CARF Working Paper

CARF-F-303

Endogenous Monetary Policy Shifts and the Term Structure: Evidence from Japanese Government Bond Yields

Junko Koeda
The University of Tokyo

First version: December 18, 2012
Current version: April 26, 2013

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Dai-ichi Mutual Life Insurance Company, Meiji Yasuda Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

Endogenous Monetary Policy Shifts and the Term Structure: Evidence from Japanese Government Bond Yields

Junko Koeda*

First version: December 18, 2012
This version: April 26, 2013

Abstract

I construct a no-arbitrage term structure model with endogenous regime shifts and apply it to Japanese government bond (JGB) yields. This model subjects the short-term interest rate to monetary regime shifts, specifically a zero interest rate policy (ZIRP) and normal regimes, which depend on macroeconomic variables. The estimates show that under the ZIRP, the deflationary effect on bond yields increases on the long end of yield curves. However, output gaps' ability to raise bond yields weakens for all maturities.

*Faculty of Economics, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan. Tel: +81-3-5841-5649, Email: jkoeda@e.u-tokyo.ac.jp.

1 Introduction

Policy shifts often depend on state variables in the real world. For example, a zero interest rate policy (ZIRP) may be introduced when the Taylor-rule policy rate, which is a function of real activity and inflation variables, hits the zero lower bound of the policy rate. Policymakers may lift a ZIRP when the exit conditions on macroeconomic variables are satisfied. Do such state-dependent regime shifts significantly affect the yield on government bonds? The answer to this question is relevant for countries currently under a ZIRP and struggling to meet macroeconomic conditions to exit the ZIRP. The answer is of particular importance for Japan, which has already experienced such regime switches several times. For Japan, small increases in bond yields can strain public finances with the ratio of public debt to GDP already exceeding 200 percent.

I examine how bond yields are affected by the state-dependent shifts in and out of a ZIRP by constructing a no-arbitrage term structure model with discrete regime shifts. The model has three key features. First, the probability of transitioning from a ZIRP depends on the state variables that appear in the monetary policy rule (e.g., output gap and inflation), allowing the entry and exit conditions of a ZIRP to depend on the state. Second, the model uses discrete regime shifts in the affine term structure framework, addressing possible nonlinearity in the conditional means of short-term yields (Ang and Bekaert, 2002).[1] Third, the state vector, which includes the policy rate, depends on the current,

[1] Alternatively, term structure models that lie outside of the affine family have been applied to the Japanese zero rate environments. See Ichiue and Ueno (2012) for an application of Black's (1995) model to JGB yields and Singleton and Kim (2012) for a comparison between Cox, Ingersoll, and Ross (1985) type affine model and non-affine models.

rather than the previous, monetary policy regime. This third feature is not trivial: today's policy rate should depend on today's monetary policy regime. If the model did not include this third feature, it would inappropriately allow a policy rate well above zero even during ZIRP periods. The model includes the policy rate in the state vector so that the lagged policy rate can affect the dynamics of macroeconomic variables, as modeled in the standard monetary VAR models (e.g., Stock and Watson, 2001) and several macro-finance term structure models (e.g, Ang, Piazzesi, and Wei, 2006, and Hördahl, Tristani, and Vestin, 2006).

This paper's model is related to the existing discrete regime-switching ATSMs[2] in the following ways. First, most existing models implement discrete regime shifts and dependence on the current policy rate (the second and third features) with a constant transition matrix (e.g., Bansal and Zhou, 2002, Ang, Bekaert, and Wei, 2008, Hamilton and Wu, 2011). This paper extends these models by introducing state dependent transition probabilities. Second, Dai, Singleton, and Yang (2007, henceforth DSY) implement state-dependent transition probabilities and discrete regime shifts (the second and third features) under the data generating or physical ($\mathbb{P}$) measure, while their state vector depends on the previous regime. This paper extends DSY's work by adding dependence on the current policy rate (the third feature) and by providing formal propositions and proofs and discussion on the link between the $\mathbb{P}$ and the risk neutral ($\mathbb{Q}$) measures.

To apply the model, I estimate the responsiveness of Japanese government bond (JGB)

[2] Alternatively, Ang, Boivin, Dong, and Loo-Kung (2011) model monetary policy shifts with continuous time-varying Taylor rule coefficients, treating these coefficients as latent factors. This paper differs from their model as it focuses on discrete and observable monetary regime shifts.

yields to a ZIRP and to macroeconomic conditions. The actual Japanese policy interest rate process (Figure 1) appears to have at least two regimes: a period during which the policy interest rate is near zero and flat (the ZIRP regime) and the remaining periods (the normal regime). Using these data as motivation, I construct a model with two regimes. Exploiting information from Bank of Japan public policy announcements, I treat the regimes as observable.[3] I model bond pricing with factor dynamics that incorporate endogenous monetary policy shifts as well as some key aspects of a ZIRP. Specifically, the model includes the zero lower bound and the Bank of Japan' forward guidance of interest rates.[4]

Following Oda and Ueda's (2007) set up with ZIRP exit rules, I consider two types of regime evolutions: one that incorporates the zero lower bound and one that additionally includes the forward guidance. I first consider a simple evolution of the regime that depends solely on a Taylor-rule policy rate (Type I evolution). If the Taylor-rule policy rate hits the lower bound of interest rates, policymakers set the policy rate at the bound under the ZIRP regime; otherwise the Taylor rule sets the policy rate under the normal regime. Type I evolution, however, does not take into account the Bank of Japan's forward guidance policy, a key feature of the ZIRP in Japan (for a discussion on this policy, see e.g., Ueda, 2012a and 2012b and Ugai's survey, 2007). Thus I extend the Type I evolution by introducing a forward guidance policy under which the ZIRP continues unless some inflation condition

[3] Another approach to modeling the regime process is to treat it as unobservable. For example, see Fujiwara (2006) and Inoue and Okimoto (2008) for Markov-switching models applied to the Japanese policy interest rate process.

[4] The Bank of Japan's forward guidance of interest rates is also called the *Jikan Jiku* (policy duration) policy.

is satisfied (Type II evolution).

I present the main results in two steps. First, I compare the empirical results of factor dynamics with and without the forward guidance. One notable difference between the two cases is that the estimated state-dependent transition probabilities are more persistent when the forward guidance is present. Furthermore, empirical evidence indicates that the evolution with the forward guidance fits the data much better than without it; out-of-sample performance results, however, are mixed. Second, I discuss the estimated yield curves and term premia using the term structure model that incorporates the forward guidance as the benchmark model. The estimated yield curves indicate that under a ZIRP output gaps' ability to raise bond yields weakens for all maturities, whereas the deflationary effect on JGB yields becomes stronger at the longer end of yield curves. Furthermore, the estimated term premia indicate that the large bond yield decline in the early 1990s was driven by expectation components, whereas that of the late 1990s was driven by both expectations and term-premium components. Term premia also declined after the introduction of the quantitative easing monetary policy (QEP) in March 2001.

This paper proceeds as follows. Section 2 describes a term structure model with endogenous regime shifts. Section 3 describes the specific regime evolutions considered. Section 4 and 5 discuss the estimation strategy and results. Section 6 concludes the paper.

2 The model

2.1 The $\mathbb{P}$ model

The state of the economy is assumed to follow a discrete time stationary Markov process $\{\mathbf{y}_t, s_t\}$ where $\mathbf{y}_t$ is a vector of continuous variables and s_t is a scalar discrete variable indicating the regime. Both $\mathbf{y}_t$ and s_t are observable. The vector $\mathbf{y}$ includes the short rate allowing the lagged short rate to directly affect the dynamics of $\mathbf{y}$. For notational convenience, the current value has tilda, and the previous period's value has no time subscript. The joint density-distribution function of $(\widetilde{\mathbf{y}}, \tilde{s})$ conditional on $(\mathbf{y}, s)$ can be expressed as the product of the conditional density of $\tilde{y}$ given $(\tilde{s}, \mathbf{y}, s)$ and the transition probability of the regime, denoted as $f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s)$ and $\rho(\tilde{s}|\mathbf{y}, s)$ respectively, i.e.,

$$p(\widetilde{\mathbf{y}}, \tilde{s}|\mathbf{y}, s) = f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s)\,\rho(\tilde{s}|\mathbf{y}, s)\,. \tag{1}$$

The $\mathbb{P}$ model satisfies two assumptions. Assumption 1 follows Hamilton (1989) that $f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s)$ depends on $\tilde{s}$ but not on s. This formulation differs from DSY, they instead assume that the conditional density of $\widetilde{\mathbf{y}}$ depends on s but not on $\tilde{s}$. Thus, the $\mathbb{P}$ model can be interpreted as an extension of DSY with Hamilton's formulation where $f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) = f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y})$. As discussed earlier, Hamilton's formulation is suitable particularly when the state vector includes a policy variable (e.g., the policy interest rate) determined by the current policy regime.

Assumption 1: $f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) = f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y})$**, and** $\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y} \sim N\left(\boldsymbol{\mu}(\tilde{s}, \mathbf{y}), \boldsymbol{\Sigma}(\tilde{s})\boldsymbol{\Sigma}(\tilde{s})'\right)$ **under** $\mathbb{P}$**.**

Assumption 2 follows DSY's assumption on the Radon-Nikodym derivative or equivalently that on the pricing kernel ($\mathcal{M}$) that accommodates both regime shift and factor risks. $\boldsymbol{\lambda}(\tilde{s}, \mathbf{y})$ and $\gamma(\tilde{s}, \mathbf{y}, s)$ are the prices-of-risk and regime-shift-risk coefficients, respectively.

Assumption 2: $\mathcal{M}(\tilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) = \exp \begin{bmatrix} -r(\mathbf{y}) - \gamma(\tilde{s}, \mathbf{y}, s) - \frac{1}{2}\boldsymbol{\lambda}(\tilde{s}, \mathbf{y})' \boldsymbol{\lambda}(\tilde{s}, \mathbf{y}) \\ -\boldsymbol{\lambda}(\tilde{s}, \mathbf{y})' \boldsymbol{\Sigma}(\tilde{s})^{-1} (\tilde{\mathbf{y}} - \boldsymbol{\mu}(\tilde{s}, \mathbf{y})) \end{bmatrix}$.

