

# Measuring Business Cycle Turning Points in Japan with a Dynamic Markov Switching Factor Model

Toshiaki Watanabe

*In the dynamic factor model, a single unobserved factor common to some macroeconomic variables is defined as a composite index to measure business cycles. This model has recently been developed by combining with the regime switching model so that the mean growth of the index may shift depending on whether the economy is in a boom regime or in a recession regime. An advantage of this dynamic Markov switching factor model is that estimating the model by a Bayesian method produces the posterior probabilities that the economy is in the recession regime, which can be used to date the business cycle turning points. This article estimates the dynamic Markov switching factor model using some macroeconomic variables in Japan. The model comparison using the Bayes factor does not provide strong evidence that the mean growth of the index shifts, but the dynamic Markov switching factor model is found to produce the estimates of turning points close to the reference dates of the Economic and Social Research Institute in the Cabinet Office, unless only weakly correlated variables are used.*

Keywords: Business cycles; Factor model; Gibbs sampling; Marginal likelihood; Markov switching; Particle filter

JEL Classification: C11, C32, E32

Department of Economics, Tokyo Metropolitan University (E-mail: twatanab@bcomp.metro-u.ac.jp)

This paper was written while the author was a visiting scholar at the Institute for Monetary and Economic Studies (IMES), Bank of Japan. The author is deeply indebted to the staff there for encouragement and numerous suggestions. Thanks are also due to Susumu Suzuki in the Economic and Social Research Institute in the Cabinet Office for useful comments. Toshiyuki Matsuura and Kiyotaka Satoyoshi provided able research assistance. All remaining errors are clearly my own.

## I. Introduction

How should we measure business cycles? This problem has long attracted the attention of many macroeconomists and econometricians, and several methods have been proposed. A well-known method is the one based on dynamic factor models proposed by Stock and Watson (1989, 1991). They define the composite index of coincident economic indicators to measure the state of overall economic activity as a single unobserved factor common to several macroeconomic variables using a dynamic factor model. Because their model can be estimated by the maximum likelihood method via the Kalman filter, their composite index can be estimated by running the Kalman filter or smoother given the maximum likelihood estimates of the parameters.

Kim and Nelson (1998) extend the dynamic factor model of Stock and Watson (1989, 1991) so that the mean growth rate of the composite index may vary depending on whether the economy is in a recession regime or in a boom regime. They specify the mean growth rate of the index using the regime-switching model of Hamilton (1989). One advantage of their model is that it produces not only the composite index but also the probabilities that the economy is in the recession regime, which can be utilized to date the business cycle turning points. It is difficult, however, to evaluate the likelihood in their model, so that they apply a Bayesian method via the Gibbs sampler. Specifically, the model parameters, the latent factor, and the regime are sampled from their posterior distribution using the Gibbs sampler, and simulated draws are used for Bayesian posterior analysis.

This article applies the Kim and Nelson (1998) model to macroeconomic data in Japan. While several researchers such as Ohkusa (1992), Mori *et al.* (1993), Kanoh and Saito (1994), and Fukuda and Onodera (2001) have already applied the Stock and Watson (1989, 1991) model to the analysis of business cycles in Japan, there are few who have applied the Kim and Nelson (1998) model. The only exception is Kaufman (2000), who applies the Kim and Nelson (1998) model to eight countries, including Japan.<sup>1</sup> While she uses the quarterly data for real GDP, consumption, and investment, we use the monthly data selected from 10 macroeconomic variables (see Table 1) used by the Economic Planning Agency (EPA), which was reorganized as the Economic and Social Research Institute (ESRI) in the Cabinet Office after January 2001, to construct its composite index.

Following Kim and Nelson (1998), we estimate the composite index and the probabilities that the economy is in recession as well as the model parameters using a Bayesian method via the Gibbs sampler. We also analyze whether the regime-shift occurs in the mean growth rate of the composite index by comparing the Kim and Nelson (1998) model with the Stock and Watson (1989, 1991) model. Classical test statistics such as the likelihood ratio statistics are not directly applicable to this analysis (see Hansen [1992] and Garcia [1998]). In a Bayesian framework, model comparisons are conducted based on the posterior odds—that is, the ratio of the marginal likelihood—which does not cause any problem in analyzing whether the

---

1. Kasuya and Shinki (2001) have applied the Kim and Nelson (1998) model to forecasting the turning points of the consumer price index in Japan.

**Table 1 Variables Used by the Economic and Social Research Institute to Construct Its Composite Index**

Definition of Variables

|    |        |   |
|----|--------|---|
| 1  | IIP95M | Index of raw materials consumption, mfg.        |
| 2  | IIP95O | Index of operating rate, mfg.                   |
| 3  | HWINMF | Index of non-scheduled hours worked, mfg.       |
| 4  | ESRAO  | Ratio of job offers to applicants               |
| 5  | SDS    | Sales of department stores                      |
| 6  | CELL9  | Electric power consumption of large users       |
| 7  | IIP95S | Index of producers' shipments, investment goods |
| 8  | SCI95  | Index of wholesale sales                        |
| 9  | SMSALE | Sales of small and medium-sized companies       |
| 10 | IIP95P | Index of industrial production, mining and mfg. |
| 11 | ZBOAS  | Business profit, all industries                 |

Note: ZBOAS is quarterly data, and the others are monthly data.

regime-shift occurs or not. We adopt this method and calculate the marginal likelihood following the method proposed by Chib (1995). A diagnostic checking is also conducted.

The model comparison using the Bayes factor does not provide strong evidence that the Kim and Nelson (1998) model is favored over the Stock and Watson (1989, 1991) model. In addition, no major differences between the composite indices produced by the two models are found. On the other hand, the Kim and Nelson (1998) model produces the estimates of turning points close to the reference dates of the ESRI in the Cabinet Office unless only weakly correlated variables are used.

The rest of this article is organized as follows. Section II explains the Kim and Nelson (1998) model and a Bayesian method for analyzing this model. Section III fits the model to macroeconomic data in Japan and summarizes the results. Conclusions are given in Section IV.

## II. Econometric Methodology

### A. Dynamic Factor Models

Since our analysis is based on the dynamic factor models proposed by Stock and Watson (1988, 1991) and developed by Kim and Nelson (1998), we start with a brief review of these models.

Suppose that we have data on  $n$  macroeconomic variables from period 0 to  $T$ . Let  $\Delta Y_{it}$  ( $i = 1, \dots, n$ ;  $t = 1, \dots, T$ ) denote the growth rate of the  $i$ -th macroeconomic variable defined as the first difference of the log of the  $i$ -th variable at time  $t$ . In the simplest version of the dynamic factor models,  $\Delta Y_{it}$  is specified as follows.

$$\Delta Y_{it} = D_i + \lambda_i \Delta C_t + e_{it},$$

where  $D_i$  and  $\lambda_i$  are constants,  $\Delta C_t$  is the component common to all variables, which is interpreted as the first difference of the composite index of coincident economic indicators  $C_t$ , and  $e_{it}$  is the idiosyncratic component of the  $i$ -th variable.

$\Delta Y_{it}$  may depend on not only the current value of  $\Delta C_t$  but also the past values. To allow for this possibility, we use the following specification.

$$\Delta Y_{it} = D_i + \lambda_{i0}\Delta C_t + \lambda_{i1}\Delta C_{t-1} + \dots + \lambda_{ir_i}\Delta C_{t-r_i} + e_{it}. \quad (1)$$

The idiosyncratic component  $e_{it}$  is assumed to follow an autoregressive (AR) process with mean zero, i.e.,

$$e_{it} = \psi_{i,1}e_{i,t-1} + \dots + \psi_{i,q_i}e_{i,t-q_i} + \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d. N(0, \sigma_i^2), \quad (2)$$

where error term  $\epsilon_{it}$  is assumed to follow a serially independent normal distribution.

The difference between the Stock and Watson (1989, 1991) and the Kim and Nelson (1998) models is the specification of the common factor  $\Delta C_t$ . While Stock and Watson (1989, 1991) specify  $\Delta C_t$  as a simple AR process, Kim and Nelson (1998) extend it so that the mean growth of the composite index may shift depending on whether the economy is in a recession or in a boom as follows.

$$\begin{aligned} \Delta C_t = & \delta + \mu_{s_t} + \phi_1(\Delta C_{t-1} - \delta - \mu_{s_{t-1}}) + \dots \\ & + \phi_p(\Delta C_{t-p} - \delta - \mu_{s_{t-p}}) + \nu_t, \quad \nu_t \sim i.i.d. N(0, 1), \end{aligned} \quad (3)$$

where  $\delta$  is the long-run growth of the composite index, which is constant, and  $\mu_{s_t}$  is the deviation from that long-run growth, which may shift depending on whether the economy is in a recession or in a boom. Error term  $\nu_t$  is assumed to follow a serially independent normal distribution. The variance of  $\nu_t$  is normalized to unity for identification of the model. Error terms  $\nu_t$  and  $\epsilon_{it}$  are assumed to be mutually independent for all  $i, t, s$ .

Using the variable  $S_t$ , which takes zero when the economy is in the recession regime and one when the economy is in the boom regime, Kim and Nelson (1998) specify  $\mu_{s_t}$  as follows.

$$\mu_{s_t} = \mu_0 + \mu_1 S_t, \quad \mu_0 < 0, \mu_1 > 0. \quad (4)$$

The reason to assume that  $\mu_1 > 0$  is that the mean growth of the composite index will be greater in a boom regime ( $S_t = 1$ ) than that in a recession regime ( $S_t = 0$ ). Although Kim and Nelson (1998) do not assume that  $\mu_0 < 0$ , we assume it because, otherwise, the average of  $\mu_{s_t}$ , which is the deviation from the long-run growth, would be positive. They assume that  $S_t$  follows a Markov process with transition probabilities

$$\begin{aligned} P(S_t = 1 | S_{t-1} = 1) &= \pi_{11}, \quad P(S_t = 0 | S_{t-1} = 1) = 1 - \pi_{11}, \\ P(S_t = 0 | S_{t-1} = 0) &= \pi_{00}, \quad P(S_t = 1 | S_{t-1} = 0) = 1 - \pi_{00}. \end{aligned} \quad (5)$$

Equations (1)–(5) constitute the Kim and Nelson (1998) model, which collapses to the Stock and Watson (1989, 1991) model if  $\mu_{s_t} = 0$  in equation (3). If  $\Delta C_t$  is an observed macroeconomic variable instead of the growth of the composite index, equations (3)–(5) constitute the regime switching model proposed by Hamilton (1989). Therefore, the Kim and Nelson (1998) model is a synthesis of the Stock and Watson (1989, 1991) and the Hamilton (1989) models. The Stock and Watson (1989, 1991) model produces the estimates of the composite index  $C_t$  but does not produce the probabilities of a recession ( $S_t = 0$ ). The regime switching model of Hamilton (1989) produces the probability of a recession, which can be used to date the business cycle turning points, but does not produce the estimates of the composite index. An advantage of the Kim and Nelson (1998) model is that it produces both the estimates of the composite index and probabilities of a recession.

$D_t$  in equation (1) and  $\delta$  in equation (3) are usually removed because they are not identified. To do so, define

$$\begin{aligned}\Delta c_t &= \Delta C_t - E(\Delta C_t) = \Delta C_t - \delta, \\ \Delta y_{it} &= \Delta Y_{it} - E(\Delta Y_{it}) = \Delta Y_{it} - D_t - (\lambda_{i0} + \dots + \lambda_{iri})\delta.\end{aligned}$$

Then, equations (1) and (3) can be written as

$$\Delta y_{it} = \lambda_{i0}\Delta c_t + \lambda_{i1}\Delta c_{t-1} + \dots + \lambda_{iri}\Delta c_{t-ri} + e_{it}, \quad (1')$$

$$\begin{aligned}\Delta c_t &= \mu_{s_t} + \phi_1(\Delta c_{t-1} - \mu_{s_{t-1}}) + \dots \\ &+ \phi_p(\Delta c_{t-p} - \mu_{s_{t-p}}) + v_t, \quad v_t \sim i.i.d. N(0, 1).\end{aligned} \quad (3')$$

The demeaned growth rate  $\Delta Y_{it} - \overline{\Delta Y}_i$ , where  $\overline{\Delta Y}_i$  is the sample average of  $\Delta Y_{i1}, \dots, \Delta Y_{iT}$ , is used for  $\Delta y_{it}$ . In what follows, we consider the model that consists of equations (1'), (2), (3'), (4), and (5) as the Kim and Nelson (1998) model.

## B. Bayesian Estimation via the Gibbs Sampler

The Stock and Watson (1989, 1991) model can be represented by a linear Gaussian state space model. The likelihood of the linear Gaussian state space model can be evaluated by executing the Kalman filter. The likelihood of the Hamilton (1989) model can also be evaluated by executing the filter proposed by Hamilton (1989). Hence, the parameters in these models can be estimated using the conventional maximum likelihood method. The likelihood of the Kim and Nelson (1998) model cannot, however, be evaluated analytically, so that the parameters cannot be estimated using the maximum likelihood method. The estimation of the Kim and Nelson (1998) model requires other estimation methods. Kim and Nelson (1998) apply a Bayesian method via the Gibbs sampler.

Let  $\theta$  denote the set of the unknown parameters. The conventional Bayesian method proceeds as follows.

1. Set the prior distribution  $f(\theta)$ , which is the distribution the researcher has in mind before observing the data.

2. Convert the prior distribution to the posterior distribution  $f(\theta|\text{data})$ , which is the distribution conditional on the data, using the Bayes theorem

$$f(\theta|\text{data}) = \frac{f(\text{data}|\theta)f(\theta)}{\int f(\text{data}|\theta)f(\theta)d\theta}. \quad (6)$$

3. Estimate the parameters  $\theta$  using the posterior distribution.

Notice that  $f(\text{data}|\theta)$  on the right-hand side of the Bayes theorem (6) is the likelihood. Therefore, the conventional Bayesian method cannot be applied to the models such as the Kim and Nelson (1998) model whose likelihood cannot be obtained analytically. In such models, steps 2 and 3 above are replaced by

- 2' Sample  $\theta$  from the posterior distribution  $f(\theta|\text{data})$ .  
3' Estimate the parameters  $\theta$  using the draws sampled in 2'.

Some readers may think it impossible to sample from the posterior distribution that cannot be obtained analytically. It is the Gibbs sampler that makes it possible.

