Economics Department

Economics Working Papers

The University of Auckland *Year* 2002

MCMC Methods for Estimating Stochastic Volatility Models with Liverage Effects: Comments on Jacquier, Polson and Rossi (2002)

Jun Yu
University of Auckland, j.yu@auckland.ac.nz

This paper is posted at ResearchSpace@Auckland.
http://researchspace.auckland.ac.nz/ecwp/241

MCMC Methods for Estimating Stochastic Volatility Models with Leverage Effects: Comments on Jacquier, Polson and Rossi (2002)*

Jun Yu†

November 29, 2002

Abstract

In this note we represent the well known discrete time stochastic volatility (SV) model with a leverage effect and the SV model of Jacquier, Polson and Rossi (JPR) (2002) using Gaussian nonlinear state space forms with uncorrelated measurement and transition errors. With the new representations, we show that the JPR specification does not necessarily lead to a leverage effect and hence is not theoretically justified. Empirical comparisons of these two models via Bayesian MCMC methods reveal that JPR's specification is not supported by actual data either. Simulation experiments are conducted to study the sampling properties of the Bayes estimator for the conventionally specified model.

JEL classification: C11, C15, G12

Keywords: Bayesian estimation; State space models; Leverage effect; Quasi maximum likelihood.

1 Introduction

Markov Chain Monte Carlo (MCMC) methods have become one of the most important tools for estimating stochastic volatility (SV) models since it was introduced in Jacquier, Polson and Rossi (1994) to analyze the basic (ie lognormal) SV model. In a Monte Carlo

*I gratefully acknowledge financial support from the Royal Society of New Zealand Marsden Fund under number 01-UOA-015. I also wish to thank Eric Jacquier, Nick Polson, Xibin Zhang and especially Neil Shephard for comments on an earlier version of the paper and Hamish Egan for his excellent research assistance.

†Department of Economics, University of Auckland, Private Bag 92019, Auckland, New Zealand; email: j.yu@auckland.ac.nz

study, Andersen, Chung and Sorensen (1999) compared the finite sample performances of various methods for estimating the lognormal SV model and found that MCMC is the most efficient tool. Their finding is not surprising since MCMC provides a fully likelihood-based inference.

Motivated by the empirical evidence that the basic SV model can be too restrictive for many financial time series, Jacquier, Polson and Rossi (2002) (JPR hereafter) extend their earlier work to analyze a SV model which generalizes the basic SV model in two different dimensions, one by replacing the Gaussian innovation by a fat-tailed distribution of innovations and one by incorporating the so-called leverage effect. The SV model with a leverage effect (which is also termed the asymmetric SV (**A-SV** hereafter) model in Harvey and Shephard (1996)) is particularly important from the finance perspective and is connected directly to the continuous time SV models widely used in the finance literature on option pricing; see for example Hull and White (1987), Wiggins (1987), and Chesney and Scott (1989). Since the parameter which captures the leverage effect is of critical importance, JPR study the sampling properties of the Bayes MCMC estimator using Monte Carlo experiments and find little loss in precision from adding the leverage parameter into the basic model. They also fit the model to many real financial time series sequences and find overwhelming evidence of a strong leverage effect in most financial time series considered.

Although we find the proposed MCMC algorithm and simulation and empirical results in JPR very interesting, their specification of the A-SV model leaves us a bit puzzled. First, their specification does not correspond to the well known continuous time A-SV model commonly used in the finance literature and hence it is less useful from theoretical (such as option pricing) viewpoints. In fact, their specification is not even consistent with the efficient market hypothesis because the model is not a martingale sequence, as noted in Harvey and Shephard (1996). Second, we find their specification is empirically inferior to the conventional specification when SP500 data are used. Third, since their specification is of little theoretical relevance as well as of limited empirical importance, the sampling properties presented in their paper are not practically very useful.

One purpose of this note is to clarify the puzzle. To achieve this objective, we derive a Gaussian nonlinear state space representation of the conventionally specified A-SV model. Using the new representation we show that the conventional specification can capture the leverage effect, but the same argument is not necessarily true for the specification in JPR. We then fit both models to a commonly used stock index based on the all purpose Bayesian software package BUGS, as described in Meyer and Yu (2000) and show that JPR's specification is inferior, judged by Bayesian statistical criteria. The remainder of the note is organized as follows. Section 2 derives the state space representation of both models and explains why the leverage effect may not be warranted in JPR's specification. Section 3 fits both models to an SP500 index. In Section 4, we present the sampling properties of the conventionally specified A-SV model. Section 5 concludes.

