

A Bayesian Infinite Hidden Markov Structural Vector Autoregressive Model

Didier Nibbering* Richard Paap Michel van der Wel

Econometric Institute, Tinbergen Institute, Erasmus University Rotterdam

February 12, 2016

Abstract

We propose a Bayesian hidden Markov model with an infinite number of states to estimate time-varying parameters in a structural vector autoregressive (SVAR) model. The Markov structure allows for heterogeneity over time while accounting for state-persistence. By modelling the transition distribution as a Dirichlet process mixture model, parameters can vary over an infinite number of regimes. The Dirichlet process favours a parsimonious model without imposing restrictions on the parameter space. In an application on a small monetary SVAR we find evidence for heterogeneity in the volatility of monetary policy shocks and impulse responses to these shocks. These features can only be captured when we allow both the coefficient and the covariance matrix in the SVAR model to be time-varying.

Keywords: Time-Varying Parameter Structural Vector Autoregressive Model, Semi-parametric Bayesian Inference, Dirichlet Process Mixture Model, Hidden Markov Chain, Monetary Policy Analysis

JEL Classification: C11, C14, C32, C51, C54

*Correspondence to: Didier Nibbering, Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: nibbering@ese.eur.nl

1 Introduction

Typical macroeconomic studies use long data samples, which run from the 1950s and even further back, and it seems highly implausible that economic relations are fixed over such a long period of time. The regimes of high inflation and unemployment in the 1970s and early 1980s are well known stylized facts of macroeconomic data. Although after the eighties the volatility of inflation and unemployment became more stable, the onset of the global financial crisis in 2007 reminded us that we cannot rely on time-invariant macroeconomic models. The federal funds rate hit the zero lower bound at the end of 2008 and, in contrast to what most econometric models predicted, was still there till the end of 2015. With this in mind, it seems very unfortunate to ignore heterogeneity over time in modelling macroeconomic relations.

Vector autoregressive (VAR) models have become the standard approach for analyzing and quantifying the effects of policy shocks to the economy. However, the large sets of (lagged) variables considered in these models, increase the complexity of estimating time-varying parameters. Many researchers have studied feasible estimation methods. These methods, almost always parametric in nature, rely invariably on a set of model restrictions. Cogley and Sargent (2005) impose the simultaneous relations among variables to be time-invariant. Chib et al. (2006) assume a factor structure for the covariance matrix. Primiceri (2005) imposes parameters to evolve smoothly over time, by modelling the evolution of the parameters in the coefficient and covariance matrices as random walks. Koop and Korobilis (2013) make use of forgetting factors to model time-variation in the parameters. By using a semi-parametric approach, we do not have to impose any restrictions on the parameters to make estimation feasible.

In this paper we contribute to the literature of time-varying parameter vector autoregressive models (TVP-VARs) by proposing a semi-parametric Bayesian model which accounts for heterogeneous parameters with a clear and relatively simple estimation algorithm. Both the coefficients and the covariances of the innovations in the model are allowed to change over time. We apply the model to a small monthly monetary VAR of the U.S. economy including unemployment, inflation, and the interest rate. We discuss the structural interpretation of the estimation results in a structural VAR (SVAR). Moreover, we investigate which

set of time-varying parameters is crucial for appropriately interpreting the effects of policy shocks to the economy, by restricting alternately the coefficients and the covariances to be time-invariant.

We employ a hidden Markov chain in combination with a Dirichlet process to allow for time-varying parameters in an SVAR. The Dirichlet process mixture encourages parameters to cluster in regimes with similar values. This feature favours a parsimonious model, which is a huge advantage in modelling parameter heterogeneity in structural time series models which, due to large sets of variables, already suffer from the curse of dimensionality. Moreover, the Dirichlet process mixture allows for estimating the (potentially infinite) number of regimes, state assignments, and parameter estimates together in one estimation. The sticky hidden Markov structure accounts for state-persistence, as is often encountered in macroeconomic data.

The sticky Dirichlet process mixture hidden Markov SVAR builds upon work of Bauwens et al. (2015). They bring a semi-parametric Bayesian model, developed by Fox et al. (2011) for speaker diarization, to the univariate time series literature. They construct autoregressive moving average (ARMA) models with an infinite number of regimes, and empirically show superior forecast performance on macroeconomic time series relative to models with fixed parameters. By introducing these semi-parametric methods to the context of structural time series models, we construct a novel alternative to existing restrictive TVP-VARs.

Although the sticky Dirichlet process mixture hidden Markov model assigns parameters to clusters with similar parameter values, it should not be confused with models which try to capture time-variation with a finite number of switching regimes (see, for example, (Sims and Zha, 2006)). These discrete break models are able to model shifts in policy but cannot account for smooth changes. However, the parameters in the sticky Dirichlet process mixture hidden Markov SVAR are distributed over an infinite number of regimes, which enables modelling of abrupt breaks in parameter values together with smoother changes.

In our empirical application, we find large differences in responses to monetary policy shocks when we compare the economy before the millennium to the economy thereafter. Especially the initial reactions of inflation and the interest rate to a policy shock show different behavior. Moreover, we can endorse that unemployment and inflation became less volatile, but evidently, the onset of the

global financial crisis causes a large volatility shock. Finally, we find that only a fully flexible model, with both a time-varying coefficient and a time-varying covariance matrix, can appropriately model changes in interactions between economic variables and monetary policy.

The outline of this paper is as follows. Section 2 discusses the model specification and explains how we infer parameter estimates by Bayesian methods. Section 3 explains this empirical application, introduces the data, and discusses how we use the model for monetary policy analysis. Section 4 discusses and interprets the inference results. We conclude with a discussion in Section 5.

2 Methods

This section discusses the specification and estimation of the sticky Dirichlet process mixture hidden Markov SVAR. Section 2.1 introduces the baseline specification of the reduced form of a TVP-VAR. From here, we explain how we capture the parameter heterogeneity over time by constructing regimes with homogeneous parameter values. The regimes and parameter values are estimated by Bayesian methods. In Section 2.2, we specify the prior distributions and set up a Markov Chain Monte Carlo (MCMC) sampler. Moreover, we show how we can restrict sets of parameters to be time-invariant by making adjustments to the sampling algorithm.

2.1 Model Specification

Consider the reduced form of the vector autoregressive model

$$y_t = B_t x_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_t), \quad t = 1, \dots, T,$$

where y_t is a $p \times 1$ vector of observed endogenous time series, B_t is a $p \times k$ matrix with time-varying coefficients, and ε_t are heteroskedastic unobservable shocks with covariance matrix Σ_t . The $(k = 1 + pl) \times 1$ vector $x_t = [1, y'_{t-1}, \dots, y'_{t-l}]'$ includes an intercept and the endogenous variables up to lag l as explanatory variables.

Both the coefficient matrix B_t and the covariance matrix Σ_t in (1) contain time-varying parameters. Equivalently, we can say that the parameters in B_t and Σ_t vary over an infinite number of regimes, where the number of regimes equals

the number of time periods T when each time period has a different parameter value. We can write (1) as

$$y_t = B_{s_t} x_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_{s_t}), \quad (1)$$

where $s_{1:T} = \{s_1, \dots, s_T\}$ take integer values indicating the regime at time t . While there is strong evidence that the behavior of macroeconomic variables changes over time, it is implausible that the economy changes in each time period with probability one, in particular when we consider quarterly or monthly data. Therefore, we can specify a potentially more parsimonious model by modelling the transition probability of moving from one state to another.

