A Bayesian Approach to Testing for Markov Switching in Univariate and Dynamic Factor Models

by

Chang-Jin Kim
Korea University

and

Charles R. Nelson
University of Washington

August 25, 1998

Kim: Dept. of Economics, Korea University, Anam-dong, Seongbuk-ku, Seoul, 136-701, Korea (cjkim@kuccnx.korea.ac.kr); **Nelson:** Dept. of Economics, University of Washington, Box 353330, Seattle, WA, 98195, USA (cnelson@u.washington.edu). This paper was written while Kim was visiting the Department of economics, University of Washington. Kim greatly appreciates support from the department.

Abstract

Though Hamilton's (1989) Markov switching model has been widely estimated in various contexts, formal testing for Markov switching is not straightforward. Univariate tests in the classical framework by Hansen (1992) and Garcia (1998) do not reject the linear model for GDP. We present Bayesian tests for Markov switching in both univariate and multivariate settings based on sensitivity of the posterior probability to the prior. We find that evidence for Markov switching, and thus the business cycle asymmetry, is stronger in a switching version of the dynamic factor model of Stock and Watson (1991) than it is for GDP by itself.

Key Words:
Bayesian Model Selection, Business Cycle Asymmetry, Dynamic Factor Model, Pseudo Prior, Model Indicator Parameter, Test of Markov Switching.

JEL Classifications: C11, C12, E32.

"The Bayesian moral is simple: Never make anything more than relative probability statements about the models explicitly entertained. Be suspicious of those who promise more!" [Poirier (1995), p. 614]

1. Introduction

As Diebold and Rudebusch (1996) have pointed out, during the first half of this century research on empirical business cycle was focused on organizing business cycle regularities within a model-free framework, leading to the two defining characteristics of the business cycle by Burns and Mitchell (1946): 'comovement' and 'asymmetry'. Modern econometric research has investigated each of these two key features of the business cycle. Stock and Watson's (1991) dynamic factor model of coincident economic variables is an example that highlights the 'comovement' feature of the business cycle. Hamilton's (1989) Markov-switching model and Tong (1983) and Potter's (1995) threshold autoregressive model of real output are the representative examples that highlight the 'asymmetric' feature of the business cycle. [1] With advances in computing and the development of numerical and simulation techniques, more recent research has been devoted to an integration of the two features of the business cycle in a comprehensive time series framework (Diebold and Rudebusch (1996) and Kim and Nelson (1998a, 1998b)).

In general, however, there seems to be less consensus on the asymmetric feature of the business cycle than on the comovement among business cycle indicators [see Diebold and Rudebusch (1996), p. 75.]. Focusing on the type of asymmetry generated by Markov-switching, we find that the literature on testing procedures is relatively new and that tests have been performed only within the univariate context. While estimation of the Markov-switching models have been well developed in both the classical and the Bayesian perspectives and their applications are abundant, there apparently seems to be a lag in the literature in the development of testing procedures for Markov-switching. In most of

[1] Unlike the Markov-switching model of Hamilton (1989), the regime switches according to the observable past observations of a time series in the 'threshold' model.

applied work, Markov-switching has just been assumed without testing for its existence. Furthermore, the literature reports mixed results on empirical tests of business cycle asymmetry or Markov-switching. For example, based on the classical approach, neither Hansen (1992) nor Garcia (1998) reject the null hypothesis of no Markov-switching in quarterly real output. On the contrary, using Garcia's (1998) test, Diebold and Rudebusch (1996) report strong evidence of Markov-switching in the composite index of monthly coincident economic indicators by the Department of Commerce (DOC) as well as each individual indicators. Using a Bayesian approach, Koop and Potter (1996) conclude that the Markov-switching model and linear AR models receive roughly equal support for quarterly real output , though Chib (1995) concludes that the data support a Markov-switching model. While univariate tests in the literature have produced somewhat conflicting evidence of Markov-switching, we speculate that tests in a multivariate framework may provide more reliable and consistent results.

In this paper, we present Bayesian tests of Markov-switching in both univariate and multivariate contexts. Along the lines of the work by Carlin and Polson (1991), George and McCulloch (1993), Geweke (1996), and Carlin and Chib (1995), our testing procedure or Bayesian model selection procedure is based on the sensitivity of the posterior probability of a model indicator parameter to the prior probability. In implementing the Markov Chain Monte Carlo (MCMC) method of Gibbs sampling, a major difficulty arises since the parameter space is not fixed in the algorithm. For example, conditional on no Markov-switching the shift parameters (the parameters of interest) are zero, and thus, the state vector and the transition probabilities that describe the dynamics of the state vector are not identified. This potentially causes a convergence problem in the Gibbs sampler, as in Carlin and Chib (1995). To overcome the problem of convergence, we employ a pseudo prior for the shift parameters that are otherwise set to zero conditional on no Markov-switching. We first present our procedure for Bayesian model selection and the modified Gibbs sampler within a relatively straightforward univariate framework. We then extend our univariate procedure to the multivariate framework of the dynamic factor models of business cycle by Stock and Watson (1991) and Diebold and Rudebusch (1996). In the multivariate framework, an additional difficulty arises since we want to test for Markov-switching in the common factor component that is unobserved. This is overcome by combining the univariate test with Kim and Nelson's (1998a) Bayesian approach to the dynamic factor models.

In Section 2, we present model specifications employed. Conditional on no Markov-switching, Hamilton's (1989) univariate model collapses to linear autoregressive model and Diebold and Rudebusch's (1996) multivariate model collapses to Stock and Watson's (1991) linear dynamic factor model. Basic issues associated with our tests are then discussed. Section 3 deals with a univariate test of Markov-switching. In Section 4, the basic test within a univariate framework is extended to a multivariate framework. In Section 5, For a univariate test, we employing as data quarterly real GDP growth (1952.II-1997.II). For a multivariate test, we employ as data the four monthly series used by the Department of Commerce (DOC) to construct its index of coincident indicators (1960.1-1995.1). Section 6 concludes the paper.

2. Model Specifications and Problem Setup

2.1. A Model Specification for a Univariate Test

For a univariate test of Markov-switching, we first consider the following model for a univariate process ΔC_t, in which a model indicator parameter (τ) is employed to represent both a linear AR process and an AR process with Markov-switching mean:

$$\phi(L)(\Delta C_t - \mu_{s_t} - \delta) = v_t, \quad v_t \sim i.i.d.N(0, \sigma^2), \quad \tau = 0 \ or \ 1 \tag{1}$$

$$\mu_{s_t} = \mu_0(1 - S_t) + \mu_1 S_t, \tag{2}$$

where the unobserved state variable S_t evolves according to a Markov-switching process with the transition probabilities given by:

$$Pr[S_t = 0|S_{t-1} = 0] = p_{00}, \ Pr[S_t = 1|S_{t-1} = 1] = p_{11}, \tag{3}$$

and the parameters μ_0 and μ_1 are defined as:

$$
\mu_0 \quad \begin{cases} = 0, & \text{if } \tau = 0; \\ = \mu_0^1 \sim N(\omega_0, -{}_0)_{1[\mu_0^1 < 0]}, & \text{if } \tau = 1, \end{cases} \tag{4}
$$

$$
\mu_1 \quad \begin{cases} = 0, & \text{if } \tau = 0; \\ = \mu_1^1 \sim N(\omega_1, -{}_1)_{1[\mu_1^1 > 0]}, & \text{if } \tau = 1, \end{cases} \tag{5}
$$

where 1[.] refers to an indicator function. Thus, conditional on $\tau = 0$, we have a linear AR model and conditional on $\tau = 1$, we have Hamilton's (1989) Markov-switching model.

In the above specification, the parameter δ determines the long-run growth rate of ΔC_t. Conditional on $\tau = 1$ (a Markov-switching model), μ_{s_t} represents a deviation of ΔC_t from its long-run growth δ. Correspondingly, the growth rate of ΔC_t during a recession is given by $\delta + \mu_0^1 < \delta$ and that during a boom is given by $\delta + \mu_1^1 > \delta$. The parameters δ, μ_0^1, and μ_1^1, however, are not separately identified due to over-parameterization, conditional on $\tau = 1$. We solve the problem of over-parameterization is by expressing the data in deviation from mean, since then the long run growth rate δ disappears from equation (1), and we have:

$$
\phi(L)(\Delta c_t - \mu_{s_t}) = v_t, \quad v_t \sim i.i.d. N(0, \sigma^2), \tag{1$'$}
$$

where $\Delta c_t = \Delta C_t - \Delta \bar{C}$. In this specification, a linear model is nested within a Markov-switching model.

An alternative way of avoiding the problem of over-parameterization in (1) conditional on $\tau = 1$ would be to specify the model as:

$$
\phi(L)(\Delta C_t - (\mu_0^* + \mu_d S_t)) = v_t, \quad v_t \sim i.i.d. N(0, \sigma^2), \quad \mu_d \geq 0, \tag{6}
$$

where $\mu_0^* = \delta + \mu_0$ and $\mu_d = \mu_1 - \mu_0$. A linear model is obtained by the constraint $\mu_d = 0$. In this specification, however, a linear model is not really nested within a Markov-switching model. This is clear by examining the μ_0^* parameter in (6). μ_0^* is not a parameter common to both models. For example, we have $\mu_0^* = \delta$ for a linear model, while we have $\mu_0^* = \delta + \mu_0^1 < \delta$ for a Markov-switching model. That is, the parameter μ_0^* is model-dependent and it has different interpretations for the two competing models.

Different specifications of the model (equation (1') and equation (6)) do not affect inferences about the parameters of alternative models and the unobserved state S_t conditional on $\tau = 1$, within either the classical or the Bayesian framework. When we come to the hypothesis testing, however, they may have different implications for the testing procedure within the Bayesian framework. [2] If one adopts the model specification in (6) within the framework discussed in this paper, for example, the model-dependent nature of the μ_0^* parameter would have to be taken into account when designing a test. Throughout this paper, we stick to the model written in deviation from mean (equation (1')).

2.2. A Model Specification for a Multivariate Test: A Dynamic Factor Model

While C_t is an observed series in the specification for a univariate test in Section 2.1, we consider a case in which C_t is an unobserved component which is common to more than one observed coincident economic variables (Y_{it}, $i = 1, 2, .., n$) for a multivariate test. If each observed variable has a unit root and the variables are not cointegrated, the ΔC_t term in equation (1) is a common factor component in the following model (Stock and Watson (1991) and Diebold and Rudebusch (1996)):

$$\Delta Y_{it} = \gamma_i(L)\Delta C_t + D_i + e_{it}, \quad i = 1, 2, .., n, \tag{7}$$

$$\psi_i(L)e_{it} = \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d.N(0, \sigma_i^2) \tag{8}$$

where roots of $\psi_i(z) = 0$, $i = 1, .., n$, lie outside the complex unit circle; ϵ_{it}, $i = 1, .., n$, and v_t are independent of one another. Each observed series ΔY_{it} consists of an individual component ($D_i + e_{it}$) and a linear combination of current and lagged values of the common factor component ($\gamma_i(L)\Delta C_t$). C_t has an interpretation of the index of coincident economic indicators. Thus, the model potentially captures the two defining features of the business cycle established by Burns and Mitchell (1946): comovement and asymmetry.