The Gibbs sampler is a Monte Carlo method for sampling from a joint distribution using conditional distributions. Suppose that it is impossible to obtain the joint posterior distribution  $f(\theta|\text{data})$  analytically using the Bayes theorem, but  $\theta$  can be divided into  $k$  partitions  $(\theta_1, \dots, \theta_k)$  where  $\theta_i$  may be a scalar or a vector, such that, for all  $i = 1, 2, \dots, k$ , it is possible to obtain conditional distribution  $f(\theta_i|\{\theta_j\}_{j \neq i}, \text{data})$  analytically and sample  $\theta_i$  from it by some methods. The Gibbs sampler is used in such cases and works as follows. Starting from an arbitrary set of initial value  $(\theta_1^{(0)}, \dots, \theta_k^{(0)})$ , we draw  $\theta_1^{(1)}$  from  $f(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}, \text{data})$ ,  $\theta_2^{(1)}$  from  $f(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}, \text{data})$ , and so on up to  $\theta_k^{(1)}$  from  $f(\theta_k|\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{k-1}^{(1)}, \text{data})$ . Let us call this procedure one iteration. After  $l$  such iterations, we obtain  $(\theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_k^{(l)})$ . Under mild conditions, it converges in distribution to be a set of random variables from  $f(\theta_1, \theta_2, \dots, \theta_k|\text{data})$  as  $l \rightarrow \infty$ . Therefore, for a sufficiently large  $M$ ,  $(\theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_k^{(l)})(l = M + 1, M + 2, \dots, M + N)$  can approximately be regarded as a sample from the joint posterior distribution  $f(\theta_1, \theta_2, \dots, \theta_k|\text{data})$ . Hence, the first  $M$  draws, called "burn-in," are discarded and the last  $N$  draws are used for parameter estimation. For instance, the expectation of a function of the parameters,  $g(\theta_1, \theta_2, \dots, \theta_k)$ , is estimated by the sample average

$$E[g(\theta_1, \theta_2, \dots, \theta_k)] \approx \frac{1}{N} \sum_{l=M+1}^{M+N} g(\theta_1^{(l)}, \theta_2^{(l)}, \dots, \theta_k^{(l)}). \quad (7)$$

The unknown parameters in the Kim and Nelson (1998) model that consist of equations (1'), (2), (3'), (4), and (5) are  $\lambda_i = [\lambda_{i0}, \dots, \lambda_{iri}](i = 1, \dots, n)$ ,  $\psi_i = [\psi_{i1}, \dots, \psi_{iq_i}](i = 1, \dots, n)$ ,  $\sigma_i^2(i = 1, \dots, n)$ ,  $\phi = [\phi_1, \dots, \phi_p]$ ,  $\mu = (\mu_0, \mu_1)$ ,  $\pi = (\pi_{00}, \pi_{11})$ . As well as these parameters, latent variables  $\Delta c_T = [\Delta c_1, \dots, \Delta c_T]$  and  $S_T = [S_1, \dots, S_T]$  are also treated as if they were unknown parameters. Then, all we have to do to sample from the joint posterior distribution using the Gibbs sampler is to sample from the following conditional distributions.

$$f(\boldsymbol{\lambda}_i | \boldsymbol{\theta}_{\lambda_i}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \quad (8)$$

$$f(\boldsymbol{\psi}_i | \boldsymbol{\theta}_{\psi_i}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \quad (9)$$

$$f(\sigma_i^2 | \boldsymbol{\theta}_{\sigma_i^2}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \quad (10)$$

$$f(\phi | \boldsymbol{\theta}_\phi, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad (11)$$

$$f(\boldsymbol{\mu} | \boldsymbol{\theta}_\mu, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad (12)$$

$$f(\boldsymbol{\pi} | \boldsymbol{\theta}_\pi, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad (13)$$

$$f(\mathbf{S}_T | \boldsymbol{\theta}, \Delta \mathbf{c}_T, \Delta \mathbf{y}_T), \quad (14)$$

$$f(\Delta \mathbf{c}_T | \boldsymbol{\theta}, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad (15)$$

where  $\boldsymbol{\theta}_\omega$  represents the set of all parameters except  $\omega$ , which does not include latent variables,  $\Delta \mathbf{y}_{iT} = [\Delta y_{i1}, \dots, \Delta y_{iT}]$ , and  $\Delta \mathbf{y}_T = [\Delta \mathbf{y}_{1T}, \dots, \Delta \mathbf{y}_{nT}]$ .

As for the prior distributions of the unknown parameters, we may use any distributions but it is convenient to assume the following distributions.

$$\boldsymbol{\lambda}_1 \sim N(M_{\lambda_1}^{(0)}, \sum_{\lambda_1}^{(0)}) I[\lambda_{10} > 0], \quad (16)$$

$$\boldsymbol{\lambda}_i \sim N(M_{\lambda_i}^{(0)}, \sum_{\lambda_i}^{(0)}), \quad i = 2, \dots, n, \quad (17)$$

$$\boldsymbol{\psi}_i \sim N(M_{\psi_i}^{(0)}, \sum_{\psi_i}^{(0)}) I_{S(\psi_i)}, \quad i = 1, \dots, n, \quad (18)$$

$$\phi \sim N(M_\phi^{(0)}, \sum_\phi^{(0)}) I_{S(\phi)}, \quad (19)$$

$$\boldsymbol{\mu} \sim N(M_\mu^{(0)}, \sum_\mu^{(0)}) I[\mu_0 < 0, \mu_1 > 0], \quad (20)$$

$$\sigma_i^2 \sim IG(\nu^{(0)}/2, \delta^{(0)}/2), \quad i = 1, \dots, n, \quad (21)$$

$$\boldsymbol{\pi}_{11} \sim \text{beta}(u_{11}^{(0)}, u_{10}^{(0)}), \quad \boldsymbol{\pi}_{00} \sim \text{beta}(u_{00}^{(0)}, u_{01}^{(0)}), \quad (22)$$

where  $I[\cdot]$  is the indicator function that takes one if the condition in the parentheses is satisfied and zero otherwise, and  $I_{S(\psi_i)}$  (or  $I_{S(\phi)}$ ) is the indicator function that takes one if the roots of the polynomial  $1 - \psi_{i1}L - \dots - \psi_{iq_i}L^{q_i} = 0$  (or  $1 - \phi_1L - \dots - \phi_pL^p = 0$ ) lie outside the unit circle and zero otherwise. At least, one parameter among  $\lambda_{ij}$  ( $i = 1, \dots, n; j = 1, \dots, n$ ) is assumed to be positive for identification of the model. Here, we assume that  $\lambda_{10} > 0$ . Hence, the prior of  $\lambda_1$  is set to be the truncated normal whose density is zero unless  $\lambda_{10} > 0$ , and that of  $\lambda_i$  ( $i = 2, \dots, n$ ) is set to be the normal. Under the assumption that equations (2) and (3) are stationary, the priors of  $\boldsymbol{\psi}_i$ ,  $\phi$  are set to be the truncated normal whose density is zero outside the stationary region. The prior of  $\boldsymbol{\mu}$  is the truncated normal whose density is zero unless  $\mu_0 < 0$  and  $\mu_1 > 0$ . The prior of  $\sigma_i^2$  is set to be the inverted gamma, which means that  $1/\sigma_i^2$  follows the gamma distribution. The priors of  $\boldsymbol{\pi}_{11}$  and  $\boldsymbol{\pi}_{00}$  are set to be beta distributions.

Under these priors, it is straightforward to obtain the conditional distributions (8)–(13), which have the same forms as the priors (16)–(22), and sample from those distributions (see Appendix 1).

The condition of (14) includes  $\Delta \mathbf{c}_T$ . Given  $\Delta \mathbf{c}_T$ , equations (3'), (4), and (5) constitute the regime switching model proposed by Hamilton (1989). Thus,

sampling  $\mathcal{S}_T$  from (14) can be conducted using the Hamilton (1989) filter. Applying the Hamilton (1989) filter to the model that consists of equations (3'), (4), and (5) produces  $p(S_t|\Delta c_t)$  and  $p(S_t|\Delta c_{t-1})$  for  $t = 1, \dots, T$ . Then, starting with  $S_T$  sampled from  $p(S_T|\Delta c_T)$ , we can proceed backwards in time. Specifically, given  $S_{t+1}$ ,  $S_t$  is generated using the probability

$$p(S_t|\Delta c_t, S_{t+1}) = \frac{p(S_{t+1}|S_t)p(S_t|\Delta c_t)}{p(S_{t+1}|\Delta c_t)},$$

where  $p(S_{t+1}|S_t)$  is the transition probability given by (5), and  $p(S_t|\Delta c_t)$  and  $p(S_{t+1}|\Delta c_t)$  are obtained from the Hamilton (1989) filter.

The condition of (15) includes  $\mathcal{S}_T$ . Once  $\mathcal{S}_T$  are given, the Kim and Nelson (1998) model can be represented by a linear Gaussian state space model. Suppose that  $n = 5$ ,  $p = 3$ , and  $r_i = q_i = 1$  ( $i = 1, \dots, n$ ). Let  $\Delta y_{it}^* = \Delta y_{it} - \psi_{i1}\Delta y_{i,t-1}$  ( $i = 1, \dots, n$ ) and  $\Delta y_t^* = [\Delta y_{1t}^*, \dots, \Delta y_{nt}^*]$ . Then, the Kim and Nelson (1998) model may be represented as

$$\Delta y_t^* = \Lambda z_t + \epsilon_t, \quad \epsilon_t \sim i.i.d. N(0, \Sigma_\epsilon), \quad (23)$$

$$z_t = M_{s_t} + \Phi z_{t-1} + v_t, \quad v_t \sim i.i.d. N(0, \Sigma_v). \quad (24)$$

Then,  $z_t$ ,  $\Lambda$ ,  $M_{s_t}$ ,  $\Phi$ ,  $\Sigma_\epsilon$ , and  $\Sigma_v$  are given by

$$z_t = [\Delta c_t, \Delta c_{t-1}, \Delta c_{t-2}],$$

$$\Lambda = \begin{bmatrix} \lambda_{10} & -\lambda_{10}\psi_{11} + \lambda_{11} & -\lambda_{11}\psi_{11} \\ \lambda_{20} & -\lambda_{20}\psi_{21} + \lambda_{21} & -\lambda_{21}\psi_{21} \\ \lambda_{30} & -\lambda_{30}\psi_{31} + \lambda_{31} & -\lambda_{31}\psi_{31} \\ \lambda_{40} & -\lambda_{40}\psi_{41} + \lambda_{41} & -\lambda_{41}\psi_{41} \\ \lambda_{50} & -\lambda_{50}\psi_{51} + \lambda_{51} & -\lambda_{51}\psi_{51} \end{bmatrix},$$

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

$$M_{s_t} = [\phi(L)\mu_{s_t}, 0, 0],$$

$$v_t = [v_t, 0, 0],$$

$$\Sigma_\epsilon = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_4^2 & 0 \\ 0 & 0 & 0 & 0 & \sigma_5^2 \end{bmatrix},$$



$$\Sigma_v = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where  $\phi(L)\mu_{s_t} = \mu_{s_t} - \phi_1\mu_{s_{t-1}} - \phi_2\mu_{s_{t-2}} - \phi_3\mu_{s_{t-3}}$ .

Since equations (23) and (24) constitute the linear Gaussian state space model, it is straightforward to sample  $\Delta c_T$  from (15) using the Kalman filter and smoother. Once  $\Delta c_T$  is obtained, it can be transformed into the composite index  $C_T = [C_1, \dots, C_T]$  as

$$C_t = \Delta c_t + C_{t-1} + \delta,$$

where  $\delta$  is the long-run growth of the composite index, which can be estimated using the steady-state Kalman gain (see Kim and Nelson [1998, 1999]).

## C. Model Comparison

### 1. Marginal likelihood

It is important to examine whether the mean growth of the composite index shifts depending on whether the economy is in the boom regime or the recession regime by comparing the Kim and Nelson (1998) model with the Stock and Watson (1989, 1991) model. Kaufman (2000) proposes a method for comparing these two models. As is a usual practice in Bayesian model comparison, his method is based on the posterior odds ratio. The posterior odds ratio between model  $i$ ,  $M_i$ , and model  $j$ ,  $M_j$ , is given by

$$POR = \frac{f(M_i|\Delta\mathbf{y}_T)}{f(M_j|\Delta\mathbf{y}_T)} = \frac{f(\Delta\mathbf{y}_T|M_i) f(M_i)}{f(\Delta\mathbf{y}_T|M_j) f(M_j)},$$

where  $f(\Delta\mathbf{y}_T|M_i)/f(\Delta\mathbf{y}_T|M_j)$  and  $f(M_i)/f(M_j)$  are called the Bayes factor and the prior odds ratio, respectively. If  $POR$  is greater than one,  $M_i$  is favored over  $M_j$ .

The prior odds ratio is usually set to be one, so that the posterior odds ratio is equal to the Bayes factor. To evaluate the Bayes factor, we must calculate  $f(\Delta\mathbf{y}_T|M_i)$  and  $f(\Delta\mathbf{y}_T|M_j)$ , called marginal likelihoods. The log of the marginal likelihood of model  $M_i$  can be written as

$$\begin{aligned} \ln f(\Delta\mathbf{y}_T|M_i) &= \ln f(\Delta\mathbf{y}_T|M_i, \boldsymbol{\theta}_i) + \ln f(\boldsymbol{\theta}_i|M_i) \\ &\quad - \ln f(\boldsymbol{\theta}_i|M_i, \Delta\mathbf{y}_T), \end{aligned} \tag{25}$$

where  $\boldsymbol{\theta}_i$  is the set of unknown parameters for model  $M_i$ ,  $f(\Delta\mathbf{y}_T|M_i, \boldsymbol{\theta}_i)$  is the likelihood,  $f(\boldsymbol{\theta}_i|M_i)$  is the prior density, and  $f(\boldsymbol{\theta}_i|M_i, \Delta\mathbf{y}_T)$  is the posterior density. The above identity holds for any value of  $\boldsymbol{\theta}_i$ , but Chib (1995) proposes to set  $\boldsymbol{\theta}_i$  equal to its posterior mean  $\hat{\boldsymbol{\theta}}_i$  calculated using the Gibbs draws. In what follows, subscript  $i$  and  $M_i$  are omitted.

The Kim and Nelson (1998) model is more general than the Stock and Watson (1989, 1991) model in the sense that setting  $\mu_{s_t} = 0$  in the former model leads to the

latter model. Hence, the likelihood that the former model cannot be smaller than that of the latter model. Notice, however, that the marginal likelihood of the Kim and Nelson (1998) model may be smaller than that of the Stock and Watson (1989, 1991) model.

## 2. Prior density

If prior distributions are given by (16)–(22) where  $\lambda_i (i = 1, \dots, n)$ ,  $\psi_i (i = 1, \dots, n)$ ,  $\sigma_i^2 (i = 1, \dots, n)$ ,  $\phi$ ,  $\mu$ ,  $\pi$  are mutually independent, we have

$$f(\hat{\theta}) = f(\hat{\lambda}_1) \times \dots \times f(\hat{\lambda}_n) \times f(\hat{\psi}_1) \times \dots \times f(\hat{\psi}_n) \times f(\hat{\sigma}_1^2) \times \dots \times f(\hat{\sigma}_n^2) \times f(\hat{\phi}) \times f(\hat{\mu}) \times f(\hat{\pi}). \quad (26)$$

It is straightforward to evaluate  $f(\hat{\lambda}_i) (i = 2, \dots, n)$ ,  $f(\hat{\mu})$ , and  $f(\hat{\pi})$ . It may be difficult to evaluate the other terms in (26) analytically, because truncation may make it difficult to calculate the normalizing constant. Even in such cases, there are some numerical methods available to evaluate the normalizing constant (see Chen *et al.* [2000]).