With the state $(\mathbf{y}, s)$, no arbitrage requires that n period bond prices satisfy

$$P_{n+1}(\mathbf{y}, s) = \sum_{\tilde{s}} \rho(\tilde{s}|\mathbf{y}, s) E[\mathcal{M}(\tilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) P_n(\tilde{\mathbf{y}}, \tilde{s}) |\tilde{s}, \mathbf{y}, s]. \tag{2}$$

For $n = 0$, since $\exp(-r(\mathbf{y})) = P_1(\mathbf{y}, s)$ and $P_0 = 1$, equation (2) becomes

$$\begin{aligned} \exp[-r(\mathbf{y})] &= \sum_{\tilde{s}} \rho(\tilde{s}|\mathbf{y}, s) E[\mathcal{M}(\tilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) |\tilde{s}, \mathbf{y}, s], \\ &= \sum_{\tilde{s}} \rho(\tilde{s}|\mathbf{y}, s) \exp[-r(\mathbf{y}) - \gamma(\tilde{s}, \mathbf{y}, s)], \end{aligned} \tag{3}$$

where that the second equality holds since

$$E[\mathcal{M}(\tilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) |\tilde{s}, \mathbf{y}, s] = \exp[-r(\mathbf{y}) - \gamma(\tilde{s}, \mathbf{y}, s)]. \tag{4}$$

Equation (3) can be simplified to

$$\sum_{\tilde{s}} \rho(\tilde{s}|\mathbf{y}, s) \exp[-\gamma(\tilde{s}, \mathbf{y}, s)] = 1. \tag{5}$$

2.2 From $\mathbb{P}$ to $\mathbb{Q}$

This subsection provides two propositions to link $\mathbb{P}$ to risk neutral (hereafter denoted by $\mathbb{Q}$) measures by finding a conditional density $f^Q(\tilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s)$ and a transition matrix $\rho^Q(\tilde{s}|\mathbf{y}, s)$, such that

$$\exp[-r(\mathbf{y})] \underbrace{f^Q(\tilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) \rho^Q(\tilde{s}|\mathbf{y}, s)}_{=p^Q(\tilde{\mathbf{y}}, \tilde{s}|\mathbf{y}, s)} = \mathcal{M}(\tilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) \underbrace{f(\tilde{\mathbf{y}}|\tilde{s}, \mathbf{y}) \rho(\tilde{s}|\mathbf{y}, s)}_{=p(\tilde{\mathbf{y}}, \tilde{s}|\mathbf{y}, s)}, \tag{6}$$

where the Radon-Nikodym derivative is given by $1/\left[\exp\left[r\left(\mathbf{y}\right)\right]\mathcal{M}\left(\tilde{\mathbf{y}},\tilde{s},\mathbf{y},s\right)\right].$ For any random variable X, define $E^{Q}\left(X|\mathbf{y},s\right)$ by

$$E^{Q}\left(X|\mathbf{y},s\right)\equiv\sum_{\tilde{s}}\rho^{Q}\left(\tilde{s}|\mathbf{y},s\right)E^{Q}\left(X|\tilde{s},\mathbf{y},s\right),\text{ where }E^{Q}\left(X|\tilde{s},\mathbf{y},s\right)\equiv\int_{\tilde{\mathbf{y}}}Xf^{Q}\left(\tilde{\mathbf{y}}|\tilde{s},\mathbf{y},s\right)d\tilde{\mathbf{y}}.$$

Then $E^{Q}\left(\exp\left[-r\left(\mathbf{y}\right)\right]X|\mathbf{y},s\right)=E\left(\mathcal{M}X|\mathbf{y},s\right)$ for f^{Q} and ρ^{Q} satisfying equation (6). In particular, the no-arbitrage condition (2) can be written as

$$P_{n+1}\left(\mathbf{y},s\right)=\sum_{\tilde{s}}\rho^{Q}\left(\tilde{s}|\mathbf{y},s\right)E^{Q}\left[\exp\left[-r\left(\mathbf{y}\right)\right]P_{n}\left(\tilde{\mathbf{y}},\tilde{s}\right)|\tilde{s},\mathbf{y},s\right]. \tag{7}$$

Proposition 1 *Under Assumptions 1 and 2, the transition probability $\rho^{Q}\left(\tilde{s}|\mathbf{y},s\right)$ is given by $\rho^{Q}\left(\tilde{s}|\mathbf{y},s\right)=\rho\left(\tilde{s}|\mathbf{y},s\right)\exp\left[-\gamma\left(\tilde{s},\mathbf{y},s\right)\right].$*

Proof. Appendix C. ■

Proposition 2 *The conditional density $f^{Q}\left(\tilde{\mathbf{y}}|\tilde{s},\mathbf{y},s\right)$ does not depend on s and is the density of $N\left(\boldsymbol{\mu}^{Q}\left(\tilde{s},\mathbf{y}\right),\boldsymbol{\Sigma}\left(\tilde{s}\right)\boldsymbol{\Sigma}\left(\tilde{s}\right)'\right)$ with $\boldsymbol{\mu}^{Q}\left(\tilde{s},\mathbf{y}\right)\equiv\boldsymbol{\mu}\left(\tilde{s},\mathbf{y}\right)-\boldsymbol{\Sigma}\left(\tilde{s}\right)\boldsymbol{\lambda}\left(\tilde{s},\mathbf{y}\right).$*

Proof. Appendix C. ■

2.3 Pricing under $\mathbb{Q}$

The no-arbitrage condition under $\mathbb{Q}$ can be rewritten as

$$1=\sum_{\tilde{s}}\rho^{Q}\left(\tilde{s}|\mathbf{y},s\right)E^{Q}\left\{\exp\left(\tilde{h}_{t+1}\right)|\tilde{s},\mathbf{y},s\right\}, \tag{8}$$

where

$$\begin{aligned}\tilde{h}_{t+1} &\equiv p_{n}\left(\tilde{\mathbf{y}},\tilde{s}\right)-p_{n+1}\left(\mathbf{y},s\right)-r\left(\mathbf{y}\right),\\ p_{n} &\equiv \log\left(P_{n}\right).\end{aligned}$$

The term $\tilde{h}_{t+1}$ is the log excess one-period return on $n+1$ period bonds.

If $\tilde{h}_{t+1}$ is conditionally normally distributed given $(\tilde{s}, \mathbf{y}, s)$ under $\mathbb{Q}$, then equation (8) becomes

$$1 = \sum_{\tilde{s}} \rho^{Q}(\tilde{s}|\mathbf{y}, s) \exp\left[E^{Q}\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right) + \frac{1}{2} Var^{Q}\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right)\right]. \tag{9}$$

Furthermore, by applying the approximation used by Bansal and Zhou (2002) and Hamilton and Wu (2012) (i.e., $\exp(x) \approx 1 + x$)[5], equation (9) becomes

$$0 \approx \sum_{\tilde{s}} \rho^{Q}(\tilde{s}|\mathbf{y}, s) \left[E^{Q}\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right) + \frac{1}{2} Var^{Q}\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right)\right]. \tag{10}$$

In order to solve bond prices, I assume two additional assumptions that are commonly assumed in the existing literature. Assumption 3 assumes that the factors that explain the yield curves follow VAR(1) under $\mathbb{Q}$.

Assumption 3:

$$E^{Q}(\tilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) \equiv \boldsymbol{\mu}^{Q}(\tilde{s}, \mathbf{y}) = \mathbf{c}^{Q}(\tilde{s}) + \boldsymbol{\Phi}^{Q}(\tilde{s})\, \mathbf{y}.$$

Assumption 4 assumes that the transition probabilities under $\mathbb{Q}$ are constant. This implies, by Proposition 1 and equation (5), $\rho(\tilde{s}|\mathbf{y}, s) \exp[-\gamma(\tilde{s}, \mathbf{y}, s)]$ does not depend on $\mathbf{y}$ and its sum over $\tilde{s}$ is 1, so that the regime-shift risk is "fully priced." I discuss the case without Assumption 4 in Appendix A.

Assumption 4:

$$\rho^{Q}(\tilde{s}|\mathbf{y}, s) = \rho^{Q}(\tilde{s}|s).$$

[5] One can note that if $\boldsymbol{\Phi}^{Q}$ does not depend on the regime s, as in DSY, then it is unnecessary to invoke the "$\exp(x) \approx 1 + x$" approximation. However, this specification does not allow the policy rule coefficients as well as coefficients in the factor dynamics to depend on the regime.

To solve bond prices, I first conjecture that

$$p_n\left(\widetilde{\mathbf{y}}, \tilde{s}\right) = -a_n\left(\tilde{s}\right) - \mathbf{b}_n\left(\tilde{s}\right)\tilde{y}.$$

Since $\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s \sim N\left(\mu\left(\tilde{s}, \mathbf{y}\right), \Sigma\left(\tilde{s}\right)\Sigma\left(\tilde{s}\right)'\right)$ under $\mathbb{Q}$, the log excess return $\tilde{h}_{t+1}$ is conditionally normal, as required above, with

$$\begin{aligned}
E^Q\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right) &= -a_n\left(\tilde{s}\right) - \mathbf{b}_n\left(\tilde{s}\right)\left[\mathbf{c}^Q\left(\tilde{s}\right) + \mathbf{\Phi}^Q\left(\tilde{s}\right)\mathbf{y}\right] + a_{n+1}\left(s\right) + \mathbf{b}_{n+1}\left(s\right)\mathbf{y} - r\left(\mathbf{y}\right), \\
Var^Q\left(\tilde{h}_{t+1}|\tilde{s}, \mathbf{y}, s\right) &= \mathbf{b}_n\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)'_n\mathbf{b}\left(\tilde{s}\right)'.
\end{aligned}$$

Substituting these equations into (10) yields

$$\begin{aligned}
0 \approx & -\left[\sum_{\tilde{s}}\rho^Q\left(\tilde{s}|s\right)\left(a_n\left(\tilde{s}\right) + \mathbf{b}_n\left(\tilde{s}\right)\mathbf{c}^Q\left(\tilde{s}\right) - \frac{1}{2}\mathbf{b}_n\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)'\mathbf{b}_n\left(\tilde{s}\right)'\right) - a_{n+1}\left(s\right)\right] \\
& -\sum_{\tilde{s}}\rho^Q\left(\tilde{s}|s\right)\left[\mathbf{b}_n\left(\tilde{s}\right)\mathbf{\Phi}^Q\left(\tilde{s}\right) + \mathbf{e}_1' - \mathbf{b}_{n+1}\left(s\right)\right]\mathbf{y},
\end{aligned}$$

where $\mathbf{e}_1$ is a vector of zeros with 1st element being one (thus $r\left(\mathbf{y}\right) = \mathbf{e}_1'\mathbf{y}$). Since this has to hold for any $\mathbf{y}$, I obtain the recursion

$$\begin{aligned}
a_{n+1}\left(s\right) &= \sum_{\tilde{s}}\rho^Q\left(\tilde{s}|s\right)\left(a_n\left(\tilde{s}\right) + \mathbf{b}_n\left(\tilde{s}\right)\mathbf{c}^Q\left(\tilde{s}\right) - \frac{1}{2}\mathbf{b}_n\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)\mathbf{\Sigma}\left(\tilde{s}\right)'_n\mathbf{b}\left(\tilde{s}\right)'\right), \\
\mathbf{b}_{n+1}\left(s\right) &= \sum_{\tilde{s}}\rho^Q\left(\tilde{s}|s\right)\left(\mathbf{b}_n\left(\tilde{s}\right)\mathbf{\Phi}^Q\left(\tilde{s}\right) + \mathbf{e}_1'\right).
\end{aligned}$$

The initial condition is $a_0\left(s\right) = 0$ and $\mathbf{b}_0\left(s\right) = 0$ for all s.

