The DIC as a Model Comparison Criterion for Stochastic Volatility Models *

Andreas Berg,[†] Renate Meyer,[†] and Jun Yu[‡]

June 27, 2001

Abstract

Bayesian methods have proven very efficient in estimating parameters of stochastic volatility (SV) models for analysing financial time series. Recent work extend the basic stochastic volatility model to include heavy-tailed error distributions, covariates, leverage effects, and jump components. Hierarchical Bayesian methods (usually implemented via state-of-the-art Markov chain Monte Carlo methods for posterior computation) allow fitting of such complex models. However, a formal model comparison is hindered by the fact that the unknowns in the model outnumber the observations. The main objective of this paper is to demonstrate that a formal model selection is possible using the deviance information criterion (DIC), a recently developed generalization of the Akaike information criterion (AIC) designed for complex hierarchical models with possibly improper prior distributions. We illustrate the performance of the DIC in discriminating between various different SV models using a data set of daily returns of the Standard & Poors 100 index. It turns out that the models which best fit the dataset are the one with a leverage effect and the one with the implied volatility as a covariate in the variance equation.

Keywords: *Stochastic Volatility, Model Selection, Bayesian Deviance, Gibbs Sampler, BUGS, Leverage Effect, Heavy-tailed Distributions, Jumps, Implied Volatility.*

---

[*] Andreas Berg gratefully acknowledges the support of his research by the DAAD, personal number D/99/19367 within a "DAAD Doktorandenstipendium im Rahmen des gemeinsamen Hochschulsonderprogramms III von Bund und Laendern". The work of Renate Meyer was supported by the Royal Society of New Zealand Marsden Fund and the University of Auckland Research Committee. Jun Yu would like to acknowledge financial support from the University of Auckland Research Committee and the University of Auckland Business School.

[†] Department of Statistics, University of Auckland

[‡] Department of Economics, University of Auckland

1 Introduction

The progress in Bayesian posterior computation due to MCMC methods has made it possible to fit increasingly complex statistical models. It is then necessary to determine the best-fitting model out of a potentially huge class of candidates. Thus, it has become more and more important to develop efficient model selection criteria. A recently developed proposal by Spiegelhalter et al. (1998) is the *Deviance Information Criterion* (DIC), a Bayesian version or generalization of the well-known AIC (Akaike Information Criterion) and also related to the BIC (Bayesian Information Criterion). The DIC is easy to calculate and applicable to a wide range of statistical models. It is based on the posterior distribution of the log-likelihood or the deviance, following the original suggestion of Dempster (1974) for model choice in the Bayesian framework. It overcomes the problem of necessity of identifying the number of parameters in the models which is required for calculation of BIC and AIC. This is accomplished through the principle of *effective number of parameters*. The DIC can therefore be applied to complex hierarchical random effects models.

This model comparison criterion has already been successfully applied to complex models in the field of medical statistics (Zhu et al. (2000)). In this paper, we demonstrate its usefulness in the model selection process for financial time series. The aim of this paper is therefore to introduce the DIC to the financial modelling community and show how to use it for the family of stochastic volatility (SV) models.

Indeed, many model checking criteria have been proposed and discussed before the development of DIC, but there was no universal criterion. While Bayes factors were viewed for many years as the only correct way to carry out Bayesian model comparison, they have come under increasing criticism of late. The most serious drawback is that they are not well-defined when using improper priors (which is typically the case in practice when employing noninformative priors). This leads to modifications, such as the *partial Bayes factor* (O'Hagan, 1991), the *intrinsic Bayes factor* (Berger and Pericchi, 1996), and the *fractional Bayes factor* (O'Hagan, 1994). These suffer from more or less arbitrary choices of training samples, weights for averaging training samples, and fractions, respectively. Furthermore, the calculation of Bayes factors can be computationally difficult due to an implicit high-dimensional integration problem. As such, the computation of Bayes factors is not a particularly user-friendly tool for the practicing statistician.

Shortcuts to the calculation of Bayes factors that avoid dependence of priors are the large sample approximations to $-2\log(\text{Bayes-factor})$, including the familiar BIC (*Bayesian Information Criterion*) or *Schwarz Criterion* (Schwarz, 1978), and the related penalized likelihood ratio model choice criterion, AIC. Both criteria demand the specification of the number of parameters. In hierarchical Bayesian models, however, the number of unknowns often outnumbers the number of observations. This makes both BIC and AIC not applicable to complex models, unfortunately.

