Economics Bulletin

Volume 29, Issue 3

A Monte Carlo comparison of Bayesian testing for cointegration rank

Katsuhiro Sugita
Faculty of Law and Letters, University of the Ryukyus

Abstract

This article considers a Bayesian testing for cointegration rank, using an approach developed by Strachan and van Dijk (2007), that is based on Koop, Leon-Gonzalez, and Strachan (2006). The Bayes factors are calculated for selecting cointegrating rank. We calculate the Bayes factors using two methods - the Schwarz BIC approximation and Chib's (1995) algorithm for calculating the marginal likelihood. We run Monte Carlo simulations to compare the two methods.

Citation: Katsuhiro Sugita, (2009) "A Monte Carlo comparison of Bayesian testing for cointegration rank", *Economics Bulletin*, Vol. 29 no.3 pp. 2145-2151.
Submitted: Jul 29 2008. **Published:** September 02, 2009.

1 Introduction

This article considers testing for cointegrating rank with Bayesian approach based on Strachan and van Dijk (2007), that extends Koop, Leon-Gonzalez, and Strachan (2006) (hereafter KLS). Bayesian approach to analyze a cointegrated system suffers from both local and global identification problems through restrictions to the linear normalization on the cointegrating vectors. Recent literature such as Strachan (2003), Strachan and Inder (2004) and Villani (2005, 2006) show that it is important to elicit a prior on the space spanned by the cointegrating vectors in order to overcome identification problems. KLS develop algorithms to carry out efficient posterior simulation to elicit a prior on the space spanned by the cointegrating vectors. Strachan and van Dijk (2007) extends the KLS method for model averaging in vector autoregressive processes.

To estimate the number of rank in a vector error correction model (VECM), we consider two methods to compute the Bayes factor - the Schwarz BIC approximation method and Chib's (1995) algorithm for calculating the marginal likelihood. We run Monte Carlo simulations to compare the performances of these methods.

This article is structured as follows. Bayesian inference for a VECM is illustrated in Section 2. Section 3 presents Monte Carlo simulations using artificially generated data for the VECM with various rank in order to examine the performances of detecting the number of rank using the Chib and the BIC methods. Section 4 concludes. All computation in this article are performed using code written by the author with Ox v4.10 for Linux (Doornik, 2006)

2 Bayesian Inference in a Vector Error Correction Model

In this section, we present a Bayesian inference in a vector error correction model (VECM) based on Strachan and van Dijk (2007) that modified method proposed by KLS. Let y_t denote a vector of $1\times n$ with r linear cointegrating relations. Then, with $\varepsilon_t \sim iidN(0,\Omega)$, a VECM with p lags is expressed as

$$\Delta y_t = (y_{t-1}\beta^{+} + d_{1,t}\theta_1)\alpha + d_{2,t}\theta_2 + \sum_{l=1}^{p} \Delta y_{t-l}\Gamma_l + \varepsilon_t \tag{1}$$

where $\Delta y_t = y_t - y_{t-1}$, ε_t is $1\times n$; Γ_l and Ω are $n\times n$, the cointegrating vector β^{+} and the adjustment term α' are $n\times r$, assumed to have rank r. The deterministic terms, $d_{1,t}\theta_1$ ($d_{1,t}$ is $1\times m_1$, and $\theta_1$is $m_1\times r$) and $d_{2,t}\theta_2$ ($d_{2,t}$ is $1\times m_2$, and θ_2 is $m_2\times n$) are defined as follows. For example, if the process contains both an intercept and linear time trend, then $\theta_{i,t} = (\mu_i', \delta')'$ and $d_{i,t} = (1,t)$, or if the process contains an intercept but no time trend, then $\theta_{i,t} = \mu_i$ and $d_{i,t} = 1$. Let the $1\times(n+m_1)$ vector $z_{1,t} = (y_{t-1}, d_{1,t})$, the $1\times(m_2+nl)$ vector $z_{2,t} = (d_{2,t}, \Delta y_{t-1}, \ldots, \Delta y_{t-l})$, the $(m_2+nl)\times n$ matrix $\Phi = \left(\theta_2', \Gamma_1', \ldots, \Gamma_l'\right)'$, and the $(m_1+n)\times r$ matrix $\beta = (\beta^{+\prime}, \theta_1')'$, then we can rewrite the equation (1) as

