A BAYESIAN ANALYSIS FOR SHORT-TERM INTEREST RATE MODELS AND ITS APPLICATIONS

Chika Kawada

Received August 4, 2009

Abstract. This paper presents an analysis of models for Japanese short-term interest rate. The models are constructed based on mean reverting model using Bayesian method to capture the dynamics of short-term interest rate. The parameters of our models are estimated by marginal likelihood and posterior expectation and we shall make model selection using the information criterion EIC(extended information criterion). An application of the models will be implemented using weekly Japanese average interest rates on certificates of deposit (new issues) less than 30 days in the period from January 2001 to December 2008.

1 Introduction Correct modeling of the short-term interest rate is essential in finace, as it is this rate that is fundamental to the pricing of securities and important for risk management. In the study of the short-term interest rate dynamics, various models have been suggested. There are examples of these models, such as Vasicek model by Vasicek(1977), CIR model by Cox, Ingersoll and Ross(1985) and so on. An empirical comparison of these models was made by Chan, Karolyi, Longstaff and Sanders(1992). In the paper, the parameters are estimated by generalized method of moments(GMM) and they implemented the hypothesis testing methods developed by Newey and West for evaluation of the models. In the parameter estimation of mean reverting model using GMM, it is known that the results are easy to be influenced by an initial values, and it becomes often unstable.

Recently, Ahangarani(2005) employed maximum likelihood method for parameter estimation and log likelihood ratio test for model selection, and Kawada(2007) implemented GMM and made model selection using the information criterion EIC(extended information criterion) for various short-term interest rate models. Another articles of the interest rate model, using Bayesian framework, Jones(2003) published study about nonlinear drift of the interest rate model in detail, and Gray(2005) studied continuous time short rate models and the parameters estimated by Markov chain Monte Carlo(MCMC) method. Sanford and Martin(2006) made estimation using MCMC algorithm and model selection is made by Bayes factors for each model calculated using Savage-Dickey density ratio. In addition, a recent study by Hong and Lin(2006) examined a variety of short-term interest rate models including the single-factor diffusion models, GARCH models, Markov regime-switching models and jump-diffusion models for Chinese rates.

However, more useful models which capture the dynamics of the short-term interest rate are needed. In this paper, we suggest two hierarchical Bayes models based on mean reverting model. The parameters of the models are treated as random variables so that we have to consider appropriate prior distributions for them. By regarding parameters as random variables, we can consider that model construction and estimation involve necessarily some uncertainty and utilize the prior information of interest rate. In the first hierarchical Bayes model, we assume that volatility is constant, and in the second one, it depends on the

2000 *Mathematics Subject Classification.* 37M10, 62F15, 93E10 .
Key words and phrases. Hierarchical Bayesian Method, GMM, EIC, Short-Term Interest Rate.

interest rate. Recently, MCMC has been widely used in model estimation because of its usability, but it takes much time for calculation. Instead, we suggest parameter estimation of our hierarchical Bayes models using its marginal likelihood and posterior expectation, whose computation time is very short.

Although the issue of comparison of the models is important, it is difficult to evaluate relative performance of models in a consistent way. In this paper, we shall make model selection by the information criterion EIC. Using EIC, we can evaluate the model performance involving bias of model estimation. Our method of analysis of the interest rate is applied to Japanese short-term interest rate data.

This paper is structured as follows. In section 2, we review mean reverting model in discrete-time and construct two hierarchical Bayes models. Section 3 gives methods of parameter estimation for three models. Firstly, we outline GMM for mean reverting model following Chan et al.(1992). Secondly, we derive marginal likelihood and posterior expectations of hierarchical Bayes models. In section 4, we estimate these models with simulation data generated from our hierarchical Bayes models, and in section 5, the methodologies are applied to Japanese short-term interest rate data series. Summary and conclusions are made in section 6.

2 Models

2.1 Mean Reverting Model In the mean reverting model, it seems that the short-term interest rate has a long-term mean and moves to revert the mean. Consider a discrete-time econometric specification and let r_t be the short-term interest rate at time t. A dynamics for r_t which has the mean reverting property can be nested within the following equation,

$$\begin{aligned} r_{t+1} &= r_t + \alpha + \beta r_t + \epsilon_{t+1}, \quad t = 1, 2, \cdots, &(1)\\ \epsilon_{t+1} &= \sigma_r r_t^{\gamma} z_t; \quad z_t \sim N(0,1), \\ E[\epsilon_{t+1}] &= 0, \quad E[\epsilon_{t+1}^2] = \sigma_r^2 r_t^{2\gamma}. &(2)\end{aligned}$$

Many well known interest rate models such as CIR model and Vasicek model can be obtained from this model by giving the restrictions on the four parameters $\alpha, \beta, \sigma_r^2$, and γ.

Now we consider about a difference of the short-term interest rate between time $t+1$ and t, and develop its distribution.

Proposition 2.1. Let y_t be

$$y_t = r_{t+1} - r_t, \tag{3}$$

then y_t is conditionally normally distributed $N(\alpha + \beta r_t, \sigma_r^2 r_t^{2\gamma})$.

Proof. From equation (1), $y_t = \alpha + \beta r_t + \epsilon_{t+1}$, $E[y_t] = \alpha + \beta r_t$ and $Var[y_t] = \sigma_r^2 r_t^{2\gamma}$. Since y_t is a linear combination of z_t, y_t is $N(\alpha + \beta r_t, \sigma_r^2 r_t^{2\gamma})$. ◇

We call $\alpha + \beta r_t$ drift term and σ_r^2 volatility. Let parameter vector be $\boldsymbol{\theta} = [\alpha, \beta, \sigma_r^2, \gamma]^\top$, likelihood of y_t with n observations is derived as follows:

$$l(\boldsymbol{\theta}) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\sigma_r^2 r_t^{2\gamma}}} \exp\left[-\frac{(y_t - \alpha - \beta r_t)^2}{2\sigma_r^2 r_t^{2\gamma}}\right].$$

2.2 Hierarchical Bayes Model I As a special case of mean reverting model, we now assume that parameter γ in equation (2) equals to 0 and construct hierarchical Bayes model. Consider mean and variance in equation (1) are random variables and y_t is $N(\mu_t, \sigma^2)$, then describe the probability density of y_t with

$$f(y_t|\mu_t, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_t-\mu_t)^2}{2\sigma^2}\right]. \tag{4}$$

Further assume that the prior distribution for μ_t is $N(\psi_t, \sigma_0^2)$, and the prior distribution of two types of uniform and inverse gamma with scale parameter $\nu_0/2$ and shape parameter $\lambda_0/2$ in $\sigma^2 + \sigma_0^2$. These probability densities are written

$$g_1(\mu_t|\psi_t, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu_t-\psi_t)^2}{2\sigma_0^2}\right], \quad \psi_t = a_1 + b_1 r_t, \tag{5}$$

$$h_1(\sigma^2 + \sigma_0^2) = \frac{1}{k}, \tag{6}$$

$$h_2(\sigma^2 + \sigma_0^2) = \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})}(\sigma^2+\sigma_0^2)^{-(\frac{\nu_0}{2}+1)} \exp\left[-\frac{\lambda_0}{2(\sigma^2+\sigma_0^2)}\right], \tag{7}$$

$$k_1 : k_2 = \sigma^2 : \sigma_0^2,$$

where $\Gamma(\cdot)$ is gamma function, and λ_0, ν_0, k_1 and k_2 are hyper parameters and given as a prior information.(See Kawada(2008) about the detail of the model.)