We let the regime indicators $s_{1:T}$ follow a first-order Markov chain, where π_{ij} denote the transition probability of moving from state i to state j under the constraint that $\sum_{j=1}^{\infty} \pi_{ij} = 1$ for all i . So each state j has a state-specific transition distribution π_j ; $s_t \sim \pi_{s_{t-1}}$. Following the framework of Fox et al. (2011) we specify the transition distributions π_j as a Dirichlet process mixture model which accounts for persistence,

$$\begin{aligned} \pi_j | \alpha, \kappa, \beta &\sim DP\left(\alpha + \kappa, \frac{\alpha\beta + \kappa\delta_j}{\alpha + \kappa}\right), \\ \beta_k &= \nu_k \prod_{l=1}^{k-1} (1 - \nu_l), \quad \nu_k | \gamma \sim \text{Beta}(1, \gamma), \quad k = 1, 2, \dots, \end{aligned}$$

where $\{\beta_k\}_{k=1}^{\infty}$ is defined as a probability mass function on a countably infinite set in a stick-breaking construction and δ_j denotes a unit-mass measure concentrated at j . The Dirichlet process is denoted by $DP\left(\alpha + \kappa, \frac{\alpha\beta + \kappa\delta_j}{\alpha + \kappa}\right)$, where $\alpha + \kappa$ is a positive concentration parameter and the term $\frac{\alpha\beta + \kappa\delta_j}{\alpha + \kappa}$ a continuous base distribution. Since the expectation of the Dirichlet process equals the base distribution, states tend to have similar transition distributions; $E[\pi_{ij}] = \frac{\alpha\beta_j + \kappa 1(i=j)}{\alpha + \kappa}$. The notation $1(A)$ represents an indicator variable that equals one if event A occurs and zero otherwise. The specification of the base distribution takes state-persistence into account, which is often encountered in macroeconomic data. An amount $\kappa > 0$ is added to the j th component of $\alpha\beta$. In other words, the expected probability of self-transition is increased by an amount proportional to κ .

We can summarize the complete model specification in the following equations,

$$y_t = B_{s_t}x_t + \varepsilon_t, \quad (2)$$

$$\varepsilon_t | \Sigma_{s_t} \sim \mathcal{N}(0, \Sigma_{s_t}), \quad (3)$$

$$s_t | s_{t-1}, \{\pi_j\}_{j=1}^{\infty} \sim \pi_{s_{t-1}}, \quad (4)$$

$$\pi_j | \alpha, \kappa, \beta \sim DP(\alpha + \kappa, \frac{\alpha\beta + \kappa\delta_j}{\alpha + \kappa}), \quad (5)$$

$$\beta_k = \nu_k \prod_{l=1}^{k-1} (1 - \nu_l), \quad (6)$$

$$\nu_k | \gamma \sim Beta(1, \gamma), \quad k = 1, 2, \dots, \quad (7)$$

$$\{B_{s_t}, \Sigma_{s_t}\} \sim H_{\theta}(\Theta). \quad (8)$$

Equations (2) and (3) specify the reduced form of the time-varying vector autoregressive model, where the parameters $\theta = \{B_{s_t}, \Sigma_{s_t}\}$ vary over an infinite number of regimes. Within these regimes the parameters are assumed to be time-invariant but across regimes the parameters are allowed to be different. To retrieve the different regimes we use a hidden Markov chain in combination with a Dirichlet process mixture model. Equation (4) specifies the hidden Markov model by introducing a first-order Markov chain with transition probability matrix π . The transition probability distribution of π_j is specified as a sticky Dirichlet process mixture in (5), (6), and (7). The Dirichlet process allows us to let the number of regimes to be possibly infinite while we can estimate the number of regimes together with the parameter values. The stickiness of this model refers to the κ parameter which accounts for state-persistence. This parameter captures the persistence in macroeconomic data by controlling the probability that parameters remain constant between time periods. Equation (8) concludes with the base distribution H_{θ} , parameterized by the hyperparameters Θ .

In contrast to an autoregressive model for the law of motion, this model does not only accounts for smooth changes in parameter values, but also for abrupt shocks between time periods. Moreover, a regime switch in this model does not necessarily mean that all parameters change. For example, a regime switch can either be the result of a change in the parameters in the covariance (coefficient) matrix while the coefficient (covariance) parameters remain constant, or a change in all parameter values.

2.2 Bayesian Inference

To estimate the parameters θ we rely on the Markov Chain Monte Carlo (MCMC) algorithm for the sticky Dirichlet process mixture hidden Markov model derived by Fox et al. (2011) in an application on the speaker diarization problem. Bauwens et al. (2015) apply a variant of this sampler in an econometric time series context; the estimation of univariate ARMA models.

Although there are sampling algorithms that can deal with an infinite number of regimes (which are also derived and discussed by Fox et al. (2011)), these algorithms suffer in general from slow mixing rates. Therefore, we follow Bauwens et al. (2015) who opt for a sampler which truncates the number of possible states to a fixed degree L , the so called degree L weak limit approximation (Ishwaran and Zarepour, 2002). When L is large enough, the error is negligible (Fox et al., 2011; Bauwens et al., 2015). The degree L weak limit approximation fosters models with less than L regimes while allowing for new regimes, bounded by L , when new data are observed.

The degree L weak limit approximation induces finite Dirichlet process priors on β and π_j ,

$$\begin{aligned}\beta|\gamma &\sim DP(\gamma/L, \dots, \gamma/L), \\ \pi_j|\alpha, \beta, \kappa &\sim DP(\alpha\beta_1, \dots, \alpha\beta_j + \kappa, \dots, \alpha\beta_L).\end{aligned}$$

We let the data determine the number of states and the degree of state-persistence, by treating the hyperparameters of the transition distributions $\{\gamma, \alpha, \kappa\}$ as unknown. We place priors on these parameters and resample them in each iteration of the sampler. The priors of the hyperparameters are

$$\alpha + \kappa \sim \text{Gamma}(a_\alpha, b_\alpha), \quad \gamma \sim \text{Gamma}(a_\gamma, b_\gamma), \quad \rho = \frac{\kappa}{\alpha + \kappa} \sim \text{Beta}(c_\rho, d_\rho).$$

The prior on the parameters $\theta_{s_t} = \{B_{s_t}, \Sigma_{s_t}\}$ is a Normal-inverse-Wishart,

$$\text{vec}(B_{s_t})|\Sigma_{s_t}, \Theta \sim \mathcal{N}(\text{vec}(b_B), V_B \otimes \Sigma_{s_t}), \quad \Sigma_{s_t}|\Theta \sim \mathcal{IW}(\nu_\Sigma, S_\Sigma),$$

where the $\text{vec}(A)$ operator stacks the columns of matrix A and Θ is the collection of hyperparameters for θ , $\{b_B, V_B, \nu_\Sigma, S_\Sigma\}$.

2.2.1 Sample Algorithm

Since Fox et al. (2011) derived the sample algorithm in detail, we present in this section only the resulting sampling steps:

Step 1. Set the truncation level L of possible hidden Markov states. Sample an initial draw for the hyperparameters of the transition distributions from their priors and do the same for the hyperparameters in the base distribution H_θ . Initialize the transition distributions β and π_j by drawing from their L -dimensional Dirichlet priors.

Step 2. Sample the regime indicators $s_{1:T}$ using the forward-backward procedure (Rabiner, 1989).

(a) First we work sequentially backwards in time. For each $k = 1, \dots, L$, $m_{T+1,T}(k) = 1$ and

$$m_{t,t-1}(k) = \sum_{j=1}^L \pi_k(j) \mathcal{N}(y_t; B_j x_t, \Sigma_j) m_{t+1,t}(j), \quad t = T, \dots, 2,$$

where $\mathcal{N}(y, \mu, \Sigma)$ denotes the probability density function of the multivariate Normal distribution with mean μ and covariance matrix Σ .

(b) Second, we work sequentially forward in time and initialize the number of transitions from state j to k observed in the state vector $s_{1:T}$, $n_{jk} = 0$ with $j, k = 1, \dots, L$. For each $k = 1, \dots, L$ compute

$$f_k(y_t) = \pi_{s_{t-1}}(k) \mathcal{N}(y_t; B_k x_t, \Sigma_k) m_{t+1,t}(k), \quad t = 1, \dots, T,$$

sample the regime indicators,

$$s_t \sim \sum_{k=1}^L f_k(y_t) 1(s_t = k), \quad t = 1, \dots, T,$$

and increment n_{s_{t-1}, s_t} .

Step 3. Sample auxiliary variables m , w , and \bar{m} .

