# Bayesian Analysis of Business Cycles in Japan by Extending the Markov Switching Model

Toshiaki Watanabe[*,†]

**Abstract**

This paper analyzes business cycles in Japan by applying Markov switching (MS) models to monthly data on the coincident indicator of composite index (CI) during the period of 1985/01–2025/05 calculated by Economic and Social Research Institute (ESRI), Cabinet Office, the Government of Japan. During the latter half of the sample period, the Japanese economy experienced major shocks such as the global financial crisis in 2008, the Great East Japan Earthquake in 2011, the consumption tax hikes in 2014 and 2019, and the COVID-19 pandemic in 2020. CI fell sharply during these periods, which make it difficult to estimate business cycle turning points using the simple MS model. In this paper, the MS model is extended by incorporating Student's $t$-error and stochastic volatility (SV). Since it is difficult to evaluate the likelihood once SV is introduced, a Bayesian method via Markov chain Monte Carlo (MCMC) is employed. The MS model with $t$-error or SV is shown to provide the estimates of the business cycle turning points close to those published by ESRI. Bayesian model comparison based on marginal likelihood provides evidence that $t$-error is not needed once SV is introduced. Using the MS model with normal error and SV, structural changes in CI's mean growth rates during booms and recessions are also analyezed and two break points are found in the both mean growth rates. One is 2008/10 and the other is 2010/02, during which the mean growth rate during recession falls and that during boom rises due to the global financial crisis.

**JEL classification**: C11, C22, C51, C52, E32.

**Keywords**: Bayesian inference, Composite index, Marginal likelihood, Markov chain Monte Carlo, Markov switching model, Stochastic volatility, Student's $t$-distribution.

[*] Graduate School of Social Data Science, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601 JAPAN. E-mail: t.watanabe@r.hit-u.ac.jp

[†] Financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research (23H00048, 24H00142) is gratefully acknowledged. All remaining errors are solely my own responsibility.

# 1 Introduction

The Markov switching (MS) model proposed by Hamilton (1989) has frequently been applied to the empirical analysis of business cycles (Kim and Nelson (1998, 1999a,b), Watanabe (2014) and Ishihara and Watanabe (2015)). This model produces the posterior probability of boom and recession for each period, which can be used for dating the business cycle turning points.

This paper analyzes business cycles in Japan by applying Markov switching (MS) models to monthly data on the coincident indicator of composite index (CI) during the period of 1985/01–2025/05 calculated by Economic and Social Research Institute (ESRI), Cabinet Office, the Government of Japan. During the latter half of the sample period, the Japanese economy experienced major shocks such as the global financial crisis in 2008, the Great East Japan Earthquake in 2011, the consumption tax hikes in 2014 and 2019, and the COVID-19 pandemic in 2020. As shown in Figure 1, the CI fell sharply during these periods. In this paper, we first show that these large shocks make it difficult to estimate business cycle turning points using the simple MS model. The simple MS model estimates only the above five periods in which those large negative shocks occurred as recessions, and estimates all other periods as booms. Hence, we extend the simple MS model by incorporating Student's $t$-error and stochastic volatility (SV) and show that these extended models provide the business cycle turning points close to those published by the ESRI.

Once SV model is introduced, it is difficult to evaluate the likelihood. We employ a Bayesian estimation where the parameters and latent variables are sampled from the posterior distribution using Markov Chain Monte Carlo (MCMC) techniques and the obtained draws are used for estimating the parameters. We use the multimove sampler proposed by Kim and Nelson (1998, 1999a,b) to sample the state variable representing the boom or recession in the MS model and the block sampler proposed by Watanabe and Omori (2004) to sample the latent volatility in the SV model. We also use the method proposed by Watanabe (2001) to sample the degree-of-freedom of the Student's $t$-distribution.

Ishihara and Watanabe (2015) has conducted a similar research using CI during the period of 1985/01–2014/05. This paper extends the sample period of 1985/01–2025/05, but the difference between this paper and Ishihara and Watanabe (2015) is not only the sample period. There is a technical problem in Ishihara and Watanabe (2015). They conduct model comparison using the marginal likelihood calculated via the modified harmonic mean method proposed by Geweke (1999) with the complete-data likelihood, i.e., the joint density of the data and latent variables given the parameters. Nakajima et al. (2011) show that this method is theoretically correct, but Chan and Grant (2015) provide evidence that it has a substantial bias and tends to select the wrong model in practice. In this paper, we use a different method for calculating the marginal likelihood. For the MS model without SV, we use the modified harmonic mean method based on the observed-data likelihood, i.e., the density of the data without the latent variables. The observed-data likelihood can be calculated using the Hamilton (1989) filter. Once SV is introduced, it is not straightforward to calculate the observed-data likelihood. Thus, we use the Chib (1995) method, where we calculate the observed-data likelihood using the particle filter. Bayesian model comparison based on marginal likelihood provides evidence that $t$-error is not needed once the SV is introduced. Using the MS model with normal error and SV, structural changes in CI's mean growth rates during booms and recessions are also analyezed and two break points are found in the both mean growth rates. One is 2008/10 and the other is 2010/02, during which the mean growth rate during recession falls and that during boom rises due to the global financial crisis.

The rest of this article is organized as follows. Next section reviews the simple MS model and extends it by incorporating with $t$-error and SV. Section 3 explains the Bayesian method using MCMC for the analysis of MS models. The simple and extended MS models are fitted to the CI in Japan in Section 4. Section 5 concludes.

## 2 Markov Switching Model

Let $y_t$ denote the growth rate of CI and $S_t$ denote a dummy variable that takes 0 when the economy is in the recession regime and 1 when the economy is in the boom regime. The simplest version of MS models assumes that the mean of $y_t$, which is denoted by $\mu_t$, may vary depending on $S_t$ as follows.

$$y_t = \mu_t + \phi(y_{t-1} - \mu_{t-1}) + e_t, \tag{1}$$

$$\mu_t = \mu^{(0)}(1 - S_t) + \mu^{(1)} S_t, \quad \mu^{(0)} < \mu^{(1)}, \tag{2}$$

$$S_t = \begin{cases} 1 & \text{boom} \\ 0 & \text{recession} \end{cases}, \tag{3}$$

where $S_t$ is assumed to follow a Markov process with transition probabilities

$$\begin{aligned} p(S_t = 1 \mid S_{t-1} = 1) &= \pi_{11}, \\ p(S_t = 0 \mid S_{t-1} = 1) &= 1 - \pi_{11}, \\ p(S_t = 0 \mid S_{t-1} = 0) &= \pi_{00}, \\ p(S_t = 1 \mid S_{t-1} = 0) &= 1 - \pi_{00}. \end{aligned} \tag{4}$$

It is straightforward to extend the lag-length in equation (1). We, however, assume that the lag-length is 1 in all analyses. It is also straightforward to extend such that the other parameters such as $\phi$ and $\sigma^2$ may also switch, but we assume that only $\mu$ may switch here.

**Model 1: MS model with normal error and constant volatility**

$e_t$ in equation (1) is the error term. In Model 1, we assume that the distribution of $e_t$ is normal with the 0 mean and the constant variance $\sigma^2$ as follows.

$$e_t = \sigma\epsilon_t, \quad \epsilon_t \sim \text{i.i.d. N}(0,1), \tag{5}$$

where "i.i.d. N(0,1)" represents the identically and independently distributed standard normal distribution.

**Model 2: MS model with $t$-error and constant volatility**

As will be shown in Section 4, this model cannot detect the business cycle turning points. Therefore, we extend the distribution of $\epsilon_t$ in equation (5) to a fat-tail distribution. As a fat-tail distribution, we use the Student's $t$-distribution standardized such that the variance is 1 as follows.

$$e_t = \sigma\epsilon_t, \quad \epsilon_t \sim \text{i.i.d. standardized } t(\nu), \tag{6}$$

where "standardized $t(\nu)$" represents the standardized Student's $t$-distribution with the degree-of-freedom $\nu$. We assume that $\nu > 2$ because the Student's $t$-distribution would not have a finite variance otherwise.

