The Korean Communications in Statistics
Vol. 13 No. 3, 2006, pp.491-501

Bayesian Change-point Model for ARCH[1)]

Seungmin Nam[2)], Juwon Kim[3)] and Sinsup Cho[4)]

Abstract

We consider a multiple change point model with autoregressive conditional heteroscedasticity (ARCH). The model assumes that all or the part of the parameters in the ARCH equation change over time. The occurrence of the change points is modelled as the discrete time Markov process with unknown transition probabilities. The model is estimated by Markov chain Monte Carlo methods based on the approach of Chib (1998). Simulation is performed using a variant of perfect sampling algorithm to achieve the accuracy and efficiency. We apply the proposed model to the simulated data for verifying the usefulness of the model.

Keywords : Bayesian change-point; Markov chain; perfect sampler.

1. Introduction

In finance and financial time series analysis, the volatility is essential to measure and forecast risks. The volatility of the financial time series has a tendency to change over time. Thus, it is difficult to explain the volatility of the series using the general time series and econometric models under the assumption of constant variance. The ARCH(Autoregressive Conditional Heteroscedasticity) model introduced in Engle (1982) has been proven to be very successful in modelling the volatility of financial time series such as stock returns and exchange rates. The ARCH model lets the conditional variance as the function of the information set to model the volatility. Many models extending the ARCH have been developed including GARCH(Generalized Autoregressive Conditional Heteroscedasticity Bollerslev (1986)), EGARCH(Exponential GARCH, Nelson (1991)) and so on. As an alternative to the ARCH type volatility models, Stochastic volatility model in which the variance is specified to follow some latent stochastic process is also widely used.

---

1) This paper was supported by the Korea Research Foundation Grant (KRF-2005-070-C00022) funded by the Korean Government (MOEHRD).
2) Automobile Pricing Department, Samsung Fire & Marine Insurance, Korea.
3) Graduate Student, Dept. of Statistics, Seoul National University, Seoul 151-742, Korea.
4) Professor, Dept. of Statistics, Seoul National University, Seoul 151-742, Korea.
Correspondence : sinsup@snu.ac.kr

In this paper, we focus on ARCH(q) models defined as follows:

$$y_t = \sqrt{h_t}\,\epsilon_t, \qquad h_t = \mu + \sum_{i=1}^{q} \alpha_i y_{t-i}^2, \tag{1}$$

where the ϵ_t are independent, identically distributed (usually Gaussian) errors with zero mean and unit variance. The ARCH(q) processes with $\mu > 0$, $\alpha_1, \cdots, \alpha_q \geq 0$ is covariance-stationary if, and only if, the associated characteristic equation has all roots outside the unit circle, (Engle, 1982 : Theorem 1). The stationary (unconditional) variance is given by $E(y_t^2) = \mu / (1 - \sum_{i=1}^{q} \alpha_i)$.

The assumption of the parameter stability over time plays a important role in statistical inference. If the parameters have changed within the observation period, then forecasts lose accuracy and the parameter estimates provide meaningless information. The existence and the location of change point might be very important information to understand the data. Therefore, the detection of possible changes in the data generating process has become an active area of research.

Kokoszka and Leipus (2000) studied the unknown change point problem for ARCH models with CUSUM type estimator. Their estimator is defined as follows:

$$\hat{k} = \min k : |R_k| = \max_{1 \leq j \leq n} |R_j|, \tag{2}$$

where

$$R_k = \frac{k(n-k)}{n^2} \left(\frac{1}{k} \sum_{j=1}^{k} y_j^2 - \frac{1}{n-k} \sum_{j=k+1}^{n} y_j^2 \right) \tag{3}$$

The consistency of the estimator is proved and its rate of convergence is established. Berkes et al. (2004) suggested a sequential detection scheme for the parameter change in a GARCH(p,q) by considering the quasi-likelihood scores.