## 3. Posterior density

Kaufman (2000) uses the method proposed by Chib (1995) to evaluate the posterior density. The posterior density is written as

$$f(\hat{\theta}|\Delta\mathbf{y}_T) = f(\hat{\pi}|\Delta\mathbf{y}_T) \times f(\hat{\lambda}_1, \dots, \hat{\lambda}_n|\hat{\pi}, \Delta\mathbf{y}_T) \times \dots \times f(\hat{\mu}|\hat{\theta}_{\mu}, \Delta\mathbf{y}_T), \quad (27)$$

and evaluates each term separately using the Gibbs sampler.

The first term can be written as

$$f(\hat{\pi}|\Delta\mathbf{y}_T) = \int f(\hat{\pi}|\mathbf{S}_T, \Delta\mathbf{y}_T) f(\mathbf{S}_T|\Delta\mathbf{y}_T) d\mathbf{S}_T. \quad (28)$$

The Gibbs sampler explained above produces draws from  $f(\mathbf{S}_T|\Delta\mathbf{y}_T)$ . All we have to do is to sample from the conditional distributions (8)–(15) sequentially. Given  $M$  draws  $(\mathbf{S}_T^{(1)}, \dots, \mathbf{S}_T^{(M)})$  from  $f(\mathbf{S}_T|\Delta\mathbf{y}_T)$ , equation (28) can be estimated by

$$f(\hat{\pi}|\Delta\mathbf{y}_T) \approx \frac{1}{M} \sum_{m=1}^M f(\hat{\pi}|\mathbf{S}_T^{(m)}, \Delta\mathbf{y}_T). \quad (29)$$

This is not true for the other terms, because some parameters included in the conditions are fixed at their posterior means. For example, the second term can be written as

$$\begin{aligned} & f(\hat{\lambda}_1, \dots, \hat{\lambda}_n|\hat{\pi}, \Delta\mathbf{y}_T) \\ &= \int f(\hat{\lambda}_1, \dots, \hat{\lambda}_n|\hat{\pi}, \sigma_1^2, \dots, \sigma_n^2, \Delta\mathbf{c}_T, \Delta\mathbf{y}_T) \\ & \quad \times f(\sigma_1^2, \dots, \sigma_n^2, \Delta\mathbf{c}_T|\hat{\pi}, \Delta\mathbf{y}_T) d\sigma_1^2 \cdots d\sigma_n^2 d\Delta\mathbf{c}_T. \end{aligned} \quad (30)$$

To sample from  $f(\sigma_1^2, \dots, \sigma_n^2, \Delta \mathbf{c}_T | \hat{\boldsymbol{\pi}}, \Delta \mathbf{y}_T)$  using the Gibbs sampler, we must sample from the following conditional distributions.

$$\begin{aligned} & f(\boldsymbol{\lambda}_i | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{(\boldsymbol{\lambda}_i, \boldsymbol{\pi})}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \\ & f(\boldsymbol{\psi}_i | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{(\boldsymbol{\psi}_i, \boldsymbol{\pi})}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \\ & f(\sigma_i^2 | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{(\sigma_i^2, \boldsymbol{\pi})}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \quad i = 1, \dots, n, \\ & f(\phi | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{(\phi, \boldsymbol{\pi})}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \\ & f(\boldsymbol{\mu} | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{(\boldsymbol{\mu}, \boldsymbol{\pi})}, \Delta \mathbf{c}_T, \mathbf{S}_T, \Delta \mathbf{y}_T), \\ & f(\mathbf{S}_T | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{\boldsymbol{\pi}}, \Delta \tilde{\mathbf{c}}_T, \Delta \tilde{\mathbf{y}}_T), \\ & f(\Delta \mathbf{c}_T | \hat{\boldsymbol{\pi}}, \boldsymbol{\theta}_{\boldsymbol{\pi}}, \mathbf{S}_T, \Delta \mathbf{y}_T). \end{aligned}$$

Given  $M$  draws from  $f(\sigma_1^2, \dots, \sigma_n^2, \Delta \mathbf{c}_T | \hat{\boldsymbol{\pi}}, \Delta \mathbf{y}_T)$ , equation (30) can be estimated by

$$\begin{aligned} & f(\hat{\boldsymbol{\lambda}}_1, \dots, \hat{\boldsymbol{\lambda}}_n | \hat{\boldsymbol{\pi}}, \Delta \mathbf{y}_T) \\ & \approx \frac{1}{M} \sum_{m=1}^M f(\hat{\boldsymbol{\lambda}}_1, \dots, \hat{\boldsymbol{\lambda}}_n | \hat{\boldsymbol{\pi}}, \sigma_i^{2(m)}, \dots, \sigma_n^{2(m)}, \Delta \mathbf{c}_T^{(m)}, \Delta \mathbf{y}_T). \end{aligned}$$

The other terms in equation (27) can be evaluated similarly.

#### 4. Likelihood

Kaufman (2000) uses a particle filter to evaluate the likelihood. A particle filter is the algorithm to sample from the filtering density  $f(\mathbf{z}_t, S_t | \Delta \mathbf{y}_t, \boldsymbol{\theta})$  sequentially starting from  $t = 0$  (see Pitt and Shephard [1999]), where  $\mathbf{z}_t$  is the state variable that appears in equations (23) and (24).

Suppose that we have  $M$  draws  $\{\mathbf{z}_{t-1}^{(m)}, S_{t-1}^{(m)}\} (m = 1, \dots, M)$  sampled from the density  $f(\mathbf{z}_{t-1}, S_{t-1} | \Delta \mathbf{y}_{t-1}, \boldsymbol{\theta})$ . Then, we can sample  $\{\mathbf{z}_t^{(m)}, S_t^{(m)}\}$  from the density  $f(\mathbf{z}_t, S_t | \Delta \mathbf{y}_t, \boldsymbol{\theta})$  as follows (see Appendix 2 for details).

Step 1. Select a  $(S_t, m)$  from  $2 \times M$  combinations ( $S_t = 0, 1; m = 1, \dots, M$ ) with probability proportional to

$$\eta S_t, m = |F|^{-1/2} \exp\left(-\frac{1}{2} \mathbf{e}'_t F^{-1} \mathbf{e}_t\right) p(S_t | S_{t-1}^{(m)}),$$

where

$$\begin{aligned} \mathbf{e}_t &= \Delta \mathbf{y}_t - \Lambda(M_{S_t} + \Phi \mathbf{z}_{t-1}^{(m)}), \\ F &= \Lambda \sum_v \Lambda' + \sum_v \epsilon. \end{aligned}$$

Step 2. Using the  $(S_t, m)$  selected in Step 1, sample from  $N(\boldsymbol{\mu}_{t|t}^{(m)}, \sum_{t|t}^{(m)})$ , where

$$\begin{aligned} \boldsymbol{\mu}_{t|t}^{(m)} &= M_{S_t} + \Phi \mathbf{z}_{t-1}^{(m)} + \sum_v \Lambda' F^{-1} \mathbf{e}_t, \\ \sum_{t|t}^{(m)} &= \sum_v - \sum_v \Lambda' F^{-1} \Lambda \sum_v. \end{aligned}$$

The likelihood can be expressed as

$$L = \prod_{t=1}^T f(\Delta y_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta}),$$

where

$$\begin{aligned} & f(\Delta y_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta}) \\ &= \int f(\Delta y_{t+1} | z_{t+1}, \boldsymbol{\theta}) f(z_{t+1} | z_t, S_{t+1}, \boldsymbol{\theta}) \\ & \quad \times p(S_{t+1} | S_t, \boldsymbol{\theta}) f(z_t, S_t | \Delta \mathbf{y}_t, \boldsymbol{\theta}) dz_t dS_t dS_{t+1} dz_{t+1}. \end{aligned} \quad (31)$$

Given  $M$  draws  $\{z_t^{(m)}, S_t^{(m)}\} (m = 1, \dots, M)$  from the density  $f(z_t, S_t | \Delta \mathbf{y}_t, \boldsymbol{\theta})$  using the above particle filter, we can evaluate  $f(\Delta y_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta})$  as follows.

Step 3. Sample  $S_{t+1}^{(m)}$  using the transition probability  $p(S_{t+1}^{(m)} | S_t)$ .

Step 4. Using  $S_{t+1}^{(m)}$  sampled in Step 3, sample  $z_{t+1}^{(m)}$  from

$$z_{t+1}^{(m)} | z_t^{(m)}, S_{t+1}^{(m)} \sim N(M_{S_t} + \Phi z_t^{(m)}, \Sigma_v).$$

Based on  $M$  draws on  $S_{t+1}$  and  $z_{t+1}$  sampled in Steps 3 and 4,  $f(\Delta y_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta})$  can be estimated by

$$f(\Delta y_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta}) \approx \frac{1}{M} \sum_{m=1}^M f(\Delta y_{t+1} | z_{t+1}^{(m)}, S_{t+1}^{(m)}, \boldsymbol{\theta}).$$

#### D. Diagnostics

Draws on  $S_{t+1}$  and  $z_{t+1}$  sampled in Steps 3 and 4 can be used also for a diagnostic test. The probability that  $\Delta y_{i,t+1}$  will be less than the observed value  $y_{i,t+1}^o$  conditional on  $\Delta \mathbf{y}_t$  and  $\boldsymbol{\theta}$  can be written as

$$\begin{aligned} & P(\Delta y_{i,t+1} \leq \Delta y_{i,t+1}^o | \Delta \mathbf{y}_t, \boldsymbol{\theta}) \\ &= \int P(\Delta y_{i,t+1} \leq \Delta y_{i,t+1}^o | z_{t+1}, S_{t+1}, \Delta \mathbf{y}_t, \boldsymbol{\theta}) f(z_{t+1}, S_{t+1} | \Delta \mathbf{y}_t, \boldsymbol{\theta}) dz_{t+1} dS_{t+1} \\ &\approx \frac{1}{M} \sum_{m=1}^M P(\Delta y_{i,t+1} \leq \Delta y_{i,t+1}^o | z_{t+1}^{(m)}, S_{t+1}^{(m)}, \boldsymbol{\theta}). \end{aligned}$$

Let  $u_{i,t+1}^M = 1/M \sum_{m=1}^M P(\Delta y_{i,t+1} \leq \Delta y_{i,t+1}^o | z_{t+1}^{(m)}, S_{t+1}^{(m)}, \boldsymbol{\theta})$ . Under the null of a correctly specified model,  $u_{i,t}^M$  converges in distribution to independently and identically distributed uniform random variables as  $M \rightarrow \infty$  (Rosenblatt [1952]). This provides a valid basis for diagnostic checking. These variables can be mapped into the normal distribution, by using the inverse of the normal distribution function  $n_{i,t}^M = F^{-1}(u_{i,t}^M)$  to give a standard sequence of independent and identically distributed normal variables.

### **III. Application to Macroeconomic Data in Japan**

#### **A. Data Description**

The ESRI uses 11 macroeconomic variables to construct its Coincident Index (see Table 1 for definitions of these 11 variables). Among them, “Business Profit” (ZBOAS) is quarterly data and the other 10 variables are monthly data. We obtained the raw data for these 10 variables from January 1975 to December 2000 and transformed them into seasonally adjusted ones by the Census-X11 method. The use of all 10 variables to estimate the Stock and Watson (1989, 1991) and the Kim and Nelson (1998) model is, however, computationally costly. Hence, our analysis is based on the following two datasets, both of which consist of five variables selected by Fukuda and Onodera (2001).

Dataset 1: (1) IIP95P (2) SCI95 (3) ESRAO (4) HWINMF (5) CELL9

Dataset 2: (1) IIP95P (2) SMSALE (3) HWINMF (4) IIP95O (5) IIP95M

Both the datasets were selected based on the principle of using variables related not only to production but also to trade sales and labor market. On one hand, dataset 1 includes “Index of Wholesale Sales” (SCI95) as a trade sales variable and “Ratio of Job Offers to Applicants” (ESRAO) and “Index of Non-Scheduled Hours Worked” (HWINMF) as labor market variables. On the other hand, dataset 2 includes “Sales of Small and Medium-Sized Companies” (SMSALE) as a trade sales variable and HWINMF as a labor market variable. These two datasets, however, differ in the sense that dataset 1 includes variables which are less correlated with each other, while all variables except HWINMF in dataset 2 are highly correlated with each other. Table 2 [1] reports the contemporaneous correlations of the growth rate of the 10 variables, showing that “Index of Industrial Production” (IIP95P), “Index of Raw Materials Consumption” (IIP95M), “Index of Operating Rate” (IIP95O), and “Sales of Small and Medium-Sized Companies” (SMSALE) have large positive correlations with each other. Dataset 2 includes all these variables. The correlations between variables in dataset 1 are less than 0.5 except the ones between IIP95P and “Electric Power Consumption of Large Users” (CELL9) and between IIP95P and SCI95. In addition, Table 2 [2] shows the serial correlations of the growth rate of the 10 variables, indicating that two labor market variables HWINMF and ESRAO have positive serial correlations and the other variables have negative serial correlations. Dataset 1 includes both of these variables, while dataset 2 includes only HWINMF.

The shifts of the mean growth of the composite index create a correlation between macroeconomic variables. Hence, we can expect that such shifts may be observed only in dataset 2, in which variables are highly correlated with each other.

**Table 2 Contemporaneous and Serial Correlations of the Growth Rate of the 10 Variables**

[1] Contemporaneous Correlations

|        | IIP95M  | IIP95O  | HWINMF  | ESRAO   | SDS     | CELL9  | IIP95S | SCI95  | SMSALE | IIP95P |
|--------|---------|---------|---------|---------|---------|--------|--------|--------|--------|--------|
| IIP95M | 1.0000  |         |         |         |         |        |        |        |        |        |
| IIP95O | 0.8820  | 1.0000  |         |         |         |        |        |        |        |        |
| HWINMF | 0.3321  | 0.3152  | 1.0000  |         |         |        |        |        |        |        |
| ESRAO  | 0.2401  | 0.2429  | 0.4157  | 1.0000  |         |        |        |        |        |        |
| SDS    | -0.0671 | -0.0943 | -0.0656 | -0.0054 | 1.0000  |        |        |        |        |        |
| CELL9  | 0.6507  | 0.6196  | 0.2549  | 0.2059  | -0.0467 | 1.0000 |        |        |        |        |
| IIP95S | 0.5250  | 0.5810  | 0.2088  | 0.1738  | 0.0298  | 0.4408 | 1.0000 |        |        |        |
| SCI95  | 0.5038  | 0.5059  | 0.1158  | 0.0974  | 0.3845  | 0.4562 | 0.4444 | 1.0000 |        |        |
| SMSALE | 0.6843  | 0.6632  | 0.2728  | 0.2288  | 0.0348  | 0.5221 | 0.6334 | 0.6070 | 1.0000 |        |
| IIP95P | 0.8673  | 0.8872  | 0.2524  | 0.2364  | -0.0727 | 0.6822 | 0.6624 | 0.6096 | 0.7756 | 1.0000 |

[2] Serial Correlations

| Variables   | IIP95M  | IIP95O  | HWINMF | ESRAO  | SDS     | CELL9   | IIP95S  | SCI95   | SMSALE  | IIP95P  |
|-------------|---------|---------|--------|--------|---------|---------|---------|---------|---------|---------|
| Correlation | -0.3492 | -0.4002 | 0.4176 | 0.5574 | -0.5628 | -0.2286 | -0.4604 | -0.3561 | -0.3550 | -0.4227 |

## B. Estimation Details

Following Fukuda and Onodera (2001), we set  $p = 3$  and  $q_i = 1 (i = 1, \dots, 5)$  for both the datasets. While Fukuda and Onodera (2001) assume  $r_i = 0$ , we set it equal to one.

