

Shen, Xiangjin; Tsurumi, Hiroki

**Working Paper**

## Comparison of Bayesian model selection criteria and conditional Kolmogorov test as applied to spot asset pricing models

Working Paper, No. 2011-26

**Provided in Cooperation with:**

Department of Economics, Rutgers University

*Suggested Citation:* Shen, Xiangjin; Tsurumi, Hiroki (2011) : Comparison of Bayesian model selection criteria and conditional Kolmogorov test as applied to spot asset pricing models, Working Paper, No. 2011-26, Rutgers University, Department of Economics, New Brunswick, NJ

This Version is available at:

<https://hdl.handle.net/10419/59503>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Comparison of Bayesian Model Selection Criteria and Conditional Kolmogorov Test as Applied to Spot Asset Pricing Models

Xiangjin Shen and Hiroki Tsurumi <sup>1</sup>  
Rutgers University

June 2, 2011

## Abstract

We compare Bayesian and sample theory model specification criteria. For the Bayesian criteria we use the deviance information criterion and the cumulative density of the mean squared errors of forecast. For the sample theory criterion we use the conditional Kolmogorov test. We use Markov chain Monte Carlo methods to obtain the Bayesian criteria and bootstrap sampling to obtain the conditional Kolmogorov test. Two non-nested models we consider are the CIR and Vasicek models for spot asset prices. Monte Carlo experiments show that the DIC performs better than the cumulative density of the mean squared errors of forecast and the CKT. According to the DIC and the mean squared errors of forecast, the CIR model explains the daily data on uncollateralized Japanese call rate from January 1 1990 to April 18 1996; but according to the CKT, neither the CIR nor Vasicek models explains the daily data.

**Keywords:** Deviance information criterion, Cumulative density of the mean squared errors of forecast, Markov chain Monte Carlo algorithms, Block bootstrap, Generalized methods of moments, Conditional Kolmogorov test, CIR and Vasicek models

---

<sup>1</sup>Address correspondence to Xiangjin Shen and Hiroki Tsurumi, Department of Economics, Rutgers University, NJ 08854, USA. Email: xshen@econ.rutgers.edu and tsurumi@econ.rutgers.edu. We thank an anonymous referee for incisive comments. Also we thank Professor Norm Swanson, Geetesh Bhardwaj, and Cai Lili for their discussions on the conditional Kolmogorov test. And we thank Dr. Junnosuke Shino from Bank of Japan about the Japanese call rate.

# 1 Introduction

Efforts to find an appropriate model with acceptable explanatory and predictive power have led to many papers on model specification tests and model selection criteria. In this paper we compare Bayesian and sample theory model specification criteria to choose an appropriate model. Although the model specification criteria we consider can be applied to any non-nested models, we focus our attention on two asset pricing models: the Cox-Ingersoll-Ross model or the CIR model (1995) and the Vasicek model (1977).

Within Bayesian model selection criteria, Bayes factors and Bayesian modifications of the Akaike information criterion (BAIC) have been frequently used. In this paper we use the deviance information criterion (DIC) of Spiegelhalter *et.al.* (2002) and the cumulative density function (cdf) of the mean squared errors of forecast (MSEF), since these model selection criteria are easily obtained by using Markov chain Monte Carlo (MCMC) algorithms.

As a sample theory model specification test, we use the conditional Kolmogorov test (CKT). The CKT is a hypothesis test whereas the Bayesian criteria are to choose the model that explains the data best. Both the CKT and Bayesian criteria rely on random number generation: in the case of the CKT bootstrap methods are used, and in the case of the Bayesian criteria Markov chain Monte Carlo algorithms are used.

The organization of the paper is as follows. In Section 2.1 we present the DIC and the cumulative density (cdf) of the mean squared errors of the forecast (MSEF). In Section 2.2 the CKT is discussed. In Section 3 using a simulated data we present the DIC, and we explain how to use the cdf of the MSEF as a model selection criterion. Also, we explain the CKT that is obtained by the in-sample prediction. We demonstrate how the  $\alpha\%$  critical value (CV) is generated by bootstrap methods. In Section 4, we conduct Monte Carlo experiments to see the performances of the Bayesian criteria and of the CKT. In Section 5, we apply the Bayesian criteria and CKT to the real data: the daily data on uncollateralized Japanese call rate during the period between January 1 1990 and April 18 1996. Concluding remarks are given in Section 6.

## 2 Bayesian Model Choice Criteria and Conditional Kolmogorov Test of Model Selection

### 2.1 Bayesian Model Choice Criteria

The spot asset price model, discretized by the Euler-Maruyama scheme, is

$$\begin{aligned} r_t &= \kappa \theta + (1 - \kappa)r_{t-1} + \sigma_t u_t \\ &= \alpha + \beta r_{t-1} + \sigma_t u_t, \quad u_t \sim \text{N}(0, 1), \end{aligned} \quad (1)$$

where  $r_t$  is the spot rate;  $\alpha = \kappa \theta$ , and  $\beta = 1 - \kappa$ . If we set  $\sigma_t = \sqrt{r_{t-1}} \sigma$  we have the CIR model. If we set  $\sigma_t = \sigma$  we have the Vasicek model. In both models, there are three unknown parameters:  $\alpha$ ,  $\beta$ , and  $\sigma$ .

Let us derive the posterior probability density function (pdf) of the parameters assuming that the prior pdf is given by

$$p(\alpha, \beta, \sigma) \propto \sigma^{-1}. \quad (2)$$

The posterior pdf's of the CIR and Vasicek models are given by

$$p(\alpha, \beta, \sigma | \text{data}) \propto \sigma^{-(n+1)} \exp \left\{ -\frac{1}{2\sigma^2} \left[ \nu s^2 + (\gamma - \hat{\gamma})' X' D^{-1} X (\gamma - \hat{\gamma}) \right] \right\}, \quad (3)$$

where

$$\gamma = (\alpha, \beta)', \quad \hat{\gamma} = (X' D^{-1} X)^{-1} X' D^{-1} y, \quad \nu s^2 = (y - X \hat{\gamma})' X' D^{-1} X (y - X \hat{\gamma})$$

$$y = \begin{bmatrix} r_1 \\ \vdots \\ r_T \end{bmatrix}, \quad X = \begin{bmatrix} 1 & r_0 \\ \vdots & \vdots \\ 1 & r_{T-1} \end{bmatrix},$$

and  $D$  is given by

$$D = \begin{cases} I_T & \text{for the Vasicek model} \\ \text{Diag}(r_0, \dots, r_{T-1}) & \text{for the CIR model.} \end{cases}$$

Let us use the deviance information criterion (DIC) (Spiegelhalter *et.al.* (2002)) and the distribution of the mean squared errors of forecast (MSEF) as the criteria to choose between equation (1) and (2).

