Research Paper Series

No. 172

Title

A Multiple DSGE-VAR Approach:
Priors from a Combination of DSGE Models and Evidence from Japan □

Name(s)†
Hirokuni Iiboshi

August, 2016

†Graduate School of Social Sciences, Tokyo Metropolitan University

A Multiple DSGE-VAR Approach:
Priors from a Combination of DSGE Models and Evidence from Japan *

Hirokuni Iiboshi†

Tokyo Metropolitan University

June 2016

Abstract

I propose a Bayesian VAR model with added priors from a combination of multiple DSGE models. The prior of the combination of multiple DSGE models improves the marginal likelihood of the DSGE-VAR with respect to a single DSGE model. This approach might be useful for model comparison between two or more DSGE models and for measuring the relative degrees of misspecification of DSGE models through comparing impulse responses of DSGE models with those of the multiple DSGE-VAR. From the data of Japanese economy including the "Bubble Boom" and the "Lost Decade", I demonstrate the multiple DSGE-VAR combined two DSGE models with and without financial frictions, and evaluate misspecification of both DSGE models from their impulse response functions.

JEL Classifications, C11, C32, C52

*The author thanks for discussion with the members of the project developing DSGE-models in Economic and Social Research Institute, Cabinet Office: Masaaki Maruyama, Tastuyoshi Matsumae, Shin-Ichi Nishiyama, Ryo Hasumi, and Daisuke Nakamura, and also thanks Sohei Kaihatsu for providing data.The author highly appreciates the comments and suggestions of an anonymous referee, which significantly contributed to improving the quality of the paper. However, remaining errors are entirely mine. The views expressed herein are of my own and do not represent those of the organizations the authors belongs to.

†E-mail: iiboshi@tmu.ac.jp,

1 Introduction

Del Negro and Schorfheide (2004, 2009) and Del Negro et al. (2007) propose to estimate Bayesian vector auto-regressive (VAR) models using priors with a dynamic stochastic general equilibrium (DSGE) model, which imposes cross-equation restrictions on multivariate time series. These are referred to as DSGE-VARs. The advantages of DSGE-VARs are their abilities to measure the misspecification of DSGE models in terms of cross-equation restrictions and to provide more plausible forecasting and policy analysis using impulse responses. And, Consolo et al. (2009) expanded them to DSGE-FAVARs.

This study considers which model weights should be used as a combination of two or more DSGE models with respect to the spirit of the DSGE-VAR approach. I refer to the new DSGE-VAR approach as multiple DSGE-VARs (MDSGE-VARs). The advantages of MDSGE-VARs are their ability to compare misspecification between two or more DSGE models in terms of (1) cross-equation restrictions and (2) contemporaneous impacts of shocks on observable variables and to conduct model selection among multiple DSGE models in terms of data-generating processes such as VAR. The idea of MDSGE-VARs is similar to the Minnesota prior approach proposed by Sims and Zha (1998), which uses multiple priors formed from several kinds of dummy variables. To show the advantage of my approach empirically, a standard New Keynesian model with financial friction invented by Kaihatsu and Kurozumi (2014a, b) along the line of Bernanke et al. (1999) is adopted as well as the model without friction.

The remainder of the paper is organized as follows. Section 2 describes the MDSGE-VAR approach. Section 3 deals with the estimation of two DSGE models as priors. Section 4 reports the estimated results, including the model weights of the two models against the data and impulse responses of the MDSGE-VAR. Section 5 concludes.

2 Multiple DSGE-VAR

2.1 Setup

To define a standard DSGE-VAR that uses a single DSGE model as a prior, I firstly describe a VAR specification for y_t, following Del Negro and Schorfheide (2004) and Del Negro et al. (2007). The p-th order VAR is represented as

$$\Delta y_t = \Phi_0 + \Phi_1 \Delta y_{t-1} + \cdots \Phi_p \Delta y_{t-p} + u_t. \tag{1}$$

n denotes the dimension of Δy_t and I let the $k \times 1$ vector x_t =[1, $\Delta y'_{t-1}$,$\cdots$, $\Delta y'_{t-p}$] and $\Phi = [\Phi_0, \Phi_1, \cdots, \Phi_p]'$. I build the DSGE model prior by generating dummy observations from the DSGE model as the following artificial sample moments. Let DSGE moments be defined as $\mathit{\Gamma}_{YY}(\theta) = \mathbb{E}^{\mathbb{D}}_{\theta}[\Delta y_t \, \Delta y'_t]$, $\mathit{\Gamma}_{XX}(\theta) = \mathbb{E}^{\mathbb{D}}_{\theta}[x_t \, x'_t]$ and $\mathit{\Gamma}_{XY}(\theta) = \mathbb{E}^{\mathbb{D}}_{\theta}[x_t \, y'_t]$ where θ is the deep parameters of the DSGE model. The derivation of the moments is described in Appendix A3. Using the DSGE moments, I obtain the mapping from the DSGE model to the VAR parameters as below:

$$\Phi^*(\theta) = \Gamma_{XX}^{-1}(\theta)\Gamma_{XY}(\theta), \tag{2}$$

$$\Sigma^*(\theta) = \Gamma_{YY}(\theta) - \Gamma_{YX}(\theta)\Gamma_{XX}^{-1}(\theta)\Gamma_{XY}(\theta), \tag{3}$$

The prior distribution from an individual DSGE model follows the inverted-Wishart multivariate normal (MNIW) form:

$$\Phi,\, \Sigma|\theta \sim MNIW\left(\Phi^*(\theta),\, [\lambda T\Gamma_{XX}(\theta)]^{-1},\, \lambda T\Sigma^*(\theta),\, \lambda T - k\right), \tag{4}$$

where $\lambda \in (0,\, \infty)$ is a hyperparameter and the ratio of the DSGE prior to the actual observation is represented as $\frac{\lambda}{1+\lambda}$. In addition, I express the standard DSGE-VAR with the hyperparameter as DSGE-VAR(λ). When $\lambda = 0$, then DSGE-VAR(0) is equivalent to a pure VAR model, and when $\lambda = \infty$, then DSGR-VAR(∞) is a pure DSGE model.

The posterior distribution with combined artificial and actual sample moments also follows the inverted-Wishart multivariate normal form:

$$\Phi,\, \Sigma|Y,\, \theta,\, \lambda \sim MNIW\left(\hat{\Phi}(\theta),\, [(\lambda+1)T\hat{\Gamma}_{XX}(\theta)]^{-1},\, (\lambda+1)T\hat{\Sigma}(\theta),\, (\lambda+1)T - k\right), \tag{5}$$

where

$$\hat{\Phi}(\theta) = (\lambda\,\Gamma_{XX}(\theta) + \hat{\Gamma}_{XX})^{-1}(\lambda\,\Gamma_{XY}(\theta) + \hat{\Gamma_{XY}}), \tag{6}$$

$$\begin{aligned}\hat{\Sigma}(\theta) &= \frac{1}{(\lambda+1)}[\,(\lambda\,\Gamma_{YY}(\theta) + \hat{\Gamma}_{YY}) - (\lambda\,\Gamma_{YX}(\theta) + \hat{\Gamma}_{YX}) \\ &\times \ (\lambda\,\Gamma_{XX}(\theta) + \hat{\Gamma}_{XX})^{-1}(\lambda\,\Gamma_{XY}(\theta) + \hat{\Gamma}_{XY})\,],\end{aligned} \tag{7}$$

and $\hat{\Gamma}_{XX}$ is the actual sample moments.