For parameter estimation, we conduct the Gibbs sampler with 12,000 iterations for each model. The first 2,000 draws are discarded and then the next 10,000 are recorded. Using these 10,000 draws for each of the parameters, we calculate the posterior means, the standard errors of the posterior means, the 95 percent intervals, and the convergence diagnostic (CD) statistics proposed by Geweke (1992). The posterior means are computed by averaging the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1,000. The 95 percent intervals are calculated using the 2.5-th and 97.5-th percentiles of the simulated draws. The convergence of the Gibbs sampler can be assessed using the method proposed by Geweke (1992). He suggests comparing values early in the sequence with those late in the sequence. Let  $\theta^{(i)}$  be the  $i$ -th draw of a parameter in the recorded 10,000 draws, and let  $\bar{\theta}_A = 1/n_A \sum_{i=1}^{n_A} \theta^{(i)}$  and  $\bar{\theta}_B = 1/n_B \sum_{i=10,001-n_B}^{10,000} \theta^{(i)}$ . Using these values, Geweke (1992) proposes the following statistic called *convergence diagnostics* (CD).

$$CD = \frac{\bar{\theta}_A - \bar{\theta}_B}{\sqrt{\hat{\sigma}_A^2/n_A + \hat{\sigma}_B^2/n_B}}, \quad (32)$$

where  $\sqrt{\hat{\sigma}_A^2/n_A}$  and  $\sqrt{\hat{\sigma}_B^2/n_B}$  are standard errors of  $\bar{\theta}_A$  and  $\bar{\theta}_B$ . If the sequence of  $\theta^{(i)}$  is stationary, it converges in distribution to the standard normal. We set  $n_A = 1,000$  and  $n_B = 5,000$  and compute  $\hat{\sigma}_A^2$  and  $\hat{\sigma}_B^2$  using Parzen windows with bandwidths of 100 and 500, respectively.

In calculating the marginal likelihood, we set the number of iterations to evaluate both the posterior densities and the likelihood equal to 2,000.

### **C. Estimation Results**

Table 3 shows the estimation results for dataset 1. Table 3 [1] and [2] are the results for the Kim and Nelson (1998) model and the Stock and Watson (1989, 1991) model, respectively. According to the CD values, the null hypothesis that the sequence of 10,000 draws is stationary is accepted at the 5 percent significance level for all parameters in both the models. The log marginal likelihood of the Kim and Nelson (1998) model of  $-2,714.16$  is smaller than that of the Stock and Watson (1989, 1991) model of  $-2,713.31$ , indicating that the latter model compares favorably with the former model, although the difference of log marginal likelihoods is small.

Table 3 [3] shows the results of diagnostic checking based on variables  $n_{i,t}^M$  explained in Section II. The table shows the mean, standard deviation, skewness, kurtosis, and Ljung-Box statistics to test the null hypothesis of no serial correlation up to the sixth lag, where the numbers in parentheses show the standard errors. If the model is correctly specified, the asymptotic distribution of  $n_{i,t}^M$  is the standard normal. For SCI95, ESRAO, and HWINMF, the null hypothesis of no serial correlation is rejected at the 1 percent level. For all variables, the kurtosis is significantly larger than three.

Figure 1 [1] depicts the composite indices (CIs) estimated by the Kim and Nelson (1998) model and the Stock and Watson (1989, 1991) model jointly with that of the ESRI. The shaded areas represent the periods of ESRI recessions (from peak to trough). There is no major difference between the CIs estimated by both the models, which is consistent with the result that the difference of log marginal likelihoods is small. Figure 1 [2] depicts the posterior probability that the economy is in the recession state in each month as inferred from the Kim and Nelson (1998) model. These probabilities can be calculated simply by averaging 10,000 draws of the state  $S_t$  sampled from its posterior distribution.

Table 4 shows the results for dataset 2. According to the CD values, the null hypothesis that the sequence of 10,000 draws is stationary is accepted at any standard level for all parameters in both the models. The log marginal likelihood of the Kim and Nelson (1998) model of  $-2,202.10$  is slightly larger than that of the Stock and Watson (1989, 1991) model of  $-2,203.91$ , providing evidence, although weak, that the mean growth rate shifts depending on whether the economy is in a recession or in a boom.

Table 4 [3] shows the results of diagnostic checking. Except for HWINMF, the null hypothesis of no serial correlation is rejected at the 1 percent level. The kurtosis is still significantly larger than three for all variables.

Figure 2 [1] depicts the CIs estimated by the Kim and Nelson (1998) model and the Stock and Watson (1989, 1991) model jointly with that of the ESRI. No major difference between the CIs estimated by both the models is found again. Figure 2 [2] depicts the posterior probability that the economy is in the recession state in each month as inferred from the Kim and Nelson (1998) model. In contrast to the probabilities based on dataset 1, they move in a wider range between 0 percent and 100 percent, compared to Figure 1 [2].

We further estimate the Kim and Nelson (1998) model by using the following dataset.

Dataset 3: (1) IIP95P (2) SCI95 (3) ESRAO (4) HWINMF

**Table 3 Estimation Results for Dataset 1**

[1] Kim and Nelson Model

Marginal likelihood = -2,714.16

| Parameter      | Mean    | Standard error | 95 percent interval | CD    |
|----------------|---------|----------------|---------------------|-------|
| $\Delta C_t$   |         |                |                     |       |
| $\pi_{00}$     | 0.9045  | 0.0023         | [ 0.7578, 0.9807]   | 0.64  |
| $\pi_{11}$     | 0.9108  | 0.0038         | [ 0.7648, 0.9815]   | -0.89 |
| $\phi_1$       | -0.0489 | 0.0164         | [-0.3711, 0.2305]   | 1.05  |
| $\phi_2$       | 0.1034  | 0.0080         | [-0.1170, 0.2604]   | 0.87  |
| $\phi_3$       | 0.3572  | 0.0056         | [ 0.1795, 0.4989]   | 1.60  |
| $\mu_0$        | -0.3460 | 0.0246         | [-0.9584, -0.0144]  | 1.55  |
| $\mu_1$        | 0.5424  | 0.0405         | [ 0.0156, 1.4138]   | -1.59 |
| $y_{1t}$       |         |                |                     |       |
| $\lambda_{10}$ | 1.0739  | 0.0026         | [ 0.9301, 1.2141]   | 1.50  |
| $\lambda_{11}$ | -0.4708 | 0.0059         | [-0.6210, -0.3100]  | -0.77 |
| $\psi_1$       | -0.3803 | 0.0030         | [-0.5192, -0.2328]  | -0.89 |
| $\sigma_1^2$   | 0.5208  | 0.0066         | [ 0.3256, 0.7779]   | -0.36 |
| $y_{2t}$       |         |                |                     |       |
| $\lambda_{20}$ | 0.9507  | 0.0028         | [ 0.7782, 1.1313]   | 1.11  |
| $\lambda_{21}$ | -0.4521 | 0.0055         | [-0.6406, -0.2628]  | -1.24 |
| $\psi_2$       | -0.3075 | 0.0010         | [-0.4232, -0.1922]  | -0.98 |
| $\sigma_2^2$   | 2.0372  | 0.0031         | [ 1.6911, 2.4346]   | -0.98 |
| $y_{3t}$       |         |                |                     |       |
| $\lambda_{30}$ | 0.4653  | 0.0031         | [ 0.2540, 0.6779]   | 1.11  |
| $\lambda_{31}$ | 0.3589  | 0.0013         | [ 0.1601, 0.5556]   | 0.69  |
| $\psi_3$       | 0.4102  | 0.0013         | [ 0.2863, 0.5319]   | -0.82 |
| $\sigma_3^2$   | 3.1206  | 0.0054         | [ 2.6375, 3.6731]   | -1.20 |
| $y_{4t}$       |         |                |                     |       |
| $\lambda_{40}$ | 0.5369  | 0.0022         | [ 0.3850, 0.6914]   | 1.22  |
| $\lambda_{41}$ | 0.4272  | 0.0025         | [ 0.2679, 0.5886]   | 0.92  |
| $\psi_4$       | 0.1370  | 0.0025         | [-0.0049, 0.2846]   | -1.02 |
| $\sigma_4^2$   | 1.6903  | 0.0051         | [ 1.4014, 2.0190]   | -0.66 |
| $y_{5t}$       |         |                |                     |       |
| $\lambda_{50}$ | 0.7666  | 0.0014         | [ 0.6472, 0.8872]   | 0.86  |
| $\lambda_{51}$ | -0.2366 | 0.0060         | [-0.3932, -0.0800]  | -1.06 |
| $\psi_5$       | -0.2060 | 0.0013         | [-0.3359, -0.0773]  | -0.15 |
| $\sigma_5^2$   | 0.7474  | 0.0042         | [ 0.6030, 0.9132]   | 0.85  |

Notes: 1.  $y_{1t}$ ,  $y_{2t}$ ,  $y_{3t}$ ,  $y_{4t}$ , and  $y_{5t}$  represent IIP95P, SCI95, ESRAO, HWINMF, and CELL9, respectively. The first 2,000 draws are discarded and then the next 10,000 are used for calculating the posterior means, the standard errors of the posterior means, 95 percent interval, and the convergence diagnostic (CD) statistics proposed by Geweke (1992). The posterior means are computed by averaging the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1,000. The 95 percent intervals are calculated using the 2.5-th and 97.5-th percentiles of the simulated draws. The CD is computed using equation (32), where we set  $n_a = 1,000$  and  $n_b = 5,000$  and compute  $\hat{\sigma}_a^2$  and  $\hat{\sigma}_b^2$  using a Parzen window with bandwidths of 100 and 500, respectively.

2. Numbers in parentheses are standard errors. LB(6) is the Ljung-Box statistic including six lags. The critical values for LB(6) are 10.64 (10 percent), 12.59 (5 percent), and 16.81 (1 percent).



