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

Chang-Jin Kim and **Charles R. Nelson**†

Revised, September 1999

† Department of Economics, Korea University and Department of Economics, University of Washington, respectively. We appreciate helpful comments from the editor, the anonymous referees, and the participants in the 1999 NSF/NBER forecasting seminar for helpful comments, but the responsibility for errors is entirely the authors'. We also acknowledge support from the National Science Foundation under grant SES-9818789.

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, Markov Switching, 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) 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 (see Diebold and Rudebusch (1996), Kim and Nelson (1998, 1999a), and for a review of this literature Diebold and Rudebusch (1998)).

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 model is well developed in both the classical and the Bayesian perspectives and

[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.

applications are abundant, there seems to be a lag in the development of procedures for testing for Markov-switching. In most applied work, Markov-switching has been assumed to exist without testing. 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). 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 have produced conflicting evidence of Markov-switching, we speculate that tests in a multivariate framework should provide more reliable and consistent results. Indeed, if the dynamic factor model we use is successful in capturing comovement across indicators, that should sharpen inference compared to univariate analysis where the information in the data may be obscured by idiosyncratic variation.

In this paper, we present Bayesian tests of Markov-switching in both univariate and multivariate contexts. Within the Bayesian framework, the main issue in hypothesis testing or model selection comes down to calculating the marginal likelihood for each model under consideration and the resulting Bayes factor, which is given by the ratio of 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), we indirectly calculate the Bayes factors using the prior and posterior probabilities of a model indicator parameter, without calculating the marginal likelihoods. [2] In implementing the Markov Chain Monte Carlo (MCMC) method of Gibbs sampling to achieve the goal, 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

[2] For a Bayesian model selection based on the predictive marginal likelihoods and the predictive Bayes factor, refer to Filardo and Gordon (1999).

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 follow Carlin and Chib's (1995) procedure and employ a pseudo prior for the shift parameters that are otherwise set to zero conditional on no Markov-switching.

We first present the 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 dynamic factor models of the business cycle of Stock and Watson (1991) and Diebold and Rudebusch (1996). A difficulty arises since we want to test for Markov-switching in the common factor that is unobserved. This is overcome by extending Carlin and Chib's (1995) procedure and by incorporating Kim and Nelson's (1998) Bayesian approach to dynamic factor models.

In Section 2, we present the model specifications employed. Conditional on no Markov-switching, Hamilton's (1989) univariate model collapses to the linear autoregressive model and Diebold and Rudebusch's (1996) multivariate model collapses to Stock and Watson's (1991) linear dynamic factor model. In Section 3, basic issues associated with our tests are then discussed. Section 4 introduces Carlin and Chib's (1995) approach to Bayesian model selection in the presence of varying parameter space. Then, the approach is extended to deal with tests of Markov-switching in both univariate and multivariate frameworks. Section 5 presents empirical results. For the univariate test we use quarterly real GDP growth (1952.II-1997.II), and for the multivariate test we use 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

2.1. 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:

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

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

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 hypothesis testing, however, they may have different implications for the testing procedure within the Bayesian framework. [3] 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 form (equation (1')).

2.2. 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

[3] 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.

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 ΔY_{it}, we first express the data as deviations from means. Also for identification purpose, we set $\sigma^2 = 1$. Then the full model on which our test will be based is given by:

Model

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

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

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

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

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

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

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

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).

3. Problem Setup

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

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

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 ($\bar{\pi}_1/(1 - \bar{\pi}_1)$):

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

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 (1998) 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' of 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 the 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 challenging in the presence of the two blocks of latent variables ($\tilde{S}_T = [\,S_1 \quad \ldots \quad S_T\,]'$ and $\Delta\tilde{c}_T = [\,\Delta c_1 \ldots \quad \Delta c_T\,]'$), 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. First, 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. Second, as discussed above, the vector of autoregressive coefficients for Δc_t in (1) and (1'), denoted by $\tilde{\phi}$, may not be the same for linear and the Markov-switching models. In this case, we have $\tilde{\phi} = \tilde{\phi}^0$ for a linear model ($\tau = 0$) and $\tilde{\phi} = \tilde{\phi}^1$ for a Markov-switching model ($\tau = 1$). Thus, conditional on $\tau = 0$, $\tilde{\phi}^1$ is forced out of model and conditional on $\tau = 1$, $\tilde{\phi}^0$ is forced out of model. For the parameters that are common to both models, we do not have such problems. Third, an additional problem arises for dynamic factor models, as the unobserved factor component $\Delta\tilde{c}_T$ is not common to linear and Markov-switching models. Denoting $\Delta\tilde{c}_T^0$ and $\Delta\tilde{c}_T^1$ to be the vectors of the factor components for linear and Markov-switching models, respectively, the usual Gibbs sampler skips the generation $\Delta\tilde{c}_T^1$ conditional on $\tau = 0$ and $\Delta\tilde{c}_T^0$ conditional on $\tau = 1$. Any these 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)).

Such convergence problem in the Gibbs sampler is solved by employing 'pseudo priors' for the parameters that are forced out of a particular model, as suggested by Carlin and Chib (1995). Thus, in the following section we first review Carlin and Chib's (1995) general approach to model selection in the presence of a varying parameter space. We then extend their approach to deal with the problems raised above in designing tests of Markov switching.

4. Testing for Markov Switching in Univariate and Dynamic Factor Models

4.1. A General Framework for Model Selection in the Presence of Varying Parameter Space: Carlin and Chib (1995)

Suppose that we are interested in selecting between two alternative models without latent variables: Model 0 and Model 1. Let $\tilde{\theta}_c$, $\tilde{\theta}_0$, $\tilde{\theta}_1$ be a vector of parameters common to both models, a vector of parameters unique to model 0, and a vector of parameters unique to model 1. Denote τ to be the model indicator parameters. Corresponding to model j we have a likelihood $p(\tilde{z}_T|\tilde{\theta}_c, \tilde{\theta}_j, \tau = j)$ and priors $p(\tilde{\theta}_c|\tau = j)$ and $p(\tilde{\theta}_j|\tau = j)$,

$j = 0, 1$, where $\tilde{\theta}_c$, $\tilde{\theta}_0$, and $\tilde{\theta}_1$ are a priori assumed independent. Thus, conditional on $\tau = j$, the vector $\tilde{\theta}_{i \neq j}$ is irrelevant and the usual Gibbs sampling skips the generation of $\tilde{\theta}_{i \neq j}$, potentially resulting in a convergence problem in the Gibbs sampler.

