Does a financial accelerator improve forecasts during financial crises?: Evidence from Japan with Prediction Pool Methods

Hasumi, Ryo and Iiboshi, Hirokuni and Matsumae, Tatsuyoshi and Nakamura, Daisuke

Japan Center for Economic Research, Graduate school of social sciences, Tokyo Metropolitan University, Graduate school of economics and managements, Tohoku University, Consumer Affairs Agency, Government of Japan

March 2018

Does a financial accelerator improve forecasts during financial crises?

-Evidence from Japan with Prediction Pool Methods-

Ryo Hasumi* Hirokuni Iiboshi†

Tatsuyoshi Matsumae‡ Daisuke Nakamura§

March 2018

Abstract

Using a Markov-switching prediction pool method (Waggoner and Zha, 2012) in terms of density forecasts, we assess the time-varying forecasting performance of a DSGE model incorporating a financial accelerator à la Bernanke et al. (1999) with the frictionless model by focusing on periods of financial crisis including the so-called "Bubble period" and the "Lost decade" in Japan. According to our empirical results, the accelerator improves the forecasting of investment over the whole sample period, while forecasts of consumption and inflation depend on the fluctuation of an extra financial premium between the policy interest rate and corporate loan rates. In particular, several drastic monetary policy changes might disrupt the forecasting performance of the model with the accelerator. A robust check with a dynamic pool method (Del Negro et al., 2016) also supports these results.

Keywords: Density forecast, Optimal prediction pool, Markov-switching prediction pool, Dynamic prediction pool, Bayesian estimation, Markov Chain Monte Carlo, Financial Friction.

JEL Classifications: C32, C53, E32, E37

*Japan Center for Economic Research

†Corresponding author, Graduate school of social sciences, Tokyo Metropolitan University , E-mail: iiboshi@tmu.ac.jp

‡Graduate school of economics, Tohoku University

§Consumer Affairs Agency, Government of Japan

1 Introduction

Since the millennium, central banks and government policymakers have increasingly paid attention to forecasting macroeconomic variables using the dynamic stochastic general equilibrium (DSGE) model, as well as to conducting policy analyses with them. In fact, there is an expanding volume of literature on evaluating the accuracy of point forecasts with DSGE models. Smets and Wouters (2007) is the pioneering work to do this for US data. Adolfson et al. (2007), Edge et al. (2010), and Kolasa et al. (2012) also pushed forward with research along the line of Smets and Wouters (2007). Edge and Gurkaynak (2010) reported that forecasts with a medium-scale DSGE model perform with similar accuracy to the competing statistical models and professional forecasts. However, the forecasts show poor performance in the absolute sense. One of the features in estimation with DSGE models is imposing contemporaneous correlations between macroeconomic variables as model-based restrictions, in contrast with alternative statistical models such as vector autoregressions (VARs), as pointed out by Del Negro and Schorfheide (2003) and Herbst and Schorfheide (2012). If the restrictions of comovements among the observed series are shown to be consistent with data, forecasts of a DSGE model with the correct restrictions are likely to prevail over those of a VAR without the restrictions.

On the other hand, there is extensive literature on empirical studies about financial crises with respect to a DSGE model. In particular, the financial accelerator mechanism of Bernanke et al. (1999), in which business cycles are amplified by the presence of asymmetric information under banks and the corporate sector, are often incorporated into DSGE models, such as Christensen and Dib (2008), De Graeve (2008), Christiano et al. (2014) and Kaihatsu and Kurozumi (2014a). Gilchrist and Zakrajsek (2012) empirically supported the mechanism by showing corporate bond credit spreads leads to significant declines in consumption, investment as well as to appreciate disinflation in the US. In Japan, a collapse of the "Bubble boom" at the beginning of 1991 and successively accruing a long stagnation called the "Lost decade" was generally believed to be due to a financial crisis. However, there is still an academic controversy over its causes. Hayashi and Prescott (2002) argue that a deep decline of total factor productivity (TFP) was the main source of the long stagnation. Kaihatsu and Kurozumi (2014b) measure the extent to which TFP and financial effects contributed to the stagnation from a historical decomposition for that period by incorporating the financial accelerator into a DSGE model. Instead of the method of historical decompositions, this paper tries to specify different comovements behind different models by comparing the density forecasts of two competing DSGE models based on Kaihatsu and Kurozumi (2014b) and to figure out the causality of the stagnation from our results.

Most studies on DSGE model forecasting have adopted point forecasts evaluated from root-mean-square error (RMSE), while density forecasts of the DSGE model, a newer concept, have recently been focused on by several papers such as Herbst and Schorfheide (2012) and Kolasa and Rubaszek (2015)[1]. The former examined density forecasts of comovements of output, inflation and interest rates of a medium-size DSGE model, while the latter reported that the DSGE model incorporating the housing market outperforms both the frictionless and financial friction models for

[1]Another approach is conducted by Amisano and Giacomini (2007), who propose a model selection using a likelihood ratio test in terms of density forecast.

US data, especially, during a period of financial turmoil. On the other hand, Geweke and Amisano (2011) propose a method to obtain the optimal combination of density forecasts generated by multiple statistical models with constant model weights. And their idea, referred to as the optimal prediction pool, is applied to a combination of macroeconomic models including DSGE models, and extended to versions with time-varying model weights by Waggoner and Zha (2012) and Del Negro et al. (2016): The former shifts the weights with a Markov switching (MS) model, while the latter changes them from the probit transformation of a latent variable following an autoregression process.

The purpose of this paper is to examine whether the financial accelerator mechanism improves density forecasts of macroeconomic variables, focusing on the collapse of the "Bubble boom" in 1991 and the "Lost decade" in the 1990's in Japan, using the optimal prediction pools. And this examination indicates that the higher forecast performance of the model with, as opposed to without, the financial friction reflects the presence of comovements predicted by the financial friction in the data during the period, and also suggests that the financial accelerator theory can explain the causality of macroeconomic dynamics rather than the frictionless model. In addition, following Waggoner and Zha (2012), we estimate when and the extent to which the comovements generated by the two DSGE models change through changes of the time-varying model weight, realizing the optimal combination of density forecasts. Furthermore, we conduct a robust check to examine whether a similar dynamic change of the weight is observed using the alternative method by Del Negro et al. (2016).

This paper shows the following findings. For the overall periods from 1981:Q1 to 1998:Q4, the model with the financial friction is predominant over the frictionless benchmark model in terms of density forecasts. The difference between them is likely to come from fluctuation of spread between corporate loan rates and the policy interest rate. In periods with a small change of the spread, the financial accelerator mechanism contributes to improve the prediction. When a drastic monetary policy was implemented, however, the loan rates that did not react to a big change of the policy rates and shifted the spread with a large step reduced the forecasting performance of the model with the friction. In particular, the frictionless model outperforms for the period from 1993 to 1995, since the spread realized with a big range despite the boom seems to be contrary to the spread predicted from the financial friction. These empirical results suggest that real spreads do not give a timely reflection of the change of the extra financial premium generated between bankers and the corporate sector and that there is a non-trivial time lag between them. The robust check also supports these results.

The rest of this paper is organized as follows. Section 2 describes two competing DSGE models with and without the financial friction. The impulse responses of both models show the difference of the comovements generated by the models. In Section 3, we deal with theoretical aspects of both the prediction score and the MS pooling method. We mention the empirical results in Section 4 and the robust check using another pool method in Section 5. Section 6 concludes. Finally, two log-linearized DSGE models are explained in the Appendix.

2 DSGE models

Our model follows Kaihatsu and Kurozumi (2014 a, b), who incorporate the financial accelerator (hereafter, FA) mechanism of Bernanke et al. (1999) into a medium-size New Keynesian model with prices and wage rigidities including consumption and investment goods as well as habit persistence of consumption and an increasing adjustment cost of investment, along the line of Christensen and Dib (2008), De Graeve (2008) and Christiano et al. (2014). And, a frictionless DSGE model is adopted in order to be compared in density forecasts as the benchmark model, and to be combined with the FA model in a prediction pool method explained later.

2.1 Frictionless DSGE model

First of all, a New Keynesian model excluding the FA mechanism is described as the benchmark model (hereafter, NK model) of this paper. The remaining parts of the model are completely the same framework as the model embedding the FA mechanism. In both model economies, there are households, four types of firms and the central bank as common agents of both models.

A. Households

Households are composed of workers and entrepreneurs whose jobs are fixed for their lives. For workers, there is a continuum of households indexed by $m \in [0, 1]$. However, they are assumed to be a representative agent when they make their intertemporal decision between consumption and leisure. The households maximize the utility function,

$$E_0 \sum_{t=0}^{\infty} \beta^t \exp(z_t^b) \left[\frac{(C_t - \theta C_{t-1})^{1-\sigma}}{1-\sigma} - (Z_t^*)^{1-\sigma} \exp(z_t^h) \int \frac{(h_t(m))^{1+\chi}}{1+\chi} dm \right],$$

subject to their budget constraint,

$$P_t C_t + B_t = r_{t-1}^n B_{t-1} + P_t \int W_t(m) h_t(m) dm + T_t., \tag{2.1}$$

where E_t is the expectation operator in period t, $\beta \in (0, 1)$ is the discount factor, σ and $\chi > 0$ are the degrees of the inverse of intertemporal elasticity of consumption and the inverse of elasticity of the labor supply, respectively. $\theta \in [0, 1]$ denotes the habit persistence of consumption. z_t^b and z_t^h are shocks of preference and the labor supply. Z_t^*, the composite technological level, is set in the disutility term to realize the balance growth path. And P_t is the price of consumption goods, B_t is the government bond, r_t^n is the gross interest rate. $W_t(m)$ is worker m 's real wage, and T_t is the total profit received from firms and lump-sum public transfer. Then, the first-order conditions for the above optimization problem are given by

$$\Lambda_t = \exp(z_t^b)(C_t - \theta C_{t-1})^{-\sigma} - \beta\theta E_t \exp(z_{t+1}^b)(C_{t+1} - \theta C_t)^{-\sigma}, \tag{2.2}$$

$$1 = E_t \beta \frac{\Lambda_{t+1}}{\Lambda_t} \frac{r_t^n}{\pi_{t+1}}, \tag{2.3}$$

where Λ_t is the marginal utility of consumption and π_t is the gross inflation rate of consumption goods, i.e., P_t/P_{t-1}.

i) Workers

The workers indexed by $m \in [0,1]$ supply their differentiated labor service with the substitution elasticity $\theta_t^W > 1$ under monopolistic competition. Based on a Calvo-style staggered wage-setting rule, the wage reoptimized in period t is decided so as to maximize

$$E\sum_{t=0}^{\infty}(\beta\xi_w)^t\left[\Lambda_{t+j}h_{t+j|t}(m)\frac{P_tW_t(m)}{P_{t+j}}\prod_{k=1}^{j}(z^*\pi_{t+k-1}^{\gamma_w}\pi^{1-\gamma_w})-\frac{\exp(z_{t+j}^b)(Z_{t+j}^*)^{1-\sigma}\ \exp(z_{t+j}^h)(h_{t+j|t}(m))^{1+\chi}}{1+\chi}\right],$$

subject to the labor demand in period $t+h$,

$$h_{t+j|t}(m)=h_{t+j}\left[\frac{P_tW_t(m)}{P_{t+j}W_{t+j}}\prod_{k=1}^{j}(z^*\pi_{t+k-1}^{\gamma_w}\pi^{1-\gamma_w})\right]^{-\theta_{t+j}^w},$$

where a fraction $1-\xi_W \in (0,1)$ of wages is reoptimized, whereas the remaining fraction ξ_W is chosen by the indexation rule made from the steady state of the gross growth rate, z^*, and a weighted average of past inflation and its steady state, $\pi_{t-1}^{\gamma_w}\pi^{1-\gamma_w}$, where $\gamma_w \in [0,1]$ is the weight on the past inflation. The first-order condition for the reoptimized real wage is given by

$$1=\frac{E_t\sum(\beta\xi_w)^j\frac{(1+\lambda_{t+j}^W)\ \exp(z_t^b)\ \exp(z_{t+j}^h)(Z_{t+j}^*)^{1-\sigma}}{\lambda_{t+j}^W}\left(h_{t+j}\left\{\frac{W_{t-j}^O(z^*)^j}{W_t}\prod\left[(\frac{\pi_{t+k-1}}{\pi})^{\gamma_w}\frac{\pi}{\pi_{t+k}}\right]\right\}^{-\frac{1+\lambda_{t+j}^W}{\lambda_{t+j}^W}}\right)^{1+\chi}}{E_t\sum(\beta\xi_w)^j\frac{\Lambda_{t+j}W_{t+j}}{\lambda_{t+j}^W}h_{t+j}\left\{\frac{W_{t-j}^O(z^*)^j}{W_t}\prod\left[(\frac{\pi_{t+k-1}}{\pi})^{\gamma_w}\frac{\pi}{\pi_{t+k}}\right]\right\}^{-\frac{1}{\lambda_{t+j}^W}}},\tag{2.4}$$

where $\lambda_t^W = 1/(\theta_t^W - 1) > 0$ stands for the wage markup. And, the aggregate wage, W_t, can be rewritten as

$$1=(1-\xi_w)\left(\left(\frac{W_t^O}{W_t}\right)^{-\frac{1}{\lambda_t^W}}+\sum(\xi_W)^j\left\{\frac{W_{t-j}^O(z^*)^j}{W_t}\prod\left[(\frac{\pi_{t-k}}{\pi})^{\gamma_w}\frac{\pi}{\pi_{t-k+1}}\right]\right\}^{-\frac{1}{\lambda_{t+j}^W}}\right),\tag{2.5}$$

from the definition of the aggregate wage,

$$W_t=\left[\int_0^1(W_t(m))^{1-\theta_t^W}\,dm\right]^{\frac{1}{1-\theta_t^W}},$$

where W_t^O is a reoptimized wage.

ii) Entrepreneurs

Entrepreneurs are owners of capital who decide the utilization rate u_t on capital K_{t-1} after purchasing K_{t-1} at the real price Q_{t-1} from capital goods firms explained later. And they provide capital service $u_t K_{t-1}$ at the real rental rate R_t^K for intermediate goods firms. The first-order condition for the optimal decision on the capital utilization rate is given by

$$R_t^k = Q_t \delta'(u_t), \tag{2.6}$$

where $\delta(u_t)$ is a depreciation rate function whose properties are $\delta' > 0$, $\delta'' > 0$, $\delta(1) \in (0,1)$, and $\delta'(1)/\delta'' = \tau > 0$. Since the real return from purchasing capital K_t is equal to that of holding the bond, the equilibrium equation between them is given by

$$E_t \Lambda_{t+1} \chi_{t+1} = E_t \Lambda_{t+1} \frac{r_t^n}{\pi_{t+1}}, \tag{2.7}$$

where the marginal return on capital is χ_t given by

$$\chi_t = \frac{u_t R_t^k + Q_t(1-\delta(u_t))}{Q_{t-1}}, \tag{2.8}$$

since the resulting capital $(1-\delta(u_t))K_t$ is evaluated at the price Q_t.