2.3 Hierarchical Bayes Model II We constructed the hierarchical Bayes model I whose volatility's fluctuation is not influenced by the interest rate. However, it is generally known that the volatility is influenced by the interest rate, and it is thought that the model which includes influence of the interest rate may capture the dynamics of the interest rate in volatility well.

Therefore, we construct the model that depends upon influence of the interest rate in volatility term as hierarchical Bayes model II. Now we assume that the prior of normal distribution $N(\varphi_t, \sigma_0^2)$ in μ_t and inverse gamma distribution which has shape parameter c and scale parameter e^{-pr_t} in $\sigma^2 + \sigma_0^2$. In addition, we also assume uniform distribution as the prior in the shape parameter c. These probability densities are written as

$$g_2(\mu_t|\varphi_t, \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu_t-\varphi_t)^2}{2\sigma_0^2}\right], \quad \varphi_t = a_2 + b_2 r_t, \tag{8}$$

$$h_3(\sigma^2+\sigma_0^2|c, e^{-pr_t}) = \frac{(e^{-pr_t})^c}{\Gamma(c)}(\sigma^2+\sigma_0^2)^{-(c+1)} \exp\left[-\frac{e^{-pr_t}}{\sigma^2+\sigma_0^2}\right], \quad k_3 : k_4 = \sigma^2 : \sigma_0^2, \tag{9}$$

$$\kappa(c) = \frac{1}{l}, \tag{10}$$

where the ratio k_3 and k_4 are given as a prior information.

3 Method of Parameter Estimation In this paper, we use GMM for mean reverting model, and marginal likelihood and posterior expectation for our hierarchical Bayes models to estimate the parameters. This section provides the details of the method of parameter estimation for the models.

3.1 GMM(Generalized Method of Moments) Using GMM, we estimate the parameters of mean reverting model (1) and (2). For the estimation, now we rewrite the equation (1),

$$\epsilon_{t+1} = r_{t+1} - r_t - \alpha - \beta r_t,$$

and define the vector $f_t(\boldsymbol{\theta})$ be

$$f_t(\boldsymbol{\theta}) \equiv \begin{bmatrix} \epsilon_{t+1} \\ \epsilon_{t+1} r_t \\ \epsilon_{t+1}^2 - \sigma_r^2 r_t^{2\gamma} \\ (\epsilon_{t+1} - \sigma_r^2 r_t^{2\gamma}) r_t \end{bmatrix}.$$

If the parameter of a model assuming is correct, we have $E[f_t(\boldsymbol{\theta})] = \emptyset$ from (1) and (2). The GMM procedure consists of replacing $E[f_t(\boldsymbol{\theta})]$ with its sample counterpart, $\boldsymbol{g}_n(\boldsymbol{\theta})$, using the n observations where

$$\boldsymbol{g}_n(\boldsymbol{\theta}) = \frac{1}{n} \sum_{t=1}^{n} f_t(\boldsymbol{\theta}),$$

and then choosing parameter estimators that minimize the quadratic form,

$$J_n(\boldsymbol{\theta}) = \boldsymbol{g}_n^\top(\boldsymbol{\theta}) W_n \boldsymbol{g}_n(\boldsymbol{\theta}) = \sum_{i=1}^{4} \sum_{j=1}^{4} g_i^n(\boldsymbol{\theta}) w_{ij}^n g_j^n(\boldsymbol{\theta}),$$

where $W_n = (w_{ij}^n)$ is a positive-definite symmetric weighting matrix, $g_i^n(\boldsymbol{\theta})$ is ith element of vector $\boldsymbol{g}_n(\boldsymbol{\theta})$ and 4 is the number of parameters. See Chan et al.(1992) about selection of weighting matrix W_n.

3.2 Marginal Likelihood and Posterior Expectations of Hierarchical Bayes Model I We estimate the parameters in hierarchical Bayes model (4), (5), (6) and (7) using marginal likelihood and posterior expectation.

Proposition 3.1. The posterior distribution for μ_t of hierarchical Bayes model I in equation (5) is

$$\mu_t | \psi_t, \sigma^2 + \sigma_0^2, y_t \sim N\left(\frac{\psi_t \sigma^2 + y_t \sigma_0^2}{\sigma^2 + \sigma_0^2}, \frac{\sigma^2 \sigma_0^2}{\sigma^2 + \sigma_0^2} \right), \tag{11}$$

and its posterior expectation is

$$E[\mu_t | \psi_t, \sigma^2 + \sigma_0^2, y_t] = \frac{\psi_t \sigma^2 + y_t \sigma_0^2}{\sigma^2 + \sigma_0^2}. \tag{12}$$

Proof. Using the following Bayesian law, we can derive

$$p(\mu_t | \psi_t, \sigma^2 + \sigma_0^2, y_t) = \frac{f(y_t | \mu_t, \sigma^2) g_1(\mu_t | \psi_t, \sigma_0^2)}{\int_{-\infty}^{\infty} f(y_t | \mu_t, \sigma^2) g_1(\mu_t | \psi_t, \sigma_0^2) d\mu_t} = \frac{D_1}{D_2}. \tag{13}$$

We divide this expression into a numerator and a denominator for the advantage of the calculation. The numerator is developed as follows:

$$\begin{aligned} D_1 &= \frac{1}{\sqrt{2\pi\sigma^2}\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(y_t-\mu_t)^2}{2\sigma^2} - \frac{(\mu_t-\psi_t)^2}{2\sigma_0^2}\right] \\ &= \frac{1}{2\pi\sigma\sigma_0} \times \exp\left[\frac{-(\sigma^2+\sigma_0^2)\left\{\mu_t - \frac{(\psi_t\sigma^2+y_t\sigma_0^2)}{\sigma^2+\sigma_0^2}\right\}^2 + \frac{(\psi_t\sigma^2+y_t\sigma_0^2)^2}{\sigma^2+\sigma_0^2} - (\psi_t^2\sigma^2+y_t^2\sigma_0^2)}{2\sigma^2\sigma_0^2}\right], \end{aligned}$$

and we set

$$\begin{aligned} \varSigma &= \sigma^2+\sigma_0^2, \\ R_0 &= \frac{(\psi_t\sigma^2+y_t\sigma_0^2)^2}{\sigma^2+\sigma_0^2} - (\psi_t^2\sigma^2+y_t^2\sigma_0^2), \end{aligned}$$

then we have the numerator of the equation of the posterior distribution of μ_t

$$\begin{aligned} D_1 &= \frac{1}{2\pi\sigma\sigma_0}\exp\left[-\frac{\varSigma\{\mu_t - \frac{(\psi_t\sigma^2+y_t\sigma_0^2)}{\varSigma}\}^2 - R_0}{2\sigma^2\sigma_0^2}\right] \\ &= \sqrt{\frac{\varSigma}{2\pi\sigma^2\sigma_0^2}}\exp\left[-\frac{\varSigma\{\mu_t - \frac{(\psi_t\sigma^2+y_t\sigma_0^2)}{\varSigma}\}^2}{2\sigma^2\sigma_0^2}\right]\frac{1}{\sqrt{2\pi\varSigma}}\exp\left(\frac{R_0}{2\sigma^2\sigma_0^2}\right). \end{aligned} \tag{14}$$

