THE UNIVERSITY OF SYDNEY

Economics Working Paper Series

2012 - 5

Bayesian Inference about the Types of Structural Breaks When There are Many Breaks

Yunjong Eo

February 2012

Bayesian Inference about the Types of Structural Breaks When There are Many Breaks

Yunjong Eo *

School of Economics
University of Sydney

This Draft: February 2012

Abstract

I propose a Bayesian approach to making an inference about complicated patterns of structural breaks in time series. Structural break models in the literature are mainly considered for a simple case in which all the parameters under the structural changes are restricted to have breaks at the same dates. Unlike the existing literature, the proposed method in this paper allows multiple parameters such as intercept, persistence, and/or residual variance to undergo mutually independent structural breaks at different dates with the different number of breaks across parameters. To estimate the complex structural break models considered in this paper, structural breaks in the multiple parameters are interpreted as regime transitions as in Chib (1998). The regime for each parameter is then indicated by a corresponding discrete latent variable which follows a first-order Markov process. A Markov-chain Monte Carlo scheme is developed to estimate and compare the complex structural break models, which are potentially non-nested, in an efficient and tractable way. I apply this approach to postwar U.S. inflation and find strong support for an autoregressive model with two structural breaks in residual variance and no break in intercept and persistence.

JEL classification: C11, C52
Keywords: Bayesian Analysis; Structural Breaks; Multiple-Parameter Change-point; Inflation Dynamics;

*Email: yunjong.eo@sydney.edu.au. Mailing address: School of Economics, University of Sydney, Merewether Bldg. (H04), NSW 2006, Australia.

1 Introduction

1.1 Motivations and main ideas

In this paper I consider how to make an inference about complicated patterns of structural breaks in time series data. Parameters undergo structural changes but some parameters could have the different number of structural breaks at different dates than others. Many macroeconomic variables in the postwar U.S. economy appear to display this pattern of structural instability. For example, Kim and Nelson (1999) and McConnell and Perez-Quiros (2000) find that there was a volatility reduction in 1984, the so-called "Great Moderation", while Perron (1989) and Zivot and Andrews (1992) find that there was a productivity growth slowdown in 1972 in the unit root literature. Hansen (2001) also examines U.S. labor productivity data and finds significant evidence for multiple structural changes in the mid-1980s and in the mid-1990s. Another interesting application is Rapach and Wohar (2005), who find substantial evidence for multiple structural breaks in the mean real interest rates for 13 industrialized countries. In addition, Stock and Watson (1996, 2002) show that most of U.S. macroeconomic data are unstable in conditional mean and residual variance of autoregressive regressions.

However, the nature and timing of structural changes are *a priori* unknown. Econometricians need to make inferences about the number and timing of the structural breaks and identify the parameters under the structural changes. Thus, it is necessary to develop an econometric framework for estimating the complex patterns of structural break models with a model selection procedure. In addition, the procedure should be able to explore all the possible patterns of structural breaks efficiently while the models are potentially non-nested (e.g. a model with one break in residual variance vs. a model with two breaks in intercept).

Thus, I propose an efficient Bayesian Markov-chain Monte Carlo (MCMC) method that allows for a number of possibilities for the nature of structural breaks. This new approach is developed to have the following distinctive features: (i) model specification of considering multiple structural changes in multiple parameters; (ii) model flexibility in allowing the multiple structural breaks to occur mutually independently at different dates across different parameters; and (iii) model selection procedure by comparing various potentially non-nested structural break models. Especially, taking the Bayesian approach makes it feasible in practice to estimate the structural break models with the feature (ii) via Gibbs sampling. In the complicated structural break models, the joint distribution of parameters of interest such as change-points and model coefficients is not known explicitly. One also encounters difficulty in considering all the possible patterns of structural breaks for inferences since the model space increases enormously with the number of structural breaks and the model parameters related to them. However, the conditional distribution for each parameter is available and the proposed approach via Gibbs sampling enables one to easily estimate the structural break models by sampling the posterior distribution from it. In addition, the feature (iii) is attained by comparing potentially non-nested structural break models based on calculations of marginal densities of data under the respective models. It is also straightforward to determine the posterior probability that each model is the true model given the observed data among all the possible structural break models from the calculations.

In more details I explain the proposed MCMC scheme for the complicated patterns of structural break models. This method extends Chib's (1998) approach in which structural breaks are interpreted as regime transitions. He introduces a latent discrete regime variable which indicates one of all the possible regimes over time. A first-order Markov process then governs the structural changes with transition probabilities constrained so that the regime indicator variable can either stay in the current regime or move to the next regime. Chib (1998) assumes that all the parameters subject to the changes are restricted to have the structural changes at the same dates. In consequence his approach requires only *one* regime indicator variable which governs the structural breaks in *all* the parameters under the changes.

Recall that unlike in Chib (1998) the proposed approach allows multiple parameters to have structural breaks mutually independently at different dates across different parameters. This more flexible assumption requires the specification that each parameter is augmented with its corresponding independent regime indicator variable. That is, each parameter is subject to an independent structural change framework in terms of the number and timing of structural breaks. The total number of the parameters subject to the structural changes is then the same as that of the regime indicator variables which are independent of one another. Note that it is possible to sort several parameters into a group and make the parameters to have structural changes at the same dates. This specification would be a simple extension of the proposed approach in this paper.

Based on this specification, a MCMC sampler can be presented through a hierarchical specification in which one draws the model parameters conditional on the regime variables and the observed data; the regime variables conditional on the model parameters and the observed data; and finally the transition probabilities conditional on the regime variables via Gibbs sampling.

I then apply the Bayesian approach proposed in this paper to an artificial data set and postwar U.S. inflation based on the GDP deflator. This new approach identifies accurately break dates from the artificial data in the sense that estimated regime changes occur around true break dates in the data generating process. In terms of the empirical application to U.S. GDP deflator inflation dynamics, I run various autoregressive regressions with structural breaks in intercept, persistence, and/or residual variance. I find that there are two structural breaks in only the residual variance parameter and no break in the intercept and persistence parameters. The residual variance switches from the low volatility regime to the high volatility regime in the early 1970s and then returns to another low volatility regime in the mid-1980s.

1.2 Recent related literature

In classical econometrics, Andrews (1993) and Andrews and Ploberger (1994) provide test statistics for one time structural change under the null hypothesis of no break. Bai and Perron (1998, 2003) then develop tests for multiple structural breaks by splitting samples and considering one more break under the null sequentially in these sub-samples whenever the null hypothesis is rejected. This test continues until it fails to reject the null hypothesis.

From the Bayesian framework, Chib (1998) considers multiple change point models by interpreting structural changes as regime transitions and the regimes follow first-order Markov processes. Wang and Zivot (2000) also consider similar models by incorporating flat priors on the structural break dates and produce posterior distributions via Gibbs sampling.

However, the econometric methods in the literature have the restriction that all the parameters subject to the structural changes undergo the structural shifts at the same dates either in pure or partial structural change models so that they are not suitable for making inferences about the complicated patterns of structural change models considered in this paper with one exception of an interesting approach proposed by Levin and Piger (2008). They allow different subsets of parameters to change at different dates. Each subset of parameters could consist of a different combination of parameters. However, their method is computationally burdensome in the presence of a large number of breaks since the space of structural break models increases enormously with the number of structural breaks. They also impose a restriction that two adjacent break dates cannot be close even if the breaks occur in different parameters.

These limitations in the literature motivate developing the new approach in this paper and the proposed method successfully resolves all the critical issues raised in the literature review.

2 Model Specification and Bayesian Inference

2.1 Model Specification

I propose a Bayesian approach to making an inference about a structural break model in which multiple parameters can have structural changes independently at different dates. Thus, the structural break model can be specified as an econometric model that allows the residual variance as well as the coefficients to undergo the parameter shifts independently. In the model specification, I consider $G-1$ coefficient parameters and one residual variance parameter. Thus, the total number of parameters is G. For $t = 1, \ldots, T$,

$$y_t = X'_{1t}\beta_{1,S_{1,t}} + X'_{2t}\beta_{2,S_{2,t}} + \cdots + X'_{G-1,t}\beta_{G-1,S_{G-1,t}} + e_t, \quad e_t \sim i.i.d.\mathcal{N}(0, \sigma^2_{S_{G,t}}) \qquad (1)$$

where $S_{g,t} \in \{1, \ldots, (M_g + 1)\}$ is a regime indicator variable at time t for the gth parameter with M_g number of structural breaks (i.e. $M_g + 1$ regimes) and unknown change point dates $\Gamma_g = \{\tau_{g,1}, \ldots, \tau_{g,M_g}\}$ for $g = 1, \ldots, G$. The value of the gth coefficient, $\beta_{g,S_{g,t}}$, for $g = 1, \ldots, G - 1$ is then given over time by

$$\beta_{g,S_{g,t}} = \begin{cases} \beta_{g,1} & \text{if } 1 \leq t \leq \tau_{g,1} \\ \beta_{g,2} & \text{if } \tau_{g,1} < t \leq \tau_{g,2} \\ & \vdots \\ \beta_{g,M_g} & \text{if } \tau_{g,(M_g-1)} < t \leq \tau_{g,M_g} \\ \beta_{g,(M_g+1)} & \text{if } \tau_{g,M_g} < t \leq T \end{cases}$$

and the residual variance is given by

$$\sigma^2_{S_{G,t}} = \begin{cases} \sigma^2_1 & \text{if } 1 \leq t \leq \tau_{G,1} \\ \sigma^2_2 & \text{if } \tau_{G,1} < t \leq \tau_{G,2} \\ & \vdots \\ \sigma^2_{M_G} & \text{if } \tau_{G,(M_G-1)} < t \leq \tau_{G,M_G} \\ \sigma^2_{(M_G+1)} & \text{if } \tau_{G,M_G} < t \leq T. \end{cases}$$

In regression (1), the variance parameter is assigned to the last Gth parameter.[1]

[1] Note that it is also possible that some of the coefficient parameters and the variance parameter are placed in the same group and all the parameters in the same parameter group are restricted to have structural

All the parameters are assumed to be independent of one another in terms of the number and timing of structural changes. Suppose for example an AR(1) model with one structural break in intercept, two breaks in persistence, and no break in residual variance. It is then necessary to develop a new approach for estimation of the complicated patterns of structural break models presented in (1).