B. Firms

There are four types of firms based on the categories of goods: intermediate goods, consumption goods, investment goods and capital goods.

i) Intermediate goods firms

There is a continuum of intermediate goods firms indexed by $f \in [0,1]$. They produce intermediate goods by demanding labor and capital inputs and provide the goods to consumption goods firms. The production function of an intermediate goods firm, f, is given by

$$Y_t(f)\, d_t = (Z_t h_t(f)\,)^{1-\alpha} (u_t K_{t-1}(f)\,)^{\alpha} - \phi y Z_t^*, \tag{2.9}$$

where Z_t stands for the level of neutral technology following the stochastic process, $\log Z_t = \log z + \log Z_{t-1} + z_t^z$, where z and z_t^z denote the steady state of the level and a neutral technology shock, respectively. After aggregating the function, the marginal rate of substitution between labor input and capital input is obtained from

$$\frac{1-\alpha}{\alpha} = \frac{W_t h_t}{R_t^k u_t K_{t-1}}, \tag{2.10}$$

and the marginal cost of the production function is written as

$$mc_t = \left(\frac{W_t}{(1-\alpha)Z_t} \right)^{1-\alpha} \left(\frac{R_t^k}{\alpha} \right)^{\alpha}. \tag{2.11}$$

The firms supply their differentiated goods with the substitution elasticity $\theta_t^P > 1$ under monopolistic competition. Based on a Calvo-style staggered price-setting rule, the price reoptimized in period t is decided so as to maximize

$$E_t \sum_{j=0}^{\infty} \xi_p^j \left(\beta^j \frac{\Lambda_{t+j}}{\Lambda_t}\right) \left[\frac{P_t(f)}{P_{t+j}} \prod_{k=1}^{j} \left(\pi_{t+k-1}^{\gamma_p} \pi^{1-\gamma_p}\right) - mc_{t+j}\right] Y_{t+j|t}(f),$$

subject to the goods demand function in period $t+j$,

$$Y_{t+j|t}(f) = Y_{t+j} \left[\frac{P_t(f)}{P_{t+j}} \prod_{k=1}^{J} (\pi_{t+k-1}^{\gamma_p} \pi^{1-\gamma_p})\right]^{-\theta_{t+j}^P},$$

where a fraction $1-\xi_P \in (0,1)$ of the price is reoptimized, whereas the remaining fraction ξ_P is chosen by the indexation rule from a weighted average of past inflation and its steady state, $\pi_{t-1}^{\gamma_p} \pi^{1-\gamma_p}$, where $\gamma_p \in [0,1]$ is the weight on the past inflation. The first-order condition for the reoptimized price is given by

$$1 = \frac{E_t \sum_{j=0}^{\infty} (\beta\xi_p)^j \frac{(1+\lambda_t^P)\, mc_{t+j}\, \Lambda_{t+j}\, Y_{t+j}}{\lambda_{t+j}^P} \left\{\frac{P_t^O}{P_t} \prod_{k=1}^{j} \left[(\frac{P_t^O}{p_t})^{\gamma_p} \frac{\pi}{\pi_{t+k}}\right]\right\}^{-\frac{1+\lambda_{t+j}^P}{\lambda_{t+j}^P}}}{E_t \sum_{j=0}^{\infty} (\beta\xi_p)^j \frac{\Lambda_{t+j} Y_{t+j}}{\lambda_{t+j}^P} \left\{\frac{P_t^O}{P_t} \prod_{k=1}^{j} \left[(\frac{P_t^O}{p_t})^{\gamma_p} \frac{\pi}{\pi_{t+k}}\right]\right\}^{-\frac{1}{\lambda_{t+j}^P}}}, \tag{2.12}$$

where $\lambda_t^W = 1/(\theta_t^W - 1) > 0$ stands for the wage markup.

ii) Consumption goods firms

Consumption goods firms produce output Y_t by using intermediate goods as input. Under perfect competition, the firms maximize,

$$P_t Y_t - \int_0^1 P_t(f)\, Y_t(f)\, df,$$

subject to transformation technology,

$$Y_t = \left(\int_0^1 Y_t(f)^{(\theta_t^p - 1)/\theta_t^P} df\right)^{\theta_t^P/(\theta_t^P - 1)},$$

with respect to Y_t . And, using Eq.(2.12), the price of consumption goods, P_t, can be rewritten as

$$1 = (1-\xi_p) \left(\left(\frac{P_t^O}{P_t}\right)^{-\frac{1}{\lambda_t^p}} - \sum (\xi_p)^j \left\{\frac{P_{t-j}^O}{P_{t-j}} \prod_{j=1}^{\infty} \left[\left(\frac{\pi_{t-k}}{\pi}\right)^{\gamma_p} \frac{\pi}{\pi_{t-k+1}}\right]\right\}^{-\frac{1}{\lambda_t^p}} \right), \tag{2.13}$$

from the definition of the price,

$$P_t = \left[\int_0^1 (P_t(f))^{1-\theta_t^p} df\right]^{\frac{1}{1-\theta_t^p}},$$

where P_t^O is a reoptimized price.

iii) Investment goods firms

There is a continuum of investment goods firms indexed by $k \in [0,1]$. They convert one unit of consumption goods into Ψ_t units of differentiated investment goods by using production technology, $I_t = (\int_0^1 I_t(k)^{(\theta_t^i-1)/\theta_t^i} dk)^{\theta^i/(\theta_t^i-1)}$ where the substitution elasticity $\theta_t^i > 1$. Under monopolistic competition, an investment-goods firm, g, maximizes its profit function,

$$\left(P_t^i(k)/P_t - 1/\Psi_t\right) I_t(k),$$

subject to the demand function,

$$I_t(k) = I_t \left(\frac{P_t^i(k)}{P_t^i}\right)^{-\theta_t^i},$$

where P_t^i is the investment goods price. The unit of investment goods follows a stochastic dynamics, $\log \Psi_t = \log \psi + \log \Psi_{t-1} + z_t^\psi$, where z_t^ψ is an investment specific (IS) shock. The first-order condition for profit maximization of investment goods firms is given by

$$P_t^i = (1 + \lambda_t^i)\frac{P_t}{\Psi_t}, \tag{2.14}$$

where $\lambda_t^i = 1/(\theta_t^i - 1) > 0$ stands for the investment goods markup.

iv) Capital goods firms

Capital goods firms produce investment I_t by using differentiated investment goods $I_t(k)$ as input. Under perfect competition, the firms maximize,

$$E_t \sum_{j=0}^{\infty} \beta^j \frac{\Lambda_{t+j}}{\Lambda_t} \left\{ Q_{t+j}\left[K_{t+1} - (1-\delta(u_{t+j}))K_{t+j-1}\right] - \frac{P_{t+j}^i}{P_{t+j}} I_{t+j} \right\},$$

subject to the capital accumulation equation,

$$K_t = (1-\delta(u_t))K_{t-1} + \exp(z_t^\nu)\left(1 - S\left(\frac{I_t/I_{t-1}}{z^*\psi}\right)\right) I_t, \tag{2.15}$$

where $S((I_t/I_{t-1})/(z^*\psi)) = (\zeta/2)[I_t/I_{t-1}/(z^*\psi) - 1]^2$, $\zeta > 0$, is an increasing adjustment cost of investment and z_t^ν is the marginal efficiency of investment (MEI) shock. And the first-order condition for profit maximization of capital goods firms is given by

$$\frac{P_t^i}{P_t} = Q_t \exp(z_t^\nu)\left[1 - S\left(\frac{I_t/I_{t-1}}{z_t^*\psi}\right) - S'\left(\frac{I_t/I_{t-1}}{z^*\psi}\right)\frac{I_t/I_{t-1}}{z^*\psi}\right] + E_t\beta\frac{\Lambda_{t+1}}{\Lambda_t} z_t^*\psi Q_{t+1}\exp(z_t^\nu)S'\left(\frac{I_{t+1}/I_t}{z^*\psi}\right). \tag{2.16}$$

2.2 Financial Accelerator Mechanism

As the DSGE model with the FA mechanism, we incorporate the bank sector (financial intermediaries) as an additional agent in the benchmark model described above. In this framework, entrepreneurs purchase capital by borrowing loans from financial intermediaries at the gross loan rate, r_t^E, aside from financing by their net worth,

$$B_t = Q_t K_t - N_t. \tag{2.17}$$

where B_t and N_t denote their real borrowing and net worth, respectively. The presence of asymmetric information between borrowers and lenders makes a loan rate greater than the deposit rate (or the policy rate), r_t^n, i.e., $r_t^E > r_t^n$. The discrepancy between the two rates is referred to as the external financial premium (EFP), expressed as

$$\frac{r_t^E}{r_t^n} = F\left(\frac{Q_t K_t}{N_t}\right) \exp(z_t^\mu).$$

where $F(\cdot)$ is a function of entrepreneurs' leverage ratio which fulfills $F' > 0$ and z_t^μ is an EFP shock. By log-linearizing this equation, the EFP is obtained from

$$r_t^E - r_t^n = \mu_E(q_t + k_t - n_t) + z_t^\mu, \tag{2.18}$$

where μ_E is set as $\mu_E = (Q_t K_t/N_t)\frac{F'(Q_t K_t/N_t)}{F(Q_t K_t/N_t)} > 0$, and represents a degree of the EFP decided from a leverage ratio, $q_t + k_t - n_t$. Instead of Eq.(2.7), the marginal return on capital χ_t is decided from the loan rate such as

$$E_t \Lambda_{t+1} \chi_{t+1} = E_t \Lambda_{t+1} \frac{r_t^E}{\pi_{t+1}}. \tag{2.19}$$

In each period, a fraction $1-\eta_t \in (0,1)$ of entrepreneurs change to workers and the same amount of workers become entrepreneurs. The remaining fraction η_t of them survive until the next period. The dynamic of their net worth is

$$N_t = \eta_t \left(\chi_t Q_{t-1} K_{t-1} - \frac{r_{t-1}^E}{\pi_t} L_{t-1}\right) + (1-\eta_t)\omega Z_t^*, \tag{2.20}$$

where ωZ_t^* is the net worth of new comers switching from workers. The survival rate, η_t, follows a stochastic process, $\eta_t = \eta \exp(z_t^\eta)/(1 - \eta + \eta \exp(z_t^\eta))$, where z_t^η is regarded as a net worth shock.

2.3 Miscellaneous

The central bank

The central bank decides the policy rate r_t^n based on a Taylor type monetary policy rule,

$$\log\, r_t^n = \phi_r \log r_{t-1}^n + (1-\phi_r)\left\{ \log\, r^n + \left(\frac{\phi_\pi}{4} \sum_{j=0}^{3} \log \frac{\pi_{t-j}}{\pi} \right) + \phi_y \log \frac{Y_t}{Y_{t-1}} \right\} + z_t^r, \qquad (2.21)$$

where $\phi_r \in (0,1)$ is a degree of persistence of the policy rate, and ϕ_π and ϕ_y stand for reaction coefficients of inflation and output growth, respectively. r^n denotes the steady state of the policy rate and z_t^r is a monetary policy shock.

The market clearing condition

The market clearing condition with respect to consumption goods is written as

$$Y_t = C_t + \int_0^1 \frac{I_t(k)}{\Psi_t} dk + gZ_t^* \exp(z_t^g) = C_t + \frac{I_t}{\Psi_t} + gZ_t^* \exp(z_t^g), \qquad (2.22)$$

where differentiated investment goods are aggregated as investment, $\int_0^1 \frac{I_t(k)}{\Psi_t} dk = I_t/\Psi_t$, and $gZ_t^* \exp(z_t^g)$ represents the exogenous demand of output except for the consumption of households and investment of firms, where z_t^g is an exogenous demand shock.

Equilibrium conditions and exogenous shocks

To solve an equilibrium of the NK model, the conditions consist of Eq.(2.2) through Eq.(2.16) and Eq.(2.21) and Eq.(2.22). Meanwhile, aside from the above conditions, Eq. (2.17) through Eq.(2.20) are additionally used as the conditions of the FA model, instead of Eq.(2.7). There are nine structural shocks consisting of three technology shocks, z_t^ν, z_t^z and z_t^ψ, three demand and policy shocks, z_t^b,z_t^g and z_t^r, and three markup shocks, z_t^p, z_t^i and z_t^w in the NK model. There are eleven shocks including two additional shocks: z_t^η, z_t^μ, in the FA model.

To get the steady state of both models in which the economy grows at the composite technological level Z_t^* given by $Z_t^* = Z_t(\Psi_t)^{\alpha/(1-\alpha)}$, we make endogenous variables detrend such as $y_t = Y_t/Z_t^*$, $c_t = C_t/Z_t^*$, $w_t = W_t/Z_t^*$, $\lambda_t = \Lambda_t(Z_t^*)^\sigma$, $i_t = I_t/(Z_t^*\Psi_t)$, $k_t = K_t/(Z_t^*\Psi_t)$, $r_t^k = R_t^k/(Z_t^*\Psi_t)$, $q_t = Q_t\Psi_t$, $n_t = N_t/Z_t^*$, and $b_t = B_t/Z_t^*$. The equilibrium conditions in terms of detrended variables log-linearized around the steady state are described in the Appendix.

2.4 Impulse Response Functions

Here, we consider the properties of the FA by comparing the impulse response functions (IRFs) of the two models in order to show differences of comovements among endogenous variables between the two DSGE models. We calibrate IRFs of both models using the same parameters and setting μ_E as 0.05. In Panels (a) to (d) of Figure 1, the IRFs of six endogenous variables in response to four structural shocks: preference, monetary policy, IS and investment price markup shocks are drawn, in which the blue solid and red dashed lines represent the FA model and the NK model, respectively.

The EFP is a key factor to represent the effects of the financial friction as seen from the difference between Eq. (2.7) and Eq.(2.19) in terms of the model description. Figure 1 (a) shows the IRF to the positive preference shock, which causes the EFP to shrink. The reduction of the EFP pushes down the loan rate and reduces the size of the decline of investment in the FA model compared with the NK model. In Panel (b), IRFs to a positive interest rate shock are depicted, in which the EFP becomes larger. We see that five variables excluding consumption are amplified and prolonged in the FA model. Panel (c) shows the IRF to the positive IS technology shock which makes the EFP expand through an increase in the leverage ratio, since the ratio of capital to output increases. A larger EFP weakens the effects to the IS shock of endogenous variables excluding consumption. Panel (d) draws the IRFs to the positive investment price markup shock, which reduces the EFP through a decrease in the demand of investment. Shrinkage of the EFP damps the amplification effect of all the endogenous variables.

[Insert Figure 1 about here]

2.5 Measurement equations

The state space models of the DSGE models consist of state equations composed of log-linearized equilibrium conditions described in the Appendix, and measurement equations. Here, we describe the measurement equations of both models as below. The equations of the NK model adopt eight observed series: output:Y_t, consumption: C_t, investment: I_t, real wage: W_t, labor input: L_t, inflation: π_t, investment price: P_t^i, and policy interest rate: r_t^n, while those of the FA model use ten series including two additional observed variables, the loan rate:r_t^E and the real borrowing: B_t.