$$\Delta y_t = z_{1,t}\beta\alpha + z_{2,t}\Phi + \varepsilon_t \tag{2}$$

Let define the $T\times n$ matrix $y = (\Delta y_1', \ldots, \Delta y_T')'$, the $T\times n$ matrix $\varepsilon = (\varepsilon_1', \ldots, \varepsilon_T')'$, the $T\times(n+m_1)$ matrix $Z = (z_{1,1}', z_{1,2}', \ldots, z_{1,T}')'$, the $T\times(m_2+nl)$ matrix $x = (z_{2,1}', z_{2,2}', \ldots, z_{2,T}')'$, the $T\times(m_2+nl+r)$ matrix $W = (Z\beta$,x), and the $(m_2+nl+r)\times n$ matrix $B = (\alpha', \Phi')'$, then (2) can be simplified as

$$\begin{aligned} y &= Z\beta\alpha + x\Phi + \varepsilon & (3)\\ &= WB + \varepsilon & (4) \end{aligned}$$

To estimate the VECM expressed in (4), we follow a Bayesian approach to analyze a VAR model with multiple structural breaks proposed by Sugita (2007). First we need to specify the prior distributions for the parameters given in the model. For the prior for the covariance-variance matrix, Ω, we assign an inverted Wishart distribution with mean Ω_0 and degree of freedom ν_0 as $\Omega \sim IW(\Omega_0, \nu_0)$. For the prior for B, we consider that the vectorized B is a multivariate normal with mean $\mathrm{vec}(B_0)$ and variance V_{B_0} as $vec(B) \sim MN(\mathrm{vec}(B_0), V_{B_0})$. For the prior for the cointegrating vector β, we adopt the KLS's efficient method of posterior simulation based on the idea of a collapsed Gibbs sampler developed in Liu (1994) and Liu, Wong and Kong (1994). The KLS's method states that the matrix of long run multipliers can be decomposed as $\beta\alpha = [\beta\kappa][\kappa^{-1}\alpha] = ba$ where β is restricted to be semi-orthogonal, $b = \beta\kappa$, $a = \kappa^{-1}\alpha$, and κ is positive definite and defined as $\kappa = (\alpha\alpha')^{1/2}$. We give b a Normal prior with mean b_0 and covariance matrix V_{b_0}.

The likelihood function is given as

$$
\begin{aligned}
\mathfrak{L}(\beta, B, \Omega \mid Y) &\propto |\Omega|^{-T/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\Omega^{-1}(y-WB)'(y-WB)\right\}\right] \\
&= |\Omega|^{-T/2} \exp\left[-\frac{1}{2}\mathrm{tr}\left\{\mathrm{vec}\,(y-WB)'(\Omega\otimes I_T)^{-1}\mathrm{vec}\,(y-WB)\right\}\right]
\end{aligned}
$$

With this likelihood and the priors, we obtain the following joint posterior:

$$
\begin{aligned}
&p(\beta, B, \Omega \mid y) \propto p(\beta, B, \Omega) L(\beta, B, \Omega \mid y) \\
&\propto p(b)\,|\Omega_0|^{\nu_0/2}\,|\Omega|^{-(T+\nu_0+n+1)/2}\,|V_{B_0}|^{-1/2} \exp\left[-\frac{1}{2}\left\{\mathrm{tr}\left(\Omega^{-1}\right) + \mathrm{vec}\,(y-WB)'\right.\right. \\
&\left.\left.\times(\Omega\otimes I_T)^{-1}\mathrm{vec}\,(y-WB)\right\} + \mathrm{vec}\,(B-B_0)'V_0^{-1}\mathrm{vec}\,(B-B_0)\right] \\
&= p(b)\,|\Omega_0|^{\nu_0/2}\,|\Omega|^{-(T+\nu_0+n+1)/2}\,|V_{B_0}|^{-1/2} \exp\left[-\frac{1}{2}\left\{\mathrm{tr}\left(\Omega^{-1}\right) + \mathrm{vec}\,(B-B_\star)'\right.\right. \\
&\left.\left.\times V_{B_\star}^{-1}\mathrm{vec}\,(B-B_\star) + \xi\right\}\right]
\end{aligned}
$$