The DIC is given by

$$\text{DIC} = \bar{D} + p_D, \quad (4)$$

where  $\theta = (\gamma', \sigma)'$ ,  $D(\bar{\theta}) = -2 \ln L(y|\bar{\theta})$ , and  $L(y|\bar{\theta})$  is the likelihood function evaluated at the posterior mean of  $\theta$ .  $\bar{D}$  is given by

$$\bar{D} = -2 \int \ln(L(y|\theta) p(\theta|\text{data})) d\theta .$$

The parameter  $p_D$  measures the model complexity and it is given by  $p_D = \bar{D} - D(\bar{\theta})$ . In the MCMC algorithm  $\bar{D}$  is evaluated by

$$-2 \left[ \frac{1}{N} \sum_{i=1}^N \ln L(y|\theta^{(i)}) \right] ,$$

where  $\theta^{(i)}$  is the  $i$ -th MCMC draw of  $\theta$ , and  $N$  is the number of MCMC draws. We choose the model with the smaller DIC. DIC is a widely used statistic for comparing models, and it is a built-in procedure in software packages such as WinBUGS. We draw  $\theta^{(i)}$  by the Gibbs sampler algorithm as well as by the Metropolis-Hastings algorithm. These two MCMC algorithms yield quite similar draws of  $\theta^{(i)}$ .

The distribution of the mean squared errors of forecast is obtained by the MCMC algorithms as follows:

Let  $\tilde{y} = (\tilde{y}_{T+1}, \tilde{y}_{T+2}, \dots, \tilde{y}_{T+m})$ . The joint pdf of  $\tilde{y}$  and  $\theta$  is given by

$$h(\tilde{y}, \theta|\text{data}) = f(\tilde{y}|\theta)p(\theta|\text{data}), \quad (5)$$

where  $f(\tilde{y}|\theta)$  is the pdf of  $\tilde{y}$  given  $\theta$  and  $\text{data}=(y_0, y_1, \dots, y_T)$ .

West (2006) discuss three ways to generate the sequences of regression estimates necessary to make predictions: recursive scheme, rolling scheme, and fixed schemes. We use the fixed scheme that is to draw  $\tilde{y}^{(i)}$ , where  $\tilde{y}^{(i)}$  is the  $i$ -th draw of  $\tilde{y}$  given the  $i$ -th draw of  $\theta$ ,  $\theta^{(i)}$ , using data from 1 to  $T$ .

Since the CIR and Vasicek models follow AR(1) processes, we draw  $\tilde{y}_{T+1}, \tilde{y}_{T+2}, \dots, \tilde{y}_{T+m}$  sequentially:

$$\begin{aligned} & h(\tilde{y}_{T+1}|y_T, \theta^{(i)}) \\ & h(\tilde{y}_{T+2}|\tilde{y}_{T+1}, \theta^{(i)}) \\ & \quad \vdots \\ & h(\tilde{y}_{T+m}|\tilde{y}_{T+m-1}, \theta^{(i)}) . \end{aligned} \quad (6)$$

After drawing  $\tilde{y}^{(i)}$  we transform it into the mean squared errors of forecast by

$$MSEF^{(i)} = \frac{1}{m} \left\{ \frac{1}{N} \sum_{j=1}^m \sum_{i=1}^N \left( y_{T+j} - \tilde{y}_{T+j}^{(i)} \right)^2 \right\}, \quad (7)$$

where  $y_{T+j}$  is the actual realized value at time  $T + j$ , and  $\tilde{y}_{T+j}^{(i)}$  is the  $i$ -th MCMC draw of the predicted value at time  $T + j$ . The mean squared errors of forecast, MSEF, is a summary statistic popular both in sample theory and Bayesian inference. By the MCMC algorithm we can obtain the distribution of this popular statistic, and plot its pdf and cumulative density function (cdf). In comparing two models, we may choose the model that has a pdf close to the origin or that has a dominating cdf.

## 2.2 Conditional Kolmogorov Test for Model Selection

The conditional Kolmogorov-Smirnov tests (CKT) have been used by Andrew (1997), Min and Hong (1997), Whang (2000), Horowitz (2003), Bai (2003), Scaillet (2005), Corradi and Swanson (2005, 2006), Bhardwaj, Corradi and Swanson (BCS) (2008), and Lee (2009), among others. BCS (2008) applied the CKT for model selections, and among the models they tested is the CIR model. Since BCS proposes an innovative CKT, let us follow their CKT. The hypotheses are

$$\begin{aligned} H_0 : \quad & F_\tau(u|r_t, \Theta) = F_\tau(u|r_t, \Theta_0) \\ H_1 : \quad & F_\tau(u|r_t, \Theta) \neq F_\tau(u|r_t, \Theta_0), \end{aligned} \quad (8)$$

where  $F_\tau(u|r_t, \Theta) = \Pr(r_{t+\tau}^\Theta \leq u | r_t^\Theta = r_t)$  is the cdf of the  $\tau$ -step ahead in-sample prediction  $r_{t+\tau}^\Theta$  given  $r_t^\Theta = r_t$ , and  $t = 1, 2, \dots, T - \tau$ . The term *in-sample prediction* is explained later.

Unlike the Bayesian inference that is parametric, the CKT is based on semi-parametric inference employing large sample (asymptotic) properties. First, the sample path,  $r_t^\Theta$  is obtained by the Milstein scheme (Seydel (2009)):

$$\begin{aligned} r_t^\Theta &= \kappa \theta + (1 - \kappa)r_{t-1}^\Theta + \sigma(\cdot)\epsilon_t \\ &\quad - \frac{1}{2}\sigma^2(\cdot) + \frac{1}{2}\sigma(\cdot)^2\epsilon_t^2 \end{aligned} \quad (9)$$

where  $\sigma(\cdot) = \sigma \sqrt{r_{t-1}^\Theta}$  if  $H_0$  is the CIR model and  $\sigma(\cdot) = \sigma$  if  $H_0$  is the Vasicek model.  $\epsilon_t$  is the standardized normal:  $\epsilon_t \sim N(0, 1)$ .

The test statistic  $V_T(\tau, \underline{u}, \bar{u})$  is computed by

$$V_T(\tau, \underline{u}, \bar{u}) = \sup_{v \in \mathcal{V}} |V_T(v|\tau, \underline{u}, \bar{u})|, \quad (10)$$

where  $V_T(v|\tau, \underline{u}, \bar{u})$  is

$$\begin{aligned} V_T(v|\tau, \underline{u}, \bar{u}) = & \\ & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left( \frac{1}{S} \sum_{s=1}^S \mathcal{I} \left\{ \underline{u} \leq r_{s,t+\tau}^{\hat{\Theta}} \leq \bar{u} \right\} - \mathcal{I}(\underline{u} \leq r_{t+\tau} \leq \bar{u}) \right) \mathcal{I}(r_t \leq v), \end{aligned} \quad (11)$$

and  $\underline{u}$ , and  $\bar{u}$  are the lower and upper bounds of the prediction interval. Since  $\underline{u}$  and  $\bar{u}$  are fixed, the prediction interval is a fixed constant. The random sample path  $r_{s,t+\tau}^{\hat{\Theta}}$  is obtained for given  $t$ , by equation (9) using a consistent estimate of  $\Theta$ ,  $\hat{\Theta}$ , under the null hypothesis and replacing  $r_t^\Theta$  with  $r_{s,t}^{\hat{\Theta}}$  for  $s = 1, 2, \dots, S$ ; where  $S$  is the number of simulated sample paths and  $\mathcal{I}(\cdot)$  is the indicator function. The value of  $S$  is chosen much larger than the sample size  $T$  and here set  $S = 5T$ . The random variable  $r_{s,t+\tau}^{\hat{\Theta}}$  is called the in-sample prediction because  $\hat{\Theta}$  is estimated by using all the sample observations from 1 to  $T$ :  $y_1, \dots, y_T$