2.2 Single-Group Change-Point Model: Review

In order to explain the basic idea about the proposed approach in this paper, I first consider a single-group change-point model as in Chib (1998). In this paper, both pure and partial structural change models in Bai and Perron (2003) are defined as a single-group change-point model in the sense that all the parameters subject to the structural changes have the structural shifts at the same dates. Thus, they are sorted into one parameter group and have the same nature of structural breaks. For more details, refer to Bai and Perron (2003). For $t = 1, \ldots, T$,

$$y_t = X_t' \beta_{S_t} + e_t, \quad e_t \sim i.i.d.\mathcal{N}(0, \sigma^2_{S_t}) \tag{2}$$

where $S_t \in \{1, \ldots, (M+1)\}$ is a regime indicator variable at time t with M number of structural breaks (i.e. $M+1$ regimes) and unknown change point dates $\Gamma = \{\tau_1, \ldots, \tau_M\}$.

changes at the same dates although this type of model specification is not considered in this paper.

The model parameters are then given by

$$(\beta_{S_t}, \sigma^2_{S_t}) = \begin{cases} (\beta_1, \sigma_1^2) & \text{if } 1 \leq t \leq \tau_1 \\ (\beta_2, \sigma_2^2) & \text{if } \tau_1 < t \leq \tau_2 \\ & \vdots \\ (\beta_M, \sigma_M^2) & \text{if } \tau_{(M-1)} < t \leq \tau_M \\ (\beta_{(M+1)}, \sigma^2_{(M+1)}) & \text{if } \tau_M < t \leq T. \end{cases}$$

This model specification is a special case of the general model presented in (1) in the sense that all the parameters are sorted into a single-group and they undergo the structural breaks at the same dates.

In Bayesian econometrics, Chib's (1998) algorithm can be applied to this single-group change-point model by incorporating a latent regime variable S_t. Chib (1998) interprets structural breaks as regime transitions.[2] This idea can be expressed in a specification that the latent state variable S_t follows a first-order Markov process with the transition probabilities constrained as

$$Pr[S_t = i|S_{t-1} = i] = p_{i,i} \text{ and } Pr[S_t = i+1|S_{t-1} = i] = p_{i,i+1} = 1 - p_{i,i} \quad \text{for } i = 1, \ldots, M \tag{3}$$

and

$$Pr[S_t = M+1|S_{t-1} = M+1] = 1 \quad \text{for the last regime } M+1. \tag{4}$$

The transition probability $Pr[S_t = j|S_{t-1} = i] = p_{i,j}$ indicates the probability of moving to

[2] Wang and Zivot (2000) also consider multiple changes of all the parameters and they sample break dates from conditional distributions of break dates with flat prior.

regime j at time t given the regime i at time $t-1$. Since a regime S_t is allowed to either stay in the current regime or move to the next regime (i.e. it never comes back to the previous regimes), the transition probability for regime i is restricted to be $p_{i,i} = 1 - p_{i,i+1}$ as in (3). In addition, the last regime $M+1$ is absorbing and its transition probability is always specified to $p_{M+1,M+1} = 1$ as in (4).

A transition probability *matrix* P can then be formed as a $(M+1)$-by-$(M+1)$ matrix with elements containing the information about the first-order Markov process in (3) and (4). This restriction can be expressed as a matrix form where $p_{i,j}$ is placed in the (i,j)th entry of the transition matrix

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & 0 & \cdots & 0 \\ 0 & p_{2,2} & p_{2,3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \vdots & 0 & p_{M,M} & p_{M,M+1} \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{5}$$

Then, the model parameters augmented with transition probability and latent regime variable are sampled via a MCMC sampler proposed in Chib (1998). Based on this algorithm, various structural break models are compared through Bayes factor calculations. The details of MCMC scheme will be explained in the next section in the context of multiple-parameter change-point model.

2.3 Multiple-Parameter Change-Point Model Algorithm

In this section, I explain a new MCMC procedure for estimating multiple-parameter change-point model as in (1). Levin and Piger (2008) also consider a similar multiple-parameter change-point model by using the flat prior for break dates and allow possibly different subsets of parameters to be subject to the respective structural breaks. However, their method is computationally very costly. To see this, suppose, for example, there are three parameters and each parameter has one break independently. Then, while the new approach in this paper requires estimation of only one model allowing for one break in each parameter independently, in their approach it would be necessary to consider thirteen possible combinations of subsets of parameters and there is a need to estimate the thirteen different models.[3] Thus, the potential model space dramatically increases with either the number of breaks or the number of parameters. Their approach also restricts possible timing of structural changes in the sense that two adjacent break dates need to allow for the minimum length to practically make inferences about change-points even though each break occurs in different parameters.

In this paper, to make inferences about multiple-parameter change-point models, I extend Chib's (1998) approach introduced in Section 2.2. Suppose a multiple-parameter change-point model allows parameters to change at different dates with the different number of breaks. This implies that a structural break in one parameter could occur independently of those in the other parameter(s). For example, consider an autoregressive model with structural breaks in which the intercept coefficient has one break, the persistence coefficient

[3] Consider three categories: three parameters at the same date; two parameters at the same date and one parameter at the other date; all three parameters at different dates. Then, $13 = 1 + 3 \times 2! + 1 \times 3!$. In addition, take into account a more complicated case in which there are m breaks and k parameters. Then, the number of potential break models is $(2^k - 1)^m$.

has two breaks, and the residual variance has one break. For the example given above, the proposed method in this paper would only need a model augmented with *three* independent latent regime indicator variables $(S_{1,t}, S_{2,t}, S_{3,t})$ as in (1) and *three* transition probability matrices corresponding to three parameters respectively while the single-group change-point model is augmented with only *one* regime indicator variable S_t and *one* transition probability matrix P. Since all the regime indicator variables are mutually independent in the new approach, the date of regime transition in a parameter is allowed to occur close to that of regime transition in other parameters without any necessary minimum distance unlike the restriction in Levin and Piger (2008).

For the purpose of illustration, consider the case in which there are *three* parameters, $(\theta_1, \theta_2, \theta_3)$. From a Bayesian perspective, a joint posterior density can be obtained as being proportional to a product of a prior density and a likelihood function of $Y_T = [y_1 \ \ \dots \ \ y_T]'$ such as

$$\pi(\boldsymbol{\theta}, \boldsymbol{P}|Y_T) \propto \pi(\boldsymbol{\theta}, \boldsymbol{P}) f(Y_T|\boldsymbol{\theta}, \boldsymbol{P})$$

where $\pi(\cdot)$ denotes a density function; $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ is a collection of model parameters; and $\boldsymbol{P} = (P_1, P_2, P_3)$ is a collection of transition probability matrices. Let $\theta_1 = (\theta_{1,1}, \dots, \theta_{1,M_1+1})$ denote the collection of parameters for all the possible regimes for the first parameter with M_1 breaks (i.e. $M_1 + 1$ regimes) and θ_2 and θ_3 are defined accordingly. Also, $\theta_{g,k}$ indicates the regime k parameter for the gth parameter for $k = 1, \dots, M_g + 1$ and $g = 1, 2, 3$. The model parameters, $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, are then augmented with the transition probability matrices, $\boldsymbol{P} = (P_1, P_2, P_3)$, and the latent regime variables $\tilde{S}_{g,T} = [S_{g,1} \ \ S_{g,2} \ \ \dots \ \ S_{g,T-1} \ \ S_{g,T}]'$ for each parameter $g = 1, 2, 3$ where the discrete regime indicator variable $S_{g,t} \in \{1, \ldots, M_g + 1\}$. The MCMC sampling algorithm can be summarized in the context of three parameter model as follows.

