Hitotsubashi University

Global COE Hi-Stat Discussion Paper Series 196

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

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

Jouchi Nakajima
Toshiaki Watanabe

July 2011

Hi-Stat Discussion Paper

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

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

Jouchi Nakajima† and Toshiaki Watanabe‡

July 11, 2011

Abstract

This paper applies the time-varying parameter vector autoregressive model to the Japanese economy. The both parameters and volatilities, which are assumed to follow a random-walk process, are estimated using a Bayesian method with MCMC. The recursive structure is assumed for identification and the reversible jump MCMC is used for the ordering of variables. The empirical result reveals the time-varying structure of the Japanese economy and monetary policy during the period from 1981 to 2008 and provides evidence that the order of variables may change by the introduction of zero interest rate policy.

Key words: Bayesian inference, Monetary policy, Reversible jump Markov chain Monte Carlo, Stochastic volatility, Time-varying parameter VAR.
JEL classification: C11, C15, E52

*Financial support from the Ministry of Education, Culture, Sports, Science and Technology of the Japanese Government through Grant-in-Aid for Scientific Research (No.22243026) and the Global COE program "Research Unit for Statistical and Empirical Analysis in Social Sciences" at Hitotsubashi University is gratefully acknowledged. I would like to thank Christopher Sims and the participants in the Conference in honor of Christopher A. Sims "Recent Developments of Dynamic Analysis in Economics– 30 Years after Macroeconomics and Reality" at Seoul National University for their useful comments. All remaining errors are solely my responsibility.

†Department of Statistical Science, Duke University. E-mail: jouchi.nakajima@stat.duke.edu
‡Institute of Economic Research, Hitotsubashi University. E-mail: watanabe@ier.hit-u.ac.jp

1 Introduction

To date, the Japanese economy has experienced several distinct periods of macroeconomic activity and monetary policy. It is relevant as existing literature has shown that the Japanese economy has faced the heteroscedasticity of the exogenous shocks and the transitions in the transmission mechanism. Nakajima et al. (2009) and Nakajima (2011a) show significant changes in relations among major macroeconomic variables in Japan in last several decades, using the time-varying parameter vector autoregressive (TVP-VAR) model,[1] originally proposed by Primiceri (2005).

The standard TVP-VAR models exploit the recursive structure in decomposition of the covariance matrix for identifying structural shocks of the system, and therefore, the ordering of variables is of interest and relevant in empirical studies of the TVP-VAR models. As discussed by Primiceri (2005), one strategy to assess this issue is to introduce uncertainty of the ordering of variables. In a Bayesian inference, we consider a prior on the model space where each model has a different ordering and explore the posterior probability of different ordering. This strategy can be accomplished by the reversible jump Markov chain Monte Carlo (RJMCMC) method (Green (1995)).[2] Although the RJMCMC is a useful approach to assess the ordering of the economic variables, little has been done in the context of the TVP-VAR models. This paper develops an efficient RJMCMC algorithm for the TVP-VAR model and provides the empirical analysis of the Japanese macroeconomy and monetary policy.

Let M denote a set of competing models, and suppose that for every model $m \in M$, we have a vector $\theta_m \in \Theta_m$ of unknown parameters. In the RJMCMC, we explore the model space by simulating the joint posterior distribution of (m, θ_m) given data using the Mertropolis-Hasting (MH) algorithm. In theory, a choice of proposal distribution does not affect the results, but it may affect the convergence speed of the MCMC. There is no general way to construct an adequate proposal distribution. One straightforward

[1]See Nakajima (2011a) for the literature survey of the TVP-VAR models. As the study for the empirical study of Japanese macroeconomy, see Miyao (2000, 2002) for standard VAR models, Fujiwara (2006) and Inoue and Okimoto (2008) for Markov-switching VAR models and Kimura et al. (2003) for time-varying coefficient VAR models, Nakajima et al. (2010) for TVP-VAR models.

[2]See Vrontos et al. (2000) and Kasuya and Takagawa (2001) for applications of the RJMCMC algorithm in econometrics. See also Brooks et al. (2003) for efficient construction of the RJMCMC algorithm

approach commonly used in literature (e.g., Vrontos et al. (2000) and Kasuya and Takagawa (2001)) is to construct the proposal distribution using the sample obtained from a preliminary MCMC run for each model in M. That is, we firstly generate the sample of θ_m via the MCMC for its posterior distribution of every single model $m \in M$. Then, we run the RJMCMC algorithm where we propose the candidate of parameters from the preliminary-constructed posterior distribution of the proposed model. This strategy may automatically assure an adequate proposal distribution, but the computational cost would become intractable when the number of competing models increase. For the TVP-VAR models, Primiceri (2005) develops the RJMCMC algorithm based on this strategy where the proposal distributions are constructed using the preliminarily MCMC run including a huge dimensional state variables. For k-variate TVP-VAR models, if we include all the permutation of the variables to examine the ordering, we explore $k!$ competing models and are forced to run $k!$ different TVP-VAR models in advance. For the preliminary run, it would not be necessary to run the MCMC until the convergence can be reached, because the proposal distribution of the RJMCMC may only require roughly approximated distribution for the conditional posterior distribution. However, there is no general guidance for the degree of preliminary runs to construct the proposal distribution for the RJMCMC.

To avoid this problem, this paper proposes a novel approach to construct the proposal distribution based on the current point of the parameters for searching the ordering of the variables in the TVP-VAR models. The candidate of the parameters is generated from the conditional posterior distribution given the permuted sample of the current point. It searches the model space by jointly generating the posterior sample of the parameters including state variables and requires no preliminary run for every single model. We illustrate our method by fitting four variables TVP-VAR model to the Japanese macroeconomic data and provide empirical results of model search and parameter estimates under model uncertainty.

The paper is organized as follows. In Section 2, we review the standard MCMC algorithm for the TVP-VAR model. Section 3 develops the new method of the RJMCMC for searching the ordering of the variables in the TVP-VAR model. Section 4 provides the empirical analysis of the proposed RJMCMC algorithm for the Japanese macroeconomic data. Finally, Section 5 concludes.