In this paper, we perform a Bayesian analysis of the multiple change point ARCH model. The joint estimation of the parameters and the unknown change point is done using the Markov chain Monte Carlo (MCMC). Geweke (1989) implemented Bayesian inference on ARCH models with importance sampling, and Geweke (1994) implemented using Metropolis-Hastings algorithm. Nakatsuma (2000) proposed an MCMC method using a multivariate Metropolis-Hastings algorithm for Bayesian estimation of the ARCH/GARCH model. The Markov-switching model, which is similar to the Bayesian change point model, for ARCH process was proposed by Kaufmann and Frühwirth-Schanatter (2002). They combined the results for hidden-Markov process (Chib, 1996) with those for MCMC estimation of Nakatsma (2000). Kaufmann and Frühwirth-Schanatter (2002) allowed only one parameter to be time-varying. They assumed the following the switching ARCH model.

$$y_t = \sqrt{h_t}\,\epsilon_t\,, \tag{4}$$

$$h_t = \gamma I_t + \alpha_1 y_{t-1}^2 + \cdots + \alpha_q y_{t-q}^2 \tag{5}$$

Section 2 presents the Bayesian multiple change point ARCH model. In section 3, we explain the implementation and MCMC estimation of the model and shows the small simulation result compared with CUSUM type estimator. The model selection are also discussed. Conclusions are given in section 4.

2. Multiple change-point model for ARCH

We assume that the part or all of the parameters in the ARCH equation can be changed at unknown change points. That is, we consider the following model, for example, ARCH(1) case :

$$\begin{aligned} y_t &= \sqrt{h_t}\,\epsilon_t\,, \\ h_t &= \mu_j + \alpha_{1j} y_{t-1}^2\,, \end{aligned} \tag{6}$$

where $j = 1,\cdots,r+1$ when the number of change points is r. ARCH(q) can be easily extended from the above equations. The multiple change point model has generated an enormous literatures in many fields such as statistics and econometrics. In this paper, we adopt the Bayesian change point model proposed by Chib (1998). Chib's algorithm introduces a latent discrete state variable as the state of the system at time t which takes values on the integer $1,2,\cdots,r+1$, when the number of change points is r.

The variable s_t is modelled as a discrete-state Markov process with the transition probability matrix that specifies that s_t can either stay at the current value or jump to the next higher value. The one-step transition probability matrix can be written as

$$P = \begin{pmatrix} p_{11} & p_{12} & 0 & \cdots & 0 \\ 0 & p_{22} & p_{23} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \cdots & 0 & p_{rr} & p_{rr+1} \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \tag{7}$$

where $p_{ij} = \Pr(s_t = j \mid s_{t-1} = i)$ is the probability of moving to state j at time t given that the state at time $t-1$ is i if we specify r as the number of change points. Note that there is only one unknown element in each row of P. The model parameters and the unknown change points are sampled in a following manner. The algorithm is implemented by simulating full conditional distribution.

$$\Theta, P \mid \boldsymbol{Y}_t, S_t \tag{8}$$

$$S_t \mid \boldsymbol{Y}_t, \Theta, P \tag{9}$$

where $\Theta = \{\alpha, \mu\}$ and $S_t = \{s_1, \cdots, s_t\}$. Firstly, the states S_t are simulated given data and the parameters, and next, the parameters are updated using the simulated states. Simulation would be performed recursively. The change points can be detected at the time points where the states have changed.

Chib's method needs a prior specification of the number of change points, but it enables the probability of jump to be nonconstant. The advantage of Chib's method is that it allows all of the change points to be sampled simultaneously without a large increase in computations and so reduces the correlation between the sampled change points. For more detailed procedures, refer to section 2 of Chib (1998).

If $\epsilon_t \sim N(0,1)$, then the conditional distribution of y_t is Gaussian with mean 0 and variance h_t.

$$f(y_t \mid \mu, \alpha, Y_{t-1}) = \frac{1}{\sqrt{2\pi h_t}} exp\left(-\frac{y_t^2}{2h_t}\right), \tag{10}$$

where $\boldsymbol{Y_t} = (y_1, \cdots, y_t)$, $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_q)$ and $\alpha_{i\cdot} = (\alpha_{i1}, \cdots, \alpha_{i,r+1})$.

The likelihood of ARCH(q) is

$$\begin{aligned} f(\boldsymbol{Y_t} \mid \mu, \boldsymbol{\alpha}, P) &= \prod_{i=1}^{n} f(y_t \mid \boldsymbol{Y_{t-1}}, \mu, \boldsymbol{\alpha}, P), \\ &= \prod_{i=1}^{n} \left\{ \sum_{k=1}^{r+1} f(y_t \mid \boldsymbol{Y_{t-1}}, \mu, \boldsymbol{\alpha}, P, s_i = k) p(s_i = k \mid \boldsymbol{Y_{t-1}}, \mu, \boldsymbol{\alpha}, P) \right\} \end{aligned} \tag{11}$$

We assume the following prior density for μ_j and α_{ij}.