**Table 3 (continued)**

[2] Stock and Watson Model

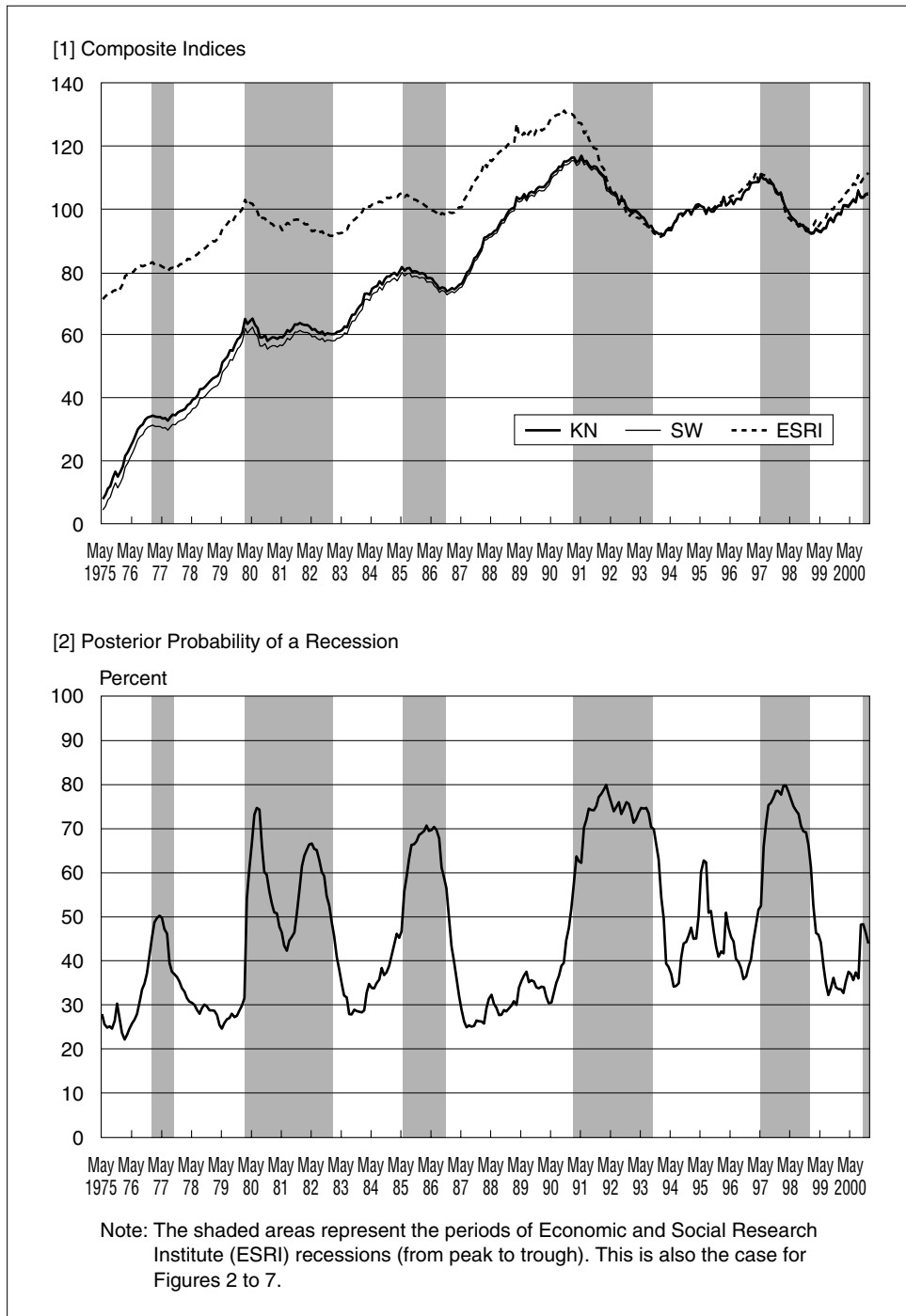
Marginal likelihood = -2,713.31

| Parameter      | Mean    | Standard error | 95 percent interval | CD    |
|----------------|---------|----------------|---------------------|-------|
| $\Delta C_t$   |         |                |                     |       |
| $\phi_1$       | 0.0164  | 0.0044         | [-0.1725, 0.2087]   | -0.93 |
| $\phi_2$       | 0.1526  | 0.0008         | [ 0.0362, 0.2683]   | -0.28 |
| $\phi_3$       | 0.3904  | 0.0006         | [ 0.2716, 0.5026]   | 0.46  |
| $y_{1t}$       |         |                |                     |       |
| $\lambda_{10}$ | 1.304   | 0.0035         | [ 0.9832, 1.2702]   | 1.65  |
| $\lambda_{11}$ | -0.5063 | 0.0026         | [-0.6642, -0.3023]  | 0.93  |
| $\psi_1$       | -0.3893 | 0.0019         | [-0.5293, -0.2335]  | 0.80  |
| $\sigma_1^2$   | 0.4775  | 0.0057         | [ 0.2632, 0.7090]   | -1.28 |
| $y_{2t}$       |         |                |                     |       |
| $\lambda_{20}$ | 0.9936  | 0.0013         | [ 0.8220, 1.1708]   | 0.90  |
| $\lambda_{21}$ | -0.4840 | 0.0028         | [ 0.4776, 0.9502]   | 0.75  |
| $\psi_2$       | -0.3025 | 0.0007         | [-0.4179, -0.1868]  | 1.05  |
| $\sigma_2^2$   | 2.0207  | 0.0025         | [ 1.6806, 2.4113]   | -0.73 |
| $y_{3t}$       |         |                |                     |       |
| $\lambda_{30}$ | 0.4735  | 0.0015         | [ 0.2629, 0.6904]   | -0.23 |
| $\lambda_{31}$ | 0.3372  | 0.0010         | [ 0.1386, 0.5358]   | 0.85  |
| $\psi_3$       | 0.4481  | 0.0013         | [ 0.3224, 0.5679]   | 0.81  |
| $\sigma_3^2$   | 3.0486  | 0.0030         | [ 2.5843, 3.5953]   | 1.43  |
| $y_{4t}$       |         |                |                     |       |
| $\lambda_{40}$ | 0.5615  | 0.0014         | [ 0.4061, 0.7167]   | 1.21  |
| $\lambda_{41}$ | 0.4258  | 0.0023         | [ 0.2692, 0.5835]   | 1.10  |
| $\psi_4$       | 0.1391  | 0.0009         | [ 0.0027, 0.2802]   | -0.12 |
| $\sigma_4^2$   | 1.6958  | 0.0030         | [ 1.4137, 2.0224]   | 1.43  |
| $y_{5t}$       |         |                |                     |       |
| $\lambda_{50}$ | 0.8001  | 0.0015         | [ 0.6848, 0.9222]   | 0.47  |
| $\lambda_{51}$ | -0.2616 | 0.0030         | [-0.4082, -0.1138]  | 1.03  |
| $\psi_5$       | -0.1975 | 0.0010         | [-0.3266, -0.0666]  | 1.19  |
| $\sigma_5^2$   | 0.7397  | 0.0028         | [ 0.5969, 0.9018]   | 1.55  |

[3] Diagnostic Check for the Kim and Nelson Model

|          | IIP95P              | SCI95               | ESRAO               | HWINMF             | CELL9              |
|----------|---------------------|---------------------|---------------------|--------------------|--------------------|
| Mean     | 0.0304<br>(0.0581)  | 0.0178<br>(0.0565)  | 0.0230<br>(0.0570)  | 0.0356<br>(0.0558) | 0.0287<br>(0.0584) |
| St. dev. | 1.0223              | 0.9950              | 1.0027              | 0.9829             | 1.0274             |
| Skewness | -0.2690<br>(0.1391) | -0.1454<br>(0.1391) | -0.1355<br>(0.1391) | 0.3565<br>(0.1391) | 0.1320<br>(0.1391) |
| Kurtosis | 4.0703<br>(0.2782)  | 7.9968<br>(0.2782)  | 5.7695<br>(0.2782)  | 5.0004<br>(0.2782) | 4.8018<br>(0.2782) |
| LB(6)    | 8.31                | 18.17               | 44.17               | 61.14              | 13.43              |

**Figure 1 Dataset 1**



**Table 4 Estimation Results for Dataset 2**

[1] Kim and Nelson Model

Marginal likelihood = -2,202.10

| Parameter      | Mean    | Standard error | 95 percent interval | CD    |
|----------------|---------|----------------|---------------------|-------|
| $\Delta C_t$   |         |                |                     |       |
| $\pi_{00}$     | 0.9178  | 0.0010         | [ 0.8334, 0.9698]   | -0.42 |
| $\pi_{11}$     | 0.9368  | 0.0018         | [ 0.8401, 0.9779]   | 0.96  |
| $\phi_1$       | -0.2761 | 0.0057         | [-0.4595, -0.0597]  | -0.00 |
| $\phi_2$       | 0.0008  | 0.0049         | [-0.1641, 0.2059]   | -0.31 |
| $\phi_3$       | 0.2657  | 0.0033         | [ 0.1236, 0.4273]   | -0.29 |
| $\mu_0$        | -0.6341 | 0.0149         | [-0.9535, -0.0767]  | -0.80 |
| $\mu_1$        | 1.0687  | 0.0242         | [ 0.1435, 1.5009]   | 0.97  |
| $y_{1t}$       |         |                |                     |       |
| $\lambda_{10}$ | 1.0930  | 0.0041         | [ 0.9875, 1.2135]   | -0.79 |
| $\lambda_{11}$ | -0.4251 | 0.0042         | [-0.5467, -0.2959]  | -1.33 |
| $\psi_1$       | -0.4006 | 0.0010         | [-0.5269, -0.2681]  | 0.34  |
| $\sigma_1^2$   | 0.2649  | 0.0010         | [ 0.1997, 0.3434]   | -1.32 |
| $y_{2t}$       |         |                |                     |       |
| $\lambda_{20}$ | 0.8629  | 0.0035         | [ 0.7463, 0.9936]   | -0.76 |
| $\lambda_{21}$ | -0.2101 | 0.0031         | [-0.3285, -0.0920]  | -1.39 |
| $\psi_2$       | -0.3417 | 0.0006         | [-0.4509, -0.2308]  | -0.44 |
| $\sigma_2^2$   | 0.9226  | 0.0009         | [ 0.7738, 1.0903]   | -0.78 |
| $y_{3t}$       |         |                |                     |       |
| $\lambda_{30}$ | 0.5736  | 0.0021         | [ 0.4377, 0.7180]   | -1.05 |
| $\lambda_{31}$ | 0.4670  | 0.0022         | [ 0.3356, 0.6070]   | -0.89 |
| $\psi_3$       | 0.0848  | 0.0017         | [-0.0478, 0.2173]   | -0.38 |
| $\sigma_3^2$   | 1.6668  | 0.0032         | [ 1.4080, 1.9668]   | -1.01 |
| $y_{4t}$       |         |                |                     |       |
| $\lambda_{40}$ | 1.1547  | 0.0045         | [ 1.0146, 1.2890]   | -0.85 |
| $\lambda_{41}$ | -0.4274 | 0.0045         | [-0.5554, -0.2921]  | -1.28 |
| $\psi_4$       | -0.4088 | 0.0014         | [-0.5588, -0.2520]  | -0.86 |
| $\sigma_4^2$   | 0.3516  | 0.0008         | [ 0.2717, 0.4494]   | -0.41 |
| $y_{5t}$       |         |                |                     |       |
| $\lambda_{50}$ | 1.1268  | 0.0043         | [ 1.0180, 1.2518]   | -1.21 |
| $\lambda_{51}$ | -0.3209 | 0.0044         | [-0.4478, -0.1810]  | -1.39 |
| $\psi_5$       | -0.2604 | 0.0013         | [-0.3971, -0.1215]  | -1.09 |
| $\sigma_5^2$   | 0.3233  | 0.0008         | [ 0.2504, 0.4097]   | 1.26  |

Notes: 1.  $y_{1t}$ ,  $y_{2t}$ ,  $y_{3t}$ ,  $y_{4t}$ , and  $y_{5t}$  represent IIP95P, SMSALE, HWINMF, IIP95O, and IIP95P, respectively. The first 2,000 draws are discarded and then the next 10,000 are used for calculating the posterior means, the standard errors of the posterior means, 95 percent interval, and the convergence diagnostic (CD) statistics proposed by Geweke (1992). The posterior means are computed by averaging the simulated draws. The standard errors of the posterior means are computed using a Parzen window with a bandwidth of 1,000. The 95 percent intervals are calculated using the 2.5-th and 97.5-th percentiles of the simulated draws. The CD is computed using equation (32), where we set  $n_A = 1,000$  and  $n_B = 5,000$  and compute  $\hat{\sigma}_A^2$  and  $\hat{\sigma}_B^2$  using a Parzen window with bandwidths of 100 and 500, respectively.

2. Numbers in parentheses are standard errors. LB(6) is the Ljung-Box statistic including six lags. The critical values for LB(6) are 10.64 (10 percent), 12.59 (5 percent), and 16.81 (1 percent).

(Continued on next page)

**Table 4 (continued)**

[2] Stock and Watson Model

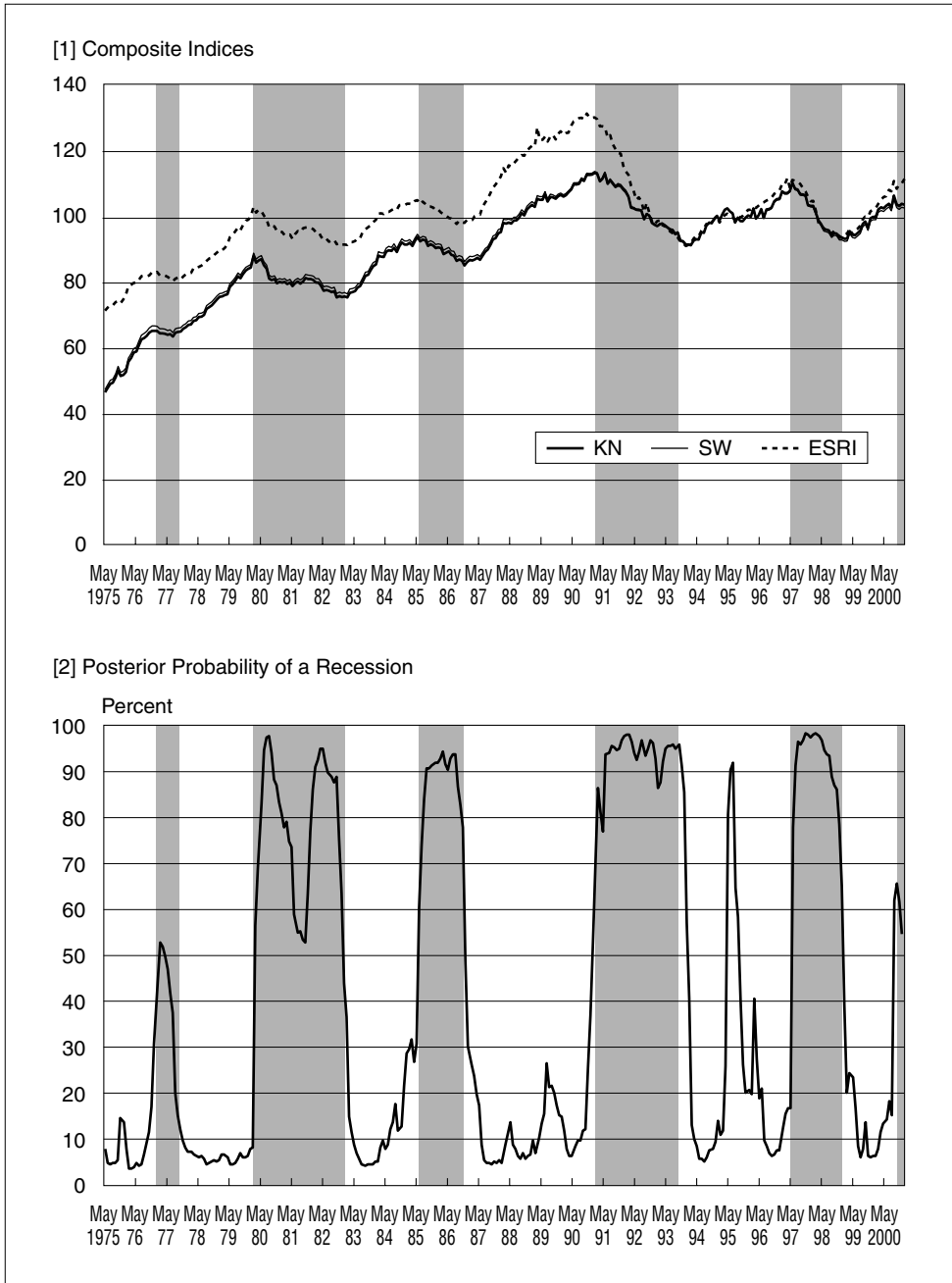
Marginal likelihood = -2,203.91

| Parameter      | Mean    | Standard error | 95 percent interval | CD    |
|----------------|---------|----------------|---------------------|-------|
| $\Delta C_t$   |         |                |                     |       |
| $\phi_1$       | -0.0869 | 0.0028         | [-0.2229, 0.0499]   | 1.23  |
| $\phi_2$       | 0.1715  | 0.0012         | [ 0.0627, 0.2783]   | 0.71  |
| $\phi_3$       | 0.3785  | 0.0007         | [ 0.2728, 0.4830]   | -0.91 |
| $y_{1t}$       |         |                |                     |       |
| $\lambda_{10}$ | 1.2304  | 0.0028         | [ 1.1229, 1.3471]   | 0.01  |
| $\lambda_{11}$ | -0.4982 | 0.0031         | [-0.6302, -0.3704]  | -1.24 |
| $\psi_1$       | -0.3874 | 0.0009         | [-0.5164, -0.2544]  | -0.16 |
| $\sigma_1^2$   | 0.2329  | 0.0007         | [ 0.1743, 0.3009]   | 0.00  |
| $y_{2t}$       |         |                |                     |       |
| $\lambda_{20}$ | 0.9695  | 0.0021         | [ 0.8475, 1.0997]   | -0.08 |
| $\lambda_{21}$ | -0.2596 | 0.0025         | [-0.3910, -0.1335]  | -1.21 |
| $\psi_2$       | -0.3395 | 0.0007         | [-0.4512, -0.2278]  | -1.05 |
| $\sigma_2^2$   | 0.9042  | 0.0009         | [ 0.7572, 1.0738]   | -1.40 |
| $y_{3t}$       |         |                |                     |       |
| $\lambda_{30}$ | 0.6409  | 0.0014         | [ 0.4902, 0.7936]   | -0.37 |
| $\lambda_{31}$ | 0.5017  | 0.0018         | [ 0.3556, 0.6511]   | -0.49 |
| $\psi_3$       | 0.0814  | 0.0013         | [-0.0476, 0.2171]   | -1.01 |
| $\sigma_3^2$   | 1.6608  | 0.0026         | [ 1.4044, 1.9578]   | -1.00 |
| $y_{4t}$       |         |                |                     |       |
| $\lambda_{40}$ | 1.3007  | 0.0030         | [ 1.1849, 1.4245]   | -0.06 |
| $\lambda_{41}$ | -0.5049 | 0.0033         | [-0.6453, -0.3658]  | -1.16 |
| $\psi_4$       | -0.3640 | 0.0008         | [-0.4894, -0.2363]  | -0.78 |
| $\sigma_4^2$   | 0.3063  | 0.0007         | [ 0.2389, 0.3831]   | 0.07  |
| $y_{5t}$       |         |                |                     |       |
| $\lambda_{50}$ | 1.2589  | 0.0030         | [ 1.1434, 1.3777]   | -0.01 |
| $\lambda_{51}$ | -0.3726 | 0.0010         | [-0.5166, -0.2282]  | -1.24 |
| $\psi_5$       | -0.2816 | 0.0013         | [-0.4095, -0.1515]  | -1.42 |
| $\sigma_5^2$   | 0.3159  | 0.0009         | [ 0.2479, 0.3945]   | 0.58  |

