# Why Can the Yield Curve Predict Output Growth, Inflation, and Interest Rates? An Analysis with Affine Term Structure Model☆

Hibiki Ichiue*

Department of Economics, University of California, San Diego

The Bank of Japan

August, 2003

**Abstract**

The literature gives evidence that term spreads help predict output growth, inflation, and interest rates. This paper integrates and explains these predictability results by using an affine term structure model with observable macroeconomic factors. The results suggest that consumers are willing to pay a higher premium for output growth risk hedge during the higher inflation regime. This causes term spreads to react to recent inflation shocks, which prove useful for prediction. We also find that term spreads using the short end of the yield curve have less predictive power than many other spreads. We attribute this to monetary policy inertia.

JEL classification: E43; E52

Keywords: Term structure, Monetary policy, VAR

---

☆ I am especially grateful to my dissertation advisor, James D. Hamilton for his valuable supports and comments. I would like to thank Marjorie Flavin, Bruce Lehmann, Alan Timmermann, Keiichi Tanaka, and participants in the presentations at UCSD. I also thank Monica Piazzesi for answering my detail questions on her papers. The views expressed here are those of the author, and not necessarily of the Bank of Japan.

* Address: 2-1-1 Nihonbashi-Hongokucho Chuo-ku Tokyo 103-8660 Japan

Tel.: +81-3-3279-1111 (The Bank of Japan), Fax: +81-3-5255-6758, *E-mail address*: hibiki.ichiue@boj.or.jp

## 1. Introduction

Many studies in the literature give evidence that term spreads of interest rates have information about three different future economic variables: output growth, inflation, and interest rates, for various sample periods and countries. But the literatures examining the predictability of these three variables have been quite distinctive. Studies of the predictability of interest rates have been mainly conducted by financial economists by testing a very popular and classic theory, the expectations hypothesis[1]. In this theory, the long rate is equal to an average of expected future short rates plus a time-invariant term premium. However, in spite of its popularity, this hypothesis typically has been rejected. Many economists argue that this expectations hypothesis failure is attributable to the failure of the assumption of time-invariant term premium[2]. The literature on the predictability of inflation also has a long history following Fama's (1975) classic study[3]. On the other hand, the history of the literature studying the predictability of output growth is relatively recent. After Stock and Watson (1989) found that a term spread plays an important role in their index of economic leading indicators, many researchers investigated this predictability[4].

---

[1] For empirical results of testing the expectations hypothesis, see, for example, Campbell and Shiller (1991), Hardouvelis (1994), Rudebusch (1995), Campbell, Lo and MacKinlay (1997), Roberds and Whiteman (1999), Bekaert, Hodrick and Marshall (2001), and Cochrane (2001).

[2] The literature gives evidence that term premium is in fact time-varying. See, for example, Mankiw and Miron (1986), Engle, Lilien and Robins (1987), Engle and Ng (1993), Dotsey and Otrok (1995), and Balduzzi, Bertola and Foresi (1997).

[3] For empirical results of the predictability of inflation, see, for example, Mishkin (1988, 1990a, b, 1991), Fama (1990), Jorion and Mishkin (1991), Estrella and Mishkin (1997), and Kozicki (1997).

[4] For empirical results of the predictability of output growth or recession, see, for example, Estrella and Hardouvelis (1991), Plosser and Rouwenhost (1994), Haubrich and Dombrosky (1996), Bonser-Neal and Morley(1997), Dueker (1997), Estrella and Mishkin (1997), Kozicki (1997), Bernard and Gerlach (1998),

Although a huge literature gives evidence and explanations for each of the predictabilities of output growth, inflation, and interest rates, no paper tries to analyze the relationship among all of these three predictabilities. The main purpose of this paper is to integrate these predictability results in an attempt to answer to an important question: why can the term structure predict future movements in economic variables? This study will help us understand the information contained in the term structure of interest rates, and the relationship between the term structure and business cycle.

We use an affine term structure model (ATSM) with observable economic factors as our main tool. After Ang and Piazzesi (2003) introduced this type of model to investigate the relationship between macroeconomic variables and the term structure, the idea has been followed by several studies, for example, Dewachter and Lyrio (2002), Hordahl, Tristani and Vestin (2002), and Wu (2002). These studies depend much on macroeconomic theories to restrict their models so that the results can be interpreted more easily. Furthermore, these models typically use latent variables other than observable variables, and interpret the latent factors as variables such as the monetary policy authority's inflation target.

Conversely, Ang, Piazzesi and Wei (2003) use only observable variables, and they do not use macroeconomic theories other than the no-arbitrage assumption to restrict their model. This type of model can be interpreted as either a VAR with no-arbitrage restrictions or ATSM with observable factors obeying VAR. In this paper, we call this type of model VAR-ATSM for convenience. Ang, Piazzesi and Wei use their VAR-ATSM to examine the predictability of output growth using term spreads. We follow this basic idea, and extend to the predictabilities of not

---

Dotsey (1998), and Hamilton and Kim (2002).

only output growth but also inflation and short rates[5]. Although their basic idea is very useful for analyzing the predictabilities, some of their assumptions and estimation method are not suitable to our purpose. Ang, Piazzesi and Wei try to find good forecasting models by comparing predictive powers, especially rolling out-of-sample forecasting performances, of various combinations of regressors. For conducting this exercise, their parsimony VAR(1) model and computationally fast, but less efficient, estimation method may be appropriate. On the other hand, we try to reveal the source of the predictability by analyzing the relationship between impulse response functions and $R^2$'s. Thus we adopt VAR with more lags and more efficient estimation method, which contribute to the reliability of impulse response functions.