Chib, Nardari and Shephard (2000) and Kim, Shephard and Chib (1998) use the marginal likelihood approach of Chib (1995) to calculate Bayes factors from the likelihood ordinates at arbitrary points (but recommended at the posterior means) using the technique of particle filtering, which requires a normal distribution for the likelihood ordinate. They convert the multivariate stochastic volatility model into a collection of Gaussian state space models, showing that the log-volatilities of the SV-process follow approximately a seven component mixture of normal distributions. Although they also show how to convert a flat-tailed (Student-t) distribution into this special form this approach is nevertheless limited and computationally intensive. The above-mentioned technique of Chib (1995) is also recommended by Han and Carlin (2000) for a wide range of hierarchical problems, but limitations of this method are outlined for models of increasing complexity.

The outline of the paper is as follows: Section 2 starts with a short description of stochastic volatility models, followed in Section 3 by the introduction of the DIC. In Section 4, the DIC is applied to compare various stochastic volatility models, fit to a dataset previously used in the literature. Section 5 concludes.

2 The Stochastic Volatility Model

In the literature of volatility of financial time series, the stochastic volatility model (Tauchen and Pitts (1983) and Taylor (1982)) has received much attention in recent years, mainly because it is able to take into account the variability of the volatility as an element for describing financial data such as interest rates, exchange rates, and stock market prices. It has become established as a powerful alternative to the ARCH and GARCH models (Engle (1982), Bollerslev (1986)).

A basic generalized stochastic volatility model (Chib et al. (1998)) consists of an observation equation

$$y_t = c \cdot x_{1t} + x_{2t}^{\gamma} \cdot exp(\frac{h_t}{2}) \cdot u_t, \qquad t = 1, \ldots, n,$$

and describes the distribution of the data given unknown states, the daily volatilities, and a state equation

$$h_t = \mu + b \cdot x_{3t} + \phi \cdot (h_{t-1} - \mu) + \sigma \cdot \eta_t, \qquad t = 1, \ldots, n,$$

which models the variation of the daily volatilities as a Markov process.

Here, y_t is the response variable, the x_{it}, $i = 1, 2, 3$, are covariates, h_t is the log-volatility process with η_t as a white noise. For the time being, there is no particuliar distribution specified for the error term u_t. We collect the model parameters in a vector $\theta = (c, \gamma, b, \phi, \mu, \sigma)$. The parameter γ plays an important role in analyzing interest rate data (for details refer for example to Chan et al. (1992) and Brenner et al. (1996)). In other applications, for example stock market data, it is common to set this parameter equal to 0 (see Section 4 below).

The SV model is a typical example of a hierarchical model, in which the number of unknowns, i.e., the parameters (θ) and the unknown states (h_t) outnumbers the number of observations.

Classical parameter estimation for this model is extremely difficult, because of the non-analytic form of the likelihood function. Whilst there are a number of valuable and widely accepted techniques, most of them require approximations.[1] A Bayesian approach provides a very promising alternative.

SV models provide an ideal context within which the benefit of the application of DIC as a model selection tool can be demonstrated. We will show that DIC provides an efficient and straightforward approach in identifying the most appropriate models.

3 The Deviance Information Criteria (DIC)

Model assessment from a frequentist point of view is based on the *deviance*, the difference in the log-likelihoods between the fitted and the saturated model (the model with as many parameters as observations, yielding a perfect fit to the data). By analogy, Dempster (1974) suggested to examine the posterior distribution of the classical deviance defined by

$$D(\theta) = -2 \log f(y \mid \theta) + 2 \log f(y)$$

for observations y and parameter vector θ. The marginal log-likelihood $\log f(y)$ serves as a standardising term. Dempster (1974) suggested plots and potential summaries such as the posterior mean of $D(\theta)$. Spiegelhalter et al. (1998) developed the DIC as model choice criterion. Based on the posterior distribution of $D(\theta)$, it consists of two

---

[1] See Andersen et al. (1999) for a review of various estimation techniques for the SV model.

components: a term that measures goodness-of-fit and a penalty term for increasing model complexity.