[3] Diagnostic Check for the Kim and Nelson Model

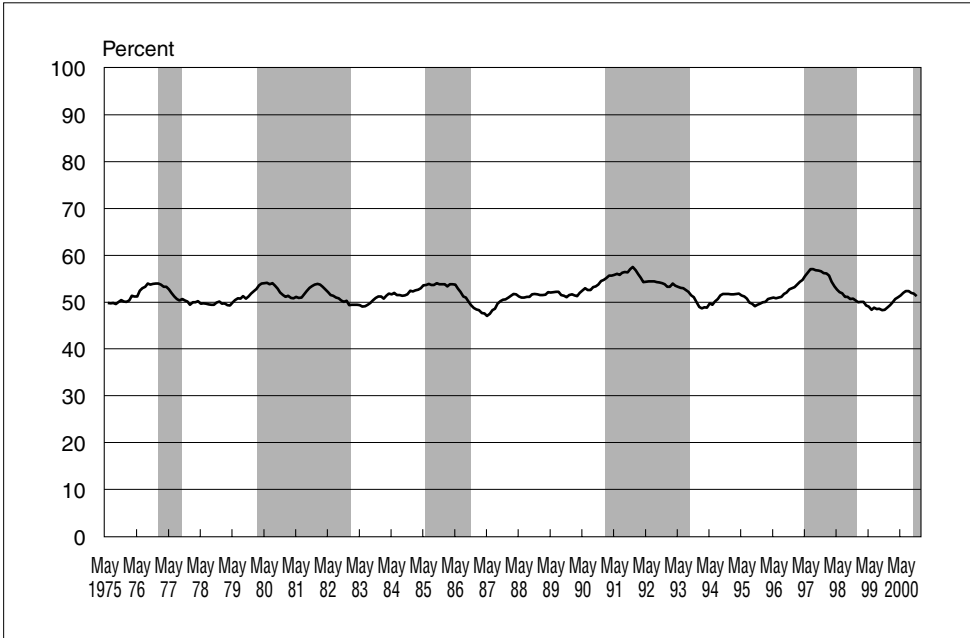
|          | IIP95P              | SMSALE             | HWINMF             | IIP95O              | IIP95P              |
|----------|---------------------|--------------------|--------------------|---------------------|---------------------|
| Mean     | 0.0129<br>(0.0590)  | 0.0107<br>(0.0571) | 0.0263<br>(0.0559) | 0.0210<br>(0.0591)  | 0.0160<br>(0.0586)  |
| St. dev. | 1.0384              | 1.0051             | 0.9843             | 1.0406              | 1.0318              |
| Skewness | -0.3275<br>(0.1391) | 0.0174<br>(0.1391) | 0.3323<br>(0.1391) | -0.2062<br>(0.1391) | -0.2058<br>(0.1391) |
| Kurtosis | 4.2933<br>(0.2782)  | 4.4246<br>(0.2782) | 4.7422<br>(0.2782) | 3.5241<br>(0.2787)  | 3.6938<br>(0.2787)  |
| LB(6)    | 4.93                | 11.92              | 69.36              | 10.82               | 4.94                |

Figure 2 Dataset 2



Dataset 3 is the one in which CELL9 is excluded from dataset 1. These four variables are used to construct the Nikkei Business Index. This index is the CI constructed by the Nihon Keizai Shimbun, Inc. using the Stock and Watson (1989, 1991) model. These four variables correspond to the four variables used by the U.S. Department of Commerce (DOC) to construct its CI: “industrial production,” “total personal income less transfer payments in 1987 dollars,” “employees on nonagricultural payrolls,” and “total manufacturing and trade sales in 1987 dollars.” We only report the posterior probabilities of a recession, which is depicted in Figure 3. Unlike datasets 1 and 2, the posterior probability moves in a very narrow range around 50 percent, so that it cannot be used to date the business cycle turning points. This may be attributed to the fact that the four variables in dataset 3 are weakly correlated with each other.

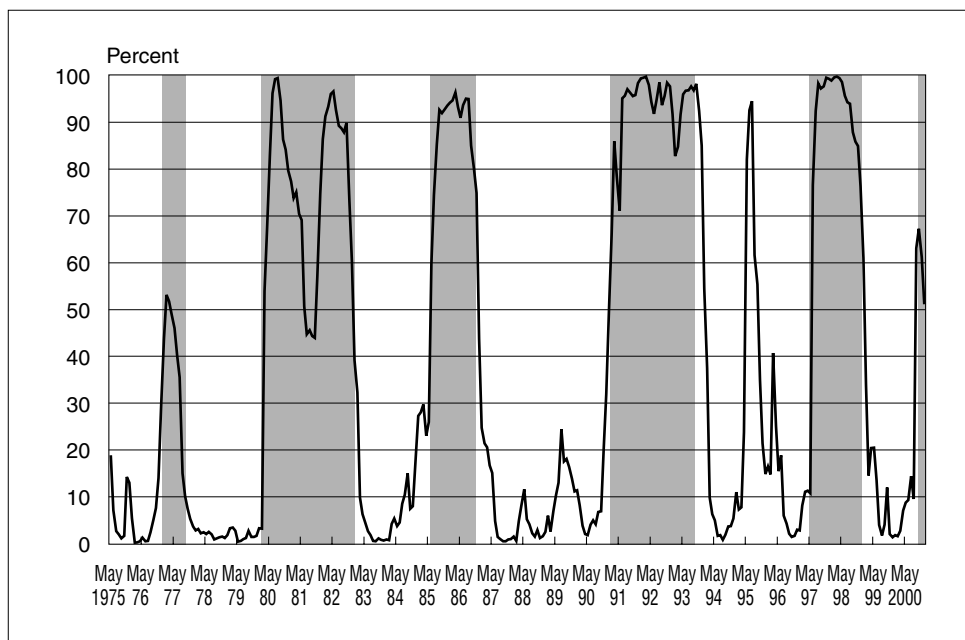
**Figure 3 Posterior Probability of a Recession: Dataset 3**



In dataset 2, the null hypothesis of no serial correlation in the diagnostic statistic is rejected for HWINMF. This may be attributed to the fact that HWINMF has a positive serial correlation while all other variables in dataset 2 have a negative serial correlation and that HWINMF is weakly correlated with other variables. Hence, we also analyze dataset 2 without HWINMF:

Dataset 4: (1) IIP95P (2) SMSALE (3) IIP95O (4) IIP95M

The posterior probability of a recession calculated by fitting the Kim and Nelson (1998) model to dataset 4 is depicted in Figure 4. Figure 2 [2] and Figure 4 look alike, demonstrating that HWINMF does not play an important role in dataset 2.

**Figure 4 Posterior Probability of a Recession: Dataset 4**


Following Kaufman (2000), we date the turning points by defining period  $t$  as a peak if the posterior probability  $P(S_t = 1 | \Delta \tilde{y}_T) > 0.5$  and  $P(S_{t+1} = 1 | \Delta \tilde{y}_T) < 0.5$  and a trough if  $P(S_t = 1 | \Delta \tilde{y}_T) < 0.5$  and  $P(S_{t+1} = 1 | \Delta \tilde{y}_T) > 0.5$ . As mentioned, the posterior probabilities estimated using dataset 3 move in a very narrow range around 50 percent, so that they cannot be used to date the business cycle turning points. Therefore, we estimate the turning points using datasets 1, 2, and 4. The estimated turning points are shown in Table 5 jointly with the reference date of the ESRI. The difference of the turning points among the three datasets is at most one month except the peak and trough in 1981, which are detected only by datasets 1 and 4, and the peak in 2000, which are detected only by datasets 2 and 4. The difference of the turning points estimated by the Kim and Nelson (1998) model from the reference date of the ESRI is larger, but at most three months except the trough in 1997, the peak and trough in 1981, and the peak in 2000.

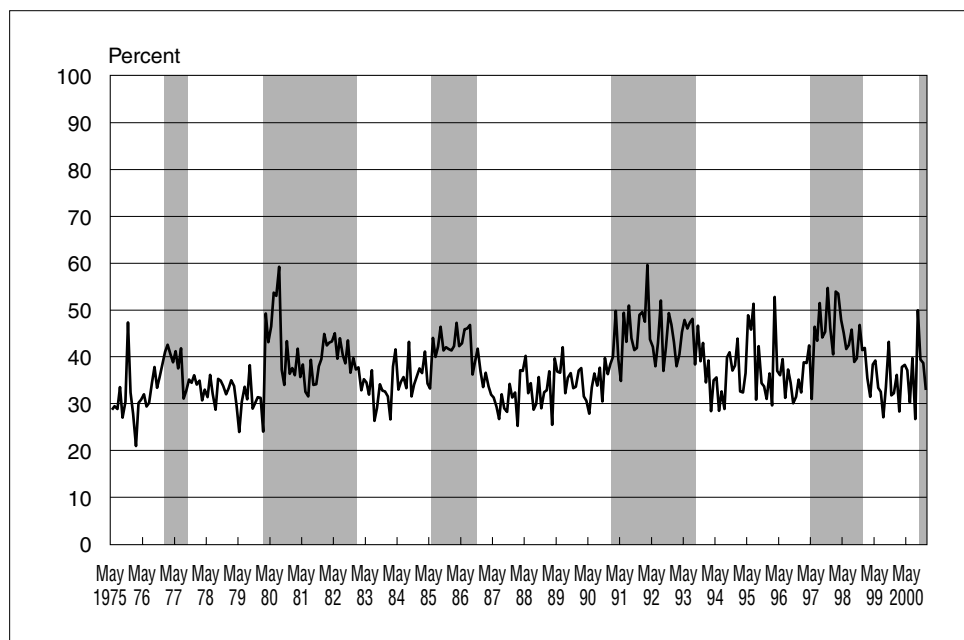
Thus far,  $beta(18, 2)$  is used as a prior distribution for the transition probabilities  $\pi_{00}$  and  $\pi_{11}$ . It is tight because the mean and standard deviation of  $beta(18, 2)$  are 0.9 and 0.065, respectively. To examine how the results are sensitive to the prior distribution of  $\pi_{00}$  and  $\pi_{11}$ , we estimate the Kim and Nelson (1998) model using datasets 1, 2, and 4 under the diffuse prior  $beta(1, 1)$ , which corresponds to a uniform distribution in  $[-1, 1]$ . The estimated posterior probabilities of a recession are depicted in Figures 5 (dataset 1), 6 (dataset 2), and 7 (dataset 4). Figure 5 shows that the posterior probabilities estimated using dataset 1 move in a very narrow range around 30 to 40 percent, indicating that dataset 1, in which five variables are weakly correlated with each other, requires a tight prior for the transition probabilities  $\pi_{00}$  and  $\pi_{11}$  to date the business cycle turning points. Figures 6 and 7 show that this is not true for datasets 2 and 4,

**Table 5 Estimates of Business Cycle Turning Points Based on the Kim and Nelson Model**

|   | ESRI      | KN        |           |           |
|---|-----------|-----------|-----------|-----------|
|   |           | Dataset 1 | Dataset 2 | Dataset 4 |
| P | Jan. 1977 | Mar. 1977 | Jan. 1977 | Jan. 1977 |
| T | Oct. 1977 | Apr. 1977 | Mar. 1977 | Mar. 1977 |
| P | Feb. 1980 | Feb. 1980 | Feb. 1980 | Feb. 1980 |
| T |           | Mar. 1981 |           | June 1981 |
| P |           | Oct. 1981 |           | Oct. 1981 |
| T | Feb. 1983 | Dec. 1982 | Dec. 1982 | Dec. 1982 |
| P | June 1985 | May 1985  | May 1985  | May 1985  |
| T | Nov. 1986 | Nov. 1986 | Nov. 1986 | Nov. 1986 |
| P | Feb. 1991 | Dec. 1990 | Dec. 1990 | Jan. 1991 |
| T | Oct. 1993 | Jan. 1994 | Jan. 1994 | Jan. 1994 |
| P |           | Mar. 1995 | Apr. 1995 | Apr. 1995 |
| T |           | Sep. 1995 | Sep. 1995 | Sep. 1995 |
| P | Mar. 1997 | Mar. 1997 | May 1997  | May 1997  |
| T | Apr. 1999 | Feb. 1999 | Jan. 1999 | Jan. 1999 |
| P |           |           | Aug. 2000 | Aug. 2000 |

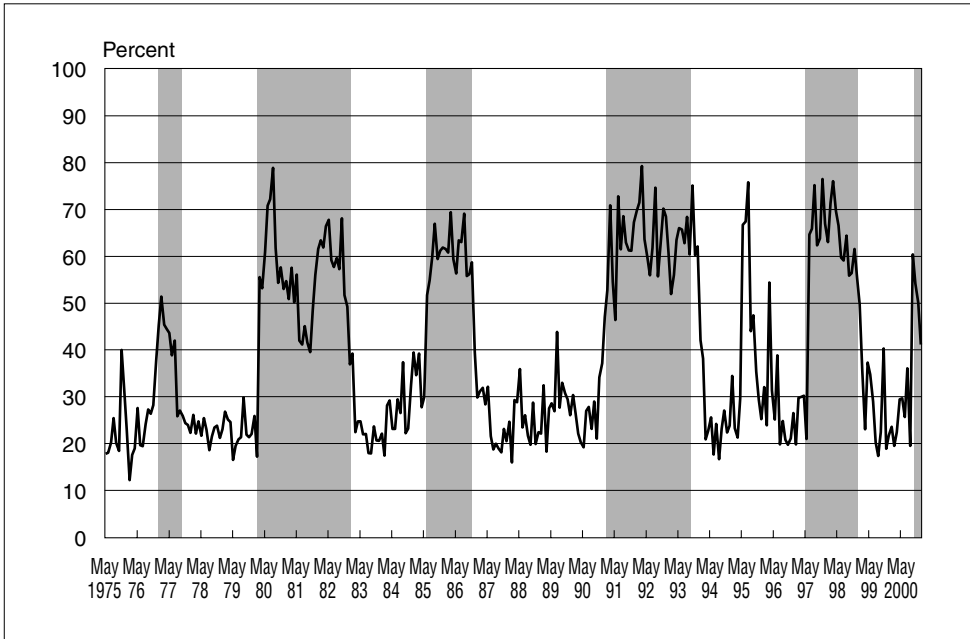
Note: "P" (peak) indicates the date when the posterior probability  $P(S_t = 1|y^T) > 0.5$  and  $P(S_{t+1} = 1|y^T) < 0.5$ . "T" (trough) indicates the date when the posterior probability  $P(S_t = 1|y^T) < 0.5$  and  $P(S_{t+1} = 1|y^T) > 0.5$ . The column "KN" is the estimates of turning points based on the Kim and Nelson (1998) model. "ESRI" is the reference date of the Economic and Social Research Institute in the Cabinet Office.