To solve the problem, Carlin and Chib (1995) consider the joint posterior distribution of $\tilde{\theta}_c$, $\tilde{\theta}_j$ and $\tilde{\theta}_{i \neq j}$ conditional on $\tau = j$, given by:

$$\begin{aligned} p(\tilde{\theta}_c, &\tilde{\theta}_j, \tilde{\theta}_{i \neq j} | \tilde{z}_T, \tau = j) \\ &\propto p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_j, \tilde{\theta}_{i \neq j} | \tau = j) \\ &= p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_j, \tilde{\theta}_{i \neq j}, \tau = j) p(\tilde{\theta}_c | \tau = j) p(\tilde{\theta}_j | \tau = j) p(\tilde{\theta}_{i \neq j} | \tau = j) \\ &= p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_j, \tau = j) p(\tilde{\theta}_c | \tau = j) p(\tilde{\theta}_j | \tau = j) p(\tilde{\theta}_{i \neq j} | \tau = j), \quad j = 0, 1. \end{aligned} \tag{12}$$

Then, they suggest employing $p(\tilde{\theta}_{i \neq j} | \tau = j)$, the pseudo prior or the 'linking density' for $\tilde{\theta}_{i \neq j}$, in order to generate $\tilde{\theta}_{i \neq j}$ without skipping its generation, conditional on $\tau = j$. By employing the pseudo priors we can avoid the convergence problem in the Gibbs sampler, but their employment does not affect the marginal likelihood for each model, since we have

$$\int \int \int p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_j, \tilde{\theta}_{i \neq j} | \tau = j) d\tilde{\theta}_c \tilde{\theta}_j \tilde{\theta}_{i \neq j} = \int \int p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_j | \tau = j) d\tilde{\theta}_c d\tilde{\theta}_j, \quad j = 0, 1. \tag{13}$$

In doing so, the full conditional distribution for τ, from which τ is to be generated, should account for the employment of the pseudo priors. The following provides the full conditional distribution of τ with pseudo priors:

$$\begin{aligned} p(\tau = 1 | \tilde{\theta}_c, \tilde{\theta}_0, &\tilde{\theta}_1, \tilde{z}_T) \\ &= \frac{p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1)}{\sum_{j=0}^{1} p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = j)} \\ &= \frac{p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 1) \underline{\pi}_1}{\sum_{j=0}^{1} p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = j) \underline{\pi}_j} \\ &= \frac{\underline{\pi}_1 C_{10}}{\underline{\pi}_0 + \underline{\pi}_1 C_{10}}, \end{aligned} \tag{14}$$

where $\underline{\pi}_0$ and $\underline{\pi}_1 = 1 - \underline{\pi}_0$ are prior probabilities of Model 0 and Model 1, respectively, and C_{10} is the conditional Bayes factor in favor of Model 1 given by:

$$
\begin{aligned}
C_{10} &= \frac{p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 1)}{p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 0)} \\
&= \frac{p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) p(\tilde{\theta}_c | \tau = 1) p(\tilde{\theta}_0 | \tau = 1) p(\tilde{\theta}_1 | \tau = 1)}{p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 0) p(\tilde{\theta}_c | \tau = 0) p(\tilde{\theta}_0 | \tau = 0) p(\tilde{\theta}_1 | \tau = 0)} \\
&= \frac{p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) p(\tilde{\theta}_0 | \tau = 1) p(\tilde{\theta}_1 | \tau = 1)}{p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 0) p(\tilde{\theta}_0 | \tau = 0) p(\tilde{\theta}_1 | \tau = 0)},
\end{aligned} \tag{15}
$$

where it is assumed that $p(\tilde{\theta}_c | \tau = 0) = p(\tilde{\theta}_c | \tau = 1)$ and where $p(\tilde{\theta}_c | \tau = j)$ and $p(\tilde{\theta}_j | \tau = j)$, $j = 0, 1$, denote the usual prior densities and $p(\tilde{\theta}_0 | \tau = 1)$ and $p(\tilde{\theta}_1 | \tau = 0)$ denote the pseudo prior densities or the linking densities employed to avoid the convergence problem with the Gibbs sampler. Carlin and Chib (1995) recommend using first-order approximations to the model-specific posterior distributions for $\tilde{\theta}_0$ and $\tilde{\theta}_1$, i.e., $p(\tilde{\theta}_0 | \tilde{z}_T, \tau = 0)$ and $p(\tilde{\theta}_1 | \tilde{z}_T, \tau = 1)$, as the pseudo prior distribution for $\tilde{\theta}_0$ conditional on $\tau = 1$ and that for $\tilde{\theta}_1$ conditional on $\tau = 0$, respectively. Notice that the joint distribution of $\tilde{\theta}_c$, $\tilde{\theta}_0$, $\tilde{\theta}_1$ and τ is given by:

$$
\begin{aligned}
p(\tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau | \tilde{z}_T) &\propto p(\tilde{z}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau) \\
&= p(\tilde{z}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau) p(\tilde{\theta}_c | \tau) p(\tilde{\theta}_0 | \tau) p(\tilde{\theta}_1 | \tau) p(\tau).
\end{aligned} \tag{16}
$$

The above joint posterior distribution and the discussion in this section, based on Carlin and Chib (1995), lead to the following procedure for Gibbs sampling in the presence of varying parameter space:

Step 1:

Generate τ from $p(\tau | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tilde{z}_T)$, given by equation (14). 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 (14), we set $\tau = 1$; otherwise, we set $\tau = 0$.

Step 2:

Generate $\tilde{\theta}_0$ and $\tilde{\theta}_1$.

If $\tau = 0$:

i) Generate $\tilde{\theta}_0$ from the usual full conditional density, $p(\tilde{\theta}_0 | \tilde{\theta}_c, \tilde{z}_T, \tau = 0)$.

ii) Generate $\tilde{\theta}_1$ from $p(\tilde{\theta}_1|\tau = 0)$, the pseudo prior density;

If $\tau = 1$:

i) Generate $\tilde{\theta}_0$ from $p(\tilde{\theta}_0|\tau = 1)$, the pseudo prior density;

ii) Generate $\tilde{\theta}_1$ from the usual full conditional density, $p(\tilde{\theta}_1|\tilde{\theta}_c, \tilde{z}_T, \tau = 1)$.

Step 3:

Generate $\tilde{\theta}_c$ from $p(\tilde{\theta}_c|\tilde{\theta}_\tau, \tilde{z}_T, \tau)$.

4.2. Testing for Markov Switching in a Univariate Model

Within the univariate model specification in Section 2.1, we assume that the vector of autoregressive parameters $\tilde{\phi}$, associated with the $\phi(L)$ term, is not common to the linear and Markov-switching ($\tau = 1$) models. That is, we assume that $\tilde{\phi} = \tilde{\phi}^0$ for the linear model ($\tau = 0$) and $\tilde{\phi} = \tilde{\phi}^1$ for the Markov-switching model ($\tau = 0$), in order to incorporate the possibility that $p(\tilde{\phi}|\tau = 0) \neq p(\tilde{\phi}|\tilde{\mu} = 0, \tau = 1)$, due to Perron (1990). Then, using the notation used in Section 4.1, we have $\tilde{\theta}_c = \sigma^2$, $\tilde{\theta}_0 = \tilde{\phi}^0$, and $\tilde{\theta}_1 = [\,\tilde{\mu}^{1\prime} \quad \tilde{\phi}^{1\prime} \quad \tilde{p}'\,]'$, where $\tilde{p}$ is the vector of transition probabilities. Unlike the framework discussed in Section 4.1, we have a vector of latent variables $\tilde{S}_T = [\,S_1 \quad S_2 \quad \ldots S_T\,]'$ that is not identified for a linear model. However, this problem is easily solved by employing a pseudo prior for $\tilde{\mu}^1$ conditional on $\tau = 0$.