MCMC sampling algorithm

Step 0: Initialize $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$, $\boldsymbol{P} = (P_1, P_2, P_3)$

Step 1: For the parameter 1,

- **(i)** Sample $\tilde{S}_{1,T}$ conditional on $(\theta_1, \theta_2, \theta_3), (\tilde{S}_{2,T}, \tilde{S}_{3,T}), (P_1, P_2, P_3), Y_T$
- **(ii)** Sample θ_1 conditional on $(\theta_2, \theta_3), (\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}), (P_1, P_2, P_3), Y_T$
- **(iii)** Sample P_1 conditional on $\tilde{S}_{1,T}$

Step 2: For the parameter 2,

- **(i)** Sample $\tilde{S}_{2,T}$ conditional on $(\theta_1, \theta_2, \theta_3), (\tilde{S}_{1,T}, \tilde{S}_{3,T}), (P_1, P_2, P_3), Y_T$
- **(ii)** Sample θ_2 conditional on $(\theta_1, \theta_3), (\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}), (P_1, P_2, P_3), Y_T$
- **(iii)** Sample P_2 conditional on $\tilde{S}_{2,T}$

Step 3: For the parameter 3,

- **(i)** Sample $\tilde{S}_{3,T}$ conditional on $(\theta_1, \theta_2, \theta_3), (\tilde{S}_{1,T}, \tilde{S}_{2,T}), (P_1, P_2, P_3), Y_T$
- **(ii)** Sample θ_3 conditional on $(\theta_1, \theta_2), (\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}), (P_1, P_2, P_3), Y_T$
- **(iii)** Sample P_3 conditional on $\tilde{S}_{3,T}$

Step 4: Repeat Steps 1-3

Note that the most recent values of the conditioning variables are used in all simulations. It is straightforward to implement simulation (ii) in each step so that it will be explained in Appendix. Only simulations (i) and (iii) in each step will be discussed below.

2.3.1 Simulation of latent regime variable $\tilde{S}_{g,T}$

In the simulation (i) in each step, the objective is to sample the discrete latent regime variable $S_{g,t} \in \{1, \ldots, M_g + 1\}$ for $t = 1, \ldots, T$ and $g = 1, 2, 3$ from the mass discrete function $p(\tilde{S}_{g,T}|\boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T)$ where $p(\cdot)$ denotes a discrete mass function and $\tilde{S}_{-g,T} = (\tilde{S}_{1,T}, \ldots, \tilde{S}_{g-1,T}, \tilde{S}_{g+1,T}, \ldots, \tilde{S}_{G,T})$. The mass function can be expressed as a joint density in reverse time order as follows.

$$
\begin{aligned}
p(\tilde{S}_{g,T}|\boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) &= p(S_{g,T}|\boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) \times p(S_{g,T-1}|S_{g,T}, \boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) \times \cdots \\
&\times p(S_{g,t}|S_g^{t+1}, \boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) \times \cdots \times p(S_{g,1}|S_g^{2}, \boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) \qquad (6)
\end{aligned}
$$

where $S_g^{t+1} = [S_{g,t+1} \ \ \ldots \ \ S_{g,T}]'$. Notice that the last regime and the first regime are always $M_g + 1$ and one respectively. These imply that for $g = 1, 2, 3$

$$
p(S_{g,T} = M_g + 1|\boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) = 1 \quad \text{and} \quad p(S_{g,1} = 1|\boldsymbol{\theta}, \tilde{S}_{-g,T}, \boldsymbol{P}, Y_T) = 1.
$$

Thus, the regimes $S_{g,t}$ for $t = 2, \ldots, t-1$ are recursively simulated from $t = T-1$ to $t = 2$ in reverse time order.

As discussed in Section 2.2, the regime transition follows a first order Markov process. It is also independent of its own parameter as well as both other parameters and their latent regime variables, as shown in Chib (1996). Thus, a term in (6) can be written that for $g = 1, 2, 3$

$$
p(S_{g,t}|Y_T, S_g^{t+1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) \propto p(S_{g,t}|Y_t, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) \times p(S_{g,t+1}|S_{g,t}, P_g).
$$

The first term in the proportion of the regime distribution is calculated recursively. Suppose $p(S_{g,t-1}|Y_{t-1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P})$ is known. Then, Bayes' rule can be applied for $k = 1, \ldots, M_g + 1$ regimes,

$$p(S_{g,t} = k|Y_t, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) = \frac{p(S_{g,t} = k|Y_{t-1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) \times f(y_t|Y_{t-1}, \theta_{g,k}, \boldsymbol{\theta}_{-g})}{\sum_{l=1}^{M_g+1} p(S_{g,t} = l|Y_{t-1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) \times f(y_t|Y_{t-1}, \theta_{g,l}, \boldsymbol{\theta}_{-g})}$$

where

$$p(S_{g,t} = k|Y_{t-1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P}) = \sum_{l=k-1}^{k} p(S_{g,t} = k|S_{g,t-1} = l, P_g) \times p(S_{g,t-1} = l|Y_{t-1}, \tilde{S}_{-g,T}, \boldsymbol{\theta}, \boldsymbol{P})$$

and $p(S_{g,t} = k|S_{g,t-1} = l, P_g)$ is the (l,k)th entry of the transition matrix P_g.

In sum, the probabilities of the regimes over dates are sampled through MCMC simulations:

$$Pr(S_{g,t} = k|Y_T) = \int p(S_{g,t} = k|Y_T, \boldsymbol{\theta}, \boldsymbol{P})\pi(\boldsymbol{\theta}, \boldsymbol{P}|Y_T)d(\boldsymbol{\theta}, \boldsymbol{P})$$

and in practice with J simulations

$$Pr(S_{g,t} = k|Y_T) = \frac{1}{J}\sum_{j=1}^{J} p(S_{g,t}^{(j)} = k|Y_T, \boldsymbol{\theta}^{(j)}, \boldsymbol{P}^{(j)}).$$

2.3.2 Simulation of transition probability matrix P_g

In simulation (iii) in each step, the transition probability matrices (P_1, P_2, P_3) are sampled only conditional on their regime variables $(S_{1,T}, S_{2,T}, S_{3,T})$ respectively. The reason is that the full conditional distribution $P_g|\boldsymbol{\theta}, S_{g,T}, S_{-g,T}, P_{-g}, Y_T$ is independent of $(\boldsymbol{\theta}, S_{-g,T}, P_{-g}, Y_T)$ where $P_{-g} = (P_1, \ldots, P_{g-1}, P_{g+1}, \ldots, P_G)$ for $g = 1, 2, 3$. Thus, it can be shown that

$$\pi\left(P_g|\boldsymbol{\theta}, S_{g,T}, S_{-g,T}, P_{-g}, Y_T\right) = \pi\left(P_g|S_{g,T}\right).$$

If Beta priors for $p_{i,i}$, $i = 1, \ldots, M_g$, are employed as

$$p_{i,i} \sim Beta(u_{i,i}, u_{i,i+1})$$

where $u_{i,i}$ and $u_{i,i+1}$ are the hyper-parameters, the posterior distribution can be derived as

$$p_{i,i}|\tilde{S}_{g,T} \sim Beta\left(u_{i,i} + n_{i,i}, u_{i,i+1} + n_{i,i+1}\right)$$

where $n_{i,j}$ refers to the total number of transitions from regime i to regime j. Note that $n_{i,i+1}$, for $i = 1, \ldots, M_g$, is always equal to one since every regime never comes back to the previous regimes and moves to the next regime only once. For details, see Albert and Chib (1993).

2.4 Model Selection

In a Bayesian framework, model selection relies on comparisons of Bayes factors. The Bayes factor is calculated from the ratio of the marginal likelihoods for two competing models under consideration. Let $\mathcal{M}$ be the model indicator parameter. For models $\mathcal{M} = i, j$, the Bayes factor in favor of model $\mathcal{M} = i$ is given by

$$B_{ij} = \frac{m(Y_T|\mathcal{M} = i)}{m(Y_T|\mathcal{M} = j)}$$

where $m(Y_T|\mathcal{M} = i)$ is the marginal likelihood or the marginal density of the data Y_T under model $\mathcal{M} = i$. The marginal likelihood of model $\mathcal{M}$ can be easily derived through the method of Chib (1995) for Gibbs sampling based on the Bayes rule identity:

$$m(Y_T|\mathcal{M}) = \frac{f(Y_T|\boldsymbol{\psi}, \mathcal{M})\pi(\boldsymbol{\psi}|\mathcal{M})}{\pi(\boldsymbol{\psi}|Y_T, \mathcal{M})} \tag{7}$$

where $\boldsymbol{\psi} = (\boldsymbol{\theta}, \boldsymbol{P})$. The above identity holds for any point $\boldsymbol{\psi}$ in the parameter space since the left hand side is free of $\boldsymbol{\psi}$. Taking the logarithm of the marginal likelihood for computational convenience, the estimate of the marginal density at any particular point $\boldsymbol{\psi}^*$ is given by

$$\begin{aligned} & ln\hat{m}(Y_T|\mathcal{M}) \\ = \; & lnf(Y_T|\boldsymbol{\theta^*}, \boldsymbol{P}^*, \mathcal{M}) + ln\pi(\boldsymbol{\theta^*}, \boldsymbol{P}^*|\mathcal{M}) - ln\pi(\boldsymbol{\theta^*}, \boldsymbol{P}^*|Y_T, \mathcal{M}). \end{aligned}$$

In this paper, I calculate the marginal likelihood at the posterior mean. I explain all the terms in equation (8) in the following subsections. For simplicity, I drop the model indicator $\mathcal{M}$ from now on.