The measure of fit consists of the posterior expectation of the deviance

$$E_{\theta|y}[D] = \bar{D}.$$

The second part measures the complexity of the model by the *effective number of parameters*, p_D, defined as the difference between posterior mean of the deviance and the deviance evaluated at the posterior mean of the parameters:

$$p_D = E_{\theta|y}[D] - D(E_{\theta|y}[\theta]) = \bar{D} - D(\bar{\theta}).$$

The DIC is then defined as the sum of both components

$$\text{DIC} = \bar{D} + p_D = 2\bar{D} - D(\bar{\theta}) = D(\bar{\theta}) + 2p_D.$$

As shown in Spiegelhalter et al. (1998), the DIC is a generalisation of the well-known AIC, defined by AIC $= D(\hat{\theta}) + 2p$ where $\hat{\theta}$ is the maximum-likelihood estimate of the parameter vector and p the number of parameters. Thus for non-hierarchical models, $p \approx p_D$, $\hat{\theta} \approx \bar{\theta}$, and DIC $\approx$ AIC. For an approximate decision-theoretic justification of DIC and asymptotic properties of p_D and $\bar{D}$ the interested reader is referred to Spiegelhalter et al. (1998).

So far, no efficient method for calculating reasonably accurate standard errors of the DIC has been developed. Zhu et al. (2000) explore this problem, but their approach using the multivariate delta method yields poor results. Their final recommendation is the "brute force" approach, which is simply replicating the calculation of the DIC N times and estimating VAR(DIC) by its sample variance

$$\widehat{VAR}(\text{DIC}) = \frac{1}{N-1}\sum_{k=1}^{N}(\text{DIC}_k - \overline{\text{DIC}})^2.$$

In spite of a painfully time-consuming approach, it at least gives an indication of the inherent variability of the DIC.

4 An Illustrative Example

In the following, we used the DIC as a model choice criterion among various different stochastic volatility models fitted to a data set from the Standard & Poor's 100 (S&P100) stock index.

4.1 The Data

The dataset consists of 1512 mean-corrected daily returns of the Standard & Poors 100 index, covering the period of time between January 1993 and December 1998. The S&P100 index returns have been used often in the literature. For instance, Blair et al. (2001) estimate the GJR-GARCH model proposed by Glosten, et al. (1993) based on the S&P100 index returns for four different sample periods from March 1984 to December 1998, one of which is identical to the period we use in this paper.

We also use data from the Chicago Board Options Exchange Market Volatility Index (VIX) for the same period of time as a covariate, measuring the so called implied volatility. For a detailed explanation of the Chicago Board Options Exchange Market Volatility Index, the reader is referred to Hol and Koopman (2000) and Fleming et al. (1995).

4.2 The Models

We fit six different stochastic volatility models to the S&P100 index. The first five models have been considered in Meyer and Yu (2000). Furthermore, we consider a model including implied volatility. The models are specified following the notation of Meyer and Yu (2000).

MODEL 1:

$$y_t \mid \theta_t = \exp(\frac{1}{2} \cdot \theta_t) \cdot u_t, \qquad u_t \overset{iid}{\sim} N(0,1), t = 1, ..., n$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2 \;=\; \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t, \qquad v_t \overset{iid}{\sim} N(0, \tau^2), t = 1, ..., n,$$

with $\theta_0 \sim N(\mu, \tau^2)$. We allow for a positive scaling factor by setting $\beta = exp(\frac{\mu}{2})$ which is often interpreted as the instantaneous volatility. In our case y_t are the observations from the mean-corrected return series with volatility process θ_t. The parameters μ and ϕ are the intercept and the persistence in the volatility process. The θ_t's are centered around μ to avoid high posterior correlations between the parameters and thus speed up convergence of the Gibbs sampler. MODEL 1 is a special case of the generalized SV model in Section 2 with b, c, and γ equal to zero.