**Model 3: MS model with normal error and SV**

We also introduce the stochastic volatility (SV) as follows.

$$\begin{aligned} e_t =& \sigma_t\epsilon_t, \quad \epsilon_t \sim \text{i.i.d. N}(0,1), \\ \log(\sigma_t^2) =& \omega + \psi\{\log(\sigma_{t-1}^2) - \omega\} + \eta_t, \\ & \eta_t \sim \text{i.i.d. N}(0, \sigma_\eta^2), \end{aligned} \tag{7}$$

where the parameter $\psi$ in equation (7) captures the autocorrelation in the log-volatility.

**Model 4: MS model with $t$-error and SV**

We also estimate the MS model with the both $t$-error and SV by combing Models 2 and 3.

$$\begin{aligned} e_t =& \sigma_t\epsilon_t, \quad \epsilon_t \sim \text{i.i.d. standardized } t(\nu), \\ \log(\sigma_t^2) =& \omega + \psi\{\log(\sigma_{t-1}^2) - \omega\} + \eta_t, \\ & \eta_t \sim \text{i.i.d. N}(0, \sigma_\eta^2). \end{aligned} \tag{8}$$

# 3 Estimation Method

Once SV is introduced, it is difficult to evaluate the likelihood analytically. We employ a Bayesian method using MCMC.

In Models 2 and 4, $\epsilon_t$ can be represented as

$$\epsilon_t = \sqrt{\lambda_t} z_t, \quad (\nu - 2)/\lambda_t \sim \chi^2(\nu). \tag{9}$$

Following Watanabe (2001), we treat $\lambda_t$ as a latent variable.

We explain our estimation method using Model 4. The parameters in Model 4 are $\boldsymbol{\theta} = (\mu^{(0)}, \mu^{(1)}, \phi, \pi_{00}, \pi_{11}, \nu, \omega, \psi, \sigma_\eta^2)$. We sample these parameters from their posterior distribution using MCMC. The Gibbs sampler, which is one of MCMC, enables us to sample from the joint posterior distribution by sampling sequentially from the full conditional posterior distributions. The full conditional posterior distribution of a parameter (or a set of parameters) is the distribution conditional on all other parameters and data. Let $h_t = \log(\sigma_t^2)$ and define $\boldsymbol{y} = (y_1, \ldots, y_T)$, $\boldsymbol{S} = (S_1, \ldots, S_T)$, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_T)$ and $\boldsymbol{h} = (h_1, \ldots, h_T)$ where $T$ is the sample size. Dealing with the latent variables $\boldsymbol{S}$, $\boldsymbol{\lambda}$ and $\boldsymbol{h}$ like parameters, we sample sequentially from the following full conditional densities.

$$f(\boldsymbol{S}|\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{10}$$
$$f(\mu^{(0)}, \mu^{(1)}|\boldsymbol{\theta}_{/(\mu^{(0)},\mu^{(1)})}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{11}$$
$$f(\phi|\boldsymbol{\theta}_{/\phi}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{12}$$
$$f(\pi_{00}, \pi_{11}|\boldsymbol{\theta}_{/\pi_{00},\pi_{11}}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{13}$$
$$f(\boldsymbol{\lambda}|\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{h}, \boldsymbol{y}) \tag{14}$$
$$f(\nu|\boldsymbol{\theta}_{/\nu}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{15}$$
$$f(\boldsymbol{h}|\boldsymbol{\theta}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{y}) \tag{16}$$
$$f(\omega|\boldsymbol{\theta}_{/\omega}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{17}$$
$$f(\psi|\boldsymbol{\theta}_{/\psi}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \tag{18}$$
$$f(\sigma_\eta^2|\boldsymbol{\theta}_{/\sigma_\eta^2}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}), \tag{19}$$

where $\boldsymbol{\theta}_{/x}$ is the set of all parameters except $x$.

In Bayesian estimation, we must set the prior distribution of the parameters. Assuming that the prior distributions of $(\mu^{(0)}, \mu^{(1)})'$, $\phi$, $\sigma_e^2$, $\pi_{00}$, $\pi_{11}$, $\nu$, $\omega$, $\psi$ and $\sigma_\eta^2$ are mutually independent, we set the prior distributions as follows.

$$\begin{aligned}
(\mu^{(0)}, \mu^{(1)})' \sim & N(M_\mu, \Sigma_\mu)\, \boldsymbol{I}\left[\mu^{(0)} < \mu^{(1)}\right], \quad \omega \sim N(m_\omega, v_\omega), \\
\frac{\phi+1}{2} \sim & \mathrm{Beta}(\alpha_\phi, \beta_\phi), \quad \frac{\psi+1}{2} \sim \mathrm{Beta}(\alpha_\psi, \beta_\psi), \\
\pi_{00} \sim & \mathrm{Beta}(u_{00}, u_{01}), \quad \pi_{11} \sim \mathrm{Beta}(u_{11}, u_{10}), \\
\sigma_\eta^2 \sim & \mathrm{IG}(k_\eta, \theta_\eta), \quad \nu \sim \mathrm{Gamma}(k_\nu, \theta_\nu)\boldsymbol{I}[\nu > 2],
\end{aligned}$$

where $\boldsymbol{I}[\,\cdot\,]$ is the indicator function that takes 1 when the inequality in the parenthesis is satisfied and 0 otherwise.

The prior of $(\mu^{(0)}, \mu^{(1)})'$ is the bivariate normal distribution truncated such that $\mu^{(0)} < \mu^{(1)}$ is satisfied. Then, the full conditional distribution (11) is also the truncated normal. The prior of $\omega$ is the normal. Then, (17) is the same. The prior of $\sigma_\eta^2$ is the inverted gamma, so that $\sigma_\eta^2 > 0$. Then, (19) is the same. It is straightforward to sample from these distributions.

For the prior distributions of $\phi$ and $\psi$, we assume that $(\phi + 1)/2$ and $(\psi + 1)/2$ follow the beta distributions for stationarity, i.e., $|\phi| < 1$ and $|\psi| < 1$. The priors of $(\pi_{00}, \pi_{11})'$ are the independent beta, distributions, so that $0 < \pi_{00} < 1$ and $0 < \pi_{11} < 1$. The prior distribution of $\nu$ is the gamma distribution truncated such that $\nu > 2$ is satisfied. The full conditional distributions of $\phi$, $\psi$, $(\pi_{00}, \pi_{11})'$ and $\nu$ distributions are non-standard. We sample $\phi$ and $\psi$ from (12) and (18) using the Metropolis-Hastings (MH) algorithm where the proposal density is selected following Chib and Greenberg (1994) and $\nu$ from (15) using the Acceptance-Rejection Metropolis-Hastings (ARMH) algorithm where the proposal density is selected following Watanabe (2001). We set

$$p(S_i|\pi_{00}, \pi_{11}) = \frac{1 - \pi_{ii}}{2 - \pi_{00} - \pi_{11}} \quad (i = 0, 1).$$

Then, the full conditional density of $(\pi_{00}, \pi_{11})'$ is given as follows.

$$\begin{aligned} f(\pi_{00}, \pi_{11}|\boldsymbol{\theta}_{/(\pi_{00},\pi_{11})}, \boldsymbol{S}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y}) \propto & \frac{(1-\pi_{00})^{S_1}(1-\pi_{11})^{1-S_1}}{2-\pi_{00}-\pi_{11}} \\ & \times \pi_{00}^{u_{00}+n_{00}}(1-\pi_{00})^{u_{01}+n_{01}} \\ & \times \pi_{11}^{u_{11}+n_{11}}(1-\pi_{11})^{u_{10}+n_{10}}, \end{aligned} \tag{20}$$

where $n_{ij}$ refers to the number of transitions from state $i$ to $j$, which can be easily computed for given $\boldsymbol{S}$. If we neglect the term $(1 - \pi_{00})^{S_1}(1 - \pi_{11})^{1-S_1}/(2-\pi_{00}-\pi_{11})$, (20) will collapse to independent Beta distributions. Since $0 < (1-\pi_{00})^{S_1}(1-\pi_{11})^{1-S_1}/(2-\pi_{00}-\pi_{11}) < 1$, we use the Acceptance-Rejection (AR) algorithm to sample from (20), where we sample a proposal from the independent Beta distributions and accept it with probability $(1 - \pi_{00})^{S_1}(1 - \pi_{11})^{1-S_1}/(2 - \pi_{00} - \pi_{11})$.