2 TVP-VAR models

2.1 Model specification

We consider the TVP-VAR model formulated by

$$y_t = c_t + B_{1t}y_{t-1} + \cdots + B_{st}y_{t-s} + e_t, \quad e_t \sim N(0, \Omega_t),$$

where y_t is a $k \times 1$ vector of observed variables, c_t is a $k \times 1$ vector of intercepts, B_{it}'s are $k \times k$ matrices of time-varying coefficients, and Ω_t is a $k \times k$ variance-covariance matrix. Following a standard recursive identification commonly used for the TVP-VAR models, we take a triangular reduction of Ω_t defined by $A_t \Omega A_t' = \Sigma_t \Sigma_t'$, where $\Sigma_t = \mathrm{diag}(\sigma_{1t}, \ldots, \sigma_{kt})$ is $k \times k$ diagonal matrix of time-varying variances for idiosyncratic shocks, and A_t is a $k \times k$ lower-triangular matrix of covariance components, defined by

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

Define the $k(k+1)s \times 1$ vector β_t by stacking the set of c_t and B_{it} by rows and by order $j = 1, \ldots s$, and the $k \times k(k+1)s$ matrix $X_t = I_k \otimes (1, y_{t-1}', \ldots, y_{t-s}')$, where $\otimes$ denotes the Kronecker product. Then the model can be written as

$$y_t = X_t \beta_t + A_t^{-1} \Sigma_t \varepsilon_t, \quad \varepsilon_t \sim N(0, I).$$

Let a_t be the vector ($q \times 1$, with $q = k(k-1)/2$) of the strictly lower-triangular elements of A_t (stacked by rows), and define $h_t = (h_{1t}, \ldots, h_{kt})'$, with $h_{it} = \log \sigma_{it}^2$, for $i = 1, \ldots, k$. The dynamics of the parameters (β_t, a_t, h_t) is specified as

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

for $t = s+1, \ldots, n$, where $\beta_{s+1} \sim N(\mu_{\beta 0}, V_{\beta 0})$, $a_{s+1} \sim N(\mu_{a0}, V_{a0})$, and $h_{s+1} \sim N(\mu_{h0}, V_{h0})$, with each of the matrices $(V_a, V_h, V_{\beta 0}, V_{a0}, V_{h0})$ diagonal.

As discussed by Primiceri (2005) and Nakajima (2011a), we remark two important aspects on modeling the TVP-VAR models. First, the assumption of a lower-triangular matrix for A_t is recursive identification for the VAR system. This specification is simple and widely used for both time-invariant and time-varying VAR models, although an estimation of structural models may require a more complicated identification to extract adequate implications for the economic structure, as pointed out by Christiano et al. (1999) and other studies. To focus on this issue, this current paper explores the ordering of the variables in y_t using the RJMCMC algorithm described below.

Second, the parameters are assumed to follow a random walk process which is non-stationary in theory. Because the TVP-VAR model has a number of parameters to estimate, we often employ the random walk assumption, which effectively decreases the number of parameters. Most of studies commonly assume the random walk process in the TVP-VAR models. The non-stationarity assumption of the time-varying parameters would be sometimes not adequate in deducing economic structural models, but as far as we fit the model for the finite sample periods, this assumption is acceptable and rather advantageous because it can capture permanent shifts of the parameter which would be plausibly observed in real data.

2.2 MCMC algorithm

We take a Bayesian inference to estimate the TVP-VAR models via the MCMC methods. The goal of the MCMC methods is to assess the joint posterior distribution of the parameters of interest under certain prior probability densities that the researchers set in advance. Given data, we repeatedly sample a Markov chain whose invariant (stationary) distribution is the posterior distribution (see e.g., Chib (2001), Koop (2003), Geweke (2005) and Gamerman and Lopes (2006)).

Define $\theta = (\beta, a, h, V)$ with $\beta = \{\beta_t\}_{t=s+1}^n$, $a = \{a_t\}_{t=s+1}^n$, $h = \{h_t\}_{t=s+1}^n$, and $V = (V_\beta, V_a, V_h)$. We assume the following priors: $V_\beta \sim IW(\nu_{\beta 0}, W_0)$, $v_{ai}^2 \sim IG(\nu_{a0}, S_{a0})$, and $v_{hi}^2 \sim IG(\nu_{h0}, S_{h0})$, where v_{ai}^2 and v_{hi}^2 are the diagonal elements in V_a and V_h, respectively, IW denotes the inverse Wishart distribution and IG denotes the inverse gamma distribution. Given observation $y = \{y_1, \dots, y_n\}$, the full joint posterior distribution $\pi(\theta|y)$ is explored using MCMC technique. We briefly describe the MCMC algorithm as follows (see the detail in Primiceri (2005) and Nakajima (2011a)).

1. Generate $\beta \sim \pi(\beta|a, h, V_\beta, y)$, based on the state space model,

$$\begin{aligned} y_t &= X_t\beta_t + A_t^{-1}\Sigma_t\varepsilon_t, & (1) \\ \beta_{t+1} &= \beta_t + u_{\beta t}. & (2) \end{aligned}$$

   The state variable β is generated using the simulation smoother (e.g. de Jong and Shephard (1995), Durbin and Koopman (2002)).

2. Generate $a \sim \pi(a|\beta, h, V_a, y)$, based on the state space model,[3]

$$\begin{aligned} \hat{y}_t &= \hat{X}_t a_t + \Sigma_t\varepsilon_t, & (3) \\ a_{t+1} &= a_t + u_{at}. & (4) \end{aligned}$$

   The state variable a is generated using the simulation smoother as in Step 1.

3. Generate $h_i \sim \pi(h_i|\beta, a, V_h, y)$ for $i = 1, \dots, k$, based on the univariate stochastic volatility model,

$$\begin{aligned} w_{it} &= \exp(h_{it}/2)\varepsilon_{it}, & (5) \\ h_{i,t+1} &= h_{it} + u_{hit}, & (6) \end{aligned}$$

   where w_{it} is the i-th element of $w_t = A_t\hat{y}_t$. The log-volatility h is generated using the multi-move sampler for the stochastic volatility models (Shephard and Pitt

[3]Define $\hat{y}_t \equiv (\hat{y}_{1t}, \dots, \hat{y}_{kt})' = y_t - X_t\beta_t$, and

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

(1997), Watanabe and Omori (2004)).