<u>MODEL 2</u>: An additional non-zero mean α is added in the observation equation:

$$\begin{aligned} y_t \mid \theta_t, \alpha &= \alpha + \exp(\frac{1}{2} \cdot \theta_t) \cdot u_t, \qquad u_t \overset{iid}{\sim} N(0, 1), t = 1, ..., n. \\ \theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2 &= \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t, \qquad v_t \overset{iid}{\sim} N(0, \tau^2), t = 1, ..., n, \end{aligned}$$

<u>MODEL 3</u>: An AR(2)-process for the state equation:

$$\begin{aligned} y_t \mid \theta_t &= \exp(\frac{1}{2} \cdot \theta_t) \cdot u_t, \qquad u_t \overset{iid}{\sim} N(0, 1), t = 1, ..., n \\ \theta_t \mid \theta_{t-1}, \mu, \phi, \psi, \tau^2 &= \mu + \phi \cdot (\theta_{t-1} - \mu) + \psi \cdot (\theta_{t-2} - \mu) + v_t, \qquad v_t \overset{iid}{\sim} N(0, \tau^2), t = 1, ..., n, \end{aligned}$$

<u>MODEL 4</u>: A central Student-t distribution with degrees of freedom, ν, for the observation error term:

$$\begin{aligned} y_t \mid \theta_t &= \exp(\frac{1}{2} \cdot \theta_t) \cdot u_t, \qquad u_t \overset{iid}{\sim} t_\nu, t = 1, ..., n, \\ \theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2 &= \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t, \qquad v_t \overset{iid}{\sim} N(0, \tau^2), t = 1, ..., n, \end{aligned}$$

<u>MODEL 5</u>: The basic SV MODEL 1 including a leverage or asymmetric effect implemented by allowing for correlation beween u_t and v_{t+1}, i.e.

$$\begin{pmatrix} u_t \\ v_{t+1} \end{pmatrix} \overset{iid}{\sim} N\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right\}, \; \theta_0 \sim N(\mu, \tau^2).$$

This effect is often observed in financial time series, e.g. in time series of exchange rates and, even stronger, in stock market data. It reveals the market behavior, first discovered by Black (1976), which "occurs when an unexpected drop in price (bad news) increases predictable volatility more than an unexpected increase in price (good news) of similar magnitude" (Engle and Ng (1993, page 1752)).

MODEL 6: The SVX model (Hol and Koopman, 2000) includes implied volatility as expressed by an additional covariate x_t:

$$\begin{aligned} y_t \mid \theta_t &= \exp\left(\frac{1}{2} \cdot \theta_t\right) \cdot u_t, \quad u_t \overset{iid}{\sim} N(0,1), t = 1, ..., n, \\ \theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2, \lambda &= \mu + \phi \cdot (\theta_{t-1} - \mu) + \lambda \cdot (x_t - \bar{x}) + v_t, \\ & \quad v_t \overset{iid}{\sim} N(0, \tau^2), t = 1, ..., n. \end{aligned}$$

The implied volatility is used in this model as an alternative source for predicting volatility and is based on calculations of option price models. The specification of the variance equation is motivated from the empirical result that implied volatilities contain useful information in forecasting future volatilities (see for example Blair et al. (2001)). Note that we demean the observations in vector x_t for convergence purposes.

4.3 Prior Distributions

We basically use the same prior distributions suggested by Kim et al. (1998).

The prior for parameter μ is a normal distribution with mean 0 and variance 10. A priori, λ (MODEL 6) is assumed to be uniformly distributed in the interval $[0, 2]$. A conjugate inverse gamma distribution with parameters 2.5 and 0.025 was chosen for τ^2. Defining $\phi = 2\phi^* - 1$, we specified a Beta-distribution with parameters 20 and 1.5 for ϕ^*. For α (MODEL 2) a flat normal distribution with mean parameter $\mu_\alpha = 0$ and variance $\sigma^2_\alpha = 10$ was specified. We used the same prior for ψ as for the AR(1) parameter ϕ. As the value for ϕ diminishes dramatically for MODEL 3 and 6, we depart from the beta-distribution for ϕ^* and choose a more general uniform (0,1)- distribution for ϕ^* and ψ^* respectively.

In Model 4, the additional parameter k, representing the unknown degree of freedom of the assumed t-distribution, has a prior χ^2-distribution with 8 degrees of freedom. The correlation parameter ρ (Model 5) is assumed uniformly distributed with support between -1 and 1.

