**Research Unit for Statistical
and Empirical Analysis in Social Sciences (Hi-Stat)**

Bayesian Analysis of Time-Varying Parameter Vector Autoregressive Model for the Japanese Economy and Monetary Policy

Jouchi Nakajima
Munehisa Kasuya
Toshiaki Watanabe

May 2009

Hi-Stat Discussion Paper

Bayesian Analysis of Time-Varying Parameter Vector Autoregressive Model for the Japanese Economy and Monetary Policy*

JOUCHI NAKAJIMA

Institute for Monetary and Economic Studies, Bank of Japan

MUNEHISA KASUYA

Research and Statistics Department, Bank of Japan

TOSHIAKI WATANABE

Institute of Economic Research, Hitotsubashi University

May 2009

Abstract

This paper analyzes the time-varying parameter vector autoregressive (TVP-VAR) model for the Japanese economy and monetary policy. The time-varying parameters are estimated via the Markov chain Monte Carlo method and the posterior estimates of parameters reveal the time-varying structure of the Japanese economy and monetary policy during the period from 1981 to 2008. The marginal likelihoods of the TVP-VAR model and other VAR models are also estimated. The estimated marginal likelihoods indicate that the TVP-VAR model best fits the Japanese economic data.

Key words: Bayesian inference, Markov chain Monte Carlo, Monetary policy, State space model, Structural vector autoregressive model, Stochastic volatility, Time-varying parameter.

1 Introduction

The Japanese economy has experienced several distinct periods of macroeconomic activity in recent decades, resulting in many of Japan's macroeconomic variables exhibiting changing behaviors over time. Since Miyao (2000, 2002) analyzed the Japanese economy using a vector autoregressive (VAR) model, the time variation of the relations among Japanese macroeconomic variables has been investigated in several studies (e.g., Fujiwara (2006), Inoue and Okimoto (2008) using a Markov-switching VAR model, and Kimura et al. (2003) using a VAR model with time-varying coefficients). In these studies, the changes in the coefficients in the VAR system are well studied, although the variance of the structural shocks is assumed constant over the sample period or subsample period. This paper estimates a time-varying parameter vector autoregressive (TVP-VAR) model for the Japanese economy and monetary policy, which allows both the coefficients and the variance of structural shock to vary over time.

From a methodological viewpoint, the TVP-VAR model has recently become increasingly popular in the macroeconomics literature following the use of this estimation technique by Primiceri (2005) using data for the US economy. The spirit of its specification is derived from Cogly and Sargent (2005). Benati and Mumtaz (2005) provide empirical results for the TVP-VAR model for the UK economy and Baumeister et al. (2008) for the Euro economy. D'Agostino et al. (2008) show the superior forecasting performance of the TVP-VAR specification over other VAR models using US macroeconomic data. We apply their method to the Japanese economy with a slight modification. Their scheme of sampling from the posterior distribution of the stochastic volatility of the TVP-VAR model uses a mixture sampler, which was originally developed by Kim et al. (1998) in the context of the stochastic volatility model in financial econometrics. The mixture sampler draws sample from the approximated posterior density and its approximation error is small enough to implement the overall model, as discussed by Omori et al. (2007). However, the multimove sampler proposed by Shephard and Pitt (1997) and modified by Watanabe and Omori (2004) can draw sample from the exact posterior density of the stochastic volatility, and this method is incorporated into the TVP-VAR model in this paper. Furthermore, Yano and Yoshino (2008) estimate the TVP-VAR model using a Monte Carlo particle filtering approach.

In our empirical analysis using Japanese data, a four-variable VAR system is estimated. The model includes the inflation rate, industrial production, nominal short-term interest rate, and money supply. The stochastic volatilities and time-varying impulse responses of the macroeco-

nomic variables are shown over time. The marginal likelihoods of the TVP-VAR specification and other VAR models are also estimated under different estimation conditions. The estimated marginal likelihood indicates the good performance of the TVP-VAR model.

The paper is organized as follows. In Section 2, we introduce the TVP-VAR specification. Section 3 illustrates the estimation procedure for the TVP-VAR model. Section 4 presents our empirical results. Finally, Section 5 concludes.

2 Structural VAR models

2.1 Preliminary

We begin with a basic structural VAR model defined as

$$Ay_t = F_1 y_{t-1} + \cdots + F_s y_{t-s} + u_t, \quad t = s+1, \ldots, n, \tag{1}$$

where y_t is an $k \times 1$ vector of observed variables, $A, F_1, \ldots, F_s$ are $k \times k$ matrices of coefficients, and u_t is a $k \times 1$ structural shock. We specify the simultaneous relations of the structural shock by recursive identification, assuming that A is lower-triangular,

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{21} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{k1} & \cdots & a_{k,k-1} & 1 \end{pmatrix}.$$

We rewrite model (1) as the following reduced form VAR model,

$$y_t = B_1 y_{t-1} + \cdots + B_s y_{t-s} + A^{-1} \Sigma \varepsilon_t, \quad \varepsilon_t \sim N(0, I_k),$$

where $B_i = A^{-1} F_i$, for $i = 1, \ldots, s$, and

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_k \end{pmatrix}.$$

The σ_i $(i = 1, \ldots, k)$ is the standard deviation of the structural shock. Stacking the elements in the rows of the B_i's to form β ($k^2 s \times 1$ vector), and defining $X_t = I_k \otimes (y'_{t-1}, \ldots, y'_{t-k})$, the model can be written as

$$y_t = X_t \beta + A^{-1} \Sigma \, \varepsilon_t. \tag{2}$$

All parameters in equation (2) are time-invariant. In the next section, we construct the model by allowing these parameters to vary over time.

2.2 Time-Varying Parameter VAR

We consider a time-varying parameter VAR (TVP-VAR) model specified by

$$y_t = X_t \beta_t + A_t^{-1} \Sigma_t \, \varepsilon_t, \quad t = s + 1, \ldots, n, \tag{3}$$

where the coefficients β_t, and the parameters A_t, and Σ_t are all time varying. There would be many ways to model the process for these time-varying parameters. Let a_t be a stacked vector of the lower-triangular elements in A_t and $h_t = (h_{1t}, \ldots, h_{kt})'$ with $h_{jt} = \log \sigma_{jt}^2$, for $j = 1, \ldots, k, t = s+1, \ldots, n$. As suggested by Primiceri (2005), we assume that the parameters in (3) follow a random walk process as follows:

$$
\begin{aligned}
\beta_{t+1} &= \beta_t + u_{\beta t}, \\
a_{t+1} &= a_t + u_{at}, \\
h_{t+1} &= h_t + u_{ht},
\end{aligned}
\qquad
\begin{pmatrix} \varepsilon_t \\ u_{\beta t} \\ u_{at} \\ u_{ht} \end{pmatrix}
\sim N \left(0,
\begin{pmatrix}
I & O & O & O \\
O & \Sigma_\beta & O & O \\
O & O & \Sigma_a & O \\
O & O & O & \Sigma_h
\end{pmatrix}
\right),
$$

for $t = s + 1, \ldots, n$, where $\beta_{s+1} \sim N(\mu_{\beta_0}, \Sigma_{\beta_0})$, $a_{s+1} \sim N(\mu_{a_0}, \Sigma_{a_0})$ and $h_{s+1} \sim N(\mu_{h_0}, \Sigma_{h_0})$. The shocks to the innovations of the time-varying parameters are assumed uncorrelated among the parameters β_t, a_t and h_t. We further assume that Σ_β, Σ_a and Σ_h is all diagonal matrices. The drifting coefficients and parameters are modeled to fully capture possible changes of the VAR structure over time. Our dynamic specification is adequate to permit the parameters to vary even if the shocks in the processes driving the time-varying parameters are uncorrelated.

Note that the log of variance (σ_t^2) for the structural shocks is modeled to follow a random walk process, which belongs to the class of stochastic volatility (e.g., Shephard (2005)). In financial literature, the log volatility (h_t) is often formulated to follow a stationary process such

as the first-order autoregressive process. The random walk process is non-stationary and it would be undesirable to analyze long-series behavior of market products such as financial daily data, while the purpose of our study is the empirical analysis for the quarterly macroeconomics data whose sample size is around one hundred. Moreover, as discussed by Primiceri (2005), the random-walk assumption can capture possible gradual (or sudden) structural changes.