We have three main findings. First, the time-varying market price of output growth risk, which is sensitive to the level of inflation, plays a key role in the predictability. When the inflation rate is higher, consumers are willing to pay a higher premium for output growth risk hedge, which may be explained by a simple model with a money in the utility function and a monetary policy rule. This causes term spreads to react to recent inflation shocks. Since the inflation shock has persistent effects on not only inflation but also output growth and interest rates, the response of term spread to the inflation shock helps predict these variables. Second, we also find that term spreads using the short end of the yield curve have less predictive power than many spreads between longer rates. This fact is attributable to the inertial character of monetary policy. Third, it is hard to predict output growth with term spreads at short horizons, because the

[5] Before Ang, Piazzesi and Wei (2003), several papers use term structure models with only latent factors for analyzing predictability using term spreads. For example, Roberds and Whiteman (1999), Dai and Singleton (2002), and Duffee (2002) examine whether ATSM's can fit to the empirical results on predictability of interest rates. Hamilton and Kim (2002) use the Longstaff and Schwarz's (1992) term structure model to explain predictability of output growth. But since these models use only latent factors, the ability to analyze the relationship among term structure and macroeconomic variables is limited.

monetary policy shock affects output growth with a lag while the term structure responds to the shock immediately.

The rest of this paper is organized as follows. Section 2 displays stylized facts from simple OLS results. In Section 3, we consider simple ATSM's to understand basic properties of ATSM's. This section will help to prepare for the more complicated VAR-ATSM introduced in Section 4. Estimation methods and results are considered in Section 5. Here we discuss the relationship between time-varying market prices of risk and information included in term structure. In Section 6, we use impulse response functions and model-implied $R^2$'s, which can be obtained from the estimated VAR-ATSM, to explain why term spreads predict well. Section 7 concludes.

## 2. Simple OLS Results

The empirical studies in the literature examine predictabilities of term spreads for future output growth, inflation, and interest rates with a common econometric method, regressions on the term spreads. However, these regressions do not have exactly the same form. For example, Estrella and Mishkin (1991) examine output growth predictability by using regressions of cumulative output growth from $t$ to $t+h$ on a fixed term spread between ten-year and three-month interest rates:

$$g_{t\to t+h} = \alpha + \beta(r_t^{(10\mathrm{Y})} - r_t^{(3\mathrm{M})}) + \varepsilon_{t+h}\,. \tag{1}$$

On the other hand, Mishkin (1990a) examines inflation predictability by using regressions of difference between $h$-period and 1-year cumulative inflation rates on maturity matching term spreads:

$$\pi_{t \to t+h} - \pi_{t \to t+1\text{Y}} = \alpha + \beta(r_t^{(h)} - r_t^{(1\text{Y})}) + \varepsilon_{t+h} . \tag{2}$$

Campbell and Shiller (1991) give evidence for short rate predictability by using the most popular expectations hypothesis test, regressions of average future short rate changes on maturity matching term spreads:

$$\frac{1}{h}\sum_{i=0}^{h-1}(r_{t+i}^{(1)} - r_t^{(1)}) = \alpha + \beta(r_t^{(h)} - r_t^{(1)}) + \varepsilon_{t+h-1} . \tag{3}$$

All three types of studies find that the slope coefficient $\beta$ is significantly different from zero in many cases, which means that term spreads have predictive power for movements in macroeconomic variables. Typically they report substantial $t$-stats and $R^2$'s for these regressions. As one can easily see, these empirical regressions do not have the same form. For example, (1) and (2) do not use the same regressor. (1) uses a fixed regressor, while the regressor of (2) depends on the forecasting horizon $h$. So for analyzing the relationship among the predictabilities, we need to put the empirical results for predicting the different variables on a consistent basis.

For this purpose, we use the regressions below,

$$g_{t+h} = \alpha + \beta(r_t^{(n)} - r_t^{(m)}) + \varepsilon_{t+h} ; \tag{4}$$

$$\pi_{t+h} = \alpha + \beta(r_t^{(n)} - r_t^{(m)}) + \varepsilon_{t+h}; \tag{5}$$

$$r_{t+h}^{(1)} = \alpha + \beta(r_t^{(n)} - r_t^{(m)}) + \varepsilon_{t+h}; \tag{6}$$

for various combinations of $h$, $n$, and $m$ ($h$ = 1,2,…,12; $n$, $m$ = 2, 4, 8, 12, 16, 20, and $n > m$), where $g_t$ is the real GDP growth rate from $t$-1 to $t$, $\pi_t$ is the inflation rate of GDP deflator from $t$-1 to $t$, and $r_t^{(n)}$ is the $n$-period discount rate of Treasury bills or bonds at end of $t$[6]. Quarterly data are used, so we interpret one period as one quarter. All of $g_t$, $\pi_t$, and $r_t^{(n)}$ are defined as rates per quarter. The sample period is 1964:1Q-2001:4Q, following Fama and Bliss (1987) who comment that long rate data before 1964 may be unreliable. There are two other properties of the set of regressions (4)-(6) worth commenting on. First, regressands are continuously compounded marginal rates or one-period short rate. Since cumulative rates are the averages of marginal rates, marginal rates are more convenient for specifying which part of future the term spreads can predict well. Second, we use various forecasting horizons $h$ and term spreads $r_t^{(n)} - r_t^{(m)}$, so we can specify which components of the yield curve predict at which future horizons.