4. Generate $V \sim \pi(V|\beta, a, h)$, based on the conjugate conditional posterior distributions $\pi(V_\beta|\beta)$, $\pi(V_a|a)$, and $\pi(V_h|h)$.

3 RJMCMC algorithm for TVP-VAR models

3.1 Concept

In a Bayesian inference, a posterior probability of the model m is given by

$$\pi(m|y) = \frac{\pi(m) \int_{\Theta_m} f(y|m, \theta_m)\pi(\theta_m|m)d\theta_m}{\sum_{m' \in M} \pi(m') \int_{\Theta_{m'}} f(y|m', \theta_{m'})\pi(\theta_{m'}|m')d\theta_{m'}}, \tag{7}$$

where $f(y|m, \theta_m)$ is the likelihood given data y under the model m, $\pi(m)$ is the prior probability for the model m, and $\pi(\theta_m|m)$ is the prior distribution of θ_m under the model m. The evaluation of the integrals in (7) is often a challenging problem. Green (1995) proposes the RJMCMC algorithm to generate sample from the joint posterior distribution $\pi(m, \theta_m|y)$ by exploiting the MH algorithm, which produces the estimates of $\pi(m|y)$. The RJMCMC algorithm searches model space and parameter space jointly based on the Markov chain whose kernel satisfies the detailed balance condition to ensure the convergence to the limiting distribution $\pi(m, \theta_m|y)$.

Given the current point (m, θ_m) of the RJMCMC, suppose that we propose a move to the model m^* with probability $j(m, m^*)$. Suppose that the models m and m^* have the unknown parameters θ_m and $\theta^*_{m^*}$ with the dimensions $d(\theta_m)$ and $d(\theta^*_{m^*})$, respectively. As an auxiliary variable to bridge between the parameter spaces of θ_m and $\theta^*_{m^*}$, we generate u from some proposal distribution $q(u|\theta_m, m, m^*)$. Then, we set $(\theta^*_{m^*}, u^*) = g_{m,m^*}(\theta_m, u)$, where g_{m,m^*} is an invertible function such that $g_{m^*,m} = g^{-1}_{m,m^*}$, which follows $d(\theta_m) + d(u) = d(\theta^*_{m^*}) + d(u^*)$. Finally, we accept the move of $(m, \theta_m) \to (m^*, \theta^*_{m^*})$ with the MH acceptance probability $\alpha(m, m^*) = \min\{1, R\}$, where

$$R = \frac{f(y|m^*, \theta^*_{m^*})\pi(\theta^*_{m^*}|m^*)\pi(m^*)j(m^*, m)q(u^*|\theta^*_{m^*})}{f(y|m, \theta_m)\pi(\theta_m|m)\pi(m)j(m, m^*)q(u|\theta_m)}|J|, \tag{8}$$

where $J = \partial(\theta^*_{m^*}, u^*)/\partial(\theta_m, u)$ is the Jacobian of the transformation.

The key issue to implement the RJMCMC algorithm is the choice of the function g for the proposal of the next point of the unknown parameters. Primiceri (2005) (and other authors, e.g., Vrontos et al. (2000), Kasuya and Takagawa (2001) in different context) suggests that $(\theta^*_{m^*}, u^*) = (\theta_m, u)$, with $d(\theta_m) = d(u^*)$, $d(\theta^*_{m^*}) = d(u)$, $q(u|\theta_m, m, m^*) = q(u|m^*)$, and $q(u^*|\theta^*_{m^*}, m^*, m) = q(u^*|m)$. That is, the proposal distribution does not depend on the current point. The independent proposal distributions $q(u|m^*)$ and $q(u^*|m)$ are constructed using the sample obtained from a preliminary MCMC run for each model in M. We generate the sample of θ_m via the MCMC for its posterior distribution of every single model $m \in M$, which requires considerable computational burden when the number of competing model is large.

We instead propose an efficient algorithm of the RJMCMC for the TVP-VAR models. We construct the proposal distribution using the current point. The candidate of the parameters is generated from the conditional posterior distribution given the permuted sample of the current point. It searches the model space by jointly generating the posterior sample of the parameters including state variables with no requirement of preliminary run for every single model. We describe the detail in the next subsection.

3.2 Proposed algorithm

We consider the RJMCMC algorithm to explore the ordering of the variables in the TVP-VAR model.[4] Our strategy to construct the efficient RJMCMC algorithm is summarized as follows.

1. Propose the new point $(m^*, \theta^*_{m^*})$.

   (a) Propose the move $m \to m^*$, with probability $j(m, m^*)$.

   (b) Propose $(\theta_m, u) \to (u^*, \theta^*_{m^*})$ with auxiliary variables $u = (\beta^*, a^*, h^*, V^*)$ and $u^* = (\beta, a, h, V)$. Note that $|J| = 1$. The proposal density is given by

$$\begin{aligned} & q(\beta^*, a^*, h^*, V^*|\beta, h, V, y) \\ = \;& q(\beta^*|a_0, \tilde{h}, \tilde{V}_\beta, \tilde{y}) \times q(a^*|\beta^*, \tilde{h}, \bar{V}_a, \tilde{y}) \times q(h^*|\beta^*, a^*, \tilde{V}_h, \tilde{y}) \times q(V^*|\beta^*, a^*, h^*), \end{aligned}$$

[4] We fix the lag length. Huerta and West (1999), Prado and Huerta (2002) develop the methodology for model order (*i.e.*, the number of autoregressive lags) uncertainty for autoregressive models.

where a_0 denotes the set of zero elements whose size is equal to that of a, and $\tilde{x}$ denotes the permutation of $x = \{y, h, V_\beta, V_h\}$ according to the order change of $m \to m^*$. Note that the proposal density is not conditional on a. Specifically, the candidate is generated by the following steps.

i. Generate $\beta^* \sim q(\beta^* | a_0, \tilde{h}, \tilde{V}_\beta, \tilde{y})$, based on the state space model (1)-(2) with $a_t = 0$ ($i.e.$, $A_t = A_t^{-1} = I$), for all t.

ii. Generate $a^* \sim q(a^* | \beta^*, \tilde{h}, \bar{V}_a, \tilde{y})$, based on the state space model (3)-(4), where $\bar{V}_a = \bar{v}_a^2 I$ with $\bar{v}_a^2 = (v_{a1}^2 + \cdots + v_{aq}^2)/q$, $i.e.$, the average of the variance in the current point of V_a.

iii. Generate $h^* \sim q(h^* | \beta^*, a^*, \tilde{V}_h, \tilde{y})$, based on the stochastic volatility model (5)-(6).

iv. Generate $V^* \sim q(V^* | \beta^*, a^*, h^*) = \pi(V^* | \beta^*, a^*, h^*)$.