(a) For $j = 1, \dots, L$ and $k = 1, \dots, L$ set $m_{jk} = 0$. For $i = 1, \dots, n_{jk}$ sample $x_i \sim \text{Bernoulli}\left(\frac{\alpha\beta_k + \kappa 1(j=k)}{i-1 + \alpha\beta_k + \kappa 1(j=k)}\right)$ and increment m_{jk} if $x_i = 1$.

- (b) For $j = 1, \dots, L$ sample $w_j \sim \text{Binomial}(m_{jj}, \frac{\rho}{\rho + \beta_j(1-\rho)})$. Set $\bar{m}_{jk} = m_{jk}$ if $j \neq k$ and $\bar{m}_{jk} = m_{jk} - w_j$ if $j = k$.

Step 4. Sample the global transition distribution

$$\beta \sim \text{Dir}(\gamma/L + \sum_j \bar{m}_{j1}, \dots, \gamma/L + \sum_j \bar{m}_{jL}).$$

Step 5. Sample the transition distribution π . For $j = 1, \dots, L$ sample

$$\pi_j \sim \text{Dir}(\alpha\beta_1 + n_{j1}, \dots, \alpha\beta_j + \kappa + n_{jj}, \dots, \alpha\beta_L + n_{jL}).$$

Step 6. Sample the regime parameters θ for $j = 1, \dots, L$. Let x_j be the $t_j \times k$ matrix with rows $x_{s_t=j}$ and t_j is the number of observations in state s_t . We define y_j as a $p \times t_j$ matrix. Sampling the parameters now boils down to

$$\begin{aligned} \bar{B} &= (x'_j x_j + V_B^{-1}), & \bar{b} &= (y_j x_j + b_B V_B^{-1}) \bar{B}^{-1}, \\ \bar{S} &= S_\Sigma + (y'_j - x'_j \bar{b}') (y'_j - x'_j \bar{b}') + (\bar{b} - b_B) V_B^{-1} (\bar{b} - b_B)', \\ \Sigma_j | y_j, \Theta &\sim \mathcal{IW}(\nu_\Sigma + t_j, \bar{S}), & \text{vec}(B_j) | y_j, \Sigma_j, \Theta &\sim \mathcal{N}(\text{vec}(\bar{b}), \bar{B}^{-1} \otimes \Sigma_j). \end{aligned}$$

Step 7. Sample the hyperparameters of the transition distributions γ , α , and κ .

- (a) Sample auxiliary variables $r_j \sim \text{Beta}(\alpha + \kappa + 1, \sum_k n_{jk})$ and $s_j \sim \text{Bernoulli}(\frac{\sum_k n_{jk}}{\sum_k n_{jk} + \alpha + \kappa})$ for $j = 1, \dots, L$.
Sample $\alpha + \kappa \sim \text{Gamma}(a_\alpha + \sum_j \sum_k m_{jk} - \sum_j s_j, b_\alpha - \sum_j \log r_j)$.
- (b) Sample $\rho = \frac{\kappa}{\alpha + \kappa} \sim \text{Beta}(c_\rho + \sum_j w_j, d_\rho + \sum_j \sum_k m_{jk} - \sum_j w_j)$.
- (c) Sample auxiliary variables $r \sim \text{Beta}(\gamma + 1, \sum_j \sum_k \bar{m}_{jk})$ and $s \sim \text{Bernoulli}(\frac{\sum_j \sum_k \bar{m}_{jk}}{\sum_j \sum_k \bar{m}_{jk} + \gamma})$. Compute $\bar{K} = \sum_k 1(\sum_j \bar{m}_{jk} > 0)$ and Sample $\gamma \sim \text{Gamma}(a_\gamma + \bar{K} - s, b_\gamma - \log r)$.

Step 8. Go to step 2.

2.2.2 Parameter Restrictions

Since it is not clear whether time variation in dependencies in the economy are caused by the relations in the coefficient matrix B or the covariances between the variables in Σ , the model in (1) allows both the parameters in B as in Σ to vary

over time. This flexible model specification may be over-parametrized in case the source of time variation is restricted to one of the two parameter matrices. A more parsimonious model may avoid the pitfall of over-fitting by restricting either the coefficient matrix or the covariance matrix to be constant over time. We can estimate the parameters in the restricted models with only slight modifications to the sampler in Subsection 2.2. In practice, we adjust Step 6 of the sampler and add an extra step to sample the fixed parameters outside the structure of the mixture model.

Instead of sampling the regime parameters as in step 6 of the sample algorithm, we sample B_j when the covariance matrix is fixed as

$$\begin{aligned}\bar{B} &= (x'_j x_j + V_B^{-1}), & \bar{b} &= (y_j x_j + b_B V_B^{-1}) \bar{B}^{-1}, \\ \text{vec}(B_j) | y_j, \Sigma, \Theta &\sim \mathcal{N}(\text{vec}(\bar{b}), \bar{B}^{-1} \otimes \Sigma),\end{aligned}$$

and sampling Σ_j when the coefficient matrix is fixed involves

$$\begin{aligned}\bar{S} &= S_\Sigma + (y'_j - x_j B'_j)(y'_j - x_j B'_j), \\ \Sigma_j | y_j, \Theta &\sim \mathcal{IW}(\nu_\Sigma + t_j, \bar{S}).\end{aligned}$$

After the seventh step, when the state assignments in the mixture model are settled down for the current iteration, we sample the time-invariant parameter matrices. In the model with a restricted covariance matrix we compute $\varepsilon = (\varepsilon'_1, \dots, \varepsilon'_T)'$ where $\varepsilon_t = y'_t - x_t B'_{s_t}$ and add the sampling step

$$\bar{S} = S_\Sigma + \varepsilon' \varepsilon, \quad \Sigma_j | y, \Theta \sim \mathcal{IW}(\nu_\Sigma + T, \bar{S}).$$

For the model with a restricted coefficient matrix we compute $\tilde{x} = (\tilde{x}'_1, \dots, \tilde{x}'_T)'$, where $\tilde{x}_t = I_p \otimes x_t$, $t = 1, \dots, T$, and I_p is the identity matrix of dimension p and, $\tilde{\Sigma} = \text{diag}(\Sigma_{s_1}, \dots, \Sigma_{s_T})$, and $\tilde{y} = (y'_1, \dots, y'_T)$. Now we can perform the sampling steps

$$\begin{aligned}\bar{B} &= (\tilde{x}' \tilde{\Sigma}^{-1} \tilde{x} + (I_p \otimes V_B)^{-1})^{-1}, & \bar{b} &= \bar{B}(\tilde{x}' \tilde{\Sigma}^{-1} \tilde{y} + (I_p \otimes V_B)^{-1} \text{vec}(b_B)), \\ \text{vec}(B) | \tilde{y}, \tilde{\Sigma}, \Theta &\sim \mathcal{N}(\text{vec}(\bar{b}), \bar{B}).\end{aligned}$$

3 Application

We apply the model on a small monetary VAR of the U.S. economy consisting of the unemployment rate, inflation rate and the federal funds rate. This section

first introduces the data which is used to estimate the parameters in the reduced form in (1). Second, we show how the structural parameters of the small monetary VAR can be inferred from the reduced form parameter estimates and discuss the impulse response functions.

3.1 Data

We use three macroeconomic time series of the U.S. economy, the unemployment rate, inflation rate, and the federal funds rate, to construct a monetary VAR. The monthly data runs over the sample period from July 1954 to November 2015, the period over which the three series were all monthly available at the time we requested the data. We use 1200 times the first difference of the logarithm of the seasonally adjusted consumer price index for all urban consumers as indicator for the inflation rate. Since we include two lags in the model and take first differences of the inflation series, we consider 734 monthly time periods starting from October 1954. The seasonally adjusted civilian unemployment rate and the effective federal funds rate in percentages are untransformed included in the model. All data are downloaded from the website of the Federal Reserve Bank of St. Louis.