2 State Space Representation of A-SV Models

In the finance literature on option pricing, the A-SV model is often formulated in terms of stochastic differential equations. The widely used lognormal A-SV model specifies the following equations for the logarithmic asset price $s(t)$ and the corresponding volatility $\sigma^2(t)$,

$$\begin{cases} ds(t) & = & \sigma(t)dB_1(t), \\ d\ln\sigma^2(t) & = & \alpha + \beta\ln\sigma^2(t)dt + \sigma_v dB_2(t), \end{cases} \tag{2.1}$$

where $B_1(t)$ and $B_2(t)$ are two Brownian motions, $corr(dB_1(t), dB_2(t)) = \rho$ and $s(t) = \ln S(t)$ with $S(t)$ being the asset price. When $\rho < 0$ we have the leverage effect. This negative correlation implies that a negative shock to the return increases the debt-equity ratio of a firm and so increases the riskiness of the firm in subsequent periods (see e.g. Black (1976)).

In the empirical literature the above model is often discretized to facilitate estimation. For instance, the Euler-Maruyama approximation leads to the following discrete time A-SV model:

$$\begin{cases} X_t & = & \sigma_t u_t, \\ \ln\sigma^2_{t+1} & = & \alpha + \phi\ln\sigma^2_t + \sigma_v v_{t+1}, \end{cases} \tag{2.2}$$

where $X_t = s_{t+1} - s_t$ is a continuously compounded return, $u_t = B_1(t+1) - B_1(t)$, $v_{t+1} = B_2(t+1) - B_2(t)$, $\phi = 1 + \beta$. Hence, u_t and v_t are independent iid $N(0,1)$ and $corr(u_t, v_{t+1}) = \rho$. This model is estimated by quasi maximum likelihood in Harvey and Shephard (1996) and by MCMC in Meyer and Yu (2000).

Comparing equation (2.2) with equation (8) in JPR, we note a small but important difference. Instead of assuming $corr(u_t, v_{t+1}) = \rho$, JPR adopt the specification of $corr(u_t, v_t) = \rho$. To fully understand the difference and also their linkage to the leverage effect, it is convenient to adopt the Gaussian nonlinear state space form with uncorrelated measurement and transition equation error terms. To do this, denote $w_{t+1} \equiv (v_{t+1} - \rho u_t)/\sqrt{1-\rho^2}$ and rewrite equation (2.2) as

$$\begin{cases} X_t & = & \sigma_t u_t, \\ \ln\sigma^2_{t+1} & = & \alpha + \phi\ln\sigma^2_t + \rho\sigma_v\sigma_t^{-1}X_t + \sigma_v\sqrt{1-\rho^2}w_{t+1}, \end{cases} \tag{2.3}$$

where w_t is iid $N(0,1)$ and $corr(u_t, w_{t+1}) = 0$.

Obviously $\ln\sigma^2_{t+1}$ is a linear function of X_t and $\partial\ln\sigma^2_{t+1}/\partial X_t = \rho\sigma_v/\sigma_t$. Therefore, if $\rho < 0$ and holding everything else constant, a fall in the stock return (ie $X_t < 0$) leads to an increase of $\ln\sigma^2_{t+1}$ and hence σ^2_{t+1}.

Using the same approach, we rewrite equation (8) in JPR in the following Gaussian nonlinear state space form:

$$\begin{cases} X_t &= \sigma_t u_t, \\ \ln \sigma_t^2 &= \alpha + \phi \ln \sigma_{t-1}^2 + \rho \sigma_v \sigma_t^{-1} X_t + \sigma_v \sqrt{1-\rho^2} w_t, \end{cases} \tag{2.4}$$

where w_t is iid $N(0,1)$ and $corr(u_t, w_t) = 0$. It is apparent in equation (2.4) that $\ln \sigma_t^2$ is related to X_t in a more complicated nonlinear way. Define

$$F(\ln \sigma_t^2, X_t) \equiv \ln \sigma_t^2 - \alpha - \phi \ln \sigma_{t-1}^2 - \rho \sigma_v \sigma_t^{-1} X_t - \sigma_v \sqrt{1-\rho^2} w_t.$$

The implicit function theorem implies that

$$\frac{\partial \ln \sigma_t^2}{\partial X_t} = -\frac{\partial F / \partial X_t}{\partial F / \ln \sigma_t^2} = \frac{\rho \sigma_v / \sigma_t}{1 + 0.5 \rho \sigma_v u_t}. \tag{2.5}$$

Similarly we can show that

$$\frac{\partial \ln \sigma_{t+1}^2}{\partial X_t} = \frac{\rho \sigma_v \sigma_{t+1} / \sigma_t}{1 + 0.5 \rho \sigma_v u_{t+1}}. \tag{2.6}$$

