

琉球大学学術リポジトリ

Model Selection for a Markov Switching Vector Error Correction Model Using the Marginal Likelihood

メタデータ	言語: 出版者: 琉球大学法文学部 公開日: 2017-01-25 キーワード (Ja): キーワード (En): 作成者: 杉田, 勝弘, Sugita, Katsuhiko メールアドレス: 所属:
URL	https://doi.org/10.24564/0002008493

Model Selection for a Markov Switching Vector Error Correction Model Using the Marginal Likelihood

Katsuhiko Sugita

Faculty of Law and Letters, University of the Ryukyus, Nishihara,

Senbaru, Okinawa, 903-0213, Japan

E-Mail: ksugita@ll.u-ryukyu.ac.jp

Abstract

In this paper we consider model selection for a Markov switching vector error correction model. We apply the algorithm proposed by Chib (1995) to calculate the marginal likelihood and the Bayes factors for this model to select the most appropriate model among all models under consideration. We perform a simple Monte Carlo simulation to illustrate the performances of the method for model selection.

1 Introduction

In this paper we consider model selection for a Markov switching vector error correction model (MS-VECM) using the marginal likelihood method proposed by Chib (1995). An MS-VECM is a nonlinear cointegration model which allows for regime shifts in VECM. Jochmann and Koop (2015) develop methods for Bayesian inference in an MS-VECM, and use a bridge sampling to calculate marginal likelihood to select a single model. In this paper, we present a procedure for calculating the marginal likelihoods for models under consideration by extending Chib's method. We show simulated experiments with artificially generated data to evaluate how the Chib's method can detect an appropriate model.

2 MS-VECM

This section introduces an MS-VECM. A linear VECM with the number of lags p for $1 \times n$ vector y_t can be written as:

$$\Delta y_t = y_{t-1} \beta \alpha + \mu + \sum_{l=1}^p \Delta y_{t-l} \Gamma_l + \varepsilon_t, \quad (1)$$

where β ($n \times r$) contains the cointegrating vectors, α ($r \times n$) contains the adjustment terms, μ ($1 \times n$) is the vector of intercept terms, and Γ_l ($n \times n$) is the lag term, and $\varepsilon_t \sim iidN(0, \Omega)$. If we assume that r , β , α , μ , Γ_l 's, and Ω in the VECM (1) are subject to an unobservable discrete state variable s_t that evolves according to an m -state, first-order Markov switching process with the transition probabilities, $p(s_t = i | s_{t-1} = j) = q_{ij}$, $i, j = 1, \dots, m$, then an MS-VECM with $\varepsilon_t \sim N(0, \Omega(s_t))$ is written as

$$\Delta y_t = y_{t-1} \beta(s_t) \alpha(s_t) + \mu(s_t) + \sum_{l=1}^p \Delta y_{t-l} \Gamma_l(s_t) + \varepsilon_t, \quad (2)$$

where $\beta(s_t)$ is $ann \times r(s_t)$, $\alpha(s_t)$ is $r(s_t) \times n$, $\mu(s_t)$ is $1 \times n$, and $\Gamma_l(s_t)$ is $n \times n$ matrix.

The MS-VECM in (2) can be rewritten as

$$\Delta y_t = y_{t-1} \beta(s_t) \alpha(s_t) + z_t \Phi + \varepsilon_t, \quad (3)$$

where z_t is $1 \times m(1 + np)$ and Φ is $m(1 + np) \times n$, and defined as

$$z_t = (\iota_t(1), \dots, \iota_t(m), \iota_t(1)\Delta y_{t-1}, \dots, \iota_t(1)\Delta y_{t-p}, \dots, \iota_t(m)\Delta y_{t-1}, \dots, \iota_t(m)\Delta y_{t-p}),$$

$$\Phi = (\mu(1)', \dots, \mu(m)', \Gamma_1(1)', \dots, \Gamma_p(1)', \dots, \Gamma_1(m)', \dots, \Gamma_p(m))',$$

and $\iota_t(i)$ in z_t is an indicator variable that equals to 1 if regime is i at t , and 0 otherwise. From equation (3), let define

$$\text{the } T \times n \text{ matrices } Y^T = (\Delta y_1', \dots, \Delta y_T')' \text{ and } E = (\varepsilon_1', \dots, \varepsilon_T')',$$