The denominator is also rewritten as

$$\begin{aligned} D_2 &= \frac{1}{\sqrt{2\pi\varSigma}}\exp\left(\frac{R_0}{2\sigma^2\sigma_0^2}\right) \times \int_{-\infty}^{\infty}\sqrt{\frac{\varSigma}{2\pi\sigma^2\sigma_0^2}}\exp\left[-\frac{\varSigma\{\mu_t - \frac{(\psi_t\sigma^2+y_t\sigma_0^2)}{\varSigma}\}^2}{2\sigma^2\sigma_0^2}\right]d\mu_t \\ &= \frac{1}{\sqrt{2\pi\varSigma}}\exp\left(\frac{R_0}{2\sigma^2\sigma_0^2}\right). \end{aligned} \tag{15}$$

By the result of (14) and (15), we can rewrite equation (13) as follows:

$$p(\mu_t|\psi_t,\sigma^2+\sigma_0^2,y_t) = \sqrt{\frac{\varSigma}{2\pi\sigma^2\sigma_0^2}}\exp\left[-\frac{\varSigma\{\mu_t - \frac{(\psi_t\sigma^2+y_t\sigma_0^2)}{\varSigma}\}^2}{2\sigma^2\sigma_0^2}\right].$$

Further we transform the equation,

$$\begin{aligned} \sigma_B^2 &= \frac{\sigma^2\sigma_0^2}{\varSigma}, \\ \varLambda_\mu &= \frac{\psi_t\sigma^2+y_t\sigma_0^2}{\varSigma}, \end{aligned}$$

then we have

$$p(\mu_t|\psi_t,\sigma^2+\sigma_0^2,y_t) = \frac{1}{\sqrt{2\pi\sigma_B^2}}\exp\left[-\frac{(\mu_t-\varLambda_\mu)^2}{2\sigma_B^2}\right].$$

We find that this is the density of normal distribution which has mean $\varLambda_\mu$ and variance σ_B^2. Therefore the posterior distribution of μ_t of hierarchical Bayes model I in equation (5) is

$$\mu_t|\psi_t,\sigma^2+\sigma_0^2,y_t \sim N\left(\frac{\psi_t\sigma^2+y_t\sigma_0^2}{\sigma^2+\sigma_0^2}, \frac{\sigma^2\sigma_0^2}{\sigma^2+\sigma_0^2}\right),$$

and its posterior expectation is

$$E[\mu_t|\psi_t, \sigma^2+\sigma_0^2, y_t] = \frac{\psi_t\sigma^2 + y_t\sigma_0^2}{\sigma^2+\sigma_0^2}. \quad \diamond$$

We estimate the parameters a_1 and b_1 in $\psi_t = a_1 + b_1 r_t$ using marginal likelihood $l_1(\sigma^2+\sigma_0^2)$ of $\sigma^2+\sigma_0^2$ in hierarchical Bayes model I and $\sigma^2+\sigma_0^2$ using its posterior expectation.

Lemma 3.1. The marginal likelihood, $l_1(\sigma^2+\sigma_0^2)$ which is used in estimation with n observations $\boldsymbol{y}_n = [y_1, \cdots, y_n]$ is written as

$$l_1(\sigma^2+\sigma_0^2) = \{2\pi(\sigma^2+\sigma_0^2)\}^{-\frac{n}{2}} \times \exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}\sum_{t=1}^{n}(y_t-\psi_t)^2\right]. \tag{16}$$

Proof. Likelihood has the information of parameters depending on data, and marginal likelihood is provided by erasing unnecessary parameters using integral calculus from its likelihood. The marginal likelihood $l_1(\sigma^2+\sigma_0^2)$ (16) is derived as follows:

$$l_1(\sigma^2+\sigma_0^2) = \prod_{t=1}^{n}\int_{-\infty}^{\infty} f(y_t|\mu_t,\sigma^2)g_1(\mu_t|\psi_t,\sigma_0^2)d\mu_t.$$

Using the result of (15) in proposition 3.1 and we have

$$\begin{aligned}
l_1(\sigma^2+\sigma_0^2) &= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[\frac{1}{2\sigma^2\sigma_0^2}\left\{\frac{(\psi_t\sigma^2+y_t\sigma_0^2)^2}{\sigma^2+\sigma_0^2} - (\psi_t^2\sigma^2 + y_t^2\sigma_0^2)\right\}\right] \\
&= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[\frac{1}{2\sigma^2\sigma_0^2}\left\{\frac{(\psi_t\sigma^2+y_t\sigma_0^2)^2 - (\psi_t^2\sigma^2+y_t^2\sigma_0^2)(\sigma^2+\sigma_0^2)}{\sigma^2+\sigma_0^2}\right\}\right] \\
&= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \\
&\quad \exp\left[\frac{1}{2\sigma^2\sigma_0^2}\left\{\frac{(\psi_t^2\sigma^4+y_t^2\sigma_0^4+2\psi_t y_t\sigma^2\sigma_0^2) - (\psi_t^2\sigma^4+y_t^2\sigma_0^4+\psi_t^2\sigma^2\sigma_0^2+y_t^2\sigma^2\sigma_0^2)}{\sigma^2+\sigma_0^2}\right\}\right] \\
&= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[-\frac{1}{2\sigma^2\sigma_0^2}\left\{\frac{y_t^2\sigma^2\sigma_0^2 - 2\psi_t y_t\sigma^2\sigma_0^2 + \psi_t^2\sigma^2\sigma_0^2}{\sigma^2+\sigma_0^2}\right\}\right] \\
&= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[-\frac{1}{2\sigma^2\sigma_0^2}\left\{\frac{(y_t-\psi_t)^2\sigma^2\sigma_0^2}{\sigma^2+\sigma_0^2}\right\}\right] \\
&= \prod_{t=1}^{n}\frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}(y_t-\psi_t)^2\right] \\
&= \{2\pi(\sigma^2+\sigma_0^2)\}^{-\frac{n}{2}} \times \exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}\sum_{t=1}^{n}(y_t-\psi_t)^2\right]. \quad \diamond
\end{aligned} \tag{17}$$

In estimation, we choose the parameters a_1, b_1 and $\sigma^2+\sigma_0^2$ to maximize this marginal likelihood. However, seeing this marginal likelihood, we find that it maximizes if we estimate $\psi_t = a_1+b_1r_t, t = 1,\cdots,n$ minimizing a value of $\sum_{t=1}^n(y_t-\psi_t)^2$ regardless of $\sigma^2+\sigma_0^2$. Therefore, we estimate only a_1 and b_1 in this marginal likelihood and $\sigma^2+\sigma_0^2$ is estimated using posterior expectation. Next, we develop the posterior distribution of $\sigma^2+\sigma_0^2$ to obtain its expectation.

Proposition 3.2.(a). Using n observations vector $\boldsymbol{y}_n = [y_1,\cdots,y_n]$ the posterior distribution for $\sigma^2+\sigma_0^2$ with the prior (6) is derived

$$\sigma^2+\sigma_0^2|\boldsymbol{y}_n \quad\sim\quad Ga^{-1}\left(\frac{n}{2}-1, \frac{\sum_{t=1}^n(y_t-\psi_t)^2}{2}\right), \tag{18}$$

where Ga^{-1} describe inverse gamma distribution. The posterior expectation for the distribution (18) is calculated as

$$E\left[\sigma^2+\sigma_0^2|\boldsymbol{y}_n\right] \quad=\quad \frac{\frac{\sum_{t=1}^n(y_t-\psi_t)^2}{2}}{\frac{n}{2}-2}. \tag{19}$$

Proof. See Appendix A.1.1. ⋄

Proposition 3.2.(b). The posterior distribution for $\sigma^2+\sigma_0^2$ with the prior (7) is derived

$$\sigma^2+\sigma_0^2|\boldsymbol{y}_n \quad\sim\quad Ga^{-1}\left(\frac{n+\nu_0}{2}, \frac{\sum_{t=1}^n(y_t-\psi_t)^2+\lambda_0}{2}\right), \tag{20}$$

and the posterior expectation for the distribution (20) is calculated as

$$E\left[\sigma^2+\sigma_0^2|\boldsymbol{y}_n\right] \quad=\quad \frac{\frac{\sum_{t=1}^n(y_t-\psi_t)^2+\lambda_0}{2}}{\frac{n+\nu_0}{2}-1}. \tag{21}$$

Proof. See Appendix A.1.2. ⋄

Because $\psi_t = a_1+b_1r_t$, substituting the estimators of a_1 and b_1 into each of the posterior expectation of $\sigma^2+\sigma_0^2$, we have the estimators of $\sigma^2+\sigma_0^2$.