While the numerators in (2.5) and (2.6) are always negative when $\rho < 0$, the denominators can be either positive or negative. As a result, there is no guarantee that the partial derivative will always be negative and hence the leverage effect is not warranted.[1]

3 Estimation of A-SV models

MCMC estimation of the conventionally specified A-SV model can be done by using the state space representation (2.3) which leads to a log-concave full conditional distribution. In consequence, one can employ the adaptive rejection sampling algorithm of Gilks and Wild (1992). Alternatively, one can make use of the all purpose Bayesian software package BUGS, based on a different representation introduced in Meyer and Yu (2000). The full conditional distribution based on this alternative representation is, however, no longer log-concave and hence a Metropolis-Hastings (MH) updating step is needed. An advantage of the latter approach is that it can be easily modified to deal with JPR's specification. This alternative representation of the A-SV model is obtained by specifying the state and observation equations as follows:

$$\begin{aligned} h_{t+1} | h_t, \alpha, \phi, \sigma_v^2 &\sim N(\alpha + \phi h_t, \sigma_v^2), \\ X_t | h_t, h_{t+1}, \alpha, \phi, \sigma_v^2, \rho &\sim N\left(\frac{\rho}{\sigma_v} e^{h_t/2} (h_{t+1} - \alpha - \phi h_t), e^{h_t}(1-\rho^2) \right). \end{aligned}$$

[1] Given empirically possible values of ρ and σ_v, however, it appears much more likely for the denominator to take a negative value than a positive value. For example, the largest leverage effect reported in the literature is -0.66 (see Harvey and Shephard (1996)) and the largest estimate of σ_v is 0.85 (see Mahieu and Schotman (1998)). To ensure a leverage effect, u_t has to take a value larger than -3.565 and this occurs with probability 0.9998.

Regarding the prior distributions, for the parameters ϕ and σ_v^2, we follow exactly the prior specifications of Kim, Shephard and Chib (1998): $\sigma_v^2 \sim$ Inverse-Gamma(2.5,0.025) which has a mean of 0.167 and a standard deviation of 0.024 and $\phi^* \sim$ Beta-distribution with parameters 20 and 1.5 which has a mean of 0.167 and a standard deviation of 0.86 and 0.11, where $\phi^* = (\phi+1)/2$. The correlation parameter ρ is assumed to be uniformly distributed with support between -1 and 1 and hence is completely flat.

In all cases we choose a burn-in period of 10,000 iterations and a follow-up period of 100,000. The MCMC sampler is initialized by setting $\phi = 0.98, \sigma_v^2 = 0.025$, and $\rho = -0.5$. BUGS code can be downloaded from my web site

http://yoda.eco.auckland.ac.nz/~jyu/research.html

4 Empirical Comparison of A-SV Models

As argued in Section 2, the SV model corresponding to the continuous time model is different from the JPR specification which does not necessarily imply a leverage effect. However, nothing says the continuous time model is some form of "truth" and hence it is interesting to compare the empirical performance of these two alternative specifications. To do this, we employ a commonly used dataset which contains 2023 daily returns of S&P500 from 1980 to 1987. The same dataset is also used in JPR.

Since neither specification is nested by each other, we cannot use the classical likelihood ratio test to compare the performances of these two alternative models. Bayesian comparison is often made using the Bayes factor which involves the calculation of the marginal likelihood of the competing models. There are various ways to calculate the marginal likelihood. For instance, Kim et al. (1998) and Chib, Nardari and Shephard (2002) have shown how to compute the marginal likelihood at the posterior mean using the approach suggested by Chib (1995). However, this marginal likelihood approach remains a computationally intensive task and is not a particularly user-friendly tool. JPR propose an interesting way to calculate the Bayes factor by making use of the special structure of the models and priors. In this paper, we follow Newton and Raftery (1994)'s suggestion which uses the harmonic mean of the sampled likelihood values as a simulation consistent estimator of the required marginal likelihood. Alternative Bayesian comparison can be made via information criteria. In this paper we employ the newly developed deviance information criterion (DIC) proposed by Spiegelhalter, Best, Carlin and van der Linde (2002). As shown in Berg, Meyer and Yu (2002), DIC is a particularly user-friendly and effective tool for comparing SV models.