NK model

$$\begin{bmatrix} 100\Delta\log Y_t \\ 100\Delta\log C_t \\ 100\Delta\log I_t \\ 100\Delta\log W_t \\ 100\log l_t \\ 100\Delta\log P_t \\ 100\Delta\log(P_t^i/P_t) \\ 100{r^n}_t \end{bmatrix} = \begin{bmatrix} \bar{z^*} \\ \bar{z^*} \\ \bar{z^*}+\bar{\psi} \\ \bar{z^*} \\ \bar{l} \\ \bar{\pi} \\ -\bar{\psi} \\ \bar{r^n} \end{bmatrix} + \begin{bmatrix} z_t^* + \hat{y}_t - \hat{y}_{t-1} \\ z_t^* + \hat{c}_t - \hat{c}_{t-1} \\ z_t^* + z_t^{\psi} + \hat{i}_t - \hat{i}_{t-1} \\ z_t^* + \hat{w}_t - \hat{w}_{t-1} \\ \hat{l}_t \\ \hat{\pi}_t \\ -z^{\psi} + z_t^{\nu} - z_t^{\nu} \\ \hat{r}_t^n \end{bmatrix} + \begin{bmatrix} e_t^Y \\ e_t^C \\ e_t^I \\ e_t^W \\ e_t^I \\ e_t^P \\ e_t^{Pi} \\ e_t^r \end{bmatrix}, \qquad (2.23)$$

FA model

$$\begin{bmatrix} 100\Delta\log Y_t \\ 100\Delta\log C_t \\ 100\Delta\log I_t \\ 100\Delta\log W_t \\ 100\log l_t \\ 100\Delta\log P_t \\ 100\Delta\log(P^i_t/P_t) \\ 100 r^n_t \\ 100\, r^E_t \\ 100\Delta\log B_t \end{bmatrix} = \begin{bmatrix} \bar{z^*} \\ \bar{z^*} \\ \bar{z^*}+\bar{\psi} \\ \bar{z^*} \\ \bar{l} \\ \bar{\pi} \\ -\bar{\psi} \\ \bar{r^n} \\ \bar{r^E} \\ \bar{z^*} \end{bmatrix} + \begin{bmatrix} z^*_t+\hat{y}_t-\hat{y}_{t-1} \\ z^*_t+\hat{c}_t-\hat{c}_{t-1} \\ z^*_t+z^\psi_t+\hat{i}_t-\hat{i}_{t-1} \\ z^*_t+\hat{w}_t-\hat{w}_{t-1} \\ \hat{l}_t \\ \hat{\pi}_t \\ -z^\psi+z^\nu_t-z^\nu_t \\ \hat{r}^n_t \\ \hat{r}^E_t \\ z^*_t+\hat{b}_t-\hat{b}_{t-1} \end{bmatrix} + \begin{bmatrix} e^Y_t \\ e^C_t \\ e^I_t \\ e^W_t \\ e^l_t \\ e^P_t \\ e^{Pi}_t \\ e^r_t \\ e^{rE}_t \\ e^B_t \end{bmatrix}, \tag{2.24}$$

where we set $\bar{z^*} = 100(z^* - 1)$, $\bar{\psi} = 100(\psi - 1)$, $\bar{r^n} = 100(r^n - 1)$, and $\bar{l}$ is normalized to be equal to zero following Kaihatsu and Kurozumi (2010, 2014b). The hatted letters indicate log-deviations from steady-state values after detrending with the level of composite technology Z^*_t. The third term of the RHS is referred to as the measurement errors of the observable variables. And we set $\bar{r^E} = 100(r^E - 1)$.

2.6 Data

The data to estimate the models are basically based on Kaihatsu and Kurozumi (2010, 2014 b). The data on the relative price of investment P^i_t/P_t, output, and consumption C_t are given by dividing the investment deflator, nominal GDP and nominal consumption with the CPI. The data on investment I_t, labor input, real wage and policy interest rate are the same as those in Sugo and Ueda (2008), except that these series are not detrended. The data on the loan rate are the average interest rate on contracted loans and discounts. The sample period is from 1981:Q1 to 1998:Q4.

3 Markov switching (MS) prediction pool

3.1 Predictive scores

From a Bayesian perspective, the *marginal likelihood* is commonly used as a criterion of model choice, since it is interpreted as the predictive density of a model obtained by integrating with respect to the prior density of the model parameters Θ. A model with the highest predictive density is thought to be the best model explaining the behaviors of observations based on information on all of the data. Let us denote a vector of future observations as y_{t+h}, where h is an h-step-ahead forecast, and its history is $Y^o_t = \{y_g, ..., y_t\}$, where $g \leq 1$ is the starting date and the superscript "o" denotes the observed data. The predictive density of a model with respect to the prior of parameters Θ is defined as

$$p^{Prior}\left(y^f_{t+h} - y^o_{t+h}|Y^o_t, \mathcal{M}\right) \equiv \int p\left(y^f_{t+h} - y^o_{t+h}|Y^o_t, \Theta, \Sigma, \mathcal{M}\right) p(\Theta|\mathcal{M})p(\Sigma)d\Theta d\Sigma,$$

where y^f_{t+h}, and y^o_{t+h} are the forecast and observed values in period $t+h$, respectively, and the difference between them is their forecasting errors ε_{t+h}. Σ is a covariance matrix of the forecasting errors, ε_{t+h}, and $\mathcal{M}$ is a prediction model. $p\left(\varepsilon_{t+h}|Y_t, \Theta, \Sigma, \mathcal{M}\right)$, and $p(\Theta|\mathcal{M})$ denote the likelihood function and the prior density of Θ of a prediction model $\mathcal{M}$, respectively. When we set $h=1$, then the density is regarded as the marginal likelihood. When replacing the prior density with the posterior density of Θ according to Geweke (2010), the predictive density can be redefined as a *posterior* predictive density,

$$p^{Post}\left(y^f_{t+h} - y^o_{t+h} \,|\, Y^o_t, \mathcal{M}\right) \equiv \int p\left(y^f_{t+h} - y^o_{t+h} \,|\, Y^o_t, \Theta, \Sigma, \mathcal{M}\right) p(\Theta \,|\, Y^o_t, \,\mathcal{M}) d\Theta d\Sigma,$$

where $p(\Theta|Y_t, \mathcal{M})$ is the posterior density of Θ conditional on the history of observations until period t, Y^O_t, and a model, $\mathcal{M}$. Following Geweke and Amisano (2011), we use the posterior predictive density in order to construct a *predictive score* for evaluating the forecasting performance of a single prediction model and of a convex combination of multiple prediction models with the optimal model weights. We define the predictive score of a model $\mathcal{M}$, $p(y^f_{t+h}\,;\, Y^o_t,\, \mathcal{M})$, for the h-step-ahead forecast as

$$p(y^f_{t+h}\,;\, Y^o_t,\, \mathcal{M}) \equiv p^{Post}\left(y^f_{t+h} - y^o_{t+h} \,|\, Y^o_t,\, \mathcal{M}\right),$$

and regard it as the key element of the following prediction pooling methods.

Forecast combination of multiple models has been known as a useful tool for improving the performance. Most of studies for the model combination have focused on point forecasts and were reviewed by Timmermann (2006) and Elliott and Timmermann (2016). Meanwhile, studies for model combination in terms of density forecasts had been much more limited, but have been recently paid more attention by macro-econometricians. Geweke and Amisano (2011) propose the optimal prediction pool with respect to density forecasts, referred to as the static prediction pool. Let us redefine $\mathcal{M}$ as the collection of competing multiple models, e.g., $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$. Given two prediction models $\mathcal{M}_1$ and $\mathcal{M}_2$, the predictive score for the h-step-ahead forecast can be constructed as the convex combination of the predictive scores of competing models,

$$p^{SP}\left(y^f_{t+h};\, Y^o_t, \mathcal{M}\right) \equiv \lambda\, p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_1\right) + (1-\lambda)\; p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_2\right), \tag{3.1}$$

where $\lambda \in (0,1)$ and $1-\lambda$ are constant values indicating model weights in favor of $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively. The optimal prediction pooling is then obtained by maximizing the cumulative log predictive score, LPS^{SP}, for the whole of the prediction periods as

$$LPS^{SP}(\lambda, h) \equiv \sum_{t=1}^{T} \log\left[\lambda p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_1\right) + (1-\lambda)\, p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_2\right)\right], \tag{3.2}$$

by choosing $\lambda^* = \arg\max LPS^{SP}(\lambda, h)$. An important assumption, as noted by Geweke and Amisano (2011), is that the two candidate prediction models have to be substantially different in terms of the functional form of their predictive densities (i.e., non-nested models). In our study, we generate a predictive density of macroeconomic observations based on each of the two DSGE models described in Section 2 from posterior estimations of their model parameters.

3.2 MS prediction pool

Waggoner and Zha (2012) extended the static prediction pool of Geweke and Amisano (2011) as Eq.(3.1) by allowing the weighting coefficient λ_{t+h} to be dependent on a regime variable, s_t, following a Markov chain as

$$\lambda_{t+h} = \lambda\left(s_{t+h}\right) = \begin{cases} \lambda_1, & s_t = 1 \\ \lambda_2, & s_t = 2 \end{cases},$$

where λ_1, $\lambda_2 \in (0,1)$ are constant weights in favor of $\mathcal{M}_1$ in the period of regimes 1 and 2, respectively. The transition probabilities matrix, Q , of the Markov chain with two regimes is given by

$$Q = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix},$$

where the element q_{ij} is a transition probability from state i in period $t-1$ to state j in period t, i.e., $q_{ij} = \Pr\left(s_t = j \middle| s_{t-1} = i\right)$ with $q_{11} + q_{12} = 1$ and $q_{21} + q_{22} = 1$. Again, we set $\mathcal{M} = (\mathcal{M}_1, \mathcal{M}_2)$. Conditional on the state s_t, the predictive score for the h-step-ahead forecast based on the MS pool can be expressed as

$$\begin{aligned} p^{MS}\left(y^f_{t+h};\, Y^o_t,\, s_t,\, \mathcal{M}\right) &= \lambda\left(s_{t+h}|s_t\right) p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_1\right) + \left(1 - \lambda\left(s_{t+h}|s_t\right)\right) p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_2\right) \\ &= \begin{bmatrix} p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_1\right) & p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_2\right) \end{bmatrix} \begin{bmatrix} \lambda\left(s_{t+h}|s_t\right) \\ 1 - \lambda\left(s_{t+h}|s_t\right) \end{bmatrix}, \end{aligned}$$

where the conditional weight is calculated from $\lambda\left(s_{t+h}|s_t\right) = \sum_{i=1}^{2} \lambda_i \Pr(s_{t+h} = i|s_t)$, and the h-step-ahead transition probability of state s_{t+h} conditional on s_t is obtained from $\Pr(s_{t+h} = i|s_t) = Q_i^h$ where Q_i is the sum of the i-th column of Q.

Hence, by integrating out the unobservable regime, s_t, for $p^{MS}\left(y^f_{t+h};\, Y^O_t, \mathcal{M},\, s_t\right)$, we have the predictive score of the MS pooling method conditional on Y_t and $\mathcal{M}$, given as

$$\begin{aligned} p^{MS}\left(y^f_{t+h};\, Y^o_t, \mathcal{M}\right) &= \sum_{s_t=1}^{2} p^{MS}\left(y^f_{t+h};\, Y^o_t, s_{t+h},\, \mathcal{M}\right) \Pr\left(s_{t+h}|\, s_t\right) \Pr\left(s_t|\, Y^o_t,\, \mathcal{M}\right) \\ &= \begin{bmatrix} p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_1\right) & p\left(y^f_{t+h};\, Y^o_t, \mathcal{M}_2\right) \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1-\lambda_1 & 1-\lambda_2 \end{bmatrix} \begin{bmatrix} q_{11} & 1-q_{11} \\ 1-q_{22} & q_{22} \end{bmatrix}^h \begin{bmatrix} \Pr(s_{1,t}) \\ 1-\Pr(s_{1,t}) \end{bmatrix}, \end{aligned} \tag{3.3}$$

where $\Pr\left(s_t|Y^o_t,\, \mathcal{M}\right)$is the posterior probability of s_t conditional on Y_t and $\mathcal{M}$ derived from the Hamilton (1989) filter described in the footnote[2]. The last term,$[Pr(s_{1,t})\; 1{-}Pr(s_{1,t})]'$, is $p\left(s_t|Y^o_t,\, \mathcal{M}\right)$ so that the expected values of them are adopted in the term.

Using Eq.(3.3), the MS prediction pool with two regimes of the log predictive score of the

[2]Using the Bayes theorem, we obtain a relation such as that $\Pr\left(s_t|Y^O_t,\, \mathcal{M}\right)$ is proportional to $p(y_t|s_t, Y^O_{t-1}, \mathcal{M})$ $\times\Pr(s_t|s_{t-1})$ $\times\Pr\left(s_{t-1}|Y^O_{t-1}, \mathcal{M}\right)$, where $p(y_t|s_t, Y^O_{t-1}, \mathcal{M})$ is a likelihood function of y_t given s_t and Y^O_{t-1}. Since we know that $\Pr(s_t|s_{t-1}) = Q$ as well as the values of $\Pr\left(s_{t-1}|Y^O_{t-1}, \mathcal{M}\right)$ and the likelihood function of y^O_t, we easily calculate value of $p\left(s_t|Y^O_t,\, \mathcal{M}\right)$. In this way, we obtain $p\left(s_t|Y^O_t,\, \mathcal{M}\right)$ for $t = 1 \cdots T$ by iterating the calculation from period 1.

h-step-ahead forecast for periods, $t = 1, \cdots, T$, is given as

$$LPS^{MS}(\lambda_1, \lambda_2, h) \equiv \sum_{t=1}^{T} \log p^{MS}\left(y^f_{t+h}\,;\, Y^o_t, \mathcal{M}\right). \tag{3.4}$$

An advantage of using the MS modeling for the weighting coefficient is that we can identify the relative importance of the models during different sample periods. Waggoner and Zha (2012) show that the DSGE model plays an important role relative to a Bayesian VAR model only in the late 1970's and the early 1980's. It is important to note that we do not incorporate an assumption of regime-switching into the economic dynamics with forward-looking agents. The regime of the MS prediction pool only reflects the particular period in history in which one model prevails over the others in terms of its density forecasts.

3.3 Estimation methodology of prediction pool

In order to estimate and compare the predictive scores of individual prediction models, say two DSGE models, and the pooling methods, we adopt a Bayesian approach with the Markov Chain Monte Carlo (MCMC) method for the MS prediction pooling as well as static and dynamic prediction poolings, which are examined in Section 5 as a robust check.

Although Waggoner and Zha (2012) simultaneously estimated two macroeconomic models and the pooling method, the simultaneous estimation of the model parameters, Θ, under a regime sustained only for a short period is thought to have only a low level of accuracy. This is because a regime generated in each MCMC iteration of a pooling method is different from that of the previous iteration, and a different regime period expands the variations of drawing Θ in the step of MCMC iteration in the DSGE model. By adopting a two-step procedure following Geweke and Amisano (2011) and Del Negro et al. (2016), we can avoid generating instability in the model parameters, Θ, estimated based on different regime periods.

The two-step procedure is described below.

Step 1. Make density forecasts of the DSGE models.

1. The posterior estimates of parameters, $p(\mathit{\Theta}|Y^o_{t-1}, \mathcal{M}_i)$, under the DSGE models, M_i, for $i = 1, \cdots, n$, are obtained for the full sample period, using the MCMC method.
2. We compute the predictive densities and predictive scores of observations, $p(y^f_{t+h}|Y^o_t, \Theta, M)$, from sampling of $p(\mathit{\Theta}|Y^o_{t-1}, \mathcal{M}_i)$ of each DSGE model, $\mathcal{M}_{\mathrm{i}}$, by the Monte Carlo simulation technique.

Step 2. Make the optimal combination of density forecasts.

1. We calculate the optimal combination of the log scores of the DSGE models obtained in the previous step, using parameters of pooling methods drawn from the Gibbs sampling method with the Hamilton filter following Albert and Chib (1993).