$$\begin{aligned} p(\mu_j) &\sim inverse\ \Gamma(m_1, m_2) \\ p\left(\alpha_{1j}, \cdots, \alpha_{qj}, 1 - \sum_{i=1}^{q} \alpha_{ij}\right) &\sim Dirichlet(v_1, \cdots, v_{q+1}) \end{aligned} \tag{12}$$

Note that the prior on $\boldsymbol{\alpha}$ is selected to satisfy the constraints that all α_{ij} are positive and $\sum_{i=1}^{q} \alpha_{ij}$ is smaller than 1. In Kaufmann and Frühwirth-Schanatter (2002) and Kim et al. (1998), the same prior was used for ARCH and GARCH models, respectively.

Since, both μ_i and $\boldsymbol{\alpha}$ have non-standard conditional posteriors Metropolis-Hastings algorithm is used to generate the samples from the posteriors. But, it is observed that the usual independent Metropolis-Hastings algorithm has some difficulties in achieving the acceptable convergence. Hence, to achieve accurate and efficient samples from the target density, we use an imperfect variant of a perfect sampling algorithm of the Metropolis-Hastings algorithm with an independent candidate density (Corcoran and Schneider, 2005). The key idea of the perfect sampling is to find a random epoch $-T$, such that, if we construct sample paths from every point in the state space starting at $-T$, then all paths will have coupled successfully by time 0. The value at zero is considered as a draw from the target density. Corcoran and Tweedie (2002) introduced the perfect sampling Metropolis-Hastings algorithm based on the backward coupling approach of Propp and Wilson (1996). Corcoran and Schneider (2005) proposed an 'imperfect' perfect sampling algorithm which is useful when it is impossible or difficult to maximize the ratio of certain densities.

3. Simulation results

In this section, we discuss the results of small simulations with the proposed change point model and the Kokoszka and Leipus (2000)'s CUSUM type estimator. The detection procedure by Berkes et al. (2004) requires rather large presamples for the estimation of parameters, which are not necessary for the implementation of our Bayesian method. Hence we only consider Kokoszka and Leipus (2000) for the comparison.

We generate various ARCH(1) series of length 1,000 which have a change point at the 500^{th} or the 750^{th} observation. For the estimation, the MCMC sampling is conducted for 3,000 iterations beyond a transient stage of 1,000 iterations. The posterior mode of the sampled change point candidates is considered as the estimator of change point in the data. The parameters in (6) of the data generation process (DGP) is following values:

Case1 : Change at 500^{th} obs., $\alpha : 0.1 \rightarrow 0.1,\ \mu : 1.0 \rightarrow 1.5,$

Case2 : Change at 750^{th} obs., $\alpha : 0.1 \rightarrow 0.1,\ \mu : 1.0 \rightarrow 1.5,$

Case3 : Change at 500^{th} obs., $\alpha : 0.1 \rightarrow 0.1,\ \mu : 1.0 \rightarrow 1.9,$

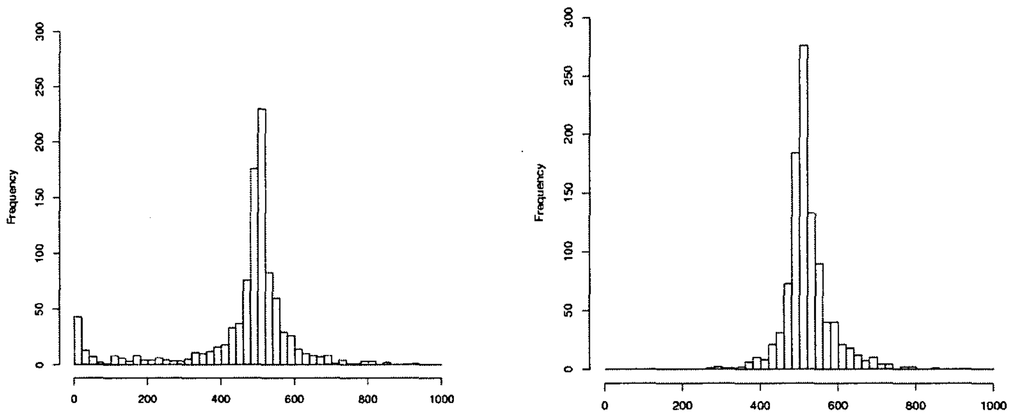
Case4 : Change at 750^{th} obs., $\alpha : 0.1 \rightarrow 0.1,\ \mu : 1.0 \rightarrow 1.9,$