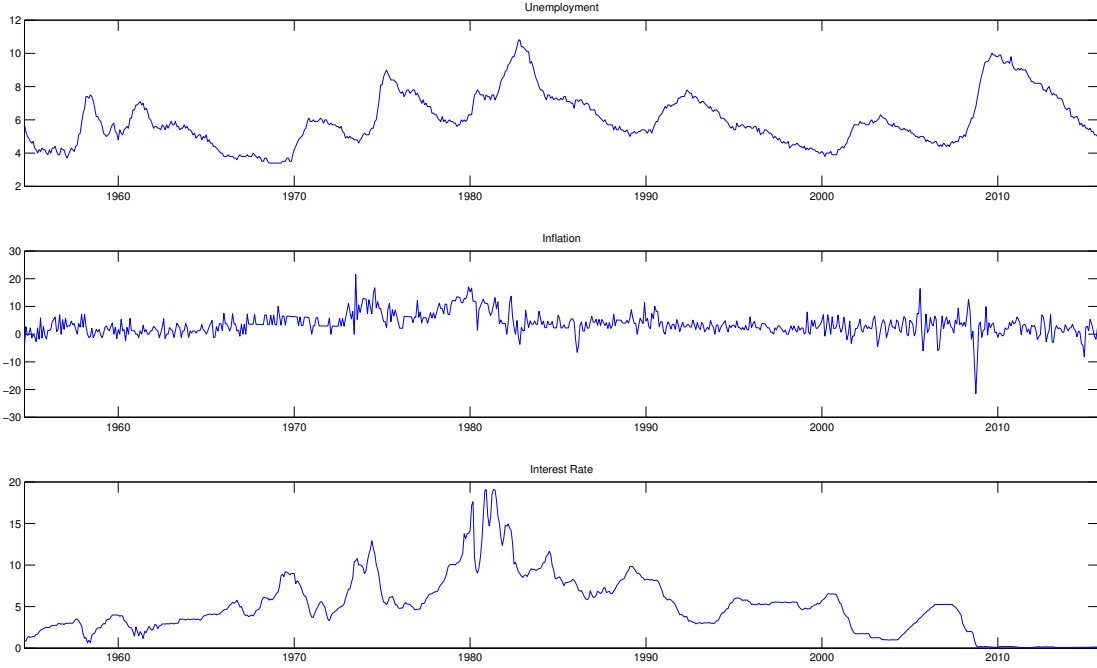
Figure 1 shows the (transformed) data series as included in the model. The simultaneous jumps in the three time series around 2007 directly attract the attention. As soon as the global financial crisis showed up, the unemployment rate rapidly increases, the price level collapses, and the federal funds rate (also indicated as the interest rate throughout this paper) drops to zero. However, where the unemployment and inflation numbers seem to recover and now show similar behaviour as before the crisis, the interest rate is still stuck to the zero lower bound.

We follow Fox et al. (2011) in the parameter values in the prior distributions on the hyperparameters of the transition distribution. In general, we opt for non-informative prior distributions. Table 1 shows the prior parameter values of the model.

3.2 Structural VAR

To discuss policy effects while incorporating changing behavior of the economy over time, we have to consider the structural form of the TVP-VAR. However, the

Figure 1: Time Series Small Monetary VAR U.S. Economy



This figure shows the (transformed) monthly data series as included in the small monetary VAR of the U.S. economy; unemployment rate, inflation rate, and interest rate. The inflation series represent 1200 times the first difference of the logarithm of the consumer price index for all urban consumers. The interest rate denotes the effective federal funds rate in percentages. The sample period runs from July 1954 to November 2015.

model in (1) and the estimation algorithm in Subsection 2.2 deal with the reduced form. Here we show how the structural form can be derived from the reduced form.

Consider the structural form of the time-varying vector autoregressive model

$$A_{yt}y_t = A_{xt}x_t + u_t, \quad u_t \sim \mathcal{N}(0, \Omega_t), \quad t = 1, \dots, T, \quad (9)$$

where the diagonal elements of A_{yt} equal one and Ω_t is a diagonal matrix. We can write the coefficient matrix in the reduced form as a function of the coefficient matrices in the structural form as $B_t = A_{yt}^{-1}A_{xt}$ and for the reduced form covariance matrix holds $\Sigma_t = A_{yt}^{-1}\Omega_t A_{yt}^{-1'}$. Since we cannot directly obtain the structural form parameter matrices from these equations, we need $p(p-1)/2$ restrictions to identify

Table 1: Prior Parameters

a_α	b_α	a_γ	b_γ	c_ρ	d_ρ	b_B	V_B	ν_Σ	S_Σ
1	10	1	10	10	1	$0_{p \times k}$	I_k	p	I_p

This table shows the parameters of the priors discussed in Section 2.2. $0_{p \times q}$ represents a zero matrix of size $p \times q$ and I_q is the identity matrix of dimension q .

the structural parameters.

For sake of simplicity, we opt for a simple Cholesky decomposition in our application on a small monetary VAR. The Cholesky decomposition on Σ_t produces an upper triangular matrix $\Sigma_t^{-\frac{1}{2}}$ which equals $\Omega_t^{-\frac{1}{2}}A_{yt}^{-1'}$ and has a main diagonal with the standard deviations of all structural shocks. So from here we can easily construct Ω_t , $A_{yt} = \Sigma_t^{\frac{1}{2}}\Omega_t^{-\frac{1}{2}}$ and $A_{xt} = A_{yt}B_t$.

Sims (1980) proposed the Cholesky decomposition to identify the structural form parameters to analyse monetary policy and discussed the implications of the Cholesky ordering. In our small monetary VAR we order the variables in y_t as {unemployment, inflation, interest rate}. This ordering implies that unemployment is exogenous, so it is only contemporaneously affected by its own shocks. Inflation is also assumed to be contemporaneously affected by its own shocks, but it is also affected by shocks in unemployment. However, inflation is not contemporaneously influenced by shocks in the interest rate. The interest rate itself, is affected by all variables. The interest rate can be moved by its own shock, a shock in unemployment, and/or a shock in inflation.

3.3 Impulse Response Functions

Policy effects can be analyzed by the response of current and future values of the variables included in the VAR, to a one unit increase in the current value of the policy variable error. When we let u_0 represent the shock at time t , that is u_0 is a $p \times 1$ vector of zeros and a one at the index corresponding to the policy variable, the impulse response function of a structural VAR with, for instance, two lags is

$$z_0^t = A_{yt}^{-1}u_0, \quad z_1^t = A_{y,t+1}^{-1}A_{x1,t+1}z_0^t, \quad z_s^t = A_{ys}^{-1}A_{x1s}z_{s-1}^t + A_{ys}^{-1}A_{x2s}z_{s-2}^t,$$

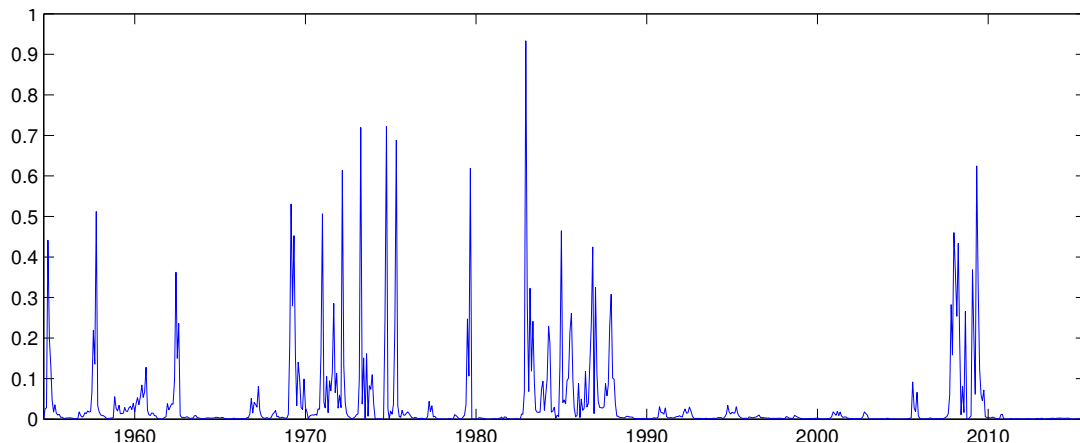
where $s > 1$ and $A_{xt} = (A_{x0t}, A_{x1t}, A_{x2t})$ consists of the coefficient matrices corresponding to the intercept, variables lagged one time period, and variables lagged two time periods, respectively.