The critical value for the in-sample prediction interval is computed by generating

$$V_T^*(\tau, \underline{u}, \bar{u}) = \sup_{v \in \mathcal{V}} |V_T^*(v|\tau, \underline{u}, \bar{u})|, \quad (12)$$

where  $V_T^*(v|\tau, \underline{u}, \bar{u})$  is

$$\begin{aligned} V_T^*(v|\tau, \underline{u}, \bar{u}) = & \\ & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left( \frac{1}{S} \sum_{s=1}^S \mathcal{I} \left\{ \underline{u} \leq r_{s,t+\tau}^{*\hat{\Theta}^*} \leq \bar{u} \right\} - \mathcal{I}(\underline{u} \leq r_{t+\tau}^* \leq \bar{u}) \right) \mathcal{I}(r_t^* \leq v) \\ & - V_T(v|\tau, \underline{u}, \bar{u}), \end{aligned} \quad (13)$$

The standard block-bootstrap method (Hall (1986), Hall and Horowitz (1996), Horowitz (2003)) is used to draw  $r_t^*$  from the data  $r_1, r_2, \dots, r_T$ , and  $\hat{\Theta}^*$  is estimated by the generalized methods of moments (GMM) or by the simulated generalized methods of moments (SGMM) given in Duffie and Singleton

(1993) using the bootstrap sample. The sample path,  $r_{s,t}^{*\hat{\Theta}^*}$ , is obtained using equation (9) under the null hypothesis. The bootstrap test statistic  $V_T^*(v|\tau, \underline{u}, \bar{u})$  of equation (13) is computed using the test statistic  $V_T(v|\tau, \underline{u}, \bar{u})$  of equation (11) and centering around  $V_T(v|\tau, \underline{u}, \bar{u})$ . The centering of the test bootstrap test statistic is suggested by Hansen (2005). From the empirical distribution of  $V_T^*(v|\tau, \underline{u}, \bar{u})$  the  $\alpha$ -% critical value is found.

There are three sets of moments to choose for the GMM: the unconditional, conditional, and simulated moments. For the CIR model the unconditional moments are derived by Jiang and Knight (2002):

$$g(r_t) = \begin{cases} \mathbb{E}(r_t) &= \theta \\ \mathbb{E}(r_t^2) &= \theta^2 + \frac{\sigma^2}{2\kappa} \theta \\ \mathbb{E}(r_t^3) &= \theta^3 + \frac{3\sigma^2}{2\kappa} \theta^2 + \frac{\sigma^4}{2\kappa^2} \theta. \end{cases} \quad (14)$$

The conditional moments are

$$g(r_t) = \begin{cases} \mathbb{E}(r_t|r_0) &= r_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \\ \mathbb{E}(r_t^2|r_0) &= (\mathbb{E}(r_t|r_0))^2 + \frac{\theta \sigma_{r_t}^2 (1 - e^{-\kappa t})^2}{2\kappa} + \frac{\sigma_{r_t}^2 r_0 (e^{-\kappa t} - e^{-2\kappa t})}{\kappa} \\ \mathbb{V}(r_t|r_0) &= \frac{\theta \sigma_{r_t}^2 (1 - e^{-\kappa t})^2}{2\kappa} + \frac{\sigma_{r_t}^2 r_0 (e^{-\kappa t} - e^{-2\kappa t})}{\kappa}. \end{cases} \quad (15)$$

The expressions for the unconditional and conditional moments for the Vasicek model are similarly obtained. The simulated moments for SGMM are simulated from sample path

$$g(r_t^*) = \begin{cases} \mathbb{E}r_t^* &= \frac{1}{N} \sum_{t=1}^N r_t^* \\ \mathbb{E}(r_t^{*2}) &= \frac{1}{N} \sum_{t=1}^N r_t^{*2} \\ \mathbb{E}(r_t^{*3}) &= \frac{1}{N} \sum_{t=1}^N r_t^{*3}, \end{cases} \quad (16)$$

where  $N$  is the bootstrap sample length.

There are two schemes to obtain the covariance matrix  $W_T$ . The first way is the heteroskedastic autocorrelation (HAC) covariance matrix:

$$W_T = \frac{1}{T} \sum_{v=-\ell_T}^{\ell_T} w_v \sum_{t=v+1+\ell_T}^{T-\ell_T} \left\{ \left( g(r_t) - \frac{1}{T} \sum_{t=1}^T g(r_t) \right)' \left( g(r_{t-v}) - \frac{1}{T} \sum_{t=1}^T g(r_t) \right) \right\}. \quad (17)$$



and  $\ell_T$  is the lag length often set at  $\ell_T = [T^{\frac{1}{6}}]$ , where  $[\cdot]$  denotes the integer value. The second scheme is the regular GMM covariance matrix (or the covariance without HAC):

$$W_T = \frac{1}{T} \sum_{t=1}^T \left\{ \left( g(r_t) - \frac{1}{T} \sum_{t=1}^T g(r_t) \right)' \left( g(r_t) - \frac{1}{T} \sum_{t=1}^T g(r_t) \right) \right\}. \quad (18)$$

In comparing the CIR and Vasicek models, the CKT is carried out as follows: Given the data  $(r_1, r_2, \dots, r_T)$  first put the CIR model as the null hypothesis  $H_0$ , and compute the test statistic  $V_T(\tau, \underline{u}, \bar{u})$ . Then we compare it with the  $\alpha$ -% critical value of the confidence interval,  $V_T^{*\alpha}(\tau, \underline{u}, \bar{u})$ . If  $V_T(\tau, \underline{u}, \bar{u}) < V_T^{*\alpha}(\tau, \underline{u}, \bar{u})$ , we accept the null hypothesis. Otherwise, reject the null hypothesis concluding that the data is not from the CIR model. Now put the Vasicek model in the null hypothesis  $H_0$ , and obtain  $V_T(\tau, \underline{u}, \bar{u})$  and  $V_T^{*\alpha}(\tau, \underline{u}, \bar{u})$ . If  $V_T(\tau, \underline{u}, \bar{u}) < V_T^{*\alpha}(\tau, \underline{u}, \bar{u})$ , we accept  $H_0$  that the data is from the Vasicek model. Otherwise, reject the null hypothesis.

### 3 Example with a Simulated Data

Before we conduct Monte Carlo experiments, let us make a simulated data analysis to illustrate the Bayesian model selection criteria as well as the conditional Kolmogorov test. We draw two samples of size  $T$ : one drawn from the Vasicek model of equation (1) and the other drawn from the CIR model of equation (2). We use the following parameter values:

$$\kappa = .3068$$

$$\theta = .0558$$

$$\sigma = .1180$$

$$r_0 = .0657$$

We generate 212 observations ( $T = 212$ ) and use the first 200 observations to estimate the parameters  $\kappa$ ,  $\theta$ , and  $\sigma$ . We use the additional 12 observations for the out-of-sample forecast. As the out-of-sample forecast periods,  $m$ , we choose  $m = 1, 2, 4, 12$ . If data is monthly, then we are examining one month, two months, 4 months and 12 months ahead predictions. If data is weekly, then we have one week, two

weeks, 4 weeks (one month), and 12 weeks (one quarter) ahead predictions. Let us call this the  $m$ -step ahead prediction.

Before we report the results of the model selection criteria, let us present the point estimates of the parameters. Table 1 gives the posterior means, while Table 2 presents the GMM estimates. As discussed in the previous section, for the GMM estimation, there are three ways to compute the moments (unconditional moments, conditional moments, and simulated moments), and two residuals (with and without HAC). Both the Bayesian and GMM point estimates of the parameters are reasonable.