3   Bayesian inference

The TVP-VAR model includes a number of parameters, while the estimation procedure can be constructed using the MCMC methods. In a Bayesian inference, under certain prior probability distributions, the MCMC algorithm produces the sample drawn from a high-dimensional posterior distribution of parameters including unobserved latent variables (see e.g., Chib (2001)). Using the time-varying parameters in our model as latent variables, the model forms a state space specification. The key to constructing an efficient sampling scheme for the TVP-VAR model is the joint sampling of $\beta = \{\beta_t\}_{t=s+1}^{n}$ (and in turn, $a = \{a_t\}_{t=s+1}^{n}$, $h = \{h_t\}_{t=s+1}^{n}$) conditioned on the rest of the parameters, which is better than the approaches that rely on one-at-a-time sampling. To accomplish this strategy, the simulation smoother (de Jong and Shephard (1995), Durbin and Koopman (2002)) is suitable for sampling the time-varying co-efficient β and parameter a because the model can be written in a linear Gaussian state space form.

Regarding the stochastic volatility h, the model forms a non-linear non-Gaussian state space form; thus, we need more technical methods of sampling. One way to sample stochastic volatility is the approach of Kim et al. (1998), called the mixture sampler. This method has been widely used in the financial and macroeconomics literature (Omori et al. (2007), Primiceri (2005)). The other way is the multimove sampler of Shephard and Pitt (1997) and Watanabe and Omori (2004). The former method deals with the approximated linear state space model. As studied by Kim et al. (1998) and Omori et al. (2007), its approximation error is small enough to implement the original model and can be corrected by the compensation step (see the details in Kim et al. (1998)). On the other hand, the latter algorithm approaches the model by drawing from the exact conditional posterior density. Both methods are appropriate to implement the volatility part in the TVP-VAR model, while we use the latter method in this paper for its direct sampling from the original form of the model.

3.1 MCMC algorithm

Let $y = \{y_t\}_{t=1}^n$ and $\omega = (\Sigma_\beta, \Sigma_a, \Sigma_h)$. We set the prior probability density as $\pi(\omega)$ for ω. Given the data y, we draw sample from the posterior distribution, $\pi(\beta, a, h, \omega | y)$, by the MCMC method. There are several ways of sampling the posterior distribution; we apply the following MCMC algorithm:

1. Initialize β, a, h, ω.

2. Sample $\beta | a, h, \Sigma_\beta, y$.

3. Sample $\Sigma_\beta | \beta$.

4. Sample $a | \beta, h, \Sigma_a, y$.

5. Sample $\Sigma_a | a$.

6. Sample $h | \beta, a, \Sigma_h, y$.

7. Sample $\Sigma_h | h$.

8. Go to 2.

As mentioned above, Steps 2 and 4 are conducted with the help of the simulation smoother, and Step 6 requires the multi-move sampler for the stochastic volatility. The assumption of the diagonal matrix of Σ_h makes the conditional posterior distribution of $\{h_{jt}\}_{t=s+1}^n$ independent with respect to the series $j\ (= 1, \ldots, k)$ and the sampling algorithm for h becomes simple. Steps 3, 5, and 7 are straightforward, drawing sample from a Wishart or Gamma distribution under conjugate priors. The details of the procedure are illustrated in Appendix.

3.2 Priors

The priors should be carefully chosen because the TVP-VAR model has many state variables in the VAR specification and their process is modeled as a non-stationary random walk process. The TVP-VAR model is so flexible that the state variables can capture both gradual and sudden changes of the underlying economic structure. On the other hand, allowing time variation in every parameter in the VAR may cause an over-identification problem in the finite period data sample. As mentioned by Primiceri (2005), the tight prior for the covariance matrix of the disturbance in the random walk process avoids implausible behaviors of the time-varying

parameters. A tighter prior should sometimes be avoided in empirical econometrics, although the TVP-VAR model needs slightly tighter priors to provide reasonable identification.

In this paper, the time-varying coefficient (β) needs a tighter prior than the simultaneous relations (a) and the volatility (h) of the structural shock for the variance of the disturbance in their time-varying process. The structural shock we consider in the model unexpectedly hits the economic system and its size would more widely fluctuate over time than the possible change of the autoregressive system of the economic variables specified by VAR coefficients. Through the estimation procedure in the following sections, we set a slightly tighter prior for Σ_β and a rather diffuse prior for Σ_a and Σ_h. Of course, a prior sensitivity analysis is necessary in this situation to check the robustness of the empirical result with respect to the prior tightness. Section 4.3 reports the robustness check.

An additional remark is required for the prior of the initial state of the time-varying parameters. When the time series to estimate is modeled as a stationary process, we often assume the initial state follows a stationary distribution of the process. However, our time-varying parameters are random walks; thus, we specify the prior of the normal distribution for the initial state of each time-varying parameter. Following Primiceri (2005), we determine the mean of these normal priors with the estimates of a time-invariance VAR model using the pre-sample period. It is possible to specify a flat prior for the initial state to draw the sample of the posterior fully using information from data. However, we consider it is reasonable to use the economic structure estimated from the pre-sample period up to the start of the main sample data.

3.3   Marginal likelihood

In a Bayesian framework, we can compare model fit given data using the posterior probabilities of the models. The posterior probability of each model is proportional to the prior probability of the model, times the marginal likelihood. The ratio of two posterior probabilities is also well known as a Bayes factor. If the prior probabilities are assumed equal, we choose the model that yields the largest marginal likelihood. The marginal likelihood is defined as the integral of the likelihood with respect to the prior density of the parameter. There are several ways to estimate the marginal likelihood for the model including state variables; we use the harmonic mean method (e.g., Geweke (1999)). The simulation-based harmonic mean estimator of the

marginal likelihood, denoted by $\hat{m}(y)$, is obtained by

$$\frac{1}{\hat{m}(y)} = \frac{1}{M} \sum_{j=1}^{M} \frac{g(\omega^{(j)})}{f(y|\omega^{(j)}, \vartheta^{(j)})\pi(\omega^{(j)})},$$

where $\vartheta = (\beta, a, h)$, $f(y|\omega^{(j)}, \vartheta^{(j)})$ and $\pi(\omega^{(j)})$ denote the likelihood function and prior density, respectively, and M is the iteration size of the MCMC. If the fraction $g(\omega)/f(y|\omega, \vartheta)\pi(\omega)$ is bounded above, then the approximation is simulation consistent and the rate of convergence is likely to be practical. While $g(\omega)$ can be any p.d.f. with support contained in the parameter space of the model, Geweke (1999) recommends the choice of g for the modified harmonic mean estimator to guarantee the boundness of this fraction. Consider the normal density with the tails truncated, namely

$$g(\omega^{(j)}) = \frac{1}{\tau(2\pi)^{p/2}|\hat{V}|^{1/2}} \exp\left\{-\frac{1}{2}(\omega^{(j)} - \hat{\omega})'\hat{V}^{-1}(\omega^{(j)} - \hat{\omega})\right\}$$
$$\times I\left[(\omega^{(j)} - \hat{\omega})'\hat{V}^{-1}(\omega^{(j)} - \hat{\omega}) \leq \chi_\tau^2(p)\right],$$

where $I[\Omega]$ is an indicator function that takes the value of one if Ω is satisfied and zero otherwise, p is the number of elements in ω, and $\chi_\tau^2(p)$ denote the τ percentile of the Chi-square distribution with p degrees of freedom. By cutting off the tails, sample that drop in the potentially problematic regions are avoided for the computation of the marginal likelihood. We set $\hat{\omega}$ and $\hat{V}$ equal to the sample mean and covariance matrix computed from the posterior draws $\{\omega^{(j)}\}_{j=1}^{M}$, and $\tau = 0.99$ in this paper. As mentioned by Schorfheide (2000), interestingly, the estimated marginal likelihood does not so serially depend on τ. We also computed the marginal likelihood of the TVP-VAR model with $\tau = 0.95, 0.90$, and found the estimated marginal likelihood is seldom sensitive to the value of τ.

4 Empirical evidences for Japanese economy and monetary policy

4.1 Data and estimation procedure

In this section, we estimate the TVP-VAR model for the Japanese economy. Our main dataset is quarterly and the sample period is from 1981/1Q to 2008/3Q. The pre-sample period from 1970/2Q to 1980/4Q is used for the prior distribution of the initial state in the process of

the time-varying parameters, which is explained in detail below. The model includes four macroeconomic variables: inflation rate, industrial production, nominal short-term interest rate, and monetary base.[1] Recursive identification is assumed in this order. These variables are used for a standard VAR model of the Japanese economy, analyzed by several papers (Miyao (2000, 2002), Fujiwara (2006), Inoue and Okimoto (2008), Yano and Yoshino (2008)). Most of these studies use monthly data for the estimation, while we change the data to a quarterly base by monthly average. The VAR estimation with monthly data often needs many lags (for example, Miyao (2000) sets 10 lags), because the changes and shocks in the economic variables are considered to affect the other variables of the system with a delay. As mentioned above, the TVP-VAR model has so many parameters and our choice of quarterly data reduces the number of parameters to estimate. The lags are determined by the estimated marginal likelihood. We estimate the model with one to four lags and choose the lags whose marginal likelihood is the highest among them. As shown below, the main empirical results are obtained with two lags.