$$\text{the } T \times n \text{ matrix } X_i = (y_0' \iota_0(i), y_1' \iota_1(i), \dots, y_{T-1}' \iota_{T-1}(i))',$$

$$\text{the } T \times m(1 + np) \text{ matrix } Z^T = (z_1', z_2', \dots, z_T')',$$

$$\text{the } T \times h \text{ (where } h = m(1 + np) + \sum_{i=1}^m r(i)) \text{ matrix } W = (X_1 \beta(1), \dots, X_m \beta(m), Z^T),$$

$$\text{the } h \times n \text{ matrix } B = (\alpha(1)', \dots, \alpha(m)', \Phi')',$$

then we can simplify the model as follows:

$$\begin{aligned}
 Y^T &= \sum_{i=1}^m X_i \beta(i) \alpha(i) + Z^T \Phi + E \\
 &= WB + E.
 \end{aligned} \tag{4}$$

Thus, to estimate the MS-VCEM in (3), we consider the simplified form given in (5).

We choose the following priors:

$$\Omega(s_t) \sim IW(\Omega_0(s_t), \nu_0(s_t)). \tag{6}$$

Let define $B = (\alpha(1)', \dots, \alpha(m)', \mu(1)', \dots, \mu(m)', \Gamma_1(1)', \dots, \Gamma_p(m)')'$, then

$$\text{vec}(B) \sim MN(\text{vec}(B_0), V_{B_0}). \tag{7}$$

where MN denotes the matricvariate-Normal and IW denotes the inverted Wishart distribution. For a prior for the cointegrating vector $\beta(s_t)$, we consider the following transformation (see Koop et al. (2010)):

$$\begin{aligned}
 \beta(s_t) \alpha(s_t) &= [\beta(s_t) \kappa(s_t)] [\kappa(s_t)^{-1} \alpha(s_t)] \\
 &= \left[\beta(s_t) (\alpha(s_t) \alpha(s_t)')^{1/2} \right] \left[(\alpha(s_t) \alpha(s_t)')^{-1/2} \alpha(s_t) \right] \equiv b(s_t) a(s_t),
 \end{aligned}$$

where $\kappa(s_t) \equiv [\alpha(s_t) \alpha(s_t)']^{1/2}$, and thus $b(s_t) = \beta(s_t) \kappa(s_t)$. We assign a prior for $\text{vec}(b(s_t))$ as

$$\tilde{b}(s_t) = \text{vec}(b(s_t)) \sim MN(\tilde{b}_0(s_t), V_{b_0}(s_t)). \tag{8}$$

A prior for the transition probabilities, q_{ij} , $i, j = 1, \dots, m$, is beta distribution as

$$q_{ij} \sim \text{beta}(u_{ij}, \bar{u}_{ij}), \tag{9}$$

where $\bar{u}_{ii} = Pr(s_t \neq i | s_{t-1} = i)$ and $\bar{u}_{ij} = Pr(s_t = j | s_{t-1} = i, s_t \neq i)$ for $i \neq j$,

With the above priors, the conditional posteriors are given as (see Sugita (2016) for details):

$$\Omega(i) | b(i), B, \tilde{S}_T, Y^T \sim IW((Y_i - W_i B)'(Y_i - W_i B) + \Omega_0(i), t_i + \nu_0(i) + n + 1), \tag{10}$$

$$\text{vec}(B) \mid \mathbf{b}, \Omega, \tilde{S}_T, Y^T \sim MN(\text{vec}(B_1), V_{B_1}), \quad (11)$$

$$\tilde{b}(i) \mid \Omega(i), \text{vec}(B), \tilde{S}_T, Y^T \sim MN(\tilde{b}_*(i), V_{\tilde{b}_*}(i)), \quad (12)$$

$$p(q_{ij} \mid \tilde{S}_T) \propto q_{ij}^{u_{ij}+m_{ij}-1} (1 - \bar{q}_{ij})^{\bar{u}_{ij}+\bar{m}_{ij}-1}. \quad (13)$$

where

$$V_{B_1} = \left\{ V_{B_0}^{-1} + \sum_{i=1}^m [\Omega(i)^{-1} \otimes (W_i' W_i)] \right\}^{-1},$$

$$\text{vec}(B_1) = V_{B_1} \left\{ V_{B_0}^{-1} \text{vec}(B_0) + \sum_{i=1}^m [(\Omega(i) \otimes I_h)^{-1} \text{vec}(W_i' Y_i)] \right\},$$

$$V_{\tilde{b}_*}(i) = \left[V_{\tilde{b}_0}(i)^{-1} + \{ (a(i) \Omega(i)^{-1} a(i)') \otimes (X_i' X_i) \} \right]^{-1},$$

$$\tilde{b}_*(i) = V_{\tilde{b}_*}(i) \left[V_{\tilde{b}_0}(i)^{-1} \text{vec}(b_0(i)) + \{ (a(i) \Omega(i)^{-1}) \otimes X_i' \} \tilde{y} \right].$$