*Table 1 and Table 2 Here*

### **Bayesian model selection criteria**

First let us discuss the MSEF's. Since we have the distributions of the MSEF's, we examine the probability density functions (pdf) and cumulative density functions (cdf) of the MSEF's. Figure 1 presents the pdf's and cdf's of {1, 2, 4, 12}-step ahead predictions for the data generated by the Vasicek model. All the pdf's of the MSEF's are skewed to the right. We see that the pdf's of the MSEF's of the Vasicek model lie close to zero while the pdf's of the MSEF's of the CIR model spread out to the right. The modes of the MSEF's of the Vasicek model are closer to zero than the modes of the MSEF's of the CIR model. The cdf's of the MSEF's of the Vasicek model dominate the cdf's of the MSEF's of the CIR model. Clearly we choose the Vasicek model over the CIR model for all of the {1, 2, 4, 12}-step ahead prediction. Although the pdf's and cdf's of the MSEF's for the sample generated by the CIR model are not presented here to save space, we obtain the similar results: the pdf's and cdf's of the CIR model dominate those of the Vasicek model.

*Figure 1 Here.*

In Table 3 we present the DIC, the means, medians, and modes of the MSEF's for the data generated by the CIR model as well as for the data generated by the Vasicek model. As discussed in the previous section, we choose the model with a smaller DIC. For the data generated by the CIR model, the DIC of the CIR model is  $-920.7102$  that is smaller than the DIC of the Vasicek model that is  $-876.5910$ . Hence we choose the CIR model. Similarly, when the data is generated by the Vasicek model, the DIC of the Vasicek model is  $-384.6277$  which is smaller than  $-315.5180$ , the DIC of the CIR model. To save space, we only present the means, medians, and modes of the MSEF's.

*Table 3 Here.*

### Conditional Kolmogorov test

Let us turn to the CKT proposed by BCS (2008). In the previous section we presented  $V_T(\tau, \underline{u}, \bar{u})$  and  $V_T^*(\tau, \underline{u}, \bar{u})$ :  $V_T(\tau, \underline{u}, \bar{u})$  in equation (11) is for computing the test statistic and  $V_T^*(\tau, \underline{u}, \bar{u})$  in equation (13) is for obtaining the critical values. Both  $V_T(\cdot)$  and  $V_T^*(\cdot)$  are computed given  $\tau$ -ahead prediction periods and the lower and upper bounds of the confidence intervals,  $\underline{u}$  and  $\bar{u}$ .

The CKT is formulated as hypothesis test. Since the CIR and Vasicek models are non-nested, the CIR and Vasicek models are put as the null hypothesis one at a time:

$$H_0 : \text{CIR model} \quad \text{vs} \quad H_1 : \text{Not CIR model}$$

and

$$H_0 : \text{Vasicek model} \quad \text{vs} \quad H_1 : \text{Not Vasicek model}$$

Given  $\tau$  and the predictive interval, we compute the test statistic  $V_T$  and we find the  $\alpha\%$  critical value  $V_T^{*\alpha}$  from the bootstrap distribution of  $V_T^*$ . We accept the null hypothesis of  $H_0$ : CIR model if  $V_T < V_T^{*\alpha}$ . Next we switch  $H_0$  to the Vasicek model and compute  $V_T$  and  $V_T^{*\alpha}$  and accept Vasicek model if  $V_T < V_T^{*\alpha}$ . As we observe in Pesaran (1974), Pesaran and Hall (1988) and Davidson and MacKinnon (1988), hypothesis test used for model selection sometimes yields the result that none of the models considered is selected.

BCS (2008) consider two prediction intervals:  $(\underline{u}, \bar{u}) = (\bar{r} \pm .5\sigma_R)$  and  $(\underline{u}, \bar{u}) = (\bar{r} \pm \sigma_R)$ , where  $\bar{r}$  is the sample mean of  $r_t$ , and  $\sigma_R$  is the sample standard deviation of  $r_t$ . As for  $\alpha\%$  BCS (2008) chooses 5%, 10%, 15% and 20%. Hence, given the null hypothesis  $H_0$ , there are eight  $V_T(\tau, \underline{u}, \bar{u})$ 's, and each  $V_T(\tau, \underline{u}, \bar{u})$  is compared against the  $\alpha\%$  critical value  $V_T^{*\alpha}(\tau, \underline{u}, \bar{u})$ . In summary, given  $H_0$ , we have 32 combinations of 4  $\tau$ 's, 2 confidence intervals, and 4  $\alpha\%$  critical values ( $V_T^{*\alpha}$ 's) as shown in Table 4 that gives the CKT results for data generated by the CIR model. In the table  $V_T^{*\alpha}$  is denoted as  $\alpha\%CV$ .

*Table 4 Here.*

The first half of Table 4 is for the CIR model as  $H_0$ . The only cases where the test statistic  $V_T(\tau, \underline{u}, \bar{u})$  are less than the critical values are for the 12-step ahead in-sample prediction ( $\tau = 12$ ) by the CIR model.

All other cases we see that  $V_T((\tau, \underline{u}, \bar{u})$ 's are greater than the critical values. Hence, eight out of thirty two cases we accept the null hypothesis that the CIR model is the correct model. Now switching the  $H_0$  to the Vasicek model, we see that all the 32 cases  $H_0$  is rejected and  $H_1$  is accepted. Since the alternative hypothesis  $H_1$  in equation (8) is presented as “not  $H_0$ ,” the acceptance of  $H_1$  does not mean that data supports the CIR model. It is up to us, the decision makers, to choose between these two models. We may say, for example, that since the CIR model is chosen 8 out of 32 cases while the Vasicek model is chosen 0 out of 32 cases we choose the CIR model over the Vasicek model.

Table 5 presents the results for the data generated by the Vasicek model. When  $H_0$  is the CIR model, all 32 cases  $V_T(\tau, \underline{u}, \bar{u})$ 's are greater than the critical values, rejecting  $H_0$ . If we switch  $H_0$  to the Vasicek model, all 32  $V_T(\tau, \underline{u}, \bar{u})$ 's are greater than the critical values, rejecting that the data is generated by the Vasicek model. Accordingly, we conclude that the CKT selects neither the CIR nor Vasicek model.

*Table 5 Here.*

## 4 Monte Carlo Experiments

Let us make Monte Carlo experiments assuming the error term  $u_t$  is non-normal. We make 500 sample replications. The parameters and the sample size are the same as those for the one sample replication. The stylized fact about financial data is that their distributions are leptokurtic and slightly skewed. If we make Monte Carlo experiments drawing  $u_t$  from a leptokurtic and slightly skewed distribution, the CKT test may perform better than the Bayesian criteria, since the CKT test is derived without assuming the distribution of  $u_t$ .

Let us assume that  $u_t$  is generated from a leptokurtic and slightly skewed distribution:

$$N(0, 1) \times \text{IVG}(\mu, \eta) \tag{19}$$

where  $\text{IVG}(\mu, \eta)$  is the inverse Gaussian distribution with parameters  $\mu$  and  $\eta$ . The algorithm for generating random numbers from  $\text{IVG}(\mu, \eta)$  is given in Michael *et.al.* (1976) and Devroye (1986). We set  $\mu = 1.8$  and  $\eta = .8$ . The skewness is  $-.687$  and kurtosis is  $24.6$ . Compared to the  $N(0,1)$  density the distribution is sharply peaked with fat tails.