We first consider the following joint posterior density of $\tilde{\mu}^1$, $\tilde{\phi}^0$, $\tilde{\phi}^1$, σ^2, $\tilde{S}_T$, and $\tilde{p}$, conditional on $\tau = 1$:

$$\begin{aligned}
&p(\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}|\Delta\tilde{c}_T, \tau = 1)\\
&\propto p(\Delta\tilde{c}_T, \tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}|\tau = 1)\\
&= p(\Delta\tilde{c}_T|\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p})p(\tilde{S}_T|\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tau = 1)p(\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}|\tau = 1) \qquad (17)\\
&= p(\Delta\tilde{c}_T|\tilde{\mu}^1, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tau = 1)p(\tilde{S}_T|\tilde{\mu}^1, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tau = 1)\\
&\qquad\qquad \times p(\tilde{\phi}^0|\tau = 1)p(\tilde{\phi}^1|\tau = 1)p(\sigma^2|\tau = 1)p(\tilde{p}|\tau = 1)p(\tilde{\mu}^1|\tau = 1),
\end{aligned}$$

where $p(\tilde{\phi}^0|\tau = 1)$ is the pseudo prior density and $p(\tilde{\phi}^1|\tau = 1)$, $p(\sigma^2|\tau = 1)$, $p(\tilde{p}|\tau = 1)$, and $p(\tilde{\mu}^1|\tau = 1)$ are the usual prior densities, and they are assumed a priori independent of one another. Thus, conditional on $\tau = 1$, $\tilde{\phi}^0$ is generated from the pseudo prior density. Generation of the other variates are based on the appropriate full conditional distributions.

Conditional on $\tau = 0$, we have:

$$\begin{aligned} & p(\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}|\Delta\tilde{c}_T, \tau = 0) \\ & \quad \propto p(\Delta\tilde{c}_T, \tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}|\tau = 0) \\ & \quad = p(\Delta\tilde{c}_T|\tilde{\phi}^0, \sigma^2, \tau = 0)p(\tilde{S}_T|\tilde{\mu}^1, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tau = 0) \\ & \qquad \times p(\tilde{\phi}^0|\tau = 0)p(\tilde{\phi}^1|\tau = 0)p(\sigma^2|\tau = 0)p(\tilde{p}|\tilde{\mu}^1, \tau = 0)p(\tilde{\mu}^1|\tau = 0), \end{aligned} \tag{18}$$

where $p(\tilde{\phi}^0|\tau = 0)$ and $p(\tilde{\sigma}^2|\tau = 0)$ are the usual prior densities; $p(\tilde{\phi}^1|\tau = 0)$ and $p(\tilde{\mu}^1|\tau = 0)$ are the pseudo prior densities or the linking densities. Conditional on $\tau = 0$, we can generate $\tilde{\phi}^1$ and $\tilde{\mu}^1$ from the pseudo prior densities. However, notice that $\tilde{p}$ is not independent of the pseudo prior for $\tilde{\mu}^1$ conditional on $\tau = 0$. This is because $\tilde{S}_T$ is pseudo-identified conditional on the pseudo prior for $\tilde{\mu}^1$ and $\tilde{p}$ is pseudo-identified conditional on $\tilde{S}_T$. [4]

Thus, $p(\tilde{p}|\tilde{\mu}^1, \tau = 0)$ may not be used as a linking density.

Finally, notice that the joint posterior distribution of $\tilde{\mu}^1$, $\tilde{\phi}^0$, $\tilde{\phi}^1$, σ^2, $\tilde{S}_T$, $\tilde{p}$, and τ is given by:

$$\begin{aligned} & p(\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}, \tau|\Delta\tilde{c}_T) \\ & \quad \propto p(\Delta\tilde{c}_T, \tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}, \tau = 0) \\ & \quad = p(\Delta\tilde{c}_T|\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{S}_T, \tilde{p}, \tau)p(\tilde{S}_T|\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tau)p(\tilde{\mu}^1, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}|\tau)p(\tau). \end{aligned} \tag{19}$$

Thus, the following procedure for the Gibbs sampler results:

Step 1:

Generate τ from $p(\tau|\tilde{\phi}^0, \tilde{\phi}^1, , \sigma^2, \tilde{\mu}^1, \tilde{p}, \Delta\tilde{c}_T)$.

[4] Throughout the paper, we use the term 'pseudo-identified' to denote that the conditional distribution of a variate exists only when the pseudo priors are given. Notice that the conditional distributions of $\tilde{S}_T$ and $\tilde{p}$ do not exist for a linear model in the absence of the pseudo prior for $\tilde{\mu}^1$.

Step 2:

If $\tau = 0$:

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

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

iii) Generate $\tilde{\phi}^0$ from $p(\tilde{\phi}^0|\sigma^2, \Delta\tilde{c}_T, \tau = 0)$.

iv) Set $\tilde{\mu} = 0$; Set $\tilde{\phi} = \tilde{\phi}^1$.

If $\tau = 1$:

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

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

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

iv) Set $\tilde{\mu} = \tilde{\mu}^1$; Set $\tilde{\phi} = \tilde{\phi}^0$.

Step 3:

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

Step 4:

Generate $\tilde{p}$ from $p(\tilde{p}|S_T)$.

Step 5:

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

To complete the above procedure, we need to specify the conditional Bayes factor to complete the full conditional distribution of τ for Step 1. After integrating $\tilde{S}_T$ out of the joint densities,

$$p(\Delta\tilde{c}_T, \tilde{S}_T, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tilde{\mu}^1|\tau = j), \quad j = 0, 1,$$

the conditional Bayes factor can be derived as:

$$
\begin{aligned}
&C_{10}\\
&= \frac{p(\Delta\tilde{c}_T, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tilde{\mu}^1 | \tau = 1)}{p(\Delta\tilde{c}_T, \tilde{\phi}^0, \tilde{\phi}^1, \sigma^2, \tilde{p}, \tilde{\mu}^1 | \tau = 0)}\\
&= \frac{p(\Delta\tilde{c}_T|\tilde{\phi}^1, \sigma^2, \tilde{p}, \tilde{\mu}^1, \tau = 1)p(\tilde{\phi}^0|\tau = 1)p(\tilde{\phi}^1|\tau = 1)p(\sigma^2|\tau = 1)p(\tilde{p}|\tau = 1)p(\tilde{\mu}^1|\tau = 1)}{p(\Delta\tilde{c}_T|\tilde{\phi}^0, \sigma^2, \tau = 0)p(\tilde{\phi}^0|\tau = 0)P(\tilde{\phi}^1|\tau = 0)p(\sigma^2|\tau = 0)p(\tilde{p}|\tilde{\mu}^1, \tau = 0)p(\tilde{\mu}^1|\tau = 0)}\\
&= \frac{p(\Delta\tilde{c}_T|\tilde{\phi}^1, \sigma^2, \tilde{p}, \tilde{\mu}^1, \tau = 1)p(\tilde{\phi}^0|\tau = 1)p(\tilde{\phi}^1|\tau = 1)p(\tilde{\mu}^1|\tau = 1)}{p(\Delta\tilde{c}_T|\tilde{\phi}^0, \sigma^2, \tau = 0)p(\tilde{\phi}^0|\tau = 0)P(\tilde{\phi}^1|\tau = 0)p(\tilde{\mu}^1|\tau = 0)},
\end{aligned}
\tag{20}
$$

and where it is a priori assumed that $p(\sigma^2|\tau = 0) = p(\sigma^2|\tau = 1)$ and $p(\tilde{p}|\tilde{\mu}^1, \tau = 0) = p(\tilde{p}|\tau = 1)$ without loss of generality. The term $p(\Delta\tilde{c}_T|\tilde{\phi}^1, \sigma^2, \tilde{\mu}^1, \tilde{p}, \tau = 1)$ can be evaluated as a byproduct of running the Hamilton filter (1989), given the conditioning parameters.