4 Empirical results

4.1 Model estimation

The frictionless DSGE model (NK model) and the model with financial friction (FA model) mentioned in Sections 2.1, 2.2 and 2.3 are estimated with the Japanese data as in Section 2.6 using Bayesian estimation via the MCMC method. The prior distributions of the parameters shown in Table A1 are decided based on Kaihatsu and Kurozumi (2010), who dealt only with the model with financial friction. The characteristic values of the prior in the NK model are the same as those of its counterpart. Notice that the NK model has just eight fewer structural parameters than the FA model.[3] To form posterior distributions of the parameters, 300,000 iterations are implemented in the MCMC. After the first 150,000 draws are discarded, the remaining draws are sampled as the posterior estimates, as shown in Tables A2 and A3. It is noteworthy that parameter μ_E, which generates the difference between the loan rate and the nominal rate, is 0.031 as the posterior mean and from 0.026 to 0.040 as the 95% credible interval, excluding 0 as in Table A3, and that this estimation indicates that financial friction exists in Japan.[4]

4.2 Density forecasts of the DSGE models

We calculate the posterior predictive distributions of the six individual observations for 1982:Q3 – 1998:Q4 from a Monte Carlo (MC) method using model parameters sampled as posterior estimates in Section 4.1. In the procedure, by generating random variables of structural shocks based on the posterior estimates, we calculate the h-step-ahead forecast of observations from a state space model with the shocks, and accumulate them as the predictive distributions. Figure 2, for instance, shows the predictive distributions of the two models for a decade as of the period 1990:Q4. This figure represents the discrepancy of the predictive means (red line) and distributions (shaded area) of the six observations between the two models based on the presence or absence of the EFP. Generally, their predictive means must correspond to a path to a steady state, and the width of the distribution could depend on the variance of the shocks in the case of a DSGE model. As seen in Panel (b), the means of inflation, investment, output, and consumption in the FA model might be amplified by the FA mechanism, compared with the NK model (Panel (a)). And two additional shocks of the FA model are likely to expand the distributions. In particular, that of inflation is significant. For the rest of the section, we focus on the predictive means and distribution for the whole of the sample periods.

[Insert Figure 2 about here]

In Table 1, the means of forecasting errors of the six observations are described in terms of booms and recessions after being classified into three periods; (1) pre-Bubble period, (2) Bub-

[3] The 8 fewer structural parameters in the NK model are represented as "NA" in Table A1 and Table A2.

[4] The posterior estimates of our estimation differ from Kaihatsu and Kurozumi (2010, 2014b) even using the same priors, since we introduce a measurement error to each of the observable variables in the above measurement equations (2.23) and (2.24).

ble period, and (3) post-Bubble period. The error-means is calculated from $E_t(y^i_{t+h}) - y^o_{t+h} = \frac{1}{NH}\sum_{i=1}^{N}\sum_{h=1}^{H}(y^i_{t+h} - y^o_{t+h})$, where superscripts "i " and "o " denote the i-th sample of the MC forecasted and realized observations, respectively, and subscript "h " is the h-step-ahead forecast. And H and N are the maximum number of horizons and total number of MC sampling, respectively. Here, we set $H = 4$ and $N = 20,000$. There are some remarks. First, the means of the nominal interest rate in both models are positive, or overestimated ($E_t y_{t+1} > y_{t+1}$) overall, except for the "Bubble" boom period. Next, the means of inflation are negative or underestimated ($E_t y_{t+1} > y_{t+1}$) in booms, whereas those of wages are basically overestimated. Finally, the means of real series such as output, consumption and investment are negative for booms and positive for recessions, except for the pre-Bubble period. These indicate that it is difficult to forecast the magnitude of fluctuation of a business cycle in terms of the DSGE model, even when incorporating the FA mechanism in the model. Recently, Comin and Gertler (2007) proposed a DSGE model with an endogenous growth model to analyze the middle term of business cycles. We might apply their model to cope with this wrinkle.

[Insert Table 1 about here]

Next, let us assess the predictive distributions in terms of the realized values of the six series. The third and fourth rows of each panel in Table 2 show the log predictive scores of the total and individual variables classified from the three periods in the NK model and the FA model, respectively. The log predictive scores are calculated from the log likelihood function of forecasting error $LS(y_{t+h}) \equiv \frac{1}{NH}\sum_{i}^{N}\sum_{h=1}^{H}\log\, p(y^i_{t+h};\, Y^o_t,\, M_j)$, where y_{t} is a single observation in period t and $p(\cdot)$ is the density function of normal distribution. Again, we set $H = 4$ and $N = 20,000$. Panel (a) represents the log predictive scores for the full sample period, and those of the three periods described in Panels (b), (c) and (d). In the table, bold numbers indicate the better performance between the two models. As concerns the four tables, the FA model is superior to the NK model in the distribution of consumption, investment, and wage for the three periods overall except for pre-Bubble investment (with a tiny difference). On the other hand, the NK model outdoes the FA model in output, inflation and nominal interest rate. Figure 3 shows the time series of the log predictive scores of the six variables. For the whole sample period, the FA model dominates in terms of wage, whereas the NK model dominates on inflation and interest rate. In the remaining three real variables, the dominance between the two models changes at a bewildering pace and depends on the period.

[Insert Figure 3 about here]

[Insert Table 2 about here]

Figure 4 shows the time series of the log predictive score of all six observations in both models with a multi-variate nominal distribution. That is $LS(y_{t+h}) = \frac{1}{NH}\sum_i^N \sum_{h=1}^H \log\, p(y_{t+h}^i;\, Y_t^o,\, M_j)$, where y_t is a 6×1 vector of the whole six observations. We observe that the log predictive score fluctuates with a large amplitude during the "Bubble" boom period; 1988 to1990, in particular for the NK model. As can be seen from Figure 3, since variations of the predictive scores in the three real variables become large for the Bubble period, the total score also reflects this. In the next subsection, we turn to analyze the MS pooling method using the log predictive score of all six variables.

[Insert Figure 4 about here]

4.3 MS prediction pool

The MS prediction pool, Eq.(3.3), is estimated with the MCMC simulation and obtained from 100,000 draws after discarding the first 40,000 burn-in draws. Table 3 shows the estimation result of the MS prediction pool explained in Section 3.2. In the pooling model, regime 1 ($s_t = 1$) indicates a regime in which the NK model beats the FA model in terms of log predictive score, whereas regime 2 ($s_t = 2$) is a regime where the FA model prevails. The model weights of the FA model in the regimes 1 and 2 are around 5 % ($\lambda_1 = 0.05$) and 88 % ($\lambda_2 = 0.88$), respectively. Using the posterior means of the parameters of the MS pooling model as shown in Table 3, the log predictive score of the MS model with the h-step-ahead forecast described in Eq.(3.4) is represented as below.

$$\begin{aligned} p^{MS}\left(y_{t+h}^f;\, Y_t^O, \mathcal{M}\right) &= \sum_{s_t=1}^{2} p^{MS}\left(y_{t+h}^f;\, Y_t^O, s_{t+h}, \mathcal{M}\right) p\left(s_{t+h}|\, s_t\right) p\left(s_t|Y_t^O,\, \mathcal{M}\right) \qquad (4.1) \\ &= \begin{bmatrix} p\left(y_{t+h}^f; Y_t^O, \mathcal{M}_1\right) & p\left(y_{t+h}^f; Y_t^O, \mathcal{M}_2\right) \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1-\lambda_1 & 1-\lambda_2 \end{bmatrix} \begin{bmatrix} q_{11} & 1-q_{11} \\ 1-q_{22} & q_{22} \end{bmatrix}^h \begin{bmatrix} \Pr(s_{1,t}) \\ 1-\Pr(s_{1,t}) \end{bmatrix} \\ &= \begin{bmatrix} p\left(y_{t+h}^f; Y_t^O, \mathcal{M}_1\right) & p\left(y_{t+h}^f; Y_t^O, \mathcal{M}_2\right) \end{bmatrix} \begin{bmatrix} 0.05 & 0.88 \\ 0.95 & 0.12 \end{bmatrix} \begin{bmatrix} 0.87 & 0.13 \\ 0.09 & 0.91 \end{bmatrix}^h \begin{bmatrix} \Pr(s_{1,t}) \\ \Pr(s_{2,t}) \end{bmatrix}, \end{aligned}$$

where regime variable $s_{1,t}$ is one when period t belongs to regime 1, and otherwise, zero. And $p(y_{t+h}^f|Y_t^o, m_{FA})$ and $p(y_{t+h}^f|Y_t^o, m_{NK})$ are the log predictive scores of the FA model and the NK model, respectively.

Like the empirical results of Waggoner and Zha (2012), each regime has an extremely high model weight with which one model overwhelms the other, such as 95% for the NK model in regime 1 and 88% for the FA model in regime 2. Figure 5 shows the time series of regime 2 (the FA model beats the NK model) for the sample period calculated from the estimation. In Panel (a), the black solid line and red dashed line are the posterior means and median of regime variable $(1 - s_{1,t})$ of regime 2, respectively. In Panel (b), these lines are the posterior mean and median model weight (λ_t) of the FA model. More precisely, we set the horizon as h=1 and the posterior probabilities of regime 2 in Panel (a) are calculated from the number of the MCMC draws of $\Pr(s_t^i = s_2|Y_t^o)$, where i denotes the i-th sample, by the Hamilton filter of the MS pooling method as explained in Section 3.2. And similarly, the posterior model weight at each period in Panel (b) is derived from the number of the

MCMC draws of $E(\lambda_t^i|Y_t^o) = \lambda_1^i \times \Pr(s_t^i = s_1|Y_t^o) + \lambda_2^i \times Pr(s_t^i = s_2|Y_t^o)$.

[Insert Table 3 about here]

Since the fluctuations of posterior regime probabilities are similar to those of the posterior model weight as shown in Figure 5, we focus on the model weight of the FA model in Panel (b). The posterior medians of the model weight are over 80% between 1982:Q1 and 1985:Q4. After that, the weight gradually decreases until 1988 and reaches 50%. From 1988:Q2, the beginning of the "Bubble" boom to 1994:Q2, the weight of the FA model bounces back to 80%. But the values decline deeply again and hover at a level as low as 20% between 1995:Q1 and 1997:Q2, before the FA model makes a recovery in 1997:Q3. In this way, the FA model outdoes the NK model for the booms and recessions of the pre-Bubble period and the Bubble period, but the boom of the post-Bubble period.

[Insert Figure 5 about here]

Next, let us consider what factors decide the size of the model weights of the FA model in terms of the financial accelerator mechanism. As described in Section 2.3, the loan rate and the real borrowing are additionally appended into the data and the loan rate directly connects with investment of the corporate sector in the FA model. On the other hand, since the NK model does not include the banking sector, the investment connects with the policy rate (or the interbank rate) instead of the loan rate. These aspects might have an influence on the forecasting of the six series. We focus on discrepancies between the loan rate and the policy rate.

In Figure 6 (a), the two representative series of corporate loan rates, say the long-term prime lending rate of long-term credit banks and the average contractual interest rate on bank loans for large-scale firms, and the policy rate, say the Bank of Japan (BOJ)'s secured overnight call rate, are depicted with a shadowed area indicating recessions. Panel (b) shows the two spreads between the loan rate and the policy rate. As the two figures show, there are three periods during which drastic monetary policies were implemented in Japan.

- The first period was 1985:Q4 when the monetary authorities implemented a policy to guide the yen higher following the Plaza Accord in the G-5 finance ministerial meeting[5] and hiked the policy rate rapidly. However, the rate reverted to the lower level once the policy had succeeded.

- The second period was between 1989:Q1 and 1991:Q1 when the BOJ had adopted a tight monetary policy to remedy the fever in the Bubble boom and raised the rate from around

[5]The Plaza Accord was an agreement between the governments of France, West Germany, Japan, the United States, and the United Kingdom, to depreciate the U.S. dollar in relation to the yen and Deutsche Mark by intervening in currency markets. The five governments signed the accord on September 22, 1985 at the Plaza Hotel in New York City.

4% to 8%. However, thanks to the Bubble boom, asset prices including securities and lands peaked and they resulted in a slow rise of the loan rate by shrinking the premium risk of corporate loans.

- The third period was between 1993:Q1 and 1995:Q4 when the Japanese economy was suffering from a long stagnation after the burst of the Bubble period and the BOJ had switched to an easy monetary policy such as driving down the policy rate gradually, reaching as low as 0.5% in 1995. However, the loan rate did not decline as much as the policy rate since leverage had not reduced in the corporate sector due to a serious bad loan problem in the banking sector.

In these three periods, the two loan rates failed to catch up with rapid fluctuations of the policy rate. As a result, the spread between them varied with big magnitude of fluctuations for those periods as shown in Panel (b). Furthermore, the three periods seem to be coincident with the timings of variations of the model weights as shown in Figure 5 (b). In the first period, the model weight of the FA model falls to nearly 50%, then it rises to about 80% in the second period, and again declines to around 20% in the third period. It might be thought that the changes of spread are closely related with the difference of forecasting performance between the two models. In the rest of this section, we analyze how the forecasting performance of the two DSGE models can be differentiated by specifying the three periods.

[Insert Figure 6 about here]

For the first period of drastic monetary policies, the spread became negative since the Plaza Accord had made the policy rate jump. As shown in Figure 4, the log predictive scores of all six series are likely to coincide between the two DSGE models after 1986:Q1 until 1988:Q1, before the beginning of the Bubble boom. This can also be seen from Figure 3, in which the log predictive scores of the individual series become close to each other in the four series: consumption, investment, interest rate and inflation, for this period. And the forecasting performance improves in the former two series of the NK model compared with the previous period, whereas it decays in the latter two series.

Next, for the Bubble period in the second period, the size of the under-estimation of the interest rate in the FA model is expanded compared with the NK model, and the predicted low interest rate also makes the estimation of inflation lower. On the other hand, the FA model has better performance in the three real series such as output, consumption and investment. As shown in the predictive distributions of Figure 2, the FA model might successfully grasp the big fluctuations of the Bubble period as the area of the distributions. In terms of wages, the NK model brings overestimation.

Finally, for the third period, forecasting of the interest rate changes from underestimation to overestimation by changing the attitude to monetary policy after the collapse of the Bubble boom. In particular, since the size of the overestimation of the rate in the FA model is much bigger than that of the NK model, the FA model makes predictions of output, consumption and inflation that are more seriously underestimated. In addition, the low interest rate policy makes the fluctuations of output, inflation and interest rate much narrower after 1995. These aspects become a disadvantage of the FA model, since the wider predictive densities of the FA model cover their smaller realized movements with too much surplus.

To sum up, although it appears paradoxical, the NK model without the financial friction performs better for the period with a bigger spread, which is thought to be compatible with the financial accelerator, such as in 1987:Q1–1988:Q1 and 1994:Q1–1997:Q1. In contrast, the FA model is predominant over its counterpart for the period generating negative spreads, in which the frictionless model seems to work well, since there is no EFP between the loan rates and the policy rate. In the light of the above consideration, we conclude that the observed spreads are not likely to be reflected in a timely way as the EFP of the corporate sector if the financial accelerator mechanism is regarded as working correctly. In particular, we observe a non-trivial time lag of the reduction of loan rates due to the rapid cutting of the policy rate. Accordingly, the FA model decays for the above two periods with big spreads despite importing the two additional categories of data.