2.4.1 Likelihood function

The logarithm of likelihood function is given by

$$ln\ f(Y_T|\boldsymbol{\psi}^*) = \sum_{t=1}^{T} lnf(y_t|Y_{t-1}, \boldsymbol{\psi}^*)$$

where

$$f(y_t|Y_{t-1},\boldsymbol{\psi}^*) = \sum_{S_{1,t}=1}^{M_1+1}\sum_{S_{2,t}=1}^{M_2+1}\sum_{S_{3,t}=1}^{M_3+1} f(y_t|Y_{t-1},S_{1,t},S_{2,t},S_{3,t},\boldsymbol{\psi}^*) \times p(S_{1,t},S_{2,t},S_{3,t}|Y_{t-1},\boldsymbol{\psi}^*)$$

is the one-step ahead prediction density and $f(y_t|Y_{t-1},S_{1,t},S_{2,t},S_{3,t},\boldsymbol{\psi}^*)$ is the conditional density of y_t given the composite of regimes $(S_{1,t},S_{2,t},S_{3,t})$ as well as the posterior mean $\boldsymbol{\psi}^*$. Define a composite of regimes for all the parameters by $S_t = (S_{1,t},S_{2,t},S_{3,t})$. Then, $p(S_{1,t},S_{2,t},S_{3,t}|Y_{t-1},\boldsymbol{\psi}^*)$ is the joint discrete mass function of the composite $S_t = (S_{1,t},S_{2,t},S_{3,t})$ and the transition probability matrix for the composite of regimes S_t is given by $P_1 \otimes P_2 \otimes P_3$ where $\otimes$ indicates the Kronecker product, P is a m-by-m square matrix, and the number of the composite of regimes S_t is given by $m = (M_1+1)\times(M_2+1)\times(M_3+1)$.

2.4.2 Prior density

All the parameters are *a priori* assumed to be independent of one another and the logarithm of prior density is given by

$$ln\ \pi(\boldsymbol{\psi}^*|M) = ln\ \pi(\theta_1^*) + ln\ \pi(\theta_2^*) + ln\ \pi(\theta_3^*) + ln\ \pi(P_1^*) + ln\ \pi(P_2^*) + ln\ \pi(P_3^*).$$

2.4.3 Posterior density

In order to estimate the posterior ordinate $\pi(\boldsymbol{\theta^*},\boldsymbol{P}^*|Y_T)$, I consider the conditional decomposition of the posterior density as in Chib (1998). Note that the latent variables $(\tilde{S}_{1,T},\tilde{S}_{2,T},\tilde{S}_{3,T})$ are integrated out in the calculation of the posterior density in each step.

$$\pi(\boldsymbol{\theta^*}, \boldsymbol{P}^*|Y_T) = \pi(\theta_1^*|Y_T) \times \pi(P_1^*|\theta_1^*, Y_T) \times \pi(\theta_2^*|\theta_1^*, P_1^*, Y_T)$$

$$\times\pi(P_2^*|\theta_1^*, \theta_2^*, P_1^*, Y_T) \times \pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, Y_T)\pi(P_3^*|\theta_1^*, \theta_2^*, \theta_3^*, P_1^*, P_2^*, Y_T)$$

where

$$\pi(\theta_1^*|Y_T) = \int \pi(\theta_1^*|\theta_2, \theta_3, P_1, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}, Y_T)$$

$$\times\pi(\theta_2, \theta_3, P_1, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}|Y_T)d\theta_2 d\theta_3 dP_1 dP_2 dP_3 d\tilde{S}_{1,T} d\tilde{S}_{2,T} d\tilde{S}_{3,T},$$

$$\pi(P_1^*|\theta_1^*, Y_T) = \int \pi(P_1^*|\theta_1^*, \theta_2, \theta_3, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}, Y_T)$$

$$\times\pi(\theta_2, \theta_3, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}|\theta_1^*, Y_T)d\theta_2 d\theta_3 dP_2 dP_3 d\tilde{S}_{1,T} d\tilde{S}_{2,T} d\tilde{S}_{3,T},$$

$$\pi(\theta_2^*|\theta_1^*, P_1^*, Y_T) = \int \pi(\theta_2^*|\theta_1^*, P_1^*, \theta_2, \theta_3, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}, Y_T)$$

$$\times\pi(\theta_3, P_2, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}|\theta_1^*, P_1^*, Y_T)d\theta_3 dP_2 dP_3 d\tilde{S}_{1,T} d\tilde{S}_{2,T} d\tilde{S}_{3,T},$$

$$\pi(P_2^*|\theta_1^*, \theta_2^*, P_1^*, Y_T) = \int \pi(P_2^*|\theta_1^*, \theta_2^*, P_1^*, \theta_3, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}, Y_T)$$

$$\times\pi(\theta_3, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}|\theta_1^*, \theta_2^*, P_1^*, Y_T)d\theta_3 dP_3 d\tilde{S}_{1,T} d\tilde{S}_{2,T} d\tilde{S}_{3,T},$$

$$\pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, Y_T) = \int \pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}, Y_T)$$

$$\times\pi(P_3, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T}|\theta_1^*, \theta_2^*, P_1^*, P_2^*, Y_T)dP_3 d\tilde{S}_{1,T} d\tilde{S}_{2,T} d\tilde{S}_{3,T},$$

and

$$\pi(P_3^*|\theta_1^*,\theta_2^*,\theta_3^{2*},P_1^*,P_2^*,Y_T)=\int\pi(P_3^*|\theta_1^*,\theta_2^*,\theta_3^*,P_1^*,P_2^*,\tilde{S}_{1,T},\tilde{S}_{2,T},\tilde{S}_{3,T},Y_T)$$

$$\times\pi(\tilde{S}_{1,T},\tilde{S}_{2,T},\tilde{S}_{3,T}|\theta_1^*,\theta_2^*,\theta_3^*,P_1^*,P_2^*,Y_T)d\tilde{S}_{1,T}d\tilde{S}_{2,T}d\tilde{S}_{3,T}.$$

The decomposition of the posterior density shows that the first ordinate $\pi(\theta_1^*|Y_T)$ can be calculated based on draws from the full Gibbs run, and $\pi(P_1^*|\theta_1^*,Y_T)$, $\pi(\theta_2^*|\theta_1^*,P_1^*,Y_T)$, $\pi(P_2^*|\theta_1^*,\theta_2^*,P_1^*,Y_T)$, $\pi(\theta_3^*|\theta_1^*,\theta_2^*,P_1^*,P_2^*,Y_T)$, and $\pi(P_3^*|\theta_1^*,\theta_2^*,\theta_3^*,P_1^*,P_2^*,Y_T)$ can be calculated from appropriate reduced Gibbs runs. The Monte Carlo estimate of each decomposition component based on draws from each Gibbs run is calculated as follows.

$$\pi(\theta_1^*|Y_T)=\frac{1}{J}\sum_{j=1}^{J}\pi(\theta_1^*|\theta_2^{(j)},\theta_3^{(j)},P_1^{(j)},P_2^{(j)},P_3^{(j)},\tilde{S}_{1,T}^{(j)},\tilde{S}_{2,T}^{(j)},\tilde{S}_{3,T}^{(j)},Y_T),$$

$$\pi(P_1^*|\theta_1^*,Y_T)=\frac{1}{J}\sum_{j_1=1}^{J}\pi(P_1^*|\theta_1^*,\theta_2^{(j_1)},\theta_3^{(j_1)},P_2^{(j_1)},P_3^{(j_1)},Y_T),$$

$$\pi(\theta_2^*|\theta_1^*,P_1^*,Y_T)=\frac{1}{J}\sum_{j_2=1}^{J}\pi(\theta_2^*|\theta_1^*,P_1^*,\theta_2^{(j_2)},\theta_3^{(j_2)},\tilde{S}_{1,T}^{(j_2)},\tilde{S}_{2,T}^{(j_2)},\tilde{S}_{3,T}^{(j_2)},Y_T),$$

$$\pi(P_2^*|\theta_1^*,\theta_2^*,P_1^*,Y_T)=\frac{1}{J}\sum_{j_3=1}^{J}\pi(P_2^*|\theta_1^*,\theta_2^*,P_1^*,\theta_3^{(j_3)},P_3^{(j_3)},\tilde{S}_{1,T}^{(j_3)},\tilde{S}_{2,T}^{(j_3)},\tilde{S}_{3,T}^{(j_3)},Y_T),$$

$$\pi(\theta_3^*|\theta_1^*,\theta_2^*,P_1^*,P_2^*,Y_T)=\frac{1}{J}\sum_{j_4=1}^{J}\pi(\theta_3^*|\theta_1^*,\theta_2^*,P_1^*,P_2^*,P_3^{(j_4)},\tilde{S}_{1,T}^{(j_4)},\tilde{S}_{2,T}^{(j_4)},\tilde{S}_{3,T}^{(j_4)},Y_T),$$

and

$$\pi(P_3^*|\theta_1^*,\theta_2^*,\theta_3^*,P_1^*,P_2^*,Y_T)=\frac{1}{J}\sum_{j_5=1}^{J}\pi(P_3^*|\theta_1^*,\theta_2^*,\theta_3^*,P_1^*,P_2^*,\tilde{S}_{1,T}^{(j_5)},\tilde{S}_{2,T}^{(j_5)},\tilde{S}_{3,T}^{(j_5)},Y_T)$$

where the superscript (j) refers to the jth draw of the full Gibbs run and the superscript (j_i), $i = 1, \ldots, 5$, refers to the j_ith draw from the appropriate reduced Gibbs run. Thus, in addition to the full Gibbs run for the usual estimation of parameters, it is required to have five additional Gibbs runs ($5 \times J$ iterations). For example, $\pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, Y_T)$ is calculated by additional J iterations from the following reduced j_4th Gibbs run.