Figures 1 and 2 display the $t$-stats and $R^2$'s of OLS regressions (4)-(6) for selected term spreads. 20Q-1Q spread has significant predictive power for all of output growth, inflation, and

[6] We use discount rate data from CRSP (Center for Research in Security Prices, Graduate School of Business, the University of Chicago: www.crsp.uchicago.edu. All rights reserved.) Monthly US Treasury Database with permission. We can construct discount rates for 1, 2, 4, 8, 12, 16, 20 quarters from the CRSP data. The 1 quarter (3 month) rate is obtained from average rates in the CRSP risk free rates file. The 2 quarter (6 month) rate is constructed by multiplying average-YTM by 12 (there is no data on 9/30/1987, so we interpolate with 3 and 12 month rates). The other rates are obtained from the Fama-Bliss discount bonds file.

short rates at least for shorter horizons. This result is consistent with the literature, which argues that term spread between 5-year (or 10-year) and 3-month rates predict well. But surprisingly we found that most term spreads without 1Q rate are better than the 20Q-1Q spread in many cases. For example, Figure 2 shows that 12Q-8Q spread is better except for predicting output growth rates at shorter horizons. In addition, 2Q-1Q spread, which also uses the 1Q rate, is almost useless. These facts seem to imply that term spreads using the short end of the yield curve reduce the predictabilities. This is surprising because the literature does not care about the spread without the short end of term structure so much, and several studies including Ang, Piazzesi and Wei (2003) argue that the maximal maturity difference is the best predictor. As another notable feature of the graphs, we can see that the $R^2$'s of output growth regressions are hump-shaped. That is, it is difficult to predict output growth rate at short horizons.

Why can term spreads predict the future in such ways? Since OLS results do not answer this question, we need a more structured model. A useful method to interpret these OLS results is proposed by Ang, Piazzesi and Wei (2003). They use a VAR-ATSM to represent the model-implied $R^2$'s for the regressions of output growth rates and compare predictive powers of various combinations of regressors. We follow their basic idea, and extend their methods to explain predictabilities of all of output growth, inflation, and short rates. Although their VAR-ATSM is very useful for examining the relationship among macroeconomic variables and the yield curve, some of their assumptions and estimation method are not suitable to our purpose. So we modify them in Section 4 and 5. Then, in Section 6, we try to reveal the source of the predictability by using impulse response functions and $R^2$'s, which can be calculated from the estimates of the VAR-ATSM.

## 3. Simple Affine Term Structure Models with Observable Factors

Before introducing our VAR-ATSM in the next section, let's consider four simpler ATSM's. Since the VAR-ATSM is too complicated to give simple interpretations, we should start from these simpler models. In particular, time-varying market prices of risk, which many classic term structure models assume constant, are the source of the complication. But since they affect the relationship between short and long rates, i.e. movements in term spreads, they are very important for examining the predictabilities of term spreads.

### *3.1. An ATSM with One Factor of Short Rate*

Suppose that quarterly data of short (3-month) rate $r_t^{(1)}$ are characterized by an AR(1) process:

$$r_{t+1}^{(1)} = c_r + \phi_r r_t^{(1)} + \sigma_r u_{r,t+1}, \tag{7}$$

where $u_{r,t+1} \sim N(0,1)$ i.i.d., and $\sigma_r > 0$. Table 1 reports the OLS estimates of (7), which show that the short rate is persistent ($\phi_r = 0.9037$). Suppose that the stochastic discount factor $M_{t+1}$ obeys a conditional log-normal distribution:

$$M_{t+1} = \exp\left( -r_t^{(1)} - \frac{1}{2}\lambda_{r,t}{}^2 - \lambda_{r,t} u_{r,t+1} \right), \tag{8}$$

where

$$\lambda_{r,t} = \gamma_r + \delta_r r_t^{(1)} . \tag{9}$$

So in this model, the market price of risk $\lambda_{r,t}$ is time-varying, depending on the factor $r_t^{(1)}$. That is, the stochastic discount factor $M_{t+1}$ is affected by not only the exogenous shock $u_{r,t+1}$ but also the level of the factor $r_t^{(1)}$ through the time-varying market price of risk. Thus the effects of the factor on the yield curve are complicated. Note that if $\delta_r = 0$, i.e. $\lambda_{r,t}$ is time-invariant, this is just the classic Vasicek (1977) model.

Let's assume there is no arbitrage opportunity in the Treasury market. Since this market is one of the largest and most highly liquid markets in the world, the no-arbitrage assumption is extremely reasonable. Under the no-arbitrage assumption, we can use the fundamental asset pricing equation for bond prices,

$$q_t^{(n)} = E_t[M_{t+1} q_{t+1}^{(n-1)}], \tag{10}$$

for $n$ = 1, 2, …, and all $t$, where $q_t^{(n)}$ is the $n$-period bond price. Note that from (8) and (10),

$$\begin{aligned} q_t^{(1)} &= E_t[M_{t+1} q_{t+1}^{(0)}] \\ &= E_t[M_{t+1}] \\ &= \exp(-r_t^{(1)}) . \end{aligned} \tag{11}$$

This is exactly the definition of the relationship between the 1-period bond price and continuously compounded discount rate. In fact, $M_{t+1}$ is chosen so that (11) holds.

By using the fundamental asset pricing equation (10), we can derive closed forms for discount rates $r_t^{(n)}$ as affine functions of the factor $r_t^{(1)}$:

$$\hat{r}_t^{(n)} = a^{(n)} + b^{(n)} r_t^{(1)}, \qquad n = 1, 2, \ldots \tag{12}$$

where

$$a^{(n)} = -A^{(n)} / n, \quad b^{(n)} = -B^{(n)}/n, \tag{13}$$

$$A^{(n+1)} = A^{(n)} + B^{(n)}(c_r - \sigma_r \gamma_r) + \frac{1}{2}\sigma_r{}^2 B^{(n)2}, \tag{14}$$

$$B^{(n+1)} = B^{(n)}(\phi_r - \sigma_r \delta_r) - 1, \tag{15}$$

$$A^{(1)} = 0, \quad B^{(1)} = -1 \text{ [7]}. \tag{16}$$

From (12), the factor loading on the short rate factor $b^{(n)}$ can be interpreted as the sensitivity of longer rates $r_t^{(n)}$ to the short rate $r_t^{(1)}$. From (13), (15), and (16), we can obtain a closed form for $b^{(n)}$:

[7] Since this is one of the simplest special cases of VAR-ATSM, it is enough to check the proof for the general model introduced in Section 4. For the proof, see Ang and Piazzesi (2003).

$$b^{(n)} = \frac{1}{n}\sum_{j=0}^{n-1}(\phi_r - \sigma_r\delta_r)^j \ . \tag{17}$$

Note that $\gamma_r$ does not appear in (17). Since movement of short rates is less volatile than that of long rates, it is reasonable that the absolute value of $b^{(n)}$ decreases as *n* increases. For satisfying this, we need parameter values such that

$$|\,\phi_r - \sigma_r\delta_r\,| < 1. \tag{18}$$

From (18),

$$-\frac{1-\phi_r}{\sigma_r} < \delta_r < \frac{1+\phi_r}{\sigma_r}. \tag{19}$$

Since point estimates in Table 1 imply $(1+\phi_r)/\sigma_r \approx 334$, $\delta_r$ can be even hundreds.

From (17), we can say that the sensitivity of long rates to the short rate is weaker, when $\delta_r$ is higher. We can relate this claim with the expectations hypothesis. From (7),

$$\begin{aligned} r_{t+j}^{(1)} &= c_r + \phi_r r_{t+j-1}^{(1)} + \sigma_r u_{r,t+j} \\ &= c_r + \phi_r(c_r + \phi_r r_{t+j-2}^{(1)} + \sigma_r u_{r,t+j-1}) + \sigma_r u_{r,t+j} \\ &\vdots \\ &= c_r \sum_{i=0}^{j-1}\phi_r^{\,i} + \phi_r^{\,j} r_t^{(1)} + \sigma_r \sum_{i=0}^{j-1}\phi_r^{\,i} u_{r,t+j-i}\,. \end{aligned} \tag{20}$$

So by taking the expectation,

$$E_t[r_{t+j}^{(1)}] = c_r \sum_{i=0}^{j-1} \phi_r^{\ i} + \phi_r^{\ j} r_t^{(1)} . \tag{21}$$

Then, from (12), (17) and (21), we can obtain the term premium:

$$r_t^{(n)} - \frac{1}{n}\sum_{j=0}^{n-1} E_t[r_{t+j}^{(1)}] = a^{(n)} - c_r \frac{1}{n}\sum_{j=0}^{n-1}\sum_{i=0}^{j-1} \phi_r^{\ i} + \frac{1}{n}\sum_{j=0}^{n-1}[(\phi_r - \sigma_r \delta_r)^j - \phi_r^{\ j}] r_t^{(1)} . \tag{22}$$

So the term premium is constant, i.e. the expectation hypothesis holds only when $\delta_r = 0$. In this case, the movements of long rates $r_t^{(n)}$ depend only on those of average expected short rates $n^{-1}\sum_{j=0}^{n-1} E_t[r_{t+j}^{(1)}]$. Since $r_t^{(1)}$ obeys a persistent AR(1) process, an increase in $r_t^{(1)}$ raises $r_t^{(n)}$. However, when $\delta_r > 0$, a rise in $r_t^{(1)}$ also has a negative effect on $r_t^{(n)}$ through a decrease in the term premium. Therefore, positive $\delta_r$ weakens the relationship between short and long rates. Then the sensitivity of the term spread to the factor $r_t^{(1)}$ is stronger when $\delta_r$ is larger.

*3.2. A one factor ATSM with a constant short rate*

Let's consider a model with constant short rate $r^{(1)}$ and one factor $x_t$ obeying AR(1):

$$x_{t+1} = c_x + \phi_x x_t + \sigma_x u_{x,t+1} , \tag{23}$$

where $u_{x,t+1} \sim N(0,1)$ i.i.d., and $\sigma_x > 0$. Suppose that the stochastic discount factor obeys

$$M_{t+1} = \exp\left(-r^{(1)} - \frac{1}{2}\lambda_{x,t}{}^2 - \lambda_{x,t}u_{x,t+1}\right) \tag{24}$$

where

$$\lambda_{x,t} = \gamma_x + \delta_x x_t \,. \tag{25}$$

By using the fundamental asset pricing equation (10), we can derive expressions for discount rates $r_t^{(n)}$:

$$\hat{r}_t^{(n)} = r^{(1)}, \qquad n = 1, 2, \ldots \tag{26}$$

That is, when short rate is constant, yield curve is always perfectly flat. More importantly, the factor $x_t$ can not affect the yield curve, even if the exogenous shock $u_{x,t+1}$ has a strong effect on the stochastic discount factor $M_{t+1}$. This implies that the stochastic factor $x_t$ affects bond prices only through the movements in short rates. So we can not conclude whether the effect of the factor on the yield curve is strong or not only from the market price of risk.

*3.3. C-CAPM*

Let's consider a simple C-CAPM, in which the stochastic discount factor obeys

$$M_{t+1} = \delta \frac{u'(C_{t+1})}{u'(C_t)} \exp(-\pi_{t+1}), \tag{27}$$

with CRRA utility function

$$u(C_t) = \frac{C_t^{1-\rho}}{1-\rho}, \tag{28}$$

where $\delta$ is the subjective discount factor, $C_t$ is consumption at *t*, and $\rho > 0$ is the coefficient of relative risk aversion. Suppose in equilibrium, the consumption $C_t$ is equal to the output $Y_t$ so that the consumption growth rate $g_{C,t+1} = g_{t+1}$ [8]. Then (27) can be rewritten as

$$\begin{aligned} M_{t+1} &= \delta \left( \frac{C_t}{C_{t+1}} \right)^{\rho} \exp(-\pi_{t+1}) \\ &= \delta \exp(-\rho g_{C,t+1} - \pi_{t+1}) \\ &= \delta \exp(-\rho g_{t+1} - \pi_{t+1}) \\ &= \delta \exp(-\rho \{E_t[g_{t+1}] + \varepsilon_{g,t+1}\} - \{E_t[\pi_{t+1}] + \varepsilon_{\pi,t+1}\}) \\ &= \exp(\log(\delta) - \rho E_t[g_{t+1}] - E_t[\pi_{t+1}] - \rho \sigma_{\varepsilon,g} u_{g,t+1} - \sigma_{\varepsilon,\pi} u_{\pi,t+1}). \end{aligned} \tag{29}$$