As the model given by (1)-(5) and (7)-(8) is not identified due to over-parameterization of the mean of data, ΔY_{it}, we first express data as deviation from means. Also for

[2] For a discussion of related issues, see Zivot (1994). In the classical framework, different specifications of the model may not affect the testing procedure, as the asymptotic distribution or a bound for the asymptotic distribution (Hansen (1992)) is obtained under the null hypothesis, which is assumed true.

identification purpose, we set $\sigma^2 = 1$. Then the full model on which our test will be based is given by:

Model

$$\Delta y_{it} = \gamma_i(L)\Delta c_t + e_{it}, \quad i = 1, 2, .., n, \tag{7'}$$

$$\psi_i(L)e_{it} = \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d.N(0, \sigma_i^2) \tag{8}$$

$$\phi(L)(\Delta c_t - \mu_{s_t}) = v_t, \quad v_t \sim i.i.d.N(0, \sigma^2), \quad \tau = 0 \ or \ 1 \tag{1'}$$

$$\mu_{s_t} = \mu_0(1 - S_t) + \mu_1 S_t, \tag{2}$$

$$Pr[S_t = 0|S_{t-1} = 0] = p_{00}, \ Pr[S_t = 1|S_{t-1} = 1] = p_{11}, \tag{3}$$

$$\mu_0 \begin{cases} = 0, & \text{if } \tau = 0; \\ = \mu_0^1 \sim N(\omega_0, \text{-}\ _0)_{1[\mu_0^1<0]}, & \text{if } \tau = 1, \end{cases} \tag{4}$$

$$\mu_1 \begin{cases} = 0, & \text{if } \tau = 0; \\ = \mu_1^1 \sim N(\omega_1, \text{-}\ _1)_{1[\mu_1^1>0]}, & \text{if } \tau = 1, \end{cases} \tag{5}$$

where $\Delta y_{it} = \Delta Y_{it} - \Delta\bar{Y}_i$; $\Delta c_t = \Delta C_t - \delta$; and $1[.]$ refers to an indicator function. Conditional on $\tau = 0$, we have a linear dynamic factor model of Stock and Watson (1991) and conditional on $\tau = 1$, we have a dynamic factor model with Markov-switching of Diebold and Rudebusch (1996).

2.3. Problem Setup

Assume that data $\Delta\tilde{z}_T = [\,\Delta z_1 \quad \ldots \quad \Delta z_T\,]'$ have arisen from either a linear model ($\tau = 0$) or a Markov-switching model ($\tau = 1$) according to a probability function (marginal likelihood) $p(\Delta\tilde{z}_T|\tau = 0)$ or $p(\Delta\tilde{z}_T|\tau = 1)$, where $\Delta z_t = \Delta c_t$ in the univariate framework of Section 2.1, and $\Delta z_t = [\,\Delta y_{1t} \quad \ldots \quad \Delta y_{nt}\,]'$ in the multivariate framework of Section 2.2. Then, given prior probabilities for the model indicator parameter, $\underline{\pi}_1 = Pr(\tau = 1)$ and $\underline{\pi}_0 = 1 - \underline{\pi}_1$, the data $\Delta\tilde{z}_T$ produce posterior probabilities, $\bar{\pi}_1 = Pr(\tau = 1|\Delta\tilde{z}_T)$ and $\bar{\pi}_0 = 1 - \bar{\pi}_1$, according to:

$$\bar{\pi}_1 = \frac{p(\Delta\tilde{z}_T|\tau = 1)\underline{\pi}_1}{p(\Delta\tilde{z}_T|\tau = 1)\underline{\pi}_1 + p(\Delta\tilde{z}_T|\tau = 0)\underline{\pi}_0} = \frac{B_{10}\underline{\pi}_1}{B_{10}\underline{\pi}_1 + (1 - \underline{\pi}_1)}, \tag{9}$$

where B_{10} is the Bayes factor in favor of a Markov-switching model. By rearranging equation (9), the Bayes factor, which is given by the ratio of the marginal likelihoods for the two alternative models, can be shown as summarizing the effect of data in modifying the prior odds ($\underline{\pi}_1/(1-\underline{\pi}_1)$) to obtain posterior odds ($\bar{\pi}_1/(1-\bar{\pi}_1)$):

$$B_{10} = \frac{p(\Delta\tilde{z}_T|\tau = 1)}{p(\Delta\tilde{z}_T|\tau = 0)} = \frac{\bar{\pi}_1/(1-\bar{\pi}_1)}{\underline{\pi}_1/(1-\underline{\pi}_1)}. \tag{10}$$

The posterior distributions of the parameters for given τ are readily available via the Markov Chain Monte Carlo (MCMC) method of Gibbs sampling as in Albert and Chib (1993) and Kim and Nelson (1998a) for the univariate model in Section 2.1 and the multivariate model in Section 2.2, respectively. However, the computation of the marginal likelihood based on the posterior distribution would be more difficult since the marginal likelihood is obtained by integrating the likelihood function with respect to the prior density, not with respect to the posterior density. See Kass and Raftery (1995) for a comprehensive review of the issues related to the Bayes factors.

Recent attempts to compute the marginal likelihoods and the Bayes factor within the univariate framework with potential Markov-switching in Section 2.1 include Koop and Potter (1996) and Chib (1995). For example, Koop and Potter (1996) employ the 'Savage density ratio method' by Dickey (1971). As the linear model is nested within the Markov-switching model, the Bayes factor in favor of the Markov-switching model may be simplified to be the ratio of the marginal posterior density of the shift parameters ($\tilde{\mu}^1 = [\,\mu_0^1 \quad \mu_1^1\,]'$) to prior density, conditional on $\tau = 1$. In order to employ the 'Savage density ratio method', one of the necessary conditions that needs to be satisfied would be:

$$p(\tilde{\phi}|\tau = 0) = p(\tilde{\phi}|\tilde{\mu} = 0, \tau = 1), \tag{11}$$

where $\tilde{\phi} = [\,\phi_1 \quad \ldots \phi_r\,]'$ is the vector of autoregressive parameters for Δc_t. However, forcing the shift parameters to be zero when a Markov-switching process is a true data generating process may potentially result in more persistent autoregressive parameters than otherwise, as implied by Perron (1990).

Chib's (1995) approach to calculating the marginal likelihoods (and the Bayes factor) that relies on the output from the Gibbs sampling algorithm would be more appropriate for our purpose. However, even though Chib's approach is readily available within the univariate framework in Section 2.1, extending the approach to the multivariate framework in Section 2.2 would be extremely challenging, in the presence of multiple blocks of parameter vectors, and especially, in the presence of the two blocks of latent variables ($\tilde{S}_T = [\, S_1 \quad \ldots \quad S_T \,]'$ and $\Delta c_T = [\, \Delta c_1 \ldots \quad \Delta c_T \,]'$), conditional on $\tau = 1$.

In this paper, we deal with such difficulties by computing the Bayes factors without attempting to calculate the marginal likelihoods. Along the lines of the work by Carlin and Polson (1991), George and McCulloch (1993), Geweke (1996), and Carlin and Chib (1995), our Bayesian model selection procedure is based on the sensitivity of the posterior probability of the model indicator parameter τ to the prior probability. Different prior probabilities for the model indicator parameter, when combined with data, could be associated with different values for the Bayes factors, suggesting different effects of data for different priors in modifying the prior odds to obtain the posterior odds. An additional advantage of the approach in this paper is that it also provides the sensitivity of the Bayes factor to different prior probabilities unlike the usual approach based on a direct calculation of the marginal likelihoods. In the usual approach, the effect of data in modifying the prior odds to obtain the posterior odds are assumed the same for different prior probabilities.

In implementing the MCMC method of Gibbs sampling to sample from an appropriate joint posterior distribution of the model indicator parameter τ, the other parameters of the models, and the latent variable(s), one potential problem is that the parameter space is not fixed in the algorithm. For example, conditional on $\tau = 1$, we have $\tilde{\mu} = \tilde{\mu}^1$ and all the variates are well identified, where $\tilde{\mu} = [\, \mu_0 \quad \mu_1 \,]'$ and $\tilde{\mu}^1 = [\, \mu_0^1 \quad \mu_1^1 \,]'$. Conditional on $\tau = 0$, however, we have $\tilde{\mu} = 0$ and a vector of transition probabilities $\tilde{p} = [\, p_{00} \quad p_{11} \,]'$ and a vector of latent state variables $\tilde{S}_T = [\, S_1 \quad \ldots \quad S_T \,]'$ are not identified. Thus, the vectors $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ are forced out of the model for both the univariate and multivariate models and the Gibbs sampler skips a generation of these vectors. This potentially creates an absorbing state, which is a violation of a condition for the convergence of Gibbs sampling (Tierney (1994) and Carlin and Chib (1995)).

The problem is solved by employing a 'pseudo prior' for $\tilde{\mu}^1$ conditional on $\tau = 0$, in the line of the work by Carlin and Chib (1995). In fact, the approach in this paper may be considered as an extension of Carlin and Chib (1995) to the case of testing for Markov-switching. Given pseudo values for $\tilde{\mu}^1$ conditional on $\tau = 0$, corresponding vectors $\tilde{p}$ and $\tilde{S}$ are pseudo-identified. Sections 3 and 4 discuss how the usual Gibbs sampling method may be modified without skipping a generation of $\tilde{\mu}^1$, $\tilde{S}_T$, and $\tilde{p}$ conditional on $\tau = 0$, by taking advantage of the pseudo prior for $\tilde{\mu}^1$.

3. A Univariate Test of Markov-Switching

3.1. Technical Development

Within the univariate framework in Section 2.1, the problem of Bayesian model selection reduces to generating τ, $\tilde{\theta}_c = [\,\tilde{\phi}' \quad \sigma^2\,]'$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$ from the joint posterior distribution:

$$
\begin{aligned}
p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tilde{S}_T, \tau | \Delta\tilde{c}_T) &\propto p(\Delta\tilde{y}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tilde{S}_T, \tau) \\
&= p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tilde{S}_T, \tau) p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau) p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu} | \tau) p(\tau),
\end{aligned}
\tag{12}
$$

where $\tilde{\theta}_c$ is a vector of parameters common to both the linear and Markov-switching models; $\tilde{p}$ and $\tilde{S}_T = [\,S_1 \quad \ldots \quad S_T\,]'$ are the variates that are identified only under the Markov-switching model; $\Delta\tilde{c}_T = [\,\Delta c_1 \quad \ldots \quad \Delta c_T\,]'$ is a vector of observed data. Thus, for given priors for the parameters, our goal would to generate, via Gibbs sampling, $\tilde{\theta}_c$, $\tilde{\mu}^1$, $\tilde{p}$, $\tilde{S}_T$, and the model indicator parameter τ from appropriate full conditional distributions.