2. Accept the candidate with MH probability $\alpha(m, m^*) = \min\{1, R\}$, where

$$R = \frac{f(y|m^*, \theta^*_{m^*})\pi(\theta^*_{m^*}|m^*)q(u^*|\theta^*_{m^*})}{f(y|m, \theta_m)\pi(\theta_m|m)q(u|\theta_m)}.$$

We generate the next point of unknown parameters basically following the original MCMC algorithm for the single TVP-VAR model as described in Section 2.2, by permuting the parameters including state variables based on the order change of $m \to m^*$. Regarding time-varying parameters, a different ordering would involve a different dynamics, but we consider it is plausible that our proposal distribution is close to the posterior distribution of the proposed model. The crucial point different between models m and m^* is the time-varying parameter a for the simultaneous effects of structural shocks. Because a different order involves a different structure of A_t by construction, it is not reliable to use the current point of a in θ. Therefore, we generate β^* conditional on $a_t = 0$ in Step 1.(b).i. Our experiments show successful implementation based on this proposed algorithm, as shown using the real data in the next section.

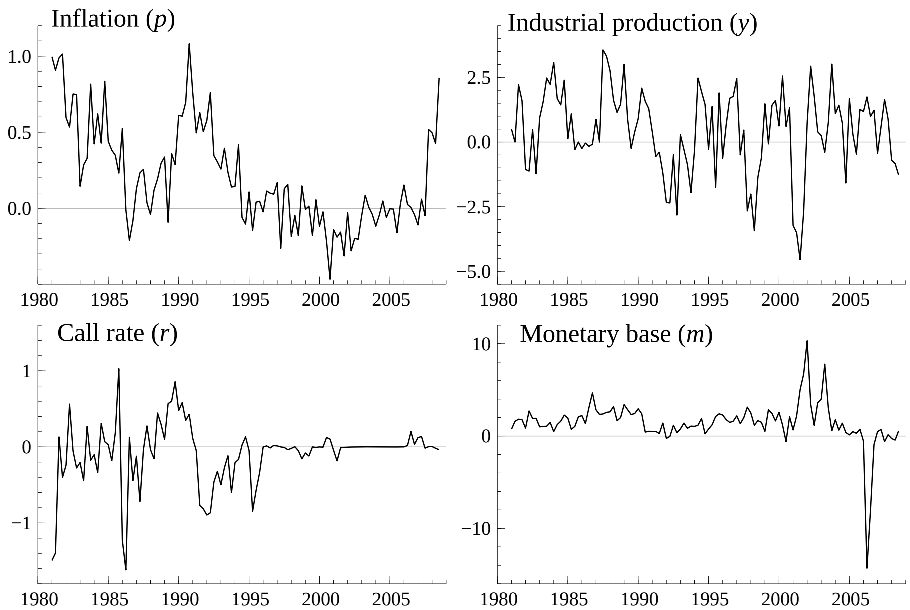

Figure 1: Japanese macroeconomic data (1981/1Q to 2008/3Q).

4 Empirical findings for Japanese economy and monetary policy

4.1 Data and setup

In this section, we apply our proposed RJMCMC algorithm for the TVP-VAR model to the Japanese macroeconomic data. The dataset is quarterly and the sample period is from 1981/1Q to 2008/3Q. We consider a four-variable TVP-VAR model which includes: inflation rate (p), industrial production (y), nominal short-term interest rate (r), and monetary base (m), exhibited in Figure 1.[5] The VAR lag is set equal to two, which yields the highest marginal likelihood for the same data in Nakajima et al. (2009).

[5]The inflation rate is taken from the CPI (consumer price index, general excluding fresh food, and seasonally adjusted). For the CPI, the effects of the increase in the consumption tax are removed for 1989/2Q and 1997/2Q. 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.

We consider all the permutation of the four variables, *i.e.*, $|M| = 4! = 24$. The model prior probability is assumed as $\pi(m) = |M|^{-1}$ for all $m \in M$. In the RJMCMC, we propose a move to the different model, $m \to m^*$ ($m \neq m^*$), with probability $j(m, m^*) = (1 - S)(|M| - 1)^{-1}$, and a stay at the same model with probability $j(m, m) = S$, where $0 \leq S < 1$. From our experiments we suggest that $S = 0.3$ for an appropriate balance of Markov chain mixing and acceptance rate of the MH algorithm in the RJMCMC, which is taken throughout this paper.[6]

The following priors are assumed: $V_\beta \sim IW(25, 0.01)$, $v_{ai}^2 \sim IG(4, 0.02)$, and $v_{hi}^2 \sim IG(4, 0.02)$. For the initial state of the time-varying parameters, $\beta_{s+1} \sim N(0, 10I)$, $a_{s+1} \sim N(0, 10I)$, and $h_{s+1} \sim N(0, 50I)$. We generate K = 100,000 sample after the initial 10,000 sample are discarded.[7] This burn-in period is determined based on the model trace and the convergence diagnostics for the common parameters.