[8] We assume this just for simplicity, and we can generalize this model to be consistent with the literature, which shows that dynamics of consumption growth rate is smoother than that of output growth rate, by assuming that consumption growth rate obeys an affine function of output growth rate with a positive and less than unity slope coefficient. Even with this generalized assumption, the main properties of the model do not change.

where $\varepsilon_{g,t+1} = \sigma_{\varepsilon,g} u_{g,t+1} = g_{t+1} - E_t[g_{t+1}]$, $\varepsilon_{\pi,t+1} = \sigma_{\varepsilon,\pi} u_{\pi,t+1} = \pi_{t+1} - E_t[\pi_{t+1}]$, and $\sigma_{\varepsilon,g}$ and $\sigma_{\varepsilon,\pi}$ are standard deviations of $\varepsilon_{g,t+1}$ and $\varepsilon_{\pi,t+1}$ so that $u_{g,t+1} \sim N(0,1)$ and $u_{\pi,t+1} \sim N(0,1)$.

Suppose $u_{g,t+1}$ and $u_{\pi,t+1}$ are uncorrelated as we often observe empirically. Since $\rho > 0$, a positive output growth shock has a negative effect on $M_{t+1}$. This is consistent with a role of bonds for consumption hedge. That is, when future output growth rate is higher, consumers feel that future cash flows are less important. Note that both of the market prices of risk corresponding to the output growth shock $u_{g,t+1}$ and inflation shock $u_{\pi,t+1}$ are constant ($\rho\sigma_{\varepsilon,g}$ and $\sigma_{\varepsilon,\pi}$ respectively).

From (10), (11), and (29),

$$\begin{aligned} \exp(-r_t^{(1)}) &= q_t^{(1)} \\ &= E_t[M_{t+1}] \\ &= \exp(\log(\delta) - \rho E_t[g_{t+1}] - E_t[\pi_{t+1}] + \{\rho^2 \sigma_{\varepsilon,g}{}^2 + \sigma_{\varepsilon,\pi}{}^2\}/2). \end{aligned} \tag{30}$$

So we can obtain

$$r_t^{(1)} = \log(\frac{1}{\delta}) + \rho E_t[g_{t+1}] + E_t[\pi_{t+1}] - \frac{\rho\sigma_g{}^2 + \sigma_\pi{}^2}{2}. \tag{31}$$

That is, this C-CAPM implicitly assumes (31) holds. This suggests that the short rate is higher when expected output growth and inflation rates are higher. An interesting special case of (31) is

$$g_{t+1} = \pi_{t+1} = 0, \tag{32}$$

without uncertainty. In this case, (31) is

$$r_t^{(1)} = \log(\frac{1}{\delta}), \tag{33}$$

which means that the short rate is equal to the subjective discount rate.

*3.4. C-CAPM with money-in-the-utility function*

Let's consider another C-CAPM in which we replace (28) with a money-in-the-utility (MIU) function $u(C_t, m_t)$ where $m_t$ is the real money holding. Even with the MIU function, the stochastic discount factor obeys the same form as (27):

$$M_{t+1} = \delta \frac{u'(C_{t+1}, m_{t+1})}{u'(C_t, m_t)} \exp(-\pi_{t+1}), \tag{34}$$

where $u'(C, m) \equiv \partial u(C, m)/\partial C$.

Suppose that the form of the utility function is

$$u(C_t, m_t) = \frac{C_t^{1-\rho}}{1-\rho} m_t^{\theta}, \tag{35}$$

where $\rho > 0$ and $0 < \theta < 1$. Then if $C_t = Y_t$ as before, (34) can be rewritten as

$$
\begin{aligned}
M_{t+1} &= \delta \left( \frac{Y_t}{Y_{t+1}} \right)^{\rho} \left( \frac{m_{t+1}}{m_t} \right)^{\theta} \exp(-\pi_{t+1}) \\
&= \delta \exp(-\rho g_{t+1} + \theta \mu_{t+1} - \pi_{t+1}) \\
&= \exp(\log(\delta) - E_t[\rho g_{t+1} - \theta \mu_{t+1} + \pi_{t+1}] - \rho \sigma_{\varepsilon,g} u_{g,t+1} + \theta \varepsilon_{\mu,t+1} - \sigma_{\varepsilon,\pi} u_{\pi,t+1}) ,
\end{aligned} \tag{36}
$$

where $\mu_{t+1}$ is the real money growth rate from *t* to *t*+1, and $\varepsilon_{\mu,t+1} = \mu_{t+1} - E_t[\mu_{t+1}]$.