Conditional on $\tau = 1$, all the variates are identified and the posterior distribution of $\tilde{\theta}_c$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$ is given by:

$$
\begin{aligned}
&p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T | \Delta\tilde{c}_T, \tau = 1) \\
&\quad \propto p(\Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T, | \tau = 1) \\
&\quad = p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T, \tau = 1) p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 1) \\
&\quad = p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T | \tau = 1) p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\tilde{\theta}_c | \tau = 1) p(\tilde{p} | \tau = 0) p(\tilde{\mu}^1 | \tau = 1),
\end{aligned}
\tag{13}
$$

where $p(\tilde{\theta}_c|\tau = 1)$, $p(\tilde{p}|\tau = 1)$, $p(\tilde{\mu}^1|\tau = 1)$ are the usual prior densities, and it is a priori assumed that $\tilde{\theta}_c$, $\tilde{p}$, and $\tilde{\mu}^1$ are independent of one another over the admissible regions.

Conditional on $\tau = 0$, however, the variates $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ are forced out of the model and the posterior distribution is given by:

$$\begin{aligned} p(\tilde{\theta}_c|\Delta\tilde{c}_T, \tau = 1) &\propto p(\Delta\tilde{c}_T, \tilde{\theta}_c|\tau = 0) \\ &= p(\Delta\tilde{c}_T|\tilde{\theta}_c, \tau = 0)p(\tilde{\theta}_c|\tau = 0), \end{aligned} \tag{14}$$

where $p(\tilde{\theta}_c|\tau = 0)$ is the usual prior density, and the usual Gibbs sampler skips a generation of $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$. To circumvent the problem of slow convergence in the presence of these variates forced out of the model and the Gibbs sampler, we adopt a pseudo prior for $\tilde{\mu}^1$ in the line of the work by Carlin and Chib (1995). Given the pseudo prior for $\tilde{\mu}^1$, the variates $\tilde{p}$ and $\tilde{S}_T$ are then pseudo-identified under the linear model ($\tau = 0$). The following joint posterior density of $\tilde{\theta}_c$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$ provides us how the Gibbs sampling may proceed without forcing these variates out of the model conditional on $\tau = 0$:

$$\begin{aligned} p(\tilde{\theta}_c, &\tilde{S}_T, \tilde{p}, \tilde{\mu}^1|\Delta\tilde{c}_T, \tau = 0) \\ &\propto p(\Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{S}_T, \tilde{p}, \tilde{\mu}^1|\tau = 0) \\ &= p(\Delta\tilde{c}_T|\theta_c, \tilde{S}_T, \tilde{p}, \tilde{\mu}^1, \tau = 0)p(\tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1|\tau = 0) \\ &= p(\Delta\tilde{c}_T|\theta_c, \tau = 0)p(\tilde{S}_T|\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0)p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1|\tau = 0) \\ &= p(\Delta\tilde{c}_T|\theta_c, \tau = 0)p(\tilde{\theta}_c|\tau = 0)p(\tilde{S}_T|\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0)p(\tilde{p}|\tilde{\mu}^1, \tau = 0)p(\tilde{\mu}^1|\tau = 0), \end{aligned} \tag{14'}$$

where $\tilde{p}$ and $\tilde{\mu}^1$ are a priori assumed independent of $\tilde{\theta}_c$ conditional on $\tau = 0$; $p(\tilde{\theta}_c|\tau = 0)$ is the usual prior density; $p(\tilde{\mu}^1|\tau = 0)$ is the pseudo prior density of interest.

Equation (14') suggests that the vector of the parameters common to both models, $\tilde{\theta}_c$, is generated in the usual fashion conditional on data, given $\tilde{\mu} = 0$. Pseudo generation of μ^1, $\tilde{S}_T$, and $\tilde{p}$ can be done as follows: first, we generate $\tilde{\mu}^1$ from an appropriate pseudo prior distribution; then, conditional on $\tilde{\theta}_c$, $\tilde{p}$, the pseudo value for $\tilde{\mu}^1$, and data, we can generate a pseudo vector of the state variables, $\tilde{S}_T$; finally, conditional on $\tilde{S}_T$, a pseudo vector of the transition probabilities $\tilde{p}$ can be generated.

Notice that, as in the case of Carlin and Chib (1995), the presence of the pseudo vectors $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ in equation (14') does not affect the marginal likelihood for the linear model ($\tau = 0$) at all, since integrating the right-hand-side of equation (14') with respect to these variates will result in the right-hand-side of equation (14), proving their irrelevance. As the pseudo prior for $\tilde{\mu}^1$ and the corresponding pseudo vectors for $\tilde{p}$ and $\tilde{S}_T$ do not affect the marginal likelihood, their existence does not affect the Bayes factor or inferences for the Bayesian model selection. The pseudo prior, or the linking density in the terminology of Carlin and Chib (1995), is employed to ensure the convergence of the Gibbs sampler.

3.2. Details of Gibbs Sampling and Full Conditional for τ

With an implementation of the pseudo prior for $\tilde{\mu}^1$ for $\tau = 0$ (linear AR model), a modified version of Gibbs sampling can be applied to generate τ, $\tilde{\theta}_c = [\,\tilde{\phi}' \quad \sigma^2\,]'$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$ from the joint posterior distribution. Given appropriate prior distributions and the pseudo prior distribution for $\tilde{\mu}^1$, and for given arbitrary starting values for all the parameters, the following provides a full description of the modified Gibbs sampling procedure, designed based on equations (12), (13) and (14'):

<u>Step 1:</u>

Generate τ from $p(\tau|\tilde{\phi}, \sigma^2, \tilde{\mu}, \tilde{p}, \Delta\tilde{c}_T)$.

<u>Step 2:</u>

<u>If $\tau = 0$:</u>

i) Generate $\tilde{\mu}^1$ from the pseudo prior distribution, $p(\tilde{\mu}^1|\tau = 0)$.

ii) Set $\tilde{\mu} = 0$.

<u>If $\tau = 1$:</u>

i) Generate $\tilde{\mu}^1$ from $p(\tilde{\mu}^1|\tilde{\phi}, \tilde{\sigma}^2, \tilde{p}, \tilde{S}_T, \Delta\tilde{c}_T, \tau = 1)$.

ii) Set $\tilde{\mu} = \tilde{\mu}^1$.

<u>Step 3:</u>

Generate $\tilde{\phi}$ and σ^2 from $p(\tilde{\phi}, \sigma^2|\tilde{\mu}, \tilde{S}_T, \Delta\tilde{c}_T)$, where, if conditional on $\tilde{\mu} = 0$, $\tilde{S}_T$ is irrelevant.

<u>Step 4:</u>

Generate $\tilde{S}_T$ from $p(\tilde{S}_T|\tilde{\phi}, \sigma^2, \tilde{p}, \tilde{\mu}^1, \Delta\tilde{c}_T)$.

Step 5:

Generate $\tilde{p}$ from $p(\tilde{p}|\tilde{S}_T)$, where, conditional on $\tilde{S}_T$, $\tilde{p}$ is independent of data and all other parameters.

Step 6: Go to Step 1.

To complete the above procedure, we need to specify the full conditional distribution for each of the variates to be generated. Full conditional distributions for $\tilde{\phi}$, σ^2, $\tilde{\mu}^1$, $\tilde{p}$ conditional on $\tau = 1$ are the same as those in Albert and Chib (1993). Concerning the generation of $\tilde{S}_T$, even though Albert and Chib's single-move algorithm is readily available, we adopt the multi-move Gibbs sampling employed in Kim and Nelson (1998a) for its simplicity.

Consider the distribution of τ conditional the parameters of the model, $\tilde{\phi}$, σ^2, $\tilde{p}$, and $\tilde{\mu}$, and data, but not on $\tilde{S}_T$:

$$p(\tau = 1|\Delta\tilde{c}_T, \tilde{\phi}, \sigma^2, \tilde{\mu}, \tilde{p}) = \frac{p(\Delta\tilde{c}_T, \tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}|\tau = 1)\underline{\pi}_1}{\sum_{j=0}^{1} p(\Delta\tilde{c}_T, \tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}|\tau = j)\underline{\pi}_j}, \tag{15}$$

where $\underline{\pi}_j$ is the prior probability that $\tau = j$. Dividing both the numerator and the denominator of the right-hand-sides of equation (15) by $p(\Delta\tilde{c}_T, \tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}|\tau = 0)$ and rearranging terms, we get:

$$p(\tau = 1|\Delta\tilde{c}_T, \tilde{\phi}, \sigma^2, \tilde{\mu}, \tilde{p}) = \frac{C_{10}\underline{\pi}_1}{C_{10}\underline{\pi}_1 + (1 - \underline{\pi}_1)}, \tag{15$'$}$$

where C_{10} is the conditional Bayes factor in favor of the Markov-switching model ($\tau = 1$) given by:

$$\begin{aligned} C_{10} &= \frac{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 1)p(\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}|\tau = 1)}{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 0)p(\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}|\tau = 0)} \\ &= \frac{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 1)p(\tilde{\phi}|\sigma^2, \tau = 1)p(\sigma^2|\tau = 1)p(\tilde{\mu}^1|\sigma^2, \tau = 1)p(\tilde{p}|\tau = 1)}{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tau = 0)p(\tilde{\phi}|\sigma^2, \tau = 0)p(\sigma^2|\tau = 0)p(\tilde{\mu}^1|\tau = 0)p(\tilde{p}|\tau = 0)} \\ &= \frac{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 1)p(\tilde{\mu}^1|\sigma^2, \tau = 1)}{p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tau = 0)p(\tilde{\mu}^1|\tau = 0)}, \end{aligned} \tag{16}$$

where it is a priori assumed that $\tilde{\mu}^1$ and $\tilde{p}$ are independent of the other parameters conditional on $\tau = 0$. It is also assumed that priors for these common parameters are the same for both models ($\tau = 0, 1$) and that $p(\tilde{p}|\tilde{\mu}^1, \tau = 0) = p(\tilde{p}|\tau = 1)$ without loss of generality. The term $p(\Delta\tilde{c}_T|\tilde{\phi}, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 1)$ can be evaluated as a byproduct of running the Hamilton filter (1989), given the conditioning parameters. To generate τ, we generate a random number from a Uniform distribution in the interval $[0, 1]$. If the generated random number is less than or equal to the value calculated using (15'), we set $\tau = 1$; otherwise, we set $\tau = 0$.