4 Results

This section discusses the estimation results of the sticky Dirichlet process mixture hidden Markov SVAR model for monetary policy variables of the U.S. economy. The model is estimated with both a time-varying coefficient matrix and a time-varying covariance matrix. Estimation results of restricted models are shown in the Appendix. All results are based on 10,000 iterations of the MCMC sampler after 10,000 burn-in iterations, and the number of possible states is truncated at 10 in the degree L weak limit approximation.

After convergence, each iteration of the sampler finds three different regimes. Figure 2 shows the estimated probability of a regime switch for each time period. We find compelling evidence of instability in the parameter estimates. Most of the breaks are detected before 1990. Thereafter, there is a stable period which is followed by higher break probabilities corresponding to the global financial crisis starting in 2007.

Figure 2: Posterior Break Probabilities



This figure shows for each time period the estimated probability of switching regime in the current period compared to the previous period, computed from the posterior samples.

Since a TVP-SVAR consists of a large amount of parameters, even in a small model with only three variables and two lags, it is cumbersome to present estimation diagnostics of all parameters. Moreover, it is difficult to give an economic interpretation to each posterior distribution. Therefore, we show the posterior results for the structural covariance matrix, which we can give a direct structural interpretation, and we report impulse response functions which summarize the economic implications of the estimated SVAR coefficients.

4.1 Structural Variance

We define monetary policy shocks as interest rate responses to variables other than unemployment and inflation. The changes in relative importance of these shocks over time can be displayed by the time-varying variance of the monetary policy shock. The third panel in Figure 3 shows the time-varying posterior mean of the variance of monetary policy shocks. The 5th and 95th percentiles of the posterior distributions provide posterior evidence that there is variation over time.

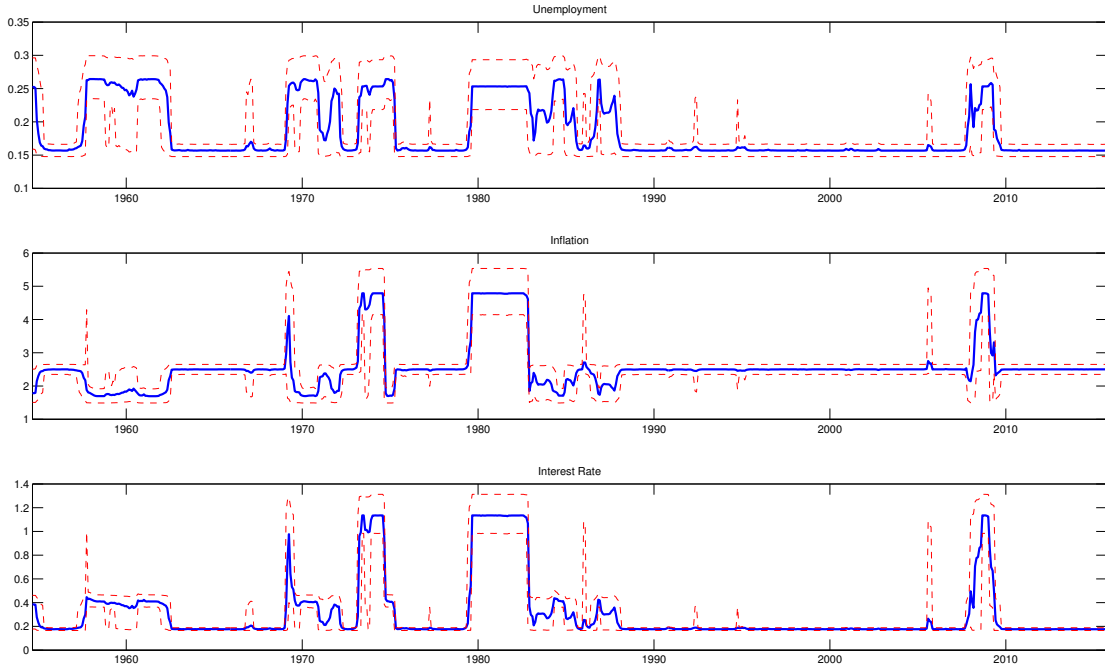
The first thing to note beside some peaks in variance in the early sixties and mid seventies, is the long period of high variance running from 1979 to 1983. This feature is well-known and can be attributed to a period with deviant monetary policy (Primiceri, 2005). After this high variance regime, the changes in variance are quite modest with one striking exception; the start of the global financial crisis provokes a large jump in volatility. However, the variance went rapidly back to a low level and has not moved since then. This is not surprising, since the interest rate never left its zero lower bound in this period.

The second panel of Figure 3 shows that the variance of the inflation shocks follow a pattern similar to the variance of monetary policy shocks. The percentiles of the posterior distribution also clearly supports modelling time variation. The variance of the shocks to unemployment, as showed in the first panel of Figure 3, seems to be more volatile over time. After the eighties it behaves more calmly, with a regime with higher variance between 2007 and 2010.

4.2 Impulse Response Functions

Figure 4 shows the impulse response functions of the SVAR to a monetary policy shock. The functions trace out the effect of the structural shocks over a time path

Figure 3: Posterior Means of the Structural Variance Parameters



This figure shows the time-varying posterior means (blue solid line) of the structural variance parameters together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The panels show from top to bottom the variance of the residuals in the unemployment equation, inflation equation, and interest equation, respectively.

of five years for each variable, conditional on the estimated states. Since these time paths are different for each date the shock hits the system, (and we do not have space to provide results for each possible shock date), we arbitrarily choose four different moments at which a monetary policy shock hits the system; mid 1970, 1985, 2000, and 2010. The posterior means of the impulse response functions are accompanied by the 5th and 95th percentiles of their posterior distributions.

The first column of Figure 4 shows the impulse responses of unemployment to a monetary policy shock for the different shock dates. We can hardly find any differences in the shape of the response functions and conclude that these responses do not vary much over time. We draw a different conclusion when we consider the impulse responses of inflation. The responses to shocks in 1970 and 1985 show

similar behavior, but the initial reaction to a monetary policy shock seems to be reversed in 2000 and 2010. First, inflation responded positively to an increase in inflation, then declined and converged to the initial level. After the millennium, the initial response is negative, subsequently increases, and converges to the old level.