where $p(b)$ denotes the prior for b, and

$$
\xi = \mathrm{vec}\,(y)'(\Omega\otimes I_T)^{-1}\mathrm{vec}\,(y) + \mathrm{vec}\,(B_0)'V_{B_0}^{-1}\mathrm{vec}\,(B_0) - \mathrm{vec}\,(B_\star)'V_{B_\star}^{-1}\mathrm{vec}\,(B_\star),
$$

$$
\mathrm{vec}\,(B_\star) = \left[V_{B_0}^{-1} + \left\{\Omega^{-1}\otimes\left(W'W\right)\right\}\right]^{-1}\left\{V_{B_0}^{-1}\mathrm{vec}\,(B_0) + (\Omega\otimes I_\kappa)^{-1}\mathrm{vec}\left(W'y\right)\right\},
$$

and

$$
V_{B_\star} = \left[V_{B_0}^{-1} + \left\{\Omega^{-1}\otimes\left(W'W\right)\right\}\right]^{-1}
$$

By integrating out above joint posterior, the conditional posterior of Ω_i is derived as an inverted Wishart distribution with mean $\Omega_\star$ and degree of freedom $T+\nu_0$ as:

$$
p(\Omega \mid \beta, B, Y) \propto |\Omega_\star|^{T/2}|\Omega|^{-(T+\nu_0+n+1)/2}\exp\left[-\frac{1}{2}\mathrm{tr}\left(\Omega^{-1}\Omega_\star\right)\right] \tag{5}
$$

where $\Omega_\star = (y-WB)'(y-WB)+\Omega_0$. The conditional posterior of $\mathrm{vec}\,(B)$ is derived as a multivariate normal density with mean $\mathrm{vec}\,(B_\star)$ and covariance $V_{B_\star}$, that is,

$$
\begin{aligned}
&p(\mathrm{vec}\,(B) \mid \beta, \Omega, y) \\
&\propto |V_{B_\star}|^{-1/2}\exp\left[-\frac{1}{2}\left\{\mathrm{vec}\,(B-B_\star)'V_{B_\star}^{-1}\mathrm{vec}\,(B-B_\star)\right\}\right]
\end{aligned} \tag{6}
$$

See Sugita (2007) for deriving above conditional posterior densities.

To obtain the conditional posterior for the cointegrating vectors, we rewrite the equation (3) as

$$
y = Zba + x\Phi + \varepsilon \tag{7}
$$

where b and a are used instead of β and α respectively. By deducting $x\Phi$ from the both sides and then vectorizing the both side, we obtain

$$
\begin{aligned}
\mathrm{vec}\,(y - x\Phi) &= \mathrm{vec}\,(Zba) + \mathrm{vec}\,(\varepsilon) \\
&= \left(a'\otimes Z\right)\mathrm{vec}\,(b) + \mathrm{vec}\,(\varepsilon),
\end{aligned} \tag{8}
$$

or $\tilde{y} = z\tilde{b} + e$ where $\tilde{y} = \text{vec}(y - x\Phi)$, $z = a' \otimes Z$, $\tilde{b} = \text{vec}(b)$, and $e = \text{vec}(\varepsilon)$. Thus, the conditional posterior distribution for $\tilde{b}$ is a multivariate normal with mean $\tilde{b}_\star$ and variance $V_{\tilde{b}_\star}$ where $V_{\tilde{b}_\star} = \left[V_{b_0}^{-1} + \left\{ \left(a\Omega^{-1}a \right)' \otimes (Z'Z) \right\} \right]^{-1}$ and $\tilde{b}_\star = V_{\tilde{b}_\star} \left[V_{b_0}^{-1} \text{vec}(b_0) + \left(a\Omega^{-1} \otimes Z \right) \tilde{y} \right]$.