Lastly, by Proposition 2 and Assumption 3, $\mathbf{c}^Q\left(\tilde{s}\right)$ and $\mathbf{\Phi}^Q\left(\tilde{s}\right)$ can be expressed with the prices of risk coefficients

$$\mathbf{c}^Q\left(\tilde{s}\right) = \mathbf{c}\left(\tilde{s}\right) - \mathbf{\Sigma}\left(\tilde{s}\right)\boldsymbol{\lambda}_0\left(\tilde{s}\right), \ \mathbf{\Phi}^Q\left(\tilde{s}\right) = \mathbf{\Phi}\left(\tilde{s}\right) - \mathbf{\Sigma}\left(\tilde{s}\right)\mathbf{\Lambda}_1\left(\tilde{s}\right), \tag{11}$$

if $\widetilde{\mathbf{y}}$ follows the VAR(1) process under $\mathbb{P}$, i.e., $\boldsymbol{\mu}\left(\tilde{s}, \mathbf{y}\right) = \mathbf{c}\left(\tilde{s}\right) + \mathbf{\Phi}\left(\tilde{s}\right)\mathbf{y}$ and the prices of risk are affine in $\mathbf{y}$, i.e., $\boldsymbol{\lambda}\left(\tilde{s}, \mathbf{y}\right) = \boldsymbol{\lambda}_0\left(\tilde{s}\right) + \mathbf{\Lambda}_1\left(\tilde{s}\right)\mathbf{y}$.

3 An application to JGB yields

In this section, I specify factor dynamics that correspond to $f(\tilde{\mathbf{y}}|\tilde{s}, \mathbf{y})$ in the $\mathbb{P}$ model and then endogenously derive the corresponding state-dependent transition probabilities $\rho(\tilde{s}|\mathbf{y}, s)$. The specified factor dynamics is VAR(1) model of three variables (i) the policy interest rate, (ii) inflation, and (iii) output gaps, with endogenous monetary policy regime shifts. The VAR explicitly models the key features of recent monetary policies in Japan, such as the zero lower bound of the policy interest rate and the Bank of Japan's forward guidance of interest rates.

3.1 Factor dynamics

The monetary policy regime (s) can be either normal ($\mathcal{P}$) or a ZIRP ($\mathcal{Z}$). I partition $\mathbf{y}_t$ as

$$\underset{3\times1}{\mathbf{y}_t} = \begin{bmatrix} r_t \\ \underset{2\times1}{\mathbf{y}_{2t}} \end{bmatrix},$$

where r is the policy interest rate (short rate) and $\mathbf{y}_2$ is a vector of macroeconomic variables (inflation and output gap).

The policy interest rate follows a regime-dependent Taylor rule

$$r_t = \underset{(1\times1)}{\alpha^{s_t}} + \underset{(1\times2)}{\boldsymbol{\beta}^{s_t\prime}} \mathbf{y}_{2t} + \underset{(1\times1)}{\delta^{s_t}} r_{t-1} + \underset{(1\times1)}{\sigma_r^{s_t}} u_{r,t}, \quad u_{r,t} \sim N(0,1), \tag{12}$$

where the ZIRP regime can be represented as

$$\begin{aligned} \alpha^{\mathcal{Z}} &\approx 0,\ \boldsymbol{\beta}^{\mathcal{Z}} = \mathbf{0},\ \delta^{\mathcal{Z}} = 0,\ \sigma_r^{\mathcal{Z}} \approx 0, \\ \mathrm{ZLB}_t &= \alpha^{\mathcal{Z}} + \sigma_r^{\mathcal{Z}} u_{r,t}, \end{aligned} \tag{13}$$

with the lower bound of policy rate (ZLB) given by $\alpha^{\mathcal{Z}} + \sigma_r^{\mathcal{Z}} u_{r,t}$. The Taylor-rule (TR) policy rate is the rate given by equation (12) under the normal regime.

The rest of system is also regime dependent

$$\mathbf{y}_{2,t} = \underset{(2\times1)}{\mathbf{c}_2^{s_t}} + \underset{(2\times3)}{\mathbf{\Phi}_2^{s_t}} \mathbf{y}_{t-1} + \underset{(2\times2)}{\mathbf{\Sigma}_{22}^{s_t}} \mathbf{u}_{2t}, \quad \mathbf{u}_{2t} \sim N(\mathbf{0}, \mathbf{I}). \tag{14}$$

All shocks are jointly standard normal and independent to each other and over time. Substituting (14) in the short rate equation yields

$$r_t = \alpha^{s_t} + \boldsymbol{\beta}^{s_t\prime} \mathbf{c}_2^{s_t} + \left[\boldsymbol{\beta}^{s_t\prime} \mathbf{\Phi}_2^{s_t} + [\delta^{s_t}, 0, 0, 0]\right] \mathbf{y}_{t-1} + \boldsymbol{\beta}^{s_t\prime} \mathbf{\Sigma}_{22}^{s_t} \mathbf{u}_{2t} + \sigma_r^{s_t} u_{r,t}.$$

Stacking the above equation over (14) results in the VAR

$$\mathbf{y}_t = \underset{(3\times1)}{\mathbf{c}^{s_t}} + \underset{(3\times3)}{\mathbf{\Phi}^{s_t}} \mathbf{y}_{t-1} + \underset{(3\times3)}{\mathbf{\Sigma}^{s_t}} \mathbf{u}_t, \quad \mathbf{u}_t \sim N(\mathbf{0}, \mathbf{I}), \tag{15}$$

where

$$\underset{(3\times1)}{\mathbf{c}^{s_t}} = \begin{bmatrix} \alpha^{s_t} + \boldsymbol{\beta}^{s_t\prime} \mathbf{c}_2^{s_t} \\ \underset{(2\times1)}{\mathbf{c}_2^{s_t}} \end{bmatrix}, \quad \underset{(3\times3)}{\mathbf{\Phi}^{s_t}} = \begin{bmatrix} \boldsymbol{\beta}^{s_t\prime} \mathbf{\Phi}_2^{s_t} + [\delta^{s_t}, 0, 0, 0] \\ \underset{(2\times3)}{\mathbf{\Phi}_2^{s_t}} \end{bmatrix}, \quad \underset{(3\times3)}{\mathbf{\Sigma}^{s_t}} = \begin{bmatrix} \sigma_r^{s_t} & \boldsymbol{\beta}^{s_t\prime} \mathbf{\Sigma}_{22}^{s_t} \\ \mathbf{0} & \mathbf{\Sigma}_{22}^{s_t} \end{bmatrix}.$$

Equation (15) is a restricted VAR because the first row of $\mathbf{\Phi}^{\mathcal{Z}}$ is a vector of zeros and the six elements of $\mathbf{\Sigma}^{\mathcal{Z}}$ are a function of four parameters.

3.2 Regime determination

Following Oda and Ueda' (2007) set up with ZIRP exit rules, I consider two types of regime evolution. I first consider a simple regime evolution that depends solely on the level of the Taylor-rule (TR) policy rate (Type I evolution). If the Taylor-rule policy hits the ZLB, the policy rate is set at the bound under the ZIRP regime ($s = \mathcal{Z}$), otherwise the policy rate is set by the Taylor rule under the normal regime ($s = \mathcal{P}$).

Type I evolution, however, does not take into account the Bank of Japan's forward guidance policy. Thus I extend the Type I evolution by introducing a forward guidance policy under which the ZIRP regime is maintained unless the expected year-on-year core inflation exceeds a certain level (Type II evolution). Such a level can be interpreted as the exit condition based on core inflation rate ($\bar{\pi}$), a parameter that must be estimated since the Bank of Japan did not commit a specific rate during the investigated period.

For notational convenience, define

$$\begin{aligned} r^e(\mathbf{y}_{t-1}) &\equiv \alpha^{\mathcal{P}} + \boldsymbol{\beta}^{\mathcal{P}\prime}\left[\mathbf{c}_2^{\mathcal{P}} + \boldsymbol{\Phi}_2^{\mathcal{P}}\mathbf{y}_{t-1}\right] + \delta^{\mathcal{P}} r_{t-1}, \\ \pi^e(\mathbf{y}_{t-1}) &\equiv [1,0]\left[\mathbf{c}_2^{\mathcal{P}} + \boldsymbol{\Phi}_2^{\mathcal{P}}\mathbf{y}_{t-1}\right]. \end{aligned}$$

so that the TR rate can be rewritten as

$$\begin{aligned} \text{TR rate}_t &= \alpha^{\mathcal{P}} + \boldsymbol{\beta}^{\mathcal{P}\prime}\mathbf{y}_{2,t} + \delta^{\mathcal{P}} r_{t-1} + \sigma_r^{\mathcal{P}} u_{r,t}, \qquad (16) \\ &= \alpha^{\mathcal{P}} + \boldsymbol{\beta}^{\mathcal{P}\prime}\left(\mathbf{c}_2^{\mathcal{P}} + \boldsymbol{\Phi}_2^{\mathcal{P}}\mathbf{y}_{t-1} + \boldsymbol{\Sigma}_{22}^{\mathcal{P}}\mathbf{u}_{2t}\right) + \delta^{\mathcal{P}} r_{t-1} + \sigma_r^{\mathcal{P}} u_{r,t}, \\ &= r^e(\mathbf{y}_{t-1}) + \boldsymbol{\beta}^{\mathcal{P}\prime}\boldsymbol{\Sigma}_{22}^{\mathcal{P}}\mathbf{u}_{2t} + \sigma_r^{\mathcal{P}} u_{r,t}. \end{aligned}$$

3.2.1 Evolution with the zero rate bound: Type I evolution

Under Type I evolution, the regime is a ZIRP if and only if the TR rate (equation (16)) lies at or below the ZLB (equation (13)); otherwise it is normal. Thus the probability that the regime is normal is given by

$$\begin{aligned} \Pr(s_t = \mathcal{P}|\mathbf{y}_{t-1}, s_{t-1}) &= \Pr\left(\underbrace{\alpha^{\mathcal{P}} + \boldsymbol{\beta}^{\mathcal{P}\prime}\mathbf{y}_{2,t} + \delta^{\mathcal{P}} r_{t-1} + \sigma_r^{\mathcal{P}} u_{r,t}}_{\text{TR rate}_t} > \underbrace{\alpha^{\mathcal{Z}} + \sigma_r^{\mathcal{Z}} u_{r,t}}_{\text{ZLB}_t} \middle| \mathbf{y}_{t-1}, s_{t-1}\right), \\ &= \Pr\left(r^e(\mathbf{y}_{t-1}) + \boldsymbol{\beta}^{\mathcal{P}\prime}\boldsymbol{\Sigma}_{22}^{\mathcal{P}}\mathbf{u}_{2t} + \sigma_r^{\mathcal{P}} u_{r,t} > \alpha^{\mathcal{Z}} + \sigma_r^{\mathcal{Z}} u_{r,t}\right), \text{ (by (16))} \\ &= \Pr\left(r^e(\mathbf{y}_{t-1}) > \alpha^{\mathcal{Z}} - \xi_t\right), \end{aligned}$$

where $\xi_t \equiv \boldsymbol{\beta}^{\mathcal{P}\prime}\boldsymbol{\Sigma}_{22}^{\mathcal{P}}\mathbf{u}_{2t} + \left[\sigma_r^{\mathcal{P}} - \sigma_r^{\mathcal{Z}}\right] u_{r,t}$. The probability that the regime is the ZIRP is given by $\Pr\left(s_t = \mathcal{Z}|\mathbf{y}_{t-1}, s_{t-1}\right) = \Pr\left(r^e\left(\mathbf{y}_{t-1}\right) \leqq \alpha^{\mathcal{Z}} - \xi_t\right)$.