We must also sample the latent variables $\boldsymbol{S}$, $\boldsymbol{\lambda}$ and $\boldsymbol{h}$ from their full conditional densities (10), (14) and (16) respectively. We sample $\boldsymbol{S}$ from (10) using the multimove sampler proposed by Kim and Nelson (1998) (see also Chapter 9 in Kim and Nelson (1999b)). We use the block sampler proposed by Watanabe and Omori (2004) to sample $\boldsymbol{h}$ from (16). It is straightforward to sample $\boldsymbol{\lambda}$ from (14) because the full conditional distributions of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_T)$ are mutually independent and given as

$$(\epsilon_t^2 + \nu - 2)/\lambda_t \sim \chi^2(\nu + 1), \quad (t = 1, \ldots, T). \tag{21}$$

In Bayesian econometrics, model comparison is conducted using marginal likelihood. A widely used method for calculating marginal likelihood is the modified harmonic mean proposed by Geweke (1999). Let $g(\boldsymbol{\theta})$ be a probability density function. Then, marginal likelihood can be estimated as follows.

$$f(\boldsymbol{y}) = \frac{1}{E\left[\frac{g(\boldsymbol{\theta})}{f(\boldsymbol{y}|\boldsymbol{\theta})f(\boldsymbol{\theta})}\right]} \approx \left[\frac{1}{M}\sum_{i=1}^{M}\frac{g(\boldsymbol{\theta}_i)}{f(\boldsymbol{y}|\boldsymbol{\theta}_i)f(\boldsymbol{\theta}_i)}\right]^{-1} \tag{22}$$

Geweke (1999) proposes to make $g(\boldsymbol{\theta})$ the truncated normal denstity as follows.

$$\begin{aligned} g(\boldsymbol{\theta}) &= \tau^{-1}(2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu})\right] \\ &\quad \times I\left[(\boldsymbol{\theta}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu}) \leq F^{-1}_{\chi^2_k}(\tau)\right] \end{aligned} \tag{23}$$

where $k$ is the number of parameters and $F^{-1}_{\chi^2_k}(\tau)$ is the inverse function of the $\chi^2$ cdf with the degree-of-freedom $k$ and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the sample mean and covariance matrix of $\boldsymbol{\theta}$ sampled from the posterior distribution using MCMC and $I[\,\cdot\,]$is the indicator function that takes 1 if the condition in the bracket is satisfied and 0 otherwise.

Watanabe (2014) and Ishihara and Watanabe (2015) use the complete-data likelihood, i.e., the joint density of the data and latent variables given the parameters instead of the observed-data likelihood, i.e., the density of the data without the latent variables $f(\boldsymbol{y}|\boldsymbol{\theta})$. Nakajima et al. (2011) show that this method is theoretically correct, but Chan and Grant (2015) provide evidence that it has a substantial bias and tends to select the wrong model in practice. For models 1 and 2, we use the observed-data likelihood, which can be calculated using the Hamilton (1989) filter. Once stochatic volatility is introduced, it is not straightforward to calculate the observed-data likelihood. Thus, we use the Chib (1995) method as follows. From the Bayes theorem, marginal likelihood can be represented as follows.

$$\log f(\boldsymbol{y}) = \log f(\boldsymbol{y}|\hat{\boldsymbol{\theta}}) + \log f(\hat{\boldsymbol{\theta}}) - \log f(\hat{\boldsymbol{\theta}}|\boldsymbol{y}) \tag{24}$$

We calculate the observed-data likelihood $f(\boldsymbol{y}|\hat{\boldsymbol{\theta}})$ using Particle filter (Pitt and Shephard (1999)). The log of the posterior density can be calculated using the following equation.

$$\log f(\hat{\boldsymbol{\theta}}|\boldsymbol{y}) = \log f(\hat{\theta}_1|\boldsymbol{y})\sum_{i=2}^{k}\log f(\hat{\theta}_i|\hat{\theta}_1,\ldots,\hat{\theta}_{i-1},\boldsymbol{y}) \tag{25}$$

Following Ishihara and Watanabe (2015), we also analyze structural changes in the mean growth rates $\mu^{(0)}$ and $\mu^{(1)}$ in equation (2). Assuming that only $\mu^{(0)}$ and $\mu^{(1)}$ in equation (2) are subject to structural changes, we give the subscript $t$ to $\mu$, $\mu^{(0)}$ and $\mu^{(1)}$ as follows.

$$\mu_t = \mu_t^{(0)}(1-S_t) + \mu_t^{(1)}S_t, \quad \mu_t^{(0)} < \mu_t^{(1)} \tag{2'}$$

Let $D_t$ denote the number of structural changes up to time$t$ and $n$ denote the total number of structural changes during the sample period. Then, we can define $\mu_t^{(0)}$ and $\mu_t^{(1)}$ as follows.

$$\mu_t^{(0)} = \begin{cases} \mu^{(00)}, & D_t = 0 \\ \mu^{(01)}, & D_t = 1 \\ \vdots & \\ \mu^{(0i)}, & D_t = i \\ \vdots & \\ \mu^{(0,n-1)}, & D_t = n-1 \\ \mu^{(0n)}, & D_t = n \end{cases}, \quad \mu_t^{(1)} = \begin{cases} \mu^{(10)}, & D_t = 0 \\ \mu^{(11)}, & D_t = 1 \\ \vdots & \\ \mu^{(1i)}, & D_t = i \\ \vdots & \\ \mu^{(1,n-1)}, & D_t = n-1 \\ \mu^{(1n)}, & D_t = n \end{cases} \tag{26}$$

Assuming that $D_t$ follows an irreversible Markov process, we express the transition probabilities as follows.

$$\begin{aligned} & p(D_t = 0 \mid D_{t-1} = 0) = q_{00}, \\ & p(D_t = 1 \mid D_{t-1} = 0) = 1 - q_{00}, \\ & p(D_t = 1 \mid D_{t-1} = 1) = q_{11}, \\ & p(D_t = 1 \mid D_{t-1} = 0) = 1 - q_{11}, \\ & \qquad \vdots \\ & p(D_t = n-1 \mid D_{t-1} = n-1) = q_{n-1,n-1}, \\ & p(D_t = n \mid D_{t-1} = n-1) = 1 - q_{n-1,n-1}, \\ & p(D_t = n \mid D_{t-1} = n) = 1 \end{aligned} \tag{27}$$

We choose the total number of structural changes $n$ using marginal likelihood calculated as mentioned above.

# 4 Empirical Analysis

## 4.1 Data

We use the monthly data on the coincident indicator of the composite index (CI) in Japan. ESRI calculates two types of CI with and without outlier replacement. We use CI with outlier replacement. This data are plotted in Figure 1 where the shadow areas are the recession periods published by ESRI. We use the growth rate of CI for $y_t$. We calculate the growth rate of CI as the percentage log difference of CIs in two consecutive periods, which is plotted in Figure 2.

[Insert Figure 1]

[Insert Figure 2]

[Insert Table 1]

The descriptive statistics of the growth rate of CI are summarized in Table 1. The mean is not significantly different from 0. The kurtosis is significantly larger than 3, showing that the growth rate of CI is more leptokurtic than the normal distribution. The skewness is significantly negative, indicating that the growth rate of CI is negatively skewed. We do not, however, take account of the skewness in this paper. The Jarque-Bera (JB) statistic is so large that it rejects the null hypothesis of normality strongly. LB(10) is the Ljung-Box statistics adjusted for heteroskedasticity following Diebold (1988) to test the null hypothesis of no autocorrelations up to 10 lags. According to the values of LB(10), the null hypothesis is rejected at the 1% significance level. The autocorrelation in $y_t$ may fully or partly be explained by the switch in its mean.

## 4.2 Estimation results for the simple MS model

We first estimate the simple MS model with normal error and constant volatility (Model 1). The parameters in Model 1 are $(\mu^{(0)}, \mu^{(1)}, \phi, \sigma^2, \pi_{00}, \pi_{11})$, whose prior distributions are set as follows.

$$(\mu^{(0)}, \mu^{(1)})' \sim N\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}\right) \boldsymbol{I}\left[\mu^{(0)} < \mu^{(1)}\right],$$

$$\frac{\phi + 1}{2} \sim \text{Beta}(1,1), \quad \sigma^2 \sim \text{IG}(6,4),$$

$$\pi_{00} \sim \text{Beta}(9,1), \quad \pi_{11} \sim \text{Beta}(9,1).$$

We throw away the first $5,000$ draws of all parameters as burn-in and use the next $10,000$ draws for the parameter estimation.