The following priors are assumed for the i-th diagonals of the covariance matrices:

$$(\Sigma_\beta)_i^{-2} \sim \text{Gamma}(10, 0.01), \ (\Sigma_a)_i^{-2} \sim \text{Gamma}(2, 0.01), \ (\Sigma_h)_i^{-2} \sim \text{Gamma}(2, 0.01). \quad (4)$$

For the initial state of the time-varying parameters, $\mu_{\beta_0} = \hat{\beta}_0$, $\mu_{a_0} = \hat{a}_0$, $\mu_{h_0} = \log \hat{\sigma}_0^2$, and $\Sigma_{\beta_0} = \Sigma_{a_0} = \Sigma_{h_0} = 4 \times I$, where $\hat{\beta}_0$, $\hat{a}_0$ and $\hat{\sigma}_0$ are the OLS estimators obtained using the pre-sample period.

We draw $M = 10,000$ sample after the initial 1,000 sample are discarded. The computational results are generated using Ox version 5.0 (Doornik (2006)). Figure 1 shows the sample autocorrelation function, the sample paths and the posterior densities for selected parameters. After discarding the sample in the burn-in period, the sample paths look stable and the sample autocorrelations drop stably, indicating our sampling method efficiently produces uncorrelated samples.

Table 1 gives the estimates for posterior means, standard deviations, the 95% credible intervals, the convergence diagnostics (CD) of Geweke (1992) and inefficiency factors. Geweke

---

[1]The inflation rate is taken from the CPI (consumer price index, general excluding fresh food, the effects of the increase in the consumption tax removed, and seasonally adjusted). Industrial production is seasonally adjusted. The nominal short-term interest rate is the overnight call rate. The monetary base is the average outstanding, adjusted for the reserve requirement ratio changes, and seasonally adjusted. For the sudden and temporal increases of the monetary base around December 1999 and February 2002, a linear interpolation is used. Except for the call rate, all the variables are transformed in logarithm, and multiplied by 100. In the estimation, we take the first difference of all variables including the call rate.

(1992) suggests a comparison between the first n_0 draws and the last n_1 draws, dropping out the middle draws. The CD statistics is computed by CD $= (\bar{x}_0 - \bar{x}_1)/\sqrt{\hat{\sigma}_0^2/n_0 + \hat{\sigma}_1^2/n_1}$, where $\bar{x}_j = \frac{1}{n_j} \sum_{i=m_j}^{m_j+n_j-1} x^{(i)}$, $x^{(i)}$ is the i-th draw, and $\sqrt{\hat{\sigma}_j^2/n_j}$ is the standard error of $\bar{x}_j$, respectively, for $j = 0, 1$. If the sequence of the MCMC sampling is stationary, it converges in distribution to a standard normal. We set $m_0 = 1$, $n_0 = 1,000$, $m_1 = 5,001$, and $n_1 = 5,000$. The $\hat{\sigma}_j^2$ is computed using a Parzen window with bandwidth, $B_m = 500$. The inefficiency factor is defined as $1 + 2 \sum_{s=1}^{B_m} \rho_s$, where ρ_s is the sample autocorrelation at lag s, which is computed to measure how well the MCMC chain mixes (see e.g., Chib (2001)). It is the ratio of the numerical variance of the posterior sample mean to the variance of the sample mean from uncorrelated draws. The inverse of the inefficiency factor is also known as relative numerical efficiency (Geweke (1992)). When the inefficiency factor is equal to m, we need to draw m times as many MCMC sample as uncorrelated samples. In our empirical result, the null hypothesis of the convergence to the posterior distribution is not rejected for the parameters at the 5% significance level based on the CD statistics, and the inefficiency factors are very low, which indicates an efficient sampling for the parameters and state variables.
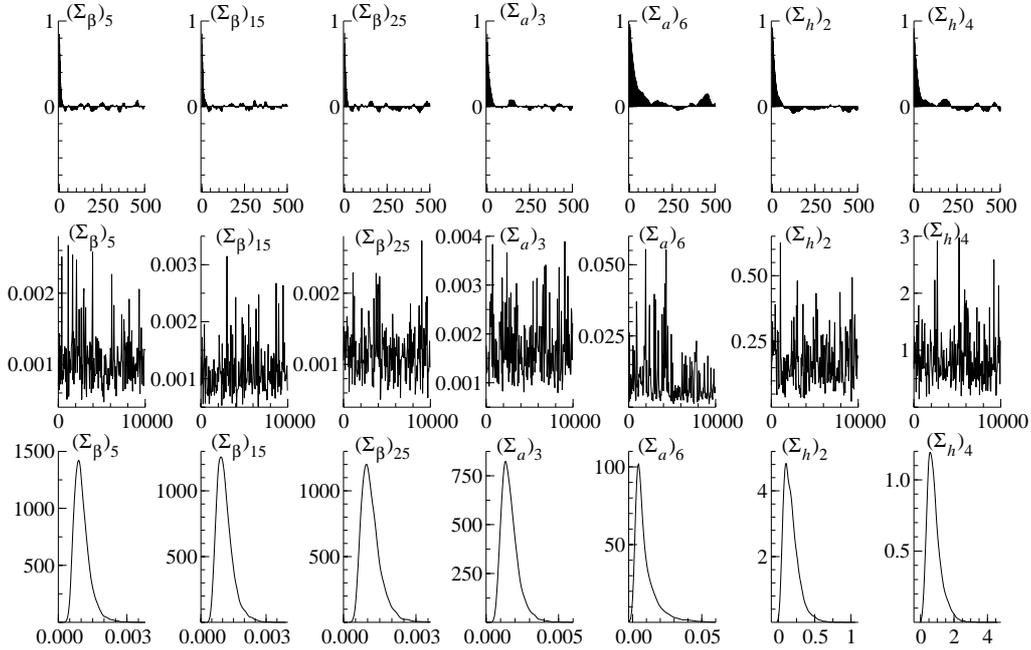

Figure 1: Estimation results for selected parameters in the TVP-VAR model. Sample auto-correlations (top), sample paths (middle) and posterior densities (bottom).

10

Parameter	Mean	Stdev.	95% interval	CD	Inefficiency
$(\Sigma_\beta)_5$	0.1037	0.0355	[0.0570, 0.1919]	0.974	8.01
$(\Sigma_\beta)_{15}$	0.1104	0.0387	[0.0579, 0.2061]	0.075	10.11
$(\Sigma_\beta)_{25}$	0.1149	0.0413	[0.0594, 0.2175]	0.700	10.90
$(\Sigma_a)_3$	0.1652	0.0610	[0.0827, 0.3178]	0.922	24.62
$(\Sigma_a)_6$	0.9300	0.8759	[0.0174, 3.4204]	0.149	68.76
$(\Sigma_h)_2$	0.1748	0.1047	[0.0411, 0.4419]	0.215	23.41
$(\Sigma_h)_4$	0.8198	0.4306	[0.2516, 1.9022]	0.622	44.64

Table 1: Estimation results for selected parameters in the TVP-VAR model (the estimates of Σ_β and Σ_a are multiplied by 100).

4.2 Empirical results

This section provides quantitative empirical results for the Japanese economy and monetary policy of the TVP-VAR model. Through the sample period from 1981 to 2008, the Japanese economy experienced several different periods. The bubble economy burst in the early 1990s, and the Bank of Japan introduced the zero interest rate policy from 1999 to 2000 and the quantitative easing policy from 2001 to 2006. Using the TVP-VAR model, we investigate the time-varying structure of the Japanese economy and monetary policy, as follows.

4.2.1 Time-varing volatility

Figure 2 plots the series for estimated stochastic volatility of the structural shock on four variables, based on the posterior mean, and the 16-th and 84-th quantile[2] intervals of the standard deviation of the shock, $\sigma_{it} = \exp(h_{it}/2)$. It presents the dynamics of the volatility over time, which differs across variables.