Since u_t and $\mathbf{u}_{2t}$ are normal and independent

$$\xi_t|\mathbf{y}_{t-1}, s_{t-1} \sim N\left(0, \sigma_\xi^2\right),\ \sigma_\xi^2 \equiv \left[\sigma_r^{\mathcal{P}} - \sigma_r^{\mathcal{Z}}\right]^2 + \boldsymbol{\beta}^{\mathcal{P}\prime}\boldsymbol{\Sigma}_{22}^{\mathcal{P}}\boldsymbol{\Sigma}_{22}^{\mathcal{P}\prime}\boldsymbol{\beta}^{\mathcal{P}}.$$

Thus the Type I transition probabilities can be rewritten as

$$\Pr\left(s_t = \mathcal{P}|\mathbf{y}_{t-1}, s_{t-1}\right) = \Pr\left(s_t = \mathcal{P}|\mathbf{y}_{t-1}\right) = F\left(\frac{r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right), \tag{17}$$

$$\Pr\left(s_t = \mathcal{Z}|\mathbf{y}_{t-1}, s_{t-1}\right) = \Pr\left(s_t = \mathcal{Z}|\mathbf{y}_{t-1}\right) = 1 - F\left(\frac{r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right), \tag{18}$$

where $F(.)$ is the cumulative distribution function of $N(0,1)$. These transition probabilities do not depend on s_{t-1}.

3.2.2 Forward guidance: Type II evolution

How does the regime evolution change when the forward guidance is introduced? Now, the determination of the regime depends on the inflation rate relative to $\bar{\pi}$, in addition to the TR policy rate.

If the previous regime is normal, then the transition probabilities are unchanged (i.e., equations (17) and (18)). On the other hand, if the previous regime is a ZIRP, the probability of it returning to normal is

$$\begin{aligned}
&\Pr\left(s_t = \mathcal{P}|\mathbf{y}_{t-1}, s_{t-1} = \mathcal{Z}\right) \\
=\ &\Pr\left(\underbrace{\alpha^{\mathcal{P}} + \boldsymbol{\beta}^{\mathcal{P}\prime}\mathbf{y}_{2,t} + \delta^{\mathcal{P}} r_{t-1} + \sigma_r^{\mathcal{P}} u_{r,t}}_{\text{TR rate}_t} > \underbrace{\alpha^{\mathcal{Z}} + \sigma_r^{\mathcal{Z}} u_{r,t}}_{\text{ZLB}_t},\ \underbrace{\pi_t > \bar{\pi}}_{\text{exit condition}} \middle| \mathbf{y}_{t-1}, s_{t-1} = \mathcal{Z}\right), \\
=\ &\Pr\left(r^e\left(\mathbf{y}_{t-1}\right) > \alpha^{\mathcal{Z}} - \xi_t, \pi^e\left(\mathbf{y}_{t-1}\right) > \bar{\pi} - \sigma_\pi u_{\pi,t}\right).
\end{aligned} \tag{19}$$

where $u_{\pi,t}$ and σ_π are the shock and volatility parameters, respectively, calculated in the inflation equation as $u_{\pi,t} = [1, 0]\, \mathbf{u}_{2t}$ and $\sigma_\pi = [1, 0]\, \mathbf{\Sigma}_{22}^{\mathcal{P}} \mathbf{\Sigma}_{22}^{\mathcal{P}\prime} [1, 0]'$.

The Type II transition probabilities can be rewritten as

$$
\begin{aligned}
\Pr\left(s_t = \mathcal{P} | \mathbf{y}_{t-1}, s_{t-1} = \mathcal{Z}\right) &= B\left(r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}, \pi^e\left(\mathbf{y}_{t-1}\right) - \bar{\pi}; \mathbf{W}\right), && (20) \\
\Pr\left(s_t = \mathcal{Z} | \mathbf{y}_{t-1}, s_{t-1} = \mathcal{Z}\right) &= 1 - B\left(r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}, \pi^e\left(\mathbf{y}_{t-1}\right) - \bar{\pi}; \mathbf{W}\right), && (21) \\
\Pr\left(s_t = \mathcal{P} | \mathbf{y}_{t-1}, s_{t-1} = \mathcal{P}\right) &= F\left(\frac{r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right), && (22) \\
\Pr\left(s_t = \mathcal{Z} | \mathbf{y}_{t-1}, s_{t-1} = \mathcal{P}\right) &= 1 - F\left(\frac{r^e\left(\mathbf{y}_{t-1}\right) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right), && (23)
\end{aligned}
$$

where $B(a, b; \mathbf{W})$ is the cumulative distribution function of the bivariate normal distribution with mean zero and the variance covariance matrix of $(\xi_t, u_{\pi,t})$. The variance covariance matrix, which is denoted as $\mathbf{W}$, is a known function of $\left(\sigma_r^{\mathcal{P}}, \sigma_r^{\mathcal{Z}}, \boldsymbol{\beta}^{\mathcal{P}}, \mathbf{\Sigma}_{22}^{\mathcal{P}}\right)$

$$
\mathbf{W} = \begin{bmatrix} \sigma_\xi^2 & \boldsymbol{\beta}^{\mathcal{P}\prime} \mathbf{\Sigma}_{22}^{\mathcal{P}} \mathbf{\Sigma}_{22}^{\mathcal{P}\prime} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \boldsymbol{\beta}^{\mathcal{P}\prime} \mathbf{\Sigma}_{22}^{\mathcal{P}} \mathbf{\Sigma}_{22}^{\mathcal{P}\prime} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{\Sigma}_{22}^{\mathcal{P}} \mathbf{\Sigma}_{22}^{\mathcal{P}\prime} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix}.
$$

4 Estimating JGB yield curves

4.1 Data

I use quarterly data on interest rates and the macro variables of inflation and output gap from 1985Q1 to 2008Q2. I use quarterly data because it may reflect Japan's overall economic activity more precisely than readily available monthly real activity measures, such as, industrial production, unemployment, and machinery orders. The sample period starts in 1985Q1 because reliable zero coupon bond yield data are available from that quarter; it ends in 2008Q2, the period prior to the Lehman shock.

The uncollaterized overnight call rate[6] is used for the short-term interest rate. Zero coupon bond yields of 4, 12, 20, and 40 quarter maturities are used for longer maturities. These bond yields are obtained from Wright's (2011) dataset and are the end of period rates expressed at annualized rates in percent.

Regarding the macro variables, inflation is measured by the percentage change in the Consumer Price Index, excluding fresh food, from the same quarter in the previous year, obtained from the Ministry of Internal Affairs and Communications[7]; real activity is measured by output gaps estimated by applying the Hodrick-Prescott filter to the logs of the seasonally adjusted GDP at 2000 prices, obtained from the Japanese Cabinet Office. Output gaps are expressed in percentage points.

The regime series is constructed based on public statements by the end of each quarter. One issue is how to differentiate the normal from the ZIRP regime; the regime can be normal even if the short rate is almost zero if the Taylor-rule interest rate indicates it. I thus identify the ZIRP regime with the forward guidance within Bank of Japan's public statements. This implies that the period from March to June 2006, when the targeted rate was zero in the absence of forward guidance policy, was under the "normal" regime.

[6] I use the average rate of the last month in each quarter to remove end-of-month fluctuations in the call rate.

[7] I use the 2000-base CPI up to mid-2006 since policy decisions were not made based on the 2005-base CPI up to that point.

4.2 Estimation strategy

The model consists of macro dynamics and static yield equations. The macro dynamics are summarized by equation (15) and the static yield equations are

$$\underset{4\times 1}{\mathbf{z}_t} = \underset{4\times 1}{A^{s_t}} + \underset{4\times 3}{\mathbf{B}^{s_t}}\mathbf{y}_t + \underset{4\times 1}{\mathbf{v}_t},$$

where $\mathbf{z}_t = [r_t^4, r_t^{12}, r_t^{20}, r_t^{40}]'$ is a 4×1 vector of bond yields with maturities corresponding to the superscript numbers (in quarters). The yield equations are an affine function of the state variables with 4×1 coefficient vectors A and a 4×3 coefficient matrix $\mathbf{B}$ corresponding to (i) a constant, (ii) the short-term interest rate, and (iii) the macro variables, respectively. The subscript numbers in A and $\mathbf{B}$ correspond to maturities, that is, $A^{s_t} = \left[\frac{a_4(s_t)}{4}, \frac{a_{12}(s_t)}{12}, \frac{a_{20}(s_t)}{20}, \frac{a_{40}(s_t)}{40}\right]'$, $\mathbf{B}^{s_t} = \left[\frac{\mathbf{b}_4(s_t)}{4}, \frac{\mathbf{b}_{12}(s_t)}{12}, \frac{\mathbf{b}_{20}(s_t)}{20}, \frac{\mathbf{b}_{40}(s_t)}{40}\right]'$.[8] The elements in A and $\mathbf{B}$ are derived from the recursive equations with the subscript numbers corresponding to maturities. Measurement errors $\mathbf{v}$ are assumed to have constant variance.

The system of equations to be estimated can be summarized as

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{c}^{s_t} + \mathbf{\Phi}^{s_t}\mathbf{y}_{t-1} + \mathbf{u}_t, && (24)\\
s_t &= h\left(s_{t-1}, \mathbf{y}_{t-1}, \mathbf{u}_{2t}, u_{r,t}\right), && (25)\\
\mathbf{z}_t &= A^{s_t} + \mathbf{B}^{s_t}\mathbf{y}_t + \mathbf{v}_t, &&\\
\mathbf{u}_t &\sim N\left(\mathbf{0}, \mathbf{\Sigma}^{s_t}\mathbf{\Sigma}^{s_t\prime}\right),\ \mathbf{v}_t \sim N\left(\mathbf{0}, \mathbf{V}\right). &&
\end{aligned}$$

where (25) is defined by (17) and (18) for Type I factor dynamics and by (20) to (23) for Type II factor dynamics. All the shocks are iid and $\mathbf{u}_t$ and $\mathbf{v}_s$ are independent for

[8] According to the basic relation between bond price and yield, the n-period bond yield is given by $\frac{a_n}{n} + \frac{\mathbf{b}_n}{b}\mathbf{y}$.

all (t, s). Thus, the observation equation linking $\mathbf{z}_t$ to the state $(\mathbf{y}_t)$ is appended to the VAR equations describing the state dynamics. I apply a two step procedure to estimate the model (e.g., Ang, Piazzesi, and Wei, 2006). Appendix B shows the derivation of the likelihood functions.

4.3 Estimated results

I present the estimated results in two steps. First, I compare the estimated factor dynamics with (Type II) and without (Type I) the forward guidance effect. Second, I discuss the estimated yield curves and term premium dynamics using the term structure model with Type II factor dynamics as the benchmark model. The estimated results using the term structure model with Type I factor dynamics are available upon request.

4.3.1 Factor dynamics: Type II versus Type I

How do the Type I and Type II factor dynamics differ from each other? Figure 2 shows the estimated state-dependent transition probabilities under each type of evolution. Under both types, $\Pr(s_t = \mathcal{P}|\mathbf{y}_{t-1}, s_{t-1} = \mathcal{P})$, that is, the estimated probability that the normal regime continues into the next period, is one until mid-1995. This is a reasonable result, since nobody would have imagined a ZIRP up to that point. A notable difference between the Type I and Type II transition probabilities is that the latter are more persistent. For example, Figure 2 shows that during the quantitative easing monetary policy (QEP) of March 2001 to February 2006, $\Pr(s_t = \mathcal{Z}|\mathbf{y}_{t-1}, s_{t-1} = \mathcal{Z})$ is estimated to be much higher under the Type II regime than under the Type I regime. This may reflect market pessimism over the recovery from deflation with the Bank of Japan's commitment that it would maintain the zero rate until some inflationary condition was satisfied.