5 Robust check by dynamic prediction pool

5.1 Robust check of model weights

Dynamic prediction pool method

IIn this section, we conduct a robust check of the previous section using another pooling method with a time-varying model weight, namely the dynamic prediction pooling method. This method was proposed by Del Negro et al. (2016). The MS model estimated in the previous section makes multiple constant model weights switch corresponding to multiple regimes, whereas the dynamic prediction pooling method has continuous values between zero and one as time-varying weights by incorporating the probit model. The model consists of the following two equations:

$$\begin{aligned} \lambda_t &= \Phi\left(x_t\right), \\ x_t &= (1-\rho)\mu - \rho x_{t-1} + \sqrt{1-\rho^2}\,\sigma\,\varepsilon_t,\ x_0 \sim N(\mu, \sigma^2), \end{aligned} \tag{5.1}$$

where $\lambda_t \in [0,1]$ is a model weight at period t , and x_t is a latent variable which is an input of a probit transformation and follows an AR(1) process. ρ is the autocorrelation coefficient. $\Phi(\cdot)$ is the cumulative density function of standard normal distribution, the disturbance term follows $\varepsilon_t \sim N(0,1)$, and x_0 is the initial value of x_t. The autocorrelation coefficient ρ captures how smoothly the weighting coefficient can change over time. The closer ρ is to one, the more slowly the model weights, λ_t , change. When $\rho = 1$, the model reduces to the case of static prediction pooling in Geweke and Amisano (2011) by taking $\lambda_t = \lambda$. When $\rho = 0$, it indicates that λ_t is serially independent and follows a random walk. μ is the mean of the unconditional distribution of the model weights, and σ is the variance of x_t, the large value of which makes the model weights fluctuate drastically. From these equations, we obtain conditional expectations and variances of the latent variables for the h-step-ahead forecast, x_{t+h},

$$
\begin{aligned}
E(x_{t+h}|x_t) &= \rho^h x_t + (1-\rho)\mu \sum_{i=0}^{h-1} \rho^i, \\
Var(x_{t+h}|x_t) &= (1-\rho^2)\sigma^2 \sum_{i=0}^{h-1} \rho^{2i},
\end{aligned}
\tag{5.2}
$$

where both conditional values converge to unconditional values, $E(x_{t+h}) = \mu$, and $Var(x_{t+h}) = \sigma^2$, when $h \to \infty$. And coefficient μ and variance σ^2 of the initial value of the latent variable are also equivalent to the unconditional values.

This study examines two versions of the above model following Del Negro et al. (2014). One is set as $\mu = 0$ and $\sigma^2 = 1$. $\mu = 0$ indicates that the unconditional expectation of the model weight is 0.5, $\Phi(0) = 0.5$, since the unconditional expectation of the weight is assumed to be equivalent between both models. And setting $\sigma^2 = 1$ comes from the assumption of the latent variable in the probit model. Accordingly, we only estimate a coefficient ρ in this version. The second sets three parameters freely and estimates them.

We obtain the dynamic prediction pooling of the log predictive score as

$$
LPS^{DP}(\lambda_{t+h},\, h) \equiv \sum_{t=1}^{T} \log \left[\lambda_{t+h}\, p\left(y^f_{t+h};\, Y^o_t,\, \mathcal{M}_1 \right) + (1-\lambda_{t+h})\, p\left(y^f_{t+h};\, Y^o_t,\, \mathcal{M}_2 \right) \right]. \tag{5.3}
$$

We adopt a particle filter for coping with a nonlinear model such as a probit model, and incorporate the nonlinear filtering method into a Bayesian estimation with the MCMC procedure, following Del Negro et al. (2016). We set the number of particles of the filter as 5,000 and calculate approximate values of the log predictive scores defined as Eq. (5.3). And we conduct 20,000 iterations as the MCMC procedure and discard the first 5,000 draws as burn-in [6].

Empirical Results

The estimation result in the version with only one flexible parameter is described in Panel (a) of Table 4 and that of the version with three flexible parameters is in Panel (b). The former has around 0.7 for the posterior mean of coefficient ρ, whereas the latter reduces to 0.6 by increasing the means of the other parameters to 1.66 and 0.5 for standard deviation σ and μ, respectively. Since ρ means persistence from the weight of the previous period, we can consider that the current weights are not so strongly influenced by the previous weight. From the posterior mean of σ, the uncertainty of the weight might be 1.66 times the second version. And the unconditional model weight of the FA model is nearly 70% since $\Phi(0.5) = 0.691$.

[Insert Table 4 about here]

[6] We code the algorithm of a particle filter following Johannes and Polson (2009). And the joint use of the MCMC procedure and the particle filter in our study also follows Andrieu et al. (2010).

[Insert Figure 7 about here]

The estimated time-varying model weights of the FA model are depicted in Figure 7. Panel (a) shows the weights of one flexible parameter, whereas Panel (b) shows those of three flexible parameters. The black solid lines represent the posterior means and the black dotted lines indicate 68% credible intervals. As shown in both graphs, like the MS model in Figure 5, for the boom and the recession of the Bubble period, the posterior means of the weights reach nearly 70% and 90% for the one parameter version and the three parameter version, respectively.

In contrast with the MS model, both versions of the dynamic model significantly make the weight of the FA model go down in the period of 1997:Q2 when the consumption tax was jacked up from 3% to 5%. In the version with one parameter, the weight dropped from nearly 50% to less than 30%, while the other version rapidly reduced the weight to 10%. And it is noteworthy that the posterior means of the model weights of the MS model drawn in Figure 5 are located between those of the two versions of the dynamic model. However, in terms of credible intervals, these two versions are in contrast with the MS model. The areas of the latter model are much narrower as a result of taking advantage of the characteristics of the discrete Markov process. Although we observe discrepancies between the two pooling methods, we successfully conduct a robust check to grasp the correlation between the three drastic monetary policy changes and the fluctuation of the model weights for the FA model.

5.2 Evaluation and validity of forecasting by pooling methods

Finally, we evaluate the forecasting performance of all the pool methods and the DSGE models. The cumulative log predictive scores of the methods are calculated from $\frac{1}{H}\sum_{h=1}^{H} LPS(\lambda_t, h)$ described as Eq.(3.2), Eq. (3.4) and Eq. (5.3), where we set $H = 4$ and posterior means of the time-varying model weights are adopted as the weight λ_t. Table 5 represents the cumulative log predictive scores of the four methods including the pooling method with the constant weight originally proposed by Geweke and Amisano (2011). As this table shows, all four pooling methods dominate the log predictive scores of both of the single DSGE models. In particular, the results for the three methods with time-varying model weights are notable, and the dynamic pooling model with three flexible parameters records the best performance.

[Insert Table 5 about here]

Using the model weight of the two pooling methods calculated from the log predictive score for all six variables, we calculate the log predictive scores for the six individual series and describe those values in the third and fourth rows of Table 2 (a) through (d). The bold numbers show the best performance out of the four models: NK model, FA model, MS model, dynamic model with three flexible parameters. As mentioned before, NK beats the FA model in terms of GDP, inflation, and interest rate for all sample periods and FA dominates for wage. For these series, the pooling methods combining both single DSGE models cannot improve the log predictive scores. On the other hand, for consumption and investment, we obtain an improvement of the forecasting performance for all three periods using the model weight of the pooling method calculated from the whole series of six. In this way, we manage to improve the predictions for several series by combining multiple DSGE models although some conditions are required. We need to further develop pooling methods with time-varying weights for predicting more accurately and for expanding to more multiple series.

6 Conclusion

Using the Markov switching prediction pool method by Waggoner and Zha (2012) in terms of density forecasts, we consider the time-varying forecasting performance of a DSGE model incorporating a financial accelerator à la Bernanke et al. (1999) with the frictionless model, by focusing on periods of financial crisis including the so-called "Bubble period" and the "Lost decade" in Japan.

One of the features in estimation with DSGE models is imposing restrictions on the comovements between macroeconomic variables from the point of view of the DSGE model. The higher forecast performance of the model with the financial friction compared to the model without friction reflects the presence of comovements generated by the friction in the data during the period. It is suggested that the causality of the financial accelerator exists with a higher probability than that of the frictionless model. And we estimated when and the extent to which the comovements generated by both DSGE models change in terms of time series through changes of the time-varying model weights, realizing the optimal combination of density forecasts. These gave us the clues to which conditions in economic situations contribute to changes of the comovements. Furthermore, we conducted a robust check to examine whether a similar dynamic change of the weight is observed when using the dynamic prediction pooling method by Del Negro et al. (2016) in this paper.

This paper showed the following findings. For the overall periods from 1981:Q1 to 1998:Q4, the model with the financial friction is predominant over the frictionless benchmark model in terms of density forecasts. The difference between them is likely to come from fluctuation of the spread between the loan rate and policy interest rate. In a period with a small change of the spread, the financial accelerator mechanism contributes to improve the prediction. However, when a drastic monetary policy was implemented, the loan rates, which did not react to the big change of the policy rates and shifted the spread with a large step, weakened the forecasting performance of the model with the friction. In particular, the frictionless model shows superior performance for the period from 1993 to 1995, since the spread realized with a big range despite the boom seems to be contrary to the spread generated from the financial friction. These empirical results suggested that real spreads do not provide a timely reflection of the change of the external financial premium generated between bankers and the corporate sector, and that there is a non-trivial time lag between them. The robust check also supported these results.

A Online Appendix

A.1 NK model

There are fourteen log-linearized equilibrium condtions in the benchmark DSGE model: Eq.(A.1) through Eq.(A.14). The hatted letters indicate log-deviations from steady-state values after detrending with the level of composite technology Z_t^*.

a. Households

$$\begin{aligned}\hat{\lambda}_t &= -\frac{1}{1-\theta\pi/r_n}\left\{\frac{\sigma}{1-\theta/z^*}\left(c_t - \frac{\theta}{z^*}(\hat{c}_{t-} - z_t^*)\right) - z_t^b\right\} \\ &+ \frac{\theta\pi/r_n}{1-\theta\pi/r_n}\left\{\frac{\sigma}{1-\theta/z^*}\left(E_t\hat{c}_{t+1} + E_t z_{t+1}^* - \theta/z^*\hat{c}_t\right) - E_t z_{t+1}^b\right\},\end{aligned} \tag{A.1}$$

$$\hat{\lambda}_t = E_t\hat{\lambda}_{t+1} - \sigma E_t\, z_{t+1}^* + \hat{r}_t^n - E_t\hat{\pi}_{t+1}, \tag{A.2}$$

i) Workers:

$$\begin{aligned}\hat{w}_t &= \hat{w}_{t-1} - \hat{\pi}_t + \gamma_w\hat{\pi}_{t-1} - z_t^* + \frac{z^*\pi}{r_n}\left(E_t\hat{w}_{t+1} - \hat{w}_t + E_t\hat{\pi}_{t+1} - \gamma_w\hat{\pi}_t + E_t z_{t+1}^*\right) \\ &+ \frac{(1-\xi_w)(1-\xi_w z^*\pi/r_n)}{\xi_w(1+\chi(1+\lambda_w)/\lambda_w)}\left(\chi\hat{l}_t - \hat{\lambda}_t - \hat{w}_t + z_t^b\right) + z_t^w,\end{aligned} \tag{A.3}$$

ii) Entrepreneurs:

$$\hat{u}_t = \tau(\hat{r}_t^k - \hat{q}_t), \tag{A.4}$$

$$\hat{\chi}_t = \left(1 - \frac{1-\delta}{r_n\psi}\right)\hat{r}_t^k + \frac{1-\delta}{r_n\psi}\hat{q}_t - \hat{q}_{t-1} - z_t^\psi, \tag{A.5}$$

$$E_t\hat{\chi}_{t+1} = \hat{r}_t^n - E_t\hat{\pi}_{t+1}, \tag{A.6}$$

b. Firms

$$\hat{y}_t = (1+\phi)\left\{(1-\alpha)\,\hat{l}_t + \alpha(\hat{u}_t + \hat{k}_{t-1} - z_t^* - z_t^\psi)\right\}, \tag{A.7}$$

$$0 = \hat{w}_t + \hat{l}_t - (\hat{r_t^k} + \hat{u}_t + \hat{k}_{t-1} - z_t^* - z_t^\psi), \tag{A.8}$$

$$\hat{mc}_t = (1-\alpha)\hat{w}_t + \alpha\hat{r}_t^k. \tag{A.9}$$

$$\hat{\pi}_t = \gamma_p \hat{\pi}_{t-1} + \frac{z^*\pi}{r_n}(E_t\hat{\pi}_{t+1} - \gamma_p\hat{\pi}_t) + \frac{(1-\xi_p)(1-\xi_p\frac{z^*\pi}{r_n})}{\xi^p}\hat{mc}_t + z_t^p, \tag{A.10}$$

$$\hat{k}_t = \frac{(1-\delta-r_n\psi/\pi)}{z^*\psi}\hat{u}_t + \frac{1-\delta}{z^*\psi}(\hat{k}_{t-1} - z_t^* - z_t^\psi) + \left(1 - \frac{1-\delta}{z^*\psi}\right)\hat{i}_t, \tag{A.11}$$

$$\hat{q}_t = \frac{1}{\zeta}(\hat{i}_t - \hat{i}_{t-1} + z_t^* + z_t^\psi) - \frac{z^*\pi}{\zeta r_n}(E_t\hat{i}_{t+1} - \hat{i}_t + E_t z_{t+1}^* + E_t z_{t+1}^\psi) - z_t^\nu + z_t^i, \tag{A.12}$$

C. Miscellaneous

$$\hat{r}_t^n = \phi_r \hat{r}_{t-1}^n + (1-\phi_r)\left\{\frac{\phi_\pi}{4}\Sigma_{j=0}^3\hat{\pi}_{t-j} + \phi_y\hat{y}_t\right\} + z_t^r, \tag{A.13}$$

$$\hat{y}_t = \frac{c}{y}\,\hat{c}_t + \frac{i}{y}\,\hat{i}_t + \frac{g}{y}\,z_t^g, \tag{A.14}$$

A.2 FA model

There are seventeen log-linearized equilibrium condtions in the DSGE model with financial accelerator, consisting of both Eq.(A.1) through Eq.(A.5) and Eq.(A.7) through Eq.(A.14), which are a common part of both models, and an additional part formd from Eq.(A.15) through Eq.(A.18) .

$$\hat{r}_t^E = \hat{r}_t^n + \mu_E(\hat{q}_t + \hat{k}_t - \hat{n}_t) + z_t^\mu, \tag{A.15}$$

$$\begin{aligned} \frac{z^*}{\eta\, r^E}\,\hat{n}_t \quad &= \quad \frac{1+\lambda_i}{n/k}\left[(1 - \frac{1-\delta}{r^E\psi})\hat{r}_t^k + \frac{1-\delta}{r^E\psi}\hat{q}_t - \hat{q}_{t-1} - z_t^\psi\right] \\ &- \quad \left(\frac{1+\lambda_i}{n/k} - 1\right) E_{t-1}\hat{r}_t^E + \hat{n}_{t-1} - z_t^* + z_t^\eta, \end{aligned} \tag{A.16}$$

$$\hat{b}_t = \frac{1+\lambda_i}{1+\lambda_i - n/k}(\hat{q}_t + \hat{k}_t) + \left(1 - \frac{1+\lambda_i}{1+\lambda_i - n/k}\right)\hat{n}_t, \tag{A.17}$$

$$E_t\hat{\chi}_{t+1} = \hat{r}_t^E - E_t\hat{\pi}_{t+1}, \tag{A.18}$$

References

[1] Adolfson, M., Linde, J., & Villani, M., (2007) "Forecasting Performance of an open economy DSGE model," Econometric Reviews, 26(2-4), 289-328.

[2] Altig, D., Christiano, L. J., Eichenbaum, & M., Linde, J. (2011) "Firm-specific capital, nominal rigidities and the business cycle ," Review of Economic Dynamics, 14(2), 225-247.

[3] Albert, J.H. & Chib, S. (1993) Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts. Journal of Business and Economic Statistics 11 (1), 1-15.

[4] Amisano, G., & Giacomini, R. (2007) "Comparing density forecasts via weighted likelihood ratio test," Journal of Business and Economic Statistics, 25, 177-190.

[5] Andrieu, C., Doucet, A., & Holenstein, R. (2010) "Particle Markov Chain Monte Carlo Methods," Journal of the Royal Statistical Society Series B, 72, 269-342.

[6] 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.