4. A Multivariate Extension of the Test Based on a Dynamic Factor Model

4.1. Technical Development

In a multivariate framework of the dynamic factor model in Section 2.2, additional difficulty arises since we want to test for Markov-switching in the common factor component that is unobserved. $\Delta\tilde{c}_T = [\,\Delta c_1 \quad \ldots \quad \Delta c_T\,]'$ is no longer a vector of observed data. It is a vector of latent factor component common to multiple observed series. We denote $\Delta\tilde{y}_T = [\,\Delta y_1' \quad \ldots \quad \Delta y_T'\,]'$ to be a $T \times n$ matrix of data on the observed series, where $\Delta y_t = [\,\Delta y_{1t} \ldots \quad \Delta y_{nt}\,]'$. We also define a vector of parameters common to both linear and Markov-switching models to be $\tilde{\theta}_c = [\,\tilde{\gamma}' \quad \tilde{\sigma}^{2\prime} \quad \tilde{\psi}' \quad \tilde{\phi}'\,]'$, where $\tilde{\gamma}$ and $\tilde{\psi}$ are the vectors of parameters associated with $\gamma_i(L)$ and $\psi_i(L)$, respectively, for $i = 1, .., n$; $\tilde{\sigma}^2 = [\,\sigma_1^2 \quad ... \quad \sigma_n^2\,]'$; $\tilde{\phi}$ is the vector of parameters associated with $\phi(L)$. The other notations used in this section are the same as in Section 3.

Consider the following decomposition of the joint posterior density of $\Delta\tilde{c}_T$, $\tilde{S}_T$, $\tilde{\theta}_c$, $\tilde{p}$, $\tilde{\mu}$, and τ, based on Bayes theorem:

$$
\begin{aligned}
& p(\Delta\tilde{c}_T, \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau|\Delta\tilde{y}_T) \\
& \propto p(\Delta\tilde{c}_T, \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau, \Delta\tilde{y}_T) \qquad (17) \\
& = p(\Delta\tilde{y}_T|\Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau)p(\Delta\tilde{c}_T|\tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau)p(\tilde{S}_T|\tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \tau)p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}|\tau)p(\tau),
\end{aligned}
$$

which allows us to sample the variates τ, $\Delta\tilde{c}_T$, $\tilde{S}_T$, $\tilde{\theta}_c$, $\tilde{\mu}$, $\tilde{p}$ from the full conditional distributions.

Conditional on $\tau = 1$, we have $\tilde{\mu} = \tilde{\mu}^1$ and all the variates $\Delta\tilde{c}_T$, $\tilde{\theta}_c$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$ are identified and the conditional joint posterior density from which these variates are to be drawn is given by:

$$
\begin{aligned}
p(\Delta\tilde{c}_T, &\tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \Delta\tilde{y}_T, \tau = 1) \\
&\propto p(\Delta\tilde{c}_T, \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \Delta\tilde{y}_T | \tau = 1) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\Delta\tilde{c}_T | \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) \\
&\qquad \times p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\tilde{\theta}_c | \tau = 1) p(\tilde{p} | \tau = 1) p(\tilde{\mu}^1 | \tau = 1),
\end{aligned} \tag{18}
$$

where $\tilde{\theta}_c$, $\tilde{p}$ and $\tilde{\mu}^1$ are a priori assumed independent conditional on $\tau = 1$, and $p(\tilde{\theta}_c|\tau = 1)$, $p(\tilde{p}|\tau = 1)$, and $p(\tilde{\mu}^1|\tau = 1)$ are the usual prior densities.

Conditional on $\tau = 0$, however, we have $\tilde{\mu} = 0$ and the variates $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ are not identified and they do not show up in the conditional joint posterior density:

$$
\begin{aligned}
p(\Delta\tilde{c}_T, \tilde{\theta}_c | \Delta\tilde{y}_T, \tau = 0) &\propto p(\Delta\tilde{y}_T, \Delta\tilde{c}_T, \tilde{\theta}_c | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tau = 0) p(\tilde{\theta}_c | \tau = 0),
\end{aligned} \tag{19}
$$

where $p(\tilde{\theta}_c|\tau = 0)$ is the usual prior density. The Gibbs sampler skips a generation of $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$.

In order not to force $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ out of the model and the algorithm conditional on $\tau = 0$, we employ a pseudo prior for $\tilde{\mu}^1$, as in Section 3.1. While $\tilde{p}$ and $\tilde{S}_T$ are not identified conditional on $\tau = 0$, pseudo vectors of $\tilde{p}$ and $\tilde{S}_T$ can be defined that correspond to the pseudo values for $\tilde{\mu}^1$. Also defined is the pseudo vector of the common factor component, $\Delta\tilde{c}_T^*$, that corresponds to the pseudo values for $\tilde{\mu}^1$ conditional on $\tau = 0$. A reason for an introduction of $\Delta\tilde{c}_T^*$ will be clear later.

Now, instead of equation (19), we consider the following joint posterior density of $\Delta\tilde{c}_T$

and $\tilde{\theta}_c$, along with the pseudo vectors, $\tilde{\mu}^1$, $\tilde{p}$, $\tilde{S}_T$, and $\Delta\tilde{c}_T^*$, conditional on $\tau = 0$:

$$\begin{aligned}
&p(\Delta\tilde{c}_T, \tilde{\theta}_c, \Delta\tilde{c}_T^*, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T | \Delta\tilde{y}_T, \tau = 0) \\
&\propto p(\Delta\tilde{c}_T, \tilde{\theta}_c, \Delta\tilde{c}_T^*, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T, \Delta\tilde{y}_T | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \Delta\tilde{c}_T^*, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T, \tau = 0) p(\Delta\tilde{c}_T, \tilde{\theta}_c, \Delta\tilde{c}_T^*, \tilde{p}, \tilde{\mu}^1, \tilde{S}_T | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T, \Delta\tilde{c}_T^*, \tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T | \Delta\tilde{c}_T^*, \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) \\
&\qquad \times p(\Delta\tilde{c}_T^*, \tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tau = 0) \\
&\qquad \times p(\Delta\tilde{c}_T^* | \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 0) \\
&= p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tau = 0) p(\tilde{\theta}_c | \tau = 0) \\
&\quad \times p(\Delta\tilde{c}_T^* | \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0) p(\tilde{p} | \mu^1, \tau = 0) p(\tilde{\mu}^1 | \tau = 0),
\end{aligned} \tag{19'}$$

where $\tilde{\theta}_c$ a priori assumed independent of $\tilde{p}$ and $\tilde{\mu}^1$ conditional on $\tau = 0$; $p(\tilde{\theta}_c | \tau = 0)$ is the usual prior density for $\tilde{\theta}_c$; $p(\tilde{\mu}^1 | \tau = 0)$ is the pseudo prior density of interest which serves as a linking density in the Gibbs sampler; $p(\tilde{p} | \tilde{\mu}^1, \tau = 0)$ may also be called a pseudo prior density but it does not serve as a linking density.

The densities $p(\Delta\tilde{c}_T^* | \tilde{S}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0)$ and $p(\tilde{S}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 0)$ describe the pseudo conditional distributions for $\Delta\tilde{c}_T$ and $\tilde{S}_T$ that correspond to pseudo prior for $\tilde{\mu}^1$. An introduction of the pseudo vectors $\Delta\tilde{c}_T^*$, $\tilde{S}_T$, and $\tilde{p}$ conditional on $\tau = 0$, along with a pseudo prior for $\tilde{\mu}^1$, does not affect the marginal likelihood and the inference at all. The is because integrating the right-hand-side of (19') with respect to $\tilde{\theta}_u$, $\tilde{S}_T$, and $\Delta\tilde{c}_T^*$ result in the right-hand-side of (19), which proves an irrelevance of $\tilde{\theta}_u$, $\tilde{S}_T$, and $\Delta\tilde{c}_T^*$. Equation (19'), however, provides us with how we may proceed with the Gibbs sampling procedure conditional on $\tau = 0$, avoiding the problem of slow convergence due to non-generation of $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$. For example, conditional on $\tau = 0$, $\Delta\tilde{c}_T$ and $\tilde{\theta}_c$ are drawn from $p(\Delta\tilde{c}_T | \tilde{\theta}_c, \Delta\tilde{y}_T, \tau = 0)$ and $p(\tilde{\theta}_c | \Delta\tilde{c}_T, \Delta\tilde{y}_T, \tau = 0)$, respectively, in the usual fashion. In order to generate $\tilde{\mu}^1$, $\tilde{p}$, and $\tilde{S}_T$ conditional on $\tau = 0$ at a particular run of the Gibbs sampler, we can proceed as follows:

i) Draw $\tilde{\mu}^{1s}$ from the pseudo prior distribution, $p(\tilde{\mu}^1 | \tau = 0)$;

ii) Draw $\Delta\tilde{c}_T^{*s}$ from $p(\Delta\tilde{c}_T|\tilde{\theta}_c, \tilde{p}, \tilde{\mu}^{1s}, \tilde{S}_T, \Delta\tilde{y}_T)$;

iii) Draw $\tilde{S}_T^s$ from $p(\tilde{S}_T|\Delta\tilde{c}_T^{*s}, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^{1s}, \Delta\tilde{y}_T)$;

iv) Draw $\tilde{p}^s$ from $p(\tilde{p}|\tilde{S}_T^s)$,

where the superscript s denotes that the variate is associated with the pseudo priors for $\tilde{\mu}^1$.

4.2. Details of Gibbs Sampling and Full Conditional for τ

In designing the Gibbs sampling procedure based on (17), (18), and (19'), the following consideration would be useful, due to the independence of the shocks and the hierarchical structure of the model for given τ: First, conditional on $\tilde{S}_T$ and all the parameters, the model is a linear Gaussian state space model and we can generate Δc_T using the procedure proposed by Carter and Kohn (1994). Second, conditional on $\Delta\tilde{c}_T$ and $\tilde{\theta}$, we can focus on (1') to generate $\tilde{S}_T$. Third, conditional on $\Delta\tilde{c}_T$ and $\tilde{S}_T$, each of the $n+1$ equations in (7') and (1') can be treated separately to generate corresponding parameters in $\tilde{\theta}$.

For given prior distributions, pseudo prior distribution, and the arbitrary starting values for $\tilde{\gamma}_i$, $\tilde{\psi}_i$, σ_i^2, $i = 1, 2, 3, 4$, $\tilde{\phi}$, $\tilde{p}$, $\tilde{\mu}^1$, and $\tilde{S}_T$, the following steps can be iterated for Gibbs sampling:

Step 1:

i) Generate $\Delta\tilde{c}_T^0$ from $p(\Delta\tilde{c}_T|\tilde{\theta}_c, \Delta\tilde{y}_T, \tilde{\mu} = 0, \tau = 0)$.

ii) Generate $\Delta\tilde{c}_T^1$ from $p(\Delta\tilde{c}_T|\tilde{\theta}_c, \tilde{p}, \tilde{\mu} = \tilde{\mu}^1, \tilde{S}_T, \Delta\tilde{y}_T, \tau = 1)$.

Step 2:

Generate τ from $p(\tau|\Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}, \Delta\tilde{y}_T)$.