The marginal likelihood (Zellner, 1971) is given by

$$\begin{aligned}p_\lambda(Y\,|\,\theta) &= (2\pi)^{-nT/2}\frac{|\lambda\,\Gamma_{XX}(\theta) + \hat{\Gamma}_{XX}|^{-\frac{n}{2}}\,|\hat{\Sigma}(\theta)|^{-\frac{T(\lambda+1)-k}{2}}}{|\lambda\,\Gamma_{XX}(\theta)|^{-\frac{n}{2}}\,|\Sigma^*(\theta)|^{-\frac{T\lambda-k}{2}}}\,\frac{2^{\frac{n(T(\lambda+1)-k)}{2}}}{2^{\frac{n(T\lambda-k)}{2}}} \\ &\times \ \frac{\prod_{i=1}^{n}\Gamma[(\,(\lambda+1)T - k + 1 - i\,)/2]}{\prod_{i=1}^{n}\Gamma[(\,\lambda T - k + 1 - i\,)/2]},\end{aligned} \tag{8}$$

Furthermore, normalization (or integration out) with respect to θ is implemented from the modified harmonic mean by Geweke (1999):

$$p_\lambda(Y) = \int p_\lambda(Y\,|\,\theta)p(\theta)d\theta. \tag{9}$$

2.2 Multiple DSGE Priors

Let two different DSGE models denote $M1$ and $M2$, respectively, and the deep parameters used in the models denote θ_1 and θ_2. Using these notations, the DSGE artificial sample moments built from the two models are given by $\mathit{\Gamma}_{YY}(\theta_1 \,|\, M_1) = \mathbb{E}^{\mathrm{D}}_{\theta_1|\mathrm{M}_1}[\Delta y_t \triangle y_t']$, $\mathit{\Gamma}_{XX}(\theta_1 \,|\, M_1) = \mathbb{E}^{\mathrm{D}}_{\theta_1|\mathrm{M}_1}[x_t\, x_t']$, $\mathit{\Gamma}_{XY}(\theta_1 \,|\, M_1) = \mathbb{E}^{\mathrm{D}}_{\theta_1|\mathrm{M}_1}[x_t\, \Delta y_t']$, for model $M1$, and $\mathit{\Gamma}_{YY}(\theta_2 \,|\, M_2) = \mathbb{E}^{\mathrm{D}}_{\theta_2|\mathrm{M}_2}[\Delta y_t \triangle y_t']$, $\mathit{\Gamma}_{XX}(\theta_2 \,|\, M_2) = \mathbb{E}^{\mathrm{D}}_{\theta_2|\mathrm{M}_2}[x_t\, x_t']$, $\mathit{\Gamma}_{XY}(\theta_2 \,|\, M_2) = \mathbb{E}^{\mathrm{D}}_{\theta_2|\mathrm{M}_2}[x_t\, \Delta y_t']$ for model $M2$.

As Eq.(6) and Eq.(7), the posterior mean and variance are formed from linear combination of the artificial and actual sample moments. We can easily extend from the single prior of dummy variables to muliple priors by redefining as,

$$\lambda\, \mathit{\Gamma}_{XX}(\theta_i) \equiv \lambda_1\, \mathit{\Gamma}_{XX}(\theta_1) + \lambda_2\, \mathit{\Gamma}_{XX}(\theta_2) + \cdots + \lambda_m\, \mathit{\Gamma}_{XX}(\theta_m), \tag{10}$$

where for each model i, hyperparameters of priors are given as $\lambda_i \in (0, \infty)$, $i = 1, 2, \cdots, m$. $\lambda\, \mathit{\Gamma}_{YY}(\theta_i)$ and $\lambda\, \mathit{\Gamma}_{XY}(\theta_i)$ are also redefined in the same way. Since the purpose of this study is model comparison, let us introduce an additional hyperparameter to compare models conveniently, instead of using λ_i. The additional hyperparameter is $\mu \in (0, 1)$, which indicates a model probability (or a model weight) of model $M1$ from both models, then the priors of multiple DSGE moments are defined as

$$\lambda\, \mathit{\Gamma}_{YY,\mu}(\theta_1, \theta_2) \equiv \lambda\, \mu \mathit{\Gamma}_{YY}(\theta_1 \,|\, M_1) + \lambda\, (1 - \mu) \mathit{\Gamma}_{YY}(\theta_2 \,|\, M_2), \tag{11}$$

$$\lambda\, \mathit{\Gamma}_{XX,\mu}(\theta_1, \theta_2) \equiv \lambda\, \mu \mathit{\Gamma}_{XX}(\theta_1 \,|\, M_1) + \lambda\, (1 - \mu) \mathit{\Gamma}_{XX}(\theta_2 \,|\, M_2), \tag{12}$$

$$\lambda\, \mathit{\Gamma}_{XY,\mu}(\theta_1, \theta_2) \equiv \lambda\, \mu \mathit{\Gamma}_{XY}(\theta_1 \,|\, M_1) + \lambda\, (1 - \mu) \mathit{\Gamma}_{XY}(\theta_2 \,|\, M_2), \tag{13}$$

where relations between hyperparameters λ_i and μ are $\lambda_1 = \lambda\mu$ and $\lambda_2 = \lambda(1 - \mu)$, and the number of models is set to $m = 2$ in this case. I express the multiple DSGE-VAR (MDSGE-VAR) with the two hyperparameters as MDSGE-VAR(λ,μ). When $\mu = 1$, then MDSGE-VAR($\lambda, 1$) is equivalent to the standard DSGE-VAR(λ) based on the prior from model $M1$, and when $\mu = 0$, then MDSGE-VAR($\lambda, 0$) is equivalent to the standard DSGE-VAR(λ) based on the prior from model $M2$.

The prior and posterior distributions from the multiple DSGE models are the same as Eq.(4) and Eq.(5) by replacing $\mathit{\Gamma}_{YY}(\theta)$, $\mathit{\Gamma}_{XX}(\theta)$ and $\mathit{\Gamma}_{XY}(\theta)$ with $\mathit{\Gamma}_{YY,\mu}(\theta_1, \theta_2)$, $\mathit{\Gamma}_{XX,\mu}(\theta_1, \theta_2)$ and $\mathit{\Gamma}_{XY,\mu}(\theta_1, \theta_2)$, respectively. In the same manner, the marginal likelihood of MDSGE-VAR(λ,μ), $p_{\lambda,\mu}(Y \,|\, \theta_1, \theta_2)$, is also given by Eq.(8) by replacing them with their counterparts.