The results of the Monte Carlo experiments with the non-normal distribution are given in Table 6 for the Bayesian model selection criteria and in Table 7 for the CKT. For the Bayesian model selection criteria the DIC criterion chooses the correct models: 92% (98.5%) of the times when the CIR (Vasicek) is the correct model. The medians of the {1, 2, 4, 12}-step ahead MSEF's choose the CIR model 58.0% to 64.0% of times when data are generated by the CIR model. When data are generated by the Vasicek model it is chosen from 57.0% to 71.5% of times.

The results for the CKT are given in Table 7. The acceptance rates of both the CIR model and the Vasicek model are very low, and this is due to the fact that the CKT test is sensitive to the non-normal error term. The reason is due to the way the test statistic  $V_T(v|\tau, \underline{u}, \bar{u})$  in equation (11) and the critical value  $V_T^*(\tau, \underline{u}, \bar{u})$  in equation (13) are computed using sample paths simulated by the Milstein scheme with the normal error  $\epsilon_t$  as given in equation (9).

*Tables 6 and 7 Here.*

## 5 Application to Japanese Call Rates

Let us apply the Bayesian model selection criteria and the conditional Kolmogorov test to the daily uncollateralized call rate of Japan. The data is daily averages from January 1 1990 to April 18 1996 for the total of 2,300 observations. This period corresponds to the beginning of the bursting of the Japanese bubble economy to the early period of the zero interest rate.

The call rate kept increasing even after the Japanese bubble bursted in early January 1990. The call rate peaked on March 14 1991 at 8.56%, and one year after the bubble bursted the call rate was still around 5.8%. It took four and a half years more for the call rate to get down to .5% on September 15 1995 which is regarded as the beginning of the zero-interest rate period. Learning from the Japanese experience, it took only one year for the Federal Reserve Board to cut the Federal Funds rate from 5.0% in late December 2007 to less than 1% in late December 2009.

The sample mean, median, skewness and kurtosis of the daily change in the Japanese call rate are

mean	median	skewness	kurtosis
-.0027	0.0	-1.064	190.61

and we see that the Japanese call rates exhibit the typical stylized fact of the return of the financial rates.

We applied the Bayesian model selection criteria and the conditional Kolmogorov test to see which model, the CIR or Vasicek, explains the Japanese call rate. Table 8 presents the DIC and the modes of the MSEF's and Table 9 presents the CKT test.

*Tables 8, 9 and Figure 2 Here.*

From Table 8 we see that the DIC and the modes of the MSEF's all choose the CIR model over the Vasicek model. The cdf's of the MSEF's of the CIR model uniformly dominate those of the Vasicek model as shown in Figure 2. From Table 9 we see that the CKT's,  $V_T$ , are all greater than all the critical values regardless of putting the CIR or Vasicek models in the null hypothesis. Accordingly we conclude the Japanese call rate follows neither the CIR nor the Vasicek model.

## 6 Concluding Remarks

Using two non-nested models of spot asset pricing models, the CIR and Vasicek models, we compared the Bayesian model choice criteria to the conditional Kolmogorov test (CKT). The Bayesian criteria are designed to choose the model that explains the data best, while the CKT is a hypothesis test to accept the null hypothesis that one model is true at a pre-specified significance level. We used two Bayesian criteria: the deviance information criterion (DIC) of Spiegelhalter *et.al.* and the cumulative density function (cdf) of the mean squared errors of forecast (MSEF).

Using a simulated data we demonstrated how to use the cdf of the MSEF as a model selection criterion: by plotting the cdf's of two MSEF's of the non-nested models we choose the model with the dominating cdf. If the two cdf's cross, we may choose the model whose median MSEF is smaller. The CKT is to obtain the in-sample prediction and compare it with the  $\alpha$  % critical value that is generated by bootstrapping. We conducted Monte Carlo experiments to see how the Bayesian model selection criteria and the CKT perform. Among the DIC, the cdf of the MSEF, and the CKT the clear winner is the DIC.

Then we applied the DIC, the cdf of the MSEF, and the CKT to the daily data of the uncollateralized Japanese call rate from January 1 1990 to April 18 1996 for the total of 2,300 observations. The sample period begins at the inception of the bursting of the Japanese bubble and ends at the beginning of the zero interest rate period. We find that the CIR model explains the data better than the Vasicek model according to the DIC and the cdf's of the MSEF. According to the CKT the data support neither the CIR nor Vasicek model.

Both the Bayesian model selection criteria and the CKT are based on random number generation. The random number generation for the Bayesian model selection criteria is used to obtain the posterior distributions of the log of the likelihood function to compute the DIC and of the mean squared errors of the forecast. In the Bayesian inference, the parameters of the model are generated while the data are treated as fixed. The random number generation for the CKT is used to generate bootstrap samples from the data. Given the bootstrap samples the distributions of the CKT are generated.

## References

- Andrews, D.W.K. (1977). A conditional Kolmogorov test. *Econometrica*, 65, 1097–1128.
- Bai, J. (2003). Testing parametric conditional distributions of dynamic models. **Review of Economics and Statistics**, 85, 531–549.
- Bhardwaj, G., V. Corradi and N.R. Swanson (2008). A simulation based specification test for diffusion processes. *Journal of Business and Economic Statistics*, 26, 176–193.
- Corradi, V. and N.R. Swanson (2005). A bootstrap specification test for diffusion processes. *Journal of Econometrics*, 124, 117–148.
- Corradi, V. and N.R. Swanson (2006). Bootstrap conditional distribution tests in the presence of dynamic misspecification. *Journal of Econometrics*, 124, 117–148.
- Cox, J.C., J.E. Ingersoll and S.A. Ross (1985). A theory of the term structure of interest rates. *Econometrica*, 53, 385–407.
- Devroye, L. (1986). *Non-uniform random variate generation*, Springer-Verlag, New York.
- Davidson, R. and J. MacKinnon (1981). Several tests for model selection in the presence of alternative hypothesis. *Econometrica*, 49, 781–793.
- Duffie, D. and K. Singleton. (1993). Simulated moment estimation of Markov models of asset prices. *Econometrica* 61, 929–952.
- Hall, P. (1986). On the bootstrap and confidence intervals. *Annals of Statistics* , 14, 1431-1452.
- Hall, P. and Horowitz. J.L. (1996). Bootstrap critical values for tests based on generalized-method-of-moments estimators. *Econometrica*, 64, 891-916.
- Hansen, P.R. (2005). A test for superior predictive ability. mimeograph

- Hong, Y.M., H. Li and F. Zhao (2004). Out of sample performance of spot interest rate models. *Journal of Business and Economic Statistics*, 22, 457–473.
- Horowitz, J. L. (2003). The bootstrap in econometrics. *Statistical Science*, 18 (2), Silver anniversary of the bootstrap, 211–218.
- Jian, G. and J. Knight (2002). Estimation of continuous-time processes via the empirical characteristic function. *Journal of Business & Economic Statistics*, 20, 198–212.
- Michael, J.R., W.R. Schucany and R.W. Haas (1976). Generating random variables using transformation with multiple roots. *American Statistician*, 30, 88–90.
- Min, C., Z.A. Hong (1997). A Kolmogorov-Smirnov type test for conditional heteroskedasticity in time series. *Statistics and Probability Letters*, 33, 321–331.
- Pesaran, H. (1974). On the general problem of model selection. *Review of Economic Studies*, 41, 153–171.
- Pesaran, M. and A. Hall (1988). Tests of non-nested linear regression models subject to linear constraints. *Economic Letters*, 27, 341–348.
- Scaillet, O. (2005). A Kolmogorov-Smirnov type test for positive quadrant dependence. *Canadian Journal of Statistics*, 33, 415–427.
- Seydel R. (2009). *Tools for computational finance*, Springer
- Spiegelhalter, D.J., N.G. Best, B.P. Carlin and A. van der Linde (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B* 4, 583–639.
- Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5, 177–188.
- West, K.D. (2006). Forecast evaluation. *Handbook of Economic Forecasting*, 2006, Vol. 1, Chapter 3.
- Whang, Y.J. (2000). Corresponding author contact information. Consistent bootstrap tests of parametric regression functions. *Journal of Econometrics*, 98, 27–46.