Step 3:

If $\tau = 0$:

i) Generate $\tilde{\mu}^1$ from the pseudo prior distribution, $p(\tilde{\mu}^1|\tau = 0)$.

ii) Set $\tilde{\mu} = 0$.

iii) Set $\Delta\tilde{c}_T = \Delta\tilde{c}_T^0$.

If $\tau = 1$:

i) Generate $\tilde{\mu}^1$ from $p(\tilde{\mu}^1|\tilde{\phi}, \tilde{S}_T, \Delta\tilde{c}_T^1, \tau = 1)$. Conditional on $\Delta\tilde{c}_T^1$, $\tilde{\mu}^1$ is independent of data.

ii) Set $\tilde{\mu} = \tilde{\mu}^1$.

i) Set $\Delta\tilde{c}_T = \Delta\tilde{c}_T^1$.

Step 4:

Generate $\tilde{S}_T$ from $p(\tilde{S}_T|\tilde{\phi}, \tilde{\mu}^1, \tilde{p}, \Delta\tilde{c}_T^1)$. Conditional on $\Delta\tilde{c}_T^1$, $\tilde{S}_T$ is independent of data.

Step 5:

Generate $\tilde{p}$ from $p(\tilde{p}|\tilde{S}_T)$. Conditional on $\tilde{S}_T$, $\tilde{p}$ is independent of data and the other parameters of the model.

Step 6:

Generate $\tilde{\phi}$ from $p(\tilde{\phi}|\tilde{\mu}, \tilde{S}_T, \Delta\tilde{c}_T)$, where, if conditional on $\tilde{\mu} = 0$, $\tilde{S}_T$ is irrelevant. Conditional on $\Delta\tilde{c}_T$, $\tilde{\phi}$ is independent of data.

Step 7:

Generate $\tilde{\gamma}_i$, $\tilde{\psi}_i$, and σ_i^2 from $p(\tilde{\gamma}_i, \tilde{\psi}_i, \sigma_i|\Delta\tilde{y}_{iT}, \Delta\tilde{c}_T)$, $i = 1, 2, 3, 4$, where $\Delta\tilde{y}_{iT}$ is a vector of data on individual coincident indicator.

Step 8:

Go to Step 1.

For the full conditional distributions from which the parameters of the model, $\Delta\tilde{c}_T$, and $\tilde{S}_T$ are to be drawn for given τ, readers are referred to Kim and Nelson (1998a). The following derives the full conditional distribution from which the model indicator parameter τ is to be drawn:

$$\begin{aligned}
&p(\tau = 1|\Delta\tilde{y}_T, \Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}) \\
&\qquad = \frac{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1|\tau = 1)\underline{\pi}_1}{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1|\tau = 1)\underline{\pi}_1 + p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1|\tau = 0)\underline{\pi}_0} \\
&\qquad = \frac{\underline{C}_{10}\pi_1}{C_{10}\underline{\pi}_1 + (1 - \underline{\pi}_1)},
\end{aligned} \tag{20}$$

where $\underline{\pi}_1$ is the prior probability of $\tau = 1$ and C_{10} is the conditional Bayes factor:

$$
\begin{aligned}
&C_{10}\\
&= \frac{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 1)}{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1 | \tau = 0)}\\
&= \frac{p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\Delta\tilde{c}_T^1 | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\tilde{\theta}_c | \tau = 1) p(\tilde{p} | \tau = 1) p(\tilde{\mu}^1 | \tau = 1)}{p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^0, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tau = 0) p(\tilde{\theta}_c | \tau = 0) p(\tilde{p} | \tilde{\mu}^1, \tau = 0) p(\tilde{\mu}^1 | \tau = 0)}\\
&= \frac{p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\Delta\tilde{c}_T^1 | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1) p(\tilde{\mu}^1 | \tau = 1)}{p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^0, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tau = 0) p(\tilde{\mu}^1 | \tau = 0)},
\end{aligned}
\tag{21}
$$

where the numerator is obtained by integrating the right-hand-side of (18) with respect to $\tilde{S}_T$ and the denominator is obtained by integrating the right-hand-side of (19') with respect to $\tilde{S}_T$ and $\Delta\tilde{c}_T^*$. It is assumed that $p(\tilde{\theta}_c | \tau = 0) = p(\tilde{\theta}_c | \tau = 0)$ and $p(\tilde{p} | \tilde{\mu}^1, \tau = 0) = p(\tilde{p} | \tau = 1)$ without loss of generality. The terms $p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tau = 0)$ and $p(\Delta\tilde{y}_T | \Delta\tilde{c}_T, \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1)$ can be computed by focusing on equations (7') and (8), by treating $\Delta\tilde{c}_T$ as a vector of data. Similarly, $p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tau = 0)$ and $p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1)$ can be computed based on (1'). For example, $p(\Delta\tilde{c}_T | \tilde{\theta}_c, \tilde{p}, \tilde{\mu}^1, \tau = 1)$ is evaluated as a byproduct of running Hamilton's (1989) basic filter using $\Delta\tilde{c}_T$.

5. Applications: Testing for Markov-Switching in the Business Cycle

5.1. Data Description

Data we employ for a univariate test of Markov-switching is the quarterly real GDP growth rate for a period of 1952.II-1997.II. The coincident variables employed for a multivariate test are the four monthly series for the United States used by the Department of Commerce (DOC) to construct its composite index of coincident indicators: industrial production (IP), total personal income less transfer payments in 1987 dollars (GMYXPQ), total manufacturing and trade sales in 1987 dollars (MTQ), and employees on nonagricultural payrolls (LPNAG). [3] The time period is 1960.1 through 1995.1, which covers Kim and Nelson's (1998a) sample period. We use the demeaned log-differences for all the series.

[3] The abbreviations IP, GMYXPQ, MTQ, and LPNAG are DRI variable names.

5.2. Specification of the Priors and the Pseudo Priors

We employ non-informative (flat) priors for all the parameters of the models expect for the shift parameters $\tilde{\mu}^1 = [\,\mu_0^1 \quad \mu_1^1\,]'$ and the transition probabilities $\tilde{p} = [\,p_{00} \quad p_{11}\,]'$. The priors employed are summarized as follows: [4]

Priors

$$p(\tilde{\gamma}_i, \tilde{\psi}_i, \sigma_i^2 | \tau = 0) = p(\tilde{\gamma}_i, \tilde{\psi}_i, \sigma_i^2 | \tau = 1) \propto \frac{1}{\sigma_i}, \qquad i = 1, 2, 3, 4 \tag{22}$$

$$p(\tilde{\phi}, \sigma^2 | \tau = 0) = p(\tilde{\phi}, \sigma^2 | \tau = 1) \propto \frac{1}{\sigma} \tag{23}$$

$$p_{00} \mid \tau = 1 \quad \sim \quad beta(\alpha_{00}, \alpha_{01}), \tag{24}$$

$$p_{11} \mid \tau = 1 \quad \sim \quad beta(\alpha_{11}, \alpha_{10}), \tag{25}$$

$$\mu_0 \mid \tau = 1 \quad \sim \quad N(0, \text{-}_{0})_{1[\mu_0 < 0]}, \tag{26}$$

$$\mu_1 \mid \tau = 1 \quad \sim \quad N(0, \text{-}_{1})_{1[\mu_1 > 0]}, \tag{27}$$

where equation (22) is relevant only within the multivariate framework, $\sigma = 1$ in the multivariate framework, $beta(.,.)$ refers to a Beta distribution, and $1[.]$ refers to an indicator function.

Notice we that cannot employ non informative priors for μ_0^1 and μ_1^1, the shift parameters being tested. A consequence of employing non informative priors for the parameters being tested will be to force the test results to favor the null hypothesis. [5] But we want their variances large enough to give support to values that are substantially different from 0, but not so large that unrealistic values are supported (George and McCulloch (1993)). For both the univariate and multivariate tests, we consider three different sets prior specifications. The three alternative sets of prior specifications we consider for the univariate tests are: i) $\alpha_{00} = 4$, $\alpha_{01} = 1$, $\alpha_{11} = 4$, $\alpha_{10} = 1$, $\text{-}_{0} = 2$, and $\text{-}_{1} = 2$ [Case 1]; ii) $\alpha_{00} = 8$, $\alpha_{01} = 2$, $\alpha_{11} = 18$, $\alpha_{10} = 2$, $\text{-}_{0} = 2$, and $\text{-}_{1} = 2$ [Case 2]; ii) $\alpha_{00} = 8$, $\alpha_{01} = 2$, $\alpha_{11} = 18$,

[4] For issues concerning the sensitivity analysis and the choice of the priors in a Bayesian model selection, refer to Kass and Raftery (1995).

[5] This is sometimes called Bartlett's (1957) paradox. For more detailed discussion, refer to Kass and Raftery (1995).

$\alpha_{10} = 2$, - $_0 = 9$, and - $_1 = 9$ [Case 3]. The three alternative sets of prior specifications for the univariate tests are: i) $\alpha_{00} = 4$, $\alpha_{01} = 1$, $\alpha_{11} = 4$, $\alpha_{10} = 1$, - $_0 = 2$, and - $_1 = 2$ [Case 1]; ii) $\alpha_{00} = 9$, $\alpha_{01} = 1$, $\alpha_{11} = 29$, $\alpha_{10} = 1$, - $_0 = 2$, and - $_1 = 2$ [Case 2]; ii) $\alpha_{00} = 9$, $\alpha_{01} = 1$, $\alpha_{11} = 29$, $\alpha_{10} = 1$, - $_0 = 9$, and - $_1 = 9$ [Case 3]. In Case 2, we have more informative priors for the transition probabilities than in Case 1. In Case 3, we have looser priors for the shift parameters than in Case 2.

The choice of the pseudo priors for μ_0^1 and μ_1^1 conditional on $\tau = 0$ is important for the convergence of the Gibbs sampler. Values for these parameters, if generated from reasonable pseudo prior distributions, would be consistent with the data. Following the recommendation of Carlin and Chib (1995), we first get preliminary estimates of marginal posterior distributions of these parameters for models with Markov-switching ($Pr(\tau = 1) = 1$). Tables 1 through 3 summarize the results for univariate Markov-switching model with different sets of priors. Tables 4 through 6 summarize the results for the dynamic factor models with Markov-switching. We use first-order (Normal) approximations to the marginal posterior distributions for μ_0^1 and μ_1^1 in Tables 1 through 6 as our pseudo prior distributions for each case: [6]

<u>**Pseudo Priors**</u>

$$\mu_0^1 \mid \tau = 0 \;\sim\; N(\mu_0^*, V_{\mu,0})_{1[\mu_0 < 0]}, \tag{28}$$

$$\mu_1^1 \mid \tau = 0 \;\sim\; N(\mu_1^*, V_{\mu,1})_{1[\mu_1 > 0]}, \tag{29}$$

where, for example, we employ $\mu_0^* = -0.617$, $V_{\mu,0} = 0.525^2$, $\mu_1^* = 0.232$, and $V_{\mu,1} = 0.200^2$ (from Table 1) for Case 1 in the univariate framework.