4.3. Testing for Markov Switching in a Dynamic Factor Model

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 unobserved factor component. $\Delta\tilde{c}_T = [\,\Delta c_1 \quad \ldots \quad \Delta c_T\,]'$ is no longer a vector of observed data. It is a vector of the 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 $\tilde{\theta}_c = [\,\tilde{\gamma}' \quad \tilde{\sigma}^{2\prime} \quad \tilde{\psi}'\,]'$, where $\tilde{\gamma}$ and $\tilde{\psi}$ are the vectors of parameters associated with $\gamma_i(L)$ and $\psi_i(L)$, respectively, for $i = 1, .., n$; $\tilde{\sigma}^2 = [\,\sigma_1^2 \quad ... \quad \sigma_n^2\,]'$. The remaining notation used in this section is the same as in Section 4.2 for a univariate test. For example, the vector of parameters unique to the linear model ($\tau = 0$) is given by $\tilde{\theta}_0 = \tilde{\phi}^0$ and that unique to the Markov-switching model ($\tau = 1$) is given by $\tilde{\theta}_1 = [\,\tilde{\mu}^{1\prime} \quad \tilde{\phi}^{1\prime} \quad \tilde{p}'\,]'$. As in the case of the univariate test, we assume $\tilde{\phi} = \tilde{\phi}^0$ for the linear model and $\tilde{\phi} = \tilde{\phi}^1$ for the Markov-switching model, in order to incorporate the possibility that $p(\tilde{\phi}|\tau = 0) \neq p(\tilde{\phi}|\tilde{\mu} = 0, \tau = 1)$.

The additional difficulty in testing for Markov switching within the multivariate framework is associated with the fact that the unobserved factor component may not be the same for the linear and the Markov-switching models. That is, we have $\Delta\tilde{c}_T = \Delta\tilde{c}_T^0$ conditional on $\tau = 0$ and we have $\Delta\tilde{c}_T = \Delta\tilde{c}_T^1$ conditional on $\tau = 1$. Thus, in order to generate both $\Delta\tilde{c}_T^0$ and $\Delta\tilde{c}_T^1$ as well as all the other variates at each run of the Gibbs sampler, we consider the joint posterior distributions of $\Delta\tilde{c}_T^0$, $\Delta\tilde{c}_T^1$, $\tilde{S}_T$, $\tilde{\theta}_c$, $\tilde{\theta}_0$, and $\tilde{\theta}_1$, conditional on $\tau = j$, $j = 0, 1$.

Conditional on $\tau = 1$, we have:

$$\begin{aligned}
&p(\Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \Delta\tilde{y}_T, \tau = 1) \\
&\quad \propto p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 1) \\
&\quad = p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) p(\Delta\tilde{c}_T^1 | \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) \\
&\qquad \times p(\tilde{S}_T | \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau = 1) p(\tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 1) \\
&\quad = p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tau = 1) p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tilde{\phi}^0, \tau = 1) p(\Delta\tilde{c}_T^1 | \tilde{S}_T, \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 1) \\
&\qquad \times p(\tilde{S}_T | \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 1) p(\tilde{\theta}_c | \tau = 1) p(\tilde{\phi}^0 | \tau = 1) p(\tilde{\phi}^1 | \tau = 1) p(\tilde{p} | \tau = 1) p(\tilde{\mu}^1 | \tau = 1),
\end{aligned} \tag{21}$$

where $p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tilde{\phi}^0, \tau = 1)$ is the pseudo density of $\Delta\tilde{c}_T^0$ that corresponds to the pseudo prior for $\tilde{\phi}^0$, conditional on $\tau = 1$. Conditional on $\tau = 1$, we can generated $\tilde{\phi}^0$ from the pseudo prior density. Then, given $\tilde{\phi}^0$ and $\tilde{\mu} = 0$, $\Delta\tilde{c}_T^0$ is pseudo-identified. All the other variates are generated in the usual way, from the appropriate full conditional distributions.

Conditional on $\tau = 0$, we have:

$$\begin{aligned}
&p(\Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \Delta\tilde{y}_T, \tau = 0) \\
&\quad \propto p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1 | \tau = 0) \\
&\quad = p(\Delta\tilde{y}_T | \Delta\tilde{c}_T^0, \tilde{\theta}_c, \tau = 0) p(\Delta\tilde{c}_T^0 | \tilde{\theta}_c, \tilde{\phi}^0, \tau = 0) p(\Delta\tilde{c}_T^1 | \tilde{S}_T, \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 0) \\
&\qquad \times p(\tilde{S}_T | \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 0) p(\tilde{\theta}_c | \tau = 0) p(\tilde{\phi}^0 | \tau = 0) p(\tilde{\phi}^1 | \tau = 0) p(\tilde{p} | \tilde{\mu}^1, \tau = 0) p(\tilde{\mu}^1 | \tau = 0),
\end{aligned} \tag{22}$$

where $p(\Delta\tilde{c}_T^1 | \tilde{S}_T, \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 0)$ and $p(\tilde{S}_T | \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \tau = 0)$ are the pseudo densities or the linking densities for $\Delta\tilde{c}_T^1$ and $\tilde{S}_T$, respectively, that corresponds to the pseudo priors for $\tilde{\phi}^1$ and $\tilde{\mu}^1$, conditional on $\tau = 0$. Thus, conditional on $\tau = 0$, we can generate $\tilde{\phi}^1$ and $\tilde{\mu}^1$ from the pseudo prior densities. Given the pseudo values for $\tilde{\phi}^1$ and $\tilde{\mu}^1$ conditional on $\tau = 0$, the vectors $\Delta\tilde{c}_T^1$ and $\tilde{S}_T$ are pseudo-identified. All the other variates are generated in the usual way, from appropriate full conditional distributions.

Finally, the joint posterior distribution for all the variates to be generated is given by:

$$
\begin{aligned}
p(\Delta\tilde{c}_T^0, &\Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau|\Delta\tilde{y}_T) \\
&\propto p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau) \\
&= p(\Delta\tilde{y}_T|\Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau)p(\Delta\tilde{c}_T^0|\tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau)p(\Delta\tilde{c}_T^1|\tilde{S}_T, \tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau) \\
&\quad \times p(\tilde{S}_T|\tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1, \tau)p(\tilde{\theta}_c, \tilde{\theta}_0, \tilde{\theta}_1|\tau)p(\tau).
\end{aligned}
\tag{23}
$$

Thus, using equations (21), (22), and (23), we can design the following procedure for the Gibbs sampler:

Step 1:

Generate τ from $p(\tau|\Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{\phi}^0, \tilde{\phi}^1, \tilde{\mu}^1, \tilde{p}, \Delta\tilde{y}_T)$.