Figure 2(i) plots the time-varying volatility of the inflation rate. The period 1981–1985 exhibits a relatively higher volatility because of the second oil shock. The additional reduction is observed around 1998 to 2006, when the Japanese economy experienced deflation. The time-varying volatility of industrial production, displayed by Figure 2(ii), shows that its standard deviation increased substantially from the mid-1990s to the first half of the 2000s compared with the preceding period. Sakura et al. (2005) point out that the variance of real GDP became more volatile during the 1990s compared with the 1980s. They estimate the time-invariant VAR model by dividing the sample into subsamples and find that both the parameters in the VAR and the shock contribute to the fluctuation of real GDP. One source of the fluctuation is considered to be the higher volatility of firm investment after the bubble economy burst, as

---

[2]As Primiceri (2005) uses, the 16-th and 84-th quantiles correspond to the one-standard-deviation band under normality.
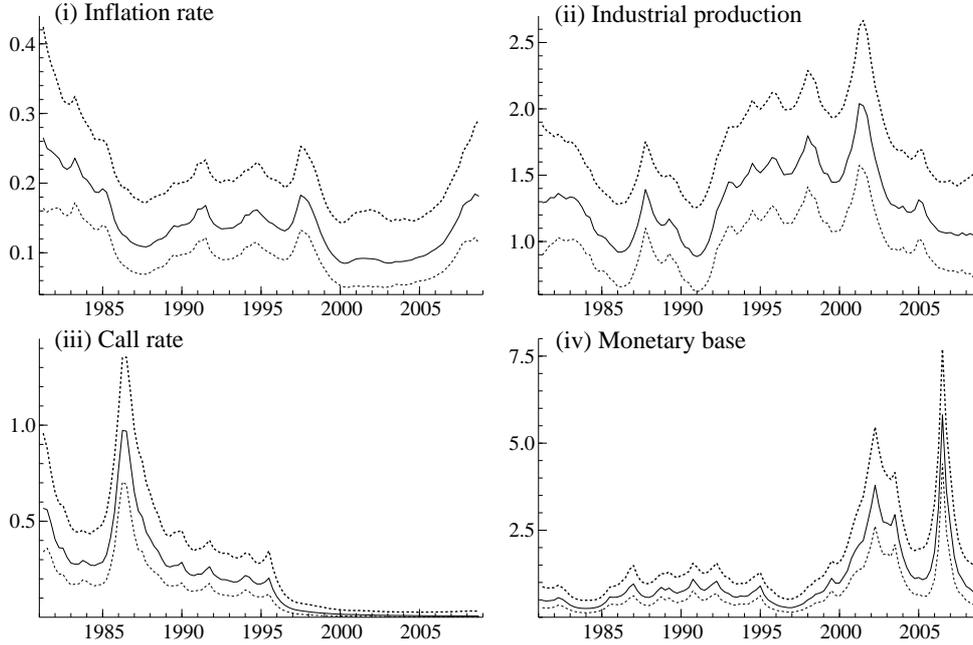
Figure 2: Posterior estimates for stochastic volatility of structural shock, $\sigma_{it} = \exp(h_{it}/2)$. Posterior mean (solid line), 16-th and 84-th percentiles (dotted line).

discussed by Sakura et al. (2005). The same fluctuation in the shock of industrial production is observed in our analysis. The TVP-VAR model has an advantage over the constant VAR models in the sense that it need not divide data into subsamples to confirm the change of the structure of the model.

Figure 2(iii) and (iv) plots the series of stochastic volatility for the call rate and the monetary base, respectively. The BOJ's policy instrument of the overnight uncollateralized call rate decreased to 0.25% in 1995. Then, the BOJ implemented the zero interest rate policy[3] in 1999 and the quantitative easing policy in 2001. During the quantitative easing policy period, the outstanding balance of the current account is the target of the BOJ's monetary policy, and expansions of the monetary base were observed by March 2006, when the BOJ altered the policy instrument to the call rate again. The estimated variances of the call rate and monetary base show these changes in monetary policy; that is, the stochastic volatility of the call rate

---

[3]In our specification of the TVP-VAR model, the zero lower bound of the nominal interest rate is not considered. While it may be possible to incorporate the lower bound for a certain variable on the TVP-VAR model, it requires a computational burden. Thus, in the current paper, this issue is left for future work. Regarding the model fit of the time-varying parameters, the estimation of the subsample period that excludes the zero interest rate regime is also examined below. One way to solve this econometric issue is well discussed by Iwata and Wu (2006) with a censored variable.

drops close to zero around the mid-1990s and the monetary base increases rapidly from the beginning of the 2000s, and a hike is marked in 2006, which corresponds to the termination of the quantitative easing policy. The spike around 1986 for the volatility of the call rate would indicate that the easing policy in this period is measured as the structural shock. After that, the volatility of the call rate is relatively low and looks moderate.

Overall, the time-varying volatility contributes to the VAR estimation, identifying the structural shock with the appropriate variance of the shock size. For the data we analyze here, the estimates of the time-invariant VAR model with constant parameters would result in biases in the covariance matrix for the disturbances and at the same time in the autoregressive coefficients because of the misspecification of the dynamics of the parameters. Again, the nominal interest rate lower bound is not explicitly incorporated into the model, although the estimated volatility for the call rate that stays low enough would help the VAR system to identify exogenous shocks better than the time-invariant VAR model.

4.2.2 Time-varying impulse responses

Impulse response analyses are provided for the time series in the TVP-VAR model. Because the coefficients are time varying, the impulse responses are calculated at each date over the sample period. In our study, the shock size for the responses is set equal to the time-series average of the stochastic volatility for each series over the sample period. Note that the shock size for the response is not based on the estimated variance each time. The impulse summarizes the effects of average-sized experimental structural shocks hitting the VAR system. The impulse responses are calculated for each iteration of the MCMC with the current draw of the parameters, and the sample mean and standard deviation of the responses are computed. Figure 3–6 plot the results of the main responses.

Figure 3 reports the impulse responses of industrial production to a positive interest rate shock. The time series of the response and the responses in three different dates with one-standard-deviation bands are shown. The dates for the comparison are (ii)1986/1Q, (iii)1994/1Q, and (iv)2002/1Q, which are chosen arbitrarily, while the overall time variation of the response is summarized in the time-series plot(i). The period around (iii)1994/1Q represents the typical economic conditions after the bubble burst and (iv)2002/1Q under the quantitative easing policy. The interest rate shock has a negative effect on industrial production, although its impact seems to exhibit time variation. It is natural that the response of industrial production becomes smaller from the second half of the 1990s to the 2000s because
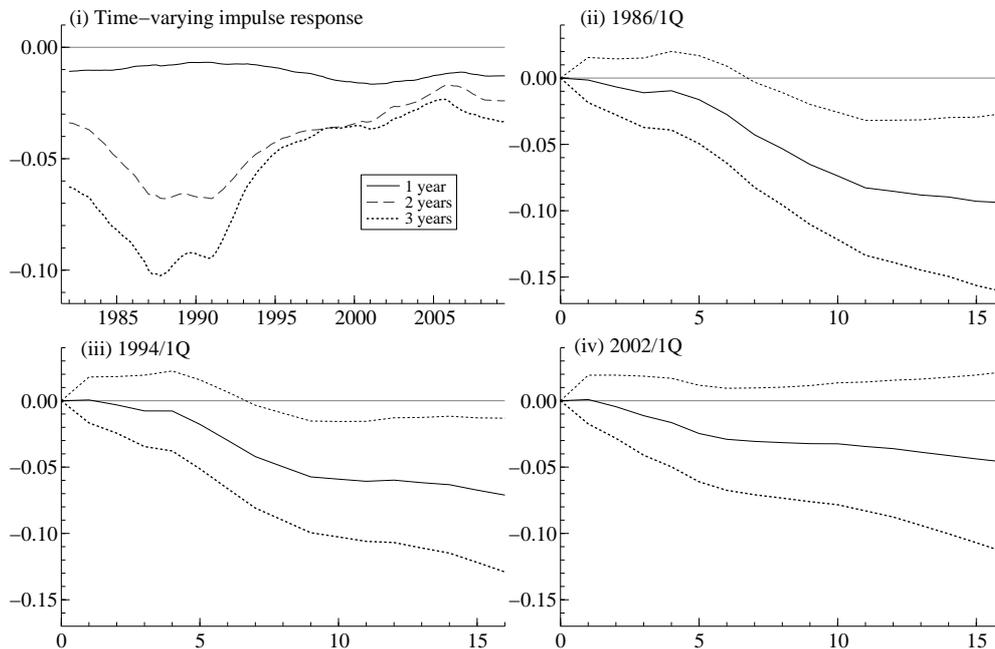
Figure 3: Impulse responses of industrial production to a positive interest rate shock. (i) Time series of response after one, two and three years, (ii)–(iv) response in each period with one-standard-deviation bands.

of the environment where the call rate stays close to zero. This behavior of the response is consistent with the result of Kimura et al. (2003). In addition, the estimation result indicates that the bottom of the impact of the interest rate shock is around 2006, when the quantitative easing policy is terminated. The time-varying impulse response implies that the increase in the policy interest rate in 2006 does not decrease industrial production.