The Taylor rule coefficients under the two types of factor dynamics are reported in Table 1. Under both types of regime evolution, the estimated coefficients have the right signs in terms of economic interpretation, and the long-run response of the short-term interest rate to a unit increase in inflation well exceeds one (3.6 and 2.2 under Types I and II, respectively). The Taylor rule coefficient with respect to inflation (i.e., the first element of $\boldsymbol{\beta}^{\mathcal{P}}$) under Type I is higher than that under Type II (0.34 versus 0.22 respectively), possibly reflecting the omission of the inflation variable in the regime evolution under Type I (see equations (26) and (27)). The estimated $\mathbf{c}_2^{\mathcal{Z}}$, $\boldsymbol{\Phi}_2^{\mathcal{Z}}$,and $\boldsymbol{\Sigma}_{22}^{\mathcal{Z}}$ are the same under Type I and II since they can be estimated separately (see Appendix B for details).

Figure 2 and Table 1 here

Which type of factor dynamics is more appropriate? I compare the empirical performances of the two types of models, with Type I and Type II factor dynamics, using two different approaches, and find that Type II has notably better fits to the data, though out-of-sample performance results are mixed. The first approach is the likelihood-ratio test. Setting Type 1 as the null and Type 2 as the alternative, a large positive test statistic (14.1) rejects the null. The second approach is one-period-ahead out-of-sample forecasting on the state variables (i.e., the short rate, inflation, and output gaps) to check these models' predictive accuracy. This exercise involves a rolling forecast covering the last three years of the QEP period. I evaluate the predictive accuracy by the following two measures: (i) the root mean square error ratios (Type I relative to Type II) and (ii) the modified Diebold-

Mariano test statistics[9] proposed by Harvey et al (1998) with differential loss based on the mean-squared errors. The results are summarized in Table 2. Type 2 weakly outperforms Type 1 for the policy rate and output gap forecasts, whereas it underperforms Type I for the inflation forecasts. Type 2's poor forecasting performance for inflation forecasts may be due to the imprecise estimate of $\bar{\pi}$.

Table 2 here

4.3.2 The yield curves and term premia

I now discuss the estimated yield curves and term premia under the benchmark model. The estimated prices-of-risk coefficients and transition probabilities under $\mathbb{Q}$ are reported in Table 3. The estimated prices-of-risk level coefficients ($\lambda_0(\mathcal{P})$, $\lambda_0(\mathcal{Z})$) differ for the two regimes, particularly those corresponding to inflation (the second element of each λ_0) which increase notably under the ZIRP, from a negative value under the normal regime to a positive value under the the ZIRP regime. In the benchmark estimation, given large standard errors with limited sample size, the remaining prices of risk coefficients are set equal to zero except for the (1,1) element of $\Lambda_1(\mathcal{P})$, that is, under the normal regime the prices of risk are allowed to change with short rate fluctuations.

Table 3 and Figure 3 here

Figure 3 shows how the yield-equation coefficients, that is the constant, short-rate, inflation, and output-gap coefficients in the yield equation, change against maturity (in

[9] Type II does not nest Type I factor dynamics. One may wonder whether Type II reduces to Type I when $\bar{\pi}$ is sufficiently negative (so that the inflation condition is always satisfied); however, this is not true since $u_\pi \sim N(0,1)$ and u_π can be $-\infty$.

quarters) under the normal regime (dashed black lines) and the ZIRP regime (red solid lines). The model-implied yields are expressed as the annualized rate in percent. The upward slopes of the constant coefficients represent the shapes of the average yield curves under the normal and ZIRP regimes. They imply that yield curves flatten on average under the ZIRP regime, consistent with the existing findings (e.g., Okina and Shiratsuka, 2004 and Oda and Ueda, 2007). The downward slopes of the short-rate coefficients imply that an increase in the short rate has a more positive impact on the shorter end of yield curves.

The bottom two charts in Figure 3 demonstrate how differently deflation and low growth contribute to lowering longer-term JGB yields between the normal and ZIRP regimes; the shapes of the inflation coefficients imply that the inflationary effect on the longer end of yield curves increases under the ZIRP; the shapes of the output-gap coefficients imply that growth effects on JGB yields weaken under the ZIRP. Quantitatively, the estimated results indicate that under the normal regime, 1-percent deflation lowers 10-year JGB yields by 14 basis points, and 1-percent output gap increase raises 10-year JGB yields by 7 basis points. On the other hand, under the ZIRP regime, 1-percent deflation lowers 10-year JGB yields by 23 basis points, and 1-percent output gap increase raises 10-year JGB yields by 2 basis points. A closer look at the recursive equations for $\mathbf{b}_n^{s_t}$ indicates that the shapes of inflation and output-gap coefficients are generated largely by the differences between macroeconomic factor coefficients (i.e., Φ^{s_t} and $\Sigma_{22}^{s_t}$) across regimes with persistent transition probabilities under $\mathbb{Q}$.

Lastly, I decompose the long term bond yields into the expectations and term premium components. Following the typical definition in the literature, the term premium of an n-

period bond yield is defined as the actual n-period bond yield minus the average expected future short-term interest rates (i.e., $\frac{1}{n}E_t\left\{\sum_{j=0}^{n-1} r_{1,t+j}\right\}$). To calculate the expectations components, I first obtain $1, 2, ..., n$ period forecasts of the future short-term interest rates at each quarter via two-regime three-variable VAR forecasting, I then use these forecasts to calculate the average expected future short-term interest rates. Figure 4 reports the model implied term premia of 10-year bonds, the corresponding averages of expected future short-term interest rates, and the actual yields. It indicates that term premia declined after the second ZIRP introduction (i.e., the QEP started in March 2001). It also indicates that the large bond yield decline in the early 1990s was driven by the expectations components, whereas that in the late 1990s was driven by both expectations and term premium components.

Figure 4 here

5 Conclusion

I construct a no-arbitrage affine term structure model with state-dependent regime shifts. In the model, the state vector depends on the current policy regime. As an application of the model, I examine how the JGB yields fluctuate with macroeconomic variables with endogenous monetary policy shifts that incorporate the key ZIRP features, specifically the zero lower bound and the Bank of Japan's forward guidance. I also analyze yield dynamics by decomposing JGB yields into the expectations and term-premium components.

The estimated results indicate that under the ZIRP, deflation plays a growing role in lowering JGB yields, especially on the long end of yield curves. On the other hand, output gaps' ability to raise bond yields weakens for all maturities. These results flag upward risks on the JGB yields when Japan exits from the ZIRP after satisfying the exit condition on inflation.

Looking forward, however, I believe it is important to understand not only "normal" bond yield responses to moderate inflation and economic growth, but also the channels that can steeply raise macroeconomic variables, especially inflation, and thus jeopardizing the JGB markets.

Acknowledgement

I thank Kazuo Ueda, Fumio Hayashi, Takatoshi Ito, Kosuke Aoki, Tatsuyoshi Okimoto, Hibiki Ichiue, Yoshihiro Yajima, Naoto Kunitomo, Iiboshi Hirokuni, Leah Brooks, and seminar participants at Bank of Japan, Hitotsubashi University, the 2012 Japan Economic Association Meeting, 13th Macro conference at Keio University, the University of Tokyo, and Yokohama National University for helpful comments on the earlier versions of this paper. I acknowledge funding from Japan Center for Economic Research and the Seimeikai Foundation.

References

[1] Ang, A., and G. Bekaert. 2002. Short Rate Nonlinearities and Regime Switches. *Journal of Economic Dynamics and Control* 26:1243–74.

[2] Ang, A, Bekaert, B, Wei, M., 2008. The Term Structure of Real Rates and Expected Inflation. *Journal of Finance*, American Finance Association, vol. 63(2), pages 797-

849, 04.

[3] Ang, A., Boivin, G., Dong, S., Loo-Kung, R, 2011. Monetary Policy Shifts and Term Structure. *Review of Economic Studies*, 78, 429-457.

[4] Ang, A., Piazzesi, M., Wei, M., 2006. What does the yield curve tell us about GDP growth? *Journal of Econometrics*, 131, 359-403.

[5] Bansal, R. and H. Zhou, 2002. Term Structure of Interest Rates with Regime Shifts. *Journal of Finance*, American Finance Association, vol. 57(5), pp. 1997-2043, October.

[6] Dai, Q., K. Singleton, and W. Yang, 2007. Regime Shifts in a Dynamic Term Structure Model of U.S. Treasury Bond Yields. *Review of Financial Studies*, Vol. 20, pp.1669-1706

[7] Fujiwara, I., 2006. Evaluating monetary policy when nominal interest rates are almost zero. *Journal of the Japanese and International Economies*, Elsevier, vol. 20(3), pp. 434-453, September.

[8] Hamilton, J. D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, Econometric Society, vol. 57(2), pages 357-84, March.

[9] Hamilton, J. D. and J. C. Wu, 2012. The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment. *Journal of Money, Credit and Banking*, Blackwell Publishing, vol. 44, pages 3-46, 02.

[10] Harvey, D.I., Leybourne, S.J., Newbold, P., 1998. Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16, 254–259.

[11] Hayashi, F., 2000. Econometrics, Princeton University Press December.

[12] Hördahl, P, O. Tristani, and D. Vestin, 2006. A Joint Econometric Model of Macroeconomic and Term Structure Dynamics. *Journal of Econometrics* 131, 405-444.

[13] Inoue, T. and Okimoto, T., 2008. Were there structural breaks in the effects of Japanese monetary policy? Re-evaluating policy effects of the lost decade. *Journal of the Japanese and International Economies*, Elsevier, vol. 22(3), pp. 320-342, September.

[14] Ichiue, H. and Ueno, Y. 2012. Monetary policy and the yield curve at zero interest. mimeo.

[15] Kim, D. H. and K. Singleton, 2012. Term structure models and the zero bound: An empirical investigation of Japanese yields. *Journal of Econometrics*. 170, 32–49.

[16] Oda, N. and K. Ueda, 2007. The Effects Of The Bank Of Japan's Zero Interest Rate Commitment And Quantitative Monetary Easing On The Yield Curve: A Macro-Finance Approach. *The Japanese Economic Review*, Japanese Economic Association, vol. 58(3), pages 303-328.

[17] Okina, K. and S. Shiratsuka, 2004. Policy Commitment and Expectation Formation: Japan's Experience Under Zero Interest Rates. *North American Journal of Economic and Finance*, Vol. 15, No. 1, pp. 75–100.

[18] Stock, J. H. and M. W. Watson, 2001. "Vector Autoregressions," *Journal of Economic Perspectives*, American Economic Association, vol. 15(4), pages 101-115, Fall.

[19] Ueda, K., 2012a. The Effectiveness of Non-Traditional Monetary Policy Measures: The Case of The Bank of Japan. *The Japanese Economic Review*, Japanese Economic Association, vol. 63(1), pages 1-22, 03.