4.2 Model uncertainty

Figure 2 exhibits the estimation results of the RJMCMC algorithm. Figure 2(i) plots the trace plot of the model search, showing a better mixing of the chain and adequately covering the model space. The MH acceptance probability is 29.7% for the move of the model, which would be enough level of the acceptance for the satisfactory model search. Figure 2(ii) plots the trajectory of the posterior model probability through the RJMCMC run, which clearly shows the convergence of the RJMCMC algorithm. Figure 2(iii) displays the posterior model probability (the model index listed in Appendix A.1). Table 1 reports the top three models selected from the RJMCMC algorithm. The result indicates considerable model uncertainty for the ordering of the variables in the TVP-VAR models, although it is evident that the similar models are selected in the top three models.

For further analysis, we divide the data into two subsample periods: (i) 1981/1Q to 1995/4Q, and (ii) 1996/1Q to 2008/3Q, based on the evidence of structural change

[6]Some authors (e.g., Vrontos et al. (2000), Kasuya and Takagawa (2001)) implement the RJMCMC where the algorithm always proposes a different model from the current model (*i.e.*, $S = 0$). In our proposed algorithm, our experience shows that the mixing is better when we take the positive stay probability ($S > 0$), which allows the procedure to update the parameters some time based on the current single model.

[7]The computational results are generated using Ox version 5.0 (Doornik (2006)).

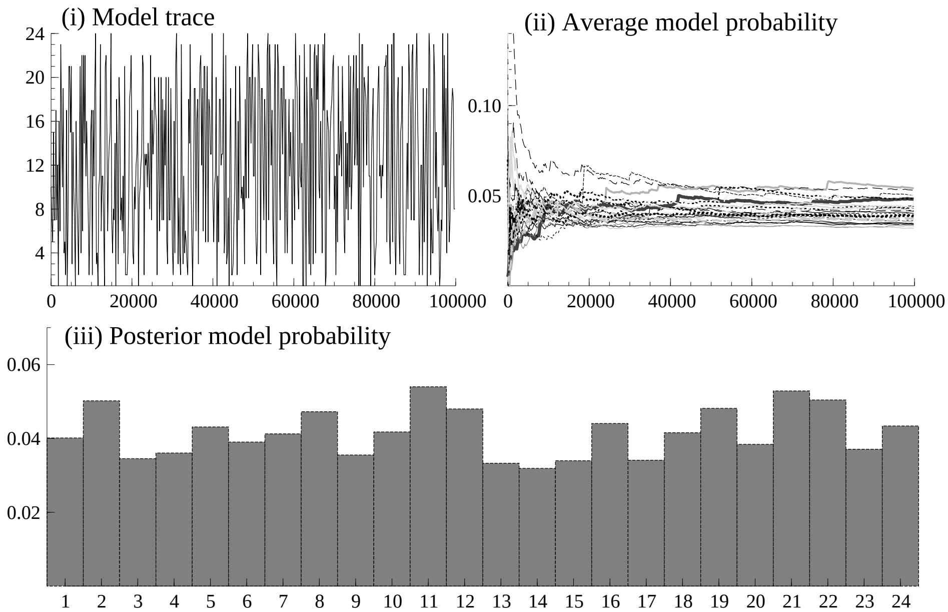

Figure 2: Estimation results of the RJMCMC algorithm: (i) trace plot of model search (top, left), (ii) on-line trajectory of the posterior model probability (top, right), and (iii) histogram of posterior model probability (bottom). The model index is listed in Appendix A.1.

in the Japanese macroeconomy and monetary policy suggested by existing literature (e.g., Fujiwara (2006), Inoue and Okimoto (2008), Kimura et al. (2003), Nakajima et al. (2009)). Obviously, the second subsample includes the periods of the zero interest rate policy (February 1999 to August 2000) and the quantitative easing policy (March 2001 to March 2006). After the Bank of Japan lowered the official discount rate from 1.0% to 0.5% in September 1995, the overnight call rate stayed in the very low level during the second subsample period until the raising of the target overnight call rate to 0.25% from the zero interest rate policy in July 2006.

We apply the RJMCMC algorithm to these two subsample periods, and the estimated top three models are reported in Table 1.[8] Evidently, the selected models differ between the subsample periods, which indicates that the ordering of variables may change by the introduction of the zero interest rate policy. The top three models in the first subsample have inflation rate at the last of the ordering, which indicates the price level would be not

[8]Figure 6 in Appendix A.2 exhibits the histograms of the posterior model probability for these subsample periods.

Rank	Full sample 1981-2008 Order	Prob.	Sub samples (i)1981-1995 Order	Prob.	(ii)1996-2008 Order	Prob.
1	(y, m, p, r)	0.054	(y, r, m, p)	0.068	(m, y, p, r)	0.071
2	(m, y, p, r)	0.053	(r, y, m, p)	0.066	(m, y, r, p)	0.048
3	(m, y, r, p)	0.050	(y, m, r, p)	0.052	(p, y, m, r)	0.046

Table 1: Posterior model probability: top 3 models.

so reactive in terms of simultaneous relation of structural shocks. The second subsample clearly prefers the top model, considerably dominating other models, where the call rate places at the last in the recursive identification, because the call rate does not play a role as the monetary policy instrument in most of the second period.

4.3 Model comparison

We explore the model comparison between the model with the ordering uncertainty and the top model selected by the RJMCMC estimation result. For this purpose, we estimate the marginal likelihood based on the harmonic mean method (e.g., Geweke (1999)). The simulation-based harmonic mean estimator, denoted by $\hat{m}(y)$, is computed by

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

where $\vartheta = (\beta, a, h)$, $f(y|V^{(j)}, \vartheta^{(j)})$ and $\pi(V^{(j)})$ denote the likelihood function and prior density, respectively and K is the iteration size of the MCMC. If the fraction $g(V)/f(y|V, \vartheta)\pi(V)$ is bounded above, the approximation is simulation consistent and the rate of convergence is likely to be practical.