To sample the state variable $\tilde{S}_T = \{s_1, s_2, \dots, s_T\}'$, we employ the multi-move Gibbs sampling method.

3 Model Selection

In this section we discuss model selection by computing Bayes factors. In Bayesian framework, Bayes factors, defined as ratios of marginal likelihoods, are often used for comparing models, see Kass and Raftery (1995). To calculate the Bayes factors, there are several methods such as Chib (1995), Gelfand and Dey (1994), the Savage-Dickey density ratio (see Verdinelli and Wasserman, 1995), Laplace's approximation method (see Tierney and Kadane, 1986), and the Schwarz Bayesian information criterion (SBIC) approximation method (Schwarz (1978)). The SBIC method is easy to implement and does not require evaluation of the prior distribution, however, it gives only a rough approximation and does not produce the correct value even for large samples, see Kass and Raftery (1995) for details. In this paper, we choose Chib (1995)'s method for the marginal likelihood calculation that utilizes the output of the Gibbs sampler. Chib (1996), Chib (1998), and

Kim and Nelson (1999) use Chib's method for a Markov switching model. Suppose we have M different models \mathcal{M}_h for $h = 1, \dots, M$. The marginal likelihood for \mathcal{M}_h , $p(Y^T | \mathcal{M}_h)$, can be expressed as

$$p(Y^T | \mathcal{M}_h) = \int f(Y^T | \Theta, \mathcal{M}_h) p(\Theta | \mathcal{M}_h) d\Theta = \frac{f(Y^T | \Theta^*, \mathcal{M}_h) p(\Theta^* | \mathcal{M}_h)}{p(\Theta^* | Y^T, \mathcal{M}_h)}. \quad (14)$$

Or, on the computationally convenient logarithm scale,

$$\ln p(Y^T | \mathcal{M}_h) = \ln f(Y^T | \Theta^*, \mathcal{M}_h) + \ln p(\Theta^* | \mathcal{M}_h) - \ln p(\Theta^* | Y^T, \mathcal{M}_h), \quad (15)$$

where $f(Y^T | \Theta^*, \mathcal{M}_h)$ is the likelihood for \mathcal{M}_h evaluated at $\Theta^* = (B^*, \mathbf{b}^*, \Omega^*, \mathbf{q}^*)$, $\mathbf{b} = (b(1), \dots, b(m))$, $\Omega = (\Omega(1), \dots, \Omega(m))$, and $\mathbf{q} = (q_{11}, q_{12}, \dots, q_{mm})$, which is the Gibbs output or the posterior mean of Θ in \mathcal{M}_h , $p(\Theta^* | \mathcal{M}_h)$ is the prior density for \mathcal{M}_h evaluated at Θ^* , and $p(\Theta^* | Y^T, \mathcal{M}_h)$ is the posterior density for \mathcal{M}_h evaluated at Θ^* . We show how each of the quantities required for the marginal likelihood calculation is obtained in the rest of the discussion. The model index \mathcal{M} is suppressed for convenience.

The log likelihood function at Θ^* , $f(Y^T | \Theta^*)$, in equation (15) can be evaluated as

$$\ln f(Y^T | \Theta^*) = \sum_{t=1}^T \ln f(y_t | Y^{t-1}, \Theta^*), \quad (16)$$

where

$$\ln f(y_t | Y^{t-1}, \Theta^*) = \sum_{i=1}^m f(y_t | Y^{t-1}, \Theta^*, s_t = i) \times p(s_t = i | Y^{t-1}, \Theta^*). \quad (17)$$

The quantity $f(y_t | Y^{t-1}, \Theta^*, s_t = i)$ in equation (17) is the conditional density at $s_t = i$ and $\Theta = \Theta^*$, and $p(s_t = i | Y^{t-1}, \Theta^*)$ is the mass.