4.4 Implementation in WinBUGS

The first part of the DIC, $\bar{D}$, is easily calculated using the MCMC output $\theta^{(i)}, i = 1, \ldots, N$. We simply calculate $D(\theta^{(i)})$ for $i = 1, \ldots, N$ and estimate $\bar{D}$ by the sample mean $\frac{1}{N}\sum_{i=1}^{N} D(\theta^{(i)})$. In practice, using BUGS, this is accomplished by adding the additional variable $D(\theta)$. For the second part, the effective number of parameters p_D, we only need to evaluate $D(\theta)$ at the sample posterior mean $\bar{\theta} = \frac{1}{N}\sum_{i=1}^{N} \theta^{(i)}$.

For a posterior analysis of our Models 1-6, we used the WinBUGS package which is available free of charge via internet from

`http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml.`

WinBUGS is an easy to learn and easy to use Bayesian software package that implements the Gibbs sampler for generating samples from a Markov chain whose equilibrium distribution is the posterior distribution. As demonstrated by Meyer and Yu (2000), it can be applied to fit stochastic volatility models. A DIC module which automatically calculates values for the DIC and related parameters is implemented in the latest WinBUGS version. WinBUGS is the BUGS-version operating under WINDOWS and offers several useful convergence checking criteria available in an attached CODA (Convergence Diagnosis and Output Analysis Software for Gibbs sampling output, Best et al. (1995)) module running for example under SPLUS. It is necessary to check whether convergence has been achieved because it is crucial that the sample is taken from the stationary distribution. The CODA package consists of a selection of model checking criteria, one of which is the Heidelberger and Welch test (Heidelberger and Welch (1983)).

4.5 Results

In Tables 1-6 we report means and variances of both prior and posterior distributions for each of the six models respectively. The results in the Tables 1-3 and 5-6 are based on 12,500 iterations. After a burn-in period of 50,000 iterations and a follow-up period of 250,000, we stored every 20th iteration. Due to higher posterior correlations amongst the parameters and thus slower convergence of the Gibbs sampler in the remaining model, we chose a burn-in period of 100,000 iterations, a follow-up period of 900,000, and stored every 40th iteration. We ran each chain 6 times to obtain a brute-force estimate of the variability of DIC. All calculations were performed on a Pentium-III PC, 550 MHz, running the WinBUGS 131 version updated with the DIC tool.

The estimated means and standard deviations for the parameters appear quite reasonable and comparable with previous estimates in the literature. For instance, in the basic SV model, the volatility process is estimated to be highly persistent. The estimate of the second autoregressive coefficient in MODEL 3 is a large positive number and consistent with results obtained by Meyer and Yu (2000) in an exchange rate series. More interestingly, it is bigger than the estimate of first autoregressive coefficient. In MODEL 4 the posterior mean of ν is 8.85 and similar to the values of 7.7 and 8.9 for the S&P500 index in Sandmann and Koopman (1998) and Chib et al. (2000) respectively. In MODEL 5 the posterior mean of ρ is -0.418 with its upper interval less than zero. It suggests that the leverage effect is important for the S&P100 index. The posterior mean of λ in MODEL 5 indicates that the implied volatility contains important information about the volatility process. Interestingly, allowing for the implied volatility as a covarite induces a negative posterior mean of the autorgressive coefficient in the model. This finding is similar to what obtained in Hol and Koopman (2000) based on a S&P100 index for a different period. Table 7 shows the smallest and largest values for the DIC, the number of effective parameters p_D and the goodness-of-fit $\bar{D}$, respectively, obtained for 6 runs for each of the MODELS 1–6.

By choosing very unreasonable initial values, we did not achieve convergence in some cases for MODEL 3 and 6. However, the values for the DIC in the non-convergent cases in MODEL 6 for example lie in between the range of the DIC values for which we have obtained convergence. For the non-convergent cases in MODEL 3 all these DIC values lie at the upper edge of the range or marginally outside the interval given in Table 3.

The most adequate models to describe the dataset according to the DIC are the implied volatility MODEL 6, and the model including the leverage effect, MODEL 5.

An additional intercept parameter does not seem to have much influence on the model performance, the values for the DIC are only marginally larger then those for MODEL 1. Quite surprising is the performance of the models with Student-t distributed errors. Although adding an additional parameter, the degrees of freedom, the effective number of parameters reduces from about 80 to 50 but their DIC values are the largest among all models that we fitted.