The $g(V)$ can be any p.d.f. with support contained in the parameter space of the model. Geweke (1999) recommends the choice of a certain g for the modified harmonic mean estimator to guarantee the boundness of this fraction as follows. Consider the

	Full sample	Sub samples	
	1981-2008	(i)1981-1995	(ii)1996-2008
(a) Under the ordering uncertainty (RJMCMC)			
TVP	-434.62	-264.56	-385.99
(b) Top model			
TVP	-291.91	-256.89	-214.11
STVP1	-360.24	-261.49	-238.61
STVP2	-498.25	-264.44	-280.66
CP	-544.54	-288.02	-256.81

Table 2: Estimated marginal likelihoods (ML) in logarithm scale. TVP refers to: time-varying parameter model, STVP: semi time-varying parameter model (1: time-varying h_t, 2: time-varying β_t and a_t), CP: constant parameter model. The top model is selected in Table 1.

normal density with the tails truncated,

$$
\begin{aligned}
g(V^{(j)}) &= \frac{1}{\tau(2\pi)^{p/2}|\hat{W}|^{1/2}} \exp\left\{-\frac{1}{2}(v^{(j)}-\hat{v})'\hat{W}^{-1}(v^{(j)}-\hat{v})\right\} \\
&\quad \times I\left[(v^{(j)}-\hat{v})'\hat{W}^{-1}(v^{(j)}-\hat{v}) \le \chi^2_\tau(p)\right],
\end{aligned}
$$

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 unknown parameters in V, v is the $p \times 1$ stacked vector of the parameters in V, and $\chi^2_\tau(p)$ denotes the τ percentile of the Chi-square distribution with p degrees of freedom. The idea is to cut off the tails so that samples that drop in that potentially problematic regions are avoided for the computation of the marginal likelihood. We set $\hat{v}$ and $\hat{W}$ equal to the sample mean and covariance matrix computed from the posterior draws $\{V^{(j)}\}_{j=1}^K$, and $\tau = 0.99$ in this paper.[9] Even under the ordering uncertainty, we can compute the marginal likelihood using the harmonic mean method through the RJMCMC algorithm.

Table 2 reports the estimated marginal likelihoods for the model under the ordering uncertainty and the top model. It is evident that the marginal likelihood of the top model is higher than the model under the ordering uncertainty for all three sample

[9] We experimented $\tau = 0.95$ and 0.90, but the difference of the estimated marginal likelihoods are negligible, as mentioned by Schorfheide (2000). For more details, see Appendix B in the updated version of Nakajima et al. (2009).

periods, which implies that the ordering uncertainty does not contribute the model fit in terms of the marginal likelihood for our data, presumably because the TVP-VAR model had a considerable uncertainty for the ordering as shown in Section 4.2.

To confirm the superiority of the time-varying parameters for the top model, we also compute the marginal likelihoods for the reduced models of TVP-VAR model. We consider semi time-varying parameter (STVP) models that partially allow the parameters time-varying in two ways. The STVP1 model includes the time-varying h_t, but β_t and a_t are constant over time. In contrast, the STVP2 model includes the time-varying β_t and a_t, but h_t is constant. A constant parameter (CP) model is defined as the standard VAR model where all the parameters are constant over time.

Table 2 reports the marginal likelihood for these models. The TVP-VAR model obviously dominates the other reduced models for all three sample periods, which implies that the Japanese macroeconomic variables yield significant changes in relations for the recent three decades. It is interesting that the STVP1 model performs better than the STVP2 model, which indicates that the changes in volatility for structural shocks are relevant, compared to the changes in the VAR coefficients and simultaneous effects of structural shocks. In particular, the marginal likelihood of the STVP2 model is lower than that of the CP model for the second subsample. We focus on the macroeconomic dynamics of the Japanese economic variables in the next subsection.

4.4 Estimation results for macroeconomic dynamics

One of the advantages using the RJMCMC algorithm is that one can obtain the parameter estimates under model uncertainty as known as model averaging. Figure 3 plots the estimated posterior means of stochastic volatility $\sigma_{it} = \exp(h_{it}/2)$ under model uncertainty based on the sample from the RJMCMC run and under the selected top model (y, m, p, r). Clearly, the levels of the volatility are almost same between two methods, although the uncertainty of the estimates is larger in the RJMCMC algorithm than that from the top model due to the model uncertainty.

Figure 4 plots the time-varying impulse response of the top model (y, m, r, p),[10] and

[10]Following Nakajima (2011a), the impulse response is computed by fixing an initial shock size equal to the time-series average of stochastic volatility over the sample period, and using the simultaneous relations at each point in time. To compute the recursive innovation of the variable, the estimated time-

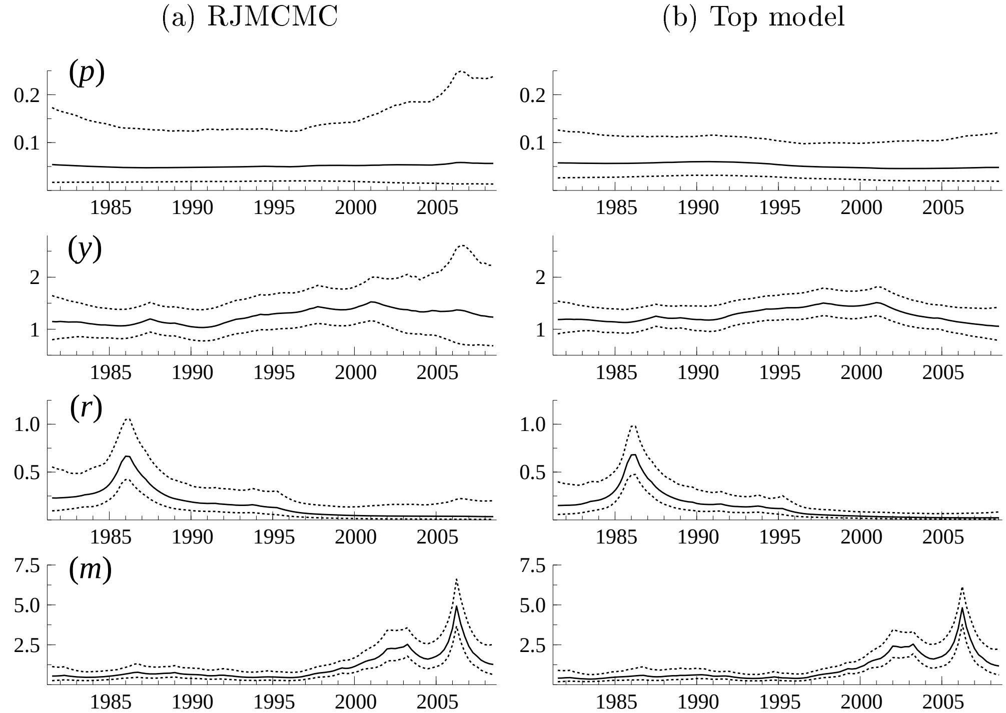

Figure 3: Estimated stochastic volatility $\sigma_{it} = \exp(h_{it}/2)$ for idiosyncratic shocks (a) under model uncertainty using the RJMCMC algorithm (left) and (b) under the top model (right). Posterior means (solid) and one-standard-deviation bands (dotted).