Second, Figure 4 shows the responses of inflation to a positive interest rate shock. A rise in inflation after a monetary tightening using VAR estimates is well known as the price puzzle (Sims (1992), Eichenbaum (1992)). In our empirical estimates of the TVP-VAR model, the impulse response exhibits a different shape in each period. For 1981–1987, the impulse response falls sufficiently in response to the positive interest rate shock and Figure 4(ii) shows a slight and temporary price puzzle in 1986/1Q. On the other hand, from 1988 to the mid-1990s, the impulse response keeps rising and stays high even at a three-year horizon. As plotted in 4(iii), the impulse response is a positive domain over all of 1990/1Q. There are several explanations for the price puzzle and the VAR estimates would depend on the lags or some omitted variables in the VAR system to some extent. Figure 4(iv) shows that the impulse response stays near
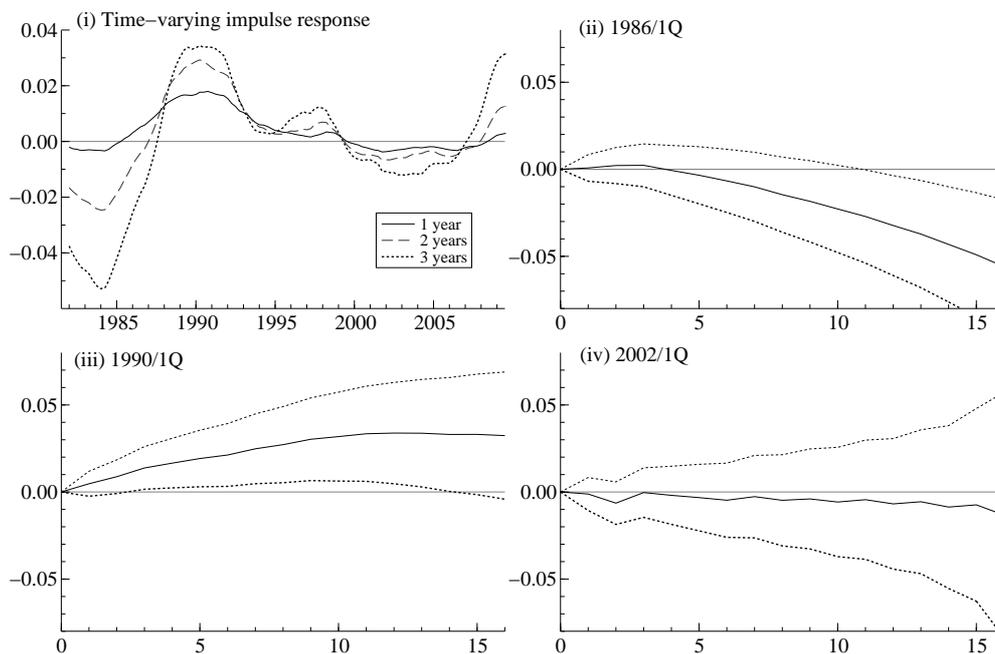
Figure 4: Impulse responses of inflation to positive interest rate shock. (i) Time series of response after one, two and three years, (ii)–(iv) response of each period with one-standard-deviation bands.

zero and the one-standard-deviation bands includes zero for 2002/1Q because the call rate is close to zero and the structural shock to the interest rate diminishes in this period. Our estimation result indicates that the TVP-VAR model has an advantage in the sense that it can assess the time-varying dynamics of the relation among economic variables.

Third, the responses of the interest rate to a positive inflation rate shock are shown in Figure 5. This response is one measure of monetary policy activism with respect to inflation. Similar to the previous figure, the response of the interest rate has time variation and the impact of the response seems to diminish for the period of the zero interest rate policy and the quantitative easing policy. In Figure 5(i), the time-series line of the response after one year does not seem to fluctuate for 1981–1998, while the response after two and three years shows a high degree of time variation commencing in the early 1990s. It means that the response increases smoothly in reaching the three-year horizon of the response. For the other periods, the response seems to reach the three-year horizon of the response in the first year. From another point of view, the time-varying impulse response shows variation of the weight of inflation in the monetary policy reaction function throughout the sample period.
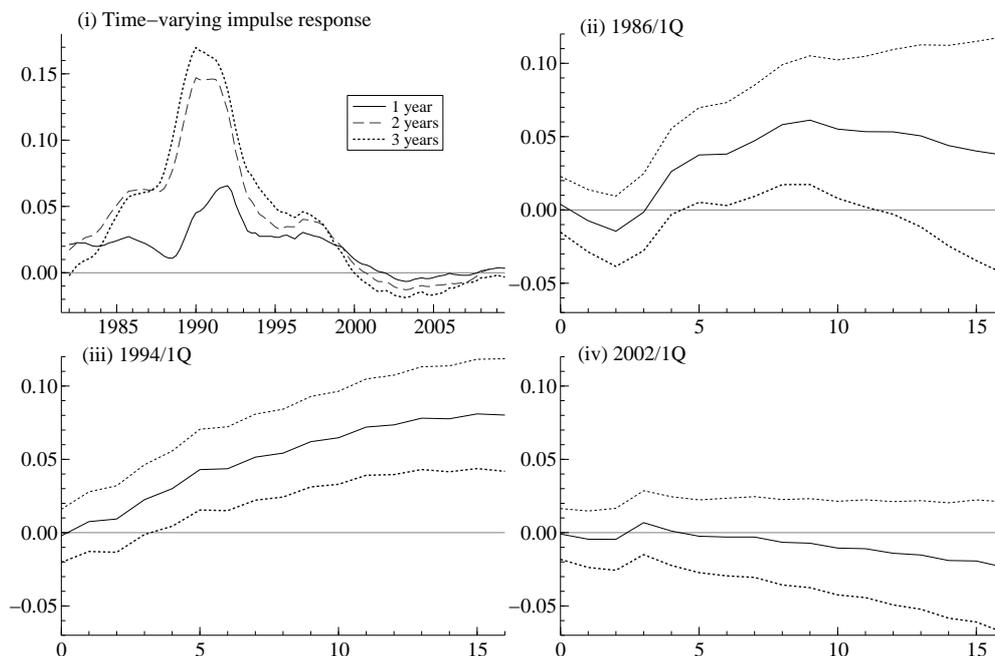
15

Figure 5: Impulse responses of interest rate to positive inflation rate shock. (i) Time series of response after one, two and three years, (ii)–(iv) response in each period with one-standard-deviation bands.

Fourth, Figure 6 reports the impulse responses of industrial production to a positive monetary base shock. Overall, the monetary base shock has a positive effect on industrial production during the sample period including the quantitative easing policy period, and in particular, the time series of the response after one year seems to be stable for 1990 to 2000, while the size of the response decreases around the mid-2000s. Throughout the sample period, industrial production seems to reach the three-year horizon response level in two years. Because we observe that the rise in the monetary base is followed by a positive interest rate shock in our estimation results (not shown here), the monetary base shock can be interpreted as a shock to the money demand function. This indicates that the money demand shock would have a positive effect through bank lending or other financial market variables. However, when the nominal interest rate is close to zero, the response becomes smaller than the one in the 1980s and the first half of the 1990s. For the impulse response of (iv)2002/1Q, the one-standard-deviation bands include zero. This result indicates that the effect of an increase in the monetary base is uncertain, as discussed by Kimura et al. (2003) and Fujiwara (2006). Similar to the interest rate shock, the impulse response of the monetary base shock reaches its bottom around 2006, when the quan-
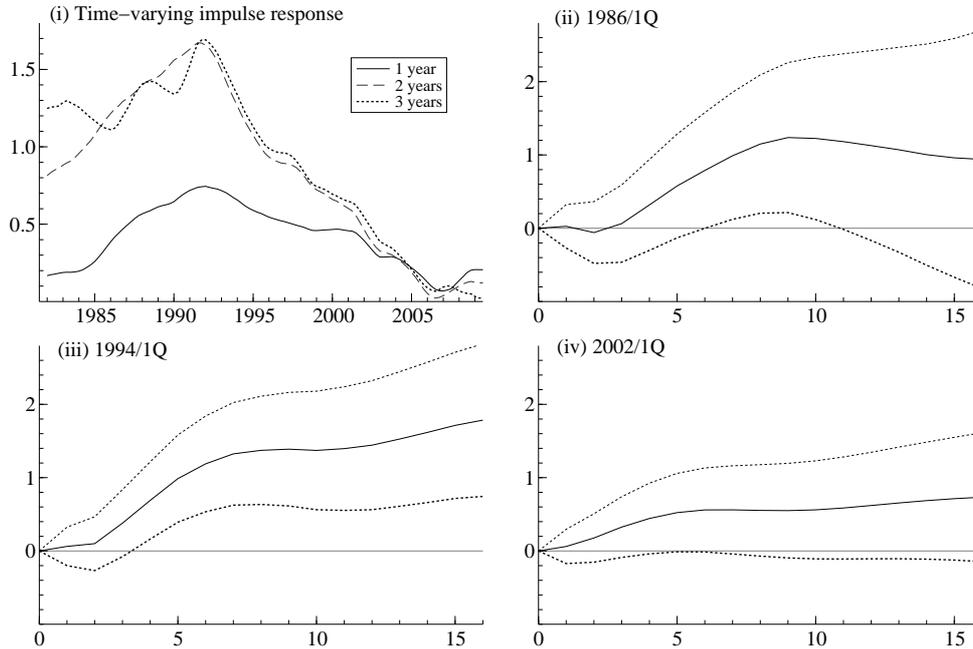
16