Algorithm for $\pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, Y_T)$

(i) Generate $\theta_3^{(j_4)}$ from $\pi(\theta_3|\theta_1^*, \theta_2^*, P_1^*, P_2^*, P_3^{(j_4-1)}, S_{1,T}^{(j_4-1)}, S_{2,T}^{(j_4-1)}, S_{3,T}^{(j_4-1)}, Y_T)$

(ii) Generate $P_3^{(j_4)}$ from $\pi(P_3|\theta_1^*, \theta_2^*, P_1^*, P_2^*, \theta_3^{(j_4)}, S_{1,T}^{(j_4-1)}, S_{2,T}^{(j_4-1)}, S_{3,T}^{(j_4-1)}, Y_T)$

(iii) Generate $\tilde{S}_{1,T}^{(j_4)}$ from $\pi(S_{1,T}|\theta_1^*, \theta_2^*, P_1^*, P_2^*, \theta_3^{(j_4)}, P_3^{(j_4)}, S_{2,T}^{(j_4-1)}, S_{3,T}^{(j_4-1)}, Y_T)$

(iv) Generate $\tilde{S}_{2,T}^{(j_4)}$ from $\pi(S_{2,T}|\theta_1^*, \theta_2^*, P_1^*, P_2^*, \theta_3^{(j_4)}, P_3^{(j_4)}, S_{1,T}^{(j_4)}, S_{3,T}^{(j_4-1)}, Y_T)$

(v) Generate $\tilde{S}_{3,T}^{(j_4)}$ from $\pi(S_{3,T}|\theta_1^*, \theta_2^*, P_1^*, P_2^*, \theta_3^{(j_4)}, P_3^{(j_4)}, S_{1,T}^{(j_4)}, S_{2,T}^{(j_4)}, Y_T)$

(vi) Evaluate $\pi(\theta_3^*|\theta_1^*, \theta_2^*, P_1^*, P_2^*, P_3^{(j_4)}, S_{1,T}^{(j_4)}, S_{2,T}^{(j_4)}, \tilde{S}_{3,T}^{(j_4)}, Y_T)$

Notice that, throughout this reduced Gibbs run, θ_1^*, θ_2^*, θ_3^*, P_1^*, and P_2^* are set equal to their posterior mean.

3 Application to Artificial Data

To demonstrate how the proposed method performs, I generate an artificial data set from the following process:

$$y_t = \alpha_{S_{1,t}} + \beta_{S_{2,t}} x_t + e_t, \ e_t \sim i.i.d.N(0, \sigma^2_{S_{3,t}}) \tag{8}$$

where $x_t \sim \mathcal{U}(0,1)$,

$$\alpha_{S_{1,t}} = \begin{cases} 1.0 \;\&\; S_{1,t}=1 & \text{if } 1 \leq t \leq 40 \\ 1.3 \;\&\; S_{1,t}=2 & \text{if } 41 \leq t \leq 200, \end{cases} \quad \beta_{S_{2,t}} = \begin{cases} 2.5 \;\&\; S_{2,t}=1 & \text{if } 1 \leq t \leq 80 \\ 1.5 \;\&\; S_{2,t}=2 & \text{if } 81 \leq t \leq 120 \\ 3.0 \;\&\; S_{2,t}=3 & \text{if } 121 \leq t \leq 200, \end{cases}$$

$$\text{and} \quad \sigma^2_{S_{3,t}} = \begin{cases} 0.4^2 \;\&\; S_{3,t}=1 & \text{if } 1 \leq t \leq 100 \\ 1.0^2 \;\&\; S_{3,t}=2 & \text{if } 101 \leq t \leq 150 \\ 0.6^2 \;\&\; S_{3,t}=3 & \text{if } 151 \leq t \leq 200. \end{cases}$$

This model consists of three parameters $(\alpha, \beta, \sigma^2)$ and has one break in the intercept α, two breaks in the response coefficient β, and two breaks in the residual variance σ^2. Note that all the true structural breaks occur at different dates. The generated data set $\{y_t\}$ is depicted in Figure 1. Looking at the generated time series, it appears difficult to make an inference about the existence and timing of the structural changes in the individual parameters.

Based on the proposed approach in Section 2, I generate posteriors for the parameters as well as regime transitions via Gibbs sampling with 10,000 simulations after 2,000 burn-ins. I assume that each parameter has the same prior distribution across regimes. This assumption ensures that the prior distributions do not affect making an inference about change points. The prior distributions are summarized as follows.

$$\begin{aligned}\alpha &\sim N(0,1)\\ \beta &\sim N(0,1)\\ \sigma^2 &\sim IG(3.01, 2.10)\\ p_{ii} &\sim Beta(1, 0.01).\end{aligned}$$

Note that the prior on the residual variance is distributed with mean 1.045 and standard deviation 1.040. Also, the mean and the standard deviation of the prior distribution for the regime transition probability are 0.99 and 0.07 respectively. The expected duration of one regime can be calculated from $D_i = 1/(1 - p_{ii})$. From the prior density for the transition probability chosen in this estimation, the regime is expected to last for 101 periods.

Simulated regime transitions for each parameter are illustrated in Figure 4. The estimated break dates are quite consistent with the true break dates although in the data generating process the pattern of the structural breaks is highly complicated.

4 Application to U.S. GDP Deflator Inflation

4.1 Data and Model Specification

For an empirical application of the new approach, I consider a univariate autoregressive model for inflation. I calculate the quarterly percentage change of the U.S. GDP deflator expressed as an annualized rate from 1953:Q1 to 2005:Q2. This data set is illustrated in Figure 3.

In this paper, I consider a pth-order autoregressive model which allows for structural breaks in three parameters such as (i) intercept α, (ii) persistence coefficient ρ and (iii) residual variance σ^2 as follows.

$$\pi_t = \alpha_{S_{1,t}} + \rho_{S_{2,t}}\pi_{t-1} + \gamma_1 \Delta\pi_{t-1} + \ldots + \gamma_{p-1}\Delta\pi_{t-(p-1)} + e_t, \quad e_t \sim i.i.d.N(0, \sigma^2_{S_{3,t}})$$

where $S_{1,t} \in \{1, .., M_\alpha + 1\}$, $S_{2,t} \in \{1, .., M_\rho + 1\}$, and $S_{3,t} \in \{1, .., M_{\sigma^2} + 1\}$ represent the regimes for the intercept parameter with M_α breaks, the persistence parameter with M_ρ breaks and the residual variance with M_{σ^2} breaks respectively. The persistence is measured as the sum of the autoregressive coefficients. The nature of the structural breaks in each parameter is independent of one another in terms of the number and timing of structural breaks as in the previous application to the artificial data set in Section 3.

In this analysis, the lag order p is set to four in order to fully capture the persistence of inflation dynamics. The diffuse and same prior distributions across different regimes are chosen in order to avoid any distortions from the choice of specific prior distributions when Bayes factors are calculated for the different structural break models. The priors of regression coefficients are distributed with mean zero and variance one (i.e., $\rho, \gamma \sim \mathcal{N}(0, 1)$) and the priors of variance parameters follow Inverse Gamma distribution such as $\sigma^2 \sim IG(3.01, 2.10)$ across different regimes. The prior of the transition probability that the current regime i stays in the same regime i in the next period is distributed as $p_{ii} \sim Beta(1, 0.01)$. The prior expected duration of a given regime is about 101 quarters. All the estimations are based on 10,000 Gibbs simulations after discarding 2,000 burn-ins.

4.2 Model Selection: Marginal Likelihood Comparison

In this application, the maximum number of structural breaks is specified to four for each parameter and in total, 125 different models are considered including a model with no break ($125 = 5 \times 5 \times 5$). Comparing the marginal likelihoods, the most preferred model has two structural breaks in the residual variance and no structural break in the intercept and persistence parameters.

Table 1 shows the model comparison results with the logarithm of marginal likelihood and the Bayes Factor. Let M_i denotes the number of breaks in the parameter i, for $i \in \{\alpha, \rho, \sigma^2\}$. Then, the Bayes factor is presented in favor of the alternative model, $\mathcal{M} = (M_\alpha, M_\rho, M_{\sigma^2})$ versus the most preferred model, $\mathcal{M} = (0, 0, 2)$ by

$$B = \frac{m\left(Y_T | \mathcal{M} = (M_\alpha, M_\rho, M_{\sigma^2})\right)}{m\left(Y_T | \mathcal{M} = (0, 0, 2)\right)}.$$

For the summary, only ten models with high marginal likelihood values are presented among 125 models. The model comparison based on calculating the Bayes Factors also shows that the most preferred model with two breaks only in the residual variance clearly dominates the other models in the sense that the Bayes factor is lower than 1/18 in favor of any alternative model.