Case5 : Change at 500^{th} obs., $\alpha : 0.5 \rightarrow 0.1,\ \mu : 1.5 \rightarrow 2.7,$

Case6 : Change at 750^{th} obs., $\alpha : 0.5 \rightarrow 0.1,\ \mu : 1.5 \rightarrow 2.7,$

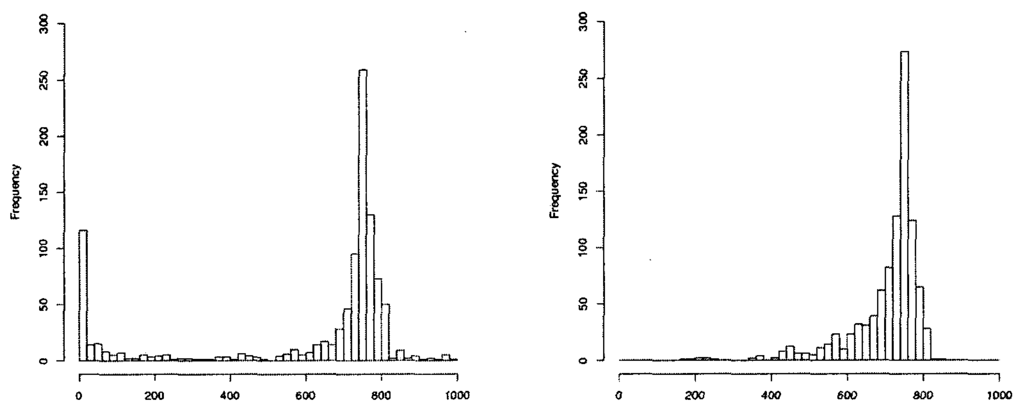
Note that the unconditional variances of the series increase by 50% at the specified change point for Case1 and Case2. For Case3 and Case4, they increase by 90%. In Case5 and Case6, the parameters are changed significantly, but the unconditional variance does not vary. The frequency distribution of the respective Cases are given in <Figure 1-6>.

As the model is assumed to have one change point, $\beta(50, 0.1)$ is used for the prior on p_{11} in (7), which is the probability of staying at state 1. This implies that the prior beliefs assume that the duration would be approximately 500 observations. We can observe that most of two estimators are distributed around the real change point when the unconditional variance changes.