In Table 1 we summarize the results from estimation and model comparison, including the posterior means, standard deviations (SD), 95% Bayes confidence intervals for all the parameters, the harmonic mean estimates of log marginal likelihood, and DIC for both models. Although the leverage effect in both models is significant, it is markedly smaller in JPR's model. This suggests that if the leverage effect were estimated from JPR's model, it would be underestimated in magnitude by about 20%. Using the log marginal likelihood values we obtain the Bayes factor of the conventional A-SV model over JPR's A-SV model which is 2.44×10^{14}. The evidence strongly favors the conventional specification against JPR's specification. The same conclusion is also drawn from the comparison of the DIC values.[2]

An alternative way for comparing these two models is to nest both the models into a single model. To do so, consider the following specification,

$$\begin{cases} X_t &= \sigma_t u_t, \\ \ln \sigma_{t+1}^2 &= \alpha + \phi \ln \sigma_t^2 + \sigma_v v_{t+1}, \end{cases} \tag{4.7}$$

with

$$\begin{pmatrix} u_t \\ v_{t+1} \end{pmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{pmatrix}\right) \text{ and } \begin{pmatrix} u_t \\ v_t \end{pmatrix} \overset{iid}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{pmatrix}\right).$$

In this model we allow correlation at both time lags, but with different degrees of correlation. If $\rho_1 = 0$, we have JPR's A-SV model. If $\rho_2 = 0$, we have the conventionally used A-SV model.

To make use of the all purpose Bayesian software package BUGS, we obtain the following state and observation equations for the encompassed model:

$$\begin{aligned} h_{t+1}|h_t, \alpha, \phi, \sigma_v^2 &\sim N(\alpha + \phi h_t, \sigma_v^2), \\ X_t|h_{t+1}, h_t, h_{t-1}, \alpha, \phi, \sigma_v^2, \rho_1, \rho_2, &\sim N\Bigg(\frac{e^{h_t/2}}{\sigma_v}(\rho_2(h_t - \alpha - \phi h_{t-1}) + \rho_1(h_{t+1} - \alpha - \phi h_t)), \\ &\qquad e^{h_t}(1 - \rho_1^2 - \rho_2^2)\Bigg). \end{aligned}$$

We adopt the same prior distributions for ϕ and σ_v^2 as before. For ρ_1 and ρ_2 we assume a uniform prior with support between -1 and 1. Table 2 reports the estimation results, including the posterior means, standard deviations, 95% Bayes confidence intervals for all the parameters and the harmonic mean estimates of log marginal likelihood. The posterior mean of ρ_1 is -0.2939 while the posterior mean of ρ_2 is -0.2140. They compare to the posterior mean of -0.3179 in the conventional A-SV model and the posterior mean of -0.2599 in JPR's A-SV model. The 95% posterior credibility interval for ρ_1 is $[-0.4490, -0.1443]$ which indicates the presence of a significant negative correlation between u_t and v_{t+1}. The 95% posterior credibility interval for ρ_2 is $[-0.3619, -0.0828]$ which suggests some but weaker evidence of negative correlation between u_t and v_t. The marginal likelihood values from the encompassed model and JPR's model differs by 2.03×10^{11}. The evidence strongly favors the encompassed specification against JPR's specification. On the other hand, the marginal likelihood values from the conventional specification and the encompassed model differs by 1200.03 which favors the conventional specification against the encompassed specification. The overall ranking of three models is the conventional A-SV model comes first, followed by the the encompassed A-SV model and then by JPR's A-SV model.

[2]One has to choose the model with the smallest value of DIC.

5 Simulation Results

Since JPR's specification of the leverage effect is neither theoretically appealing nor empirically supported, the sampling properties of Bayes estimators reported in their paper are less practically relevant. Although the sampling properties of MCMC estimates of the continuous time SV model with the leverage effect are examined in Eraker, Johannes and Polson (2002), to the best of our knowledge, the sampling properties of MCMC estimates of the discrete time SV model with a correctly specified leverage effect remain unknown. On the other hand, understanding the finite sample performance of MCMC estimates is important from several aspects. First, it provides the reliability of MCMC estimates of the A-SV models, in particular of the new parameter, ρ. Second, since many more estimation tools have been developed to estimate the discrete time A-SV model than to the continuous time A-SV model, it is interesting to compare directly the performance of MCMC estimates with other estimates in the discrete time context. In this section, sampling experiments are designed to obtain sampling properties of the proposed MCMC estimates in the conventionally specified discrete time A-SV model.[3]

In the first experiment we use the similar parameter setting as in JPR. We simulate 100 samples of 1000 observations from the A-SV model (2.2). Simulation results such as the sample average and sample root mean square error (RMSE) are given in Table 3. The evidence proposed in Table 3 shows that Bayes estimates have very good sampling properties.