Table 1: Bayesian Posterior Means of the Parameters

Data Generated by CIR Model			
	$\kappa$	$\theta$	$\sigma$
CIR Model	.2866	.0591	.1084
Vasicek Model	.3755	.0586	.0268
Data Generated by Vasicek Model			
	$\kappa$	$\theta$	$\sigma$
CIR Model	.6663	.1598	.3107
Vasicek Model	.7335	.1579	.0918

Notes: The true parameters are  $\kappa = .3068$ ,  $\theta = .0558$ ,  $\sigma = .1180$ .  
Sample size is 200.  
Metropolis-Hastings algorithm is used.

Table 2: Point Estimates of Parameters by GMM and SGMM

		GMM with unconditional moments			GMM with conditional moments			SGMM		
		$\kappa$	$\theta$	$\sigma$	$\kappa$	$\theta$	$\sigma$	$\kappa$	$\theta$	$\sigma$
Estimates from Data Generated by CIR Model										
CIR Model	HAC	.770	.059	.170	.830	.0775	.102	.538	.111	.248
	w/o HAC	.370	.055	.120	.512	.075	.111	.572	.088	.214
Vasicek Model	HAC	.550	.055	.03	.690	.0055	.050	.608	.071	.228
	w/o HAC	.570	.055	.03	.650	.045	.030	.514	.090	.209
Estimates from Data Generated by Vasicek Model										
CIR Model	HAC	.150	.150	.120	.690	.195	.225	.692	.160	.289
	w/o HAC	.390	.155	.190	.552	.175	.337	.442	.105	.222
Vasicek Model	HAC	.590	.150	.090	.710	.005	.050	.614	.083	.222
	w/o HAC	.450	.155	.090	.590	.005	.030	.5470	.095	.184

Notes: True parameter values are  $\kappa = .3068$ ,  $\theta = .0558$ ,  $\sigma = .0657$   
Sample size is 200.  
w/o HAC = without HAC.

Table 3: Bayesian Model Selection Criteria: Example with Simulated Data

Data Generated by CIR Model			
		CIR Model	Vasicek Model
DIC		-920.7102	-876.5910
MSEF			
Mean	1-step-ahead	.00021	.00051
	2-step-ahead	.00030	.00061
	4-step-ahead	.00051	.00074
	12-step-ahead	.00113	.00107
Median	1-step-ahead	.000107	.00025
	2-step-ahead	.000200	.00039
	4-step-ahead	.000382	.000549
	12-step-ahead	.00080	.00090
Mode	1-step-ahead	.000036	.000105
	2-step-ahead	.000087	.000227
	4-step-ahead	.000236	.000386
	12-step-ahead	.00069	.00075
Data Generated by Vasicek Model			
		CIR Model	Vasicek Model
DIC		-315.5180	-384.6227
Mean	1-step-ahead	.02475	.00866
	2-step-ahead	.02637	.016334
	4-step-ahead	.02965	.015343
	12-step-ahead	.031529	.016426
Median	1-step-ahead	.011880	.004203
	2-step-ahead	.020591	.012972
	4-step-ahead	.021904	.013573
	12-step-ahead	.024776	.015558
Mode	1-step-ahead	.005156	.001531
	2-step-ahead	.008453	.005131
	4-step-ahead	.016925	.010492
	12-step-ahead	.017842	.014615

Notes: Metropolis-Hastings algorithm is used.

1-step-ahead means that the MSEF of 1 period out-of-sample prediction.

2-step-ahead means that the MSEF of 2 periods out-of-sample prediction.

4-step-ahead means that the MSEF of 4 periods out-of-sample prediction.

12-step-ahead means that the MSEF of 12 periods out-of-sample prediction.

Table 4: Conditional Kolmogorov Test (Data generated by the CIR model)

$\tau$	$\underline{u}, \bar{u}$	$V_T$	5%CV	10%CV	15%CV	20%CV
H <sub>0</sub> : CIR Model						
1	$\bar{R} \pm .5\sigma_R$	2.6101	1.0163	1.0163	.8812	.8812
	$\bar{R} \pm \sigma_R$	2.8640	1.0589	1.0589	.7675	.7675
2	$\bar{R} \pm .5\sigma_R$	2.0429	1.0571	1.0163	.9071	.8812
	$\bar{R} \pm \sigma_R$	2.4643	1.0589	.8929	.8143	.8071
4	$\bar{R} \pm .5\sigma_R$	1.2052	1.0571	1.0163	.8812	.8805
	$\bar{R} \pm \sigma_R$	1.2413	1.1042	1.0589	.9526	.8929
12	$\bar{R} \pm .5\sigma_R$	.3090	1.1114*	1.0553*	.9724*	.8812*
	$\bar{R} \pm \sigma_R$	.9121	1.1114*	1.0589*	.9526*	.9219*
H <sub>0</sub> : Vasicek Model						
1	$\bar{R} \pm .5\sigma_R$	5.7209	1.0447	1.0447	.9878	.9878
	$\bar{R} \pm \sigma_R$	4.2490	1.0447	1.0447	.8670	.8670
2	$\bar{R} \pm .5\sigma_R$	7.1857	1.2286	1.1000	1.0447	.9878
	$\bar{R} \pm \sigma_R$	4.9929	1.4498	1.4286	1.0571	1.0447
4	$\bar{R} \pm .5\sigma_R$	7.8592	1.2918	1.1286	1.0753	1.0447
	$\bar{R} \pm \sigma_R$	5.0518	1.4286	1.0825	1.0447	1.0000
12	$\bar{R} \pm .5\sigma_R$	8.4423	1.2918	1.1000	1.0553	1.0031
	$\bar{R} \pm \sigma_R$	2.8640	.6893	.6893	.6680	.6680

Notes: The GMM with unconditional moments is used.  
 Two confidence intervals ( $\underline{u}, \bar{u}$ ) are  $\bar{R} \pm .5\sigma_R$   
 and  $\bar{R} \pm \sigma_R$  are created for each of  $\tau = 1, 2, 4, 12$ .  
 $V_T$  is the test statistic.  $CV$ 's are the critical values at 5%, 10%, 15% and 20%.  
 A number with \* indicates that the test statistics  $V_T$ 's lead to the acceptance  
 of the null hypothesis.