2.3 Choice of Hyperparameters

The marginal data density of the MDSGE-VAR(λ,μ) is given by

$$p_{\lambda,\mu}(Y) = \int p_{\lambda,\mu}(Y \,|\, \theta_1, \theta_2) p(\theta_1, \theta_2) d\theta, \tag{14}$$

where $p(\theta_1, \theta_2) = p(\theta \mid M1)\, p(\theta \mid M2)$, meaning that the priors of the structural parameters of model 1 and 2 are assumed to be independent of each other. Additionally, the normalization (or integration out) in terms of the hyperparameters is implemented using the modified harmonic mean by Geweke (1999). The choice of the hyperparameters, λ and μ, is derived from the following maximization of the marginal likelihood conditional on the hyperparameters:

$$[\hat{\lambda},\, \hat{\mu}] = \underset{\lambda \in \Lambda, \mu \in M}{\arg\max}\; p_{\lambda,\mu}(Y). \tag{15}$$

2.4 Impulse Response Functions

The impulse response functions (IRFs) of the MDSGE-VAR are calculated as well as the IRFs of the DSGE-VAR, following Del Negro and Schorfheide (2004). Here, I show how to derive the IRFs, step by step. Firstly, the contemporaneous impact of structural shocks on the endogenous variables in the DSGE model $M1$ is given by

$$\left(\frac{\partial y_t}{\partial \varepsilon_t}\right)_{DSGE(M1)} = \underbrace{A_0}_{J\times J}(\theta_1 \mid M1) = \Sigma^*_{tr}(\theta_1 \mid M1)\, \Omega^*(\theta_1 \mid M1),$$

where a square matrix A_0 is the contemporaneous impact matrix. The matrix A_0 is built from multiplication of a matrices Λ and $H(\theta)$ of the state space model described in Appendix A1. However, since $\frac{\partial y_t}{\partial \varepsilon_t} = \frac{\partial y_t}{\partial \Delta S_t}\frac{\partial \Delta S_t}{\partial \varepsilon_t} = \underbrace{\Lambda}_{J\times J}\underbrace{H(\theta)}_{J\times L}$, $(J \leq L)$, we need to change matrix $H(\theta)$ to a square matrix by reducing the number of columns of $H(\theta)$ to the number of variables: J. Since the structural shocks are independent of each other, we must identify the impact matrix even cutting some of columns of $H(\theta)$. We select common shocks between the two DSGE models as the columns of the matrices $H(\theta_1)$ and $H(\theta_2)$ in order to evaluate combination of the models. And, the impact matrix A_0 is *uniquely* decomposed into the two factors $\Sigma^*_{tr}(\theta_1 \mid M1)$ and $\Omega^*(\theta_1 \mid M1)$ by using ***QR decomposition***. Here, $\Sigma^*_{tr}(\theta_1 \mid M1)$ is a lower triangular matrix and $\Omega^*(\theta_1 \mid M1)$ is an orthogonal matrix, that is, $\Omega^*(\theta_1 \mid M1)\Omega^*(\theta_1 \mid M1)' = I$.

Secondly, the QR decomposition is also useful for identifing the shocks of structural VARs. Using the decomposition, Rubio-Ramirez et al. (2010) propose an efficient and fast identification method of structural VARs by imposing sign restrictions or zero restrictions for their impulse response function. The contemporaneous impact of structural shocks in the version of structural VARs is the same structure given as

$$\left(\frac{\partial y_t}{\partial \varepsilon_t}\right)_{SVAR} = \Sigma_{tr}\, \Omega,$$

where Σ_{tr} is a lower triangular matrix generated from Choleski decomposition of variance-covariance matrix of one-step ahead forecasting errors of a reduced VAR. Under a given Σ_{tr}, an orthogonal matrix Ω can be calculated to satify sign restrictions or zero restrictions. Although Σ_{tr} would influence to determine sizes of impulse reponses, Ω would determine shape of those.

Thirdly, in the case of the DSGE-VAR, we replace the factor Ω with $\Omega^*(\theta \mid M1)$, and the impact matrix is rewritten as

$$\left(\frac{\partial y_t}{\partial \varepsilon_t}\right)_{DSGE-VAR} = \Sigma_{tr}\, \Omega^*(\theta_1|M_1),$$

where Σ_{tr} is obtained from the DSGE-VAR.

Finally, in the case of the MDSGE-VAR, instead of the factor $\Omega^*(\theta_1 \mid M1)$, I use a combination of factors of the two DSGE models given by:

$$\begin{aligned} \left(\frac{\partial y_t}{\partial \varepsilon_t}\right)_{MDSGE-VAR} &= \Sigma_{tr}\, \mathbb{E}\Omega^*(\theta_1, \theta_2), \\ \mathbb{E}\Omega^*(\theta_1, \theta_2) &= \mu\Omega^*(\theta_1 \mid M1) + (1-\mu)\Omega^*(\theta_2 \mid M2), \end{aligned} \tag{16}$$

where $\mathbb{E}\Omega^*(\theta_1, \theta_2)$ is the expected values weighted on model weight μ. Although the expected values are not orthogonal such that $\mathbb{E}\Omega^*\mathbb{E}\Omega^*(\theta)' \neq I$, the realized values are an orthogonal matrix, $\Omega^*(\theta_1, \theta_2)\Omega^*(\theta_1, \theta_2)' = I$, where $\Omega^*(\theta_1, \theta_2) = S(M1)\Omega^*(\theta_1 \mid M1) + (1 - S(M1))\Omega^*(\theta_2 \mid M2)$, using the regime variable weighted on model $M1$, such that $S(M1)$= 1, if $model = M1$, and $S(M1)$= 0, otherwise.

As mentioned above, the QR decomposition must be widely applicable for identification of VARs, although the decomposition would not be used as perfect identification. In fact, we simulate IRFs to a shock unidentified by VARs but identified by DSGE models, using $\Omega^*(\theta_1 \mid M1)$ and $\Omega^*(\theta_2 \mid M2)$ derived from impact matrices A_0 of DSGE models. In Section 4.4, I consider usefulness of the method again, by checking impulse responses of MDSGE-VAR.

3 DSGE Models with/without Financial Friction

3.1 Benchmark Model

The standard New Keynesian model, into which Calvo-style staggered price and wage settings are incorporated, is selected as the benchmark model. In this model, the economy consists of three representative agents: households, firms and the central bank. There are fourteen log-linearized structural equations in the benchmark model, the following Eq.(17) through Eq.(30).[1]

Households

The marginal utility of the households, λ_t, is represented as:

$$\begin{aligned} \lambda_t &= -\frac{1}{1-\theta\pi/r_n}\left\{\frac{\sigma}{1-\theta/z^*}\left(c_t - \frac{\theta}{z^*}(c_{t-} - z_t^*)\right) - z_t^b\right\} \\ &+ \frac{\theta\pi/r_n}{1-\theta\pi/r_n}\left\{\frac{\sigma}{1-\theta/z^*}\left(E_t c_{t+1} + E_t z_{t+1}^* - \theta/z^* c_t\right) - E_t z_{t+1}^b\right\}, \end{aligned} \tag{17}$$

[1]A detailed explanation of the DSGE model and the derivation of log-linearized structural equations is proveded by Kaihatsu and Kurozumi(2014a,b).

$$\lambda_t = E_t\lambda_{t+1} - \sigma E_t\, z^*_{t+1} + r^n_t - E_t\pi_{t+1}, \tag{18}$$

where c_t is consumption, and r_n and π are the steady states of the nominal interest rate and inflation, respectively. In addition, r^n_t and π_t are the nominal interest rate and inflation. z^b_t and z^*_t are a preference shock and a composite technological shock, respectively. The real wage version of the New Keynesian Phillips curve is as follows:

$$\begin{aligned} w_t &= w_{t-1} - \pi_t + \gamma_w\pi_{t-1} - z^*_t + \frac{z^*\pi}{r_n}(E_t w_{t+1} - w_t + E_t\pi_{t+1} - \gamma_w\pi_t + E_t z^*_{t+1}) \\ &+ \frac{(1-\xi_w)(1-\xi_w z^*\pi/r_n)}{\xi_w(1+\chi(1+\lambda_w)/\lambda_w)}/(\chi l_t - \lambda_t - w_t + z^b_t) + z^w_t, \end{aligned} \tag{19}$$

where w_t is the real wage, l_t: the labor supply, and z^w_t : the real wage markup shock. Futhermore, z^*_t is a composite technology shock, such as $z^*_t = z^z_t + \alpha/(1-\alpha)z^\psi_t$, where z^z_t is a neutral technology shock.