In the second experiment we adopt a parameter setting as in Harvey and Shephard (1996) that enables us to compare the relative efficiency of our Bayes estimate to the quasi-maximum likelihood (QML) estimate of Harvey and Shephard (1996). Table 4 reports the means and RMSEs of all the estimates. The simulation results for the QML estimates are obtained directly from Harvey and Shephard (1996). Our results are computed using 100 replications whereas Harvey and Shephard's results are based on 1000 replications. As expected, since MCMC is a fully likelihood-based method, it always performs better than QML. For example, relative efficiency of QML to MCMC in terms of the RMSE's are 0.5633, 0.7071 and 0.5909 respectively for ρ, ϕ and $\ln \sigma_v^2$.

6 Conclusions

In this note we propose a Gaussian nonlinear state space representation of both the classical A-SV model and the A-SV model of JPR. Using the new representation, we show that the leverage effect is not warranted in JPR's model. Moreover, our empirical analysis demonstrates that JPR's model is dominated by the classical A-SV model. Combined, the results necessitate a revaluation of the sampling properties of Bayes estimators.

[3]The sampling properties of MCMC estimates for the SV model with the fat-tailed error distribution have been obtained in Chib et al. (2002).

References

[1] Andersen, T., H. Chung and B. Sorensen (1999). Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics* 91, 61–87.

[2] Berg, A., R. Meyer and J. Yu (2002). Deviance Information Criterion for Comparing Stochastic Volatility Models. Working Paper, Department of Economics, University of Auckland.

[3] Black, F. (1976). Studies of stock market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section* 177–181.

[4] Chesney, M. and L.O. Scott (1989). Pricing European currency options: A comparison of the modified Black-Scholes model and a random variance model. *Journal of Financial and Quantitative Analysis* 24, 267–284.

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

[6] Chib, S., Nardari, F. and N. Shephard (2002). Markov Chain Monte Carlo methods stochastic volatility models. *Journal of Econometrics* 108, 281–316.

[7] Eraker, B., Johannes, M. and N. Polson (2002) The impact of jumps in volatility and returns. *Journal of Finance*, forthcoming.

[8] Gilks, W.R. and P. Wild (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics* 41, 337–48.

[9] Harvey, A.C. and N. Shephard (1996). The estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics* 14, 429–434.

[10] Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.

[11] Jacquier, E., N.G. Polson and P.E. Rossi (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* 12, 371–389.

[12] Jacquier, E., N.G. Polson and P.E. Rossi (2002). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, forthcoming.

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

[14] Mahieu, R. J. and P.C. Schotman (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics* 13, 330–360.

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

[16] Newton, M. and A.E. Raftery (1994). Approximate Bayesian inferences by the weighted likelihood bootstrap. *Journal of the Royal Statistical Society, Series B*, 56, 3–48 (with discussion).

[17] Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B*, 64, 583-639.

[18] Wiggins, J.B. (1987) Option values under stochastic volatility: Theory and empirical estimate. *Journal of Financial Economics* 19, 351–372.

Table 1: Empirical Results for S&P500

	Conventional A-SV			JPR A-SV		
	Mean	SD	95% CI	Mean	SD	95% CI
ϕ	.972	.0091	(.9511, .9871)	.9769	.0081	(.9587, .9902)
σ_v	.1495	.020	(.1139, .1928)	.1347	.0183	(.1031, .1759)
ρ	-.3179	.0855	(-.4749, -.1428)	-.2559	.0941	(-.4384, -.07295)
Log Marg Lik	-2801.6626			-2832.4874		
DIC	5441.740			5453.140		

Table 2: Empirical Results of Encompassed Model

	Mean	SD	95% CI
ϕ	0.9758	0.0085	(0.9568, 0.9898)
σ_v	0.1435	0.0199	(0.1097, 0.1859)
ρ_1	-0.2939	0.078	(-0.4490, -0.1443)
ρ_2	-0.2140	0.0715	(-0.3619, -0.0828)
Log Marg Lik	-2808.7527		

Table 3: Simulations for MCMC Estimates of the Conventional A-SV Model when the Sample Size is 1000

	True Value	Mean	RMSE
ρ	-0.6	-0.564	0.085
ϕ	0.95	0.945	0.0145
σ_v	0.26	0.254	0.037

Table 4: Simulations to Compare MCMC and QML Estimates of the Conventional A-SV Model when the Sample Size is 1000

		MCMC		QML	
	True Value	Mean	RMSE	Mean	RMSE
ρ	-0.9	-0.8815	0.0445	-0.911	0.079
ϕ	0.975	0.9732	0.00495	0.974	0.007
$\ln \sigma_v^2$	-4.605	-4.595	0.2086	-4.617	0.353