Realized Box-Cox Stochastic Volatility Models

Didit B. Nugroho and Takayuki Morimoto

Department of Mathematical Sciences, School of Science and Technology
Kwansei Gakuin University

The Third International Conference
High-Frequency Data Analysis in Financial Markets

November 16-18, 2012

Abstract

We extend the realized stochastic volatility (RSV) model proposed by Takahashi et al. (2009) to the realized Box-Cox stochastic volatility (RBCSV) model by applying the Box-Cox transformation to the volatility equation. Our empirical applications use daily returns and intraday returns of the TOPIX. To analyze the RBCSV models, we employ Chibs marginal likelihood method.

1. Introduction

- Yu et al. (2006) and Zhang and King (2008) have proposed the Box-Cox stochastic volatility (BCSV) models, where the basic model is expressed as

$$
\begin{aligned}
R_t &= \sqrt{g(\alpha_t, \delta)}\epsilon_t, \\
\alpha_{t+1} &= \mu + \phi(\alpha_t - \mu) + \tau\eta_{t+1},
\end{aligned}
$$

in which

$$
g(\alpha_t, \delta) = \begin{cases} (1 + \delta\alpha_t)^{\frac{1}{\delta}} & \\ \exp(\alpha_t) & \text{if } \delta = 0 \end{cases}.
$$

- To understand the interpretation and effect of δ, see Yu et al. (2006) and Zhang and King (2008).
- Employing daily real data, they found that the BCSV model is strongly favored by Bayes factors against the SV model.

- An alternative approach, intra-day high frequency data are used to construct an observable proxy for the latent volatility. This proxy is *realized volatility* (RV) estimator proposed by Andersen and Bollerslev (1998).
- Recently, Takahashi et al. (2009) have proposed an asymmetric SV model which utilizes returns and RV simultaneously:

$$\begin{aligned} R_t &= \exp\left(\tfrac{1}{2}h_t\right)\epsilon_t, \\ \log RV_t &= \beta + h_t + \sigma u_t \\ h_{t+1} &= \mu + \phi(h_t - \mu) + \tau\eta_t. \end{aligned}$$

- Our purpose is to study Box-Cox transformation for Takahashi's model, which is labelled realized stochastic volatility (RSV) hereafter.

2. RBCSV model

- The basic realized Box-Cox stochastic volatility (RBCSV) model is expressed as

$$
\begin{aligned}
R_t &= \sqrt{g(\alpha_t, \delta)}\epsilon_t, \\
\log RV_t &= \beta_0 + \beta_1 \alpha_t + \sigma u_t, \\
\alpha_{t+1} &= \mu + \phi(\alpha_t - \mu) + \tau \eta_{t+1}, \\
\alpha_1 &\sim \mathcal{N}\left(\mu, \tau^2 / \left(1 - \phi^2\right)\right)
\end{aligned}
$$

in which

$$
g(\alpha_t, \delta) = \begin{cases} (1 + \delta \alpha_t)^{\frac{1}{\delta}} & \\ \exp(\alpha_t) & \text{if } \delta = 0 \end{cases} .
$$

- As pointed out in Yu (2005), if we assume $corr\left(\epsilon_t, \eta_{t+1}\right) = \rho$, the specification implies that the above model is a martingale difference sequence and clear to interpret the leverage effect instead of assuming $corr\left(\epsilon_t, \eta_t\right) = \rho$ in Jacquire et al.(2004).

- The standard definition (for an equally spaced returns series) of the RV over a time interval of one day is

$$RV_t = \sum_{k=2}^{N_t} \left(p_{t_k} - p_{t_{k-1}}\right)^2,$$

where p_{t_k} denotes the log-price at the k'th observation in day t. Under no microstructure noise, the standard RV is a consistent estimator (see, Andersen et al. (2001)).

- To control for microstructure noise, Zhang et al. (2005) proposed a two scales realized volatility (TSRV) estimator based on subsampling, averaging and bias-correction:

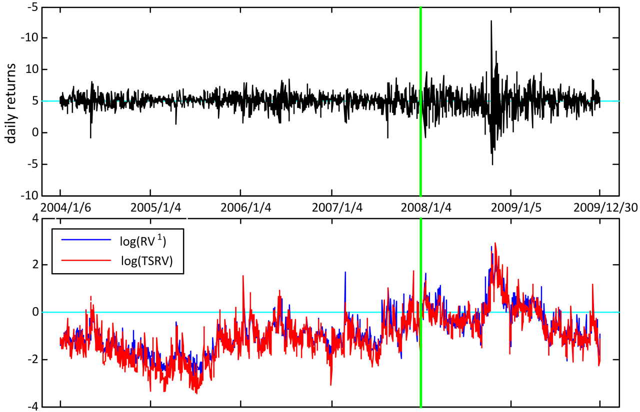
$$TSRV_t = \frac{1}{K}\sum_{k=1}^{K} RV_t^k - \frac{\overline{N}}{N_t} RV_t^{all},$$

where $\overline{N} = N_t / K$.