The log of prior density at Θ^* , $\ln p(\Theta^*)$, in (15) is given by

$$\ln p(\Theta^*) = \ln p(\text{vec}(B^*)) + \ln p(\mathbf{b}^*) + \ln p(\Omega^*) + \ln p(\mathbf{q}^*). \quad (18)$$

To estimate the marginal posterior density at Θ^* , $\ln p(\Theta^* | Y^T)$, in (15) using the conditional posteriors, first block Θ into L segments as $\Theta = (\theta'_1, \dots, \theta'_L)'$, and define $\varphi_{k-1} = (\theta'_1, \dots, \theta'_{k-1})$ and $\varphi^{k+1} = (\theta'_{k+1}, \dots, \theta'_L)$. Since

$$p(\Theta^* | y) = \prod_{k=1}^L p(\theta_k^* | y, \varphi_{k-1}^*)$$

and

$$p(\Theta^* | y, \Phi_{k-1}^*) = \int p(\theta_k^* | y, \Phi_{k-1}^*, \Phi^{k+1}) p(\Phi^{k+1} | y, \Phi_{k-1}^*) d\Phi^{k+1},$$

we can draw $\theta_k^{(g)}, \Phi^{k+1,(g)}$, where g indicates the Gibbs output $g = 1, \dots, G$, from $(\theta_k, \dots, \theta_L) = (\theta_k, \Phi^{k+1}) \sim p(\theta_k, \Phi^{k+1} | y, \Phi_{k-1}^*)$, and estimate $p(\theta_k^* | y, \Phi_{k-1}^*)$ as

$$\hat{p}(\theta_k^* | y, \Phi_{k-1}^*) = \frac{1}{G} \sum_{g=1}^G p(\theta_k^* | y, \Phi_{k-1}^*, \Phi^{k+1,(g)}) \rightarrow p(\theta_k^* | y, \Phi_{k-1}^*) \text{ as } G \rightarrow \infty,$$

where the superscript (g) refers to the g -th draw of the full Gibbs sampler. Thus, the log of the posterior density $p(\Theta^* | Y)$ can be estimated as

$$\ln \hat{p}(\Theta^* | Y^T) = \sum_{k=1}^L \ln \hat{p}(\theta_k^* | \Phi_{k-1}^*, Y^T) = \sum_{k=1}^L \ln \left\{ \frac{1}{G} \sum_{g=1}^G p(\theta_k^* | \Phi_{k-1}^*, \Phi^{k+1,(g)}, Y^T) \right\}. \quad (19)$$

In our MS-VECM case, the log of the posterior density evaluated at Θ^* is given by

$$\begin{aligned} \ln p(\Theta^* | Y^T) &= \ln p(\text{vec}(B^*) | Y^T) + \ln p(\mathbf{b}^* | \text{vec}(B^*), Y^T) \\ &\quad + \ln p(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, Y^T) + \ln p(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, Y^T) \end{aligned}$$

where

$$p(\text{vec}(B^*) | Y^T) = \int \int \int \int p(\text{vec}(B^*) | \mathbf{b}^*, \Omega^*, \mathbf{q}^*, \tilde{S}_T, Y^T) p(\mathbf{b}^*, \Omega^*, \mathbf{q}^*, \tilde{S}_T | Y^T) d\mathbf{b}^* d\Omega^* d\mathbf{q}^* d\tilde{S}_T \quad (20)$$

$$p(\mathbf{b}^* | \text{vec}(B^*), Y^T) = \int \int \int p(\mathbf{b}^* | \text{vec}(B^*), \Omega^*, \mathbf{q}^*, \tilde{S}_T, Y^T) p(\Omega^*, \mathbf{q}^*, \tilde{S}_T | \text{vec}(B^*), Y^T) d\Omega^* d\mathbf{q}^* d\tilde{S}_T \quad (21)$$

$$p(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, Y^T) = \int \int p(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, \mathbf{q}^*, \tilde{S}_T, Y^T) p(\mathbf{q}^*, \tilde{S}_T | \text{vec}(B^*), \mathbf{b}^*, Y^T) d\mathbf{q}^* d\tilde{S}_T \quad (22)$$

$$p(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, Y^T) = \int p(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, \tilde{S}_T, Y^T) p(\tilde{S}_T | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, Y^T) d\tilde{S}_T \quad (23)$$