From these conditional posterior distributions derived above, we summarize the following Gibbs sampling scheme:

- Step 1: Set $j = 1$. Specify starting values for the parameters of the model, $\beta^{(0)}$, $B^{(0)}$, and $\Omega^{(0)}$.
- Step 2: Generate $\Omega^{(j)}$ from $p(\Omega \mid \beta^{(j-1)}, B^{(j-1)}, y)$.
- Step 3: Generate $vec(B)^{(j)}$ from $p(vec(B) \mid \beta^{(j-1)}, \Omega^{(j)}, y)$ to obtain α^*. Compute a^* using $a^* = (\alpha^* \alpha^{*\prime})^{-\frac{1}{2}} \alpha^*$.
- Step 4: Generate b^* from $p(\widetilde{b} \mid \Omega^{(j)}, B^{(j)}, Yy)$. Then, compute $\beta^{(j)} = b^*(b^{*\prime}b^*)^{-\frac{1}{2}}$ and $\alpha^{(j)} = a^*(b^{*\prime}b^*)^{\frac{1}{2}}$.
- Step 5: Set $j = j + 1$, and go back to Step 2.

Step 2 through Step 5 can be iterated N times to obtain the posterior densities. Note that the first L iterations are discarded in order to remove the effect of the initial values.

There are several methods to calculate the Bayes factor such as Chib (1995), Gelfand and Dey (1994), the Savage-Dickey density ratio (see Verdinelli and Wasserman, 1995), and the Schwarz Bayesian information criterion (BIC) approximation method (Schwarz, 1978). Among these, we choose the BIC and Chib's method to select the cointegrating rank. The BIC can give a rough approximation to the Bayes factors. It is, however, easy to implement and does not require evaluation of the prior distribution, as Kass and Raftery (1995) note. The BIC to approximate the Bayes factors is employed by Wang and Zivot (2000) for detecting the number of structural breaks. The BIC for model j, $\mathcal{M}_j$, is calculated as

$$\text{BIC}_j = -2\ln \mathfrak{L}\left(\widehat{\theta_j} \mid y; \mathcal{M}_j\right) + q_j \ln(t) \tag{9}$$

where $\mathfrak{L}\left(\widehat{\theta_j} \mid y; \mathcal{M}_j\right)$ denotes the likelihood function under the model j; q_j denotes the total number of estimated parameters in $\mathcal{M}_j$. The likelihood function $\mathfrak{L}\left(\widehat{\theta_j} \mid y; \mathcal{M}_j\right)$ is evaluated at $\widehat{\theta_j}$, the posterior means of the parameters for $\mathcal{M}_j$.

With the BICs for $\mathcal{M}_j$ and $\mathcal{M}_i$, the Bayes factor for $\mathcal{M}_j$ against $\mathcal{M}_i$ can be approximated by

$$BF_{ji} = \exp\left[-0.5\left(\text{BIC}_i - \text{BIC}_j\right)\right]. \tag{10}$$

With the prior odds, defined as $p(\mathcal{M}_j)/p(\mathcal{M}_i)$, the posterior odds can be computed by multiplying the Bayes factor by the prior odds as $\text{PosteriorOdds}_{ji} = BF_{ji} \times \text{PriorOdds}_{ji}$. By using the BIC to approximate to logarithm of the Bayes factor, it is easy to detect the number of the rank in the VECM as a problem of model selection.