[20] Ueda, K., 2012b. Japan's Deflation and the Bank of Japan's Experience with Nontraditional Monetary Policy. *Journal of Money, Credit and Banking*, Blackwell Publishing, vol. 44, pages 175-190, 02.

[21] Ugai, H., 2007. Effects of the Quantitative Easing Policy: A Survey of Empirical Analysis. *Monetary and Economic Studies*, Vol. 25, No. 1, Bank of Japan, pp. 1–47.

[22] Wright, J., 2011. Term Premia and Inflation Uncertainty: Empirical Evidence from an International Panel Dataset. *American Economic Review*, 101 June.

A A non-ATSM without Assumption 4

One may want to do away with Assumption 4 of fully priced regime risk. Under Assumptions 1-3, $p_1(\tilde{\mathbf{y}}, \tilde{s}) = -r(\tilde{\mathbf{y}})$ is conditionally normally distributed under $\mathbb{Q}$. So, (9) holds for $n = 1$. It can be rewritten as

$$\exp\left[p_2(\tilde{\mathbf{y}}, \tilde{s})\right] = \sum_{\tilde{s}} \rho^Q(\tilde{s}|\mathbf{y}, s) \exp\left[E^Q\left(p_1(\tilde{\mathbf{y}})|\tilde{s}, \mathbf{y}, s\right) + \frac{1}{2} Var^Q\left(p_1(\tilde{\mathbf{y}})|\tilde{s}, \mathbf{y}, s\right) - r(\mathbf{y})\right].$$

The term in the brackets can be rewritten as

$$
\begin{aligned}
& -E^Q\left(r\left(\tilde{\mathbf{y}}\right)|\tilde{s},\mathbf{y},s\right)+\frac{1}{2}Var^Q\left(p_1\left(\tilde{\mathbf{y}}\right)|\tilde{s},\mathbf{y},s\right)-r\left(\mathbf{y}\right), \\
= \quad & -\mathbf{e}_1'E^Q\left(\tilde{\mathbf{y}}|\tilde{s},\mathbf{y},s\right)+\frac{1}{2}\mathbf{e}_1'Var^Q\left(p_1\left(\tilde{\mathbf{y}}\right)|\tilde{s},\mathbf{y},s\right)\mathbf{e}_1-\mathbf{e}_1'\mathbf{y}, \\
= \quad & -\mathbf{e}_1'\left(\mathbf{c}^Q\left(\tilde{s}\right)+\mathbf{\Phi}^Q\left(\tilde{s}\right)\mathbf{y}\right)+\frac{1}{2}\mathbf{e}_1'\Sigma\left(\tilde{s}\right)\Sigma\left(\tilde{s}\right)'\mathbf{e}_1-\mathbf{e}_1'\mathbf{y},
\end{aligned}
$$

that is,

$$
p_2\left(\tilde{\mathbf{y}},\tilde{s}\right)=\log\left(\sum_{\tilde{s}}\rho^Q\left(\tilde{s}|\mathbf{y},s\right)\exp\left[-\mathbf{e}_1'\left(\mathbf{c}^Q\left(\tilde{s}\right)+\mathbf{\Phi}^Q\left(\tilde{s}\right)\mathbf{y}\right)+\frac{1}{2}\mathbf{e}_1'\Sigma\left(\tilde{s}\right)\Sigma\left(\tilde{s}\right)'\mathbf{e}_1-\mathbf{e}_1'\mathbf{y}\right]\right).
$$

The whole term structure can be generated by the no-arbitrage condition for $n = 2, 3, ...$, once the model parameters are estimated by using only p_1 and p_2.

B The likelihood function

B.1 Separating yield information from factor dynamics

The goal is to derive the likelihood of the data, i.e.,

$$
\mathcal{L}\equiv p\left(\mathbf{z}_0,...,\mathbf{z}_T,\mathbf{y}_1,...,\mathbf{y}_T,s_1,...,s_T|\mathbf{y}_0,s_0\right),
$$

which can be decomposed as

$$
\mathcal{L}\equiv\underbrace{f\left(\mathbf{z}_0,...,\mathbf{z}_T|\mathbf{y}_0,...,\mathbf{y}_T,s_0,...,s_T\right)}_{\mathcal{L}_1}\underbrace{p\left(\mathbf{y}_1,...,\mathbf{y}_T,s_0,...,s_T|\mathbf{y}_0,s_0\right)}_{\mathcal{L}_2}.
$$

B.2 Likelihood for yields ($\mathcal{L}_1$)

The model-implied static yield equation

$$\mathbf{z}_t = A^{s_t} + \mathbf{B}^{s_t}\mathbf{y}_t + \mathbf{v}_t,$$

where

$$A^{s_t} \equiv \begin{bmatrix} a_1^{s_t} \\ \vdots \\ a_{40}^{s_t}/40 \end{bmatrix}, \ \mathbf{B}^{s_t} \equiv \begin{bmatrix} \mathbf{b}_1(s_t) \\ \vdots \\ \mathbf{b}_{40}(s_t)/40 \end{bmatrix}.$$

The usual assumption that the error is iid normal can be stated precisely as

$$\left.\begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_T \end{bmatrix}\right| (\mathbf{y}_0, ..., \mathbf{y}_T, s_0, ..., s_T) \sim N(\mathbf{0}, \mathbf{I}_T \otimes \mathbf{V}), \ \ \mathbf{V} \equiv Var(\mathbf{v}_t).$$

So

$$\begin{aligned} \mathcal{L}_1 &\equiv f(\mathbf{y}_0, ..., \mathbf{y}_T, s_0, ..., s_T) = \prod g_1(\mathbf{z}_t - A^{s_t} - \mathbf{B}^{s_t}\mathbf{y}_t; \mathbf{V}), \\ \text{Or } L_1 &\equiv \log(\mathcal{L}_1) = \sum \log[g_1(\mathbf{z}_t - A^{s_t} - \mathbf{B}^{s_t}\mathbf{y}_t; \mathbf{V})], \end{aligned}$$

where g_1 is the density of $N(0, \mathbf{V})$. Furthermore, the log likelihood can be concnetreated (Hayashi (2000), eq. (8.5.23)) as follows

$$L_1 = const. - \frac{1}{2}\log\left|\frac{1}{T+1}\sum_{t=1}^{T}(\mathbf{z}_t - A^{s_t} - \mathbf{B}^{s_t}\mathbf{y}_t)(\mathbf{z}_t - A^{s_t} - \mathbf{B}^{s_t}\mathbf{y}_t)'\right|.$$

B.3 Likelihood for factor dynamics ($\mathcal{L}_2$)

Since $\{\mathbf{y}, s\}$ is Markov, the usual sequantial factorization argument yields

$$\mathcal{L}_2 \equiv p(\mathbf{y}_1, ..., \mathbf{y}_T, s_1, ...s_T | \mathbf{y}_0, s_0) = \prod_{t=1}^{T} p(\mathbf{y}_t, s_t | \mathbf{y}_{t-1}, s_{t-1}).$$

As mentioned in the text, $p(\mathbf{y}_t, s_t|\mathbf{y}_{t-1}, s_{t-1}) = f(\mathbf{y}_t|s_t, \mathbf{y}_{t-1}, s_{t-1})\,\rho(s_t|\mathbf{y}_{t-1}, s_{t-1})$. The component $f(\mathbf{y}_t|s_t, \mathbf{y}_{t-1}, s_{t-1})$ can be decomposed as

$$\begin{aligned}
f(\mathbf{y}_t|s_t, \mathbf{y}_{t-1}, s_{t-1}) &= f(r_t|\mathbf{y}_{2,t}, s_t, \mathbf{y}_{t-1}, s_{t-1}) \times f(\mathbf{y}_{2,t}|s_t, \mathbf{y}_{t-1}, s_{t-1}), \\
f(r_t|\mathbf{y}_{2,t}, s_t, \mathbf{y}_{t-1}, s_{t-1}) &= g_2\left(r_t - \alpha^{s_t} - \delta^{s_t} r_{t-1} - \boldsymbol{\beta}^{s_t\prime}\mathbf{y}_{2,t}; (\sigma_r^{s_t})^2\right), \\
f(\mathbf{y}_{2,t}|s_t, \mathbf{y}_t, s_t) &= g_3\left(\mathbf{y}_{2,t} - \mathbf{c}_2^{s_t} - \boldsymbol{\Phi}_2^{s_t}\mathbf{y}_{t-1}; \boldsymbol{\Sigma}_{22}^{s_t}\boldsymbol{\Sigma}_{22}^{s_t\prime}\right),
\end{aligned}$$

where g_2 is the density of $N\left(0, (\sigma_r^{s_t})^2\right)$ and g_3 is the density of $N(0, \boldsymbol{\Sigma}^{s_t}\boldsymbol{\Sigma}^{s_t\prime})$. The other component in (1) $\rho(s_t|\mathbf{y}_{t-1}, s_{t-1})$ was derived in the text. Putting all pieces about $\mathcal{L}_2$ together, $L_2 \equiv \log(\mathcal{L}_2)$ can be written as

Type I:
$$\begin{aligned}
L_2 = &\sum_{t=1}^{T} \log\left[g_2\left(r_t - \alpha^{s_t} - \delta^{s_t} r_{t-1} - \boldsymbol{\beta}^{s_t\prime}\mathbf{y}_{2,t}; (\sigma_r^{s_t})^2\right)\right] \\
&+ \sum_{t=1}^{T} \log\left[g_3\left(\mathbf{y}_{2,t} - \mathbf{c}_2^{s_t} - \boldsymbol{\Phi}_2^{s_t}\mathbf{y}_{t-1}; \boldsymbol{\Sigma}_{22}^{s_t}\boldsymbol{\Sigma}_{22}^{s_t\prime}\right)\right] \\
&+ \sum_{t=1}^{T} \left\{ s_t \log\left[F\left(\frac{r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right)\right] + (1-s_t)\log\left[1 - F\left(\frac{r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right)\right]\right\}.
\end{aligned} \tag{26}$$

Type II:
$$\begin{aligned}
L_2 = &\sum_{t=1}^{T} \log\left[g_2\left(r_t - \alpha^{s_t} - \delta^{s_t} r_{t-1} - \boldsymbol{\beta}^{s_t\prime}\mathbf{y}_{2,t}; (\sigma_r^{s_t})^2\right)\right] \\
&+ \sum_{t=1}^{T} \log\left[g_3\left(\mathbf{y}_{2,t} - \mathbf{c}_2^{s_t} - \boldsymbol{\Phi}_2^{s_t}\mathbf{y}_{t-1}; \boldsymbol{\Sigma}_{22}^{s_t}\boldsymbol{\Sigma}_{22}^{s_t\prime}\right)\right] \\
&+ \sum_{t=1}^{T} \left\{ \begin{array}{c}
s_{t-1}s_t \log\left[F\left(\frac{r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right)\right] + s_{t-1}(1-s_t)\log\left[1 - F\left(\frac{r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}}{\sigma_\xi}\right)\right] \\
+(1-s_{t-1})\,s_t \log\left[B\left(r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}, \pi^e(\mathbf{y}_{t-1}) - \bar{\pi}; \mathbf{W}\right)\right] \\
+(1-s_{t-1})(1-s_t)\log\left[1 - B\left(r^e(\mathbf{y}_{t-1}) - \alpha^{\mathcal{Z}}, \pi^e(\mathbf{y}_{t-1}) - \bar{\pi}; \mathbf{W}\right)\right]
\end{array}\right\}.
\end{aligned} \tag{27}$$

Maximization of L_2 can be simplified because $\mathbf{c}_2^{\mathcal{Z}}$, $\boldsymbol{\Phi}_2^{\mathcal{Z}}$,and $\boldsymbol{\Sigma}_{22}^{\mathcal{Z}}$ appear only in the g_3 component of the second summation in L_2.