The utilization rate of capital, u_t, is represented as:

$$u_t = \mu(r^k_t - q_t), \tag{20}$$

where r^k_t is the rental rate, and q_t is the price of capital. The current and expected marginal returns on capital, χ_t, are given as:

$$\chi_t = \left(1 - \frac{1-\delta}{r_n\psi}\right) r^k_t + \frac{1-\delta}{r_n\psi} q_t - q_{t-1} - z^\psi_t, \tag{21}$$

$$E_t\chi_{t+1} = r^n_t - E_t\pi_{t+1}, \tag{22}$$

where z^ψ_t is an IS technology shock.

Firms

The aggregate production function and marginal rate of substitution are given by

$$y_t = (1+\phi)\left\{(1-\alpha)\, l_t + \alpha(u_t + k_{t-1} - z^*_t - z^\psi_t)\right\}, \tag{23}$$

$$0 = w_t + l_t - (r^k_t + u_t + k_{t-1} - z^*_t - z^\psi_t). \tag{24}$$

where y_t, and k_t are the output and capital, respectively. The real marginal cost, mc_t, is given by

$$mc_t = (1-\alpha)w_t + \alpha r^k_t. \tag{25}$$

The consumption goods version of the New Keynesian Phillips curve is represented by

$$\pi_t = \gamma_p\pi_{t-1} + \frac{z^*\pi}{r_n}(E_t\pi_{t+1} - \gamma_p\pi_t) + \frac{(1-\xi_p)(1-\xi_p\frac{z^*\pi}{r_n})}{\xi^p} mc_t + z^p_t, \tag{26}$$

where z^p_t is a price markup shock.

The capital accumulation equation is given by

$$k_t = \frac{(1-\delta - r_n\psi/\pi)}{z^*\psi}u_t + \frac{1-\delta}{z^*\psi}(k_{t-1} - z_t^* - z_t^{\psi}) + \left(1 - \frac{1-\delta}{z^*\psi}\right) i_t, \tag{27}$$

Furthermore, the price of capital is represented by

$$q_t = \frac{1}{\zeta}(i_t - i_{t-1} + z_t^* + z_t^{\psi}) - \frac{z^*\pi}{\zeta r_n}(E_t i_{t+1} - i_t + E_t z_{t+1}^* + E_t z_{t+1}^{\psi}) + z_t^{\nu} + z_t^{i}, \tag{28}$$

where i_t is investment. z_t^{ν} and z_t^{i} are a shock to MEI and an investment-good price markup shock, respectively.

The central bank and the market clearing condition

The monetary policy rule conducted by the central bank is represented as:

$$r_t^n = \phi_r r_{t-1}^n + (1-\phi_r)\left\{\frac{\phi_\pi}{4}\Sigma_{j=0}^{3}\pi_{t-j} + \phi_y y_t\right\} + \phi_{\Delta y}(y_t - y_{t-1} + z_t^*) + z_t^r, \tag{29}$$

where z_t^r is a monetary policy shock. The market clearing condition is given by

$$y_t = \frac{c}{y}\,c_t + \frac{i}{y}\,i_t + \frac{g}{y}\,z_t^g, \tag{30}$$

where $\frac{c}{y}$, $\frac{i}{y}$ and $\frac{g}{y}$ are ratios of consumption, investment and government expenditure to output in terms of steady states, respectively. z_t^g is a government expenditure shock.

3.2 DSGE Model with Financial Friction

Financial intermediaries

In the economy with financial friction, financial intermediaries are added as additional agents. The loan rate of financial intermediaries, r_t^E , and the net worth of entrepreneurs, n_t are represented as

$$r_t^E = r_t^n + \mu_E(q_t + k_t - n_t) + z_t^{\mu}, \tag{31}$$

$$\begin{aligned} \frac{z^*}{\eta\, r^E}\, n_t \;&=\; \frac{1+\lambda_i}{n/k}\left[(1 - \frac{1-\delta}{r^E\psi})r_t^k + \frac{1-\delta}{r^E\psi}q_t - q_{t-1} - z_t^{\psi}\right] \\ &-\; \left(\frac{1+\lambda_i}{n/k} - 1\right) E_{t-1}r_t^E + n_{t-1} - z_t^* + z_t^{\eta}, \end{aligned} \tag{32}$$

where μ_E is the degree of external financial premium (EPF) decided by the leverage ratio, $q_t + k_t - n_t$, and z_t^{μ} is an EPF shock. The real borrowing of entrepreneurs, b_t is represented as

$$b_t = \frac{1+\lambda_i}{1+\lambda_i - n/k}(q_t + k_t) + \left(1 - \frac{1+\lambda_i}{1+\lambda_i - n/k}\right) n_t, \tag{33}$$

In addition, the expected marginal return on capital is replaced as

$$E_t \chi_{t+1} = r_t^E - E_t \pi_{t+1}, \tag{34}$$

with Eq. (22).

To sum up, there are seventeen log-linearized structural equations in the DSGE model with financial friction, consisting of both Eq.(17) through Eq.(21) and Eq.(23) through Eq.(34) based on Kaihatsu and Kurozumi (2014a, b).

4 Empirical Results

4.1 Data

The data to estimate the models are based on Kaihatsu and Kurozumi (2014 b). The data on output, Y_t, and consumption, C_t, are obtained by dividing nominal GDP and nominal consumption with the CPI. The sample period is from 1981:Q1 to 1998:Q4, after the hyperinflation caused by the oil shock in 1979 and before the zero interest rate policy conducted by the Bank of Japan starting from 1999Q1. For the VAR model, I use six variables: (1) output, (2) consumption, (3) investment, (4) real wage, (5) inflation and (6) nominal interest rate.

4.2 Model Misspecification

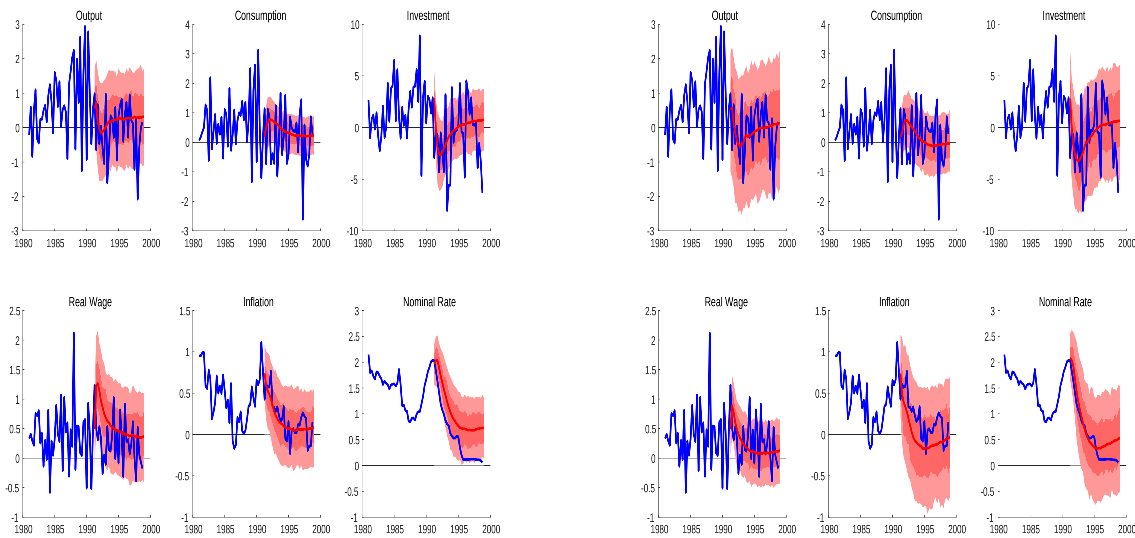
Before evaluating model combination of the two DSGE models by the MDSGE-VAR approach, I measure their model misspecifications in terms of predictive densities.[2] One model with small degree of misspecification might have better performance of forecast than another with the large degree. Figure 1 shows predictive densities of the models at the point of 1990:Q4. Panel (a) and (b) are those of model without and with the financial friction, respectively. The blue lines are actual data, and the red lines and the light red shade areas denote posterior means and 68% interval of predictive densities. From comparison of the red lines and shade areas in these figures, FF model makes the forecasts of key macroeconomic variables such as output, consumption, investment, inflation and interest rates more fluctuate and more volatile than those of NK model, because financial accelerator mechanism works.