The other method to calculate the Bayes factors is proposed by Chib (1995) providing a method of computing the marginal likelihood that utilizes the output of the Gibbs sampler. The marginal likelihood can be expressed from the Bayes rule as

$$p(y \mid \mathcal{M}_i) = \frac{p(y \mid \theta_i^\star) p(\theta_i^\star)}{p(\theta_i^\star \mid y)} \tag{11}$$

where $p(y \mid \theta_i^\star)$ is the likelihood for Model i evaluated at $\theta_i^\star$, which is the Gibbs output or the posterior mean of θ_i, $p(\theta_i^\star)$ is the prior density and $p(\theta_i^\star \mid y)$ is the posterior density. If the exact forms of the marginal posteriors are not known like our case, $p(\theta_i^\star \mid y)$ cannot be calculated. To estimate the marginal posterior density evaluated at $\theta_i^\star$ using the conditional posteriors, first block θ into l segments as $\theta = (\theta_1', \ldots, \theta_l')'$, and define $\varphi_{i-1} = (\theta_1', \ldots, \theta_{i-1}')$ and $\varphi^{i+1} = (\theta_{i+1}', \ldots, \theta_l')$. Since $p(\theta^\star \mid y) = \prod_{i=1}^{l} p(\theta_i^\star \mid y, \varphi_{i-1}^\star)$, we can draw $\theta_i^{(j)}$, $\varphi^{i+1,(j)}$, where j indicates the Gibbs output $j = 1, \ldots, N$, from $(\theta_i, \ldots, \theta_l) = (\theta_i, \varphi^{i+1}) \sim p(\theta_i, \varphi^{i+1} \mid y, \varphi_{i-1}^\star)$, and then estimate $\widehat{p}(\theta_i^\star \mid y, \varphi_{i-1}^\star)$ as

$$\widehat{p}(\theta_i^\star \mid y, \varphi_{i-1}^\star) = \frac{1}{N} \sum_{j=1}^{N} p(\theta_i^\star \mid y, \varphi_{i-1}^\star, \varphi^{i+1,(j)}).$$

Thus, the posterior $p(\theta_i^\star \mid Y)$ can be estimated as

$$\widehat{p}(\theta^\star \mid y) = \prod_{i=1}^{l} \left\{ \frac{1}{N} \sum_{j=1}^{N} p(\theta_i^\star \mid y, \varphi_{i-1}^\star, \varphi^{i+1,(j)}) \right\}. \tag{12}$$

3 Monte Carlo Simulation

In this section we perform some Monte Carlo simulations to illustrate the performances of Bayesian tests for the rank of cointegration described in the previous section. The data generating processes (DGPs) consist of two and three-variable VECMs with an intercept term having various number of cointegrating vectors.

We consider the following DGPs for two-variable VECM:

$$\begin{aligned}
&\text{DGP 1 with } r=0: \quad \Delta y_t = \mu + \varepsilon_t \\
&\text{DGP 2 with } r=1: \quad \Delta y_t = \mu + y_{t-1} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} -0.1 & -0.1 \end{bmatrix} + \varepsilon_t \\
&\text{DGP 3 with } r=2: \quad \Delta y_t = \mu + y_{t-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -0.1 & 0.1 \\ -0.1 & -0.1 \end{bmatrix} + \varepsilon_t
\end{aligned}$$

where $\mu = \begin{bmatrix} 0.1 & 0.1 \end{bmatrix}$, and $\varepsilon_t \sim iidN(0, 0.1I_2)$.

We also consider the following DGPs for three-variable VECM:

$$\begin{aligned}
&\text{DGP 4 with } r=0: \quad \triangle y_t = \mu + \varepsilon_t \\
&\text{DGP 5 with } r=1: \quad \triangle y_t = \mu + y_{t-1} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \begin{bmatrix} -0.1 & -0.1 & -0.1 \end{bmatrix} + \varepsilon_t \\
&\text{DGP 6 with } r=2: \quad \triangle y_t = \mu + y_{t-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} -0.1 & 0.1 & -0.1 \\ -0.1 & -0.1 & 0.1 \end{bmatrix} + \varepsilon_t \\
&\text{DGP 7 with } r=3: \quad \triangle y_t = \mu + y_{t-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -0.1 & 0.1 & -0.1 \\ -0.1 & -0.1 & 0.1 \\ -0.1 & -0.1 & -0.1 \end{bmatrix} + \varepsilon_t
\end{aligned}$$

where $\mu = \begin{bmatrix} 0.1 & 0.1 & 0.1 \end{bmatrix}$, and $\varepsilon_t \sim NID(0, 0.1I_3)$.