Step 2:

If $\tau = 0$:

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

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

iii) Generate $\tilde{\phi}^0$ from $p(\tilde{\phi}^0|\Delta\tilde{c}_T^0, \tau = 0)$, where, conditional on $\Delta\tilde{c}_T^0$, data $\Delta\tilde{y}_T$ and all the other variates are irrelevant.

iv) Generate $\Delta\tilde{c}_T^1$ from $p(\Delta\tilde{c}_T^1|\tilde{S}_T, \tilde{\theta}_c, \tilde{\phi}^1, \tilde{\mu}^1, \Delta\tilde{y}_T, \tau = 0)$, where $\tilde{\phi}^1$ and $\tilde{\mu}^1$ are generated from the pseudo prior distributions.

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

vi) Set $\tilde{\mu} = 0$; $\tilde{\phi} = \tilde{\phi}^0$; and $\Delta\tilde{c}_T = \Delta\tilde{c}_T^0$.

If $\tau = 1$:

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

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

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

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

v) Generate $\Delta\tilde{c}_T^0$ from $p(\Delta\tilde{c}_T^0|\tilde{\theta}_c, \tilde{\phi}^0, \Delta\tilde{y}_T, \tau = 1)$, where $\tilde{\phi}^0$ is generated from the pseudo prior distribution.

vi) Set $\tilde{\mu} = \tilde{\mu}^1$; $\tilde{\phi} = \tilde{\phi}^1$; and $\Delta\tilde{c}_T = \Delta\tilde{c}_T^1$.

Step 4:

Generate $\tilde{S}_T$ from $p(\tilde{S}_T|\tilde{\phi}^1, \tilde{\mu}^1, \tilde{p}, \Delta\tilde{c}_T^1)$. 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 $\tilde{\theta}_c$ from $p(\tilde{\theta}_c|\Delta\tilde{c}_T, \Delta\tilde{y}_T)$, where, conditional on $\Delta\tilde{c}_T$, $\tilde{\theta}_c$ is independent of the other variates.

The conditional Bayes factor in the presence case can easily be derived as in the case of the univariate test. After integrating $\tilde{S}_T$ from the joint densities,

$$p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{S}_T, \tilde{\theta}_c, \tilde{\phi}^0, \tilde{\phi}^1, \tilde{\mu}^1, \tilde{p}|\tau = j), \quad j = 0, 1,$$

the conditional Bayes factor in favor of model can be derived as:

$$\begin{aligned} &C_{10} \\ &= \frac{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{\phi}^0, \tilde{\phi}^1, \tilde{\mu}^1, \tilde{p}|\tau = 1)}{p(\Delta\tilde{y}_T, \Delta\tilde{c}_T^0, \Delta\tilde{c}_T^1, \tilde{\theta}_c, \tilde{\phi}^0, \tilde{\phi}^1, \tilde{\mu}^1, \tilde{p}|\tau = 0)} \\ &= \frac{p(\Delta\tilde{y}_T|\Delta\tilde{c}_T^1, \tilde{\theta}_c, \tau = 1)p(\Delta\tilde{c}_T^0|\tilde{\theta}_c, \tilde{\phi}^0\tau = 1)p(\Delta\tilde{c}_T^1|\tilde{\theta}_c, \tilde{\phi}^1, \mu^1, \tilde{p}, \tau = 1)}{p(\Delta\tilde{y}_T|\Delta\tilde{c}_T^0, \tilde{\theta}_c, \tau = 0)p(\Delta\tilde{c}_T^0|\tilde{\theta}_c, \tilde{\phi}^0\tau = 0)p(\Delta\tilde{c}_T^1|\tilde{\theta}_c, \tilde{\phi}^1, \mu^1, \tilde{p}, \tau = 0)} \times RR, \end{aligned} \tag{24}$$

with

$$RR = \frac{p(\tilde{\phi}^0|\tau = 1)p(\tilde{\phi}^1|\tau = 1)p(\tilde{\mu}^1|\tau = 1)}{p(\tilde{\phi}^0|\tau = 0)p(\tilde{\phi}^1|\tau = 0)p(\tilde{\mu}^1|\tau = 0)}. \tag{25}$$

Notice that, as in the case of the univariate test, 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. Empirical Tests of Markov Switching

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). [5] The time period is 1960.1 through 1995.1, which covers Kim and Nelson's (1998) sample period. We use the demeaned log-differences for all the series. [6]

5.2. Specification of the Priors and the Pseudo Priors

Since our goal is the computation of the Bayes factors, we assume that each prior is proper. A consequence of employing non informative priors for the parameters being tested will be to force the test results to favor the null hypothesis. [7] 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)). The priors employed are summarized as follows: [8]

<u>Priors</u>

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

[6] We have done some limited experiments allowing for possible shift in mean in the demeaning process. For example, in order to take into account the post-1973 productivity slowdown, the pre-1973 subsample and the post-1973 subsample have been demeaned separately. However, the empirical results were qualitatively robust with respect to different demeaning processes.

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

[8] For issues concerning the sensitivity analysis and the choice of the priors in a Bayesian model selection, refer to Kass and Raftery (1995). The sensitivity analysis suggests that qualitative results are robust with respect to different priors employed for the parameters. Thus, we do not report the details of sensitivity analysis.

$$\sigma^2 \sim IG(4,4), \tag{26}$$

$$\tilde{\phi}^0|\tau = 0 \sim N(0, I_2), \tag{27}$$

$$\tilde{\phi}^1|\tau = 1 \sim N(0, I_2), \tag{28}$$

$$\mu_0^1 \mid \tau = 1 \quad \sim \quad N(-0.5, 1)_{1[\mu_0^1<0]}, \tag{29}$$

$$\mu_1^1 \mid \tau = 1 \quad \sim \quad N(0.2, 1)_{1[\mu_1^1>0]}, \tag{30}$$

$$p_{00} \mid \tau = 1 \quad \sim \quad beta(4, 1), \tag{31}$$

$$p_{11} \mid \tau = 1 \quad \sim \quad beta(4, 1), \tag{32}$$

$$\gamma_i \sim N(0, 1), \quad i = 1, 2, 3, 4, \tag{33}$$

$$\tilde{\psi}_i \sim N(0, I_2), \quad i = 1, 2, 3, 4 \tag{34}$$

$$\sigma_i^2 \sim IG(4, 4), \quad i = 1, 2, 3, 4, \tag{35}$$

where $beta(.,.)$ refers to a Beta distribution; IG refers to inverse Gamma distribution; and $1[.]$ refers to an indicator function. Priors in (26)-(31) are relevant for the univariate tests and those in (27)-(35) are relevant for the multivariate tests.