[7] Bates, J. & Granger, C. W. (1969) "The Combination of Forecasts," Operations Reserach, 20, 451-468.

[8] Christensen, I., & Ali Dib (2008) The financial accelerator in an estimated New Keyesian model, Review of Economic Dynamics, 11, 155-178.

[9] Christiano, L.J., Motto, R., & Rostagno, M.V., (2014) "Risk Shocks," American Economic Review, 104(1), 27-65.

[10] Christiano, L.J., Tranbandt, M., & Walentin, K. (2010) "DSGE models for Monetary Policy Analysis," Handbook of Monetary Economics.

[11] Christiano, L.J., Tranbandt, M., & Walentin, K. (2011) "Introducing Financial Frictions and Unemployment into a Small Open Economy Model," Journal of Economic and Dynamic Control, 35(12) 1999-2041.

[12] Del Negro, M., Hasegawa, R.B. & Schorfheide, F. (2016) Dynamic Prediction Pools: an Investigation of Financial Frictions and Forecasting Performance. Journal of Econometrics, Vol.192 (2), p.391-405,

[13] Del Negro, M. & Schorfheide, F. (2004) "Priors from General Equilibrium Models for VARs, " International Economic Reviews, 45(2), 643-673.

[14] Del Negro, M. & Schorfheide, F. (2013) "DSGE model-based forecasting," In G.Elliott and A. Timmermann (Eds.) Handbook of Economic Forecasting, Vol.2, 57-140, Elsevier.

[15] Edge, Rochelle M. & Refet S. Gurkaynak (2010) How Useful are Estimated DSGE Model Forecasts? Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 41(2 (Fall)), pages 209-259.

[16] Edge, R. M., Kiley, M. T., & Laforte, J.-P. (2010) "A Comparison of Forecast Performance between Federal Reserve Staff Forecast, Simple Reduced-Form Models, and a DSGE Model, " Journal of Applied Econometrics, 25(4), 720-754.

[17] Elliott, G., & Timmermann, A., (2016) *Economic Forecasting*, Princeton University Press.

[18] Geweke J. (1992) Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments.In Bayesian Statistics 4, ed. J. M. Bernardo, J. Berger, A. P. Dawid, and A. F. M. Smith (Oxford: Oxford University Press) pp. 169-193.

[19] Geweke J. (2010) Complete and Incomplete Econometric Models. The Econometric and Tinbergen Institutes Lectures, Princeton University Press.

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

[21] Geweke, J. & Amisano, G. (2012) Prediction with Misspecified Models. American Economic Review: Papers & Proceedings 2012, 102 (3), 482-486.

[22] Gilchrist, S., & Zakrajsek, E., (2012) "Credit Spreads and Business Cycle Fluctuations," American Economic Review, 102(4), 1692-1720.

[23] Hayashi, F., & Prescott, E. C., (2002) "The 1990s in Japan: the Lost decade," Review of Economic Dynamics, 5(1), 206-235.

[24] Herbst, E. & Schorfheide, F. (2012) "Evaluating DSGE model forecast of Comovement," Journal of Econometrics, 171(2), 152-166.

[25] Hoeting, J., Madigan, D., Raftey, A. & Volinsky, C. (1999) Bayesian Model Averaging. Statistical Science 14, 382-401.

[26] Johanness, M. & Polson, N. (2009) Particle Filtering, Handbook of Financial Time Series, edited by T.G. Anderson et al. pp.1015-1029. Springer-Verlag Berlin Heidelberg.

[27] Kaihatsu, S., & Kurozumi, T., 2010. Sources of Business Fluctuations: Finanical or Technology Shocks? Bank of Japan Working Paper Series, No 10-E-12.

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

[29] 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.

[30] Kolasa, M., & Rubaszek, M. 2015. Forecasting using DSGE models with financial frictions, *International Journal of Forecasting*, 31, 1-19.

[31] Kolasa, M., Rubaszek, M., & Skirzypczynski, P. 2012, Putting the New Keynesian DSGE model to real-time forecasting test, *Journal of Money, Credit and Banking*, 44(7), 1301-1324.

[32] Rafty, A.E., Madigan, D. & Hoeting, J.A. (1997) Bayesian Model Averaging for Linear Regression Models. Journal of the American Statistical Association 92, 179-191.

[33] Smet, F. & Wouters, R. (2007) "Shocks and Frictions in US Business Cycles: a Bayesian DSGE approach," American Economic Review, 97(3), 586-606.

[34] Sugo, T. & Ueda, K. (2008) "Estimating a Dynamic Stochastic General Equilibrium Model for Japan," *Journal of Japanese and International Economies*, **22**, 476-502.

[35] Timmermann, A., (2006) "Forecast combination ," in: Elliott, G., Granger , C.W.J., Timmermann, A., (Eds.) Handbook of Economic Forecasting, chapter 4, 135-196.

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

B Tables

Table 1: Means of Forecasting Errors of Observations

			Output	Consumption	Investment	Real Wage	Inflation	Nominal Rate
Booms	Full Sample	NK	-0.276	-0.265	-0.979	0.172	-0.047	0.033
		FA	-0.326	-0.151	-1.278	-0.010	-0.057	0.040
	Pre Bubble	NK	-0.248	-0.566	-0.156	-0.125	-0.038	0.041
		FA	0.198	-0.093	0.263	-0.013	-0.051	0.034
	Bubble	NK	-0.236	-0.087	-1.785	0.481	-0.064	-0.090
		FA	-0.568	-0.133	-2.591	0.073	-0.064	-0.107
	Post Bubble	NK	-0.386	-0.262	-0.862	0.082	-0.062	0.153
		FA	-0.581	-0.298	-1.262	-0.112	-0.085	0.202
Recessions	Full Sample	NK	0.135	0.237	0.708	0.280	0.018	0.187
		FA	0.041	0.182	0.577	0.032	-0.003	0.184
	Pre Bubble	NK	-0.019	-0.309	0.419	0.011	0.245	0.124
		FA	0.185	-0.093	0.552	-0.107	0.231	0.080
	Bubble	NK	0.062	0.386	0.106	0.414	-0.064	0.195
		FA	-0.191	0.173	-0.077	0.055	-0.128	0.135
	Post Bubble	NK	0.382	0.472	1.901	0.299	-0.048	0.228
		FA	0.283	0.433	1.628	0.113	-0.006	0.351

Notes: Full Smaple: 1981:Q1-1998:Q4. Boom and recession in the pre-Bubble: 1981:Q1-1985:Q2 and 1985:Q3-1986:Q2. Boom and recession in the Bubble period: 1986:Q3-1991:Q1 and 1991:Q2-1993:Q2. And boom and recession of the post-Bubble: 1993:Q3-1997:Q2 and 1997:Q3-1998:Q4.

Table 2: Log Scores of Single and the Whole of Observations

(a) Full Smaple: 1981:Q1-1998:Q4

	Log Score						
Models	Whole	Output	Cons	Inv	Real Wage	Inflation	Nominal Rate
NK	-522.45	**-95.49**	-114.16	-185.58	-65.36	**-20.43**	**-21.43**
FA	**-492.66**	-102.16	**-106.57**	**-175.42**	**-51.66**	-28.56	-28.10
MS	-482.08	-98.51	-106.34	-174.90	-52.88	-25.19	-25.73
D3	**-477.16**	-97.84	**-104.32**	**-172.30**	-53.16	-25.16	-25.16

(b) Pre Bubble Cycle: 1981:Q1-1986:Q2

	Log Score						
Models	Whole	Output	Cons	Inv	Real Wage	Inflation	Nominal Rate
NK	-135.89	**-20.79**	-30.35	**-42.45**	-15.15	**-6.97**	**-5.14**
FA	**-126.42**	-23.81	**-26.15**	-43.70	**-13.85**	-8.73	-6.76
MS	-125.95	-23.14	-25.79	-43.22	-13.92	-8.46	-6.52
D3	**-124.33**	-22.14	**-24.49**	-42.77	-13.99	-8.10	-6.22

(c) Bubble Boom and Recession: 1986:Q3-1993:Q2

	Log Score						
Models	Whole	Output	Cons	Inv	Real Wage	Inflation	Nominal Rate
NK	-231.29	**-46.81**	-51.51	-82.21	-33.89	**-7.57**	**-9.08**
FA	**-207.87**	-47.49	**-48.56**	**-75.35**	**-23.33**	-11.40	-11.17
MS	-205.33	-47.12	-48.62	-75.34	-23.97	-10.76	-10.78
D3	**-202.18**	-46.84	**-48.06**	**-74.90**	-24.07	-10.24	-10.35

(d) Post Bubble Cycle: 1993:Q3-1998:Q4

	Log Score						
Models	Whole	Output	Cons	Inv	Real Wage	Inflation	Nominal Rate
NK	**-149.92**	**-27.03**	-31.00	-59.42	-15.69	**-5.38**	**-6.84**
FA	-152.33	-29.80	**-30.49**	**-54.51**	**-14.07**	-7.83	-9.74
MS	**-144.95**	-27.24	-30.59	-54.57	-14.55	-5.93	-7.78
D3	-145.11	-27.94	**-30.45**	**-53.00**	-14.55	-6.27	-8.00

Notes:

1. FA model and NK model stand for the DSGE models with financial friction and without financial friction, respectively.
2. MS and D3 stand for Markov-swiching pool and dynamic pool with three flexible parameters, respectively.
3. A boldface type of the the third and forth rows represents the best value out of two DSGE models, while that of the fifth and sixth rows represent the best value out of all four methods in terms of each observation of the colums.

Table 3: Markov Switching Prediction Pool

Parameter	Prior	Mean	Std. Dev.	90 % Band	Inefficiency
λ_1	$G(0.05, 0.1)$	0.049	0.080	[0.000 0.126]	261.73
λ_2	$G(0.95, 0.1)\, I(\lambda_2 > \lambda_1)$	0.884	0.089	[0.773 0.984]	572.56
q_{11}	$Beta(1, 9)$	0.865	0.094	[0.736 0.968]	580.43
q_{22}	$Beta(1, 9)$	0.909	0.073	[0.810 0.983]	593.56

Notes:

1. For estimation of MS prediction pool method, we conduct 100,000 MCMC iterations, the first 40,000 iterations are discarded.

2. In prior, G, $Beta$ stand for gamma and beta distributions, respectively. $I(\cdot)$ represent an indicator function which returns one if a condition of inside are hold, otherwise zero.

Table 4: Dynamic Prediction Pool

(i) One flexible parameter version

Parameter	Prior	Mean	Std. Dev.	90 % band	Inefficiency
ρ	$U(0,1)$	0.700	0.227	[0.288 0.984]	102.085

(ii) Three flexible parameters version

Parameter	Prior	Mean	Std. Dev.	90 % band	Inefficiency
ρ	$Beta(5,5)$	0.590	0.094	[0.433 0.703]	4.875
μ	N(0.5,1)	0.500	0.011	[0.480 0.514]	9.295
σ	IG(1,10)	1.657	0.137	[1.380 1.845]	13.336

Notes:

1. For estimation of Dynamic prediction pool method, we conduct 20,000 MCMC iterations with 5,000 particles, the first 5,000 iterations are discarded.
2. In prior, U, $Beta$, N and IG stand for uniform, beta, normal and inverse gamma distributions, respectively.

Table 5: Cumulative Log Scores

Component Models		Model Pooling	
Model	Log Score	Methods	Log Score
NK	-522.45	Static Pool	-489.18
FA	-492.66	MS Pool	-482.08
		Dynamic Pool (1)	-484.68
		Dynamic Pool (3)	-477.16

Notes:

1. The prediction scores $p(y_{t+h}; Y_t, M_i)$ for the NK model: M_1, and FA model: M_2, are obtained by simulation using the MCMC draws of the posterior model parameters.
2. Dynamic pool (1) and (3) stand for those with one and three flexible parameters, respectively.
3. Log scores of Static Pool, MS Pool and Dynamic Pool are calculated from Eq.(3.2), Eq.(3.4), and Eq.(5.3), respectively.

Table A1. Priors of the Parameters in both the Models

Parameter		Dist. Type	Benchmark Model Mean	Benchmark Model S.D.	Financial Friction Mean	Financial Friction S.D.
σ	Risk aversion	G	1.5	0.375	1.5	0.375
θ	Habit persistence	B	0.7	0.1	0.7	0.1
χ	Inverse of elesiticity of labor supply	G	2.0	0.75	2.0	0.75
$1/\zeta$	Elasticity of investment adjustment cost	G	4.0	1.5	4.0	1.5
τ	Inverse of elasticity of utilization rate adjustment cost	G	1.0	0.2	1.0	0.2
ϕ/y	Fixed production cost-output ratio	G	0.25	0.125	0.25	0.125
γ_w	Wage indexation	B	0.5	0.15	0.5	0.15
ξ_w	Wage stickiness	B	0.5	0.1	0.5	0.1
γ_p	Intermediate-goods price indexation	B	0.5	0.15	0.5	0.15
ξ_p	Inermediate-goods price stickiness	B	0.5	0.1	0.5	0.1
ϕ_r	Monetary policy rate smoothing	B	0.75	0.1	0.75	0.1
ϕ_π	Monetary policy response to inflation	G	1.5	0.25	1.5	0.25
ϕ_y	Monetary policy response to output	G	0.125	0.05	0.125	0.05
z^*	Steady-state rate of balanced growth	G	0.36	0.1	0.36	0.1
ψ	Steady-state rate of IS technological change	G	0.32	0.1	0.32	0.1
η	Entrepreneur survival probability	B	NA	NA	0.973	0.02
n/k	Steady-state net worth-capital ratio	B	NA	NA	0.5	0.07
μ_E	Elasticity of EF premium	G	NA	NA	0.07	0.02
r^E	Steady-state loan rate	G	NA	NA	1.19	0.05
ρ_b	Persistence of preference shock	B	0.5	0.2	0.5	0.2
ρ_g	Persistence of exogenous demand shock	B	0.5	0.2	0.5	0.2
ρ_w	Persistence of wage shock	B	0.5	0.2	0.5	0.2
ρ_p	Persistence of intermediate-goods price markup shock	B	0.5	0.2	0.5	0.2
ρ_i	Persistence of investment-goods price markup shock	B	0.5	0.2	0.5	0.2
ρ_r	Persistence of monetary policy shock	B	0.5	0.2	0.5	0.2
ρ_z	Persistence of neutral technology shock	B	0.5	0.2	0.5	0.2
ρ_ψ	Persistence of IS technology shock	B	0.5	0.2	0.5	0.2
ρ_ν	Persistence of MEI shock	B	0.5	0.2	0.5	0.2
ρ_μ	Persistence of EFP shock	B	NA	NA	0.5	0.2
ρ_η	Persistence of net worth shock	B	NA	NA	0.5	0.2
σ_b	S.D. of preference shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_g	S.D. of exogenous demand shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_w	S.D. of wage shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_p	S.D. of intermediate-goods price markup shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_i	S.D. of investment-goods price markup shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_r	S.D. of monetary policy shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_z	S.D. of neutral technology shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_ψ	S.D. of IS technology shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_ν	S.D. of MEI shock innovation	IG	0.5	Inf.	0.5	Inf.
σ_μ	S.D. of EFP shock innovation	IG	NA	NA	0.5	Inf.
σ_η	S.D. of net worth shock innovation	IG	NA	NA	0.5	Inf.

Notes: Regarding the type of prior distributions, B,G and IG stand for Beta, Gamma, and Inverse Gamma distributions, respectively.