Table 5: Conditional Kolmogorov Test (Data generated by the Vasicek model)

$\tau$	$\underline{u}, \bar{u}$	$V_T$	5%CV	10%CV	15%CV	20%CV
$H_0$ : CIR Model						
1	$\bar{R} \pm .5\sigma_R$	2.0467	.8559	.8559	.6183	.6183
	$\bar{R} \pm \sigma_R$	2.0325	.6965	.6965	.6325	.6325
2	$\bar{R} \pm .5\sigma_R$	1.6714	.8599	.8599	.8000	.7857
	$\bar{R} \pm \sigma_R$	1.1786	.7888	.7500	.7143	.6965
4	$\bar{R} \pm .5\sigma_R$	1.3279	.8599	.8599	.7857	.6357
	$\bar{R} \pm \sigma_R$	1.4795	.8588	.7888	.7361	.7143
12	$\bar{R} \pm .5\sigma_R$	.9799	.8732	.8599	.7857	.6784
	$\bar{R} \pm \sigma_R$	1.6583	1.0553	.8857	.8216	.7500
$H_0$ : Vasicek Model						
1	$\bar{R} \pm .5\sigma_R$	2.4447	.7817	.7817	.6680	.6680
	$\bar{R} \pm \sigma_R$	2.8640	.6893	.6893	.6680	.6680
2	$\bar{R} \pm .5\sigma_R$	2.9786	.7817	.6714	.6680	.6538
	$\bar{R} \pm \sigma_R$	4.0429	.6893	.6680	.5970	.5143
4	$\bar{R} \pm .5\sigma_R$	4.9003	.9598	.8083	.7145	.6714
	$\bar{R} \pm \sigma_R$	4.7054	.7178	.7073	.6680	.5970
12	$\bar{R} \pm .5\sigma_R$	4.2965	1.0376	.9598	.7817	.6714
	$\bar{R} \pm \sigma_R$	4.2212	.7538	.7538	.7178	.6893

Notes: The GMM with unconditional moments is used.  
 Two confidence intervals ( $\underline{u}, \bar{u}$ ) are  $\bar{R} \pm .5\sigma_R$   
 and  $\bar{R} \pm \sigma_R$  are created for each of  $\tau = 1, 2, 4, 12$ .  
 $V_T$  is the test statistic.  $CV$ 's are the critical values at 5%, 10%, 15% and 20%.

Table 6: Results of Monte Carlo Experiments, Bayesian Model Selection Non-normal Error Term

	% of Choosing CIR Model	% of Choosing Vasicek Model
Data generated by the CIR Model		
DIC	92%	8%
Median of MSEF		
1-step-ahead	64%	36%
2-step-ahead	62%	38%
4-step-ahead	59.5%	40.5%
12-step-ahead	58%	42%
Data generated by the Vasicek Model		
DIC	1.5%	98.5%
Median of MSEF		
1-step-ahead	40%	60%
2-step-ahead	43%	57%
4-step-ahead	40.5%	59.5%
12-step-ahead	28.5%	71.5%

Notes: Non-normal error term is the Normal  $\times$  Inverse Gaussian error term.  
The median of the MSEF is used for the model selection  
“#-step-ahead“ denotes the # periods out-of-sample prediction.

Table 7: Results of Monte Carlo Experiments, CKT, Non-normal Error Term: Acceptance rates (%) of  $H_0$

$\tau$	$\underline{u}, \bar{u}$	5%CV	10%CV	15%CV	20%CV
Data Generated by the CIR Model $H_0$ : CIR Model					
1	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
2	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
4	$\bar{R} \pm .5\sigma_R$	0.5%	0.5%	1.0%	1.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
12	$\bar{R} \pm .5\sigma_R$	5.0%	5.0%	5.0%	4.5%
	$\bar{R} \pm \sigma_R$	1.0%	0.5%	1.0%	0.0%
Data Generated by the Vasicek Model $H_0$ : Vasicek Model					
1	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
2	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
4	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%
12	$\bar{R} \pm .5\sigma_R$	0.0%	0.0%	0.0%	0.0%
	$\bar{R} \pm \sigma_R$	0.0%	0.0%	0.0%	0.0%

Notes: The GMM with unconditional moments is used.  
Two confidence intervals ( $\underline{u}, \bar{u}$ ) are  $\bar{R} \pm .5\sigma_R$   
and  $\bar{R} \pm \sigma_R$  are created for each of  $\tau = 1, 2, 4, 12$ .  
Non-normal error term is the Normal  $\times$  Inverse Gaussian error term.

Table 8: Bayesian Model Selection Criteria: Japanese Call Rates

Bayesian model selection		
	CIR Model	Vasicek Model
DIC	-6,031.0344	-3,168.2471
Mode of MSEF		
1-step ahead	.000307	.00250
2-step ahead	.000721	.00562
4-step ahead	.00141	.01062
12-step ahead	.00378	.02790

Notes: The Metropolis-Hastings algorithms are used.  
#-step ahead denotes the # periods out-of-sample prediction.

Table 9: Conditional Kolmogorov Test of Model Specification: Japanese Call Rates

$\tau$	$\underline{u}, \bar{u}$	$V_T$	5%CV	10%CV	15%CV	20%CV
CIR Model						
1	$\bar{R} \pm .5\sigma_R$	2.8069	2.1137	2.1137	1.7575	1.7575
	$\bar{R} \pm \sigma_R$	7.2926	1.9943	1.9943	1.6807	1.6807
2	$\bar{R} \pm .5\sigma_R$	4.0545	2.6022	2.3580	2.1552	2.1137
	$\bar{R} \pm \sigma_R$	9.9847	2.9896	2.4134	2.0464	1.9943
4	$\bar{R} \pm .5\sigma_R$	4.7523	2.3986	2.3580	2.2811	2.2405
	$\bar{R} \pm \sigma_R$	12.7235	3.1717	2.9896	2.8257	2.8151
12	$\bar{R} \pm .5\sigma_R$	5.8824	4.2445	3.2970	3.0719	2.6732
	$\bar{R} \pm \sigma_R$	16.2259	5.9308	3.5907	3.2306	3.1534
Vasicek Model						
1	$\bar{R} \pm .5\sigma_R$	2.8923	1.9687	1.9687	1.8108	1.8108
	$\bar{R} \pm \sigma_R$	7.4782	2.0412	2.0412	1.7106	1.7106
2	$\bar{R} \pm .5\sigma_R$	4.1057	2.5958	2.3750	2.1894	1.9687
	$\bar{R} \pm \sigma_R$	10.1575	2.9512	2.4775	2.1963	2.0412
4	$\bar{R} \pm .5\sigma_R$	4.7459	2.3750	2.3729	2.3409	2.1894
	$\bar{R} \pm \sigma_R$	12.8879	3.1632	3.0564	2.8001	2.6378
12	$\bar{R} \pm .5\sigma_R$	5.9209	4.2896	3.3313	3.1641	2.6282
	$\bar{R} \pm \sigma_R$	16.4702	5.7194	3.5521	3.2091	3.1534

Notes: The simulated GMM (SGMM) is used.