The good performance of the leverage and implied volatility model are even more surprising taking into account their high values for the effective number of parameters, which leads to the result of an astonishing small value for the posterior expectation of the deviance. The effective number of parameters raises to values around 115-125 in the AR(2)-process and implied volatility model. Even almost three times higher than the effective numbers of parameters in these models are those in the leverage effect model. Therefore, under all the examined models, we obtain overwhelming results in terms of the goodness-of-fit $\overline{D}$ for the leverage effect model. With a value of around 2990 it outperforms clearly the second best model with a value of around 3190.

5 Conclusion

In this paper we have explored the practical performance of the DIC as model selection criterion for comparing various stochastic volatility models. The DIC is a Bayesian version of the classical deviance for model assessment. It is particularly suited to compare Bayesian models whose posterior distributions have been obtained using MCMC simulation. Similar to AIC and BIC, the DIC comprises two parts, a goodness-of-fit measure, the posterior distribution of the deviance, and a penalty term, the effective number of parameters, measuring complexity. Using this concept of *effective number of parameters*, the DIC can be used in complex hierarchical models where the actual number of paramaters often exceeds the number of observations. This is in contrast to AIC and BIC, where the number of parameters needs to be specified. The DIC has been implemented as a tool in the BUGS software package.

By comparing 6 different stochastic volatility models for the S&P100 index, comprising 1512 observations from 1993 to 1998, the SV model including a leverage effect turned out to be the most adequate as indicated by the smallest DIC values. The SV model with implied volatility also performs reasonably well. The Monte Carlo error of the DIC was fairly low for all the models, thus indicating a stable performance for model comparison purposes.

6 References

Andersen, T.G. and Benzoni, L. and Lund, J. (2000): Estimating jump-diffusions for equity returns, working paper, Kellogg School of Management, Northwestern University.

Andersen, T.G. and Chung, H.-J. and Sørensen B.E. (1999): Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics* **91**, 61–87.

Berger, J.O. and Pericchi, L.R. (1996): The intrinsic Bayes factor for model selection and prediction, *The Journal of the American Statistical Association*, **91**, 109–122.

Best, N. and Cowles, M.K. and Vines, K. (1995): CODA Convergence diagnosis and output analysis software for Gibbs sampling output version 0.30, MRC Biostatistics Unit, Cambridge, UK.

Blair, B.J. and Poon, S.-H. and Taylor, S.J. (2000): Modelling S&P 100 volatility: The information content of stock returns, Forthcoming, *Journal of Banking and Finance.*

Blair, B.J. and Poon, S.-H. and Taylor, S.J. (2001): Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high frequency returns, Forthcoming, *Journal of Econometrics.*

Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.

Brenner, R.J. and Harjes, R.H. and Kroner, K.F. (1996): Another look at models of the short-term interest rate, *Journal of Financial and Quantitative Analysis*, **31**, 85–107.

Chan, K.C. and Karolyi, G.A. and Longstaff, F.A. and Sanders A.B. (1992): An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance*, **47**, 1209–1227.

Chib, S. (1995): Marginal likelihood from the Gibbs output, *The Journal of the American Statistical Association*, **90**, 1313-1321.

Chib, S. and Nardari, F. and Shephard, N. (1999): Analysis of high dimensional multivariate Stochastic Volatility models, unpublished paper: John M. Olin School of Business, Washington University, St. Louis.

Chib, S. and Nardari, F. and Shephard, N. (2000): Markov Chain Monte Carlo methods for Generalized Stochastic Volatility Models, unpublished paper: Nuffield College, Oxford.

Dempster, A.P. (1974): The direct use of likelihood for significance testing, *Proceedings of Conference on Foundational Questions in Statistical Inference*, Department of Theoretical Statistics: University of Aarhus, 335–352.

Engle, R.F. (1982): Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.

Engle, R.F. and Ng, V.K. (1993): Measuring and testing the impact of news on volatility, *The Journal of Finance*, Vol. XLVIII, No. 5, 1749–1778.

Eraker, B. and Johannes, M.S. and Polson, N.G. (2000): The impact of jumps in volatility and returns, working paper, Graduate School of Business, University of Chicago.

Fleming, J. and Ostdiek, B. and Whaley, R.E. (1995): Predicting stock market volatility: a new measure, *Journal of Futures Markets*, **15**, 265–302.