**Figure 5 Posterior Probability of a Recession (Diffuse Prior for the Transition Probabilities): Dataset 1**

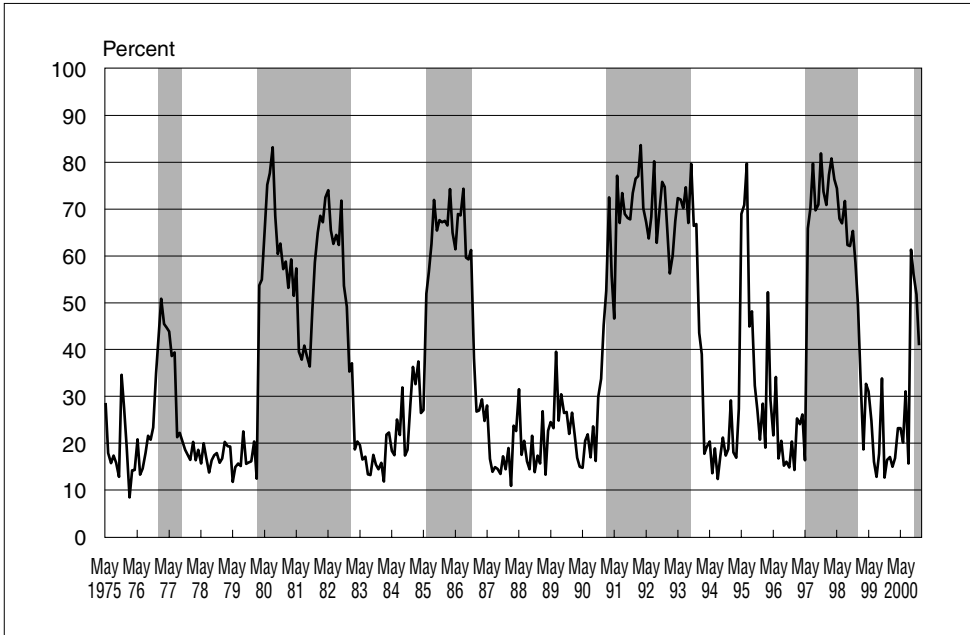




**Figure 6 Posterior Probability of a Recession (Diffuse Prior for the Transition Probabilities): Dataset 2**



**Figure 7 Posterior Probability of a Recession (Diffuse Prior for the Transition Probabilities): Dataset 4**



in which the correlations of variables are not so weak. The turning points estimated using datasets 2 and 4 with the diffuse prior for  $\pi_{00}$  and  $\pi_{11}$  are shown in Table 6. The effects of a prior for  $\pi_{00}$  and  $\pi_{11}$  on the estimated turning points are small, especially for dataset 4. When dataset 4 is used, the turning points estimated with the diffuse prior are the same as those with the tight prior except the trough in 2000, which is not detected when the tight prior is used, and the trough in 1985, whose difference is only one month. The conclusion must be that the estimation results of the Kim and Nelson (1998) model are insensitive to the prior distribution when highly correlated variables are used, but it is not true when weakly correlated variables are used.

**Table 6 Estimates of Business Cycle Turning Points Based on the Kim and Nelson Model (Diffuse Prior for Transition Probabilities  $\pi_{00}$  and  $\pi_{11}$ )**

|   | KN        |           |
|---|-----------|-----------|
|   | Dataset 2 | Dataset 4 |
| P |           | Jan. 1977 |
| T |           | Mar. 1977 |
| P | Feb. 1980 | Feb. 1980 |
| T | Mar. 1981 | May 1981  |
| P | Nov. 1981 | Oct. 1981 |
| T | Nov. 1982 | Dec. 1982 |
| P | May 1985  | May 1985  |
| T | Nov. 1986 | Nov. 1986 |
| P | Jan. 1991 | Jan. 1991 |
| T | Dec. 1993 | Jan. 1994 |
| P | Apr. 1995 | Apr. 1995 |
| T | July 1995 | Sep. 1995 |
| P | May 1997  | May 1997  |
| T | Dec. 1998 | Jan. 1999 |
| P | Aug. 2000 | Aug. 2000 |
| T | Oct. 2000 | Nov. 2000 |

Note: "P" (peak) indicates the date when the posterior probability  $P(S_t = 1|y^T) > 0.5$  and  $P(S_{t+1} = 1|y^T) < 0.5$ . "T" (trough) indicates the date when the posterior probability  $P(S_t = 1|y^T) < 0.5$  and  $P(S_{t+1} = 1|y^T) > 0.5$ .

The ESRI announces the date of a turning point one year and a few months after the date of the turning point and may revise the date a few months after the first announcement. For example, the ESRI announced the peak in 1997 as March 1997 in June 1998, and revised it as May 1997 in December 2001. The trough in 1999 was first announced as April 1994 in June 2000 and revised as April 1999 in December 2001. The peak in 2000 was announced as October 2001 in December 2001 and may be revised in the future. We examine how quickly the Kim and Nelson (1998) model can detect the date of a turning point. Specifically, we examine when the peaks in 1997 and 2000 and the trough in 1999 are first detected by estimating the Kim and Nelson

(1998) model using dataset 4 up to one to three months after those turning points. Surprisingly enough, all three turning points are detected only two months after the dates of turning points, and the detected dates are the same as those estimated using the data up to the three months after the turning points and those using the full sample.

We find that the Kim and Nelson (1998) model performs well when dataset 4 (or 2) is used. All variables except SMSALE in dataset 4 are production-related variables. Hence, the Kim and Nelson (1998) model using only IIP95P, which is representative of production-related variables, may also perform well. To examine whether this is true, we estimate the Kim and Nelson (1998) model using only IIP95P. The estimated posterior probabilities of a recession move in a narrow range around 50 percent. This result indicates that the Kim and Nelson (1998) does not perform well when only IIP95P is used and requires other production-related variables.

## **IV. Conclusions**

---

This article fits the Markov switching dynamic factor model proposed by Kim and Nelson (1998) to some macroeconomic variables in Japan. We do not find strong evidence that the Kim and Nelson (1998) model is favored over the Stock and Watson (1989, 1991) model, nor major differences between the composite indices estimated by the two models. The Kim and Nelson (1998) model, however, produces the estimates of turning points close to the reference dates of the ESRI in the Cabinet Office unless only weakly correlated variables are used.

In this article, we focus on the in-sample fit of the models. Needless to say, it is worthwhile examining the out-of-sample forecasting ability.

## APPENDIX 1: SAMPLING FROM CONDITIONAL DISTRIBUTIONS (8)–(13)

Conditional distributions (8)–(12) can be derived based on the following theorem, which is well known in Bayesian econometrics.

THEOREM. *Consider the linear regression model*

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim N(0, \sigma^2 \mathbf{I}_T), \quad (\text{A.1})$$

where  $\mathbf{Y}$  is the  $T \times 1$  vector of the dependent variable,  $\mathbf{X}$  is the  $T \times k$  matrix of independent variables,  $\boldsymbol{\beta}$  is the  $k \times 1$  vector of regression coefficients,  $\mathbf{u}$  is the  $T \times 1$  vector of the error term that follows the independent normal distribution with variance  $\sigma^2$ , and  $\mathbf{I}_T$  is the  $T \times T$  identity matrix.

*Under the prior distributions*

$$\boldsymbol{\beta} \sim N(\mathbf{M}^{(0)}, \boldsymbol{\Sigma}^{(0)}), \quad \sigma^2 \sim IG(\nu^{(0)}/2, \delta^{(0)}/2),$$

the conditional distributions  $f(\boldsymbol{\beta}|\sigma^2, \mathbf{X}, \mathbf{Y})$  and  $f(\sigma^2|\boldsymbol{\beta}, \mathbf{X}, \mathbf{Y})$  are given by

$$\boldsymbol{\beta}|\sigma^2, \mathbf{Y} \sim N(\mathbf{M}^{(1)}, \boldsymbol{\Sigma}^{(1)}), \quad (\text{A.2})$$

where

$$\begin{aligned} \mathbf{M}^{(1)} &= (\boldsymbol{\Sigma}^{(0)-1} + \sigma^{-2} \mathbf{X}'\mathbf{X})^{-1} (\boldsymbol{\Sigma}^{(0)-1} \mathbf{M}^{(0)} + \sigma^{-2} \mathbf{X}'\mathbf{Y}), \\ \boldsymbol{\Sigma}^{(1)} &= (\boldsymbol{\Sigma}^{(0)-1} + \sigma^{-2} \mathbf{X}'\mathbf{X})^{-1}, \end{aligned}$$

and

$$\sigma^2|\boldsymbol{\beta}, \mathbf{Y} \sim IG(\nu^{(1)}/2, \delta^{(1)}/2), \quad (\text{A.3})$$

where

$$\begin{aligned} \nu_1 &= \nu_0 + T, \\ \delta_1 &= \delta_0 + (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}). \end{aligned}$$

*Proof.*

$$\begin{aligned} &f(\boldsymbol{\beta}|\sigma^2, \mathbf{Y}) \\ &\propto f(\mathbf{Y}|\boldsymbol{\beta}, \sigma^2) f(\boldsymbol{\beta}) \\ &\propto \exp \left[ -\frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})' (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \right] \exp \left[ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{M}^{(0)})' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\beta} - \mathbf{M}^{(0)}) \right] \\ &\propto \exp \left[ -\frac{1}{2} (\boldsymbol{\beta} - \mathbf{M}^{(1)})' \boldsymbol{\Sigma}^{(1)-1} (\boldsymbol{\beta} - \mathbf{M}^{(1)}) \right]. \end{aligned}$$

$$\begin{aligned}
 f\left(\frac{1}{\sigma^2} \mid \beta, \mathbf{Y}\right) & \\
 & \propto f(\mathbf{Y} \mid \beta, \sigma^2) f\left(\frac{1}{\sigma^2}\right) \\
 & \propto \left(\frac{1}{\sigma^2}\right)^{(T/2)} \exp\left[-\frac{1}{2\sigma^2} (\mathbf{Y} - \mathbf{X}\beta)' (\mathbf{Y} - \mathbf{X}\beta)\right] \left(\frac{1}{\sigma^2}\right)^{\nu_0/2-1} \exp\left[-\frac{\delta_0}{2\sigma^2}\right] \\
 & \propto \left(\frac{1}{\sigma^2}\right)^{\nu_1/2-1} \exp\left[-\frac{\delta_1}{2\sigma^2}\right].
 \end{aligned}$$

### A. Conditional Distribution (8)

The condition of (8) includes  $\psi_i$ ,  $\Delta \mathbf{y}_T$ , and  $\Delta \mathbf{c}_T$ . Given them, we can calculate

$$\begin{aligned}
 \Delta y_{it}^* &= \Delta y_{it} - \psi_{i1} \Delta y_{i,t-1} - \dots - \psi_{iq_i} \Delta y_{i,t-q_i}, \\
 \Delta c_i^* &= \Delta c_i - \psi_{i1} \Delta c_{i,t-1} - \dots - \psi_{iq_i} \Delta c_{i,t-q_i}.
 \end{aligned}$$

Using them, define

$$\begin{aligned}
 \mathbf{Y} &= [\Delta y_{i,q_i+r_i}^*, \dots, \Delta y_{iT}^*], \\
 \mathbf{X} &= \begin{bmatrix} \Delta c_{q_i+r_i}^* & \dots & \Delta c_{q_i}^* \\ \vdots & \ddots & \vdots \\ \Delta c_T^* & \dots & \Delta c_{T-r_i}^* \end{bmatrix}, \\
 \beta &= \boldsymbol{\lambda}_i, \\
 \mathbf{u} &= [\boldsymbol{\epsilon}_{i,q_i+r_i}, \dots, \boldsymbol{\epsilon}_{i,T}], \\
 \sigma^2 &= \sigma_i^2.
 \end{aligned}$$

Then, equation (1') will be the linear regression model (A.1) in the above theorem, so that the conditional distribution (8) is given by equation (A.2). Specifically,

$$\lambda_1 | \cdot \sim N(M_{\lambda_1}^{(1)}, \sum_{\lambda_1}^{(1)}) I[\lambda_{10} > 0], \quad (\text{A.4})$$

$$\lambda_i | \cdot \sim N(M_{\lambda_i}^{(1)}, \sum_{\lambda_i}^{(1)}), \quad i = 2, \dots, n, \quad (\text{A.5})$$

where

$$\begin{aligned}
 M_{\lambda_i}^{(1)} &= \left(\sum_{\lambda_i}^{(0)-1} + \sigma_i^{-2} \mathbf{X}'\mathbf{X}\right)^{-1} \left(\sum_{\lambda_i}^{(0)-1} M_{\lambda_i}^{(0)} + \sigma_i^{-2} \mathbf{X}'\mathbf{Y}\right), \\
 \sum_{\lambda_i}^{(1)} &= \left(\sum_{\lambda_i}^{(0)-1} + \sigma_i^{-2} \mathbf{X}'\mathbf{X}\right)^{-1}.
 \end{aligned}$$

It is straightforward to sample from the normal distribution (A.5). We can sample from the truncated normal distribution (A.4) by sampling from the normal distribution  $N(M_{\lambda_1}^{(1)}, \sum_{\lambda_1}^{(1)})$  and accepting it only if it is positive.