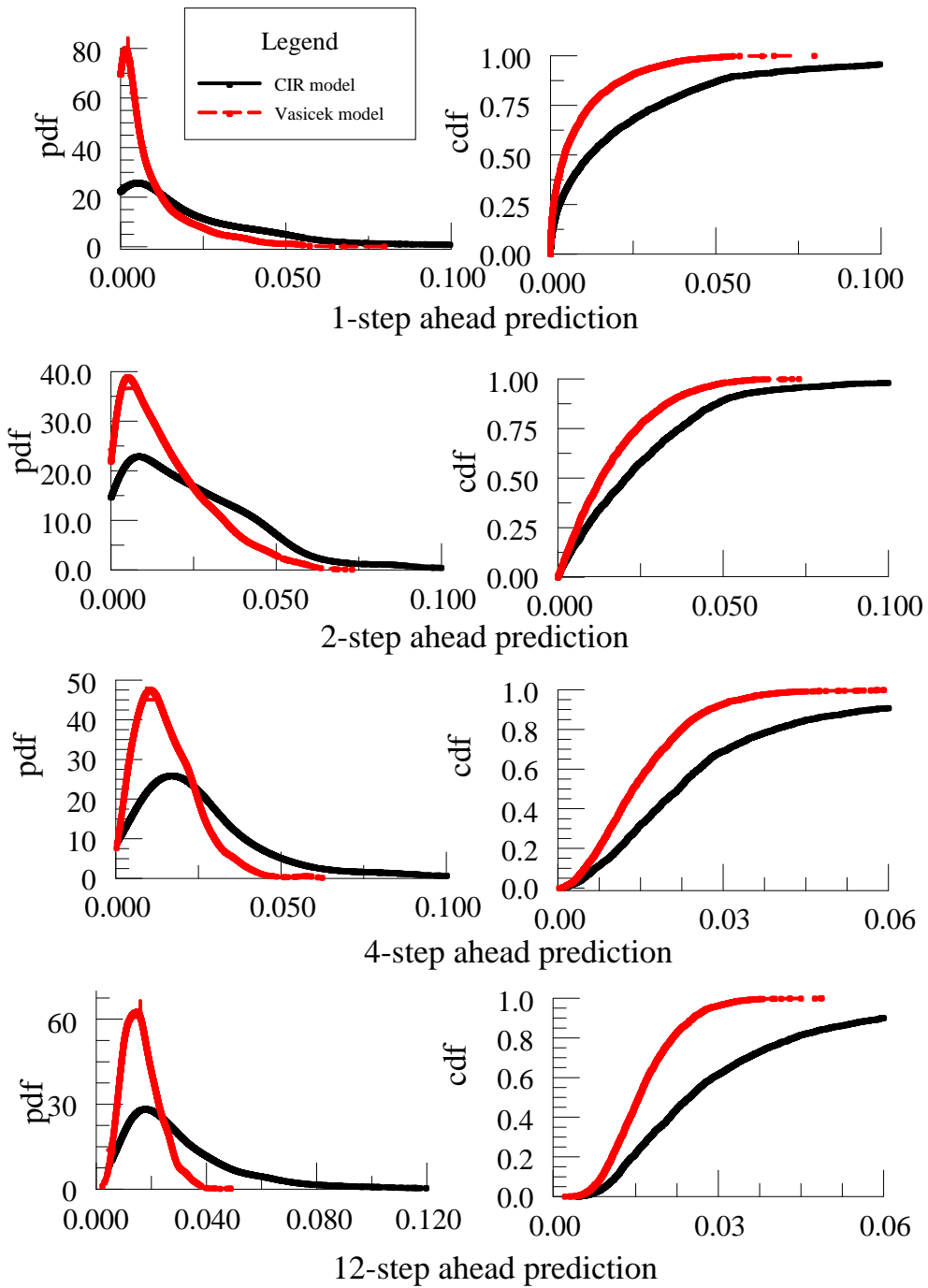


Figure 1: 1-,2-, 4-, and 12-period ahead MSEF's: Data generated by the Vasicek model



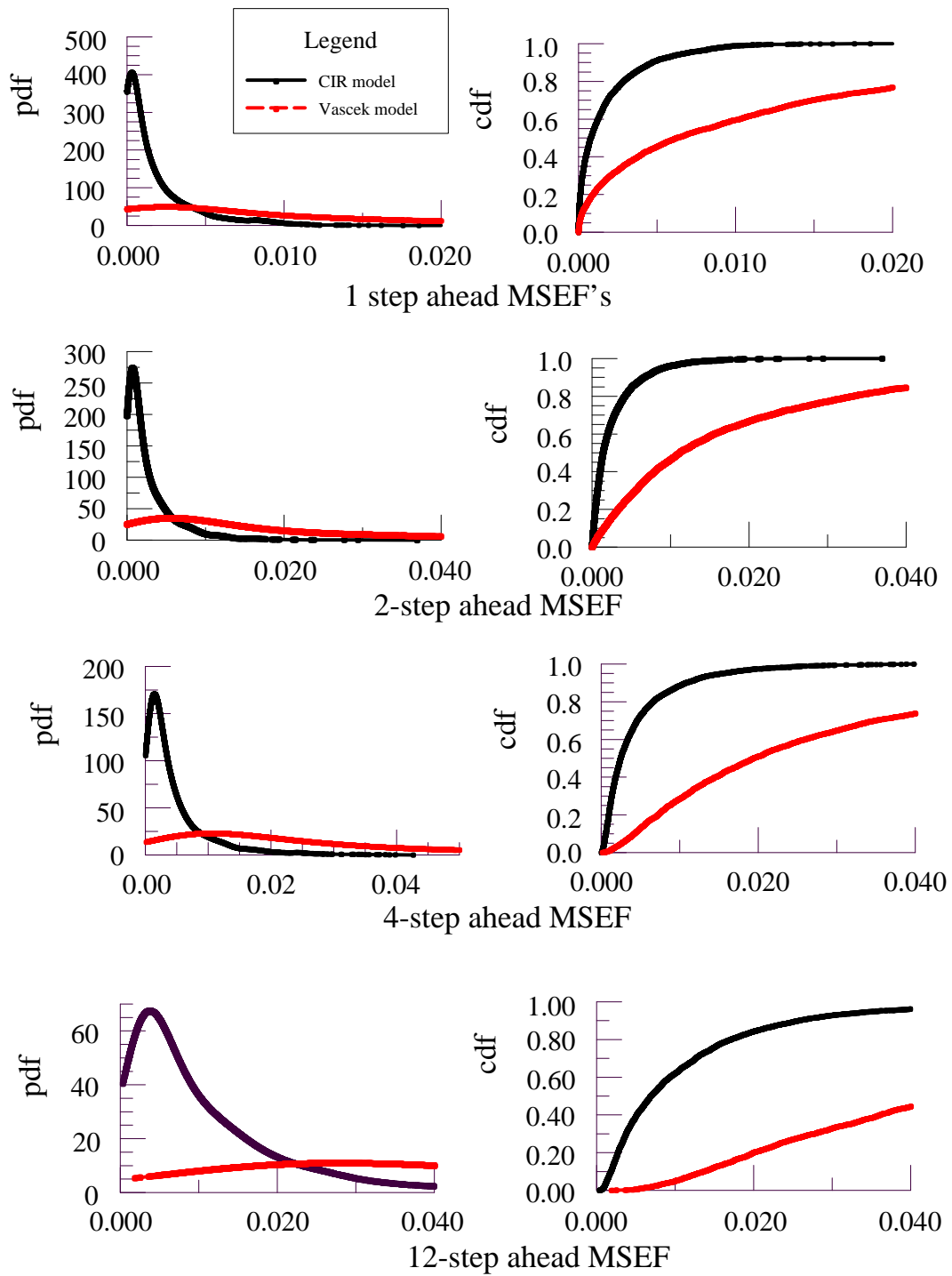


Figure 2: 1-, 2-, 4-, 12-period ahead MSEF's: CKLS Model for the Japanese Call Rate