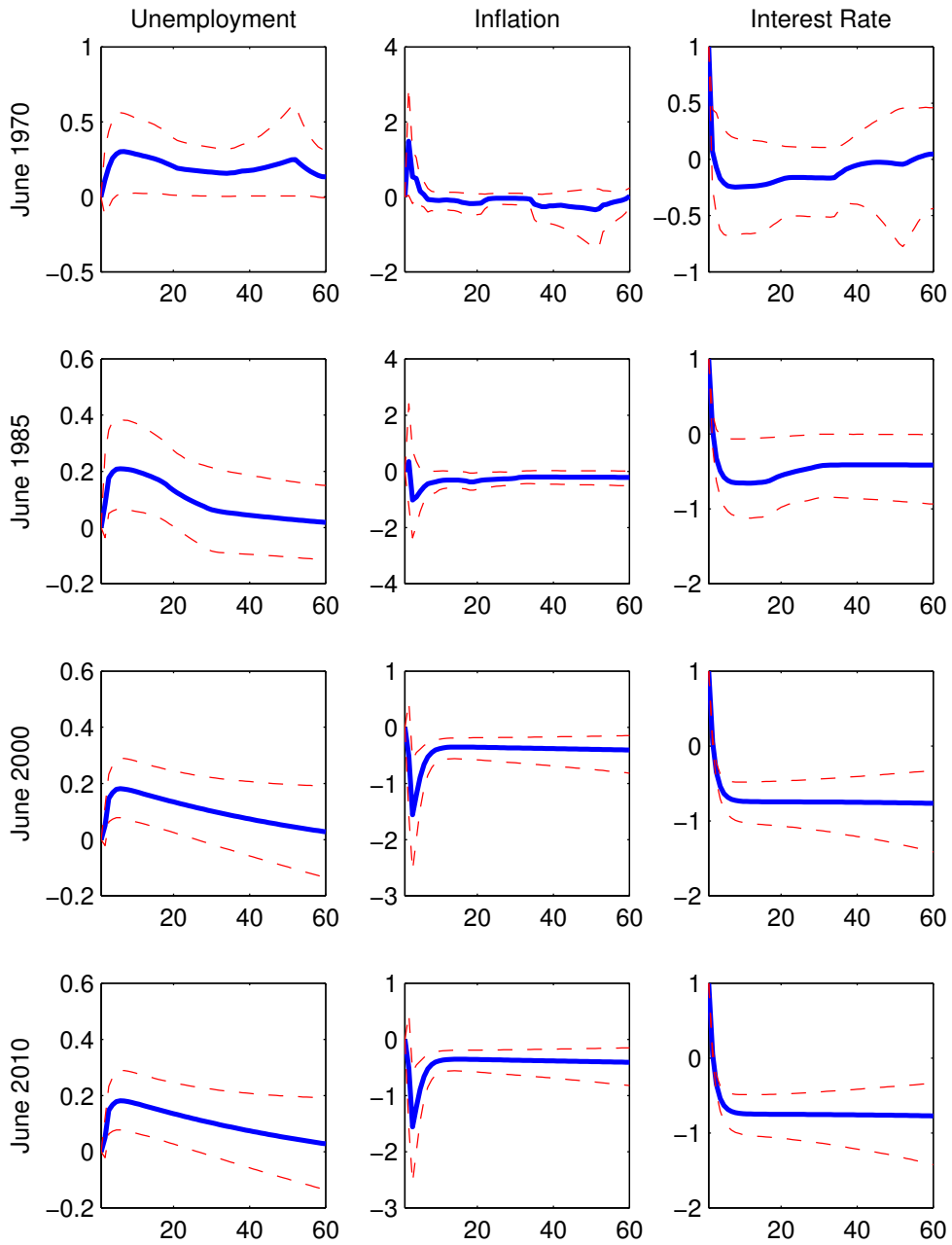
The response of the interest rate to a shock to itself shows also differences over time. A shock in 1970 or 1985 has a small impact on the interest rate after five years, where there is a clear negative long-term effect of a shock in 2000 or 2010. Based on impulse response function analyses in a sample period before 2000, Primiceri (2005) concludes that there is no evidence of nonlinearities in the responses of the economy to monetary policy shocks. We find here the opposite, by comparing impulse responses to shocks in 1970 and 1985 to responses to shocks in 2000 and 2010. In particular, the behavior of inflation and interest rate as reaction to a monetary policy shock is different before and after the millennium.

4.3 Parameter Restrictions

The results discussed above are estimated in a model with both a time-varying coefficient matrix and a time-varying covariance matrix. We can examine whether this high degree of flexibility is necessary to appropriately model the structural relations in the economy, by restricting the coefficient and/or the covariance matrix in (1) to be time invariant. Appendix A shows the results of a model with a time-invariant coefficient matrix, Appendix B results based on a model with a fixed covariance matrix, and Appendix C results of a complete time-invariant model.

Figure A1 shows that fixing the coefficient matrix over time, results in a much more volatile estimate for the variance parameter of the residuals in the inflation and interest rate equation. Apparently, when the coefficients cannot account for time variation in the transmission mechanism of shocks, the variance estimates have to account for these nonlinearities and become more volatile. We find less similarity in the time-varying pattern of the variance estimates between the different equations. The impact of the financial crisis in 2007 on the variances is estimated much smaller in the interest equation. A time-invariant coefficient matrix results with our identification scheme to time-invariant impulse response functions, as Figure A2 shows. We saw in Figure 4 that impulse responses to monetary policy

Figure 4: Impulse Response Functions



This figure shows the posterior means (blue solid line) of the impulse response functions together with the 5th and 95th percentiles of their posterior distributions (red dotted line). From top to bottom we have the impulse responses over a period of 60 months to monetary policy shocks in June 1970, June 1985, June 2000, and June 2010, respectively. The columns correspond to the responses in the unemployment equation, inflation equation, and interest equation, respectively. The monetary policy shock is defined as a one unit shock to interest rate.

shocks show different behavior when we change the timing of the shocks.

As a result of our identification restrictions, the estimated structural variances are fixed over time when we model the reduced form covariance matrix time-invariant. Figure B1 shows the posterior means of the time-invariant structural variance parameters. Although the impulse response functions are still allowed to vary over time, Figure B2 shows that monetary policy shocks at different dates result in almost identical impulse response functions. Thus, heterogeneity in responses is not only driven by time variation in the coefficient parameters. Finally, we consider a complete time-invariant model, which provides constant structural variance estimates and impulse response functions independent of the time of the shock. Figure C1 and Figure C2 show these structural variance estimates and impulse response, respectively. The posterior means of the structural variances of this model and the model with only a fixed covariance matrix are close to each other, except for the interest rate. A time-varying coefficient matrix results in a posterior mean of almost half the posterior mean in the complete time-invariant model. The impulse response function of the time-invariant model and the model with only a time-invariant coefficient matrix are much alike.

5 Conclusion

In this paper we propose a new method to estimate time-varying parameters in an SVAR. To avoid the curse of dimensionality, we opt for a semi-parametric approach. The Dirichlet process mixture model encourages estimation of a parsimonious model by clustering parameter values over time, without restricting the parameter space. To accommodate for persistence in macroeconomic data, we impose the Dirichlet process mixture on the transition probabilities in a hidden Markov-switching framework. Parameter values are assigned to a possibly infinite number of states, with a potentially increased probability of self-transition. Except from the degree L weak limit approximation, which comes with negligible costs, the estimation algorithm of the model does not impose any restrictions on or (linear) approximations to the parameters.

We apply the model on a small monetary SVAR consisting of the monthly unemployment rate, inflation rate, and interest rate of the U.S. economy. Although

we use a relatively simple estimation algorithm, we find economically interpretable results which are comparable to results based on similar SVAR's in earlier research. For instance, results show a high volatility regime in the early eighties corresponding to a period of deviant monetary policy and a volatility spike associated with the onset of the global financial crisis in 2007. Moreover, we present new findings regarding impulse responses to a monetary policy shock, which show different behavior after the millennium in comparison with the sample period before the millennium.

The results suggest that the semi-parametric Bayesian framework is a promising alternative for parametric approaches to TVP-VAR modelling. We find new evidence for heterogeneity over time, in both the volatility of monetary policy shocks and the impulse responses to a monetary policy shock. The fact that models with either a time-invariant coefficient or covariance matrix cannot capture both of these features, supports the need for time-varying parameters.

Based on this study we provide recommendations for further research. The drop of the interest rate to the zero lower bound is a new phenomenon in the U.S. economy which is an interesting research direction. Extending the SVAR with more variables, for example financial indicators, can probably shine some light on the effects of the zero lower bound and unconventional monetary policies to the economy. In general, we encourage to apply the method to large TVP-VARs. Because of the parsimony of the model, we expect the computational burden to decrease and make inference on larger sets of variables feasible. Due to the parsimony of the model, a forecast exercise can be another promising research direction. Furthermore, future research can perhaps investigate whether extending and controlling the lag length and prior sensitivity analysis can be optimized in a Bayesian semi-parametric context. Finally, sampling algorithms can be examined which guarantee a stationary development of the time-varying parameter values.