Let's consider a case in which $\varepsilon_{\mu,t+1}$ can be represented as a linear combination of $u_{g,t+1}$ and $u_{\pi,t+1}$ with time-varying weights:

$$\varepsilon_{\mu,t+1} = w_{g,t} u_{g,t+1} + w_{\pi,t} u_{\pi,t+1} , \tag{37}$$

where the weights $w_{g,t}$ and $w_{\pi,t}$ are affine functions of $g_t$ and $\pi_t$:

$$w_{g,t} = \omega_g + \omega_{gg} g_t + \omega_{g\pi} \pi_t , \tag{38}$$

$$w_{\pi,t} = \omega_\pi + \omega_{\pi g} g_t + \omega_{\pi\pi} \pi_t . \tag{39}$$

The idea behind (37) is similar as Taylor's rule. But (37) uses real money growth rate instead of target short rate, and has time-varying weights. The time-varying weights can be interpreted, for example, as follows. Suppose that the monetary policy authority (the Fed) can observe $u_{g,t+1}$ and

$u_{\pi,t+1}$ before their decisions, by which they can perfectly control the real money growth rate $\mu_{t+1}$ (i.e. $\varepsilon_{\mu,t+1}$). When output growth rate surprisingly increases ($u_{g,t+1} > 0$), the Fed may accommodate the real money growth rate to an increase in money demand caused by the output growth shock. Conversely, the Fed may suppress the real money growth rate in responses to the shock, if they consider that this output growth shock may cause serious inflation in the future. These two plausible stories imply that the weight on the output growth shock $w_{g,t}$ can be either positive or negative. Equation (38) implies that the weight depends on $g_t$ and $\pi_t$. We can also discuss the weight on the inflation shock $w_{\pi,t}$ in a similar way.

With (37)-(39), (36) can be rewritten as

$$\begin{aligned} M_{t+1} = \exp(\log(\delta) - E_t[\rho g_{t+1} - \theta\mu_{t+1} + \pi_{t+1}] \\ -[(\rho\sigma_{\varepsilon,g} - \theta\omega_g) - \theta\omega_{gg} g_t - \theta\omega_{g\pi}\pi_t]u_{g,t+1} \\ -[(\sigma_{\varepsilon,\pi} - \theta\omega_\pi) - \theta\omega_{\pi g} g_t - \theta\omega_{\pi\pi}\pi_t]u_{\pi,t+1}) \ . \end{aligned} \tag{40}$$

Now in contrast with the simple C-CAPM discussed in previous subsection, the market prices of risk corresponding to $u_{g,t+1}$ and $u_{\pi,t+1}$ are time-varying, depending on $g_t$ and $\pi_t$. Similarly to (31), we can obtain

$$\begin{aligned} r_t^{(1)} = \log(\frac{1}{\delta}) + E_t[\rho g_{t+1} - \theta\mu_{t+1} + \pi_{t+1}] \\ - \frac{[(\rho\sigma_{\varepsilon,g} - \theta\omega_g) - \theta\omega_{gg} g_t - \theta\omega_{g\pi}\pi_t]^2 + [(\sigma_{\varepsilon,\pi} - \theta\omega_\pi) - \theta\omega_{\pi g} g_t - \theta\omega_{\pi\pi}\pi_t]^2}{2} \ . \end{aligned} \tag{41}$$

Although this type of MIU functions is often used in the literature, the validity of this theoretical model is under criticisms. The utility function may not depend on money directly. The time-separable utility function may be unreasonable due to, for example, habit formation. In Section 4, we will introduce a more general and less restricted model, which nests all of four models discussed in Section 3.

## 4. The VAR-ATSM

Now let's introduce the VAR-ATSM used for later analyses. This type of model is used by Ang, Piazzesi and Wei (2003) to examine the predictability of output growth rate using term spreads. We use the VAR-ATSM for examining the predictabilities of not only output growth, but also inflation and short rates. The VAR-ATSM can be interpreted as either a VAR model with no-arbitrage restrictions or ATSM with observable factors obeying VAR. Let's start by considering the VAR of factors.

We use four variables: output growth rate $g_t$, inflation rate $\pi_t$, short rate $r_t^{(1)}$, and a benchmark term spread $s_t$ as factors. As $s_t$, we use the term spread between ten-year Treasury bond YTM at end of quarter *t* and $r_t^{(1)}$. These four macroeconomic variables are assumed to obey VAR(4),

$$\mathbf{x}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{x}_{t-1} + \mathbf{\Phi}_2 \mathbf{x}_{t-2} + \mathbf{\Phi}_3 \mathbf{x}_{t-3} + \mathbf{\Phi}_4 \mathbf{x}_{t-4} + \boldsymbol{\varepsilon}_t , \tag{42}$$

where $\mathbf{x}_t = (g_t, \pi_t, r_t^{(1)}, s_t)'$ and $\boldsymbol{\varepsilon}_t = (\varepsilon_{g,t}, \varepsilon_{\pi,t}, \varepsilon_{r,t}, \varepsilon_{s,t})'$. Following the VAR literature, let's interpret $r_t^{(1)}$ as a proxy for the monetary policy instrument. Ang, Piazzesi and Wei (2003) use a simpler model than ours. They use a VAR with only one lag and three variables which do not include inflation rate. But the VAR literature usually uses at least four lags for quarterly data, and indicates that the inflation rate plays an important role. So we follow the VAR literature to generalize Ang, Piazzesi and Wei's model.

To give a structural interpretation to the VAR, we need identifying assumptions. We use a recursive structure with the variables ordered as $(g_t, \pi_t, r_t^{(1)}, s_t)$. That is,

$$\boldsymbol{\varepsilon}_t = \boldsymbol{\Sigma}\mathbf{u}_t \tag{43}$$

where exogenous shocks $\mathbf{u}_t = (u_{g,t}, u_{\pi,t}, u_{r,t}, u_{s,t})' \sim N(\mathbf{0}, \mathbf{I})$ i.i.d., and $\boldsymbol{\Sigma}$ is lower-triangular with positive diagonal elements. Since it is not plausible that $g_t$ and $\pi_t$ respond to contemporaneous interest rates, we order them before $r_t^{(1)}$ and $s_t$. The order between $g_t$ and $\pi_t$ should not have serious effects on the empirical results, since the correlation between $\varepsilon_{g,t}$ and $\varepsilon_{\pi,t}$ is small as shown later. But the correlation between $\varepsilon_{r,t}$ and $\varepsilon_{s,t}$ is too large to be ignored. For identifying the last two exogenous shocks $u_{r,t}$ and $u_{s,t}$, typically we need adopt one of two assumptions: the short rate (the monetary policy authority) does not respond to the term spread (bond market) contemporarily, or vice versa. Since we often observe that long rates move immediately after changes in monetary policy, the second assumption seems to be unreasonable. On the other hand, there is no clear evidence that the monetary policy authority responds to the bond market contemporaneously. Moreover the literature gives evidence for the