Table A2. Posterior Estimates of the Parameters in the Benchmark model

Parameter	Mean	Stdev	[95% Band]		Geweke	Inef.
σ	1.3384	0.2254	[0.994	1.741]	0.169	175.166
θ	0.3662	0.0576	[0.277	0.469]	0.636	162.026
χ	3.5377	0.724	[2.326	4.716]	0.028	191.738
$1/\zeta$	0.5825	0.1492	[0.365	0.87]	0.24	217.022
τ	1.0423	0.1823	[0.758	1.366]	0.539	141.557
ϕ/y	0.3642	0.1114	[0.195	0.558]	0	277.987
γ_w	0.4511	0.134	[0.21	0.673]	0.124	261.432
ξ_w	0.3101	0.0809	[0.195	0.466]	0	269.981
γ_p	0.6807	0.1217	[0.462	0.856]	0.022	140.977
ξ_p	0.4811	0.0281	[0.435	0.527]	0.815	94.937
ϕ_r	0.5264	0.0803	[0.385	0.643]	0.002	175.321
ϕ_π	2.1025	0.22	[1.762	2.46]	0.671	98.36
ϕ_y	0.1048	0.0461	[0.043	0.192]	0.418	323.954
z^*	0.3939	0.0929	[0.25	0.557]	0.014	251.441
ψ	0.4277	0.0766	[0.306	0.564]	0.04	161.517
η	NA	NA	[NA	NA]	NA	NA
n/k	NA	NA	[NA	NA]	NA	NA
μ_E	NA	NA	[NA	NA]	NA	NA
r^E	NA	NA	[NA	NA]	NA	NA
ρ_b	0.8256	0.067	[0.688	0.895]	0.003	98.55
ρ_g	0.9159	0.0246	[0.872	0.952]	0.001	27.284
ρ_w	0.5906	0.1609	[0.326	0.842]	0	343.84
ρ_p	0.9495	0.0275	[0.899	0.983]	0	52.996
ρ_i	0.9112	0.0424	[[0.833	0.967]	0	80.725
ρ_r	0.3991	0.1234	[0.203	0.589]	0.255	279.234
ρ_z	0.0633	0.0304	[0.024	0.131]	0.23	304.232
ρ_ψ	0.1456	0.0635	[0.041	0.252]	0	301.843
ρ_ν	0.8252	0.1005	[0.621	0.95]	0.874	89.026
ρ_μ	NA	NA	[NA	NA]	NA	NA
ρ_η	NA	NA	[NA	NA]	NA	NA
σ_b	1.9911	0.3553	[1.496	2.61]	0.109	195.043
σ_g	0.9987	0.0817	[0.874	1.143]	0.058	43.062
σ_w	0.754	0.1027	[0.606	0.945]	0.006	141.629
σ_p	0.5865	0.0783	[0.475	0.734]	0.58	121.738
σ_i	1.8523	0.2901	[1.422	2.373]	0.418	180.29
σ_r	0.535	0.0543	[0.455	0.633]	0.64	52.711
σ_z	1.6481	0.1646	[1.406	1.924]	0.283	91.421
σ_ψ	0.5684	0.0486	[0.506	0.66]	0.109	62.768
σ_ν	0.5148	0.0142	[0.501	0.543]	0.099	5.916
σ_μ	NA	NA	[NA	NA]	NA	NA
σ_η	NA	NA	[NA	NA]	NA	NA

Notes: To form posterior distributions of the parameters, 300,000 iterations are implemented in MCMC. After the first 150,000 draws are discarded, the remaining draws are sampled as the posterior estimates. Mean and Stdev stand for the posterior mean and standard deviation, respectively. Geweke and Inef. refer to the p-value associated with the convergence diagnostic of Geweke (1992) and the simulation inefficient statistics of Kim, Shephard and Chib (1998).

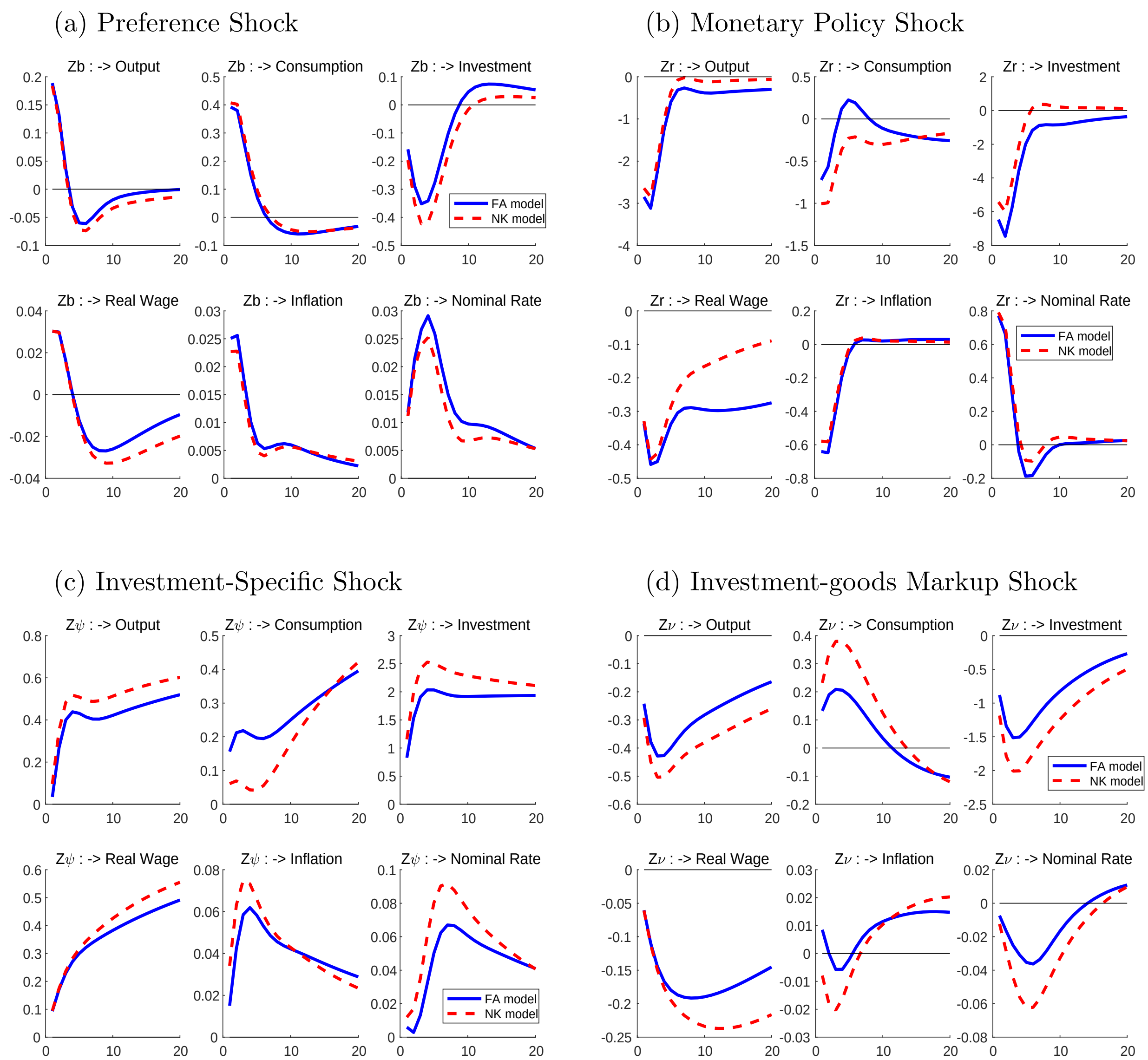
Table A3. Posterior Estimates of the Parameters in the Financial Friction model

Parameter	Mean	Stdev	[95% Band]		Geweke	Inef.
σ	0.9114	0.1005	[0.744	1.085]	0.007	336.418
θ	0.4907	0.0604	[0.393	0.582]	0	340.524
χ	1.8996	0.3357	[1.128	2.315]	0.77	344.756
$1/\zeta$	0.4113	0.0617	[0.31	0.518]	0.724	342.076
τ	1.797	0.1373	[1.532	1.974]	0.003	287.338
ϕ/y	0.9585	0.035	[0.886	0.997]	0.746	163.002
γ_w	0.4309	0.0593	[0.294	0.512]	0.024	323.808
ξ_w	0.5149	0.0415	[0.452	0.592]	0.202	280.793
γ_p	0.6908	0.1116	[0.507	0.872]	0	337.899
ξ_p	0.3749	0.0484	[0.309	0.471]	0.378	335.937
ϕ_r	0.4527	0.0478	[0.359	0.528]	0.604	309.306
ϕ_π	1.7541	0.1114	[1.57	1.933]	0.432	254.574
ϕ_y	0.1026	0.0155	[0.079	0.127]	0	339.599
z^*	0.2892	0.0306	[0.245	0.345]	0.03	314.2
ψ	0.407	0.0664	[0.291	0.519]	0	347.066
η	0.984	0.0094	[0.966	0.997]	0.191	54.085
n/k	0.5827	0.0411	[0.513	0.654]	0.02	297.98
μ_E	0.0331	0.0044	[0.026	0.04]	0	318.722
r^E	1.1845	0.0482	[1.107	1.264]	0.965	161.831
ρ_b	0.3442	0.0955	[0.178	0.48]	0.049	322.948
ρ_g	0.9245	0.0229	[0.883	0.958]	0.931	84.04
ρ_w	0.1625	0.0328	[0.108	0.217]	0	352.195
ρ_p	0.377	0.0803	[0.254	0.514]	0.002	337.322
ρ_i	0.9765	0.0031	[0.97	0.98]	0	31.383
ρ_r	0.2005	0.1144	[0.056	0.421]	0.081	361.944
ρ_z	0.2231	0.0194	[0.187	0.251]	0.498	311.019
ρ_ψ	0.096	0.0323	[0.037	0.141]	0	361.778
ρ_ν	0.8219	0.0475	[0.732	0.888]	0.176	215.383
ρ_μ	0.9066	0.0217	[0.871	0.941]	0.956	87.187
ρ_η	0.7047	0.0599	[0.598	0.799]	0.057	292.095
σ_b	1.4693	0.153	[1.245	1.763]	0	327.772
σ_g	1.0141	0.0969	[0.872	1.196]	0.697	308.772
σ_w	0.6174	0.0546	[0.531	0.722]	0.272	277.323
σ_p	0.5845	0.0656	[0.479	0.71]	0.876	326.688
σ_i	2.1229	0.1952	[1.793	2.442]	0.265	313.628
σ_r	0.6136	0.0462	[0.539	0.692]	0.005	278.955
σ_z	0.6318	0.0619	[0.537	0.741]	0.037	304.696
σ_ψ	0.5921	0.0479	[0.519	0.673]	0	298.752
σ_ν	0.5174	0.0152	[0.501	0.548]	0.209	110.049
σ_μ	0.4135	0.0336	[0.359	0.473]	0.001	282.062
σ_η	1.011	0.1471	[0.781	1.276]	0	342.149

Notes: To form posterior distributions of the parameters, 300,000 iterations are implemented in MCMC. After the first 150,000 draws are discarded, the remaining draws are sampled as the posterior estimates. Mean and Stdev stand for the posterior mean and standard deviation, respectively. Geweke and Inef. refer to the p-value associated with the convergence diagnostic of Geweke (1992) and the simulation inefficient statistics of Kim, Shephard and Chib (1998).

C Figures

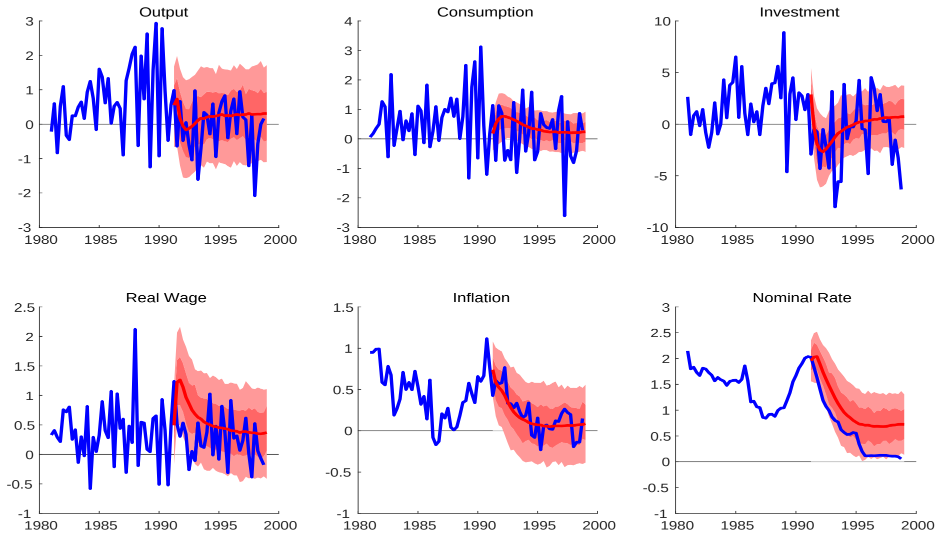
Figure 1: Impulse Response Functions



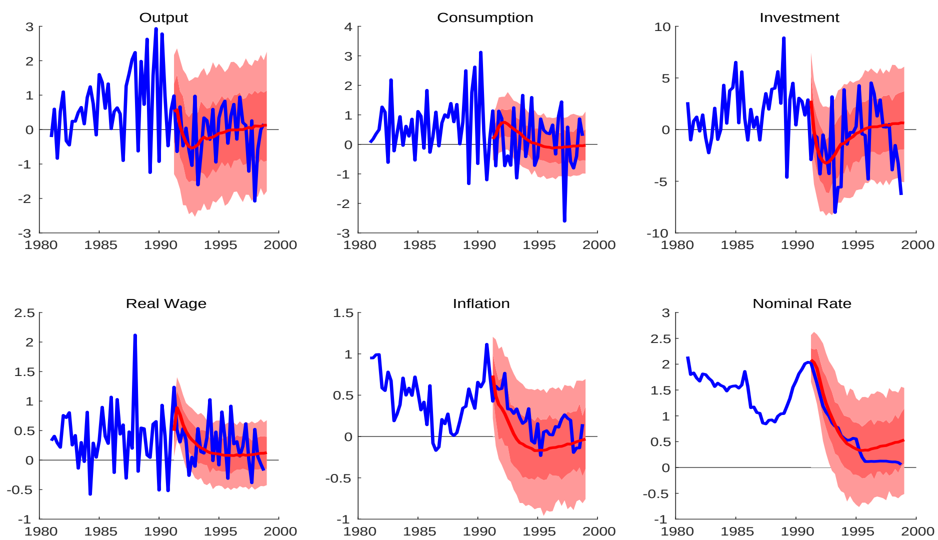
Notes: FA model and NK model stand for the DSGE models with financial friction and without financial friction, respectively. The IRFs of both the DSGE models are calculated from the structural parameters whose values are prior mean shown in Table A1, except the parameter μ_E whose value is 5 times bigger than that of the prior.