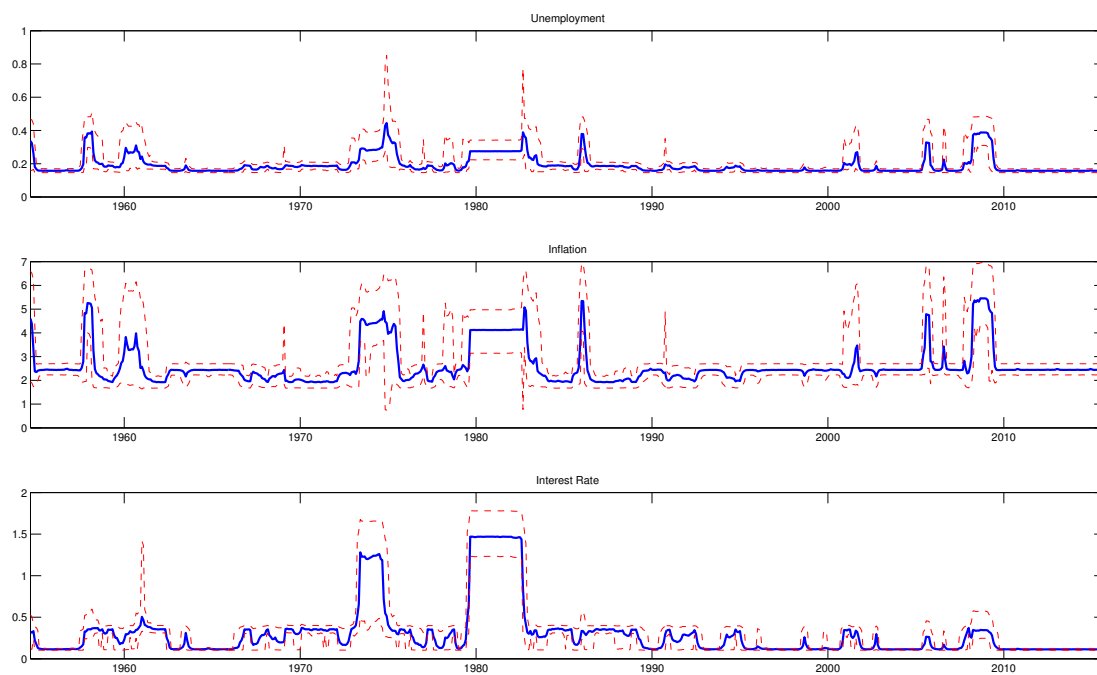
References

Bauwens, L., Carpentier, J.-F., and Dufays, A. Autoregressive moving average infinite hidden markov-switching models. *Journal of Business & Economic Statistics*, (just-accepted), 2015.

- Chib, S., Nardari, F., and Shephard, N. Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics*, 134(2):341–371, 2006.
- Cogley, T. and Sargent, T. J. Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302, 2005.
- Fox, E. B., Sudderth, E. B., Jordan, M. I., and Willsky, A. S. A sticky hdp-hmm with application to speaker diarization. *The Annals of Applied Statistics*, pages 1020–1056, 2011.
- Ishwaran, H. and Zarepour, M. Exact and approximate sum representations for the dirichlet process. *Canadian Journal of Statistics*, 30(2):269–283, 2002.
- Koop, G. and Korobilis, D. Large time-varying parameter vars. *Journal of Econometrics*, 177(2):185–198, 2013.
- Primiceri, G. E. Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852, 2005.
- Rabiner, L. R. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- Sims, C. A. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, pages 1–48, 1980.
- Sims, C. A. and Zha, T. Were there regime switches in us monetary policy? *The American Economic Review*, pages 54–81, 2006.

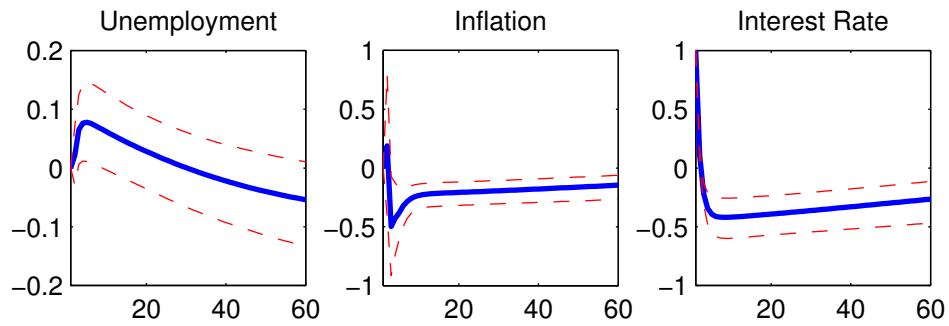
A Results Time-Invariant Coefficient Matrix

Figure A1: Posterior Means of the Structural Variance Parameters



This figure shows the time-varying posterior means (blue solid line) of the structural variance parameters together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The panels show from top to bottom the variance of the residuals in the unemployment equation, inflation equation, and interest equation, respectively.

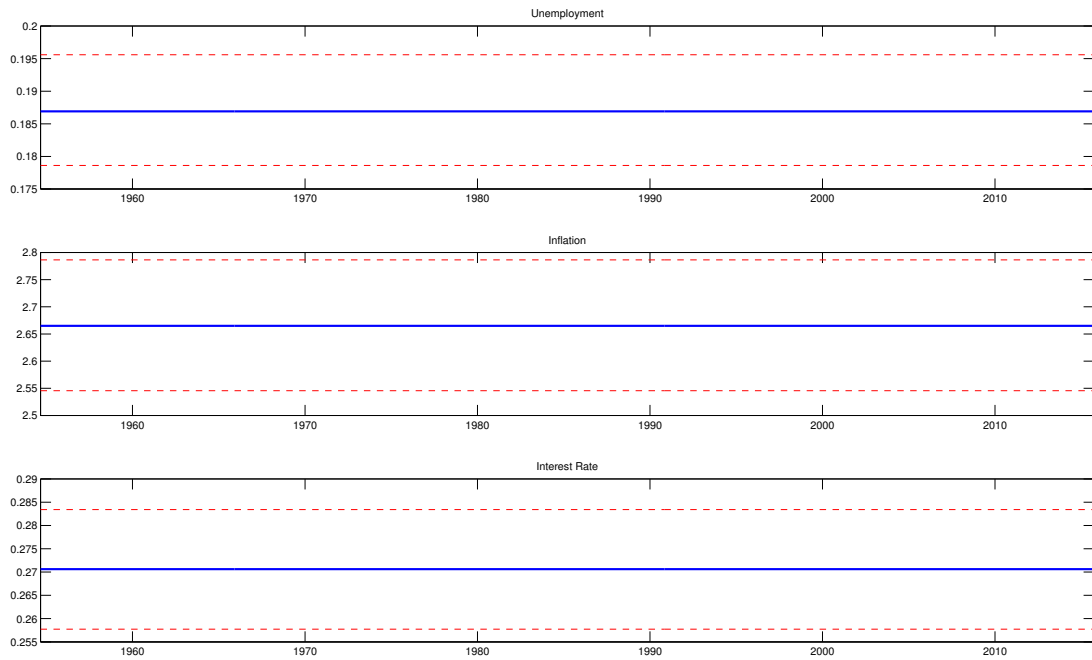
Figure A2: Impulse Response Functions



This figure shows the posterior means (blue solid line) of the impulse response functions together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The columns correspond to the responses in the unemployment equation, inflation equation, and interest equation, respectively. The monetary policy shock is defined as a one unit shock to interest rate. Since we estimate a time-invariant coefficient matrix, the impulse response functions are invariant to the timing of the shock.