3.3 Marginal Likelihood and Posterior Expectations of Hierarchical Bayes Model II The parameters of hierarchical Bayes Model II are also estimated by its marginal likelihood and posterior expectation. We derive the marginal likelihood and the posterior distributions.

Proposition 3.3. The posterior distribution for μ_t of hierarchical Bayes model II in equation (8) is

$$\mu_t|\varphi_t,\sigma^2+\sigma_0^2,y_t \sim N\left(\frac{\varphi_t\sigma^2+y_t\sigma_0^2}{\sigma^2+\sigma_0^2}, \frac{\sigma^2\sigma_0^2}{\sigma^2+\sigma_0^2}\right), \tag{22}$$

and its posterior expectation is

$$E[\mu_t|\varphi_t, \sigma^2+\sigma_0^2, y_t] = \frac{\varphi_t\sigma^2 + y_t\sigma_0^2}{\sigma^2+\sigma_0^2}. \tag{23}$$

Proof. This proof can lead similarly Proposition 3.1 using Bayesian law. ◇

We need to develop the posterior distribution of $\sigma^2+\sigma_0^2$ in equation (9) because it is estimated using posterior expectation.

Proposition 3.4. The posterior distribution of $\sigma^2+\sigma_0^2$ is derived by following Bayesian law,

$$p(\sigma^2+\sigma_0^2|c, e^{-pr_t}, \boldsymbol{y}_n) = \frac{l_2(\sigma^2+\sigma_0^2)h_3(\sigma^2+\sigma_0^2)}{\int_0^\infty l_2(\sigma^2+\sigma_0^2)h_3(\sigma^2+\sigma_0^2)d(\sigma^2+\sigma_0^2)}, \tag{24}$$

where $l_2(\sigma^2+\sigma_0^2)$ is marginal likelihood of $\sigma^2+\sigma_0^2$ in hierarchical Bayes model II. We have

$$\sigma^2+\sigma_0^2|c, e^{-pr_t}, \boldsymbol{y}_n \sim Ga^{-1}\left(\frac{n}{2}+c, \frac{\sum_{t=1}^n (y_t-\varphi_t)^2}{2} + e^{-pr_t}\right), \tag{25}$$

and find that the expectation

$$E\left[\sigma^2+\sigma_0^2|c, e^{-pr_t}, \boldsymbol{y}_n\right] = \frac{\frac{\sum_{t=1}^n (y_t-\varphi_t)^2}{2} + e^{-pr_t}}{\frac{n}{2}+c-1}, \tag{26}$$

where $\varphi_t = a_2 + b_2 r_t$.

Proof. This proof can lead similarly Lemma 3.1 and Proposition 3.2.(b) using Bayesian law. ◇

The parameters a_2, b_2, c and p are estimated by marginal likelihood of c and e^{-pr_t} in hierarchical Bayes model II.

Proposition 3.5. The marginal likelihood $l(c, e^{-pr_t})$ is derived as follows:

$$l(c, e^{-pr_t}) = \left[\frac{\Gamma\left(c+\frac{1}{2}\right)}{\Gamma(c)}\right]^n (2\pi)^{-\frac{n}{2}} \prod_{t=1}^n (e^{-pr_t})^c \left[\frac{2}{(y_t-\varphi_t)^2 + 2e^{-pr_t}}\right]^{(c+\frac{1}{2})}, \tag{27}$$

for convenience of estimation,

$$\begin{aligned} \log l(c, e^{-pr_t}) &= -\frac{n}{2}\log(2\pi) + n\log\frac{\Gamma(c+\frac{1}{2})}{\Gamma(c)} \\ &\quad + \sum_{t=1}^n \left[c(-pr_t) + \left(c+\frac{1}{2}\right)\log\left\{\frac{2}{(y_t-\varphi_t)^2+2e^{-pr_t}}\right\}\right], \end{aligned} \tag{28}$$

Proof. See Appendix A.2. ◇

In estimation, we choose the parameters a_2, b_2, c and p to maximize this marginal likelihood. Substituting (a_2, b_2, c, p) estimated in (28) into the posterior expectations (26) and we have a estimator of $\sigma^2+\sigma_0^2$ in hierarchical Bayes model II.

4 Simulation Analysis In this section, we show the effectiveness and the validity of usage of marginal likelihood and posterior expectation in parameter estimation and the information criterion EIC for model selection. We generate simulation data from the models that we suggested and implement the parameter estimation of each model for the simulation data. Using the estimators, we calculate EIC and choose the model that the value is smallest as the most suitable one for the data which we analyzed. (See Ishiguro, Sakamoto and Kitagawa(1997) about the detail of EIC.) We generate 1000 simulation data for three of hierarchical Bayes models.

For the comparison, we also estimate mean reverting model (1) and (2) using GMM for the simulation data. Estimation results are shown in from Table 1 to Table 3. The hierarchical Bayse models used in the simulation are as follows.

(i)Simulation data 1Fhierarchical Bayes model I(h_1) with uniform distribution (see (4), (5) and (6))

$$\begin{aligned} y_t &\sim N(\mu_t, \sigma^2), \\ \mu_t &\sim N(a_1 + b_1 r_t, \sigma_0^2), \\ \sigma^2 + \sigma_0^2 &\sim Uni(0, k), \quad \sigma^2 : \sigma_0^2 = k_1 : k_2, \end{aligned}$$

$$a_1 = 1.6, b_1 = -0.2, k = 1.0, k_1 : k_2 = 7 : 3.$$

The estimation result is given in Table 1. It is shown that the most suitable model is hierarchical Bayes model I(h_2), though the true model is hierarchical Bayes model I(h_1). This can be considered due to the estimation bias.

(ii)Simulation data 2Fhierarchical Bayes model I(h_2) with inverse gamma distribution (see (4), (5) and (7))

$$\begin{aligned} y_t &\sim N(\mu_t, \sigma^2), \\ \mu_t &\sim N(a_1 + b_1 r_t, \sigma_0^2), \\ \sigma^2 + \sigma_0^2 &\sim Ga^{-1}\left(\frac{\nu_0}{2}, \frac{\lambda_0}{2}\right), \quad \sigma^2 : \sigma_0^2 = k_1 : k_2, \end{aligned}$$

$$a_1 = 0.6, b_1 = -0.15, \nu_0 = 3.0, \lambda_0 = 6.0, k_1 : k_2 = 7 : 3.$$

The estimation result is given in Table 2. It is shown that the most suitable model is hierarchical Bayes model I(h_2).