5.3. Empirical Results [7]

We first examine the sensitivities of the inferences on the regime probabilities to three alternative sets of priors employed for the Markov-switching models with $Pr(\tau = 1) = 1$.

[6] As Carlin and Chib (1995) note, we are not using the data to select the prior, but only the pseudo prior.

[7] All the inferences in this section are based on 9,000 Gibbs simulations, after discarding the first 1,000 out of 10,000 Gibbs simulations.

The univariate results in Figure 1 suggest that with tighter priors for the transition probabilities in Case 2 than in Case 1, inferences on the regime probabilities get much sharper. [8] A comparison of the results for Cases 2 and 3 in Figure 1 also suggests that, given the same tight priors for the transition probabilities, different priors for the shift parameters affect the inferences on regime probabilities very little. The multivariate results in Figure 2 lead us to similar conclusion. In the multivariate framework, however, inferences on regime probabilities are significantly less sensitive to the priors for the transition probabilities.

The posterior probability of Markov-switching ($Pr(\tau = 1|\Delta\tilde{z}_T)$, where $\Delta\tilde{z}_T$ is data) is obtained by the proportion of the posterior simulations in which $\tau = 1$. Tables 7 and 8 summarize the sensitivities of the posterior probabilities of Markov-switching to different prior probabilities for the univariate tests and the multivariate tests, respectively. Figures 3 and 4 visually summarize the same results. For the univariate tests, the posterior probability of Markov-switching is quite sensitive to the prior probabilities. As we change the prior probability from 0.1 to 0.9, the posterior probability ranges between 0.072 and 0.724 in Case 1. However, the implied Bayes factor, which summarizes the effect of the data in modifying the prior odds to obtain posterior odds, is consistently lower than 1, ranging between 0.163 and 0.724. With a prior probability of 0.5, for example, the posterior probability is 0.300 and the implied Bayes factor is 0.429. These results may be interpreted as sample evidence being against Markov-switching, even though the posterior probability is quite sensitive to the prior probability. Different cases (Case 2 and Case 3) considered with different priors for $\tilde{\mu}$ and $\tilde{p}$ do not seem to affect the results significantly.

For the multivariate tests, we get somewhat qualitatively different results. The posterior probability of Markov-switching are not very sensitive to the prior probability as depicted in Figure 4. In case 1, as we change the prior probability from 0.1 to 0.9, the posterior probability ranges from 0.522 to 0.591. The posterior probabilities are consistently higher, but only slightly higher, than 0.5. For the prior probability of 0.5, the posterior probability is 0.560 and the implied Bayes factor is 1.273. These considerations might suggest that the data slightly favors Diebold and Rudebusch's (1996) dynamic factor model

[8] The shaded areas represent the periods of National Bureau of Economic Research (NBER) recessions.

with Markov-switching over Stock and Watson's (1991) linear dynamic factor model. The results are not sensitive to alternative priors employed for $\tilde{\mu}^1$ and $\tilde{p}$ (Case 2 and Case 3). However, unlike the univariate test results, the Bayes factor is quite sensitive to the prior probability and ranges from 9.828 to 0.161. This suggests that the multivariate test results leave more room for subjective interpretations than the univariate test results.

While the univariate and multivariate test results leave plenty of room for subjective interpretations when examined separately, a comparison of the two results allows us to draw a conclusion which is objective enough: Evidence of Markov-switching, if exists, is much more compelling in the multivariate tests.

6. Summary and Discussion: Is the Business Cycle Asymmetric?

In this paper, we present Bayesian tests of Markov-switching within both univariate and multivariate frameworks. In the univariate framework, we design a procedure for testing for Markov-switching in an observed time series. With no Markov-switching, Hamilton's (1989) model collapses to a linear autoregressive model. In the multivariate framework, we deal with testing for Markov-switching in an unobserved factor component which is common to multiple observed time series. With no Markov switching, Diebold and Rudebusch's (1996) model collapses to Stock and Watson's (1991) linear dynamic factor model. The tests are based on the sensitivity of the posterior probability to the prior probability of the model indicator parameter which is employed to represent both a linear model and a Markov-switching model within a unified framework.

We apply the proposed testing procedure to the quarterly real GDP series and four monthly coincident economic indicators in order to investigate Markov-switching in the business cycle. For the univariate tests which are based on quarterly real GDP growth, the data in general seem to be against Markov-switching. However, we do not interpret the univariate test results as rejecting the business cycle asymmetry. For example, in a test of structural break in the shift parameters of a Markov-switching model for the real GDP growth, Kim and Nelson (1998c) find strong sample evidence in favor of a narrowing gap between the growth rates during booms and recessions. They report significantly sharper regime probabilities than those in Figure 1, after a structural break is taken into account with an unknown changepoint. While we investigate Markov-switching in the growth rate of the GDP series in this paper, Kim and Nelson (1998d) raise a possibility of Markov-switching in the cyclical component of the real GDP series, as implied by Friedman's (1964, 1993) 'plucking' model. The threshold autoregressive model of Tong (1983) and Potter (1995) is another type of asymmetry not considered here. It is possible that a linear model may be less favored against these alternatives.

Besides, of the two defining characteristics of the business cycle by Burns and Mitchell (1946), namely 'comovement' and 'asymmetry', the univariate tests of Markov-switching (or asymmetry) fail to take into account the 'comovement' feature of the business cycle. The multivariate tests, which explicitly take into account comovement among economic variables through the business cycle, seem to provide sample evidence that slightly favors a Markov-switching model over a linear model. Even though the multivariate test results are open to more subjective interpretations than the univariate test results, a comparison of the two results allows us to draw a conclusion which is objective enough: Evidence of Markov-switching or asymmetry in the business cycle, if exists, is much more compelling in the multivariate tests.

References

[1] Albert, James H. and Siddhartha Chib, 1993, "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business and Economic Statistics,* Vol. 11, No. 1, 1-15.

[2] Bartlett, M. S., 1957, "Comment on 'a Statistics Paradox," by D. V. Lindley, *Biometrika,* 44, 533-544.

[3] Burns, A.F. and W.C. Mitchell, 1946, "Measuring Business Cycles," New York: National Bureau of Economic Research.

[4] Carlin, B. P., and Chib, S, 1995, "Bayesian Model Choice via Markov Chain Monte Carlo Methods," *Journal of the Royal Statistical Society,* Series B, 57, No. 3, 473-484.

[5] Carlin, B. P., and Polson, N. G., 1991, "Inference for Nonconjugate Bayesian Models Using the Gibbs Sampler," *Canadian Journal of Statistics,* 19, 399-405.

[6] Carter, C. K. and P. Kohn, 1994, "On Gibbs Sampling for State Space Models," *Biometrica,* 81, p. 541-553.

[7] Chib, Siddhartha, 1995, "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association,* Vol. 90, No. 432, Theory and Methods, 1313-1321.

[8] Dickey, J. M, 1971, "The Weighted Likelihood Ratio, Linear Hypothesis on Normal Location Parameters," *Annals of Mathematical Statistics,* 42, 204-223.

[9] Diebold, Francis X. and Glenn D. Rudebusch, 1996, "Measuring Business Cycles: A Modern Perspective," *Review of Economics and Statistics*, 78, 67-77.

[10] Friedman, Milton, 1964, "Monetary Studies of the National Bureau," *The National bureau Enters Its 45th Year,* 44th Annual Report, 7-25; Reprinted in Milton Friedman, 1969, *The Optimum Quantity of Money and Other Essays,* Chicago: Aldine, Chap. 12, 261-284.

[11] Friedman, Milton, 1993, "The 'Plucking Model' of Business Fluctuations Revisited," *Economic Inquiry,* April, 171-177.

[12] Geweke, John, 1996, "Variable Selection and Model Comparison in Regression," *Bayesian Statistics 5,* ed. J.O. Berger, J. M. Bernardo, A. P. Dawid, and A. F. M. Smith), Oxford: Oxford University Press.

[13] George, E. I. and McCulloch, R. E., 1993, "Variable Selection via Gibbs Sampling," *Journal of American Statistical Association,* 88, 881-889.

[14] Garcia, Rene, 1998, "Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models," *International Economic Review*, Vol. 39, No. 3, 763-788.

[1] Hamilton, James D., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica,* Vol. 57, No. 2, 357-384.

[2] Hamilton, James D., 1990, "Analysis of Time Series Subject to Changes in Regime, *Journal of Econometrics,* 45, 39-70.

[3] Hansen, B. E., 1992, "The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP," *Journal of Applied Econometrics,* Vol. 7, S61-S82.

[4] Jeffreys, H., 1961, Theory of Probability, 3rd ed., Oxford: Clarendon Press.

[5] Kass, Robert E. and Raftery, Adrian E., 1995, "Bayes Factors," *Journal of the American Statistical Association,* Vol. 90, No. 430, 773-795.

[6] Kim, Chang-Jin and Charles R. Nelson, 1998a, "Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching," *Review of Economics and Statistics,* 80, 188-201.

[7] Kim, Chang-Jin and Charles R. Nelson, 1998b, *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* forthcoming, MIT Press.

[8] Kim, Chang-Jin and Charles R. Nelson, 1998c, "A Test for Structural Change in Markov-Switching Models: Has the U.S. Economy Become More Stable?" Working Paper, Dept. of Economics, University of Washington.

[9] Kim, Chang-Jin and Charles R. Nelson, 1998d, "Friedman's Plucking Model of Business Fluctuation: Tests and Estimates of Permanent and Transitory Components," forthcoming, *Journal of Money, Credit, and Banking.*

[10] Koop, Gary and Simon M. Potter, 1996, "Bayes Factors and Nonlinearity: Evidence from Economic Time Series," Department of Economics, UCLA.

[11] McCulloch, R. and R. Tsay, 1994, "Statistical Analysis of Economic Time Series via Markov Switching Models," *Journal of Time Series Analysis," 15, 523-539.*

[12] Newton, Michael A. and Raftery, Adrian E., 1994, "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society,* Series B, 3-48.

[13] Perron, Pierre, 1990, "Testing for a Unit Root in a Time Series with a Changing Mean," *Journal of Business & Economic Statistics,* Vol. 8, No. 2, 153-162.

[14] Poirier, Dale J., 1995, *Intermediate Statistics and Econometrics,* Cambridge: The MIT Press.

[15] Potter, Simon M., 1995, "A Nonlinear Approach to U.S. GNP," *Journal of Applied Econometrics,* 10, 109-125.

[16] Stock, James H. and Mark W. Watson, 1991, "A Probability Model of the Coincident Economic Indicators," in K. Lahiri and G.H. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records,* Cambridge: Cambridge University Press.