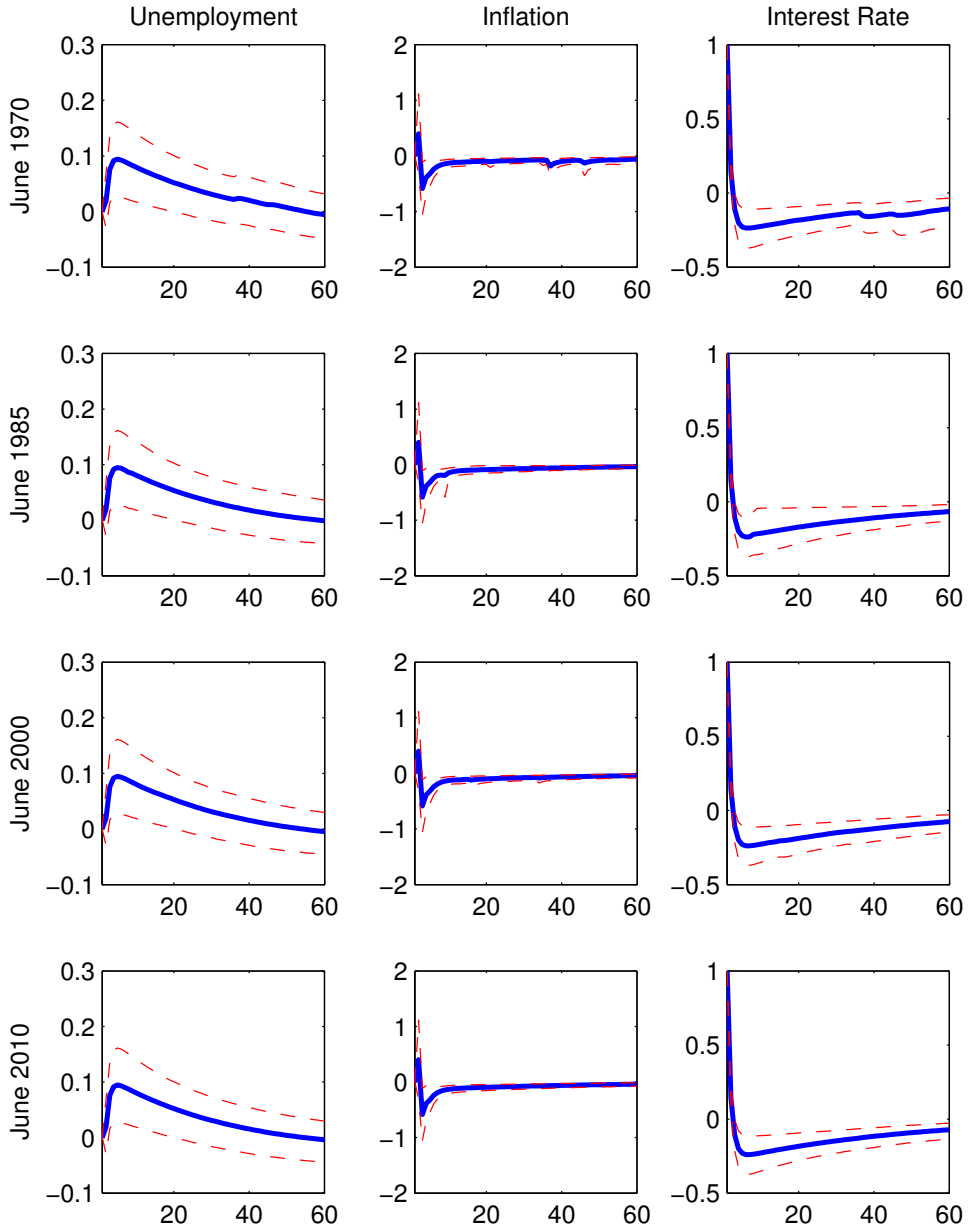
B Results Time-Invariant Covariance Matrix

Figure B1: Posterior Means of the Structural Variance Parameters



This figure shows the posterior means (blue solid line) of the structural variance parameters together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The panels show from top to bottom the variance of the residuals in the unemployment equation, inflation equation, and interest equation, respectively. Since we estimate a time-invariant covariance matrix, the estimated variances are time-invariant.

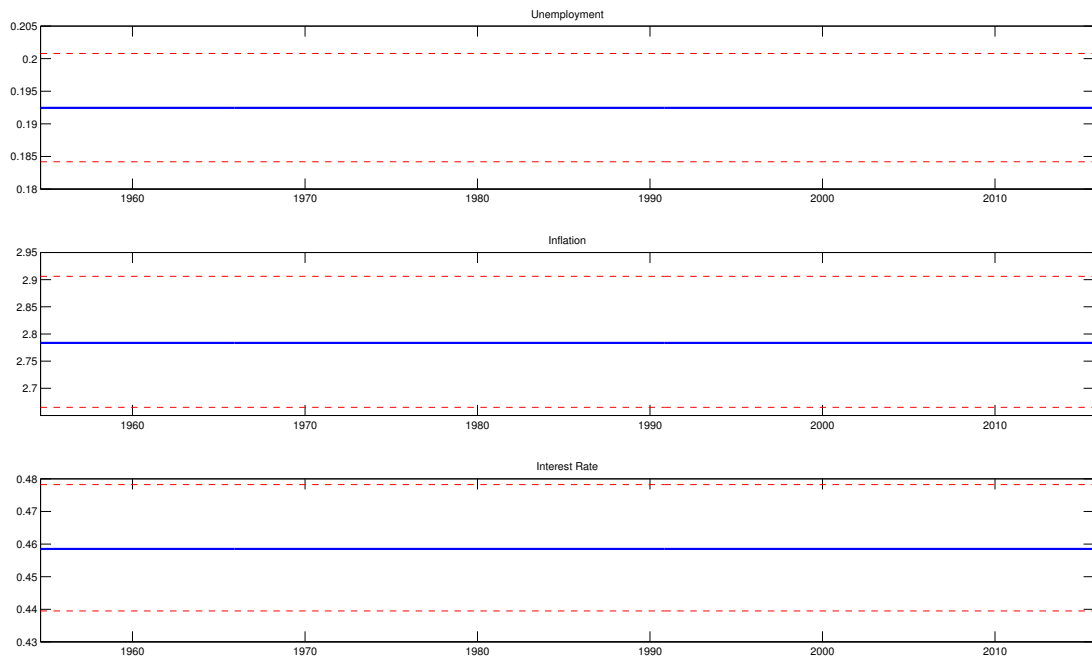
Figure B2: Impulse Response Functions



This figure shows the posterior means (blue solid line) of the impulse response functions together with the 5th and 95th percentiles of their posterior distributions (red dotted line). From top to bottom we have the impulse responses over a period of 60 months to monetary policy shocks in June 1970, June 1985, June 2000, and June 2010, respectively. The columns correspond to the responses in the unemployment equation, inflation equation, and interest equation, respectively. The monetary policy shock is defined as a one unit shock to interest rate.

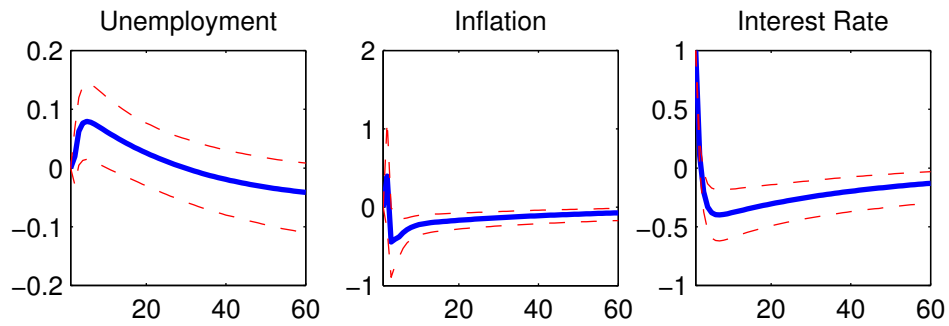
C Results Time-Invariant Model

Figure C1: Posterior Means of the Structural Variance Parameters



This figure shows the posterior means (blue solid line) of the structural variance parameters together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The panels show from top to bottom the variance of the residuals in the unemployment equation, inflation equation, and interest equation, respectively. Since we estimate a time-invariant covariance matrix, the estimated variances are time-invariant.

Figure C2: Impulse Response Functions



This figure shows the posterior means (blue solid line) of the impulse response functions together with the 5th and 95th percentiles of their posterior distributions (red dotted line). The columns correspond to the responses in the unemployment equation, inflation equation, and interest equation, respectively. The monetary policy shock is defined as a one unit shock to interest rate. Since we estimate a time-invariant coefficient matrix, the impulse response functions are invariant to the timing of the shock.