(iii)Simulation data 3Fhierarchical Bayes model II (see (4), (8), (9) and (10))

$$\begin{aligned} y_t &\sim N(\mu_t, \sigma^2), \\ \mu_t &\sim N(a_2 + b_2 r_t, \sigma_0^2), \\ \sigma^2 + \sigma_0^2 &\sim Ga^{-1}\left(c, e^{-pr_t}\right), \quad \sigma^2 : \sigma_0^2 = k_3 : k_4, \\ c &\sim Uni(0, l) \end{aligned}$$

$$a_2 = 0.7, b_2 = -0.2, p = 0.1, l = 10.0, k_3 : k_4 = 8 : 2.$$

The estimation result is given in Table 3. It is shown that the most suitable model is hierarchical Bayes model II.

Note that the estimate of $\sigma^2 + \sigma_0^2$ in hierarchical Bayes model II depends on time t and we have estimates of the same number as the data, therefore it is described as expectation.

5 Data Analysis The weekly Japanese average interest rates on certificates of deposit (new issues) less than 30 days in the period from January 2001 to December 2008 which is 418 observations are used in this practical analysis. Figure 1 shows the data r_t and Figure 2 shows $y_t = r_{t+1} - r_t$. Because the values of the data are so small, we extend the data 100 times and use it for the analysis. In this estimation, the values of hyper parameter are $\lambda_0 = 1.0$, $\nu_0 = 5.0$, $k_1 = k_3 = 0.8$, and $k_2 = k_4 = 0.2$ respectively. The result is shown in Table 4, in which the most suitable model is hierarchical Bayes model II.

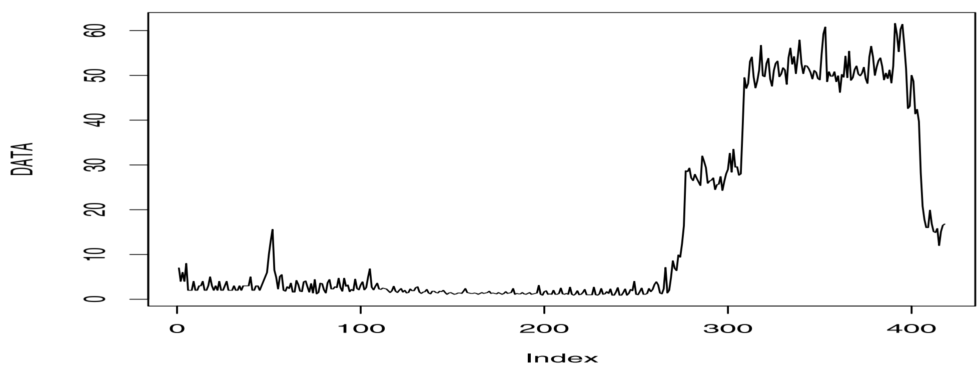

Figure 1: Short-Term Interest Rate Data

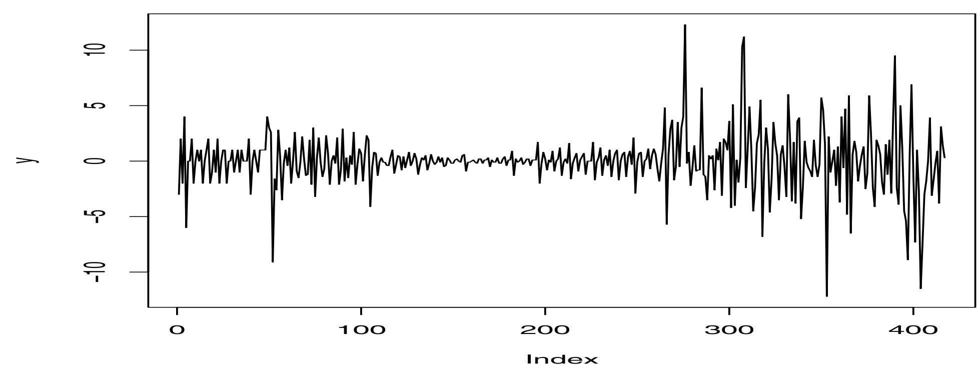

Figure 2: First Differenced Rate Data

6 Summary and Conclusions In this paper, we constructed more useful models based on mean reverting model to express the dynamics of short-term interest rate using hierarchical Bayes method. We further suggested a method of parameter estimation for the models using its marginal likelihood and posterior expectation and made model selection by EIC. In adopting Bayesian approach uncertainty relative to all unkown parameters is infered from likelihood and posterior probabilities, which reflect both data and prior information about the interest rate process.

Simulation results showed that our method is effective and valid in the modeling of the short-term interest rate and estimation. Seeing the estimation results of Japanese interest rate data, we found that our models are more suitable than the mean reverting model for the interest rate data which we analyzed in the model selection using EIC. In comparison with GMM and MCMC, large reduction of calculation time is realized by using marginal likelihood and posterior expectation for estimation. And we can apply the information criterion EIC that is an inclusive model evaluation method.

Hierarchical Bayes method allows to infer an influence of drift and volatility in the interest rate at each hierarchy. For future prospects, it is necessary to consider more useful models of drift term, and to further examine how to take in influence of the interest rate in volatility.

A Appendix

A.1 Proof of Proposition 3.2 In this Appendix, we develop the posterior distributions and expectations of $\sigma^2+\sigma_0^2$ of hierarchical Bayes model I in proposition 3.2.

A.1.1 Proof of Proposition 3.2.(a) Firstly, we consider the posterior distribution of $\sigma^2+\sigma_0^2$ which has uniform distribution as the prior in equation (6). It is derived using following Baysian law:

$$p(\sigma^2+\sigma_0^2|\boldsymbol{y}_n)=\frac{l_1(\sigma^2+\sigma_0^2)h_1(\sigma^2+\sigma_0^2)}{\int_0^\infty l_1(\sigma^2+\sigma_0^2)h_1(\sigma^2+\sigma_0^2)d(\sigma^2+\sigma_0^2)}=\frac{D_3}{D_4}, \tag{29}$$

where $l(\sigma^2+\sigma_0^2)$ is the marginal likelihood in Lemma 3.1. We divide this expression into numerator and denominator for the advantage of the calculation. The denominator advanced as follows:

$$\begin{aligned}D_4 &= \int_0^\infty \{2\pi(\sigma^2+\sigma_0^2)\}^{-\frac{n}{2}}\times\frac{1}{k}\times\exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}\sum_{t=1}^n(y_t-\psi_t)^2\right]d(\sigma^2+\sigma_0^2)\\ &= -(2\pi)^{-\frac{n}{2}}\times\frac{1}{k}\times\int_0^\infty(\sigma_0^2+\sigma^2)^{-\frac{n}{2}}\times\exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}\sum_{t=1}^n(y_t-\psi_t)^2\right]d(\sigma^2+\sigma_0^2).\end{aligned}$$

Now we set

$$\begin{aligned}P&=\frac{1}{\sigma^2+\sigma_0^2},\\ dP&=-\frac{1}{\sigma^2+\sigma_0^2}d(\sigma^2+\sigma_0^2),\quad d(\sigma^2+\sigma_0^2)=-P^{-2}dP,\\ T&=\sum_{t=1}^n(y_t-\psi_t)^2,\end{aligned}$$

then we rewrite

$$D_4=-(2\pi)^{-\frac{n}{2}}\times\frac{1}{k}\int_0^\infty P^{\frac{n}{2}-2}e^{-\frac{T}{2}P}dP.$$