To analyze the impact of large shocks, we estimate the model using the subsample 1985/02–2008/08, which is prior to large shocks, as well as the full sample 1985/02–2015/05.

[Insert Figure 3]

Figures 3 depicts the posterior probabilities of recession for each period estimated using the subsample and the full sample respectively. The posterior probabilities of recession for the subsample almost coincide with the recession periods published by ESRI. Those for the full sample are close to 0 except the periods of the global financial crisis in 2008, the Great East Japan Earthquake in 2011, the consumption tax hikes in 2014 and 2019, and the COVID-19 pandemic in 2020 when they are close to 1. The negative impact of these events is so large that the simple MS model leads to the result that only those five periods are in the recession regime and all other periods are in the boom regime.

The estimation result of each parameter in Model 1 is summarized in Table 2. The mean and standard deviation (SD) are calculated as the sample mean and standard deviation of the $10,000$ draws of each parameter after burn-in. The standard error (SE) of sample mean is calculated using the Parzen window because the draws sampled using MCMC are autocorrelated. The 95% Bayesian credible interval is obtained as the 2.5th and 97.5th percentiles of the $10,000$ draws of each parameter. CD is the convergence diagnostic statistic proposed by Geweke (1992), whose asymptotic distribution is the standard normal if the draws have converged to the ones from the posterior distribution. The standard error of CD statistic is also calculated using the Parzen window. IF is the inefficiency factor proposed by Chib (2001). If this is equal to 2, it implies that the number of draws must be twice as much as that of random sampling to make the both standard errors the same. The inefficiency factor increases with the autocorrelation.

[Insert Table 2]

According to the CD values, the null hypothesis that the $10,000$ draws used for estimation have converged to the ones from the posterior distribution is accepted for all parameters in the both subsample and full sample. The IF values for all parameters are small, indicating that the sampling method is efficient. The mean and 95% interval of $\mu^{(0)}$ in the full sample are much smaller than those in the subsample, showing that the negative impact of the financial crisis and the Tohoku earthquake on the growth rate of CI is large. The mean and 95% interval of $\phi$ are negative in the subsample while they are positive in the full sample.

### 4.3 Estimation results for the extended MS model

Next, we estimate the extended models. The new parameter in the MS model with $t$-error and constant volatility (Model 2) is $\nu$, which is the degree-of-freedom of the Student's $t$-distribution. Its prior is set as follows.

$$\nu \sim \text{Gamma}(1, 0.1)\boldsymbol{I}[\nu > 2].$$

The priors for all other parameters are set as the same as those in Model 1. The new parameters in the MS model with normal error and SV (Model 3) are $(\omega, \phi, \sigma_\eta^2)$. Their priors are set as follows.

$$\omega \sim N(0, 10), \quad \frac{\psi + 1}{2} \sim \text{Beta}(2, 1), \quad \sigma_\eta^2 \sim \text{IG}(6, 4).$$

The priors for all other parameters are set as the same as those in Model 1. The priors of the parameters in the MS model with $t$-error and SV (Model 4) are set as the same as before. We throw away the first $10,000$ draws of all parameters as burn-in and use the next $10,000$ draws for the parameter estimation.

Tables 3–5 summarize the estimation results for the extended MS models using the full sample. According to the CD values, the null hypothesis that the $10,000$ draws used for estimation have converged to the ones from the posterior distribution is accepted at the 5% significance level for all parameters in all models. The IF values for all parameters are larger than those in Model 1, but they are still not so large indicating that the sampling method we use is efficient. The mean and 95% interval of $\mu^{(0)}$ and $\pi_{00}$ are small in Model 1 for the full sample, but in all extended models, they recover close to those in Model 1 for the subsample prior to large shocks. In all extended models, the mean of $\phi$ is negative. The mean of $\nu$ in Model 4 is 30.2283 while that in Model 2 is 2.8527. If volatility changes, the distribution of $e_t = \sigma_t \epsilon_t$ becomes leptokurtic even if the distribution of $\epsilon_t$ is normal. The leptokurtosis may be captured by changes in volatility. This is the reason why the mean of $\nu$ increases once SV is introduced. The means of $\psi$ are 0.8640 and 0.8540 in Models 3 and 4 respectively, indicating that shocks to volatility are persistent although the persistence is small compared with financial volatility.

[Insert Table 3]

[Insert Table 4]

[Insert Table 5]

For model comparison, we calculate the log marginal likelihoods for all models using the methods explained in Section 3. Table 6 shows the result. We do not report their standard errors because they are close to 0. The log marginal likelihood in Model 3 is significantly larger than those of other models. The conclusion must be that the MS model with normal error and SV (Model 3) fits the data best.

[Insert Table 6]

[Insert Figure 4]

[Insert Figure 5]

[Insert Figure 6]

Figure 4–6 depicts the posterior probabilities of recession for each period estimated by Models 2–4. All models provide the similar posterior probabilities. As is shown in Figure 3, the simple MS model with normal error and constant volatility predicts that the periods of large shocks are in the recession regime, which does not hold true once $t$-error or SV is introduced. Using the posterior probabilities of recession, we estimate the business cycle turning points as follows. Let $\hat{p}(S_t = 0|\boldsymbol{y})$ be the estimated probability of recession for period $t$. We define $t$ as peak if $\hat{p}(S_{t-1} = 0|\boldsymbol{y}) < 0.5$ and $\hat{p}(S_t = 0|\boldsymbol{y}) > 0.5$ and as trough if $\hat{p}(S_{t-1} = 0|\boldsymbol{y}) > 0.5$ and $\hat{p}(S_t = 0|\boldsymbol{y}) < 0.5$. Table 7 shows the result. The turning points estimated by Models 2–4 are close to those by ESRI except that Model 4 estimates 2015/08 as peak and 2016/03 as trough.

[Insert Table 7]

[Insert Figure 7]

Figure 7 plots the posterior mean of volatility $\sigma_t^2$ estimated using Model 3 (solid line) and Model 4 (dotted line). During the periods of large shocks, the posterior means of volatility estimated by the both models jumps up. The difference in the posterior mean of volatility between Models 3 and 4 is negligible because the degree of freedom $\nu$ of the $t$-distribution in Model 4 is so large that it is no different from using the standard normal distribution.

Since Model 3 (normal error and SV) fits the data best, using this model, we analyze structural changes in in the mean growth rates $\mu^{(0)}$ and $\mu^{(1)}$ in Table 8 shows the log marginal likelihood for each number of break points between 0 and 4. According to the table, the log marginal likelihood of the model with two break points is the highest. Hereinafter, the model with two structural change points added to model 4 will be called model 5.

[Insert Table 8]

[Insert Figure 8]

Figure 8 shows the posterior distribution of each structural change point estimated from Model 5 with two structural change points. These are the 10,000 samples after the burn-in of $\tilde{D}_T$, and since there are two $t$s where $D_{t-1} < D_t$, these were extracted and their respective histograms were drawn. The peak of the posterior distribution of the first structural change point is in 2008/10 after the Lehman Shock, and the peak of the posterior distribution of the second structural change point is in 2010/02.

Table 9 shows the parameter estimation results for model 5 with two structural change points. Comparing the posterior means of $\mu^{(00)}$ and $\mu^{(01)}$ and those of $\mu^{(10)}$ and $\mu^{(11)}$, we can see that at the first structural change point, the mean growth rate during recessions declines and that during booms rises. Comparing the posterior means of $\mu^{(01)}$ and $\mu^{(02)}$ and those of $\mu^{(11)}$ and $\mu^{(12)}$, we can see that at the second structural change point, the mean growth rate during recessions rises and that during expansions falls. These results show that this structural change captures the instability of economic fluctuations during financial crises and the subsequent recovery periods. $\mu^{(01)}$ and $\mu^{(02)}$ have large standard errors and standard deviations, but this is thought to be because the period of the economic downturn between the first and second structural turning points and the period of the economic downturn after the second structural turning point are short, and the number of samples that can be used to estimate $\mu^{(01)}$ and $\mu^{(02)}$ is small.