The choice of the pseudo priors for μ_0^1, μ_1^1, $\tilde{\phi}^0$, and $\tilde{\phi}^1$ is important for the convergence of the Gibbs sampler. [9] 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 both linear and Markov-switching models. Tables 1 and 2 summarize the results for univariate linear model ($\tau = 0$) and for the univariate Markov-switching model ($\tau = 1$). Tables 3 and 4 summarize the results for linear dynamic factor model ($\tau = 0$) and for the dynamic factor model with Markov switching ($\tau = 1$). The Normal approximation to the marginal posterior distribution for $\tilde{\phi}^0$ in Tables 1 and 3 are employed as the pseudo prior for $\tilde{\phi}^0$ conditional on $\tau = 1$. Likewise, the pseudo

[9] Note that, even though the vector of transition probabilities $\tilde{p}$ is not identified under the linear model, we do not employ a pseudo prior. Given the pseudo priors for $\tilde{\mu}^1$ and $\tilde{\phi}^1$ conditional on $\tau = 0$, the state vector $\tilde{S}_T$ is pseudo-identified, and thus, $\tilde{p}$ is pseudo-identified.

prior distributions for $\tilde{\phi}^1$, μ_0^1 and μ_1^1, conditional on $\tau = 0$, are obtained from the Normal approximations to the marginal posterior distributions for these variates in Tables 2 and 4. [10] Thus, the pseudo priors employed are as follows:

Pseudo Priors for Univariate Test [From Tables 1 and 2]

$$\mu_0^1 \mid \tau = 0 \;\sim\; N(-0.703, 0.510^2)_{1[\mu_0^1<0]}, \tag{36}$$

$$\mu_1^1 \mid \tau = 0 \;\sim\; N(0.194, 0.158^2)_{1[\mu_1^1>0]}, \tag{37}$$

$$\tilde{\phi}^1 \mid \tau = 0 \;\sim\; N\left(\begin{pmatrix} 0.272 \\ 0.041 \end{pmatrix}, \begin{pmatrix} 0.093^2 & 0 \\ 0 & 0.080^2 \end{pmatrix}\right), \tag{38}$$

$$\tilde{\phi}^0 \mid \tau = 1 \;\sim\; N\left(\begin{pmatrix} 0.321 \\ 0.043 \end{pmatrix}, \begin{pmatrix} 0.074^2 & 0 \\ 0 & 0.075^2 \end{pmatrix}\right), \tag{39}$$

Pseudo Priors for Multivariate Test [From Tables 3 and 4]

$$\mu_0^1 \mid \tau = 0 \;\sim\; N(-1.833, 0.540)_{1[\mu_0^1<0]}, \tag{40}$$

$$\mu_1^1 \mid \tau = 0 \;\sim\; N(0.320, 0.144)_{1[\mu_1^1>0]}, \tag{41}$$

$$\tilde{\phi}^1 \mid \tau = 0 \;\sim\; N\left(\begin{pmatrix} 0.332 \\ 0.021 \end{pmatrix}, \begin{pmatrix} 0.097^2 & 0 \\ 0 & 0.070^2 \end{pmatrix}\right), \tag{42}$$

$$\tilde{\phi}^0 \mid \tau = 1 \;\sim\; N\left(\begin{pmatrix} 0.645 \\ 0.024 \end{pmatrix}, \begin{pmatrix} 0.075^2 & 0 \\ 0 & 0.065^2 \end{pmatrix}\right), \tag{43}$$

5.3. Empirical Results

For each of the univariate and multivariate cases, we perform two tests under alternative assumptions. In **Test #1**, we allow for the possibility that $\tilde{\phi}$ may be different for linear and Markov-switching models. That is, we assume $\tilde{\phi} = \tilde{\phi}^0$ for the linear model

[10] As Carlin and Chib (1995) note, we are not using the data to select the prior, but only the pseudo prior. Figures 1 and 2 depict inferences about recession probabilities obtained from a univariate Markov-switching model and a dynamic factor model with Markov-switching, respectively.

and $\tilde{\phi} = \tilde{\phi}^1$ for the Markov-switching model. Except for $\tilde{\phi}^0$, $\tilde{\phi}^1$, $\tilde{\mu}^1$, and $\tilde{p}$, all the other parameters are assumed common to both models. In this test, we need pseudo priors for $\tilde{\phi}^0$, $\tilde{\phi}^1$, and $\tilde{\mu}^1$. In **Test #2**, we force $\tilde{\phi}^0 = \tilde{\phi}^1$. That is, we treat $\tilde{\phi}$ is common to both linear and Markov-switching models. All the other parameters including $\tilde{\phi}$, except for $\tilde{\mu}^1$ and $\tilde{p}$, are common to both models. Notice that the pseudo priors specified in Section 5.2 are for Test #1. When $\tilde{\phi}$ is assumed common to both linear and nonlinear models in Test #2, however, we need a pseudo prior only for $\tilde{\mu}^1$. The Gibbs sampling procedures in Sections 4.2 and 4.3 can accordingly modified to deal with Test #2.

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 posterior probability of Markov-switching ($Pr(\tau = 1|\tilde{z}_T)$, where $\tilde{z}_T$ is data) is obtained by the proportion of the posterior simulations in which $\tau = 1$. Tables 5 and 6 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 0.054 and 0.713 for Test #1, in which $\tilde{\phi}^0 \neq \tilde{\phi}^1$. 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 0.276 and 0.514. With a prior probability of 0.5, for example, the posterior probability is 0.269 and the implied Bayes factor is 0.368. These results may be interpreted as sample evidence being against Markov-switching, even though the posterior probability is quite sensitive to the prior probability. Forcing $\tilde{\phi}^0 = \tilde{\phi}^1$ (that is, treating $\tilde{\phi}$ to be common to both models) in Test #2 does not seem to affect the results much.

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 shown in Table 6. In Test #1, as we change the prior probability from 0.1 to 0.9, the posterior probability ranges from 0.628 to 0.732. For the prior probability of 0.5, the posterior probability is 0.683 and the implied Bayes factor is 2.155. These considerations might suggest that the data slightly favors Diebold and Rudebusch's (1996) dynamic factor model with Markov-switching over Stock and Watson's (1991) linear dynamic factor model. However, unlike the univariate test results, the Bayes factor is quite sensitive to the prior probability and ranges from 15.194 to 0.303. This suggests that the multivariate test results leave more room for subjective interpretation than do the univariate test results. By treating $\tilde{\phi}$ to be common to both linear and Markov-switching models in Test #2, we get consistently lower posterior probabilities. This suggests that imposing the assumption $p(\tilde{\phi}|\tau = 0) = p(\tilde{\phi}|\tilde{\mu} = 0, \tau = 1)$ in equation (11) could result in a test results that are biased toward a direction less favorable to the Markov-switching model.

While the univariate and multivariate test results leave room for subjective interpretation when examined separately, a comparison of the two allows us to draw a conclusion which is objective enough: Evidence of Markov-switching, if exists, is much more compelling in the multivariate tests. This should not surprise us; if the dynamic factor model is correct then the multivariate data contain more information about whether Markov switching occurs than do individual series.

6. Summary and Discussion

In this paper, we have presented 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 either 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 (1999b) find strong sample evidence in favor of a narrowing gap between the growth rates during booms and recessions. Such structural break in the shift parameters has not been taken into account in this paper. In addition, while we investigate Markov-switching in the growth rate of the GDP series in this paper, Kim and Nelson (1999c) 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 test results are open to subjective interpretation, a comparison of the two allows us to draw a conclusion which is objective enough: Evidence of Markov switching or regime switching 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] Diebold, Francis X. and Glenn D. Rudebusch, 1998, *Business Cycles: Durations, Dynamics, and Forecasting* Princeton: Princeton University Press. Forthcoming.