Futher we set

$$Q = \frac{T}{2}P,$$
$$dQ = \frac{T}{2}dP,$$

then

$$\begin{aligned} D_4 &= -(2\pi)^{-\frac{n}{2}} \times \frac{1}{k}\int_0^\infty \left(\frac{2}{T}Q\right)^{(\frac{n}{2}-2)} e^{-Q} \times \left(\frac{2}{T}\right) dQ \\ &= -(2\pi)^{-\frac{n}{2}} \times \left(\frac{2}{T}\right)^{\frac{n}{2}-1} \times \frac{1}{k}\int_0^\infty Q^{(\frac{n}{2}-1)-1} e^{-Q} dQ, \end{aligned}$$

we find that

$$\int_0^\infty Q^{(\frac{n}{2}-1)-1} e^{-Q} dQ = \Gamma\left(\frac{n}{2}-1\right).$$

Therefore we have

$$D_4 = -(2\pi)^{-\frac{n}{2}} \times \left(\frac{2}{T}\right)^{\frac{n}{2}-1} \times \frac{1}{k} \times \Gamma\left(\frac{n}{2}-1\right).$$

Do it likewise, the numerator is rewritten as follows:

$$D_3 = (2\pi)^{-\frac{n}{2}} \times \frac{1}{k} \times P^{\frac{n}{2}-1} \times \exp\left[-\frac{T}{2}P\right].$$

Consequently, the equation of the posterior distribution of $\sigma^2 + \sigma_0^2$ is

$$\begin{aligned} \frac{D_3}{D_4} &= \frac{-(2\pi)^{-\frac{n}{2}} \times \frac{1}{k} \times P^{\frac{n}{2}-1} \times \exp\left[-\frac{T}{2}P\right]}{-(2\pi)^{-\frac{n}{2}} \times (\frac{2}{T})^{\frac{n}{2}-1} \times \frac{1}{k} \times \Gamma(\frac{n}{2}-1)} \\ &= \left(\frac{T}{2}\right)^{\frac{n}{2}-1} \times \frac{1}{\Gamma(\frac{n}{2}-1)} \times P^{\frac{n}{2}-2} \times e^{-\frac{T}{2}P}. \end{aligned}$$

We set

$$s = \frac{n}{2} - 1,$$
$$u = \frac{T}{2},$$

then it is rewritten as follows:

$$\begin{aligned} \frac{D_3}{D_4} &= \frac{u^s}{\Gamma(s)} P^{s-1} e^{-uP} \\ &= \frac{u^s}{\Gamma(s)} (\sigma^2 + \sigma_0^2)^{-(s+1)} e^{-\frac{u}{\sigma^2+\sigma_0^2}}, \end{aligned}$$

We find that this is the density of inverse gamma $Ga^{-1}(s, u)$, therefore the posterior distribution of $\sigma^2 + \sigma_0^2$ is derived as follows:

$$\sigma^2 + \sigma_0^2 | \boldsymbol{y}_n \sim Ga^{-1}\left(\frac{n}{2} - 1, \frac{\sum_{t=1}^n (y_t - \psi_t)^2}{2}\right).$$

Because of the expectation of inverse gamma distribution $Ga^{-1}(s,u)$ is $u/(s-1)$, the posterior expectation of (18) is

$$E\left[\sigma^2+\sigma_0^2|\boldsymbol{y}_n\right] \quad = \quad \frac{\frac{\sum_{t=1}^n (y_t-\psi_t)^2}{2}}{\frac{n}{2}-2}.$$

A.1.2 Proof of Proposition 3.2.(b) Secondly, we consider the posterior distribution of $\sigma^2+\sigma_0^2$ which has the prior distribution of inverse gamma in equation (7). The Bayesian law is applied to this derivation:

$$p(\sigma^2+\sigma_0^2|\boldsymbol{y}_n) = \frac{l_1(\sigma^2+\sigma_0^2)h_2(\sigma^2+\sigma_0^2)}{\int_0^\infty l_1(\sigma^2+\sigma_0^2)h_2(\sigma^2+\sigma_0^2)d(\sigma^2+\sigma_0^2)} = \frac{D_5}{D_6}. \tag{30}$$

The numerator is derived as follows:

$$\begin{aligned} D_5 \quad &= \quad \{2\pi(\sigma^2+\sigma_0^2)\}^{-\frac{n}{2}} \times \exp\left[-\frac{\sum_{t=1}^n (y_t-\psi)^2}{2(\sigma^2+\sigma_0^2)}\right] \\ &\quad \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})}(\sigma^2+\sigma_0^2)^{-(\frac{\nu_0}{2}+1)} \exp\left[-\frac{\lambda_0}{2(\sigma^2+\sigma_0^2)}\right], \end{aligned}$$

and the denominator is derived as follows:

$$\begin{aligned} D_6 \quad &= \quad \int_0^\infty \{2\pi(\sigma^2+\sigma_0^2)\}^{-\frac{n}{2}} \times \exp\left[-\frac{\sum_{t=1}^n (y_t-\psi)^2}{2(\sigma^2+\sigma_0^2)}\right] \\ &\quad \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})}(\sigma^2+\sigma_0^2)^{-(\frac{\nu_0}{2}+1)} \exp\left[-\frac{\lambda_0}{2(\sigma^2+\sigma_0^2)}\right] d(\sigma^2+\sigma_0^2) \\ &= \quad (2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} \int_0^\infty P^{\frac{n+\nu_0}{2}+1} \exp\left[-\frac{T+\lambda_0}{2}P\right] \times (-P^2)dP \\ &= \quad -(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} \int_0^\infty P^{\frac{n+\nu_0}{2}-1} \exp\left[-\frac{T+\lambda_0}{2}P\right] dP. \end{aligned}$$

Now we set for advantage,

$$\begin{aligned} M &= \frac{T+\lambda_0}{2}P, \\ P &= \frac{2}{T+\lambda_0}M, \\ dP &= \frac{2}{T+\lambda_0}dM, \end{aligned}$$

then we have

$$\begin{aligned} D_6 &= -(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} \times \int_0^\infty \left(\frac{2}{T+\lambda_0}M\right)^{\frac{n+\nu_0}{2}-1} e^{-M}\frac{2}{T+\lambda_0}dM \\ &= -(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma\left(\frac{\nu_0}{2}\right)} \times \left(\frac{2}{T+\lambda_0}\right)^{\frac{n+\nu_0}{2}} \times \int_0^\infty M^{\frac{n+\nu_0}{2}-1}e^{-M}dM \\ &= -(2\pi)^{-\frac{n}{2}} \times \frac{\left(\frac{\lambda_0}{2}\right)^{\frac{\nu_0}{2}}}{\Gamma\left(\frac{\nu_0}{2}\right)} \times \left(\frac{2}{T+\lambda_0}\right)^{\frac{n+\nu_0}{2}} \times \Gamma\left(\frac{n+\nu_0}{2}\right). \end{aligned}$$

Do it likewise, we rewrite the equation of the numerator and have

$$D_5 = -(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} P^{\frac{n+\nu_0}{2}-1} \times \exp\left[-\frac{T+\lambda_0}{2}P\right].$$

Therefore, the equation of the posterior distribution of $\sigma^2+\sigma_0^2$ which has the prior distribution of inverse gamma,

$$\begin{aligned}\frac{D_5}{D_6} &= \frac{-(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} \times P^{\frac{n+\nu_0}{2}-1} \times \exp\left[-\frac{T+\lambda_0}{2}P\right]}{-(2\pi)^{-\frac{n}{2}} \times \frac{(\frac{\lambda_0}{2})^{\frac{\nu_0}{2}}}{\Gamma(\frac{\nu_0}{2})} \times (\frac{2}{T+\lambda_0})^{\frac{n+\nu_0}{2}} \times \Gamma(\frac{n+\nu_0}{2})} \\ &= \left(\frac{T+\lambda_0}{2}\right)^{\frac{n+\nu_0}{2}} \times \frac{1}{\Gamma(\frac{n+\nu_0}{2})} \times P^{\frac{n+\nu_0}{2}-1} \times \exp\left[-\frac{T+\lambda_0}{2}P\right].\end{aligned}$$