4. Estimation and Tools for Comparison

- We study the estimation of RBCSV models using daily returns of TOPIX with two alternative volatility proxies: RV 1min and TSRV 5min.
- To sample parameters, we employ the single-move MCMC algorithm for 25,000 iterations but discard the first 5,000 draws.
- The convergence and mixing performance of chain are respectively measured by convergence diagnostic test from Geweke (1991) and simulation inefficiency factors (SIF) based on Monte Carlo standard error (MCSE) with 50 batches.
- For model comparison, we calculate the Bayes factors, where the marginal likelihood is estimated by Chib's method (1995, 2001) from 20 iterations.
- The auxiliary particle filter algorithm with 50,000 particle points is employed to approximate the likelihood ordinate.

- We estimate the proposed model using daily data of TOPIX:
    1. from January 2004 to December 2007 and
    2. from January 2008 to December 2009 (including crash period).
- Time series plot of daily returns, RV^1, and TSRV:



6.1 Empirical Results

We found evidence of a negative correlation between returns and RV[1] (or TSRV):

Statistic	TOPIX 2004 - 2007		TOPIX 2008 - 2009	
	RSV	RBCSV	RSV	RBCSV
Model with correlation between returns and RV[1]				
Mean (SD)	-0.226 (0.039)	-0.226 (0.038)	-0.122 (0.054)	-0.120 (0.054)
95% HPD	(-0.302,-0.150)	(-0.305,-0.153)	(-0.228,-0.015)	(-0.226,-0.011)
SIF	10.476	8.653	8.843	6.920
Model with both heavy-tailed error and correlation between returns and RV[1]				
Mean (SD)	-0.227 (0.039)	-0.227 (0.039)	-0.122 (0.054)	-0.123 (0.055)
95% HPD	(-0.301,-0.149)	(-0.302,-0.148)	(-0.227,-0.013)	(-0.233,-0.013)
SIF	13.347	12.064	7.666	9.895
Model with both heavy-tailed error and correlation between returns and TSRV				
Mean (SD)	-0.320 (0.036)	-0.320 (0.036)	-0.176 (0.051)	-0.176 (0.053)
95% HPD	(-0.392,-0.250)	(-0.380,-0.261)	(-0.276,-0.075)	(-0.280,-0.069)
SIF	34.549	25.162	11.666	14.670

6.2 Empirical Results: TOPIX 2004 - 2007

Log marginal likelihood of RSV and RBCSV models using RV 1min.

Model	log-likelihood	log-marginal	$2\ln B_{10}$*	δ	
				Mean (SD)	90% HPD (MCSE)
Model with normal error					
RSV_0	-1797.34 (0.40)	-1817.74 (0.40)	4.39 (Positive)	–	–
$RBCSV_1$	-1793.58 (0.84)	-1815.55 (0.87)		-0.183 (0.138)	(-0.408,0.040) (0.015)
Model with correlation between returns and RV					
RSV_0	-1780.33 (0.67)	-1806.13 (0.70)	3.41 (Positive)	–	–
$RBCSV_1$	-1777.16 (0.47)	-1804.42 (0.58)		-0.194 (0.137)	(-0.421,0.025) (0.013)
Model with correlation between returns and volatility					
RSV_0	-1749.69 (0.24)	-1775.08 (0.26)	2.61 (Positive)	–	–
$RBCSV_1$	-1747.74 (0.24)	-1773.78 (0.38)		-0.115 (0.152)	(-0.354,0.140) (0.017)

*: Evidence for RBCSV model.

Log marginal likelihood of RSV and RBCSV models using RV 1min.

Model	log-likelihood	log-marginal	$2\ln B_{10}$	δ Mean (SD)	δ 90% HPD (MCSE)
Model with heavy-tailed error					
RSV_0	-1763.35 (0.52)	-1790.03 (0.51)	2.90 (Positive)	–	–
$RBCSV_1$	-1760.29 (0.42)	-1788.58 (0.37)		-0.192 (0.128)	(-0.408,0.005) (0.012)
Model with both heavy-tailed error and correlation between returns and RV					
RSV_0	-1745.06 (0.60)	-1771.02 (0.65)	2.89 (Positive)	–	–
$RBCSV_1$	-1742.22 (0.37)	-1769.57 (0.42)		-0.197 (0.147)	(-0.439,0.043) (0.013)
Model with both heavy-tailed error and correlation between returns and volatility					
RSV_0	-1712.27 (0.27)	-1738.85 (0.50)	48.06 (Very strong)	–	–
$RBCSV_1$	-1686.91 (0.59)	-1714.82 (0.70)		-0.050 (0.186)	(-0.353,0.257) (0.021)

6.2 Empirical Results: TOPIX 2004 - 2007

Log marginal likelihood of RSV and RBCSV models using TSRV 5min.