Using predictive densities of the models generated from multi-step forecasting, I calculate *log scores* of the models as measure of their model misspecifications as below.

[2]There are large literature on model misspecification. The index of measuring model misspecification in this study is provided from the concept of optimal prediction pool proposed by Geweke and Amisano (2010). In their method, the pool combines predictive densities of alternative models where each model may be misspecified. The predictive densities are measured from log score. The log score calculated from multi-step forecasting are also used by Del Negro et al. (2014), who propose dynamic prediction pool method by extending constant model weight by Geweke and Amisano (2010) to time-varying model weight. My calculation of log score is based on Waggoner and Zha (2012) as well as Del Negro et al. (2014).

Figure 1: Forecasts of Two Models

(a) NK model: Forecast 1990:Q4 (b) FF model: Forecast 1990:Q4

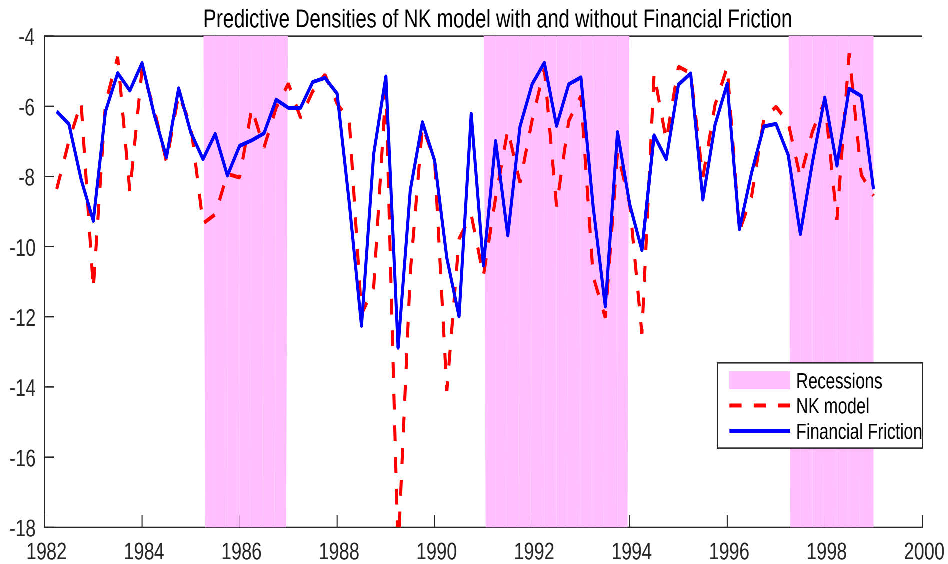


Notes: The blue lines are actual data, and the red lines and the light red shade areas denote posterior means and 68% interval of predictive densities. The posterior prediction distributions of the two models are calculated based on the posterior estimates of parameters of the models, using 10,000 draws of posterior estimates over the full sample.

$$\log \text{score}\, y_t = \sum_{h=1}^{4} \log p(y_t^O - y_t^F | Y_{t-h}^O, \theta),$$

where function $p(\cdot)$ is likelihood function, and h is forecast horizon. I forecast from 1Q ahead to 4Q ahead. The superscript "O" and "F" denote observed and forecasted variables, respectively. Figure 2 shows time series of the log scores of the two models. The blue solid and red dashed lines are FF and NK models, respectively. Basically, from 1988:Q1 to 1994:Q4, so-called the "bubble boom and burst" periods, the log scores of FF model are higher than those of NK model. It indicates that the former's model misspecication is smaller than the latter's. In the period, it is thought that many firms of corporate sector have faced credit constraints and the financial accelerator explains the business cycle much better. In contrast, after 1995:Q4, the NK model's log scores are better than its counterpart, since financial market might come to work more smoothly than the previous period. In this way, one model does not dominate the other model for the whole sample period. We can conclude that each model may be misspecified and model combination of the two model would be significant.

Figure 2: Log Scores of Two Models



Notes: The log score at each period is calculated from log likelihood function of forecasting errors of the individual model such as the NK model and the FF model as explained in Sec 4.2.

4.3 Posterior Distributions of Hyperparameters

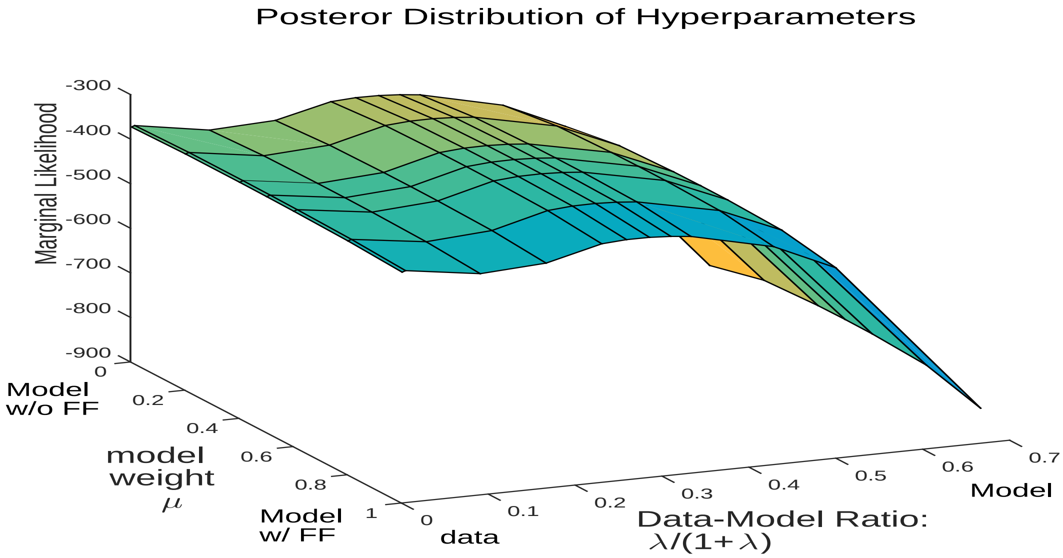
Using the posterior, Eq.(5), with the multiple DSGE moments, Eq.(11), Eq.(12) and Eq.(13) derived from estimations of the deep parameters in the two DSGE models with and without financial friction, the MDSGE-VAR (λ, μ) is estimated.[3] The benchmark (NK) model is set as model $M1$ and the financial friction (FF) model as model $M2$, so that the weights of models $M1$ and $M2$ against the data are $\frac{\lambda\mu}{1+\lambda}$ and $\frac{\lambda(1-\mu)}{1+\lambda}$, respectively. The marginal likelihood, Eq.(8), is calculated using the posterior estimates of the MDSGE-VAR (λ, μ) given the values of the hyperparameters. As long as$\lambda \geq (k+n)/T$ (in this study, $(k+n)/T \doteqdot 0.264$ since $n = 6$, $k = 1 + pn$ and $T = 72$), the prior is proper. If the prior is improper, the marginal likelihood, Eq. (8) cannot be calculated. Accordingly, the values of λ and μ are selected as 0.3, 0.35, 0.4, 0.45, 0.5, 0.75, 1, 2 and 0, 0.2, 0.4, 0.5, 0.6, 0.8, 1, respectively. I set the lag order of the MDSGE-VAR as 2, that is, $p = 2$, for the estimation for Japan. When $p = 4$, the prior is improper because of the small λ. For each λ and μ, Markov chain Monte Calro (MCMC) simulation is conducted with $30,000$ iterations. The first 5000 iterations are discarded and the next $25,000$ iterations are sampled as posterior estimates. Figure 3 show the posterior distributions of λ and μ. In Panel (a) of Figure 3, the marginal likelihoods of MDSGE-VARs are drawn in terms of two hyperparameters from three