B.4 Choices of parameters

The model parameters are

$$\begin{aligned}
\boldsymbol{\theta} &\equiv (\boldsymbol{\theta}_1, \boldsymbol{\theta}_2), \\
\boldsymbol{\theta}_1 &\equiv \left(\rho^Q(\mathcal{P}|\mathcal{P}), \rho^Q(\mathcal{Z}|\mathcal{Z}), \boldsymbol{\lambda}_0(\mathcal{P}), \boldsymbol{\lambda}_0(\mathcal{Z}), \boldsymbol{\Lambda}_1(\mathcal{P}), \boldsymbol{\Lambda}_1(\mathcal{Z})\right),
\end{aligned}$$

$$\text{Type I:} \qquad \boldsymbol{\theta}_2 \equiv \begin{pmatrix} \alpha^{\mathcal{P}}, \alpha^{\mathcal{Z}}, \boldsymbol{\beta}^{\mathcal{P}}, \delta^{\mathcal{P}}, \sigma_r^{\mathcal{P}}, \sigma_r^{\mathcal{Z}}, \\ \mathbf{c}_2^{\mathcal{P}}, \mathbf{c}_2^{\mathcal{Z}}, \boldsymbol{\Phi}_2^{\mathcal{P}}, \boldsymbol{\Phi}_2^{\mathcal{Z}}, \boldsymbol{\Sigma}_{22}^{\mathcal{P}}, \boldsymbol{\Sigma}_{22}^{\mathcal{Z}} \end{pmatrix},$$

$$\text{Type II:} \qquad \boldsymbol{\theta}_2 \equiv \begin{pmatrix} \alpha^{\mathcal{P}}, \alpha^{\mathcal{Z}}, \boldsymbol{\beta}^{\mathcal{P}}, \delta^{\mathcal{P}}, \sigma_r^{\mathcal{P}}, \sigma_r^{\mathcal{Z}}, \\ \mathbf{c}_2^{\mathcal{P}}, \mathbf{c}_2^{\mathcal{Z}}, \boldsymbol{\Phi}_2^{\mathcal{P}}, \boldsymbol{\Phi}_2^{\mathcal{Z}}, \boldsymbol{\Sigma}_{22}^{\mathcal{P}}, \boldsymbol{\Sigma}_{22}^{\mathcal{Z}}, \bar{\pi} \end{pmatrix}.$$

C Proofs for propositions 1 and 2

Proof for proposition 1 Pin down ρ^Q. Integrate both sides of (6) over $\widetilde{\mathbf{y}}$.

$$\begin{aligned}
LHS &= \int_{\widetilde{\mathbf{y}}} \exp[-r(\mathbf{y})] f^Q(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) \rho^Q(\tilde{s}|\mathbf{y}, s) \, d\widetilde{\mathbf{y}}, \\
&= \exp[-r(\mathbf{y})] \rho^Q(\tilde{s}|\mathbf{y}, s) \int_{\widetilde{\mathbf{y}}} f^Q(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) \, d\widetilde{\mathbf{y}}, \\
&= \exp[-r(\mathbf{y})] \rho^Q(\tilde{s}|\mathbf{y}, s) \quad (\text{since} \int_{\widetilde{\mathbf{y}}} f^Q(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}, s) \, d\widetilde{\mathbf{y}} = 1).
\end{aligned}$$

$$\begin{aligned}
RHS &= \int_{\widetilde{\mathbf{y}}} \mathcal{M}(\widetilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}) \rho(\tilde{s}|\mathbf{y}, s) \, d\widetilde{\mathbf{y}}, \\
&= \rho(\tilde{s}|\mathbf{y}, s) \int_{\widetilde{\mathbf{y}}} \mathcal{M}(\widetilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) f(\widetilde{\mathbf{y}}|\tilde{s}, \mathbf{y}) \, d\widetilde{\mathbf{y}}, \\
&= \rho(\tilde{s}|\mathbf{y}, s) E[\mathcal{M}(\widetilde{\mathbf{y}}, \tilde{s}, \mathbf{y}, s) | \tilde{s}, \mathbf{y}, s], \\
&= \rho(\tilde{s}|\mathbf{y}, s) \exp[-r(\mathbf{y}) - \gamma(\tilde{s}, \mathbf{y}, s)], \quad (\text{by equation (4)}) \\
&= \exp[-r(\mathbf{y})] \rho(\tilde{s}|\mathbf{y}, s) \exp[-\gamma(\tilde{s}, \mathbf{y}, s)].
\end{aligned}$$

Q.E.D.

Proof for proposition 2 Pin down f^Q. Divide both sides of equation (6) by the expression of ρ^Q in Proposition 1 and use Assumption 2 to obtain

$$f^Q(\widetilde{\mathbf{y}}|\tilde{s},\mathbf{y},s) = \exp\left[-\frac{1}{2}\boldsymbol{\lambda}(\tilde{s},\mathbf{y})'\boldsymbol{\lambda}(\tilde{s},\mathbf{y}) - \boldsymbol{\lambda}(\tilde{s},\mathbf{y})'\Sigma(\tilde{s})^{-1}(\widetilde{\mathbf{y}}-\boldsymbol{\mu}(\tilde{s},\mathbf{y}))\right] \times f(\widetilde{\mathbf{y}}|\tilde{s},\mathbf{y}).$$

So the conditional moment-generating function of $\tilde{\mathbf{y}}$ under $\mathbb{Q}$ can be written as

$$E^Q[\exp(\zeta'\widetilde{\mathbf{y}})|\tilde{s},\mathbf{y},s] = \int_{\tilde{\mathbf{y}}} \exp(\zeta'\widetilde{\mathbf{y}}) f^Q(\widetilde{\mathbf{y}}|\tilde{s},\mathbf{y})\,d\tilde{\mathbf{y}} = \int_{\tilde{\mathbf{y}}} \exp(X) f(\widetilde{\mathbf{y}}|\tilde{s},\mathbf{y})\,d\widetilde{\mathbf{y}} = E[\exp(X)|\tilde{s},\mathbf{y}],$$

where $X \equiv \zeta'\widetilde{\mathbf{y}} - \frac{1}{2}\boldsymbol{\lambda}'\boldsymbol{\lambda} - \boldsymbol{\lambda}'\boldsymbol{\Sigma}^{-1}(\widetilde{\mathbf{y}}-\boldsymbol{\mu})$ with $\boldsymbol{\mu}$ here being $\boldsymbol{\mu}(\tilde{s},\mathbf{y})$, $\boldsymbol{\lambda}$ being $\boldsymbol{\lambda}(\tilde{s},\mathbf{y})$, and $\boldsymbol{\Sigma}$ being $\boldsymbol{\Sigma}(\tilde{s})$. Since $\widetilde{\mathbf{y}}|\tilde{s},\mathbf{y} \sim N(\boldsymbol{\mu},\boldsymbol{\Sigma\Sigma}')$ under $\mathbb{P}$ by Assumption 1, the conditional distribution of X is normal with

$$E(X|\tilde{s},\mathbf{y}) = \zeta'\boldsymbol{\mu} - \frac{1}{2}\boldsymbol{\lambda}'\boldsymbol{\lambda},\ Var(X|\tilde{s},\mathbf{y}) = \left(\zeta - \boldsymbol{\Sigma}^{-1\prime}\boldsymbol{\lambda}\right)\boldsymbol{\Sigma\Sigma}'\left(\zeta - \boldsymbol{\Sigma}^{-1\prime}\boldsymbol{\lambda}\right).$$

Thus

$$E[\exp(X)|\tilde{s},\mathbf{y}] = \exp\left[E(X|\tilde{s},\mathbf{y}) + \frac{1}{2}(X|\tilde{s},\mathbf{y})\right] = \exp\left[\zeta'(\boldsymbol{\mu} - \boldsymbol{\Sigma\lambda}) + \frac{1}{2}\zeta\boldsymbol{\Sigma\Sigma}'\zeta\right],$$

which is the moment-generating function of a normal random variable with mean $\mu - \Sigma\lambda$ and variance $\Sigma\Sigma'$. Q.E.D.

Type I factor dynamics

Taylor rule

α^P	δ^P	$\boldsymbol{\beta}^{P'}$		σ_r^P
0.06	0.91	0.34	0.11	0.49
(0.098)	(0.029)	(0.168)	(0.041)	(0.040)

α^Z	σ_r^Z
0.01	0.02
(0.004)	(0.003)

Dynamics of macroeconomic variables

$\mathbf{c}_2^P$	$\boldsymbol{\Phi}_2^P$			$\boldsymbol{\Sigma}_{22}^P$	
0.17	-0.01	0.73	0.09	0.18	---
(0.044)	(0.012)	(0.070)	(0.017)	(0.016)	---
-0.22	0.06	-0.10	0.83	-0.05	0.93
(0.223)	(0.063)	(0.359)	(0.086)	(0.112)	(0.079)

$\mathbf{c}_2^Z$	$\boldsymbol{\Phi}_2^Z$			$\boldsymbol{\Sigma}_{22}^Z$	
0.01	-0.11	0.92	3.8E-03	0.06	---
(0.019)	(0.091)	(0.043)	(0.007)	(0.004)	---
-0.27	-1.36	1.39	0.77	-0.09	0.57
(0.191)	(0.903)	(0.422)	(0.069)	(0.059)	(0.042)

Type II factor dynamics

Taylor rule

α^P	δ^P	$\boldsymbol{\beta}^{P'}$		σ_r^P
0.20	0.90	0.22	0.10	0.50
(0.110)	(0.030)	(0.178)	(0.043)	(0.041)

α^Z	σ_r^Z
0.01	0.02
(0.004)	(0.003)

Dynamics of macroeconomic variables

$\mathbf{c}_2^P$	$\boldsymbol{\Phi}_2^P$			$\boldsymbol{\Sigma}_{22}^P$	
0.17	-0.01	0.72	0.09	0.18	---
(0.043)	(0.012)	(0.070)	(0.016)	(0.015)	---
-0.19	0.06	-0.15	0.82	-0.05	0.93
(0.224)	(0.063)	(0.360)	(0.086)	(0.113)	(0.080)

$\mathbf{c}_2^Z$	$\boldsymbol{\Phi}_2^Z$			$\boldsymbol{\Sigma}_{22}^Z$	
0.01	-0.11	0.92	3.8E-03	0.06	---
(0.019)	(0.091)	(0.043)	(0.007)	(0.004)	---
-0.27	-1.36	1.39	0.77	-0.09	0.57
(0.191)	(0.903)	(0.422)	(0.069)	(0.059)	(0.042)

$\overline{\pi}$
0.11
(0.104)

Table 1. The factor dynamics coefficients. This table reports the estimated coefficients of Type I (left panel) and Type II (right panel) factor dynamics.