The above decomposition of the posterior density suggests that $p(\text{vec}(B^*) | Y^T)$ can be calculated based on draws for the full Gibbs run, and $p(\mathbf{b}^* | \text{vec}(B^*), Y^T)$, $p(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, Y^T)$, and $p(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, Y^T)$ can be calculated based on draws from the reduced Gibbs runs as follows:

$$\hat{p}(\text{vec}(B^*) | Y^T) = \frac{1}{G} \sum_{g=1}^G p(\text{vec}(B^*) | \mathbf{b}^{(g)}, \Omega^{(g)}, \mathbf{q}^{(g)}, \tilde{S}_T^{(g)}, Y^T), \quad (24)$$

$$\hat{p}(\mathbf{b}^* | \text{vec}(B^*), Y^T) = \frac{1}{G} \sum_{g_1=1}^G p(\mathbf{b}^* | \text{vec}(B^*), \Omega^{(g_1)}, \mathbf{q}^{(g_1)}, \tilde{S}_T^{(g_1)}, Y^T), \quad (25)$$

$$\hat{p}(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, Y^T) = \frac{1}{G} \sum_{g_2=1}^G p(\Omega^* | \text{vec}(B^*), \mathbf{b}^*, \mathbf{q}^{(g_2)}, \tilde{S}_T^{(g_2)}, Y^T), \quad (26)$$

$$\hat{p}(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, Y^T) = \frac{1}{G} \sum_{g_3=1}^G p(\mathbf{q}^* | \text{vec}(B^*), \mathbf{b}^*, \Omega^*, \tilde{S}_T^{(g_3)}, Y^T), \quad (27)$$

where the superscript (g) refers to the g -th draw of the full Gibbs sampler and the superscript g_i , $i = 1, 2, 3$, refers to the g_i -th draw from the appropriate reduced Gibbs sampler. Thus, we need the usual G iterations for the full Gibbs sampler for (24) and additional $3 \times G$ iterations for the appropriate reduced Gibbs sampler for (25), (26), and (27) to calculate the log of the posterior density $\ln \hat{p}(\Theta^* | Y^T)$ in (18). From the estimated log of the posterior density $\ln \hat{p}(\Theta^* | Y^T)$ with the log of the likelihood $\ln f(Y^T | \Theta^*)$ in (16) and the log of the prior density $\ln p(\Theta^*)$ in (18), we can calculate the marginal likelihood in (15). In this way, we calculate the marginal likelihoods for all models under consideration. Then, we compute the Bayes factor for \mathcal{M}_h and \mathcal{M}_i as $BF_{hi} = \exp \{ \ln \hat{p}(Y^T | \mathcal{M}_h) - \ln \hat{p}(Y^T | \mathcal{M}_i) \}$, where $h = 2, \dots, M$, and \mathcal{M}_i ($i = 1$) is a base model for comparison. The posterior odds PO_{hi} can be obtained by $PO_{hi} = BF_{hi} \times p(\mathcal{M}_h) / p(\mathcal{M}_i)$, where $p(\mathcal{M}_h) / p(\mathcal{M}_i)$ is the prior odds. In this paper we set the prior odds as $p(\mathcal{M}_j) / p(\mathcal{M}_i) = 1$ for noninformative choice. Then, we can obtain posterior model probability for each model as $p(\mathcal{M}_h | Y) = PO_{hi} / \sum_{k=1}^M PO_{ki}$.

4 Monte Carlo Simulation

In this section we perform a simple Monte Carlo simulation to illustrate the performances of the marginal likelihood method for model selection described in the previous section. The data generating processes (DGPs) consist of two-state and two-variable MS-VECMs as follows:

$$\text{M1: } \Delta y_t = y_{t-1} \beta \alpha_1 + \mu_1 + \varepsilon_t, \varepsilon_t \sim iidN(0, \Omega_1)$$

$$\text{M2: } \Delta y_t = y_{t-1} \beta \alpha(s_t) + \mu(s_t) + \varepsilon_t, \varepsilon_t \sim N(0, \Omega_1)$$

$$\text{M3: } \Delta y_t = y_{t-1} \beta \alpha(s_t) + \mu(s_t) + \varepsilon_t, \varepsilon_t \sim N(0, \Omega(s_t))$$