[3] Posterior estimates of parameters are obtained from Bayesian estimation via MCMC, using the state space model in Appendix A1, and measurement equations are in A2. MCMC simulation is conducted with $300,000$ iterations. The first $100,000$ iterations are discarded and the next $200,000$ iterations are sampled as posterior estimates. Out of $200,000$ samples, I use $30,000$ samples for estimateing the MSDSGE-VAR model.

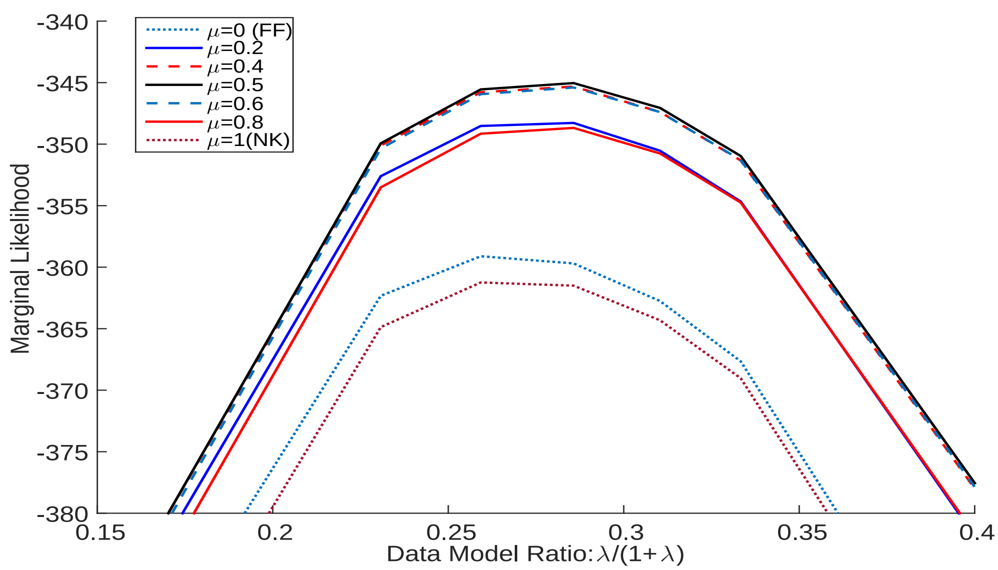
dimensions, whereas Panel (b) is reduced to two dimensions version. The vertical axis of both figures denotes the marginal likelihood and it is convenient that the horizontal axis of λ is modified to that of data-model ratio such as $\lambda/(1+\lambda) \in [0,1]$, since the range of data-model ratio is between zero and one. As λ becomes larger after $\lambda = 0.4$, the marginal likelihood drops monotonically. It indicates evidence of misspecification of both DSGE models. However, these DSGE models are still useful for explaining the data, since the marginal likelihood is maximized when $\lambda = 0.4$ (or data-model ratio: $\lambda/(1+\lambda) = 0.28$). Furthermore, when $\mu = 0.5$, the marginal likelihood is maximized. It indicates that the combination of the two models is useful and the model probability of the FF model is almost the same as that of the NK model. This result is consistent with the log scores of the two models in Section 4.2, where one DSGE model does not dominate the other over all of the sample period.

Figure 3: Posterior Distributions of Hyperparameters (Lag=2)

(a) Three-Dimension Version



(b) Two-Dimension Version



Notes: μ denotes the hyperparameter of model weight between two DSGE models, and λ the hyperparameter of the DSGE models prior to the actual observation. FF denotes the DSGE model with financial friction, NK the DSGE model without the friction. When $\mu = 0.5$ (NK model:50%, FF model:50%) and $\lambda = 0.4$ (or $\frac{\lambda}{1+\lambda} = 0.28$: Model:28%, Data:72%), the marginal likelihood is the largest one.

4.4 Comparison of IRFs

Using the structural shocks identified by the QR factorization described in Section 2.4, the contemporaneous impact in the MDSGE-VAR($\lambda = 0.4$,$\mu = 0.5$), maximizing the marginal likelihood as shown in Figure 3, is calculated from

$$\left(\frac{\partial y_t}{\partial \varepsilon_t}\right)_{MDSGE-VAR} = \underbrace{\Sigma_{tr}\,\mathbb{E}\Omega^*(\theta)}_{6\times 6}. \tag{35}$$

where $\mathbb{E}\Omega^*(\theta) = 0.5\Omega^*(\theta_1\,|\,NK\,model) + 0.5\Omega^*(\theta_2\,|\,FF\,model)$, using Eq.(16).

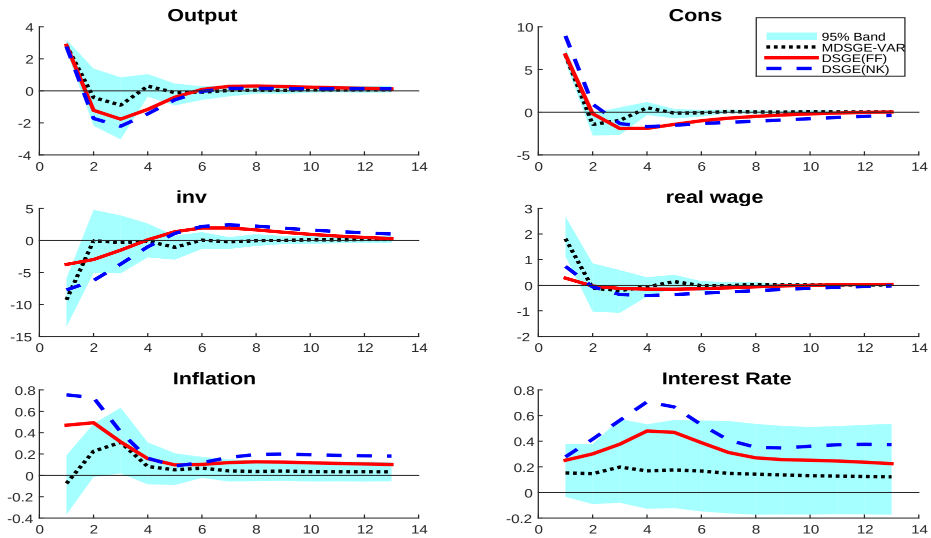
I select six shocks: (1) preference, (2) monetary policy, (3) neutral technology, (4) IS technology, (5) price markup, (6) government expenditure shocks, as common shocks of both DSGE models. Panel (a) and (b) of Figure 4 show that the IRF to preference and monetary policy shocks in the MDSGE-VAR ($\lambda = 0.4$, $\mu = 0.5$), respectively. The black dotted line and light blue shaded area denote the means and 95% credible interval, respectively. These figures also depict those of the two individual DSGE models; red solid lines and blue dashed lines are the financial friction model and the benchmark model.