Setting

$$\begin{aligned} s' &= \frac{n+\nu_0}{2}, \\ u' &= \frac{T+\lambda_0}{2}, \end{aligned}$$

then we have

$$\begin{aligned}\frac{D_5}{D_6} &= \frac{u'^{s'}}{\Gamma(s')} \times P^{s'-1} \times e^{u'P} \\ &= \frac{u'^{s'}}{\Gamma(s')} \times (\sigma^2+\sigma_0^2)^{-(s'+1)} \times e^{-\frac{u'}{\sigma^2+\sigma_0^2}}.\end{aligned}$$

This is the density of inverse gamma distribution $Ga^{-1}(s', u')$, and we find that the posterior distribution of $\sigma^2+\sigma_0^2$ which has the prior distribution of inverse gamma is

$$\sigma^2+\sigma_0^2|\boldsymbol{y}_n \sim Ga^{-1}\left(\frac{n+\nu_0}{2}, \frac{\sum_{t=1}^n (y_t-\psi_t)^2+\lambda_0}{2}\right).$$

Consequently, we find that the posterior expectation of $\sigma^2+\sigma_0^2$ in (20) is

$$E[\sigma^2+\sigma_0^2|\boldsymbol{y}_n] = \frac{\frac{\sum_{t=1}^n (y_t-\psi_t)^2+\lambda_0}{2}}{\frac{n+\nu_0}{2}-1}.$$

A.2 Proof of Proposition 3.5 The marginal likelihood $l(c, e^{-pr_t})$ of hierarchical Bayes model II is derived as follows:

$$(31)\quad l(c, e^{-pr_t}) = \prod_{t=1}^n \int_0^\infty \int_{-\infty}^\infty f(y_t|\mu_t, \sigma^2) g_2(\mu_t|\varphi_t, \sigma_0^2) h_3(\sigma^2+\sigma_0^2|c, e^{-pr_t}) d\mu_t d(\sigma^2+\sigma_0^2).$$

Firstly, we consider the integral in equation (31).

$$\begin{aligned}\int_0^\infty \int_{-\infty}^\infty & f(y_t|\mu_t, \sigma^2) g_2(\mu_t|\varphi_t, \sigma_0^2) h_3(\sigma^2+\sigma_0^2|c, e^{-pr_t}) d\mu_t d(\sigma^2+\sigma_0^2) \\ = & \int_0^\infty \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_t-\mu_t)^2}{2\sigma^2}\right] \\ & \times \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu_t-\varphi_t)^2}{2\sigma_0^2}\right] \\ & \times \frac{(e^{-pr_t})^c}{\Gamma(c)} (\sigma^2+\sigma_0^2)^{-(c+1)} \exp\left[-\frac{e^{-pr_t}}{\sigma^2+\sigma_0^2}\right] d\mu_t d(\sigma^2+\sigma_0^2) = D_7.\end{aligned}$$

In proposition 3.1, we found that

$$\begin{aligned} D_2 &= \int_{-\infty}^{\infty} f(y_t|\mu_t,\sigma^2) g_2(\mu_t|\varphi_t,\sigma_0^2) d\mu_t \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(y_t-\mu_t)^2}{2\sigma^2}\right] \times \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{(\mu_t-\varphi_t)^2}{2\sigma_0^2}\right] d\mu_t \\ &= \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}(y_t-\varphi_t)^2\right], \end{aligned}$$

substituting this into D_7 and we have

$$\begin{aligned} D_7 &= \int_0^{\infty} \frac{1}{\sqrt{2\pi(\sigma^2+\sigma_0^2)}} \times \exp\left[-\frac{1}{2(\sigma^2+\sigma_0^2)}(y_t-\varphi_t)^2\right] \\ &\qquad \times \frac{(e^{-pr_t})^c}{\Gamma(c)} (\sigma^2+\sigma_0^2)^{-(c+1)} \exp\left[-\frac{e^{-pr_t}}{\sigma^2+\sigma_0^2}\right] d(\sigma^2+\sigma_0^2) \\ &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \int_0^{\infty} (\sigma^2+\sigma_0^2)^{-\frac{1}{2}} (\sigma^2+\sigma_0^2)^{-(c+1)} \\ &\qquad \times \exp\left[-\frac{(y_t-\varphi_t)^2}{2(\sigma^2+\sigma_0^2)} - \frac{e^{-pr_t}}{\sigma^2+\sigma_0^2}\right] d(\sigma^2+\sigma_0^2) \\ &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \int_0^{\infty} (\sigma^2+\sigma_0^2)^{-(c+\frac{3}{2})} \exp\left[-\frac{(y_t-\varphi_t)^2+2e^{-pr_t}}{2(\sigma^2+\sigma_0^2)}\right] d(\sigma^2+\sigma_0^2). \end{aligned}$$

Secondly, setting

$$\begin{aligned} \frac{1}{\sigma^2+\sigma_0^2} &= L, \quad \frac{-1}{(\sigma^2+\sigma_0^2)^2} d(\sigma^2+\sigma_0^2) = dL, \\ -\frac{1}{L^2} dL &= d(\sigma^2+\sigma_0^2), \\ (y_t-\varphi_t)^2 + 2e^{-pr_t} &= M, \end{aligned}$$

then

$$\begin{aligned} D_7 &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \int_{\infty}^{0} L^{(c+\frac{3}{2})} \exp\left[-\frac{ML}{2}\right] \times (-\frac{1}{L^2}) dL \\ &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \int_0^{\infty} L^{(c-\frac{1}{2})} \exp\left[-\frac{ML}{2}\right] dL. \end{aligned}$$

Futher we set

$$\frac{ML}{2} = N, \quad L = \frac{2N}{M}, \quad dL = \frac{2}{M} dN,$$

and transform the equation as follows:

$$\begin{aligned} D_7 &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \int_0^{\infty} \left(\frac{2N}{M}\right)^{(c-\frac{1}{2})} e^{-N} \left(\frac{2}{M}\right) dN \\ &= \frac{(e^{-pr_t})^c}{\Gamma(c)} (2\pi)^{-\frac{1}{2}} \left(\frac{2}{M}\right)^{(c+\frac{1}{2})} \int_0^{\infty} N^{(c+\frac{1}{2})-1} e^{-N} dN, \end{aligned}$$

the integral is gamma function and it is rewritten as

$$\int_0^\infty N^{(c+\frac{1}{2})-1}e^{-N}dN = \Gamma\left(c+\frac{1}{2}\right).$$