[17] Tierney, L., 1994, "Markov Chains for Exploring Posterior Distributions," *The Annals of Statistics,* 22, 1701-1762.

[18] Tong, Howell, *Threshold Models in Non-Linear Time-Series Models,* New York: Springer-Verlag, 1983.

[19] Zellner, A., 1971, *An Introduction to Bayesian Inference in Econometrics,* Wiley, New York, NY.

[20] Zellner, A., 1984, *Basic Issues in Econometrics,* University of Chicago Press, Chicago, IL.

[21] Zivot, Eric, 1994, "A Bayesian Analysis of the Unit Root Hypothesis within an Unobserved Components Model," *Econometric Theory,* 13(3).

Table 1. Posterior Distribution of the Parameters of the Markov-Switching Model ($Pr(\tau = 1) = 1$): Case 1

$$\mu_0^1 \sim N(0,2)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,2)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(4,1); \quad p_{11} \sim beta(4,1),$$

	Mean	**SD**	**Median**	**95% Bands**
p_{00}	0.725	0.156	0.741	(0.437, 0.955)
p_{11}	0.864	0.126	0.910	(0.589, 0.983)
μ_0^1	**-0.671**	**0.525**	-0.540	(-1.663, -0.050)
μ_1^1	**0.232**	**0.200**	0.191	(0.021, 0.569)
ϕ_1	0.253	0.095	0.253	(0.096, 0.405)
ϕ_2	0.068	0.085	0.070	(-0.079, 0.212)
ϕ_3	-0.014	0.085	-0.014	(-0.153, 0.131)
ϕ_4	-0.068	0.083	-0.069	(-0.208, 0.072)
σ^2	0.818	0.121	0.815	(0.625, 1.017)

1. Non-informative (flat) priors are used for all parameters except μ_0^1, μ_1^1, p_{00}, and p_{11}.

2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 90,000 Gibbs simulations.

3. SD and MD refer to standard deviation and median, respectively.

4. 95% Bands refers to 95% posterior probability bands.

Table 2. Posterior Distribution of the Parameters of the Markov-Switching Model ($Pr(\tau = 1) = 1$): Case 1

$$\mu_0^1 \sim N(0,2)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,2)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(8,2); \quad p_{11} \sim beta(18,2),$$

	Mean	**SD**	**Median**	**95% Bands**
p_{00}	0.742	0.113	0.748	(0.543, 0.912)
p_{11}	0.916	0.049	0.927	(0.823, 0.978)
μ_0^1	**-0.813**	**0.480**	-0.779	(-1.618, -0.105)
μ_1^1	**0.220**	**0.138**	**0.206**	(0.026, 0.456)
ϕ_1	0.233	0.095	0.234	(0.073, 0.386)
ϕ_2	0.059	0.089	0.058	(-0.087, 0.206)
ϕ_3	-0.016	0.086	-0.016	(-0.157, 0.127)
ϕ_4	-0.058	0.086	-0.058	(-0.200, 0.083)

1. Non-informative (flat) priors are used for all parameters except μ_0^1, μ_1^1, p_{00}, and p_{11}.

2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 90,000 Gibbs simulations.

3. SD and MD refer to standard deviation and median, respectively.

4. 95% Bands refers to 95% posterior probability bands.

Table 3. Posterior Distribution of the Parameters of the Markov-Switching Model ($Pr(\tau = 1) = 1$): Case 1

$$\mu_0^1 \sim N(0,9)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,9)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(8,2); \quad p_{11} \sim beta(18,2),$$

	Mean	**SD**	**Median**	**95% Bands**
p_{00}	0.740	0.115	0.750	(0.533, 0.916)
p_{11}	0.920	0.049	0.931	(0.828, 0.978)
μ_0^1	**-0.896**	**0.556**	-0.856	(-1.802, -0.112)
μ_1^1	**0.213**	**0.138**	0.196	(0.028, 0.461)
ϕ_1	0.235	0.093	0.237	(0.073, 0.387)
ϕ_2	0.058	0.090	0.056	(-0.086, 0.205)
ϕ_3	-0.012	0.088	-0.012	(-0.154, 0.133)
ϕ_4	-0.062	0.085	-0.061	(-0.201, 0.083)

1. Non-informative (flat) priors are used for all parameters except μ_0^1, μ_1^1, p_{00}, and p_{11}.

2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 90,000 Gibbs simulations.

3. SD and MD refer to standard deviation and median, respectively.

4. 95% Bands refers to 95% posterior probability bands.

Table 4. Bayesian Inference of Diebold and Rudebusch's (1996) Dynamic Factor Model with Markov-switching ($Pr(\tau = 1) = 1$): Case 1

$$\mu_0^1 \sim N(0,2)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,2)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(4,1); \quad p_{11} \sim beta(4,1),$$

		Mean	SD	MD	95%Bands
Δc_t	p_{00}	0.861	0.061	0.870	(0.750, 0.945)
	p_{11}	0.957	0.072	0.972	(0.866, 0.987)
	μ_0^1	**-1.692**	**0.553**	-1.784	(-2.354, -0.185)
	μ_1^1	**0.338**	**0.128**	0.345	(0.127, 0.518)
	ϕ_1	0.337	0.093	0.330	(0.199, 0.502)
	ϕ_2	0.011	0.066	0.011	(-0.100, 0.122)
IP	γ_1	0.562	0.041	0.560	(0.496, 0.632)
	ψ_{11}	-0.020	0.068	-0.020	(-0.130, 0.093)
	ψ_{12}	-0.022	0.066	-0.021	(-0.130, 0.084)
	σ_1^2	0.224	0.033	0.223	(0.171, 0.280)
GMYZPQ	γ_2	0.211	0.023	0.210	(0.174, 0.249)
	ψ_{21}	-0.307	0.051	-0.307	(-0.390, -0.221)
	ψ_{22}	-0.069	0.051	-0.069	(-0.152, 0.016)
	σ_2^2	0.320	0.024	0.319	(0.282, 0.359)
MTQ	γ_3	0.435	0.037	0.434	(0.377, 0.499)
	ψ_{31}	-0.360	0.054	-0.359	(-0.449, -0.268)
	ψ_{32}	-0.160	0.053	-0.160	(-0.246, -0.074)
	σ_3^2	0.662	0.052	0.658	(0.583, 0.753)
LPNAG	γ_4	0.116	0.010	0.116	(0.100, 0.132)
	ψ_{41}	-0.021	0.059	-0.021	(-0.119, 0.075)
	ψ_{42}	0.277	0.061	0.280	(0.174, 0.373)
	σ_4^2	0.022	0.002	0.022	(0.018, 0.025)
	γ_{41}	0.007	0.009	0.007	(-0.008, 0.023)
	γ_{42}	0.021	0.009	0.021	(0.007, 0.036)
	γ_{43}	0.030	0.008	0.030	(0.017, 0.043)

1. Non-informative (flat) priors are used for all parameters except μ_0^1, μ_1^1, p_{00}, and p_{11}.
2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 90,000 Gibbs simulations.
3. SD and MD refer to standard deviation and median, respectively.
4. 95% Bands refers to 95% posterior probability bands.

Table 5. Bayesian Inference of Diebold and Rudebusch's (1996) Dynamic Factor Model with Markov-switching ($Pr(\tau = 1) = 1$): Case 2

$$\mu_0^1 \sim N(0,2)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,2)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(9,1); \quad p_{11} \sim beta(29,1),$$

		Mean	SD	MD	95%Bands
Δc_t	p_{00}	0.866	0.052	0.872	(0.770, 0.937)
	p_{11}	0.975	0.010	0.976	(0.957, 0.989)
	μ_0^1	**-1.799**	**0.342**	-1.803	(-2.339, -1.225)
	μ_1^1	**0.341**	**0.107**	0.345	(0.157, 0.509)
	ϕ_1	0.319	0.079	0.319	(0.192, 0.452)
	ϕ_2	0.008	0.067	0.007	(-0.104, 0.118)
IP	γ_1	0.561	0.038	0.559	(0.501, 0.625)
	ψ_{11}	-0.015	0.069	-0.015	(-0.127, 0.094)
	ψ_{12}	-0.020	0.066	-0.019	(-0.129, 0.087)
	σ_1^2	0.223	0.033	0.223	(0.170, 0.278)
GMYZPQ	γ_2	0.211	0.022	0.210	(0.176, 0.247)
	ψ_{21}	-0.303	0.051	-0.304	(-0.389, -0.220)
	ψ_{22}	-0.067	0.051	-0.066	(-0.152, 0.017)
	σ_2^2	0.319	0.023	0.317	(0.283, 0.361)
MTQ	γ_3	0.435	0.035	0.433	(0.380, 0.496)
	ψ_{31}	-0.358	0.056	-0.357	(-0.447, -0.270)
	ψ_{32}	-0.161	0.054	-0.160	(-0.248, -0.073)
	σ_3^2	0.661	0.051	0.660	(0.581, 0.751)
LPNAG	γ_{41}	0.115	0.010	0.115	(0.099, 0.131)
	ψ_{41}	-0.022	0.058	-0.021	(-0.120, 0.072)
	ψ_{42}	0.277	0.060	0.278	(0.176, 0.374)
	σ_4^2	0.022	0.002	0.022	(0.018, 0.025)
	γ_{41}	0.008	0.009	0.008	(-0.007, 0.023)
	γ_{42}	0.021	0.009	0.021	(0.007, 0.036)
	γ_{43}	0.030	0.008	0.030	(0.017, 0.043)

1. Non-informative (flat) priors are used for all parameters except μ_0^1 and μ_1^1.
2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 9,000 Gibbs simulations.
3. SD and MD refer to standard deviation and median, respectively.
4. 95% Bands refers to 95% posterior probability bands.