### B. Conditional Distributions (9) and (10)

The conditions of (9) and (10) include  $\lambda_i$ ,  $\Delta y_T$ , and  $\Delta c_T$ . Given them, we can calculate

$$e_{it} = \Delta y_{it} - \lambda_{i0} \Delta c_t - \lambda_{i1} \Delta c_{t-1} - \dots - \lambda_{ir_i} \Delta c_{t-r_i}, \quad t = r_i + 1, \dots, T.$$

Define

$$\begin{aligned} Y &= [e_{i, r_i+q_i+1}, \dots, e_{i, T}], \\ X &= \begin{bmatrix} e_{i, r_i+q_i} & \cdots & e_{i, r_i+1} \\ \vdots & \ddots & \vdots \\ e_{i, T-1} & \cdots & e_{i, T-q_i} \end{bmatrix}, \\ \beta &= \psi_i, \\ u &= [\epsilon_{i, r_i+q_i+2}, \dots, \epsilon_{i, T}], \\ \sigma^2 &= \sigma_i^2. \end{aligned}$$

Then, equation (2) will be the linear regression model (A.1), so that the conditional distributions (9) and (10) are given by equations (A.2) and (A.3). Specifically,

$$\psi_i | \cdot \sim N(M_{\psi_i}^{(1)}, \Sigma_{\psi_i}^{(1)}) I_{S(\psi_i)}, \quad i = 1, \dots, n, \quad (\text{A.6})$$

$$\sigma_i^2 | \cdot \sim IG(\nu^{(1)}/2, \delta^{(1)}/2), \quad i = 1, \dots, n, \quad (\text{A.7})$$

where

$$\begin{aligned} M_{\psi_i}^{(1)} &= (\Sigma_{\psi_i}^{(0)-1} + \sigma_i^{-2} X'X)^{-1} (\Sigma_{\psi_i}^{(0)-1} M_{\psi_i}^{(0)} + \sigma_i^{-2} X'Y), \\ \Sigma_{\psi_i}^{(1)} &= (\Sigma_{\psi_i}^{(0)-1} + \sigma_i^{-2} X'X)^{-1}. \end{aligned}$$

We can sample from the truncated normal distribution (A.6) by sampling from the normal distribution  $N(M_{\psi_i}^{(1)}, \Sigma_{\psi_i}^{(1)})$  and accepting it only if it is in the stationary region. It is straightforward to sample from the gamma distribution (see Ripley [1987]). Thus, we can sample  $\sigma_i^2$  from the inverted gamma distribution (A.7) by sampling  $1/\sigma_i^2$  from the gamma distribution and taking its reciprocal.

### C. Conditional Distribution (11)

The condition of (11) includes  $\Delta c$  and  $S_T$ . Given them, we can calculate

$$\begin{aligned} Y &= [(\Delta c_{p+1} - \mu_{s_{p+1}}), \dots, (\Delta c_T - \mu_{s_T})], \\ X &= \begin{bmatrix} (\Delta c_p - \mu_{s_p}) & \cdots & (\Delta c_1 - \mu_{s_1}) \\ \vdots & \ddots & \vdots \\ (\Delta c_{T-1} - \mu_{s_{T-1}}) & \cdots & (\Delta c_{T-p} - \mu_{s_{T-p}}) \end{bmatrix}, \end{aligned}$$

If we further define

$$\beta = \phi, \quad \mathbf{u} = [\nu_{p+1}, \dots, \nu_T], \quad \sigma^2 = 1,$$

equation (3') will be the linear regression model (A.1). Hence, conditional distribution (11) is given by

$$\phi | \cdot \sim N(M_\phi^{(1)}, \Sigma_\phi^{(1)}) I_{S(\phi)},$$

where

$$M_\phi^{(1)} = (\Sigma_\phi^{(0)-1} + \mathbf{X}'\mathbf{X})^{-1} (\Sigma_\phi^{(0)-1} M_\phi^{(0)} + \mathbf{X}'\mathbf{Y}),$$

$$\Sigma_\phi^{(1)} = (\Sigma_\phi^{(0)-1} + \mathbf{X}'\mathbf{X})^{-1}.$$

#### D. Conditional Distribution (12)

The condition of (12) includes  $\phi$ ,  $\Delta c_T$ , and  $S_T$ . Given them, we can calculate

$$\mathbf{Y} = \begin{bmatrix} \Delta c_{p+1} - \phi_1 \Delta c_p - \dots - \phi_p \Delta c_1 \\ \vdots \\ \Delta c_T - \phi_1 \Delta c_{T-1} - \dots - \phi_p \Delta c_{T-p} \end{bmatrix},$$

$$\mathbf{X} = \begin{bmatrix} 1 - \phi_1 - \dots - \phi_p & S_{p+1} - \phi_1 S_p - \dots - \phi_p S_1 \\ \vdots & \vdots \\ 1 - \phi_1 - \dots - \phi_p & S_T - \phi_1 S_{T-1} - \dots - \phi_p S_{T-p} \end{bmatrix}.$$

If we further define

$$\beta = \mu, \quad \mathbf{u} = [\nu_{p+1}, \dots, \nu_T], \quad \sigma^2 = 1,$$

equation (3') will be the linear regression model (A.1). Hence, the conditional distribution (12) is given by

$$\mu' | \cdot \sim N(M_\mu^{(1)}, \Sigma_\mu^{(1)}) I[\mu_0 < 0, \mu_1 > 0],$$

where

$$M_\mu^{(1)} = (\Sigma_\mu^{(0)-1} + \mathbf{X}'\mathbf{X})^{-1} (\Sigma_\mu^{(0)-1} M_\mu^{(0)} + \mathbf{X}'\mathbf{Y}),$$

$$\Sigma_\mu^{(1)} = (\Sigma_\mu^{(0)-1} + \mathbf{X}'\mathbf{X})^{-1}.$$

### E. Conditional Distribution (13)

Conditional distribution (13) can be written as

$$f(\boldsymbol{\pi}_{00}, \boldsymbol{\pi}_{11}|\cdot) \propto f(\mathcal{S}_T|\boldsymbol{\pi}_{00}, \boldsymbol{\pi}_{11})f(\boldsymbol{\pi}_{00}, \boldsymbol{\pi}_{11}),$$

where

$$f(\boldsymbol{\pi}_{00}, \boldsymbol{\pi}_{11}) \propto \boldsymbol{\pi}_{00}^{u_{00}}(1 - \boldsymbol{\pi}_{00})^{u_{01}}\boldsymbol{\pi}_{11}^{u_{11}}(1 - \boldsymbol{\pi}_{11})^{u_{10}}.$$

Once  $\mathcal{S}_T$  are given, we can obtain the number of transitions from  $S_{t-1} = i$  to  $S_t = j$ , which is denoted by  $n_{ij}$ . Then,

$$f(\mathcal{S}_T|\boldsymbol{\pi}_{00}, \boldsymbol{\pi}_{11}) \propto \boldsymbol{\pi}_{00}^{n_{00}}(1 - \boldsymbol{\pi}_{00})^{n_{01}}\boldsymbol{\pi}_{11}^{n_{11}}(1 - \boldsymbol{\pi}_{11})^{n_{10}}.$$

Hence,

$$\boldsymbol{\pi}_{00}|\cdot \sim \text{beta}(u_{00} + n_{00}, u_{01} + n_{01}),$$

$$\boldsymbol{\pi}_{11}|\cdot \sim \text{beta}(u_{11} + n_{11}, u_{10} + n_{10}).$$

We can sample from  $\text{beta}(\alpha_1, \alpha_2)$  as the ratio  $x_1/(x_1 + x_2)$ , where  $x_1$  and  $x_2$  are draws sampled from  $\text{gamma}(\alpha_1, 1)$  and  $\text{gamma}(\alpha_2, 1)$ , respectively.

### APPENDIX 2: PARTICLE FILTER

The filter density  $f(\mathbf{z}_t, S_t|\Delta\mathbf{y}_t)$ , where  $\boldsymbol{\theta}$  in the condition is omitted, can be written as

$$\begin{aligned} f(\mathbf{z}_t, S_t|\Delta\mathbf{y}_t) & \propto f(\Delta\mathbf{y}_t|\mathbf{z}_t)f(\mathbf{z}_t, S_t|\Delta\mathbf{y}_{t-1}) \\ & = f(\Delta\mathbf{y}_t|\mathbf{z}_t)\int f(\mathbf{z}_t, S_t|\mathbf{z}_{t-1}, S_{t-1})f(\mathbf{z}_{t-1}, S_{t-1}|\Delta\mathbf{y}_{t-1})d\mathbf{z}_{t-1}dS_{t-1}. \end{aligned} \quad (\text{A.8})$$

Suppose that we have  $M$  draws  $\{\mathbf{z}_{t-1}^{(m)}, S_{t-1}^{(m)}\}$  ( $m = 1, \dots, M$ ) sampled from the density  $f(\mathbf{z}_{t-1}, S_{t-1}|\Delta\mathbf{y}_{t-1})$ . Using these draws, the integral in (A.8) can be estimated as follows.

$$\begin{aligned} & \int f(\mathbf{z}_t, S_t|\mathbf{z}_{t-1}, S_{t-1})f(\mathbf{z}_{t-1}, S_{t-1}|\Delta\mathbf{y}_{t-1})d\mathbf{z}_{t-1}dS_{t-1} \\ & \approx \frac{1}{M}\sum_{m=1}^M f(\mathbf{z}_t, S_t|\mathbf{z}_{t-1}^{(m)}, S_{t-1}^{(m)}). \end{aligned} \quad (\text{A.9})$$

Substituting (A.9) into (A.8) yields

$$f(\mathbf{z}_t, S_t|\Delta\mathbf{y}_t) \approx \frac{1}{M}\sum_{m=1}^M f(\Delta\mathbf{y}_t|\mathbf{z}_t)f(\mathbf{z}_t, S_t|\mathbf{z}_{t-1}^{(m)}, S_{t-1}^{(m)}), \quad (\text{A.10})$$



where  $f(\Delta y_t | z_t) f(z_t, S_t | z_{t-1}^{(m)}, S_{t-1}^{(m)})$  on the right-hand side can be written as

$$\begin{aligned}
 & f(\Delta y_t | z_t) f(z_t, S_t | z_{t-1}^{(m)}, S_{t-1}^{(m)}) \\
 &= f(\Delta y_t | z_t) f(z_t | S_t, z_{t-1}^{(m)}) p(S_t | S_{t-1}^{(m)}) \\
 &= f(\Delta y_t, z_t | S_t, z_{t-1}^{(m)}) p(S_t | S_{t-1}^{(m)}) \\
 &= f(z_t | S_t, z_{t-1}^{(m)}, \Delta y_t) f(\Delta y_t | S_t, z_{t-1}^{(m)}) p(S_t | S_{t-1}^{(m)}) \\
 &= \eta_{S_t, m} f(z_t | S_t, z_{t-1}^{(m)}, \Delta y_t),
 \end{aligned} \tag{A.11}$$

where

$$\begin{aligned}
 \eta_{S_t, m} &= f(\Delta y_t | S_t, z_{t-1}^{(m)}) p(S_t | S_{t-1}^{(m)}) \\
 &\propto |F|^{-1/2} \exp\left(-\frac{1}{2} e_t' F^{-1} e_t\right) p(S_t | S_{t-1}^{(m)}),
 \end{aligned}$$

and where

$$\begin{aligned}
 e_t &= \Delta y_t - \Lambda(M_{S_t} + \Phi z_{t-1}^{(m)}), \\
 F &= \Lambda \sum_v \Lambda' + \sum_e \epsilon.
 \end{aligned}$$

Substituting (A.11) into (A.10) yields

$$f(z_t, S_t | \Delta y_t) \approx \frac{1}{M} \sum_{m=1}^M \eta_{S_t, m} f(z_t | S_t, z_{t-1}^{(m)}, \Delta y_t) \tag{A.12}$$

We can sample from this mixture distribution by first selecting the indices  $(S_t, m)$  with probability proportional to  $\eta_{S_t, m}$  and then sampling from  $f(z_t | S_t, z_{t-1}^{(m)}, \Delta y_t)$ , which is the normal whose mean and variance are given by

$$\begin{aligned}
 \mu_{z_t}^{(m)} &= M_{S_t} + \Phi z_{t-1}^{(m)} + \sum_v \Lambda' F^{-1} e_t, \\
 \Sigma_{z_t}^{(m)} &= \Sigma_v - \sum_v \Lambda' F^{-1} \Lambda \Sigma_v.
 \end{aligned}$$

## References

- Chen, M.-H., Q.-M. Shao, and J. G. Ibrahim, *Monte Carlo Methods in Bayesian Computation*, New York: Springer, 2000.
- Chib, S., "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, 1995, pp. 1313–1321.
- Fukuda, S., and T. Onodera, "A New Composite Index of Coincident Economic Indicators in Japan: How Can We Improve the Forecast Performance?" *International Journal of Forecasting*, 17, 2001, pp. 483–498.
- Garcia, R., "Asymptotic Null Distribution of the Likelihood Ratio Statistics," *International Economic Review*, 39, 1998, pp. 763–788.
- Geweke, J., "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," in J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds. *Bayesian Statistics 4*, Oxford: Oxford University Press, 1992, pp. 169–193.
- Hamilton, J. D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 1989, pp. 357–384.
- Hansen, B. E., "The Likelihood Ratio Test under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics*, 7, 1992, pp. S61–S82.
- Kanoh, S., and S. Saito, "Extracting Actuality from Judgement: A New Index of Business Cycle," *Monetary and Economic Studies*, 12, Institute for Monetary and Economic Studies, Bank of Japan, 1994, pp. 77–93.
- Kasuya, M., and K. Shinki, "Bukka Hendo no Tenkanten Yosoku ni Tsuite (On Forecasting the Turning Points of Price Changes)," Working Paper Series 01-20, Research and Statistics Department, Bank of Japan, 2001 (in Japanese).
- Kaufman, S., "Measuring Business Cycles with a Dynamic Markov Switching Factor Model: An Assessment Using Bayesian Simulation Methods," *Econometrics Journal*, 3, 2000, pp. 39–65.
- Kim, C.-J., and C. R. Nelson, "Business Cycle Turning Points, a New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime-Switching," *The Review of Economics and Statistics*, 80, 1998, pp. 188–201.
- , and ———, *State-Space Models with Regime Switching*, Cambridge, Massachusetts: MIT Press, 1999.
- Mori, K., M. Satake, and Y. Ohkusa, "Stock-Watson Type no Keiki Shisu: Nihon Keizai he no Oyo (Stock-Watson Type Business Cycles Index: Application to Japanese Economy)," *Doshisha University Keizaigaku Ronshu*, 45, 1993, pp. 28–50 (in Japanese).
- Ohkusa, Y., "Nihon ni Okeru Kakuritsu-Teki Keiki Shisu no Kaihatsu (Constructing a Stochastic Business Index in Japan)," *Doshisha University Keizaigaku Ronshu*, 44, 1992, pp. 25–60 (in Japanese).
- Pitt, M. K., and N. Shephard, "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association*, 94, 1999, pp. 590–599.
- Ripley, B. D., *Stochastic Simulation*, New York: Wiley, 1987.
- Rosenblatt, M., "Remarks on a Multivariate Transformation," *Annals of Mathematical Statistics*, 23, 1952, pp. 470–472.
- Stock, J. H., and M. W. Watson, "New Indexes of Coincident and Leading Macroeconomic Indicators," in O. Blanchard and S. Fischer, eds. *NBER Macroeconomic Annual*, Cambridge, Massachusetts: MIT Press, 1989, pp. 351–394.
- , and ———, "A Probability Model of the Coincident Economic Indicators," in K. Lahiri and G. Moore, eds. *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge: Cambridge University Press, 1991, pp. 63–89.