4.3 Model Selection: Posterior Probability of the Alternative Number of Breaks

To examine the robustness of the model selection procedure in Section 4.2, I also calculate the posterior probability for the number of structural breaks in the *individual* parameters by integration. For example, the posterior probability for l structural breaks in the parameter α, denoted by $M_{\alpha,l}$, is given by

$$Pr(M_{\alpha,l}|Y_T) = \frac{m(Y_T|M_{\alpha,l})\pi(M_{\alpha,l})}{\sum_{i=0}^{4} m(Y_T|M_{\alpha,i})\pi(M_{\alpha,i})} \tag{9}$$

where

$$m(Y_T|M_{\alpha,i}) = \sum_{j=0}^{4}\sum_{k=0}^{4} m(Y_T|M_{\alpha,i}, M_{\rho,j}, M_{\sigma^2,k})\pi(M_{\rho,j}, M_{\sigma^2,k}|M_{\alpha,i})$$

is the integrated likelihood of $M_{\alpha,i}$; $\pi(M_{\alpha,i})$ is the prior probability of $M_{\alpha,i}$; and $\pi(M_{\rho,j}, M_{\sigma^2,k}|M_{\alpha,i})$ is the joint prior probability of $M_{\rho,j}$ and $M_{\sigma^2,k}$ conditional on $M_{\alpha,i}$. Since all the models are considered *a priori* equally likely as well as independent, $\pi(M_{\alpha,i})$ is equal to 1/5 and $\pi(M_{\rho,j}, M_{\sigma^2,k}|M_{\alpha,i})$ is equal to 1/25 when the maximum number of structural breaks is specified to four as in this analysis. In fact, the posterior probability of l structural breaks in the parameter α is simply the sum of posterior probabilities for all the models which has the same number of structural breaks in the parameter α. For other parameters, ρ and σ^2, $Pr(M_{\rho,l}|Y_T)$ and $Pr(M_{\sigma^2,l}|Y_T)$ can be easily obtained by using the same approach in (9). Note that all the terms in (9) are readily available when the marginal likelihood calculations are completed.

Table 2 lists the posterior probabilities for all possible values of the number of structural breaks in the individual parameters. The posterior probability for two breaks in the residual variance (σ^2) is 0.934 and that for three breaks is 0.043 while the posterior probabilities for no break in the intercept parameter (α) and in the persistence parameter (ρ) are 0.992 and 0.994 respectively.

Thus, not only the marginal likelihood calculation comparisons but also the posterior probability calculations produce very strong evidence that the autoregressive model for the inflation rate has two breaks in the residual variance and no break in the intercept and the persistence parameters. This result is also consistent with the finding that evidence for shifts in persistence is not statistically significant, particularly once allowing for shifts in the residual variance in Pivetta and Reis (2007) and Stock (2001).

4.4 Posterior Summary for the Most Preferred Model

Having found the substantial evidence for two breaks in the residual variance and no break in the intercept and the persistence, I focus more on the estimation results from the most preferred model.

Table 3 and Figure 4 present the posterior distribution for the timing of the structural changes in the residual variance. The break dates with the highest posterior probabilities are 1970:Q2 for the first structural change and 1985:Q2 for the second respectively. The shape of the posterior distributions is very sharp in Figure 4. This finding is also confirmed by the narrow 90% credible intervals: 1966:Q4 to 1971:Q3 for the first break date and 1983:Q2 to 1988:Q1 for the second break date. This implies that the inflation rate data provide quite precise information about the timing of the structural changes in the residual variance when utilizing the proposed approach.

Table 4 summarizes the posterior distributions for the parameters in the most preferred model and Figure 5 plots the median and the 90% credible interval for the model parameters over the sample period. They show that the residual variance switches from the low volatility regime to the high volatility regime around 1970 and then returns to another low volatility regime around 1985. Comparing two low volatility regimes, the residual variance before 1970 appears a bit bigger than that after 1985. Also, note that the inflation rate is highly persistent in the sense that the mean of the persistence parameter is 0.9212. This empirical result on high inflation persistence is consistent with finding in the literature (e.g. Fuhrer and Moore (1995)). However, the 90% credible interval doesn't cover the unit root.

5 Conclusion

Most macroeconomic variables in the postwar U.S. economy have experienced structural instability in conditional mean and variance, but the nature of the structural breaks is different across parameters of interest. The possibility of various complicated patterns of structural breaks is what motivates this paper.

Thus, I propose an efficient Bayesian MCMC method that allows for a number of possibilities for the nature of structural breaks. This new approach is developed to have the following distinctive features: (i) model specification of considering multiple structural changes in multiple parameters; (ii) model flexibility in allowing the multiple structural breaks to occur mutually independently at different dates across different parameters; and (iii) model selection procedure by comparing various potentially non-nested structural break models.

This method extends Chib's (1998) approach in which structural breaks are interpreted as regime transitions. He introduces a latent discrete regime variable which indicates one of all the possible regimes over time. A first-order Markov process then governs the structural changes with transition probabilities constrained so that the regime indicator variable can either stay in the current regime or move to the next regime. Chib (1998) assumes that all the parameters under the structural changes are restricted to have the structural shifts at the same dates. In consequence his approach requires only *one* regime indicator variable which governs the structural breaks in *all* the parameters subject to the changes.

The more flexible assumption in this paper requires the specification that each parameter is augmented with its corresponding independent regime indicator variable. The total number of the parameters subject to structural changes is then the same as that of the regime indicator variables which are independent of one another. In this paper, the MCMC sampler is presented through a hierarchical specification in which one draws the model parameters conditional on the regime variables and the observed data; the regime variables conditional on the model parameters and the observed data; and finally the transition probabilities conditional on the regime variables via Gibbs sampling.

I then apply this approach to an artificial data set generated by a model which has a complicated pattern of structural breaks in parameters. The new approach identifies accurately break dates from the artificial data in the sense that estimated regime transitions occur around true break dates in the data generating process. As an empirical application of the method proposed in this paper, I run various autoregressive regressions with structural breaks in intercept, persistence, and/or residual variance for U.S. GDP deflator inflation. I find that there are two structural changes in the residual variance and no break in the intercept and the persistence parameters. The residual variance switches from the low volatility regime to the high volatility regime in the early 1970s and then returns to another low volatility regime in the mid-1980s

References

Albert, J. H., and S. Chib (1993): "Bayes inference iva Gibbs Sampling of autoregressive time series subject to Markov mean and variance shifts," *Journal of Business and Economic Statistics*, 11, 1–15.

Andrews, D. W. K. (1993): "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61(4), 821–56.

Andrews, D. W. K., and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter Is Present Only under the Alternative," *Econometrica*, 62(6), 1383–1414.

Bai, J., and P. Perron (1998): "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66(1), 47–78.

——— (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics*, 18(1), 1–22.

Chib, S. (1996): "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics*, 75(1), 79–97.

——— (1998): "Estimation and comparison of multiple change-point models," *Journal of Econometrics*, 86(2), 221–241.

Fuhrer, J., and G. Moore (1995): "Inflation Persistence," *The Quarterly Journal of Economics*, 110(1), 127–159.

Hansen, B. E. (2001): "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity," *The Journal of Economic Perspectives*, 15(4), 117–128.

Kim, C.-J., and C. R. Nelson (1999): "Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle," *The Review of Economics and Statistics*, 81(4), 608–616.

Levin, A., and J. Piger (2008): "Bayesian Model Selection for Structural Break Models," *Manuscript*.

McConnell, M. M., and G. Perez-Quiros (2000): "Output Fluctuations in the United States: What Has Changed since the Early 1980's?," *American Economic Review*, 90(5), 1464–1476.

Perron, P. (1989): "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Econometrica*, 57(6), 1361–1401.

Pivetta, F., and R. Reis (2007): "The Persistence of Inflation in the United States," *Journal of Economic Dynamics and Control*, 31, 1326–1358.

Rapach, D. E., and M. E. Wohar (2005): "Regime Changes in International Real Interest Rates: Are They a Monetary Phenomenon?," *Journal of Money, Credit and Banking*, 37 (5), 887–906.

Stock, J. (2001): "Comment on Evolving Post-World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual 2001*, pp. 379–387.

Stock, J. H., and M. W. Watson (1996): "Evidence on Structural Instability in Macroeconomic Time Series Relations," *Journal of Business & Economic Statistics*, 14(1), 11–30.

——— (2002): "Has the Business Cycle Changed, and Why?," *NBER Macroeconomics Annual*, 17, 159–218.

Wang, J., and E. Zivot (2000): "A Time Series Model of Multiple Structural Changes in Level, Trend and Variance," *Journal of Business and Economic Statistics*, 18(3), 374–386.

Zivot, E., and D. W. K. Andrews (1992): "Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis," *Journal of Business & Economic Statistics*, 10(3), 251–70.

Appendix

In this appendix, I describe how to sample the posterior of model parameters conditional on the regime transition probabilities, $\boldsymbol{P} = (P_1, P_2, P3)$ and the latent regime variables, $(\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T})$ in the case of *three* parameters given in Section 2.

Consider a regression model with structural breaks in intercept, response coefficient, and residual variance independently as follows.

$$y_t = \alpha_{S_{1,t}} + x_t' \beta_{S_{2,t}} + e_t, \quad e_t \sim N(0, \sigma^2_{s_{3,t}})$$

$$S_{1,t} \in \{1, \ldots, (M_1+1)\}; \quad S_{2,t} \in \{1, \ldots, (M_2+1)\}; \quad S_{3,t} \in \{1, \ldots, (M_3+1)\}$$

A.1 Sampling of intercept α

Conditional on $(\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T})$, $\tilde{\beta} = (\beta_1, \ldots, \beta_{M_2+1})$, and $\tilde{\sigma}^2 = (\sigma_1^2, \ldots, \sigma^2_{M_3+1})$, intercept parameter for regime j, α_j, for $j = 1, \ldots, (M_1+1)$ can be sampled as follows.