Model	log-likelihood	log-marginal	$2 \ln B_{10}$	δ	
				Mean (SD)	90% HPD (MCSE)
Model with both heavy-tailed error and correlation between returns and RV					
RSV_0	-1967.68 0.23	-1995.08 0.21	3.48 (Positive)	–	–
RBCSV_1	-1965.09 0.21	-1993.34 0.36		-0.213 (0.160)	(-0.477,0.047) (0.018)
Model with both heavy-tailed error and correlation between returns and volatility					
RSV_0	-1963.52 (0.15)	-1991.51 (0.37)	36.14 (Very strong)	–	–
RBCSV_1	-1944.29 (0.18)	-1973.44 (0.34)		-0.171 (0.186)	(-0.472,0.131) (0.022)

6.3 Empirical Results: TOPIX 2008 - 2009

Log marginal likelihood of RSV and RBCSV models using RV 1min.

Model	log-likelihood	log-marginal	$2 \ln B_{10}$	δ	
				Mean (SD)	90% HPD (MCSE)
Model with normal error					
RSV_0	-1269.08 (0.09)	-1288.92 (0.10)	3.12 (Positive)	–	–
$RBCSV_1$	-1269.35 (0.10)	-1287.36 (0.44)		-0.233 (0.091)	(-0.397,-0.106) (0.015)
Model with correlation between returns and RV					
RSV_0	-1265.20 (0.11)	-1290.18 (0.11)	3.49 (Positive)	–	–
$RBCSV_1$	-1264.48 (0.09)	-1288.44 (0.49)		-0.210 (0.089)	(-0.354,-0.067) (0.017)
Model with correlation between returns and volatility					
RSV_0	-1254.91 (0.11)	-1278.43 (0.22)	0.97 (Not worth)	–	–
$RBCSV_1$	-1253.87 (0.09)	-1277.94 (0.39)		-0.200 (0.064)	(-0.309,-0.096) (0.011)

Log marginal likelihood of RSV and RBCSV models using RV 1min.

Model	log-likelihood	log-marginal	$2 \ln B_{10}$	δ Mean (SD)	δ 90% HPD (MCSE)
Model with heavy-tailed error					
RSV_0	-1254.07 (0.09)	-1280.16 (0.09)	1.11 (Not worth)	–	–
$RBCSV_1$	-1253.26 (0.15)	-1279.61 (0.30)		-0.139 (0.089)	(-0.288,-0.012) (0.014)
Model with both heavy-tailed error and correlation between returns and RV					
RSV_0	-1250.50 (0.11)	-1275.82 (0.22)	3.07 (Positive)	–	–
$RBCSV_1$	-1249.81 (0.14)	-1274.29 (0.35)		-0.150 (0.143)	(-0.440,0.012) (0.028)
Model with both heavy-tailed error and correlation between returns and volatility					
RSV_0	-1239.26 (0.10)	-1264.83 (0.23)	9.24 (Strong)	–	–
$RBCSV_1$	-1236.70 (0.08)	-1260.21 (0.42)		-0.210 (0.086)	(-0.363,-0.067) (0.014)

6.3 Empirical Results: TOPIX 2008 - 2009

Log marginal likelihood of RSV and RBCSV models using TSRV 5min.

Model	log-likelihood	log-marginal	$2\ln B_{10}$	δ	
				Mean (SD)	90% HPD (MCSE)
Model with both heavy-tailed error and correlation between returns and RV					
RSV_0	-1243.20 (0.11)	-1267.70 (0.20)	2.40 (Positive)	–	–
RBCSV_1	-1242.49 0.14	-1266.50 0.20		-0.141 (0.144)	(-0.375,0.098) (0.029)
Model with both heavy-tailed error and correlation between returns and volatility					
RSV_0	-1242.53 (0.10)	-1266.67 (0.40)	10.74 (Very strong)	–	–
RBCSV_1	-1236.55 (0.14)	-1261.30 (0.40)		-0.149 (0.082)	(-0.287,-0.021) (0.014)

7. Comment

Based on the empirical results, we conclude:

- There is an evidence of leverage effect between returns and RV (or TSRV).
- The log-Bayes factors indicate the importance of incorporating both heavy-tailedness and leverage effects into RSV and RBCSV models.
- The log-Bayes factors also indicate very strong evidence supporting the BC transformation of squared volatility in the full RSV model, even though 90% Bayesian credible interval includes 0 for TOPIX of 2004-2007.

Some References

Andersen TG and Bollerslev T (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, **39**, 885–905.

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

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

Chib S and Jeliazkov I (2001). Marginal likelihood from the Metropolis-Hastings output. *Journal of the American Statistical Association*, **96**, 270–281.

Geweke J (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*. Paper presented at the Fourth Valencia International Meeting on Bayesian Statistics, Peiscola, Spain.

Some References (Continued)

Takahashi M, Omori Y and Watanabe T (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously. *Computational Statistics and Data Analysis*, **53**, 2404–2426.

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

Yu J, Yang Z and Zhang X (2006). A class of nonlinear stochastic volatility models and its implications for pricing currency options. *Computational Statistics and Data Analysis*, **51**, 2218–2231.

Zhang L, Mykland Per A and Aït-Sahalia Y (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, **100**(472), 1394–1411.

Zhang X and King ML (2008). Box-Cox stochastic volatility models with heavy-tails and correlated errors. *Journal of Empirical Finance*, **15**, 549–566.

Thanks for your attention!