Fed's inertial behavior, in which the Fed's responses to new information tend to delay. Thus we adopt the first assumption[9]. As will be seen in Section 6, the impulse responses to estimated monetary policy shock $u_{r,t}$ and spread shock $u_{s,t}$ seem to be reasonable, and support our recursive assumption. With this ordering, each component of $\mathbf{u}_t$ can be interpreted as the exogenous shock to each corresponding variable. We call them, output growth, inflation, monetary policy, and spread shocks. Now we may interpret the first three rows of the system (42) as IS curve, Phillips curve, and monetary policy rule. The last row can be interpreted as endogenous response function of bond market.

We can rewrite the VAR in (42) into companion form,

$$\begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_{t-1} \\ \mathbf{x}_{t-2} \\ \mathbf{x}_{t-3} \end{bmatrix} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{\Phi}_1 & \mathbf{\Phi}_2 & \mathbf{\Phi}_3 & \mathbf{\Phi}_4 \\ \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{t-1} \\ \mathbf{x}_{t-2} \\ \mathbf{x}_{t-3} \\ \mathbf{x}_{t-4} \end{bmatrix} + \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{44}$$

or

$$\mathbf{X}_t = \tilde{\mathbf{c}} + \tilde{\mathbf{\Phi}}\mathbf{X}_{t-1} + \tilde{\mathbf{\Sigma}}\tilde{\mathbf{u}}_t, \tag{45}$$

where $\mathbf{X}_t = (g_t, \pi_t, r_t^{(1)}, s_t, \ldots, g_{t-3}, \pi_{t-3}, r_{t-3}^{(1)}, s_{t-3})'$ is the 16×1 state vector.

The stochastic discount factor is defined as

[9] Most studies in the VAR literature using both short and long rates choose the first assumption. For example, Leeper, Sims, and Zha (1996) discuss this issue in detail, and conclude that the first assumption is less harmful than the second one.

$$
\begin{aligned}
M_{t+1} &= \exp\left(-r_t^{(1)} - \frac{1}{2}\boldsymbol{\lambda}_t{}'\boldsymbol{\lambda}_t - \boldsymbol{\lambda}_t{}'\mathbf{u}_{t+1}\right) \\
&= \exp\left(-r_t^{(1)} - \frac{1}{2}\boldsymbol{\lambda}_t{}'\boldsymbol{\lambda}_t - \lambda_{g,t}u_{g,t+1} - \lambda_{\pi,t}u_{\pi,t+1} - \lambda_{r,t}u_{r,t+1} - \lambda_{s,t}u_{s,t+1}\right),
\end{aligned}
\tag{46}
$$

where $\boldsymbol{\lambda}_t = (\lambda_{g,t}, \lambda_{\pi,t}, \lambda_{r,t}, \lambda_{s,t})'$ is the market prices of risk. The vector $\boldsymbol{\lambda}_t$ is an affine function of the current economic variables $\mathbf{x}_t = (g_t, \pi_t, r_t^{(1)}, s_t)'$:

$$
\boldsymbol{\lambda}_t = \boldsymbol{\gamma} + \boldsymbol{\delta}\mathbf{x}_t, \tag{47}
$$

for a 4×1 vector $\boldsymbol{\gamma}$ and a 4×4 matrix $\boldsymbol{\delta}$. Equation (46) is a generalization of the examples of stochastic discount factors considered in Section 3. For example, if we restrict the last two elements in $\boldsymbol{\gamma}$ and all elements in $\boldsymbol{\delta}$ to be equal to zero, we obtain the same stochastic discount factor as the simple C-CAPM introduced in the previous section.

By using the fundamental asset pricing equation (10), we can obtain closed forms for $r_t^{(n)}$ [10]:

$$
\hat{r}_t^{(n)} = a^{(n)} + \mathbf{b}^{(n)}{}'\mathbf{X}_t, \qquad n = 1, 2, \ldots \tag{48}
$$

[10] We derive the closed forms for discount rates so that a restriction $\hat{r}_t^{(1)} = r_t^{(1)}$ holds. Since we can derive YTM's and $s_t$ from the discount rates, we could also restrict the model-implied spread $\hat{s}_t$ to be equal to $s_t$. But since there may be measurement error of $s_t$, we do not use this restriction.

where

$$a^{(n)} = -A_n / n, \quad \mathbf{b}^{(n)} = -\mathbf{B}^{(n)} / n, \tag{49}$$

$$A^{(n+1)} = A^{(n)} + \mathbf{B}^{(n)}{}'(\tilde{\mathbf{c}} - \tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\gamma}}) + \frac{1}{2}\mathbf{B}^{(n)}{}'\tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\Sigma}}'\mathbf{B}^{(n)}, \tag{50}$$

$$\mathbf{B}^{(n+1)}{}' = \mathbf{B}^{(n)}{}'(\tilde{\boldsymbol{\Phi}} - \tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\delta}}) - \mathbf{e}_3', \tag{51}$$

$$A^{(1)} = 0, \quad \mathbf{B}^{(1)}{}' = -\mathbf{e}_3', \tag{52}$$

$$\tilde{\boldsymbol{\gamma}} = \begin{bmatrix} \boldsymbol{\gamma} \\ \mathbf{0} \end{bmatrix} \text{ and } \tilde{\boldsymbol{\delta}} = \begin{bmatrix} \boldsymbol{\delta} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \tag{53}$$

$\mathbf{e}_j$ is the $j$ th column of the 16×16 identity matrix.