Glosten, L.R. and Jagannathan, R. and Runkle, D.E. (1993): On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, **48**, 1779–1801.

Han, C. and Carlin, B.P. (2000): MCMC Methods for Computing Bayes Factors: A Comparative Review, Working paper, Division of Biostatistics, School of Public Health, University of Minnesota.

Heidelberger, P. and Welch, P. (1983): Simulation run length control in the presence of an initial transient, *Operations Research*, **31**, 1109–1144.

Hol, E. and Koopman, S.J. (2000): Forecasting the variability of stock index returns with Stochastic Volatility Models and implied volatility, Tinbergen Institute Discussion paper, No. TI 2000-104/4.

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

Meyer, R. and Yu, J. (2000): BUGS for a Bayesian Analysis of Stochastic Volatility Models, *The Econometrics Journal*, **3**, 198–215.

O'Hagan, A. (1991): Contribution to the discussion of "Posterior Bayes Factors", *Journal of the Royal Statistical Society, Series B*, **53**, 136.

O'Hagan, A. (1994): *Kendall's Advanced Theory of Statistics, Volume 2b: Bayesian Inference*, Edward Arnold, London.

O'Hagan, A. (1995): Fractional Bayes factors for model comparison (with discussion), *Journal of the Royal Statistical Society, Series B*, **57**, 99–138.

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

Schwarz, G. (1978): Estimating the dimension of a model, *Annals of Statistics*, **6**, 461-464.

Spiegelhalter, D.J. and Best, N.G. and Carlin, B.P. (1998): Bayesian deviance, the effective number of parameters, and the comparison of arbitrarily complex models, Research Report 98-009, http://www.med.ic.ac.uk/divisions/60/biointro.asp.

Spiegelhalter, D.J. and Thomas, A. and Best, N.G. and Gilks, W.R. (1996): *BUGS 0.5, Bayesian inference using Gibbs sampling. Manual (version ii).* MRC Biostatistics

Unit, Cambridge, UK.

Tauchen, G. and Pitts, M. (1983): The price variability-volume relationship on speculative markets, *Econometrica*, **51**, 485–505.

Taylor, S.J. (1982): Financial returns modelled by the product of two stochastic processes — a study of the daily sugar prices 1961-75. In Anderson, O.D., Editor, *Time Series Analysis: Theory and Practice, 1*, pages 203–226. North-Holland, Amsterdam.

Zhu, L. and Carlin, B. P. (2000): Comparing hierarchical Models for spatio-temporally misaligned data using the deviance information criterion, *Statistics in Medicine*, **19**, 2265–2278.

Appendix: Tables

The tables contain the parameter estimations for mean and standard deviation for every model based on the standard priors described in Section 4. The values for the DIC and p_D are summarized only for the set of modified models passing the Heidelberger and Welch convergence test for all parameters. The total running time for each of the models is also included.

Table 1: Parameter estimations and MC error for Model 1:

$$y_t \mid \theta_t = exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \quad u_t \overset{i.i.d.}{\sim} N(0,1), \quad t = 1, ..., n.$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2 = \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t, \quad v_t \overset{i.i.d.}{\sim} N(0, \tau^2), \quad t = 1, ..., n.$$

Model 1					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.695	0.082	0.001962
μ	0.00	3.16	-0.745	0.244	0.005914
ϕ	0.86	0.11	0.980	0.009	0.000511
τ	0.12	0.05	0.167	0.031	0.002275
time	9,763 sec				

Table 2: Parameter estimations and MC error for Model 2:

$$y_t \mid \theta_t, \alpha = \alpha + exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \quad u_t \overset{i.i.d.}{\sim} N(0,1), \quad t = 1, ..., n.$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2 = \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t, \quad v_t \overset{i.i.d.}{\sim} N(0,\tau^2), \quad t = 1, ..., n.$$

Model 2					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.695	0.080	0.001604
μ	0.00	3.16	-0.742	0.237	0.004826
ϕ	0.86	0.11	0.980	0.008	0.000384
τ	0.12	0.05	0.170	0.028	0.001726
α	0.00	3.16	0.010	0.017	0.000217
time	10,973 sec				