Figure 2: Posterior Predictive Disribulions in the Bubble Period

(a) NK model: Forecast as of 90:Q4

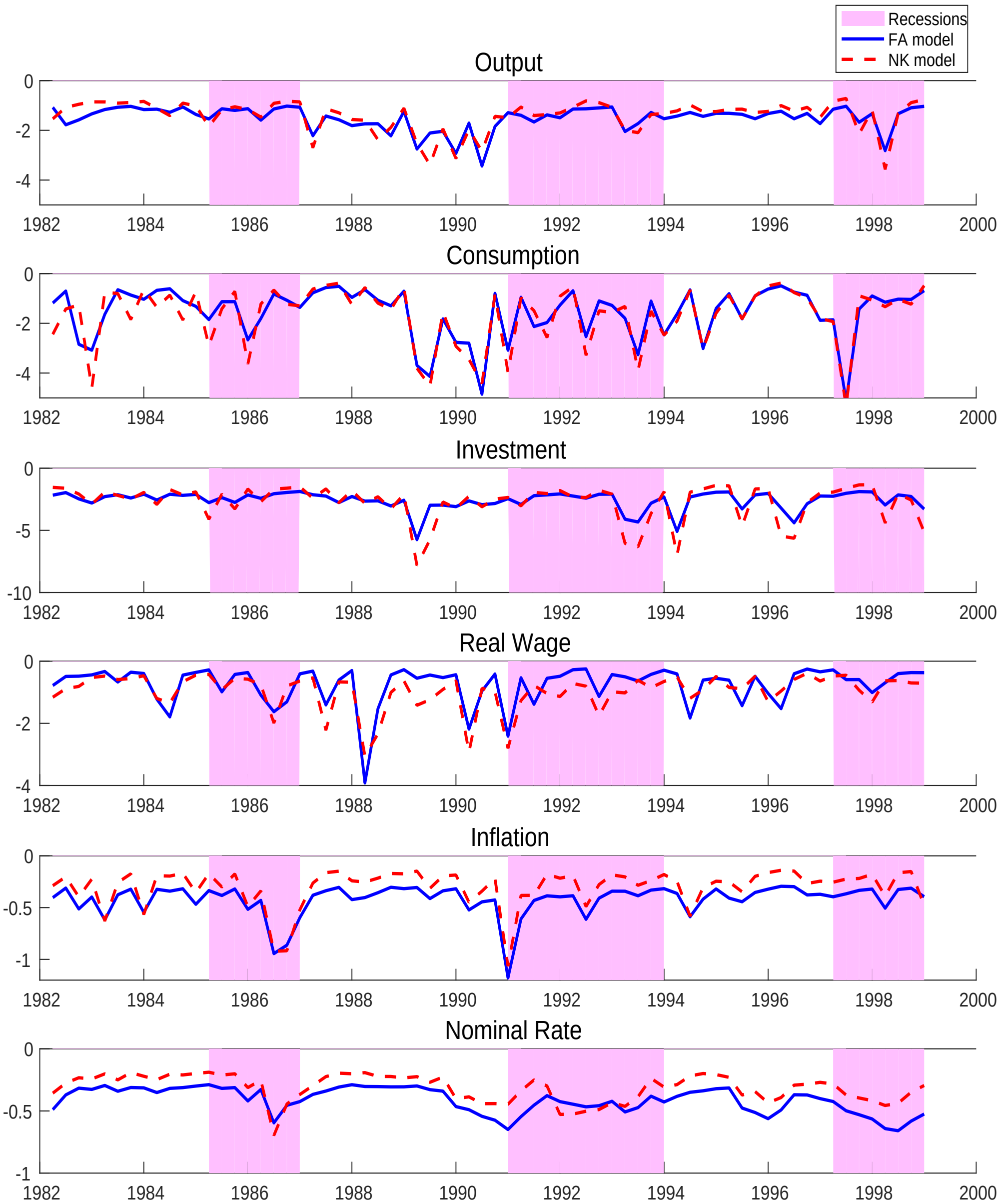


(b) FA model: Forecast as of 90:Q4



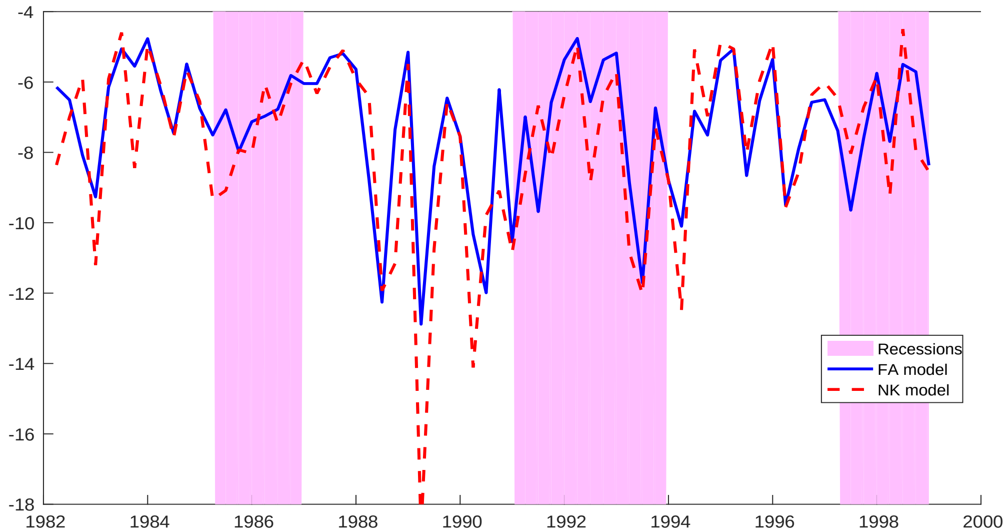
Note: The posterior prediction distributions of the DSGE models are calculated based on the Monte Calro procedure as described in Section 4.2, using 10,000 draws of posterior estimates over the full sample. FA model and NK model stand for the DSGE models with financial friction and without financial friction, respectively.

Figure 3: Predictive Log Scores of Individual Observations



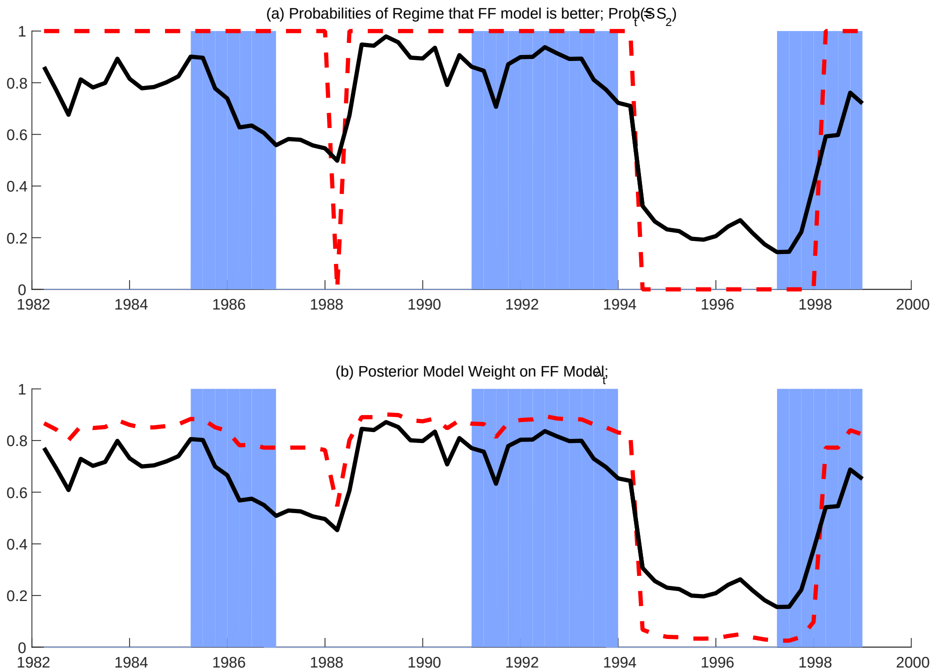
Note: The log score at each period is calculated from log $p(y_t^O; Y_{t-1}^O, M_i)$ for i = 1,2, of the individual model out of the NK model and the FA model as explained in Sec 2.

Figure 4: Predictive Log Scores of the Whole Six Observations



Note: The log score at each period is calculated from log $p(y^f_{t+h}; Y^O_t, M_i)$ for i = 1,2, of the individual model out of the NK model and the FA model as explained in Sec 2.

Figure 5: Markov-switching Prediction Pool of Whole Six Observations



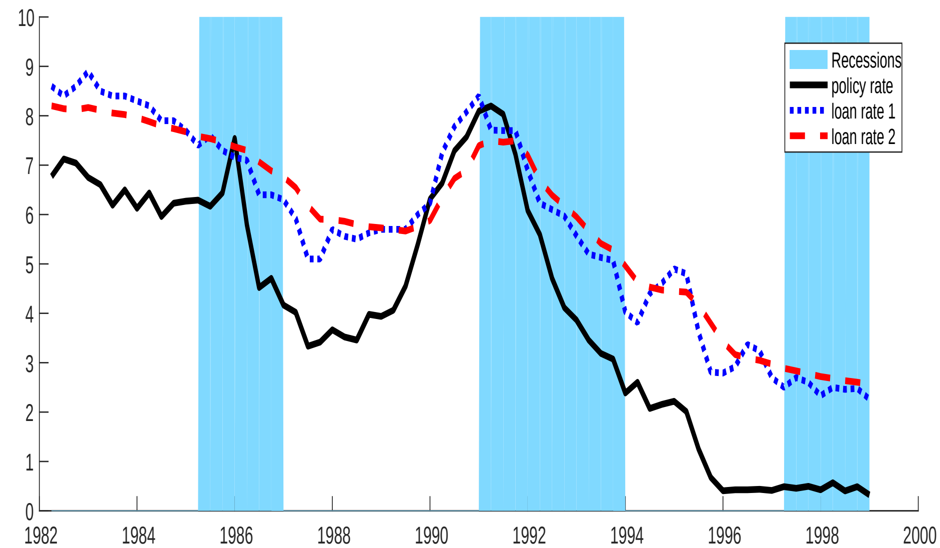
Notes:

1. Markov-switching pooling model is calculated from Eq.(3.3). The weighting coefficients on the FA model, λ_i, are estimated with MCMC simulation and obtained from 100,000 draws after discarding the first 40,000 burn-in draws.

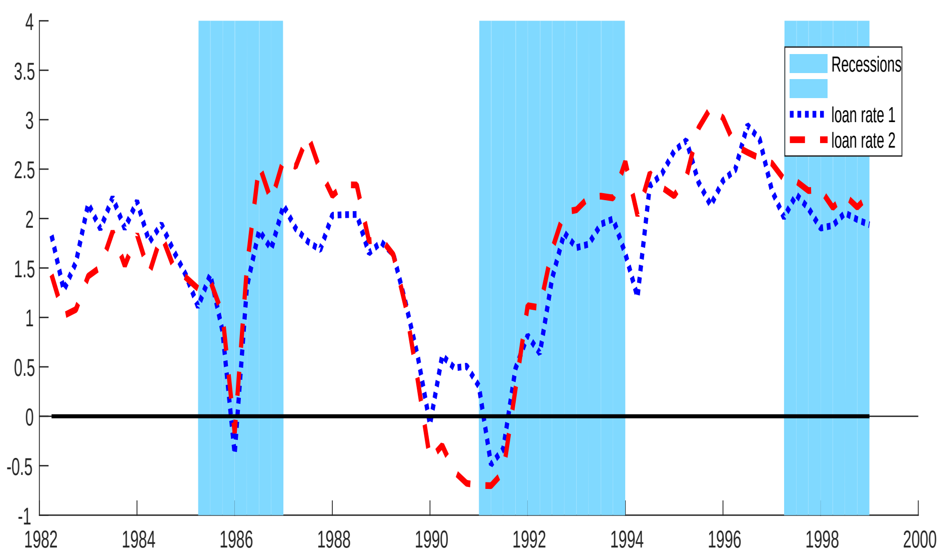
2. The solid black and red lines denote their posterior means and medians, respectively, and the blue shaded area represents recessions reported by the Cabinet Office, government of Japan.

Figure 6: Levels and Speads of Corporate Loan Rates

(a) Levels of Loan Rates and Policy Rate



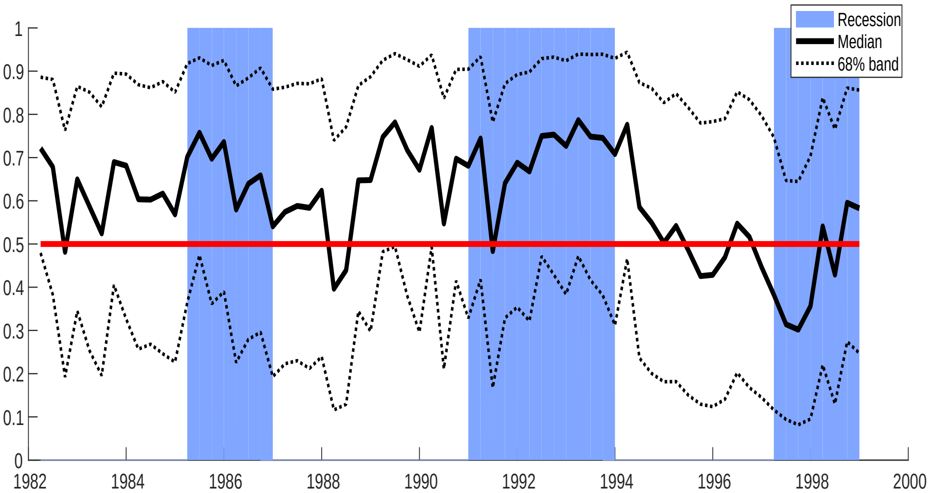
(b) Spreads between Loan Rates and Policy Rate



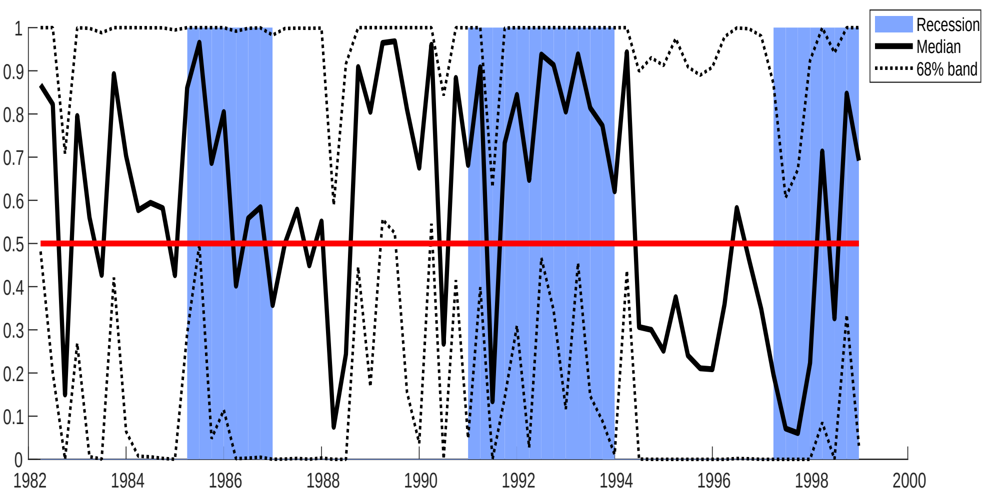
Notes: Policy rate stand for the Bank of Japan's secured overnight call rate. Loan rate 1 and loan rate 2 represent the long-term prime lending rate of Long-term credit banks and the average contractural interest rate on bank loan for large scale firms, respectively.

Figure 7: Dynamic Prediction Pool of the Whole Six Observations

(a) One flexible parameter



(b) Three flexible parameters



Notes:

1. Dynamic pooling model is calculated from Eq.(3). The time-varying coefficient is estimated from 20,000 draws by the particle MCMC simulation with 5,000 particles, after the first 5,000 draws are discarded.

2. In one flexible parameters model of panel (a), the constant coefficients μ and σ are fixed as $\mu = 0$ and $\sigma = 1$, following Del Negro et al. (2014).

3. In three felexible parameters model of panel (b), the coefficients μ and σ are estimated as well as ρ, following Del Negro et al. (2014).

4. The solid black line denotes their posterior means and the blue shaded area represents their 68% confidence interval.