Figure 5 illustrates the impulse response in selected time points; 1988/2Q, 1995/4Q and 2005/4Q. The responses indicate considerable changes in macroeconomic behavior related to the monetary policy in these three decades, as discussed by Nakajima et al. (2009). The responses of call rate to the shocks of other variables are diminishing towards zero during the zero interest rate periods. In Figure 5, it is remarkable that the response of industrial production to monetary base shock ($\varepsilon_m \to y$) is considerably small in 2005/4Q when the quantitative easing policy was implemented.

varying coefficients are used from the current date to future periods. Around the end of the sample period, the coefficients are set constant in future periods for convenience.

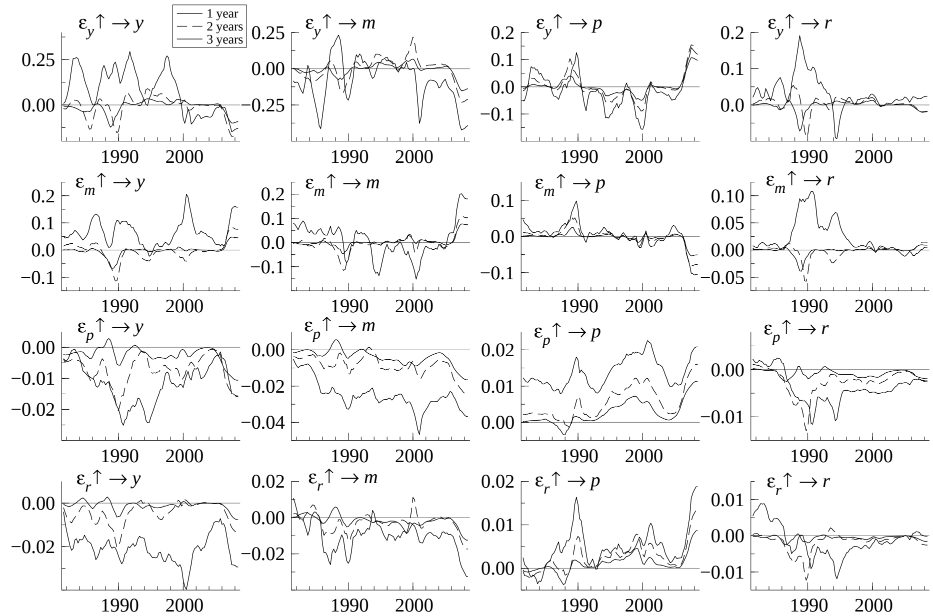

Figure 4: Time-varying impulse response from the TVP-VAR model of the top ordering (y, m, p, r) for one-year (solid), two-year (dashed) and three-year (dotted) horizons.

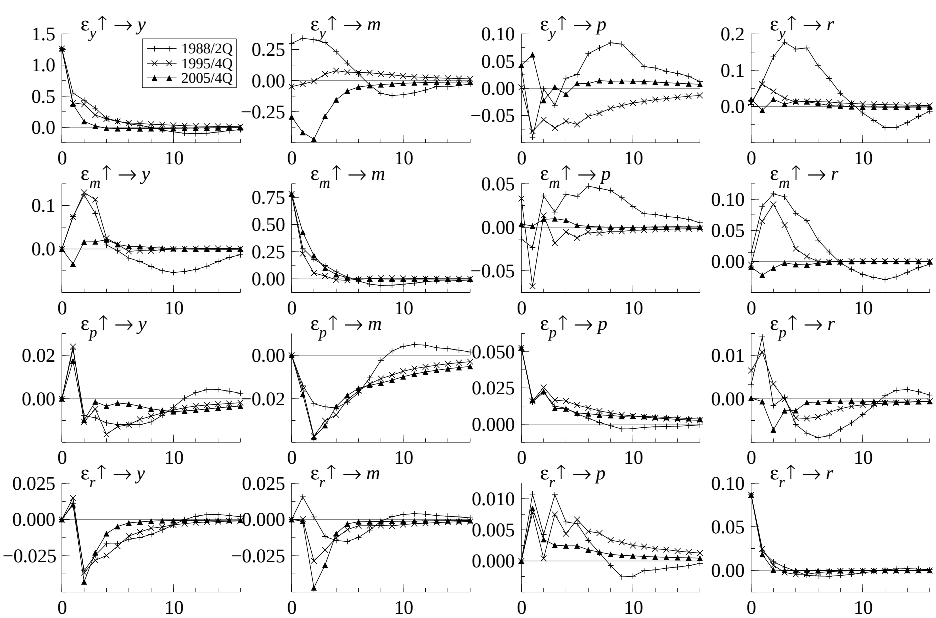

Figure 5: Impulse response from the TVP-VAR model of the top ordering (y, m, p, r) in 1988/2Q, 1995/4Q and 2005/4Q.

5 Concluding remarks

This paper proposes the efficient RJMCMC algorithm for the TVP-VAR model to explore the ordering of the variables and provides empirical evidence of the Japanese macroeconomy and monetary policy. The empirical result reveals the time-varying structure of the Japanese economy and monetary policy during the period from 1981 to 2008 and reveals that the order of variables may change by the introduction of zero interest rate policy.