We demonstrate the performance of the test by varying the true rank. Each simulation of DGPs for various ranks is repeated 1000 times. The sample size T is 50, 100, 200, 500, and 1000. The prior parameter specifications for the natural conjugate prior are given with $\Omega_0 = I$, $\nu_0 = 3.01$, $\mathrm{vec}(B_0) = 0$, $V_{B_0} = 100I$, $b_0 = 0$, and $V_{b_0} = 100I$. These prior variances are chosen for relatively noninformative prior.

Table 1 summarizes the results of Monte Carlo simulation for two-variable VECMs ($n = 2$) by computing the BIC and the Chib's method. Table 2 reports the results when $n = 3$. Each value in the Table represents the average posterior probabilities of 1,000 iterations for each true rank. For each iteration, the Gibbs sampling is performed with 10,000 draws and the first 1,000 discarded.

Our findings from Table 1 and 2 are as follows. When $T = 50$, the highest average posterior probabilities do not indicate the correct rank except when the true rank is 0. Increasing the sample size to 100 improves the performances of test by the BIC as all highest average posterior probabilities indicate the true rank. When $T = 200$, the BIC method chooses the correct rank with more than 90% for each of the rank. For the Chib's method, only when $T = 1000$, each of the true rank is correctly chosen.

Table 1: Monte Carlo Results: The Average Posterior Probabilities with $n = 2$

		$T = 50$		$T = 100$		$T = 200$		$T = 500$		$T = 1000$	
DGP	rank r	BIC	Chib	BIC	Chib	BIC	Chib	BIC	Chib	BIC	Chib
$r = 0$	0	0.945	1.000	0.968	1.000	0.982	1.000	0.993	1.000	1.000	1.000
	1	0.048	0.000	0.029	0.000	0.017	0.000	0.007	0.000	0.000	0.000
	2	0.008	0.000	0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000
$r = 1$	0	0.416	0.996	0.044	0.747	0.000	0.049	0.000	0.000	0.000	0.000
	1	0.391	0.004	0.720	0.253	0.975	0.951	0.998	1.000	0.999	1.000
	2	0.193	0.000	0.236	0.000	0.025	0.000	0.002	0.000	0.001	0.000
$r = 2$	0	0.632	1.000	0.297	1.000	0.005	0.999	0.000	0.279	0.000	0.000
	1	0.130	0.000	0.106	0.000	0.012	0.001	0.000	0.253	0.000	0.000
	2	0.237	0.000	0.597	0.000	0.984	0.000	1.000	0.468	1.000	1.000

Table 2: Monte Carlo Results: The Average Posterior Probabilities with $n = 3$

		$T = 50$		$T = 100$		$T = 200$		$T = 500$		$T = 1000$	
DGP	rank r	BIC	Chib	BIC	Chib	BIC	Chib	BIC	Chib	BIC	Chib
$r = 0$	0	0.972	1.000	0.993	1.000	0.995	1.000	1.000	1.000	1.000	1.000
	1	0.027	0.000	0.007	0.000	0.005	0.000	0.000	0.000	0.000	0.000
	2	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$r = 1$	0	0.686	1.000	0.208	0.991	0.000	0.746	0.000	0.000	0.000	0.000
	1	0.241	0.000	0.732	0.009	0.951	0.254	0.986	1.000	0.995	1.000
	2	0.048	0.000	0.039	0.000	0.038	0.000	0.012	0.000	0.004	0.000
	3	0.025	0.000	0.021	0.000	0.011	0.000	0.002	0.000	0.001	0.000
$r = 2$	0	0.554	1.000	0.114	0.998	0.000	0.867	0.000	0.000	0.000	0.000
	1	0.175	0.000	0.123	0.002	0.006	0.118	0.000	0.022	0.000	0.000
	2	0.228	0.000	0.679	0.000	0.919	0.016	0.945	0.978	0.956	1.000
	3	0.043	0.000	0.084	0.000	0.074	0.000	0.055	0.000	0.044	0.000
$r = 3$	0	0.557	1.000	0.126	0.996	0.000	0.805	0.000	0.001	0.000	0.000
	1	0.115	0.000	0.066	0.004	0.000	0.189	0.000	0.197	0.000	0.000
	2	0.078	0.000	0.094	0.000	0.004	0.006	0.000	0.740	0.000	0.000
	3	0.250	0.000	0.714	0.000	0.996	0.000	1.000	0.062	1.000	1.000