[11] Filardo, Andrew J. and Stephen F. Gordon, "Business Cycle Turning Points: Two Empirical Business Cycle Model Approaches," in Rothman Philip (ed.), *Nonlinear Time Series Analysis of Economic and Financial Data,* 1-32, Kluwer Academic Press: Boston.

[12] 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.

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

[14] 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.

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

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

[17] 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.

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

[19] 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.

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

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

[22] Kim, Chang-Jin and Charles R. Nelson, 1998, "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.

[23] Kim, Chang-Jin and Charles R. Nelson, 1999a, *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications,* MIT Press: Cambridge.

[24] Kim, Chang-Jin and Charles R. Nelson, 1999b, "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle" forthcoming, *Review of Economics and Statistics* 81(4).

[25] Kim, Chang-Jin and Charles R. Nelson, 1999c, "Friedman's Plucking Model of Busi-

ness Fluctuation: Tests and Estimates of Permanent and Transitory Components," *Journal of Money, Credit, and Banking* 31(3), Part 1, 317-334.

[26] Koop, Gary and Simon M. Potter, 1999, "Bayes Factors and Nonlinearity: Evidence from Economic Time Series," *Journal of Econometrics*, 88, 251-281.

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

[28] 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.

[29] 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.

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

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

[32] 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.

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

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

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

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

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

Table 1. Summary of Results for Univariate Linear Model

	Prior		Posterior	
	Mean	**Std Dev**	**Mean**	**Std Dev**
p_{00}	–	–	–	–
p_{11}	–	–	–	–
μ_0^1	–	–	–	–
μ_1^1	–	–	–	–
ϕ_1^0	0	1	**0.321**	**0.074**
ϕ_2^0	0	1	**0.043**	**0.075**
σ^2	1.33	0.943	0.909	0.098

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

Table 2. Summary of Results for Univariate Markov-Switching Model

	Prior		**Posterior**	
	Mean	**Std Dev**	**Mean**	**Std Dev**
p_{00}	0.8	0.163	0.716	0.157
p_{11}	0.9	0.090	0.901	0.081
μ_0^1	-0.5	1	**-0.703**	**0.510**
μ_1^1	0.2	1	**0.194**	**0.158**
ϕ_1^1	0	1	**0.272**	**0.093**
ϕ_2^1	0	1	**0.041**	**0.080**
σ^2	1.33	0.943	0.831	0.122

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

Table 3. Summary of Results for Multivariate Linear Model

		Prior		Posterior	
		Mean	Std Dev	Mean	Std Dev
Δc_t	p_{00}	–	–	–	–
	p_{11}	–	–	–	–
	μ_0^1	–	–	–	–
	μ_1^1	–	–	–	–
	ϕ_1^0	0	1	**0.545**	**0.075**
	ϕ_2^0	0	1	**0.024**	**0.065**
IP	γ_1	0	1	0.618	0.044
	ψ_{11}	0	1	-0.017	0.097
	ψ_{12}	0	1	-0.057	0.066
	σ_1^2	1.33	0.943	0.260	0.033
GMYZPQ	γ_2	0	1	0.228	0.024
	ψ_{21}	0	1	-0.294	0.053
	ψ_{22}	0	1	-0.053	0.051
	σ_2^2	1.33	0.943	0.333	0.024
MTQ	γ_3	0	1	0.482	0.041
	ψ_{31}	0	1	-0.367	0.054
	ψ_{32}	0	1	-0.162	0.052
	σ_3^2	1.33	0.943	0.665	0.052
LPNAG	γ_{40}	0	1	0.115	0.012
	ψ_{41}	0	1	0.029	0.068
	ψ_{42}	0	1	0.296	0.068
	σ_4^2	1.33	0.943	0.045	0.003
	γ_{41}	0	1	0.010	0.014
	γ_{42}	0	1	0.021	0.013
	γ_{43}	0	1	0.029	0.011

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

Table 4. Summary of Results for Multivariate Markov-Switching Model

		Prior		Posterior	
		Mean	Std Dev	Mean	Std Dev
Δc_t	p_{00}	0.8	0.163	0.833	0.070
	p_{11}	0.9	0.090	0.967	0.030
	μ_0^1	-0.5	1	**-1.833**	**0.540**
	μ_1^1	0.2	1	**0.320**	**0.144**
	ϕ_1^1	0	1	**0.332**	**0.097**
	ϕ_2^1	0	1	**0.021**	**0.070**
IP	γ_1	0	1	0.565	0.042
	ψ_{11}	0	1	0.006	0.068
	ψ_{12}	0	1	-0.048	0.062
	σ_1^2	1.33	0.943	0.257	0.033
GMYZPQ	γ_2	0	1	0.210	0.022
	ψ_{21}	0	1	-0.291	0.052
	ψ_{22}	0	1	-0.055	0.052
	σ_2^2	1.33	0.943	0.334	0.024
MTQ	γ_3	0	1	0.443	0.038
	ψ_{31}	0	1	-0.359	0.053
	ψ_{32}	0	1	-0.158	0.054
	σ_3^2	1.33	0.943	0.657	0.051
LPNAG	γ_{40}	0	1	0.103	0.012
	ψ_{41}	0	1	0.014	0.068
	ψ_{42}	0	1	0.282	0.067
	σ_4^2	1.33	0.943	0.044	0.003
	γ_{41}	0	1	0.013	0.012
	γ_{42}	0	1	0.019	0.011
	γ_{43}	0	1	0.027	0.010

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

Table 5. 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 ($\bar{\pi}_1 = Pr(\tau = 1 \mid \Delta\tilde{c}_T)$)				
Test #1	0.054	0.147	0.269	0.460	0.713
(BF)	(0.514)	(0.402)	(0.368)	(0.369)	(0.276)
Test #2	0.060	0.165	0.270	0.452	0.739
(BF)	(0.574)	(0.461)	(0.370)	(0.353)	(0.315)

1. **Test #1**: $\tilde{\phi}$ is not a vector of parameters common to both linear and Markov-switching models. [$\tilde{\phi}^0 \neq \tilde{\phi}^1$]

2. **Test #2**: $\tilde{\phi}$ is a vector of parameters common to both linear and Markov-switching models. [$\tilde{\phi}^0 = \tilde{\phi}^1$]

3. BF refers to the Bayes Factor in favor of the Markov-switching Model ($\tau = 1$).

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

Table 6. 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 ($\bar{\pi}_1 = Pr(\tau = 1 \mid \Delta\tilde{y}_T)$)				
Test #1	0.628	0.660	0.683	0.716	0.732
(BF)	(15.194)	(4.529)	(2.155)	(1.080)	(0.303)
Test #2	0.545	0.574	0.593	0.614	0.649
(BF)	(10.780)	(3.144)	(1.457)	(0.682)	(0.205)

1. **Test #1**: $\tilde{\phi}$ is not a vector of parameters common to both linear and Markov-switching models. [$\tilde{\phi}^0 \neq \tilde{\phi}^1$]
2. **Test #2**: $\tilde{\phi}$ is a vector of parameters common to both linear and Markov-switching models. [$\tilde{\phi}^0 = \tilde{\phi}^1$]
3. BF refers to the Bayes Factor in favor of the Markov-switching Model ($\tau = 1$).
4. 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