As Panel (a) of Figure 4, the contemporaneous impact of output, consumption and investment to the preference shock in MDSGE-VAR are almost same as both DSGE models. In contrast, the impact inflation of the VAR is negative or almost zero, and that of real wage is obviously positive unlike those of the DSGE models. In the monetary policy shock as Panel (b), the contemporaneous impacts of endogenous variables except interest rate are not matched between the MSDSGE-VAR and the DSGE models. It seems to be ambiguous about effect of monetary policy from MDSGE-VAR unlike original DSGE approach.

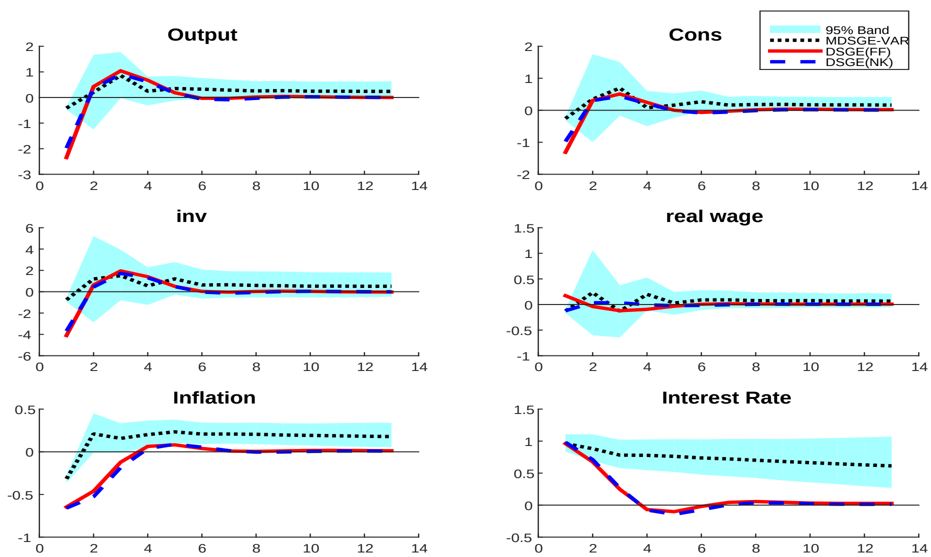
However, overall the IRFs of endogenous variables except the inflation and interest rate in the two theoretical models are within the blue shaded interval of the empirical VAR model. During the estimation period, the interest rate had already been persistent even before the zero interest rate policy, since the Japanese economy had suffered from deflation. Finally, it is noteworthy that coefficients of VAR determine path of IRFs after the second horizon, and the contemporaneous impacts matrix determines only values of endogenous variables at the first horizon. Accordingly, the gaps of IRFs between the VAR and the DSGE models after the second horizon show misspecification of the DSGE models from the VAR model regarded as the reference model. The gaps indicate that the preference shock simulates good performance in the both DSGE models, while simulations of the monetary policy shock might not be good.

Figure 4: Impulse Response Functions

(a) Preference Shock



(b) Monetary Policy Shock



Notes: The impulse response function of the MDSGE-VAR is calculated from Eq.(35), in which the hyperparameters, μ and λ, are set to 0.5 and 0.4, respectively. FF denotes the DSGE model with financial friction, NK the DSGE model without the friction.

5 Conclusion

This study considers the model weights that should be used as the model combination in terms of the spirit of the DSGE-VAR approach. The prior of the combination of multiple DSGE models makes the marginal likelihood of the DSGE-VAR better than that of a single DSGE model. From the data of Japanese economy including the "Bubble Boom" and the "Lost Decade", I demonstrate the multiple DSGE-VAR combined two DSGE models with and without financial frictions, and evaluate misspecification of both DSGE models from their impulse response functions. The estimation results show that the multiple DSGE-VAR might be useful for model comparison between two or more DSGE models and for measuring the relative degrees of misspecification as additional advantages.

A Appendix

A.1 State Space Representation

From observable variables, y_t, and state variables, s_t , a state space model for the DSGE model with the first difference series $\triangle y_t$ is represented as:

$$\underbrace{\triangle \mathbf{y}_t}_{J\times 1} = \underbrace{\mathbf{D}}_{\mathbf{J\times 1}} + \underbrace{\Lambda}_{\mathbf{J\times K}} \underbrace{\triangle \mathbf{S}_t}_{K\times 1} + \underbrace{\mathbf{e}_t}_{J\times 1}, \tag{36}$$

$$\underbrace{\mathbf{S}_t}_{K\times 1} = \underbrace{\mathbf{G}(\theta)}_{K\times K} \underbrace{\mathbf{S}_{t-1}}_{K\times 1} + \underbrace{\mathbf{H}(\theta)}_{K\times L} \underbrace{\varepsilon_t,}_{L\times 1} \quad \varepsilon_t \sim N(\mathbf{0}, \mathbf{Q}(\theta)), \tag{37}$$

where Eq. (36) is a measurement eq. and Eq. (37) is a state eq. D is a vector of constant terms. e_t and ε_t are measurement errors and structural shocks, respectively.

Using matrix expression, we rewrite Eq. (36) and Eq. (37) as below.

$$\tilde{\triangle \mathbf{y}}_t = \mathbf{D} + \underbrace{\begin{bmatrix} \Lambda & -\Lambda \end{bmatrix}}_{\tilde{\Lambda}} \underbrace{\begin{bmatrix} \mathbf{S_t} \\ \mathbf{S}_{t-1} \end{bmatrix}}_{\tilde{S}_t} + \nu_t, \qquad \nu_t \sim N(0, \mathbf{R}), \tag{38}$$

$$\underbrace{\begin{bmatrix} \mathbf{S_t} \\ \mathbf{S}_{t-1} \end{bmatrix}}_{\tilde{S}_t} = \underbrace{\begin{bmatrix} G(\theta) & 0 \\ I & 0 \end{bmatrix}}_{\tilde{G}} \underbrace{\begin{bmatrix} \mathbf{S}_{t-1} \\ \mathbf{S}_{t-2} \end{bmatrix}}_{\tilde{S}_{t-1}} + \underbrace{\begin{bmatrix} \mathbf{H}(\theta) \\ 0 \end{bmatrix}}_{\tilde{H}} \varepsilon_t, \qquad \varepsilon_t \sim N(0, \mathbf{Q}(\theta)), \tag{39}$$

A.2 Measurement Equations

State space models of the DSGE models consist of a measurement equation and a state equation, which are described in Appendix A1. In the following subsection, I explain about the measurement equations.