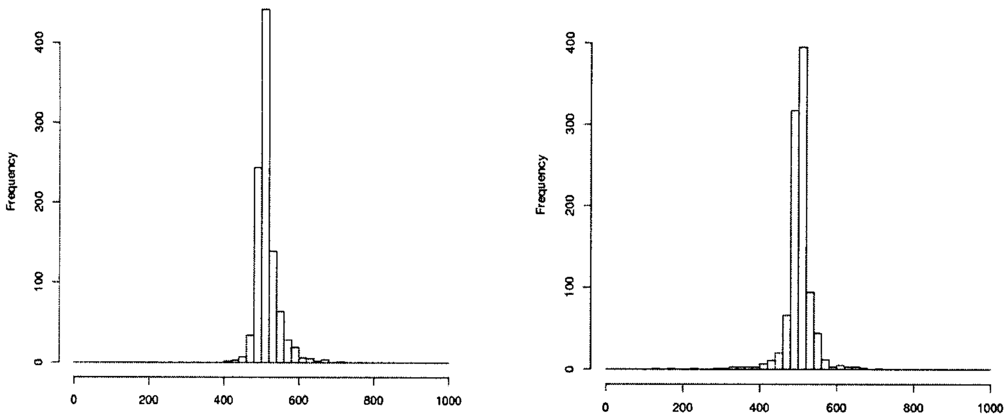
(a) Posterior mode (b) CUSUM estimator

< Figure 1> Histogram of the estimators for Case 1



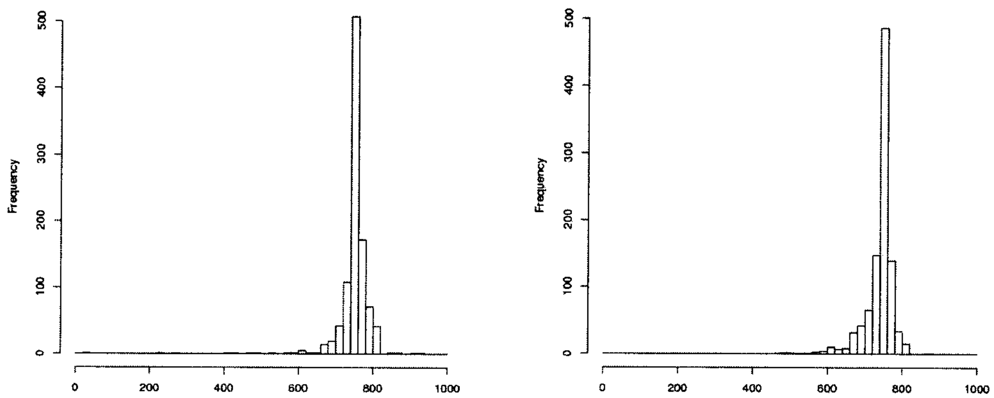
(a) Posterior mode (b) CUSUM estimator

<Figure 2> Histogram of the estimators for Case 2



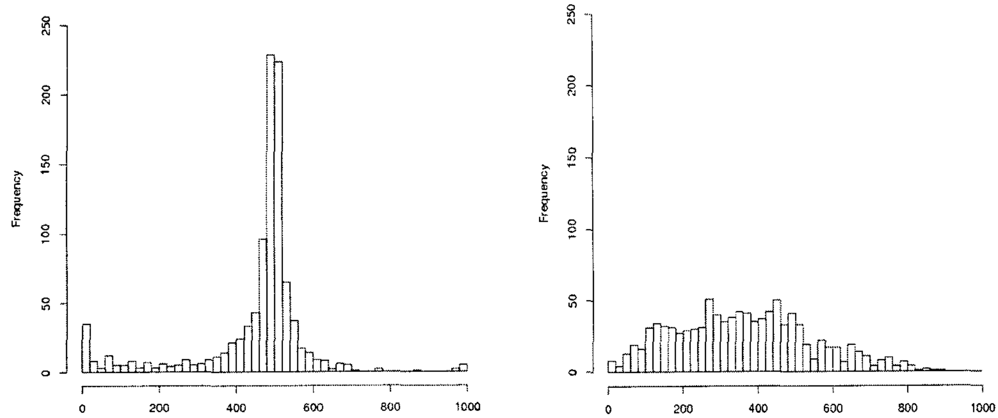
(a) Posterior mode (b) CUSUM estimator

<Figure 3> Histogram of the estimators for Case 3



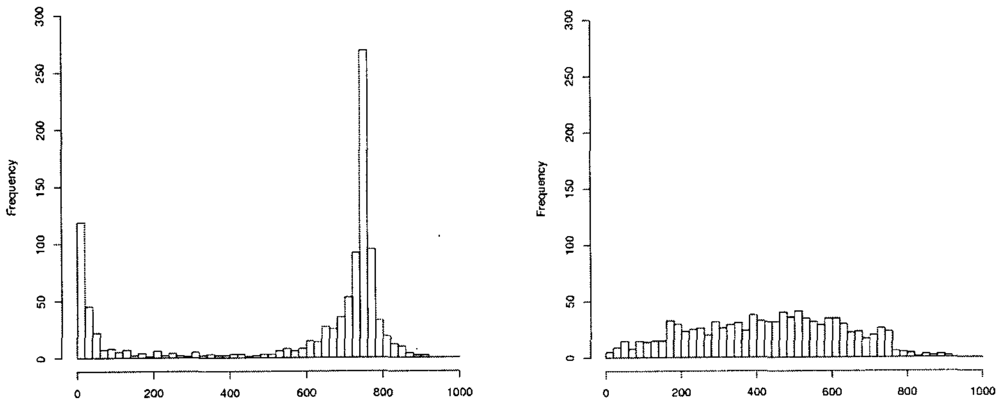
(a) Posterior mode (b) CUSUM estimator

<Figure 4> Histogram of the estimators for Case 4



(a) Posterior mode (b) CUSUM estimator

<Figure 5> Histogram of the estimators for Case 5



(a) Posterior mode (b) CUSUM estimator

<Figure 6> Histogram of the estimators for Case 6

<Table 1> Summaries of estimators

		Mean	Median	S.D.
Case1	Posterior mode	459.569	500	152.372
	CUSUM estimator	520.182	509	59.387
Case2	Posterior mode	604.913	745	282.375
	CUSUM estimator	706.524	740	92.284
Case3	Posterior mode	502.195	502	37.721
	CUSUM estimator	514.056	506.5	29.284
Case4	Posterior mode	749.112	752	48.156
	CUSUM estimator	739.435	750	38.430
Case5	Posterior mode	451.550	495	146.170
	CUSUM estimator	362.338	355	181.073
Case6	Posterior mode	562.026	728	297.903
	CUSUM estimator	430.286	442.5	197.833