	RMSE ratios	DM test
r	1.30	-1.33
π	0.77	2.89**
g	1.43	-1.52

Table 2. (Pseudo) out of sample performance. The second column reports the root mean square error (RMSE) ratios of the Type I factor dynamics relative to the Type II factor dynamics. The third column reports modified Diebold-Mariano test statistics. Significantly negative statistics indicate that the Type II specification outperforms the Type I specification. The out-of-sample period consists of the last three years of the QEP period. The superscript ** indicate significance at the 1% level.

	The prices of risk			**Transition probabilities under Q**	
	$\lambda_0(Z)$	$\lambda_0(P)$	$\Lambda_1(P)$	$\rho^Q(Z \mid Z)$	$\rho^Q(P \mid P)$
r	-15.50	2.93	-0.50	0.94	0.77
	(0.10)	(50.44)	(0.001)	(2.7e-6)	(4.9e-6)
π	10.81	-22.20			
	(0.56)	(0.25)			
g	-9.25	-4.78			
	(0.24)	(0.79)			

Table 3. Yield curve coefficients. This table reports the coefficients of the prices of risk and transition probabilities under **Q** estimated for the benchmark model.

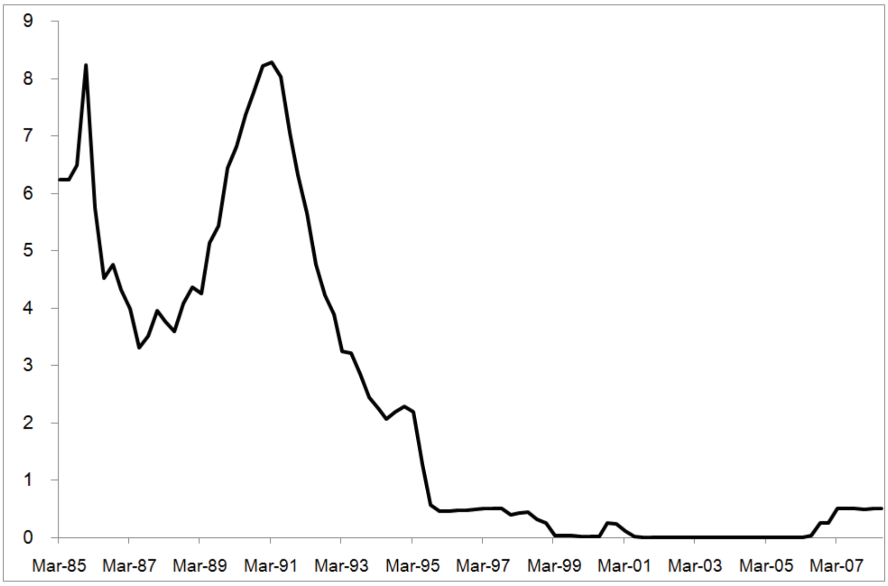

Figure 1. Uncollarterized overnight call rate in Japan (annualized rate in percent).

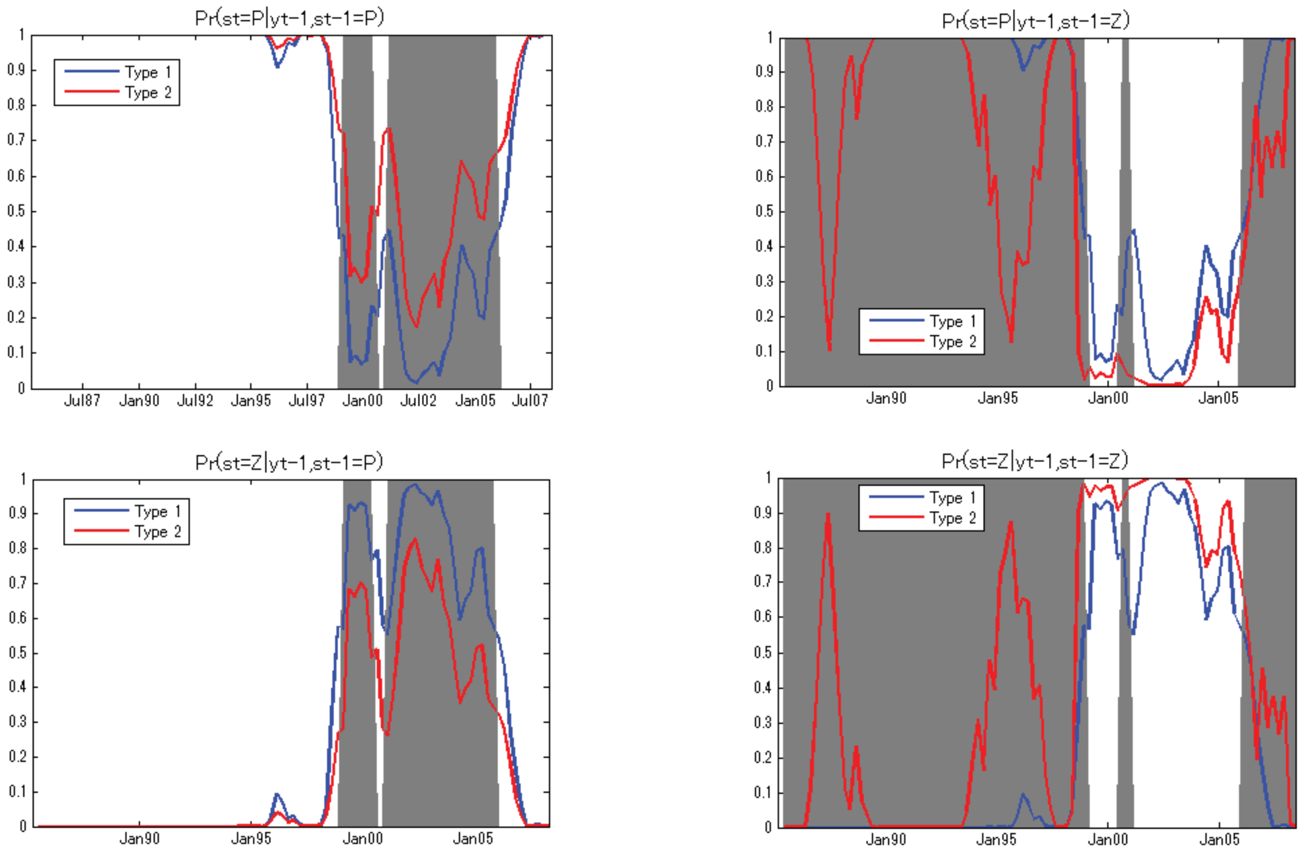

Figure 2. State-dependent transition probabilities. The left column reports the probability that the regime is normal (top) or a ZIRP (bottom) given that the previous regime is normal. The right column reports the probability that the regime is normal (top) or a ZIRP (bottom) given that the previous regime is ZIRP. The blue lines are transition probabilities under Type I evolution and red lines are those under Type II evolution. The periods that the current regime is not normal (left column) or a ZIRP (right column) are shaded in gray.

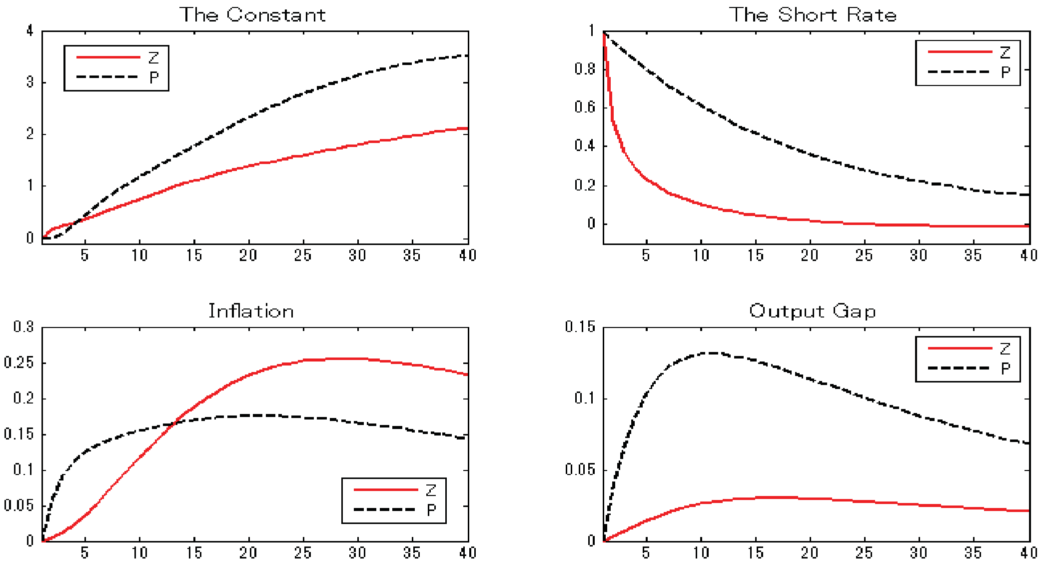

Figure 3. Factor weights against maturity. This figure plots the coefficients of the yield equation against maturity (in quarters) estimated for the benchmark model. The coefficients correspond to the constant, short-rate, inflation, and output-gap terms in the yield equation under the normal regime (dashed black lines) and the ZIRP regime (red solid lines). The model-implied yields are expressed as the annualized rate in percent.

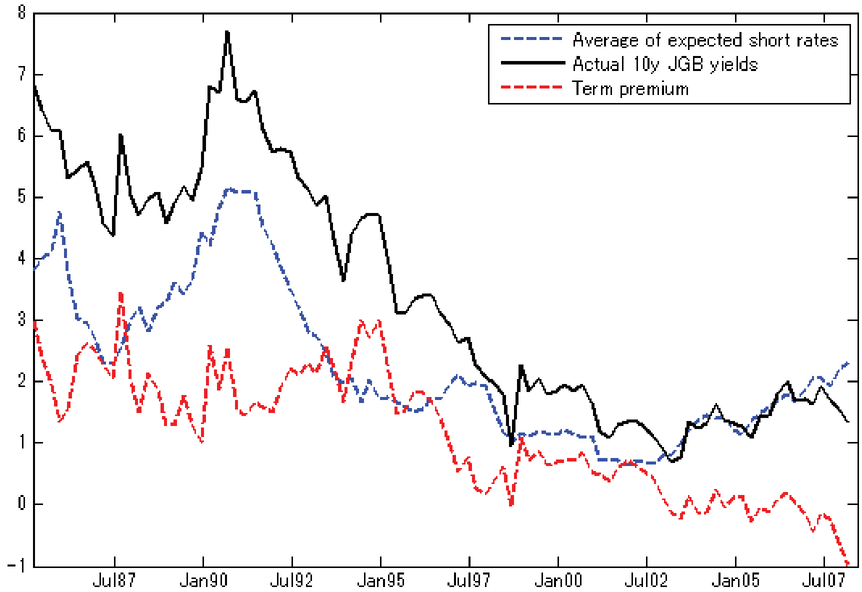

Figure 4. Estimation of expectations and term premium components of 10-year bond yields (annualized rates in percent). This figure plots the actual 10-year bond yields, the average expected future short-term interest rates over the life of the bond, and its difference from the actual yields (i.e., term premium) obtained via two-regime three-variable VAR forecasting for the benchmark model.