(1) Benchmark Model The measurement equation of the benchmark model is represended as below. There are eight observed variables such as ouyput;Y_t, consumption; C_t, investment; I_t, real wage; W_t, labor input; L_t, inflation; π_t, investment price; P_t^i, nominal rate; r_t^n.

$$
\begin{bmatrix} 100\Delta\log Y_t \\ 100\Delta\log C_t \\ 100\Delta\log I_t \\ 100\Delta\log W_t \\ 100\log l_t \\ 100\Delta\log P_t \\ 100\Delta\log(P_t^i/P_t) \\ 100{r^n}_t \end{bmatrix} = \begin{bmatrix} \bar{z^*} \\ \bar{z^*} \\ \bar{z^*}+\bar{\psi} \\ \bar{z^*} \\ \bar{l} \\ \bar{\pi} \\ -\bar{\psi} \\ \bar{r^n} \end{bmatrix} + \begin{bmatrix} z_t^* + y_t - y_{t-1} \\ z_t^* + c_t - c_{t-1} \\ z_t^* + z_t^{\psi} + i_t - i_{t-1} \\ z_t^* + w_t - w_{t-1} \\ l_t \\ \pi_t \\ -z^{\psi} + z_t^{\nu} - z_t^{\nu} \\ r_t^n \end{bmatrix} + \begin{bmatrix} e_t^Y \\ e_t^C \\ e_t^I \\ e_t^W \\ e_t^I \\ e_t^P \\ e_t^{Pi} \\ e_t^r \end{bmatrix}, \tag{40}
$$

where $\bar{z^*} = 100(z^*-1)$, $\bar{\psi} = 100(\psi-1)$, $\bar{r^n} = 100(r^n-1)$, and $\bar{l}$ is normaized to be equal to zero following Kaihatsu and Kurozumi (2014a,b). Small letters indicate log-deviations from steady state values after detrending by level of composite technology Z_t^*.

(2) DSGE model with Financial Friction The measurement equation of the financial friction model is represended as below. There are additional two observed variables such as loan rate;r_t^E, and real borrowing; B_t.

$$
\begin{bmatrix} 100\Delta\log Y_t \\ 100\Delta\log C_t \\ 100\Delta\log I_t \\ 100\Delta\log W_t \\ 100\log l_t \\ 100\Delta\log P_t \\ 100\Delta\log(P_t^i/P_t) \\ 100 r_t^n \\ 100\, r_t^E \\ 100\Delta\log B_t \end{bmatrix} = \begin{bmatrix} \bar{z^*} \\ \bar{z^*} \\ \bar{z^*}+\bar{\psi} \\ \bar{z^*} \\ \bar{l} \\ \bar{\pi} \\ -\bar{\psi} \\ \bar{r^n} \\ \bar{r^E} \\ \bar{z^*} \end{bmatrix} + \begin{bmatrix} z_t^* + y_t - y_{t-1} \\ z_t^* + c_t - c_{t-1} \\ z_t^* + z_t^{\psi} + i_t - i_{t-1} \\ z_t^* + w_t - w_{t-1} \\ l_t \\ \pi_t \\ -z^{\psi} + z_t^{\nu} - z_t^{\nu} \\ r_t^n \\ r_t^E \\ z_t^* + b_t - b_{t-1} \end{bmatrix} + \begin{bmatrix} e_t^Y \\ e_t^C \\ e_t^I \\ e_t^W \\ e_t^I \\ e_t^P \\ e_t^{Pi} \\ e_t^r \\ e_t^{rE} \\ e_t^B \end{bmatrix}, \tag{41}
$$

where$\bar{r^E} = 100(r^E - 1)$.

A.3 Derivation of DSGE moments

Using the state space model, Eq.(36) and Eq.(37), the DSGE moments are given by

$$
\mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_t'] = \Lambda \Omega_{ss} \Lambda' + \Sigma_{ee} + DD',
$$

$$
\mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-h}'] = \Lambda G(\theta)^h \Omega_{ss} \Lambda' + DD'.
$$

where $\Omega_{ss} = \mathbb{E}_\theta^{\mathbb{D}}[S_t\, S_t']$, and can be obtained by solving the following Lyapunov equation,

$$
\Omega_{ss} = \Lambda \Omega_{ss} \Lambda' + H(\theta) Q(\theta) H(\theta)'.
$$

And $\mathbb{E}_\theta^{\mathbb{D}}[x_t\, x_t']$ is defined as

$$
\mathbb{E}_\theta^{\mathbb{D}}[x_t\, x_t'] = \begin{bmatrix} \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_t'] & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-1}'] & \cdots & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-p+1}'] & D \\ \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-1}'] & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_t'] & \cdots & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-p+2}'] & D \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-p+1}'] & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_{t-p+2}'] & \cdots & \mathbb{E}_\theta^{\mathbb{D}}[\Delta y_t\, \Delta y_t'] & D \\ D & D & \cdots & D & 1 \end{bmatrix},
$$

where D is the vector in Eq.(36).

References

[1] Barnanke, B.S., Gertler, M., Gilchrist, S., (1999) “The Financial Accelerator in Quantitative Business Cycle Framework,” In: Taylor, J.B., Woodford, M. (Eds), Handbook of Macroeconomics, Vol 1C, 1341-1393.

[2] Consolo, A., C. A. Favero, A. Paccagnini (2009) “On the Statistical Identification of DSGE Models,” Journal of Econometrics, 150, 99-115.

[3] Del Negro, M., Hasegawa, R.B. and Schorfheide, F. (2014) Dynamic Prediction Pools: an Investigation of Financial Frictions and Forecasting Performance. NBER Working Paper Series 20575.

[4] Del Negro, M. and F. Schorfheide (2004) “Priors From General Equilibrium Models For VARs,” International Economic Review, 45 (2), 643-673.

[5] Del Negro, M. and F. Schorfheide (2009) “Monetary Policy Analysis with Potentially Misspecified Models,” American Economic Review, 99(4), 1415-1450.

[6] Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007) “On the Fit of New Keynesian Models,” Journal of Business & Economic Statistics, 25 (2), 123-143.

[7] Geweke, J. (1999) "Using simulation methods for bayesian econometric models: inference, development,and communication," Econometric Reviews, vol. 18(1), p1-73.

[8] Geweke, J. and Amisano, G. (2011) Optimal Prediction Pools. Journal of Econometrics 164, 130-141.

[9] Kaihatsu, S., Kurozumi, T., (2014a). Sources of Business Fluctuations: Finanical or Technology Shocks? *Review of Economic Dynamics*, **17**, 224-242.

[10] Kaihatsu, S., Kurozumi, T., (2014b). What caused Japan’s Great Stagnation in the 1990s? Evidence from an estimated DSGE model. *Journal of Japanese and International Economies*, **34**, 217-235

[11] Rubio-Ramirez, Waggoner, D.F. and Zha, T. (2010) “Structural Vector Autoregressions: Thory of Identification and Algorithms for Inference,” Review of Economic Studies, 77, 665-696.

[12] Sims, C.A. and T. Zha (1998) “Bayesian Methods for Dynamic Multivariate Models,” International Economic Review, 39(4), 949-968.

[13] Waggoner, D.F. and Zha, T. (2012) Confronting Model Misspecification in Macroeconomics. Journal of Econometrics 171, 167-184.

[14] Zellner, A. (1971) Introduction to Bayesian Inference in Econometrics, Wiley & Sons, New York.