Figure 6: Impulse responses of industrial production to positive monetary base shock. (i) Time series of response after one, two and three years, (ii)–(iv) response of each period with one-standard-deviation bands.

titative easing policy is terminated and the monetary base decreases. The estimation results show that the rapid decrease in the monetary base does not decrease industrial production after 2006. It implies that this period would have been an appropriate time to terminate the quantitative easing policy.

4.3 Model selection and robustness

In the previous section, we showed the empirical results of the TVP-VAR model for the main dataset with the specified priors. Below, we examine whether the TVP-VAR model also fits the data better than other VAR models. We estimate the marginal likelihood using the modified harmonic mean method in Section 3.3 for different model specifications, lags, and priors. The competing model specifications are a constant parameter VAR model (CP-VAR) and a semi time-varying parameter VAR model (STVP-VAR). The CP-VAR model refers to the equation (2). The STVP-VAR model allows only the coefficients (β) and the simultaneous relations of structural shock (a) to vary over time, and the volatility of structural shock is time invariant with $\Sigma_t = \Sigma$, for all $t = s + 1, \ldots, n$. By comparing the model fit between the TVP-VAR

Lag	CP-VAR	STVP-VAR	TVP-VAR	Lag	CP-VAR	STVP-VAR	TVP-VAR
(i) Full sample, Prior1				(ii) Full sample, Prior2			
1	-519.01	-457.41	-339.11	1	-535.90	-515.79	-369.87
2	-548.76	-430.67	-311.24	2	-577.73	-527.24	-382.85
3	-581.30	-404.67	-327.12	3	-622.73	-535.14	-397.44
4	-621.13	-442.36	-346.30	4	-675.30	-567.05	-407.87
(iii) Subsample, Prior1				(iv) Subsample, Prior2			
1	-267.08	-266.23	-265.31	1	-284.13	-298.07	-263.43
2	-292.63	-282.30	-278.17	2	-321.82	-307.57	-287.26
3	-315.38	-292.20	-279.30	3	-357.18	-322.95	-305.14
4	-350.11	-315.50	-285.77	4	-404.59	-347.79	-329.82

Table 2: Estimated marginal likelihood for competing VAR models (logarithm scale; CP: constant parameter, STVP: semi time-varying parameter, TVP: time-varying parameter).

and STVP-VAR models, the advantage of the stochastic volatility of the structural shock, as emphasized by Cogly and Sargent (2005), is examined. The lag is set one up to four. We specify additional priors for the CP-VAR and STVP-VAR models and alternative priors for the TVP-VAR model, as follows:

Prior1 :

(4) and $\beta \sim N(0, 2 \times I)$, $\quad \bar{\alpha} \sim N(0, 2 \times I)$, $\quad \sigma_i^{-1} \sim \text{Gamma}(2, 0.02)$,

Prior2 :

$(\Sigma_\beta)_i^{-2} \sim \text{Gamma}(20, 0.01)$, $(\Sigma_a)_i^{-2} \sim \text{Gamma}(5, 0.01)$, $(\Sigma_h)_i^{-2} \sim \text{Gamma}(2, 0.01)$,

$\beta \sim N(0, 10 \times I)$, $\quad \bar{\alpha} \sim N(0, 10 \times I)$, $\quad \sigma_i^{-1} \sim \text{Gamma}(2, 0.02)$,

where $\bar{\alpha}$ denotes the stacked vector of the elements in A.

In addition to the estimation with the full sample period, we estimate the marginal likelihood using the subsample period from 1981/1Q to 1995/4Q. The period of the zero interest rate policy and the monetary easing policy is omitted in this subsample period. We estimate three competing models for the subsample data with Prior1 and Prior2.

Table 2 reports the estimated marginal likelihoods for competing models under four conditions: (i) full sample, Prior1, (ii) full sample, Prior2, (iii) subsample, Prior1, and (iv) subsample, Prior2. The iteration size of MCMC is the same as in the main empirical results. For all the conditions, the TVP-VAR model has the highest marginal likelihood. The STVP-VAR model is favored over the CP-VAR model model, while the marginal likelihood of the TVP-VAR model is much higher than for the STVP-VAR model for the full-sample period.

Interestingly, the difference in the marginal likelihood between the competing models is smaller for the subsample period than for the full-sample period under both priors. In particular, the marginal likelihood of the STVP-VAR model becomes close to the one for the TVP-VAR model in the subsample estimation. One possible reason for this result is that the full-sample period includes the zero interest rate policy and the monetary easing policy; thus, the posterior distribution of constant volatility estimated through the full-sample period would be biased and unable to follow the volatility dynamics, especially that seen for the call rate and money supply, as shown in Figure 2. We understand that the TVP-VAR specification is a very flexible model for analyzing the economic variables including the period when the nominal interest rate is close to zero.

Overall, the time-varying parameter contributes to the VAR model based on the estimated marginal likelihoods. The time-varying coefficients increase the marginal likelihood, although not by enough, and the time-varying simultaneous relations and stochastic volatility of structural shock contributes the model fit over our data set, especially for the sample period including the zero interest rate policy. These results are robust as shown by the estimations under different priors and lags.

5   Conclusion

This paper analyzes the TVP-VAR models of the Japanese economy and monetary policy. The time-varying parameters are estimated via the Markov chain Monte Carlo method and the posterior estimates of parameters reveal the time-varying structure of the Japanese economy and monetary policy during the period from 1981 to 2008. The marginal likelihoods of the TVP-VAR model and other VAR models are estimated under different priors, lags, and sample periods. The estimated marginal likelihoods indicate that the TVP-VAR model best fits the Japanese economic data over; in particular, it marks much higher marginal likelihoods for the sample period that includes the zero interest rate policy.

Appendix. MCMC algorithm for the TVP-VAR model

A.1 Sampling β

To sample β from the conditional posterior distribution $\pi(\beta|a, h, \Sigma_\beta, y)$, we write the model in the state space form as

$$
\begin{aligned}
y_t &= X_t\beta_t + A_t^{-1}\Sigma_t\varepsilon_t, & t &= s+1, \ldots, n, \\
\beta_{t+1} &= \beta_t + u_{\beta t}, & t &= s, \ldots, n-1,
\end{aligned}
$$

where $\beta_s = \mu_{\beta_0}$, and $u_{\beta s} \sim N(0, \Sigma_{\beta_0})$. It is possible to draw sample from the joint posterior distribution $\pi(\beta_{s+1}, \ldots, \beta_n|a, h, \Sigma_\beta, y)$. Following de Jong and Shephard (1995), we show the algorithm of the simulation smoother on the state space model

$$
\begin{aligned}
y_t &= Z_t\alpha_t + G_t u_t, & t &= s+1, \ldots, n, \\
\alpha_{t+1} &= T_t\alpha_t + H_t u_t, & t &= s, \ldots, n-1,
\end{aligned}
$$

where $u_t \sim N(0, I)$ and $G_t H_t' = O$. The simulation smoother draws $\eta = (\eta_s, \ldots, \eta_{n-1}) \sim \pi(\eta|y, \theta)$ where $\eta_t = H_t u_t$ for $t = s, \ldots, n-1$, and θ denotes the rest of the parameters in the model. We run the Kalman filter:

$$
\begin{aligned}
e_t &= y_t - Z_t a_t, & D_t &= Z_t P_t Z_t' + G_t G_t', & K_t &= T_t P_t Z_t' D_t^{-1}, \\
L_t &= T_t - K_t Z_t, & a_{t+1} &= T_t a_t + K_t e_t, & P_{t+1} &= T_t P_t L_t' + H_t H_t',
\end{aligned}
$$

for $t = s+1, \ldots, n$, where $a_{s+1} = T_s\alpha_s$ and $P_{s+1} = H_s H_s'$. Then, letting $\Lambda_t = H_t H_t'$, we run the simulation smoother:

$$
\begin{aligned}
C_t &= \Lambda_t - \Lambda_t U_t \Lambda_t, & \eta_t &= \Lambda_t r_t + \varepsilon_t, & \varepsilon_t &\sim N(0, C_t), & V_t &= \Lambda_t U_t L_t, \\
r_{t-1} &= Z_t' D_t^{-1} e_t + L_t' r_t - V_t' C_t^{-1}\varepsilon_t, & U_{t-1} &= Z_t' D_t^{-1} Z_t + L_t' U_t L_t + V_t' C_t^{-1} V_t,
\end{aligned}
$$

for $t = n, n-1, \ldots, s+1$, with $r_n = U_n = 0$. Finally, we can draw $\eta_s = \Lambda_s r_s + \varepsilon_s$, $\varepsilon_s \sim N(0, C_s)$ with $C_s = \Lambda_s - \Lambda_s U_s \Lambda_s$. We construct the sample of $\{\alpha_t\}_{t=s+1}^n$ via the state equation using $\{\eta_t\}_{t=s}^{n-1}$ drawn through the simulation smoother. In the case of sampling β, we coordinate the parameters as $Z_t = X_t$, $T_t = I$, $G_t = (A_t^{-1}\Sigma_t, O)$, $H_t = (O, \Sigma_\beta^{1/2})$, for $t = s+1, \ldots, n$, $T_s\alpha_s = \mu_{\beta_0}$, and $H_s = (O, \Sigma_{\beta_0}^{1/2})$.

A.2 Sampling a

For sampling a from its conditional posterior distribution $\pi(a|\beta, h, \Sigma_a, y)$, the expression of the state space form with respect to a is beneficial to see the implementation, namely,

$$\hat{y}_t = \hat{X}_t a_t + \Sigma_t \varepsilon_t, \qquad t = s+1, \ldots, n,$$

$$a_{t+1} = a_t + u_{at}, \qquad t = s, \ldots, n-1,$$

where $a_s = \mu_{a_0}$, $u_{as} \sim N(0, \Sigma_{a_0})$, $\hat{y}_t = y_t - X_t \beta_t$, and

$$\hat{X}_t = \begin{pmatrix} 0 & \cdots & & & & & & 0 \\ -\hat{y}_{1t} & 0 & 0 & \cdots & & & & \vdots \\ 0 & -\hat{y}_{1t} & -\hat{y}_{2t} & 0 & \cdots & & & \\ 0 & 0 & 0 & -\hat{y}_{1t} & \cdots & & & \\ \vdots & & & & \ddots & 0 & \cdots & 0 \\ 0 & \cdots & & & 0 & -\hat{y}_{1t} & \cdots & -\hat{y}_{k-1,t} \end{pmatrix},$$

for $t = s+1, \ldots, n$. Similar to sampling β, we use the simulation smoother with $Z_t = \hat{X}_t$, $T_t = I$, $G_t = (\Sigma_t, O)$, $H_t = (O, \Sigma_a^{1/2})$ for $t = s+1, \ldots, n$, $T_s \alpha_s = \mu_{a_0}$, and $H_s = (O, \Sigma_{a_0}^{1/2})$.

A.3 Sampling h

The state space equations for state variable h belong to a non-linear state space class. Because we assume Σ_h and Σ_{h_0} are diagonal matrices, we make the inference for $\{h_{jt}\}_{t=s+1}^n$ separately for $j \ (= 1, \ldots, k)$. Let y_{it}^* denote the i-th element of $A_t \hat{y}_t$. Then, we can write

$$y_{it}^* = \exp(h_{it}/2)\varepsilon_{it}, \quad t = s+1, \ldots, n,$$

$$h_{i,t+1} = h_{it} + \eta_{it}, \quad t = s, \ldots, n-1,$$

$$\begin{pmatrix} \varepsilon_{it} \\ \eta_{it} \end{pmatrix} \sim N\left(0, \begin{pmatrix} 1 & 0 \\ 0 & v_i^2 \end{pmatrix}\right),$$

where $\eta_{is} \sim N(0, v_{i0}^2)$, and v_i^2 and v_{i0}^2 are the i-th diagonal elements of Σ_h and Σ_{h_0}, respectively, and η_{it} is the i-th element of u_{ht}. We sample $(h_{i,s+1}, \ldots, h_{in})$ from its conditional posterior density by running the multi-move sampler for the non-linear Gaussian state space model, developed by Shephard and Pitt (1997) and Watanabe and Omori (2004).

We divide $(h_{i,s+1}, \ldots, h_{in})$ into $K+1$ blocks and sample each block from its joint posterior density conditioned on the other blocks and parameters. Now we focus on sampling a typical block $(h_{ir}, \ldots, h_{i,r+d})$ from its joint posterior density, (note that $r \geq s+1$, $d \geq 1$, $r+d \leq n$). The conditional posterior distribution is given by

$$\pi(\eta_{i,r-1}, \ldots, \eta_{i,r+d-1}|\theta_i) \propto \prod_{t=r}^{r+d} \frac{1}{e^{h_{it}/2}} \exp\left(-\frac{y_{it}^{*2}}{2e^{h_{it}}}\right) \times \prod_{t=r-1}^{r+d-1} f(\eta_{it}) \times f(h_{i,r+d}), \qquad (5)$$

where

$$f(\eta_{it}) = \begin{cases} \exp\left(-\dfrac{\eta_{is}^2}{2v_{i0}^2}\right) & \text{(if } t = s\text{)}, \\[2ex] \exp\left(-\dfrac{\eta_{it}^2}{2v_i^2}\right) & \text{(if } t \geq s+1\text{)}, \end{cases}$$

$$f(h_{i,r+d}) = \begin{cases} \exp\left\{-\dfrac{(h_{i,r+d+1} - h_{i,r+d})^2}{2v_i^2}\right\} & \text{(if } r+d < n\text{)}, \\[2ex] 1 & \text{(if } r+d = n\text{)}, \end{cases}$$

and $\theta_i = (h_{i,r-1}, h_{i,r+d+1}, y_{ir}^*, \ldots, y_{i,r+d}^*, v_i, v_{i0})$. The posterior draw of $(h_{ir}, \ldots, h_{i,r+d})$ can be obtained by running the state equation with this draw of $(\eta_{i,r-1}, \ldots, \eta_{i,r+d-1})$ given $h_{i,r-1}$.

We sample $(\eta_{i,r-1}, \ldots, \eta_{i,r+d-1})$ from (5) using the AR-MH algorithm (Tierney (1994), Chib and Greenberg (1995)) with the following proposal distribution. Our construction of the proposal density begins with the second-order Taylor expansion of

$$g(h_{it}) \equiv -\frac{h_{it}}{2} - \frac{y_{it}^{*2}}{2e^{h_{it}}},$$

around the certain point $\hat{h}_t$ discussed later, namely,

$$\begin{aligned} g(h_{it}) &\approx g(\hat{h}_{it}) + g'(\hat{h}_{it})(h_{it} - \hat{h}_{it}) + \frac{1}{2}g''(\hat{h}_{it})(h_{it} - \hat{h}_{it})^2 \\ &\propto \frac{1}{2}g''(\hat{h}_{it})\left\{h_{it} - \left(\hat{h}_{it} - \frac{g'(\hat{h}_{it})}{g''(\hat{h}_{it})}\right)\right\}^2, \end{aligned}$$

where

$$g'(\hat{h}_{it}) = -\frac{1}{2} + \frac{y_{it}^{*2}}{2e^{h_{it}}}, \quad g''(\hat{h}_{it}) = -\frac{y_{it}^{*2}}{2e^{h_{it}}}.$$