(a) Prior

$$\alpha_j \sim N\left(\underline{\alpha}_j, \underline{D}_{\alpha_j}\right)$$

(b) Posterior

$$\overline{\alpha}_j | \tilde{\beta}, \tilde{\sigma}^2, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T} \sim N\left(\overline{\alpha}_j, \overline{D}_{\alpha_j}\right)$$

where

$$\overline{\alpha}_j = \left(\underline{D}^{-1}_{\alpha_j} + \sum_{\{S_{1,t}=j\}} 1/\sigma^2_{S_{3,t}}\right)^{-1} \left(\underline{D}^{-1}_{\alpha_j}\underline{\alpha}_j + \sum_{\{S_{1,t}=j\}} (y_t - x_t'\beta_{S_{2,t}})/\sigma^2_{S_{3,t}}\right)$$

and

$$\overline{D}_{\alpha_j} = \left(\underline{D}^{-1}_{\alpha_j} + \sum_{\{S_{1,t}=j\}} 1/\sigma^2_{S_{3,t}}\right)^{-1}.$$

A.2 Sampling of coefficient β

Conditional on $(\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T})$, $\tilde{\alpha} = (\alpha_1, \ldots, \alpha_{M_1+1})$, and $\tilde{\sigma}^2 = (\sigma_1^2, \ldots, \sigma^2_{M_3+1})$, response coefficient parameter for regime j, β_j, for $j = 1, \ldots, (M_2+1)$ can be sampled as follows.

(a) Prior

$$\beta_j \sim N\left(\underline{\beta}_j, \underline{D}_{\beta_j}\right)$$

(b) Posterior

$$\overline{\beta}_j | \tilde{\alpha}, \tilde{\sigma}^2, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T} \sim N\left(\overline{\beta}_j, \overline{D}_{\beta_j}\right)$$

where

$$\overline{\beta}_j = \left(\underline{D}_{\beta_j}^{-1} + \sum_{\{S_{2,t}=j\}} x_t x_t' / \sigma^2_{S_{3,t}} \right)^{-1} \left(\underline{D}_{\beta_j}^{-1} \underline{\beta}_j + \sum_{\{S_{2,t}=j\}} x_t (y_t - \alpha_{S_{1,t}}) / \sigma^2_{S_{3,t}} \right)$$

and

$$\overline{D}_{\beta_j} = \left(\underline{D}_{\beta_j}^{-1} + \sum_{\{S_{2,t}=j\}} x_t x_t' / \sigma^2_{S_{3,t}} \right)^{-1}.$$

A.3 Sampling of variance σ^2

Conditional on $(\tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T})$, $\tilde{\alpha} = (\alpha_1, \ldots, \alpha_{M_1+1})$, and $\tilde{\beta} = (\beta_1, \ldots, \beta_{M_2+1})$, residual variance for regime j, σ_j^2, for $j = 1, \ldots, (M_3 + 1)$ can be sampled as follows.

(a) Prior

$$\sigma_j^2 \sim IG\left(\frac{\nu_j}{2}, \frac{\delta_j}{2}\right)$$

(b) Posterior

$$\sigma_j^2 | \tilde{\alpha}, \tilde{\beta}, \tilde{S}_{1,T}, \tilde{S}_{2,T}, \tilde{S}_{3,T} \sim IG\left(\frac{\nu_j + n_j}{2}, \frac{\delta_j + \sum_{\{S_{3,t}=j\}} (y_t - \alpha_{S_{1,t}} - x_t' \beta_{S_{2,t}})^2}{2}\right)$$

where $n_j = \sum_{t=1}^{T} \mathbf{1}[S_{3,t} = j]$ is the number of observations ascribed to regime j and $\mathbf{1}[\cdot]$ is an indicator function.

Table 1: Model Selection: Marginal Likelihood and Bayes Factor

Model (# of breaks) $\mathcal{M} = (m_\alpha, m_\rho, m_{\sigma^2})$	Log marginal likelihood	Bayes Factor
(0,0,2)	-308.03	1.0000
(0,0,3)	-310.96	0.0534
(0,0,1)	-311.87	0.0214
(1,0,2)	-312.93	0.0074
(0,1,2)	-313.26	0.0053
(0,0,4)	-313.97	0.0026
(0,1,3)	-315.26	0.0007
(1,1,2)	-315.62	0.0005
(1,0,3)	-315.71	0.0005
(0,1,1)	-316.38	0.0000

Note: M_α, M_ρ, and M_{σ^2} denote the number of breaks in the intercept (α), the persistence (ρ), and the residual variance (σ^2) respectively. Bayes factors are calculated in favor of the alternative model, $\mathcal{M} = (M_\alpha, M_\rho, M_{\sigma^2})$ versus the most preferred model, $\mathcal{M} = (0, 0, 2)$:

$$B = \frac{m\left(Y_T|\mathcal{M} = (M_\alpha, M_\rho, M_{\sigma^2})\right)}{m\left(Y_T|\mathcal{M} = (0, 0, 2)\right)}$$

Table 2: Posterior Probability of the Number of Structural Breaks in Individual Parameters

Parameter	# of breaks	Posterior Probability
Intercept	**0**	**0.9917**
	1	0.0079
	2	0.0004
	3	0.0000
	4	0.0000
Persistence	**0**	**0.9937**
	1	0.0061
	2	0.0001
	3	0.0000
	4	0.0000
Residual Variance	0	0.0000
	1	0.0205
	2	**0.9341**
	3	0.0433
	4	0.0020

Table 3: Posterior Distribution of Break Dates for the Most Preferred Model: # of Breaks (intercept, Pers., Var.)=(0,0,2)

	Break Date	90% Credibility Interval
First Structural Break Date	1970:Q2	1966:Q4 $\sim$ 1971:Q3
Second Structural Break Date	1985:Q2	1983:Q2 $\sim$ 1988:Q1

Note: Estimates of break dates are based on the highest posterior probability. Credibility intervals are based on the posterior distribution of break dates.

Table 4: Posterior Distributions for Parameters in the Most Preferred Model: # of Breaks (intercept, Pers., Var.)=(0,0,2)

Parameter	Regimes		
	Regime 1	Regime 2	Regime 3
α	0.2314		
	(0.1114)		
ρ	0.9212		
	(0.0372)		
σ^2	0.7031	2.2972	0.4204
	(0.1291)	(0.4688)	(0.0784)
p_{ii}	0.9849	0.9845	
	(0.0148)	(0.0157)	
γ_1	-0.3329		
	(0.0730)		
γ_2	-0.1873		
	(0.0746)		
γ_3	-0.1364		
	(0.0674)		

Note: Standard deviations are reported in parentheses. The posterior distributions are generated based on the most preferred autoregressive model of U.S. GDP Deflator inflation with two structural breaks in the residual variance (σ^2) and no break in the intercept (α) and the persistence (ρ) parameters:

$$\pi_t = \alpha + \rho\pi_{t-1} + \gamma_1\Delta\pi_{t-1} + \gamma_2\Delta\pi_{t-2} + \gamma_3\Delta\pi_{t-3} + e_t, \quad e_t \sim N(0, \sigma^2_{S_t}), \quad S_t = 1, 2, 3$$

and

$$Pr[S_t = i|S_{t-1} = i] = p_{i,i}, \text{ for } i = 1, 2; \quad Pr[S_t = 3|S_{t-1} = 3] = 1$$

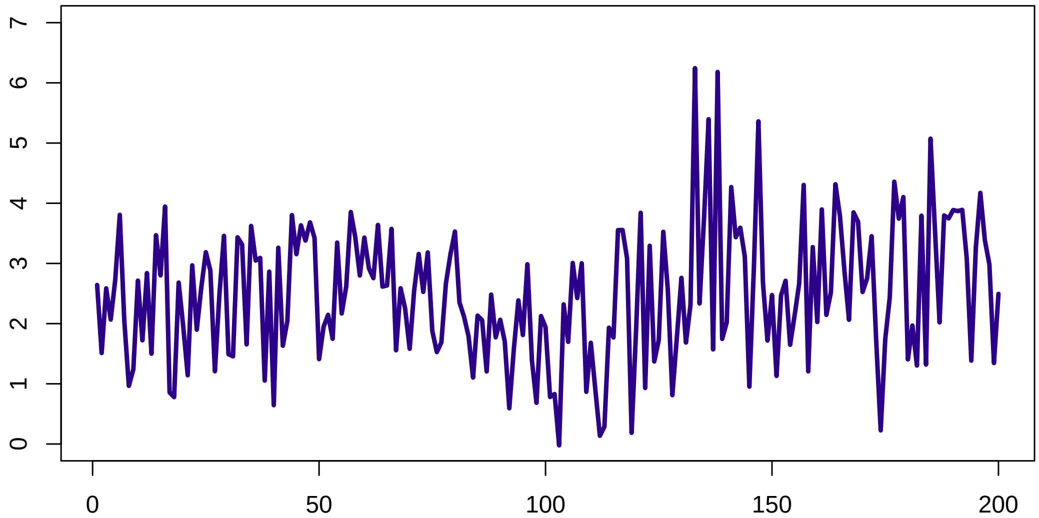
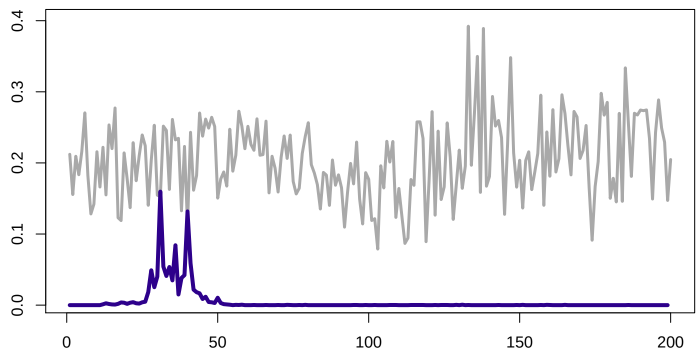
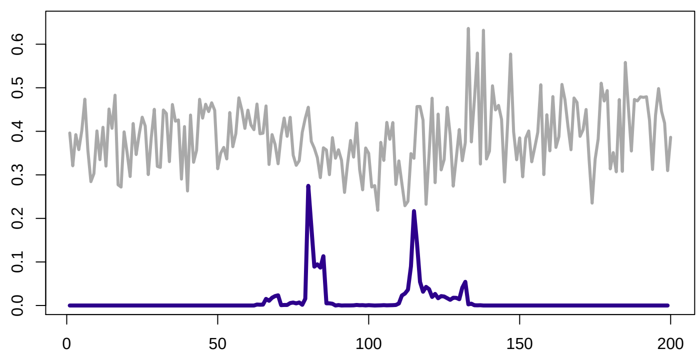