[Insert Table 9]

[Insert Figure 10]

Table 9 shows the business cycle turning points estimated by Model 5 (normal error error + stochastic volatility + 2 change points). This model produces the estimates of peak and trough similar to those published by ESRI and estimated by other models except that this model estimates 2015/02 as peak and 2016/06 as trough.

## 5 Conclusion

We analyze the business cycles in Japan by applying MS models to the growth rate of the coincident indicator of CI during the period of 1985/01–2025/05 calculated by ESRI. We first show that the impact of the global financial crisis in 2008, the Great East Japan Earthquake in 2011, the consumption tax hikes in 2014 and 2019, and the COVID-19 pandemic in 2020 on this index is so large that the simple MS model with the normal error and constant volatility cannot detect the business cycle turning points properly. We extend the MS model by incorporating $t$-error and SV and employ a Bayesian method via MCMC for the analysis of the extended models. We show that the MS model with $t$-error or SV provides the estimates of the business cycle turning points close to those published by ESRI. The marginal likelihoods provide evidence that the MS model with normal error and SV fits the data best.

We did not take account of skewness in this article although the growth rate of CI is negatively slewed (see Table 1). We should also use skewed distributions such as skew $t$-distributions (see Aas (2005), Azzalini and Capitano (2003) and Fernández and Steel (1998)).

# References

Aas, K. (2005), "The Generalized Hyperbolic Skew Student's $t$-distribution," *Journal of Financial Econometrics*, 4(2), pp.275–309.

Azzalini, A. and Capitano, A. (2003), "Distributions Geberated by Perturbation of Symmetry with Emphasis on a Multivariate Skew $t-$distribution," *Journal of the Royal Statistical Society*, Series B, 65(2), 367–389.

Chan, J. C. C. and Grant, A. L. (2015), "Pitfalls of estimating the marginal likelihood using the modified harmonic mean," *Economics Letters*, 131(1), 29–33.

Chib, S. (1995), "Marginal likelihood from the Gibbs output," *Journal of the American Statistical Association*, 90(432), 1313–1321.

Chib, S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference," in J. J. Hechman and E. Leeper (eds.), *Handbook of Econometrics*, Vol.5, Chapter 57, Elsevier, pp.3569–3649.

Diebold, F. (1988), *Empirical Modeling of Exchange Rate Dynamics*, Springer-Verlag.

Fernández, C. and Steel, M. F. J. (1998), "On Bayeasian Modeling of Fat Tail and Skewness," *Journal of the American Statistical Association*, Vol.93, No.441, pp.359–371.

Geweke, J. (1992), "Evaluating the Accuracy of Sampling-based Approaches to the Calculation of Posterior moments," (with discussion), in J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith (eds.), *Bayesian Statistics* 4, Oxford University Press, pp.169–191.

Geweke, J. (1999), "Using simulation methods for Bayesian econometric models: inference, developments, and communications," *Econometric Reviews*, 18(1), 1–127.

Hamilton, J. D. (1989), "A new approach to the economic analysis of nonstationary time series and business cycle," *Econometrica*, 57(2), 357–384.

Ishihara, T. and Watanabe, T. (2015), "Econometric analysis of business cycles: A survey with the application to the composite index in Japan," (in Japanese), *Economic Review*, Institute of Economic Research, Hitotsubashi University, 66(2), 145–168.

Kim, C.-J. and Nelson, C. R. (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(2), 188–201.

Kim, C.-J. and Nelson, C. R. (1999a), "Has the U.S. economy become more stable? A Bayesian approach based on a Markov-Switching model of the business cycle," *Review of Economics and Statistics*, 81(2), 608–616.

Kim, C.-J. and Nelson, C. R. (1999b), *State-Space Models with Regime Switching*, MIT Press.

Nakajima, J., Kasuya, M. and Watanabe, T. (2011), "Bayesian analysis of time-varying parameter vector autoregressive model for the Japanese economy and monetary policy," *Journal of the Japanese and International Economies*, 25(3), .225–245.

Pitt, M. K. and Shephard, N. (1999), "Filtering via simulation: Auxiliary particle filters," *Journal of the American Statistical Association*, 94(446), 590–599

Watanabe, T. (2001), "On sampling the degree-of-freedom of student's $t$–disturbances," *Statistics and Probability Letters*, 52(2), 177–181.

Watanabe, T. (2014), "Bayesian analysis of business cycle in Japan using Markov switching model with stochastic volatility and fat-tail distribution," *Economic Review*, Institute of Economic Research, Hitotsubashi University, 65(2), 156–167.

Watanabe, T. and Omori, Y. (2004), "A multi-move sampler for estimating non-Gaussian time series model: Comments on Shephard & Pitt (1997)," *Biometrika*, 91(1), .246–248.

Table 1: Descriptive statistics of the growth rate of CI

| Mean | SD | Skewness | Kurtosis | JB | LB(10) |
|---|---|---|---|---|---|
| 0.0511 | 1.4889 | -2.2838 | 16.9305 | 4334.24 | 22.38 |
| (0.0677) | | (0.1113) | (0.2227) | | |

The sample period is 1985/2–2025/05 and the sample size is 484. The numbers in parentheses are standard errors. SD is the standard deviation. JB is the Jarque-Bera statistic to test the null hypothesis of normality. LB(10) is the Ljung-Box statistics adjusted for heteroskedasticity following Diebold (1988) to test the null hypothesis of no autocorrelations up to 10 lags.

Table 2: Estimation result for Model 1 (normal error and constant volatility)

Subsample (1985/02–2008/08)

| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(0)}$ | -0.7481 | 0.0020 | 0.1132 | [-0.9841, -0.5381] | -0.73 | 2.05 |
| $\mu^{(1)}$ | 0.3830 | 0.0009 | 0.0550 | [0.2768, 0.4928] | -0.63 | 2.55 |
| $\phi$ | -0.2941 | 0.0007 | 0.0610 | [-0.4109, -0.1714] | 1.71 | 1.55 |
| $\sigma^2$ | 0.6954 | 0.0007 | 0.0617 | [0.5821, 0.8267] | -1.70 | 1.46 |
| $\pi_{00}$ | 0.9323 | 0.0005 | 0.0335 | [0.8517, 0.9806] | -1.36 | 1.12 |
| $\pi_{11}$ | 0.9698 | 0.0003 | 0.0139 | [0.9368, 0.9903] | 0.45 | 3.67 |

Full Sample (1985/02–2025/05)

| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(0)}$ | -6.7753 | 0.0070 | 0.4863 | [-7.7374, -5.8170] | -0.56 | 2.28 |
| $\mu^{(1)}$ | 0.1741 | 0.0004 | 0.0605 | [0.0535, 0.2931] | -0.20 | 1.11 |
| $\phi$ | 0.1257 | 0.0004 | 0.0472 | [0.0318, 0.2165] | -0.52 | 1.77 |
| $\sigma^2$ | 1.3114 | 0.0010 | 0.0846 | [1.1554, 1.4870] | 1.82 | 1.30 |
| $\pi_{00}$ | 0.7337 | 0.0009 | 0.1024 | [0.5147, 0.9064] | 0.54 | 1.46 |
| $\pi_{11}$ | 0.9900 | 0.0000 | 0.0046 | [0.9793, 0.9969] | -0.43 | 1.63 |

SE is the standard error of mean and SD is the standard deviation. CD is the convergence diagnostic statistics proposed by Geweke (1992). IF is the inefficiency factor proposed by Chib (2001).

Table 3: Estimation result for Model 2 ($t$-error and constant volatility)

| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(0)}$ | -0.5870 | 0.0061 | 0.1536 | [-0.9150, -0.3114] | -0.79 | 14.60 |
| $\mu^{(1)}$ | 0.4486 | 0.0028 | 0.0740 | [0.3108, 0.6022] | -0.65 | 12.54 |
| $\phi$ | -0.0526 | 0.0013 | 0.0523 | [-0.1546, 0.0501] | -0.79 | 7.55 |
| $\sigma^2$ | 1.9650 | 0.0025 | 0.1467 | [1.6980, 2.2728] | -1.24 | 2.09 |
| $\pi_{00}$ | 0.9192 | 0.0005 | 0.0296 | [0.8483, 0.9653] | 0.11 | 3.84 |
| $\pi_{11}$ | 0.9652 | 0.0004 | 0.0154 | [0.9288, 0.9881] | 1.16 | 8.25 |
| $\nu$ | 2.8527 | 0.0103 | 0.1558 | [2.6013, 3.2064] | 0.28 | 34.16 |

SE is the standard error of mean and SD is the standard deviation. CD is the convergence diagnostic statistics proposed by Geweke (1992). IF is the inefficiency factor proposed by Chib (2001).