Table 3: Parameter estimations and MC error for Model 3:

$$y_t \mid \theta_t = exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \quad u_t \overset{i.i.d.}{\sim} N(0,1), \quad t = 1, ..., n.$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \psi, \tau^2 = \mu + \phi \cdot (\theta_{t-1} - \mu) + \psi \cdot (\theta_{t-2} - \mu) + v_t,$$

$$v_t \overset{i.i.d.}{\sim} N(0,\tau^2), \quad t = 1, ..., n.$$

Model 3					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.679	0.086	0.001199
μ	0.00	3.16	-0.790	0.261	0.003790
ϕ	0.00	0.58	0.188	0.165	0.013230
τ	0.12	0.05	0.295	0.055	0.003499
ψ	0.00	0.58	0.779	0.163	0.012970
time	17,095 sec				

Table 4: Parameter estimations and MC error for Model 4:

$$y_t \mid \theta_t = exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \quad u_t \overset{i.i.d.}{\sim} t_\nu, \quad t = 1, ..., n.$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2, \lambda = \mu + \phi \cdot (\theta_{t-1} - \mu) + v_t,\ v_t \overset{i.i.d.}{\sim} N(0, \tau^2), \quad t = 1, ..., n,$$

Model 4					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.585	0.093	0.003090
μ	0.00	3.16	-1.100	0.326	0.011060
ϕ	0.86	0.11	0.991	0.005	0.000186
τ	0.12	0.05	0.108	0.021	0.001131
ν	8.00	16.00	8.85	2.31	0.122800
time	96,598 sec				

Table 5: Parameter estimations and MC error for Model 5:

$$y_t \mid \theta_t, \rho = exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \quad t = 1, ..., n.$$

$$\theta_{t+1} \mid \theta_t, \mu, \phi, \tau^2, \rho = \mu + \phi \cdot (\theta_t - \mu) + \tau \cdot v_t, \quad t = 1, ..., n-1,$$

$$\text{with} \begin{pmatrix} u_t \\ v_{t+1} \end{pmatrix} \overset{i.i.d.}{\sim} N \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right\}, \quad \theta_0 \sim N(\mu, \tau^2).$$

Model 5					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.711	0.063	0.000687
μ	0.00	3.16	-0.690	0.177	0.001937
ϕ	0.86	0.11	0.974	0.010	0.000337
τ	0.12	0.05	0.194	0.032	0.001376
ρ	0.00	0.58	-0.415	0.088	0.001756
time	71,015 sec				

Table 6: Parameter estimations and MC error for MODEL 6:

$$y_t \mid \theta_t = exp\ (\tfrac{1}{2} \cdot \theta_t) \cdot u_t, \ \ u_t \overset{i.i.d.}{\sim} N(0,1), \ \ t = 1, ..., n.$$

$$\theta_t \mid \theta_{t-1}, \mu, \phi, \tau^2, \lambda = \mu + \phi \cdot (\theta_{t-1} - \mu) + \lambda \cdot (x_t - \bar{x}) + v_t, \ v_t \overset{i.i.d.}{\sim} N(0, \tau^2), \ \ t = 1, ..., n,$$

Model 6					
	Prior		Posterior		
Parameter	Mean	SD	Mean	SD	MC Standard error
β	3.49	11.67	0.695	0.016	0.001325
μ	0.00	3.16	-0.730	0.047	0.003720
ϕ	0.00	0.58	-0.276	0.117	0.004328
τ	0.12	0.05	0.435	0.080	0.008935
λ	1.00	0.58	0.153	0.016	0.000582
time	12,824 sec				

Table 7: Deviance summaries for the S&P100 data:

Model		$\overline{D}_{min}$	$\overline{D}_{max}$	p_{Dmin}	p_{Dmax}	DIC_{min}	DIC_{max}
1	basic SV model	3307.6	3313.0	83.2	84.9	3392.5	3396.2
2	intercept	3307.8	3311.1	84.9	87.0	3394.8	3396.5
3	AR(2)-process	3260.3	3264.3	115.5	118.6	3378.9	3379.8
4	t-distribution	3350.6	3353.0	49.1	50.4	3400.7	3402.7
5	leverage effect model	2989.2	2990.7	322.7	324.9	3313.4	3314.
6	implied volatility	3183.2	3196.5	116.2	124.2	3307.4	3312.7