(y[t,indexclon] >= 0, y[t,indexclon]/skew, y[t,indexclon]*skew)
loglik[t,indexclon] <- logdensity.t(z[t,indexclon], 0, ysigma2[t,indexclon], nu) +
log(2) - log(skew + 1/skew)
lambda[t,indexclon] <- 10000 - loglik[t,indexclon]
zero[t,indexclon] ~ dpois(lambda[t,indexclon])
}

}
for (t in 1:t_max) {
httot[t] <- mean(ht[t,])
}
```

```
}
```

References

Abanto-Valle, C. A., V. H. Lachos, and D. K. Dey (2015). Bayesian estimation of a skew-Student-t stochastic volatility model. *Methodology and Computing in Applied Probability 17*(3), 721–738.

Andersen, T. G., T. Bollerslev, F. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*(453), 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactor volatility models. *Journal Empirical Finance 15*, 332–341.

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.

Barra, I., L. Hoogerheide, S. J. Koopman, and A. Lucas (2017). Joint Bayesian analysis of parameters and states in nonlinear non-Gaussian state space models. *Journal of Applied Econometrics 32*(5), 1003–1026.

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

Catania, L. and M. Bernardi (2015). *MCS: Model Confidence Set Procedure*. R package version 0.1.1.

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

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.

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

Demos, A. (2002). Moments and dynamic structure of a time-varying parameter stochastic volatility in mean model. *Econometrics Journal 5*(2), 345–357.

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.

Fernández, C. and M. F. J. Steel (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association 93*(441), 359–371.

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.

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

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

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

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.

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.

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.

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

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

Mao, X., E. Ruiz, H. Veiga, and V. Czellar (2015). Asymmetric stochastic volatility models: Properties and estimation. *Available at SSRN 2544473*.

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

Meyer, R. and J. Yu (2000). BUGS for a Bayesian analysis of stochastic volatility models. *The Econometrics Journal 3*(2), 198–215.

Nakajima, J. and Y. Omori (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution. *Computational Statistics & Data Analysis 56*(11), 3690–3704.

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

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling.

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.

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

Solymos, P. (2010). dclone: Data cloning in R. *The R Journal 2*(2), 29–37.

Solymos, P. (2016). *dcmle: Hierarchical Models Made Easy with Data Cloning*. R package version 0.3-1.

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

Taylor, S. J. (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical Finance 4*, 183–204.

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.

Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics 127*(2), 165–178.

Yu, J. (2012). A semiparametric stochastic volatility model. *Journal of Econometrics 167*(2), 473–482.