Substituting it into the integral of the marginal likelihood,

$$\begin{aligned}
D_7 &= \int_0^\infty \int_{-\infty}^\infty f(y_t|\mu_t,\sigma^2)g_2(\mu_t|\varphi_t,\sigma_0^2)h_3(\sigma^2+\sigma_0^2|c,e^{-pr_t})d\mu_t d(\sigma^2+\sigma_0^2) \\
&= \frac{(e^{-pr_t})^c}{\Gamma(c)}(2\pi)^{-\frac{1}{2}}\left(\frac{2}{M}\right)^{(c+\frac{1}{2})}\Gamma\left(c+\frac{1}{2}\right) \\
&= \frac{(e^{-pr_t})^c}{\Gamma(c)}(2\pi)^{-\frac{1}{2}}\left[\frac{2}{(y_t-\varphi_t)^2+2e^{-pr_t}}\right]^{(c+\frac{1}{2})}\Gamma\left(c+\frac{1}{2}\right) \\
&= \frac{\Gamma\left(c+\frac{1}{2}\right)}{\Gamma(c)}(2\pi)^{-\frac{1}{2}}(e^{-pr_t})^c\left[\frac{2}{(y_t-\varphi_t)^2+2e^{pr_t}}\right]^{(c+\frac{1}{2})},
\end{aligned}$$

then

$$\begin{aligned}
l(c,e^{-pr_t}) &= \prod_{t=1}^n \int_0^\infty \int_{-\infty}^\infty f(y_t|\mu_t,\sigma^2)g_2(\mu_t|a_2,b_2,\sigma_0^2)h_3(\sigma^2+\sigma_0^2|c,p)d\mu_t d(\sigma^2+\sigma_0^2) \\
&= \prod_{t=1}^n \frac{\Gamma\left(c+\frac{1}{2}\right)}{\Gamma(c)}(2\pi)^{-\frac{1}{2}}(e^{-pr_t})^c\left[\frac{2}{(y_t-\varphi_t)^2+2e^{-pr_t}}\right]^{(c+\frac{1}{2})} \\
&= \left[\frac{\Gamma\left(c+\frac{1}{2}\right)}{\Gamma(c)}\right]^n (2\pi)^{-\frac{n}{2}}\prod_{t=1}^n (e^{-pr_t})^c\left[\frac{2}{(y_t-\varphi_t)^2+2e^{-pr_t}}\right]^{(c+\frac{1}{2})}. \qquad (32)
\end{aligned}$$

We take a logarithm of (32) and have

$$\begin{aligned}
\log l(c,e^{-pr_t}) &= -\frac{n}{2}\log(2\pi) + n\log\frac{\Gamma(c+\frac{1}{2})}{\Gamma(c)} \\
&\quad + \sum_{t=1}^n \left\{c(-pr_t) + \left(c+\frac{1}{2}\right)\log\left[\frac{2}{(y_t-\varphi_t)^2+2e^{-pr_t}}\right]\right\}.
\end{aligned}$$

Acknowledgements. The author would like to thank professors Masatoshi Fujisaki, Kensaku Kikuta and Yoshimi Inouchi for their valuable advices.

Table 1: Estimation Result of Simulation data 1 ($k = 1.0, k_1 : k_2 = 7 : 3$)

	True values	mean reverting model	Bayes model I(h_1)	Bayes model I(h_2)	Bayes model II
α		1.2356			
β		-0.1658			
σ_r^2		0.292			
γ		-0.1295			
a_1	1.6		1.5485	1.5485	
b_1	-0.2		-0.2084	-0.2084	
$\sigma^2 + \sigma_0^2$			0.1753	0.1785	
a_2					1.5086
b_2					-0.2025
c					1.3773
p					0.3164
$E[\sigma^2 + \sigma_0^2]$					0.1750
EIC		1113.775	1105.908	1096.829	1111.197

Table 2: Estimation Result of Simulation data 2 ($\nu_0 = 3.0, \lambda_0 = 6.0, k_1 : k_2 = 7 : 3$)

	True values	mean reverting model	Bayes model I(h_1)	Bayes model I(h_2)	Bayes model II
α		0.5734			
β		-0.1304			
σ_r^2		1.6036			
γ		-0.6532			
a_1	0.6		0.6144	0.6144	
b_1	-0.15		-0.1386	-0.1386	
$\sigma^2 + \sigma_0^2$			0.2291	0.2322	
a_2					0.6211
b_2					-0.1404
c					3.8289
p					0.1061
$E[\sigma^2 + \sigma_0^2]$					0.2287
EIC		1430.378	1391.636	1364.476	1368.669

Table 3: Estimation Result of Simulation data 3 ($l = 10.0, k_3 : k_4 = 8 : 2$)

	True values	mean reverting model	Bayes model I(h_1)	Bayes model I(h_2)	Bayes model II
α		0.5577			
β		-0.1630			
σ_r^2		0.3825			
γ		-0.4638			
a_1			0.5583	0.5583	
b_1			-0.1607	-0.1607	
$\sigma^2 + \sigma_0^2$			0.1208	0.1243	
a_2	0.7				0.5640
b_2	-0.2				-0.1649
c					3.6597
p	0.1				0.3584
$E[\sigma^2 + \sigma_0^2]$					0.1206
EIC		774.511	736.478	725.966	723.830

Table 4: Estimation Result of Short-Term Interest Rate Data

	mean reverting model	Bayes model I(h_1)	Bayes model I(h_2)	Bayes model II
α	0.1925			
β	-0.0101			
σ_r^2	0.0510			
γ	0.7334			
a_1		0.1490	0.1490	
b_1		-0.0077	-0.0077	
$\sigma^2 + \sigma_0^2$		6.3939	6.2902	
a_2				0.1039
b_2				-0.0072
c				1.4542
p				-0.0522
$E[\sigma^2 + \sigma_0^2]$				6.3729
EIC	2883.2780	1975.1840	1970.6630	1785.5820

References

[1] Ahangarani,P.M.(2005): An Empirical Estimation and Model Selection of the Short-Term Interest Rates, Working Paper, Presented in 2005 Winter Simulation Conference.

[2] Chan,K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders(1992): An Empirical Comparison of Alternative Models of the Short-Term Interest Rate, *Journal of Finance*, **47**, 1209-1227.

[3] Cox,J.C., J.E. Ingersoll and S.A. Ross(1985): A Theory of the Term Structure of Interest Rates, *Econometrica*,**53**,385-407.

[4] Gray,P.(2005): Bayesian Estimation of the Short-Rate Models, *Australian Journal of Management*, **30**, 1-22.

[5] Hong,Y. and H. Lin(2006): Nonparametric Specification Tests of Discrete Time Spot Interest Rate Models in China, Wise Working Paper Series, WISEWP0609.

[6] Ishiguro,M., Y. Sakamoto and G. Kitagawa(1997) : Bootstrapping Log Likelihood and EIC, an Extension of AIC, *Annals of the Institute of Statistical Mathematics*, **49**, 411-434.

[7] Jones,C.(2003): Nonlinear Mean Reversion in the Short-Term Interest Rate, *The Review of Financial Studies*, **16**, 793-843.

[8] Kawada,C.(2007): Information Criterion Statistical Analysis for Interest Rate Models Using GMM, THE SEIRYODAI RONSYU,*Journal of University of Hyogo*, **40-2**, 1-17. (in Japanese)

[9] Kawada,C.(2008): Information Criterion Statistical Analysis of Interest Data Based on Hierarchical Bayesian Mean Reverting Models, THE SEIRYODAI RONSYU,*Journal of University of Hyogo*, **40-3**, 43-58. (in Japanese)

[10] Sanford,A.D. and G.M. Martin(2006): A Bayesian Comparison of Several Continuous Time Models of the Australian Short Rate, *Accounting and Finance*, **46**, 309-326.

[11] Vasicek,O.(1977): An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, **5**, 177-188.

Graduate School of Business Administration, University of Hyogo,
8-2-1, Gakuen-Nishi-Machi, Nishi-ku,Kobe, 651-2197, JAPAN