We use the proposal density formed as

$$q(\eta_{i,r-1},\ldots,\eta_{i,r+d-1}|\theta_i) \; \propto \; \prod_{t=r}^{r+d} \exp\left\{-\frac{(h_{it}^* - h_{it})^2}{2\sigma_{it}^{*2}}\right\} \times \prod_{t=r-1}^{r+d-1} f(\eta_{it}),$$

where

$$\sigma_{it}^{*2} = -\frac{1}{g''(\hat{h}_{it})}, \quad h_{it}^* = \hat{h}_{it} + \sigma_{it}^{*2} g'(\hat{h}_{it}), \tag{6}$$

for $t = r,\ldots,r+d-1$ and $t = r+d$ (when $r+d = n$). For $t = r+d$ (when $r+d < n$),

$$\sigma_{i,r+d}^{*2} = \frac{1}{-g''(\hat{h}_{i,r+d}) + 1/v_i^2}, \tag{7}$$

$$h_{i,r+d}^* = \sigma_{i,r+d}^{*2}\left\{g'(\hat{h}_{i,r+d}) - g''(\hat{h}_{i,r+d})\hat{h}_{i,r+d} + h_{i,r+d+1}/v_i^2\right\}. \tag{8}$$

The choice of this proposal density is derived from its correspondence to the state space model

$$
\begin{aligned}
h_{it}^* &= h_{it} + \zeta_{it}, && t = r,\ldots,r+d, \\
h_{i,t+1} &= h_{it} + \eta_{it}, && t = r-1,\ldots,r+d-1,
\end{aligned}
\tag{9}
$$

$$\begin{pmatrix} \zeta_{it} \\ \eta_{it} \end{pmatrix} \sim N\left(0, \begin{pmatrix} \sigma_{it}^{*2} & 0 \\ 0 & v_i^2 \end{pmatrix}\right), \quad t = r,\ldots,r+d,$$

with $\eta_{i,r-1} \sim N(0, v_i^2)$ when $r \geq 2$, and $\eta_{is} \sim N(0, v_{i_0}^2)$. Given θ_i we draw the candidate point of $(\eta_{i,r-1},\ldots,\eta_{i,r+d-1})$ for the AR-MH algorithm by running the simulation smoother over the state space representation (9).

Now we come to find $(\hat{h}_{ir},\ldots,\hat{h}_{i,r+d})$ which is near the mode of the posterior density for an efficient sampling. We loop the following steps several times to reach near the mode:

1. Initialize $(\hat{h}_{ir},\ldots,\hat{h}_{i,r+d})$.

2. Compute $(h_{ir}^*,\ldots,h_{i,r+d}^*)$, $(\sigma_{ir}^{2*},\ldots,\sigma_{i,r+d}^{2*})$ by (6) and (8).

3. Run the simulation smoother using current $(h_{ir}^*,\ldots,h_{i,r+d}^*)$, $(\sigma_{ir}^{2*},\ldots,\sigma_{i,r+d}^{2*})$ on (9) and obtain $\hat{h}_{it}^* \equiv \mathrm{E}(h_{it}|\theta_i)$ for $t = r,\ldots,r+d$.

4. Replace $(\hat{h}_{ir},\ldots,\hat{h}_{i,r+d})$ by $(\hat{h}_{ir}^*,\ldots,\hat{h}_{i,r+d}^*)$.

5. Go to 2.

Note that the $\mathrm{E}(h_{it}|\theta_i)$ is the product in the simulation smoother as $\Lambda_t r_t$ with $\varepsilon_t = 0$.

Finally, we remark the selection of the blocks. A typical block can be denoted as $(h_{i,k_{j-1}+1}, \ldots, h_{i,k_j})$ for $j = 1, \ldots, K + 1$ with $k_0 = s$ and $k_{K+1} = n$ and Shephard and Pitt (1997) suggests the determination of $(k_1, \ldots, k_K)$, called stochastic knots. It leads to

$$k_j = \mathrm{int}\left[\frac{n(j + U_j)}{K + 2}\right],$$

for $j = 1, \ldots, K$, where U_j is the random sample from the uniform distribution $U[0, 1]$. We randomly choose $(k_1, \ldots, k_K)$ for every iteration of MCMC sampling to obtain a flexible draw of $(h_{i,s+1}, \ldots, h_{in})$.

A.4 Sampling ω

Given β, we can derive the conditional posterior density of Σ_β. Let σ_{β_i} denote the i-th diagonal element of Σ_β. Because we assume that Σ_β is a diagonal matrix, we sample σ_{β_i} independently for $i = 1 \ldots, k$. If we specify the prior as $\sigma_{\beta_i}^{-2} \sim \mathrm{Gamma}(s_{\beta_0}/2, S_{\beta_0}/2)$, we obtain the conditional posterior distribution as $\sigma_{\beta_i}^{-2}|\beta \sim \mathrm{Gamma}(\hat{s}_{\beta_i}/2, \hat{S}_{\beta_i}/2)$ where

$$\hat{s}_{\beta_i} = s_{\beta_0} + n - s - 1, \qquad \hat{S}_{\beta_i} = S_{\beta_0} + \sum_{t=s+1}^{n-1} (\beta_{i,t+1} - \beta_{it})^2,$$

and β_{it} is the i-th element of β_t. The gamma prior is conjugate in this case and the posterior draw is straightforward. We sample the diagonal elements of $\Sigma_a|a$ and $\Sigma_h|h$ in the same way.

References

Baumeister, C., E. Durinck, and G. Peersman (2008). Liquidity, inflation and asset prices in a time-varying framework for the Euro area. Manuscript.

Benati, L. and H. Mumtaz (2005). The 'Great Stability' in the U.K.: good policy or good luck? Manuscript, Bank of England.

Chib, S. (2001). Markov chain Monte Carlo methods: Computation and inference. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, pp. 3569–3649. Amsterdam: North-Holland.

Chib, S. and E. Greenberg (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician 49*, 327–335.

Cogly, T. and T. J. Sargent (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII U.S. *Review of Economic Dynamics 8*, 262–302.

D'Agostino, A., L. Gambetti, and D. Giannone (2008). Macroeconomic forecasting and structural change. Manuscript.

de Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika 82*, 339–350.

Doornik, J. (2006). *Ox: Object Oriented Matrix Programming*. London: Timberlake Consultants Press.

Durbin, J. and S. J. Koopman (2002). Simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*, 603–616.

Eichenbaum, M. (1992). Interpreting the macroeconomic time series facts: the effects of monetary policy – comments. *European Economic Review 36*, 1001–1011.

Fujiwara, I. (2006). Evaluating monetary policy when nominal interest rates are almost zero. *Journal of Japanese and International Economies 20*, 434–453.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics*, Volume 4, pp. 169–188. New York: Oxford University Press.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: inference, development and communication. *Econometric Reviews 18*, 1–126.

Inoue, T. and T. Okimoto (2008). Were there structural breaks in the effects of Japanese monetary policy? Re-evaluating policy effects of the lost decade. *Journal of Japanese and International Economies 22*, 320–342.

Iwata, S. and S. Wu (2006). Estimating monetary policy effects when interest rates are close to zero. *Journal of Monetary Economics 53*, 1395–1408.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies 65*, 361–393.

Kimura, T., H. Kobayashi, J. Muranaga, and H. Ugai (2003). The effect of the increase in monetary base on Japan's economy at zero interest rates: an empirical analysis. In Bank for International Settlement (Ed.), *Monetary policy in a changing environment*, 276–312.

Miyao, R. (2000). The role of monetary policy in Japan: a break in the 1990s? *Journal of Japanese and Interenational Economies 14*, 366–384.

Miyao, R. (2002). The effects of monetary policy in Japan. *Journal of Money, Credit, and Banking 34*, 376–392.

Omori, Y., S. Chib, N. Shephard, and J. Nakajima (2007). Stochastic volatility with leverage: fast likelihood inference. *Journal of Econometrics 140*, 425–449.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies 72*, 821–852.

Sakura, K., H. Sasaki, and M. Higo (2005). Japan's economic fluctuation since the 1990s: fact finding (in Japanese). Bank of Japan Working Paper Series, 05-J-10.

Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics 15*, 645–670.

Shephard, N. (Ed.) (2005). *Stochastic Volatility: Selected Readings*. Oxford: Oxford University Press.

Shephard, N. and M. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika 84*, 653–667.

Sims, C. A. (1992). Interpreting the macroeconomic time series facts: the effects of monetary policy. *European Economic Review 36*, 975–1001.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *Annals of Statistics 21*, 1701–1762.

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

Yano, K. and N. Yoshino (2008). Japanese monetary policy reaction function and time-varying structural vector autoregressions: a Monte Carlo particle filering approach. Manuscript.