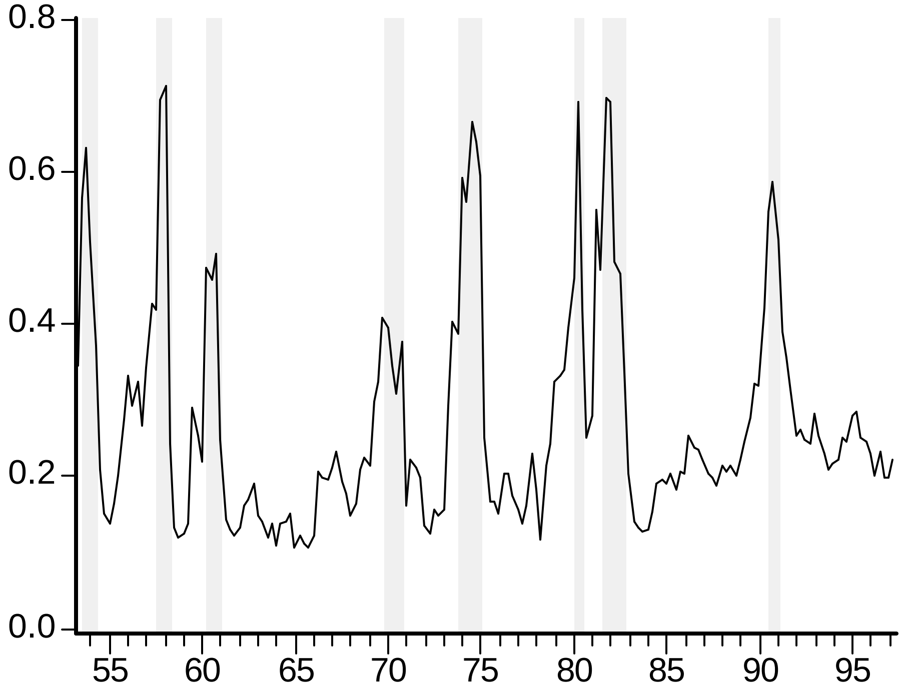

Figure 2. Probability of a Recession: Multivariate Model with Markov Switching

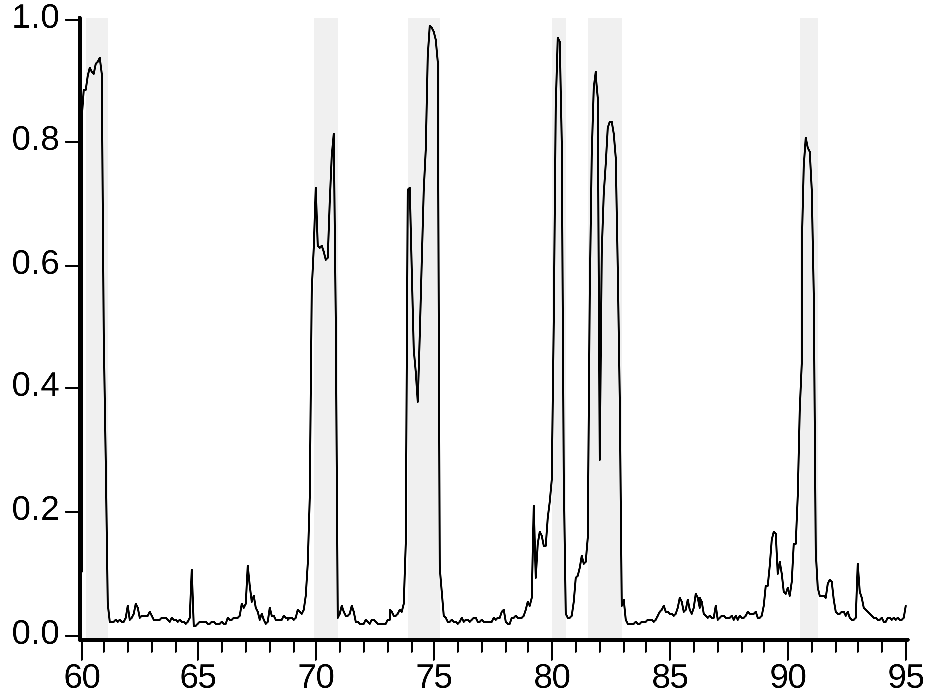

Figure 3. Sensitivity of Posterior Probabilities of Markov-Switching to Prior Probabilities: Univariate Tests

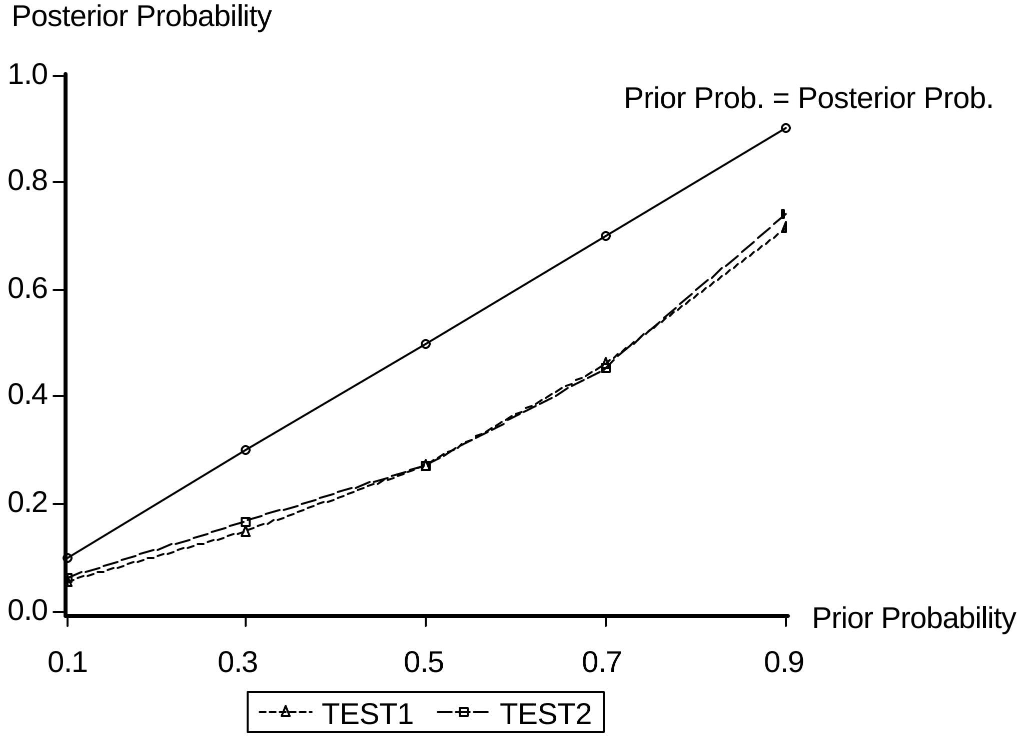

Figure 4. Sensitivity of Posterior Probabilities of Markov-Switching to Prior Probabilities: Multivariate Tests