Table 4: Estimation result for Model 3 (normal error and stochastic volatility)

| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(0)}$ | -0.4597 | 0.0080 | 0.1884 | [-0.8770, -0.1455] | -0.29 | 31.75 |
| $\mu^{(1)}$ | 0.3892 | 0.0022 | 0.0646 | [0.2710, 0.5219] | -0.34 | 20.52 |
| $\phi$ | -0.1368 | 0.0011 | 0.0584 | [-0.2520, -0.0240] | 0.79 | 7.34 |
| $\pi_{00}$ | 0.9175 | 0.0006 | 0.0325 | [0.8399, 0.9670] | 0.67 | 5.86 |
| $\pi_{11}$ | 0.9617 | 0.0007 | 0.0190 | [0.9158, 0.9892] | 1.46 | 13.93 |
| $\omega$ | -0.0113 | 0.0020 | 0.2183 | [-0.4472, 0.4223] | 0.28 | 1.11 |
| $\psi$ | 0.8640 | 0.0017 | 0.0394 | [0.7791, 0.9334] | -0.51 | 25.13 |
| $\sigma_\eta^2$ | 0.3640 | 0.0057 | 0.0930 | [0.2111, 0.5832] | 0.0082 | 54.69 |

SE is the standard error of mean and SD is the standard deviation. CD is the convergence diagnostic statistics proposed by Geweke (1992). IF is the inefficiency factor proposed by Chib (2001).

Table 5: Estimation result for Model 4 ($t$-error and stochastic volatility)

| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(0)}$ | -0.4353 | 0.0082 | 0.1512 | [-0.7875, -0.1837] | 0.60 | 32.71 |
| $\mu^{(1)}$ | 0.4026 | 0.0027 | 0.0614 | [0.2883, 0.5314] | -0.01 | 23.26 |
| $\phi$ | -0.1396 | 0.0014 | 0.0571 | [-0.2486, -0.0256] | -1.52 | 5.84 |
| $\pi_{00}$ | 0.9177 | 0.0005 | 0.0305 | [0.8466, 0.9654] | 0.99 | 4.99 |
| $\pi_{11}$ | 0.9597 | 0.0007 | 0.0187 | [0.9147, 0.9872] | -0.20 | 16.45 |
| $\omega$ | -0.1821 | 0.0069 | 0.2254 | [-0.6232, 0.2723] | -1.06 | 6.42 |
| $\psi$ | 0.8540 | 0.0027 | 0.0443 | [0.7522, 0.9274] | -0.76 | 33.22 |
| $\sigma_\eta^2$ | 0.3737 | 0.0082 | 0.0996 | [0.2220, 0.6025] | 0.74 | 55.69 |
| $\nu$ | 30.2283 | 2.7244 | 18.7448 | [7.3251, 80.6351] | 0.39 | 116.53 |

SE is the standard error of mean and SD is the standard deviation. CD is the convergence diagnostic statistics proposed by Geweke (1992). IF is the inefficiency factor proposed by Chib (2001).

Table 6: Log marginal likelihood

| Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|
| -798.85 | -781.33 | -681.47 | -702.37 |

Marginal likelihoods for Models 1 and 2 arecalculated using the modified harmonic mean method proposed by Geweke (1999) and those for Models 3 and 4 are calculated following Chib (1995). We do not report standard errors of log marginal likelihood because they are close to zero.

Table 7: Business cycle turning points: Models 2–4

| ESRI | Model 2 | Model 3 | Model 4 |
|---|---|---|---|
| Peak | | | |
| 85/06 | 85/10 | 85/11 | 85/08 |
| 91/02 | 90/11 | 90/11 | 90/11 |
| 97/05 | 97/07 | 97/07 | 97/07 |
| 00/10 | 01/01 | 01/01 | 01/01 |
| 08/02 | 07/09 | 07/09 | 07/07 |
| 12/03 | 12/04 | 12/04 | 12/04 |
| – | – | – | 15/08 |
| 18/10 | 18/06 | 18/06 | 18/05 |
| Trough | | | |
| 86/11 | 86/11 | 86/09 | 86/09 |
| 93/10 | 94/01 | 94/01 | 94/01 |
| 99/01 | 98/11 | 98/11 | 98/12 |
| 02/01 | 02/01 | 02/01 | 02/01 |
| 09/03 | 09/04 | 09/03 | 09/03 |
| 12/11 | 12/11 | 12/11 | 12/12 |
| – | – | – | 16/03 |
| 20/05 | 20/06 | 20/05 | 20/05 |

Table 8: Number of break points and log marginal likelihood

| 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| -681.47 | -677.41 | -676.59 | -699.73 | -706.933 |

Marginal likelihoods are calculated following Chib (1995). We do not report standard errors of log marginal likelihood because they are close to zero.

Table 9: Estimation result for Model 5 (normal error error + stochastic volatility + 2 change points)

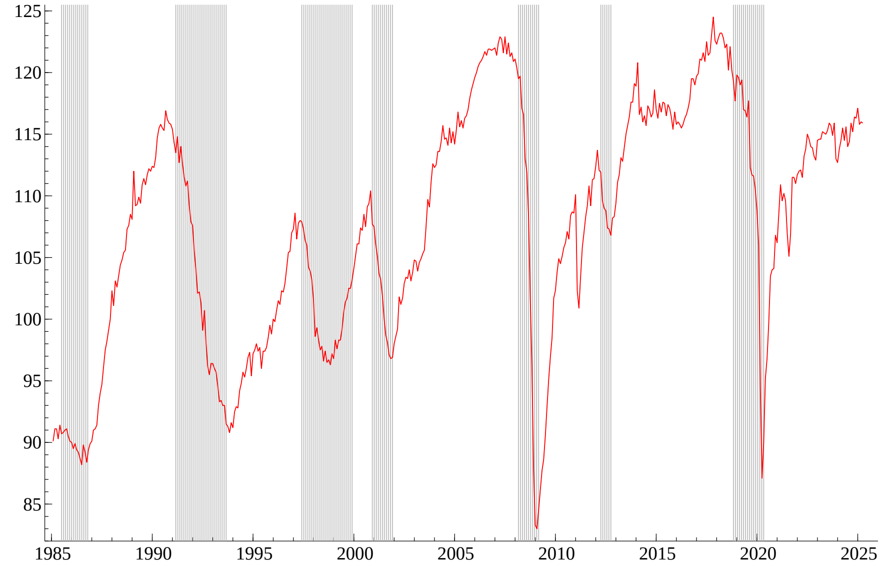
| | Mean | SE | SD | 95% Interval | CD | IF |
|---|---|---|---|---|---|---|
| $\mu^{(00)}$ | -0.5372 | 0.0091 | 0.1516 | [-0.8781, -0.2819] | -1.83 | 32.11 |
| $\mu^{(01)}$ | -5.5976 | 0.0338 | 1.2258 | [-7.3781, -2.4430] | 0.64 | 2.94 |
| $\mu^{(02)}$ | -0.2869 | 0.0359 | 0.9460 | [-2.7238, 0.1630] | -0.66 | 3.46 |
| $\mu^{(10)}$ | 0.4098 | 0.0021 | 0.0574 | [0.2976, 0.5236] | 0.62 | 9.04 |
| $\mu^{(11)}$ | 1.8311 | 0.0146 | 0.4554 | [1.2562, 2.3585] | 0.86 | 7.90 |
| $\mu^{(12)}$ | 0.4528 | 0.0181 | 0.2381 | [0.1287, 0.8646] | -0.60 | 61.55 |
| $\phi$ | -0.2050 | 0.0013 | 0.0575 | [-0.3168, -0.0912] | -0.024 | 5.79 |
| $p_{00}$ | 0.9366 | 0.0009 | 0.0284 | [0.8697, 0.9788] | -1.73 | 13.77 |
| $p_{11}$ | 0.9622 | 0.0005 | 0.0168 | [0.9232, 0.9882] | -1.34 | 2.75 |
| $q_{00}$ | 0.9962 | 0.0000 | 0.0036 | [0.9864, 0.9999] | -0.48 | 1.40 |
| $q_{11}$ | 0.9575 | 0.0006 | 0.0395 | [0.8528, 0.9985] | 1.17 | 1.15 |
| $\omega$ | -0.3102 | 0.0030 | 0.1793 | [-0.6625, 0.0449] | -0.89 | 3.42 |
| $\psi$ | 0.8007 | 0.0035 | 0.0575 | [0.6722, 0.8966] | -0.89 | 46.15 |
| $\sigma_\eta^2$ | 0.4173 | 0.0088 | 0.1211 | [0.2270, 0.6983] | 0.98 | 57.44 |

SE is the standard error of mean and SD is the standard deviation. CD is the convergence diagnostic statistics proposed by Geweke (1992). IF is the inefficiency factor proposed by Chib (2001).