4 Conclusion

This article considers a Bayesian method to test for cointegrating rank. Estimating the cointegrated model is based on Strachan and van Dijk (2007) and KLS, that is an efficient and valid method. The cointegrating rank is chosen by computing the Bayes factors. The Bayes factors are calculated using the BIC method or the Chib's method. We compare the both methods in simulation and find that the BIC method, although it gives only an approximation of the Bayes factor, is useful even when the sample size is not large, while the Chib's method can be used only when the sample size is large. In our simulation, when the sample size is 1000, the Chib's method performs better than the BIC's method.

References

[1] Chib, S. (1995) "Marginal Likelihood from the Gibbs Output" *Journal of the American Statistical Association* **90**, 432, Theory and Methods, 1313-1321.

[2] Doornik, J.A. (2006) *OX an Object-Oriented Matrix Programming Language.* London: Timberlake Consultants Press.

[3] Gelfand, A. and D. Dey (1994) "Bayesian Model Choice: Asymptotics and Exact Calculations" *Journal of the Royal Statistical Society Series B* **56**, 501-14.

[4] Kass, R. E. and A. E. Raftery (1995) "Bayes Factors" *Journal of the American Statistical Association* **90**, 773-795.

[5] Koop, G., R. Leon-Gonzalez and R. Strachan (2006) "Efficient Posterior Simulation for Cointegrated Models with Priors on the Cointegration Space" manuscript.

[6] Liu, J.S. (1994) "The Collapsed Gibbs Sampler with Application to a Gene Regulation Problem" *Journal of the American Statistical Association* **89**, 958-966.

[7] Liu, J.S. W.H. Wong, and A. Kong (1994) "Covariance Structure of the Gibbs Sampler with Application to Comparison of Estimators and Augmentation Schemes" *Biometrika* **81**, 27-40.

[8] Schwarz, G. (1978) "Estimating the Dimension of a Model" *Annals of Statistics* **6**, 461-464.

[9] Strachan, R.W. (2003) "Valid Bayesian Estimation of the Cointegrating Error Correction Model" *Journal of Business and Economic Statistics* **21**, 185-195.

[10] Strachan, R.W. and B. Inder (2004) "Bayesian Analysis of the Error Correction Model" *Journal of Econometrics* **123**, 307-325.

[11] Strachan, R. W. and H.K. van Dijk (2007) "Bayesian Model Averaging in Vector Autoregressive Processes with an Investigation of Stability of the US Great Ratios and Risk of a Liquidity Trap in the USA, UK and Japan"

[12] Sugita, K. (2007) "Bayesian Analysis of a Vector Autoregressive Model with Multiple Structural Breaks" *Economics Bulletin* **3**, 22, 1-7.

[13] Verdinelli, I. and L. Wasserman (1995) "Computing Bayes Factors Using a Generalization of the Savage-Dickey Density Ratio" *Journal of the American* Statistical *Association* **90**, 614-618.

[14] Villani, M. (2005) "Bayesian Reference Analysis of Cointegration" *Econometric Theory* **21**, 326-357.

[15] Villani, M. (2006) "Bayesian Point Estimation of the Cointegration Space" *Journal of Econometrics* **134**, 645-664.

[16] Wang, J. and E. Zivot (2000) "A Bayesian Time Series Model of Multiple Structural Changes in Level, Trend, and Variance" *Journal of Business and Economic Statistics* **18**, 374-386.