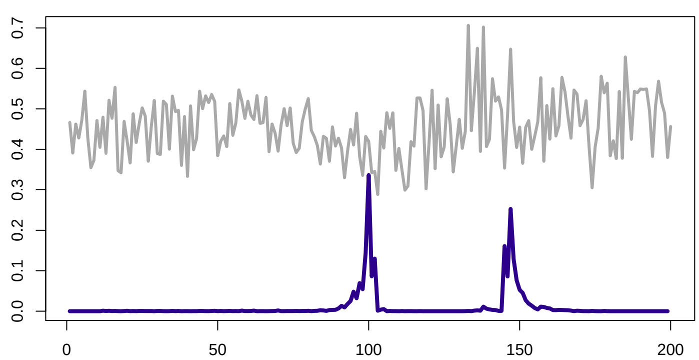
Figure 1: An Artificial Data Set



(a) α with true break point 40



(b) β with true break points 80 and 120



(c) σ^2 with true break points 100 and 150

Figure 2: Densities of Regimes Changes for an Artificial Data Set

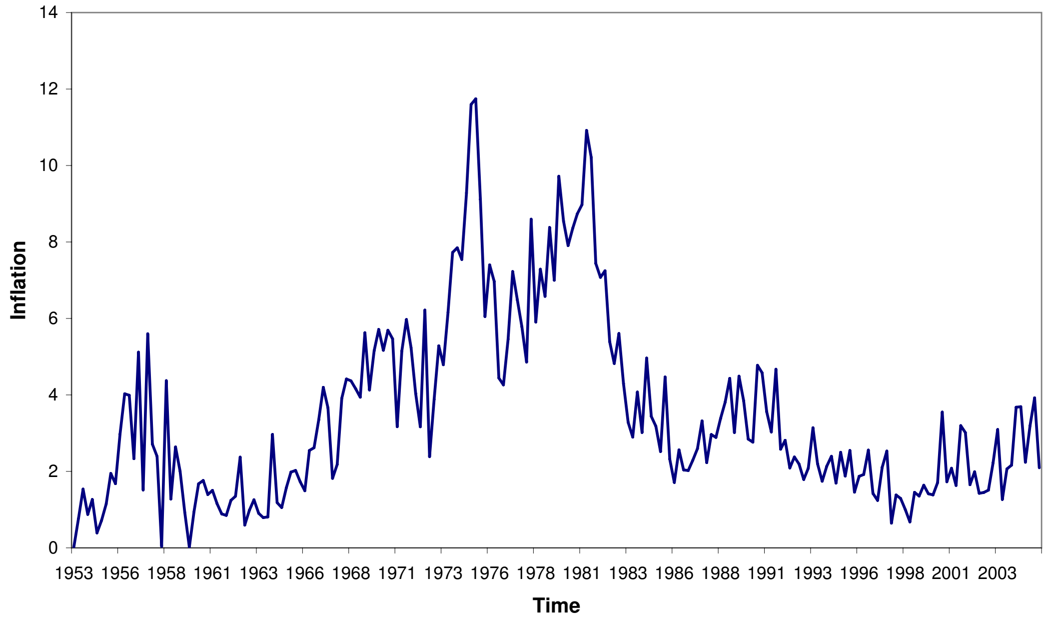
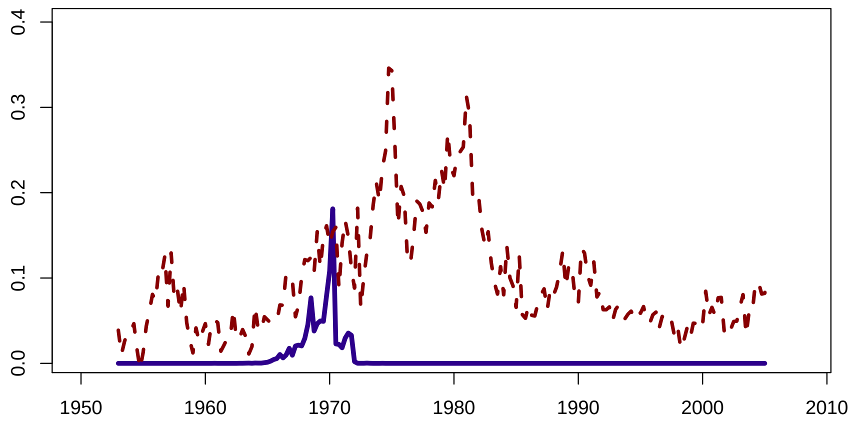
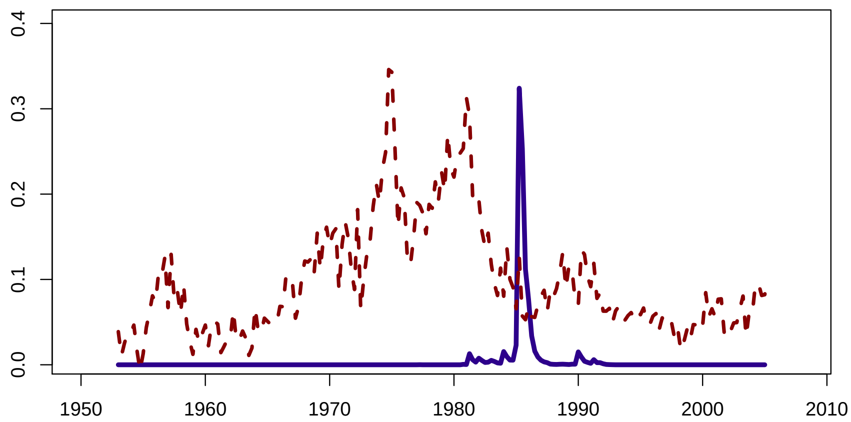


Figure 3: U.S. GDP Deflator Inflation (Quarterly percentage change at an annual rate): 1953:Q2-2005:Q2

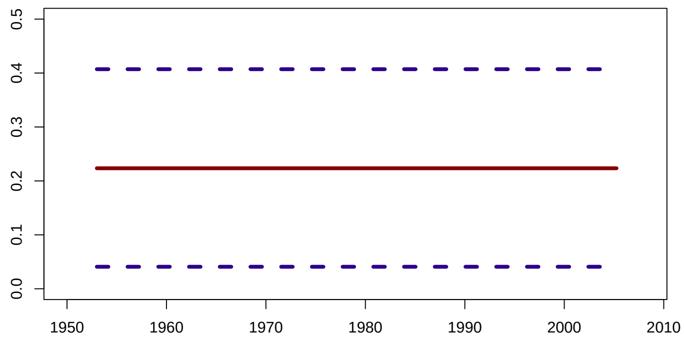


(a) First Break

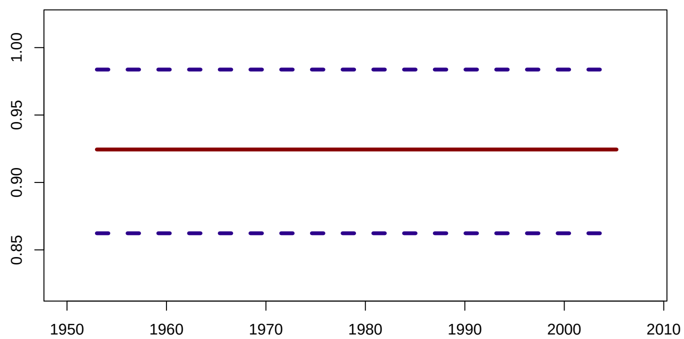


(b) Second Break

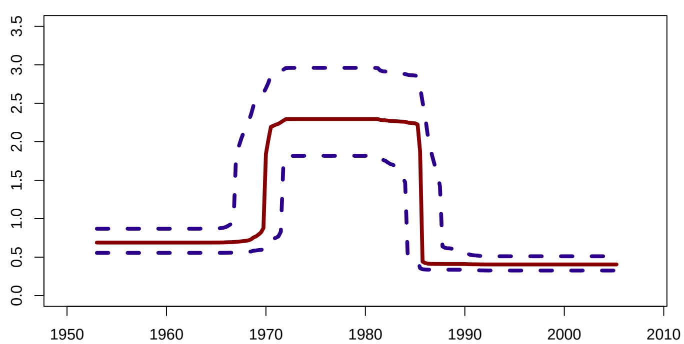
Figure 4: Posterior Probability of the Timing of Structural Break in Residual Variance



(a) Intercept



(b) Persistence



(c) Residual Variance

Figure 5: Posterior Distribution of Parameters