Table 10: Business cycle turning points: Model 5 (normal error error + stochastic volatility + 2 change points)

| ESRI | Model 5 |
|---|---|
| Peak | |
| 85/06 | 85/11 |
| 91/02 | 90/11 |
| 97/05 | 97/07 |
| 00/10 | 01/01 |
| 08/02 | 07/11 |
| 12/03 | 12/04 |
| – | 15/02 |
| 18/10 | 18/02 |
| Trough | |
| 86/11 | 86/09 |
| 93/10 | 94/01 |
| 99/01 | 98/11 |
| 02/01 | 02/01 |
| 09/03 | 09/03 |
| 12/11 | 12/10 |
| – | 16/06 |
| 20/05 | 20/06 |

Figure 1: CI




The coincident indicator of composite index published by Economic and Social Research Institute (ESRI), Cabinet Office, the Government of Japan. Shadow areas are the recession periods published by ESRI.

Figure 2: Growth rate of CI (%)

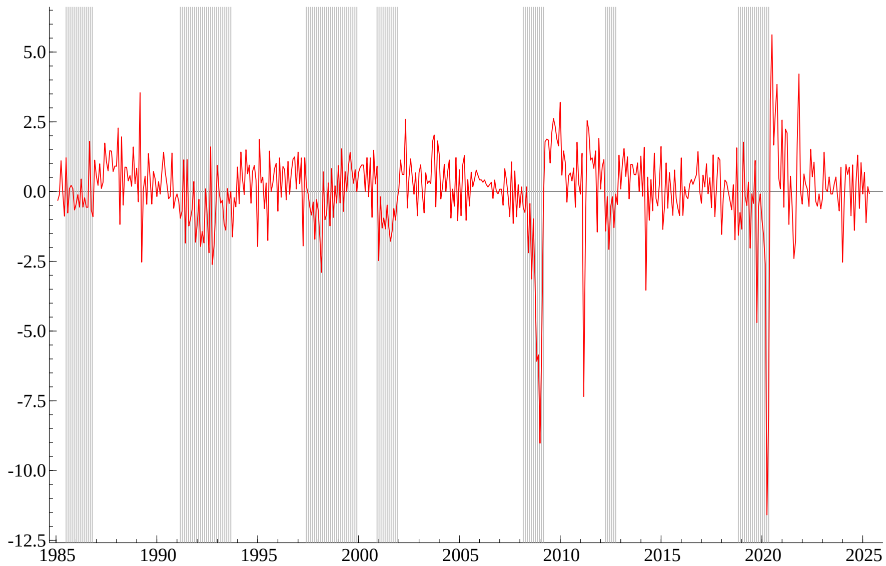
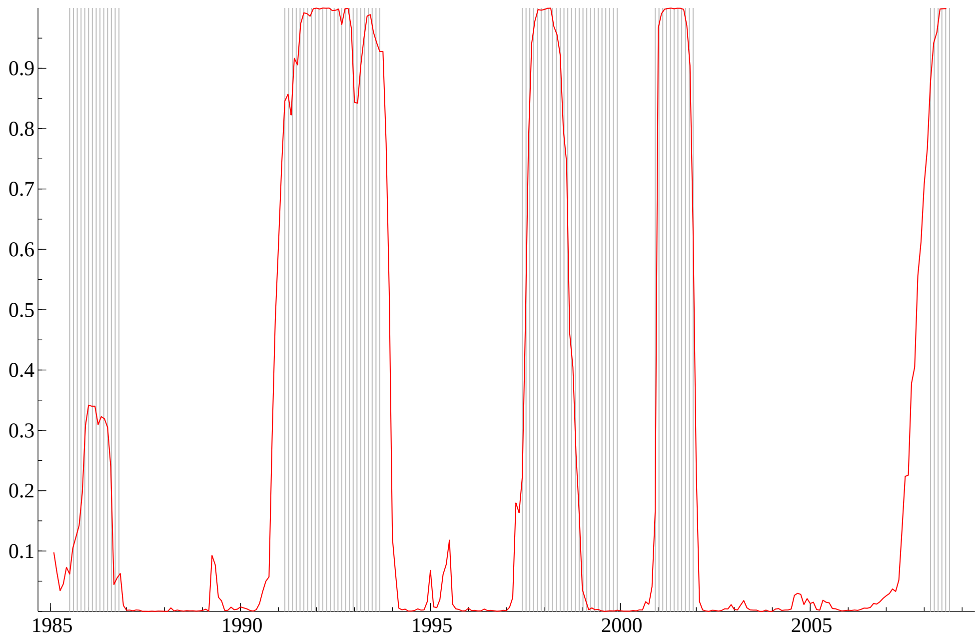


Figure 3: Posterior probabilities of recession: Model 1 (normal error and constant volatility)

Subsample (1985/02–2008/08)




Full Sample (1985/02–2025/05)

Figure 4: Posterior probabilities of recession: Model 2 ($t$-error and constant volatility)

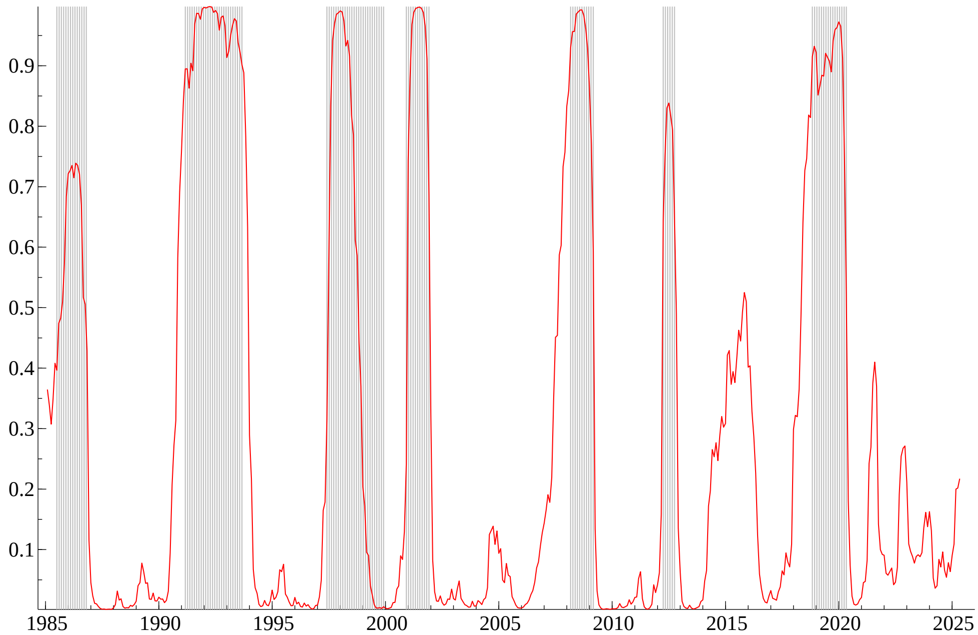


Figure 5: Posterior probabilities of recession: Model 3 (normal error and stochastic volatility)