where $\beta = (1, -1)'$, $\alpha_1 = \alpha(s_t = 1) = (-0.1, -0.1)$, $\alpha(s_t = 2) = (0, 0)$, $\mu_1 = \mu(s_t = 1) = (-0.1, -0.1)$, $\mu(s_t = 2) = (0.1, 0.1)$, $\Omega_1 = \Omega(s_t = 1) = 0.5I_2$, $\Omega(s_t = 2) = 0.1I_2$, the transition probability $(p_{11}, p_{22}) = (0.95, 0.95)$, and the sample size $T = \{100, 200, 500, 1000\}$. M1 is a linear VECM, while M2 and M3 are MS-VECMs. We compute the Bayes factors for these three model to calculate the posterior probability for each model. For prior parameters, we set $\tilde{b} = (1, 0)'$, $V_{\tilde{b}_0} = I_2$ in (8), $q_{11} \sim beta(u_{11}, u_{12}) = beta(9, 1)$ and $q_{22} \sim beta(u_{22}, u_{21}) = beta(9, 1)$ in (9), $\Omega_0(i) = I_2$ and $v_0(i) = 10$ for $i = 1$ or 2 in (6), $V_B = 10I_{k_T}$ and $B_0 = 0$ in (7) favoring the absence of cointegration. These values are assigned to ensure fairly large variances for representing prior ignorance. The simulation is replicated 100 times. For each iteration, the full Gibbs sampler is run with $G = 1,000$ and additional $3 \times G = 3,000$ draws for the reduced Gibbs sampler after 500 discarded to calculate the marginal likelihood by the Chib's method. The number of cointegration rank and the number of the lags is assumed to be known.

Table 1 summarizes the results of Monte Carlo simulations. The values in the columns are the average posterior probabilities. Increasing the sample size to 500 improves the performances considerably.

5 Conclusion

In this paper we show the Chib's method for marginal likelihood calculation for the MS-VECM to compute the Bayes factors to select an appropriate model. A simple Monte Carlo experiment shows that the Chib's method is useful to select the model. It is of interest that we compare the Chib's method to a bridge sampling method used by Jochmann and Koop (2015) by conducting Monte Carlo simulation for future research.

Table 1: Average posterior probabilities for model selection

Model	$T = 100$	$T = 200$	$T = 500$	$T = 1000$
True model: M1				
M1	0.811	0.867	0.922	0.960
M2	0.164	0.131	0.078	0.040
M3	0.025	0.002	0.000	0.000
True model: M2				
M1	0.066	0.011	0.000	0.000
M2	0.530	0.725	0.917	0.957
M3	0.403	0.264	0.083	0.043
True model: M3				
M1	0.020	0.002	0.000	0.000
M2	0.188	0.100	0.059	0.038
M3	0.792	0.898	0.941	0.962

References

- Chib, S. (1995). “Marginal likelihood from the Gibbs output.” *Journal of the American Statistical Association*, 90(432): 1313–1321.
- (1996). “Calculating posterior distributions and modal estimates in Markov mixture models.” *Journal of Econometrics*, 75(1): 79–97.
- (1998). “Estimation and comparison of multiple change-point models.” *Journal of Econometrics*, 86(2): 221–241.
- Gelfand, A. E. and Dey, D. K. (1994). “Bayesian model choice: asymptotics and exact calculations.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 501–514.
- Jochmann, M. and Koop, G. (2015). “Regime-switching cointegration.” *Studies in Nonlinear Dynamics & Econometrics*, 19(1): 35–48.
- Kass, R. E. and Raftery, A. E. (1995). “Bayes factors.” *Journal of the American Statistical Association*, 90(430): 773–795.

- Kim, C.-J. and Nelson, C. R. (1999). “Has the US economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle.” *Review of Economics and Statistics*, 81(4): 608–616.
- Koop, G., León-González, R., and Strachan, R. W. (2010). “Efficient posterior simulation for cointegrated models with priors on the cointegration space.” *Econometric Reviews*, 29(2): 224–242.
- Schwarz, G. (1978). “Estimating the dimension of a model.” *The Annals of Statistics*, 6(2): 461–464.
- Sugita, K. (2016). “Bayesian inference in Markov switching vector error correction model.” *Economics Bulletin*, 36(3): 1534–1546.
- Tierney, L. and Kadane, J. B. (1986). “Accurate approximations for posterior moments and marginal densities.” *Journal of the American Statistical Association*, 81(393): 82–86.
- Verdinelli, I. and Wasserman, L. (1995). “Computing Bayes factors using a generalization of the Savage-Dickey density ratio.” *Journal of the American Statistical Association*, 90(430): 614–618.