For identification issues in the TVP-VAR models, Benati and Surico (2008), Baumeister and Benati (2010) and Franta (2011) develop the sign restriction approach. From another perspective, Nakajima (2011b) proposes an explicit zero lower bound of the short-term interest rate in the TVP-VAR models, and Nakajima and West (2010) provide the latent threshold technique to induce an implicit zero interest rate constraint. Including these models, there remain further analyses of model search using the RJMCMC algorithm for future research.

Appendix

A.1. Model index for the Japanese macroeconomic data

No.	Order
1	(p, y, r, m)
2	(p, y, m, r)
3	(p, r, y, m)
4	(p, r, m, y)
5	(p, m, y, r)
6	(p, m, r, y)
7	(y, p, r, m)
8	(y, p, m, r)
9	(y, r, p, m)
10	(y, r, m, p)
11	(y, m, p, r)
12	(y, m, r, p)
13	(r, p, y, m)
14	(r, p, m, y)
15	(r, y, p, m)
16	(r, y, m, p)
17	(r, m, p, y)
18	(r, m, y, p)
19	(m, p, y, r)
20	(m, p, r, y)
21	(m, y, p, r)
22	(m, y, r, p)
23	(m, r, p, y)
24	(m, r, y, p)

Table 3: Model index for the ordering of the Japanese macroeconomic data (p: inflation rate, y: industrial production, r: call rate, m: monetary base).

A.2. Posterior model probability for the subsample periods

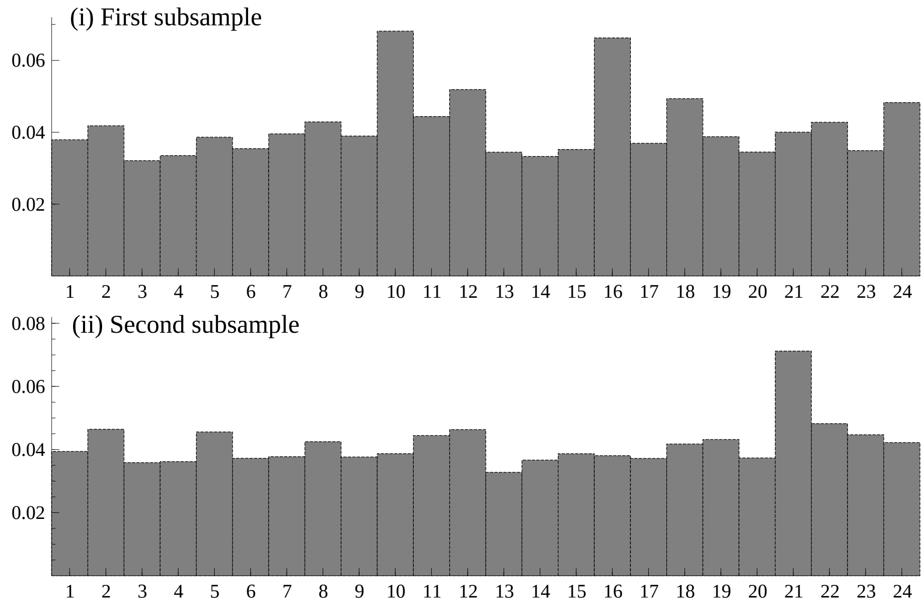

Figure 6: Histograms of posterior model probability for the first (top) and second (bottom) subsample periods. The model index is listed in Appendix A.1.

References

Baumeister, C. and L. Benati (2010). Unconventional monetary policy and the great recession – Estimating the impact of a compression in the yield spread at the zero lower bound. Working Paper Series 1258, European Central Bank.

Benati, L. and P. Surico (2008). Evolving U.S. monetary policy and the decline of inflation predictability. *Journal of the European Economic Association 6*, 643–646.

Brooks, S. P., P. Giudici, and G. O. Roberts (2003). Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions. *Journal of the Royal Statistical Society, Series B 65*, 3–55.

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.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999). Monetary policy shocks: What have we learned and to what end? In J. B. Taylor and M. Woodford (Eds.), *Handbook of Macroeconomics*, Volume 3, pp. 65–148. Amsterdam: Elsevier Science B. V.

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.

Franta, M. (2011). Identification of monetary policy shocks in Japan using sign restrictions within the TVP-VAR framework. Mimeo.

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

Gamerman, D. and H. F. Lopes (2006). *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

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

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics.* Wiley.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*, 711–732.

Huerta, G. and M. West (1999). Priors and component structures in autoregressive time series models. *Journal of the Royal Statistical Society, Series B 61*, 881–899.

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.

Kasuya, M. and I. Takagawa (2001). Model uncertainty of real exchange rate forecast over mid-term horizons. Working Paper, No.01-E-23, Bank of Japan.

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.

Koop, G. (2003). *Bayesian Econometrics*. Hemel Hempstead: Wiley-Interscience.

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.

Nakajima, J. (2011a). Monetary policy transmission under zero interest rates: An extended time-varying parameter vector autoregression approach. IMES Discussion Paper, 2011-E-8, Bank of Japan.

Nakajima, J. (2011b). Monetary policy transmission under zero interest rates: An extended time-varying parameter vector autoregression approach. IMES Discussion Paper, 2011-E-8, Bank of Japan.

Nakajima, J., M. Kasuya, and T. Watanabe (2009). Bayesian analysis of time-varying parameter vector autoregressive model for the Japanese economy and monetary policy. IMES Discussion Paper, 2009-E-13, Bank of Japan.

Nakajima, J., S. Shiratsuka, and Y. Teranishi (2010). The effects of monetary policy commitment: Evidence from time-varying parameter VAR analysis. IMES Discussion Paper, 2010-E-6, Bank of Japan.

Nakajima, J. and M. West (2010). Bayesian analysis of latent threshold dynamic models. Working Paper 10-25, Department of Statistical Science, Duke University.

Prado, R. and G. Huerta (2002). Time-varying autoregressions with model order uncertainty. *Journal of Time Series Analysis 23*, 599–618.

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

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

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

Vrontos, I. D., P. Dellaportas, and D. N. Politis (2000). Full Bayesian inference for GARCH and EGARCH models. *Journal of Business and Economic Statistics 18*, 187–198.

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.