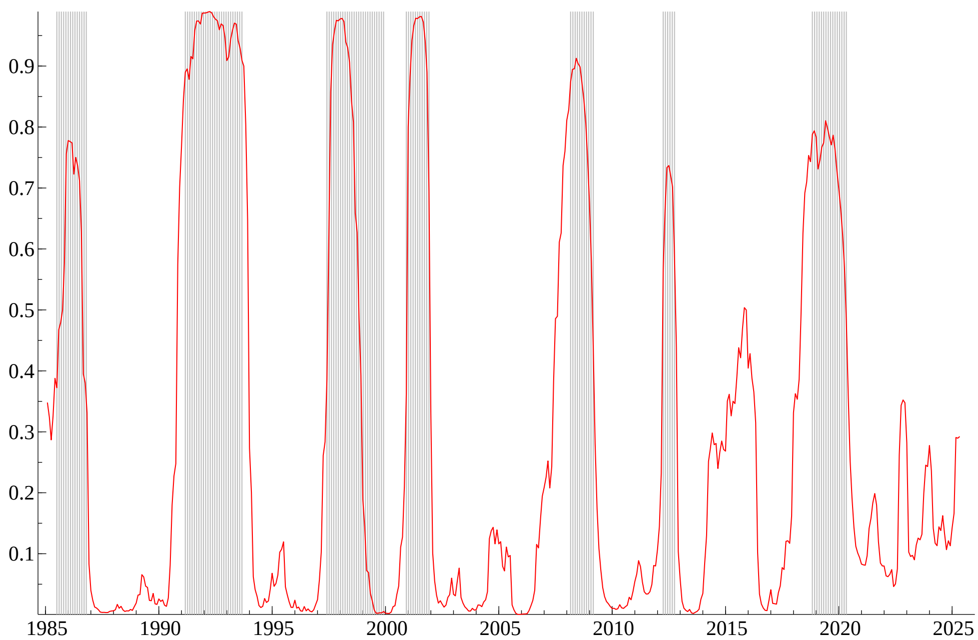


Figure 6: Posterior probabilities of recession: Model 4 ($t$-error and stochastic volatility)

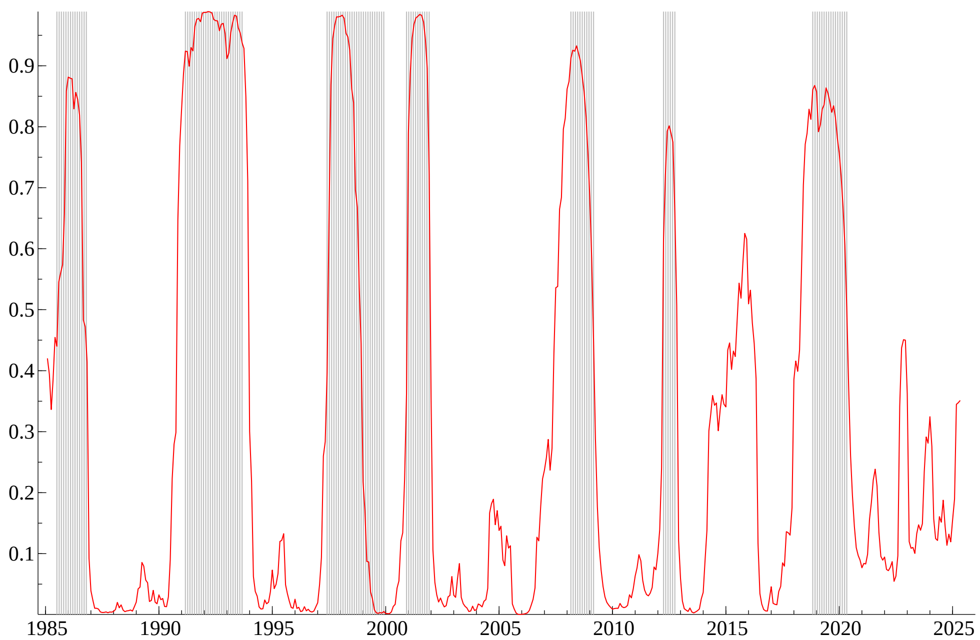
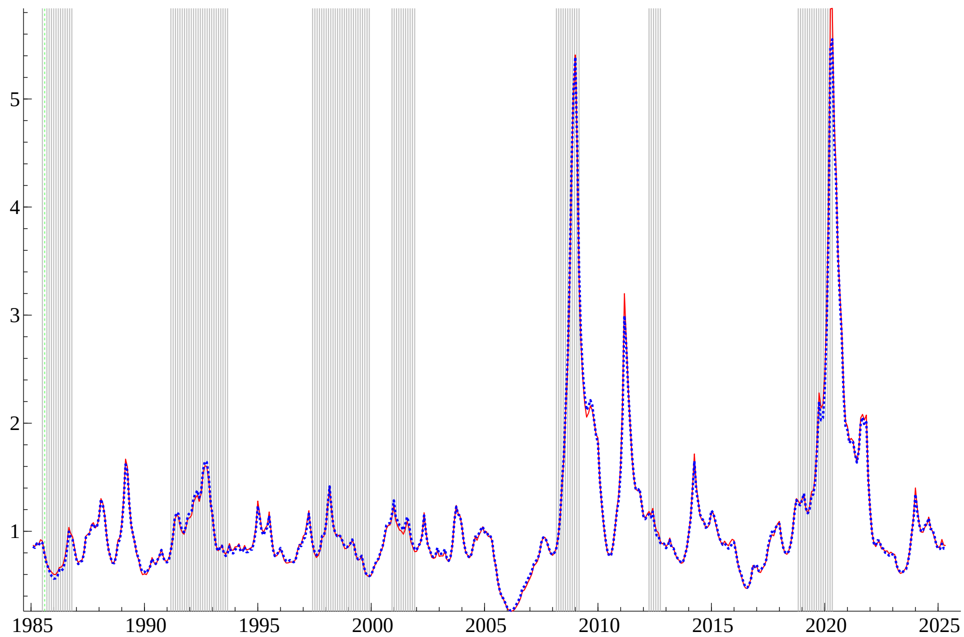
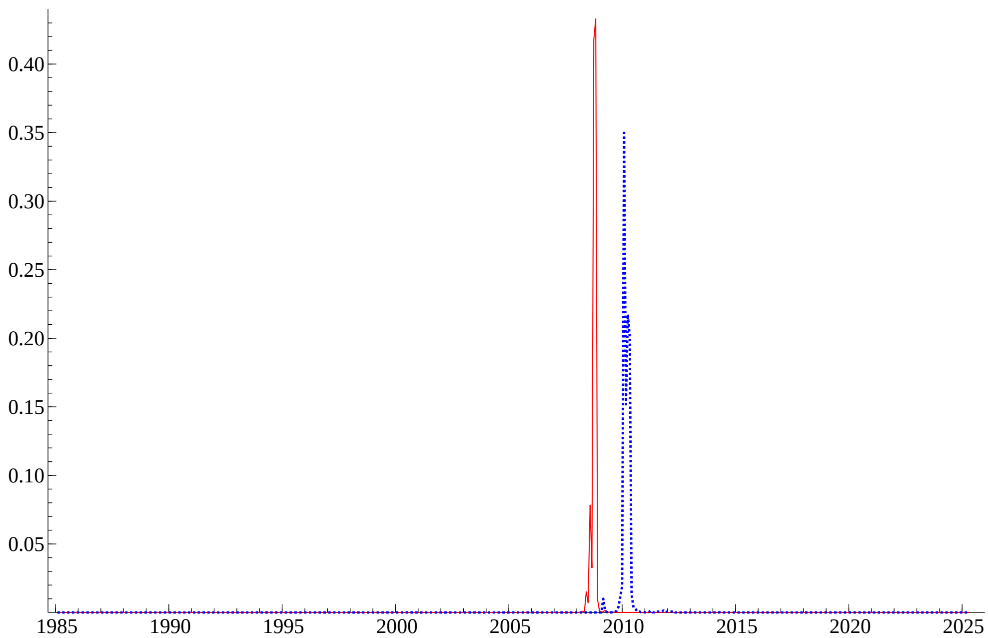


Figure 7: Posterior mean of volatility: Models 3 (normal error and stochastic volatility) and Model 4 ($t$-error and stochastic volatility)




The solid line is the posterior mean of volatility $\sigma_t^2$ from Model 3, and the dotted line is that from Model 4.

Figure 8: Posterior probabilities of change point: Model 5 ($t$-error+stochastic volatility+2 change points)




The solid line is the first change point and the dotted line is the second change point.

Figure 9: Posterior probabilities of recession: Model 5 ($t$-error + stochastic volatility+2 change points)