Table 6. Bayesian Inference of Diebold and Rudebusch's (1996) Dynamic Factor Model with Markov-switching ($Pr(\tau = 1) = 1$): Case 3

$$\mu_0^1 \sim N(0,9)_{I[\mu_0^1<0]}, \quad \mu_1^1 \sim N(0,9)_{I[\mu_1^1>0]}$$
$$p_{00} \sim beta(9,1); \quad p_{11} \sim beta(29,1),$$

		Mean	SD	MD	95%Bands
	p_{00}	0.862	0.054	0.868	(0.763, 0.934)
	p_{11}	0.975	0.010	0.976	(0.957, 0.989)
Δc_t	μ_0^1	**-1.893**	**0.372**	-1.884	(-2.476, -1.310)
	μ_1^1	**0.337**	**0.111**	0.338	(0.154, 0.512)
	ϕ_1	0.323	0.079	0.322	(0.199, 0.457)
	ϕ_2	0.009	0.069	0.010	(-0.104, 0.124)
	γ_1	0.554	0.037	0.554	(0.493, 0.616)
IP	ψ_{11}	-0.015	0.068	-0.015	(-0.128, 0.096)
	ψ_{12}	-0.018	0.066	-0.018	(-0.127, 0.088)
	σ_1^2	0.226	0.033	0.225	(0.173, 0.281)
	γ_2	0.208	0.021	0.208	(0.174, 0.244)
GMYZPQ	ψ_{21}	-0.306	0.052	-0.305	(-0.390, -0.221)
	ψ_{22}	-0.064	0.051	-0.064	(-0.150, 0.017)
	σ_2^2	0.320	0.024	0.318	(0.283, 0.360)
	γ_3	0.429	0.035	0.429	(0.374, 0.488)
MTQ	ψ_{31}	-0.358	0.055	-0.357	(-0.446, -0.268)
	ψ_{32}	-0.161	0.054	-0.161	(-0.245, -0.073)
	σ_3^2	0.663	0.053	0.660	(0.580, 0.752)
	γ_{41}	0.114	0.009	0.114	(0.099, 0.130)
	ψ_{41}	-0.025	0.060	-0.024	(-0.123, 0.075)
	ψ_{42}	0.273	0.060	0.274	(0.172, 0.370)
LPNAG	σ_4^2	0.022	0.002	0.022	(0.018, 0.025)
	γ_{41}	0.008	0.009	0.008	(-0.008, 0.023)
	γ_{42}	0.021	0.009	0.021	(0.007, 0.035)
	γ_{43}	0.030	0.008	0.030	(0.017, 0.043)

1. Non-informative (flat) priors are used for all parameters except μ_0^1 and μ_1^1.
2. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 9,000 Gibbs simulations.
3. SD and MD refer to standard deviation and median, respectively.
4. 95% Bands refers to 95% posterior probability bands.

Table 7. Bayesian Model Selection Based on Sensitivity of the Posterior Probability of Markov-Switching Model to Prior Probability: Univariate Tests

	Prior Probabilities ($\underline{\pi}_1 = Pr(\tau = 1)$)				
	0.1	0.3	0.5	0.7	0.9
	Posterior Probabilities ($\bar{\pi}_1 = Pr(\tau = 1 \mid \Delta\tilde{c}_T)$)				
Case 1	0.072	0.183	0.300	0.487	0.724
(Bayes Factor)	(0.701)	(0.524)	(0.429)	(0.407)	(0.292)
Case 2	0.074	0.194	0.325	0.511	0.756
(Bayes Factor)	(0.724)	(0.561)	(0.482)	(0.447)	(0.343)
Case 3	0.037	0.114	0.207	0.319	0.594
(Bayes Factor)	(0.348)	(0.302)	(0.261)	(0.201)	(0.163)

Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 9,000 Gibbs simulations.

Table 8. Bayesian Model Selection Based on Sensitivity of the Posterior Probability of Markov-Switching to Prior Probability: Multivariate Tests.

	Prior Probabilities ($\underline{\pi}_1 = Pr(\tau = 1)$)				
	0.100	0.300	0.500	0.700	0.900
	Posterior Probabilities ($\bar{\pi}_1 = Pr(\tau = 1 \mid \Delta\tilde{y}_T)$)				
Case 1	0.522	0.544	0.560	0.574	0.591
(Bayes Factor)	(9.828)	(2.784)	(1.273)	(0.577)	(0.161)
Case 2	0.544	0.572	0.580	0.602	0.632
(Bayes Factor)	(10.737)	(3.118)	(1.380)	(0.648)	(0.191)
Case 3	0.502	0.534	0.564	0.577	0.618
(Bayes Factor)	(9.072)	(2.674)	(1.294)	(0.585)	(0.180)

1. Out of 10,000 Gibbs simulations, the first 1,000 are discarded and inferences are based on the remaining 9,000 Gibbs simulations.

Figure 1. Probability of a Recession: Univariate Model with Markov-Switching

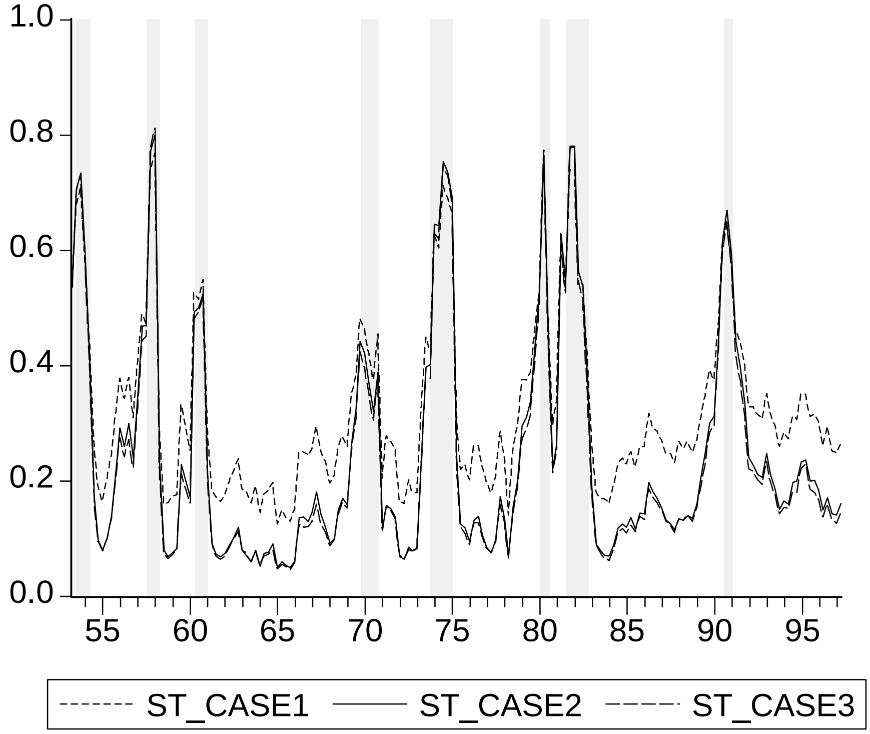

Figure 2. Probabilities of a Recession: Dynamic Factor Model with Markov-Switching

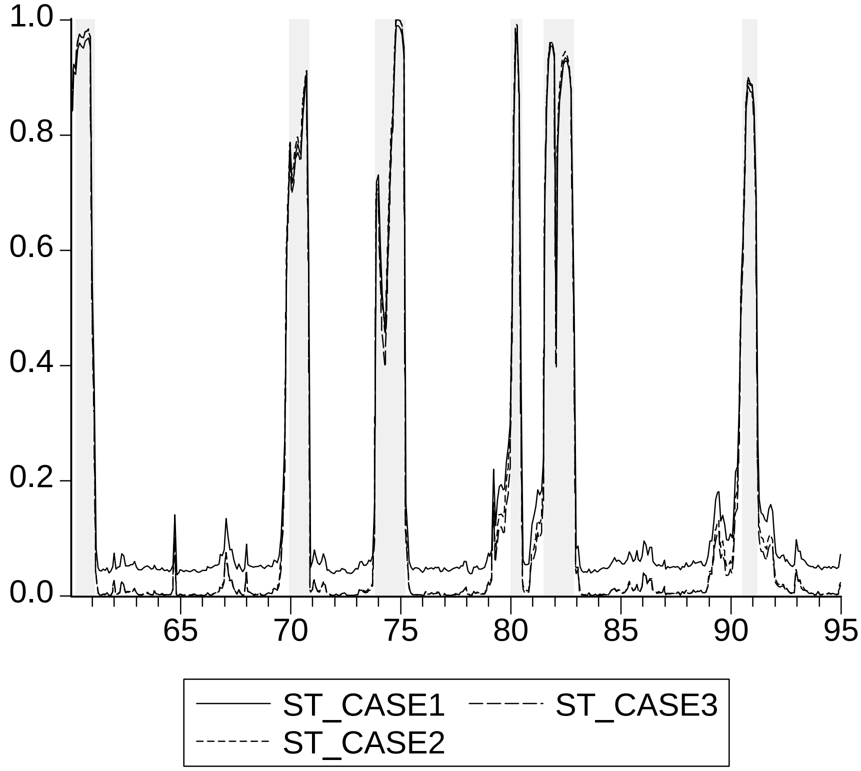

Figure 3. Sensitivity of Posterior Probability of Markov-Switching to Prior Probability

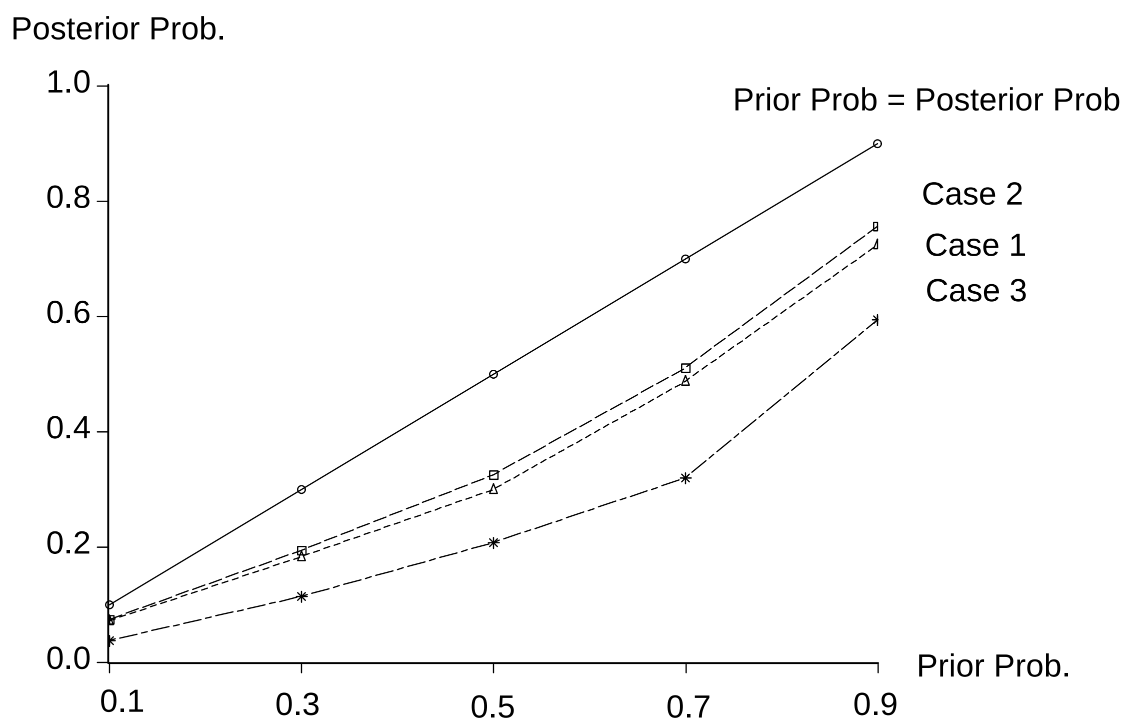

Figure 4. Sensitivity of Posterior Probabilities Markov-Switching to Prior Probabilities: Multivariate Tests