In the Case1 and Case2, that is, when relatively small increase in variances happens (from our model's viewpoint, it is the small change in parameters), the posterior mode may be unable to detect any change point for a few series. In case of relatively large increase in variances, the posterior mode shows slightly better performance than CUSUM type estimator. The results of the Case5 and Case6 indicate big difference between two estimators. The CUSUM type estimator try to detect change point using the difference of the unconditional variance before and after change point. On the other hand, the proposed model detect change points by obtaining the estimates of the model parameters and calculating the conditional variance of the series. If the proposed model is applied to the data which has change point in parameters with insignificant unconditional variance change, it would improve the accuracy of the forecasts for the volatilities considerably. In <Table 1>, we have the summaries of change point estimators for 6 cases.

Generally speaking, the number of change points is not known in real examples. Since Chib(1998)'s Bayesian change point approach assumes that the number of change points are known before the estimation, it is necessary to determine the number of change points by comparing models with different number of change points. The Bayes factor which is the ratio of marginal likelihoods of the two models under comparison is the formal Bayesian model comparison criterion. The Bayes factor for two models M_1 and M_2 is defined as

$$B_{12} = \frac{p(Y_n | M_1)}{p(Y_n | M_2)} \tag{13}$$

where $p(Y_n | \bullet)$ is the marginal likelihood. The marginal likelihood is defined as,

$$p(Y_n | M) = \int p(Y_n | \theta, M) dp(\theta | M) \tag{14}$$

where $p(Y_n | \theta, M)$ is likelihood of observed data and $p(\theta | M)$ is the prior distribution of all parameters under model M.

4. Conclusion

In this paper, we discuss the Bayesian change point model for ARCH processes. To detect the change point of parameters, we adopt the hidden Markov model in which the transition probabilities of the hidden state are dependent on the regime. For the more exact estimation of the model, we use an imperfect variant of a perfect sampling algorithm. To show the applicability of the proposed model, we perform simulations of various settings. The reasonable results are obtained from the comparison with the CUSUM type estimator.

References

[1] Berkes, I., Gombay, E., Horváth, L., Kokoszka, P. (2004). Sequential change-point detection in GARCH(p,q) models. *Econometric Theory*, Vol. 20, 1140-1167.

[2] Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, Vol. 31, 307-327.

[3] Cassella, G., Lavine, M. and Robert, C. (2000), Explaining the perfect sampler. *Working paper 00-16*, State University of New York at Stony Brook, Duke University, Durham.

[4] Chib, S. (1996), Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, Vol. 75, 79-98.

[5] Chib, S. (1998), Estimation and Comparison of Multiple Change-point Models. *Journal of Econometrics*, Vol. 86, 221-241.

[6] Corcoran, J. and Tweedie, R. (2002), Perfect sampling from Independent Metropolis-Hastings Chains. *Journal of Statistical Planning and Inference*, Vol. 104, 297-314.

[7] Corcoran, J. and Schneider, U. (2005), Pseudo-perfect and adaptive variants of the Metropolis-Hastings algorithm with an independent candidate density. *Journal of Statistical Computation and Simulation*, Vol. 75,

459-475.

[8] Engle, R. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, Vol. 50, 987-1008.

[9] Geweke, J. (1989), Exact predictive densities in linear models with ARCH disturbances. *Journal of Econometrics*, Vol. 40, 63-86.

[10] Geweke, J. (1994), Bayesian comparison of econemetric models. *Working Paper 532*, Research Department, Federal Reserve Bank of Minneapolis.

[11] Kaufmann, S. and FrÜhwirth-Schanatter, S. (2002), Bayesian Analysis of Switching ARCH models. *Journal of Time Series Analysis*, Vol. 23, 425-458.

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

[13] Kokoszka, P. and Leipus, R. (2000), Change-point estimation in ARCH models. *Bernoulli*, Vol. 6, 513-539.

[14] Nakatsuma, T. (2000), Bayesian analysis of ARMA-GARCH models: A Markov chain sampling approach. *Journal of Econometrics*, Vol. 95, 57-69.

[15] Nelson, D. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, Vol. 59, 347-370.

[16] Propp, J. and Wilson, D. (1996), Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, Vol. 9, 223-252.

[Received October 2005, Accepted July 2006]