From (49), (51) and (52), we can obtain

$$\mathbf{b}^{(n)}{}' = \frac{1}{n}\mathbf{e}_3'\sum_{j=0}^{n-1}(\tilde{\boldsymbol{\Phi}} - \tilde{\boldsymbol{\Sigma}}\tilde{\boldsymbol{\delta}})^j \ . \tag{54}$$

This is a quite similar form to (17), and again the term premium is constant only when $\boldsymbol{\delta} = \mathbf{0}$.

# 5. Estimation

*5.1. Estimation Methods*

The VAR-ATSM has 98 parameters consisting of 78 from the VAR ( $\mathbf{c}$ ,

$\mathbf{\Phi} \equiv [\mathbf{\Phi}_1 \, \mathbf{\Phi}_2 \, \mathbf{\Phi}_3 \, \mathbf{\Phi}_4]$, and $\mathbf{\Sigma}$) and 20 in market prices of risk ($\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$). We use GMM to estimate all parameters simultaneously[11]. Moment conditions are constructed by assuming that three types of errors are orthogonal to instruments. The first type of the errors are the errors of the VAR,

$$\boldsymbol{\varepsilon}_t = \mathbf{x}_t - (\mathbf{c} + \mathbf{\Phi}_1 \mathbf{x}_{t-1} + \mathbf{\Phi}_2 \mathbf{x}_{t-2} + \mathbf{\Phi}_3 \mathbf{x}_{t-3} + \mathbf{\Phi}_4 \mathbf{x}_{t-4}), \tag{55}$$

with instruments a constant, $\mathbf{x}_{t-1}$, $\mathbf{x}_{t-2}$, $\mathbf{x}_{t-3}$, and $\mathbf{x}_{t-4}$. The second type is the error of the covariance matrix of the VAR,

$$\boldsymbol{\xi}_t = \mathrm{vech}(\mathbf{\Sigma\Sigma}' - \boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t'). \tag{56}$$

We assume that the sample mean of $\boldsymbol{\xi}_t$ is exactly equal to zero. Note that the moment conditions corresponding to (55) and (56) are exactly same as OLS. The third type is the pricing errors of discount rates

$$\mathbf{v}_t = [v_t^{(2)} \; v_t^{(4)} \; v_t^{(8)} \; v_t^{(12)} \; v_t^{(16)} \; v_t^{(20)}]' \tag{57}$$

[11] Ang, Piazzesi and Wei (2003) use two-step estimation, in which the VAR parameters are estimated by OLS and then given these point estimates, $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are estimated by minimizing the sum of squared pricing errors of discount rates. This estimation method has an advantage of less computational burden over our one-step estimation. On the other hand, since their estimation method does not use efficient weights on moment conditions, this is less efficient than ours. In particular, their estimates for VAR parameters can not have the efficiency gains from the no-arbitrage assumption at all. Since our later analyses are based on impulse response functions calculated from the estimates of VAR parameters, the efficiency gains are crucial.

where

$$
\begin{aligned}
v_t^{(n)} &= r_t^{(n)} - \hat{r}_t^{(n)} \\
&= r_t^{(n)} - (a^{(n)} + \mathbf{b}^{(n)}{}'\mathbf{X}_t)
\end{aligned} \quad (58)
$$

We use as instruments a constant, $\mathbf{x}_{t-1}$, and $\mathbf{x}_{t-2}$ for this type of moment. Now we have 132 moment conditions, which are sufficient for identifying 98 parameters. We use the sample period 1964:1Q-2001:4Q, the same as was used for the OLS regressions in Section 2.

We restrict the parameter space by two types of restrictions. First, the modulus of eigenvalues of $\tilde{\mathbf{\Phi}}$ are restricted to less than unity. Since the state vector $\mathbf{X}_t$ follows the VAR(1) of (45) with the autocorrelation coefficient matrix $\tilde{\mathbf{\Phi}}$, this restriction guarantees the stationarity of $\mathbf{X}_t$. In fact, estimation results show that this restriction does not bind. Second, the modulus of eigenvalues of $\tilde{\mathbf{\Phi}} - \tilde{\mathbf{\Sigma}}\tilde{\boldsymbol{\delta}}$ are restricted to be less than or equal to unity. From (54), the factor loading $\mathbf{b}^{(n)}$ can be considered as the average of $\mathbf{e}_3{}'(\tilde{\mathbf{\Phi}} - \tilde{\mathbf{\Sigma}}\tilde{\boldsymbol{\delta}})^j$; $j$ = 0, 1,…, $n$-1. So this second restriction guarantees the factor loading not to diverge with the maturity $n$. Note that this restriction is the generalization of (18). In our estimation results, only one of the restrictions binds[12].

[12] When a restriction binds, the spectral density matrix at frequency zero is not guaranteed to be the optimal weighting matrix in GMM. For solving this problem, we use the binding restriction to substitute out a parameter in advance. Then we use the obtained non-restricted GMM to estimate parameters with correct inference. The estimate and standard error of the substituted parameter are obtained by substituting out another parameter and re-estimating.

*5.2. Estimation Results*

The VAR estimates have great efficiency gains from the no-arbitrage assumption, although point estimates are not so different from OLS results. 42 out of 68 estimates for $\mathbf{c}$ and $\mathbf{\Phi}$ (not reported) are significantly different from zero at size of 5%, while OLS without the no-arbitrage assumption gives only 17 significant estimates. These efficiency gains contribute to the reliability of impulse response functions used later.

The estimate of $\mathbf{\Sigma}$ is reported in Table 2. The diagonal elements of $\mathbf{\Sigma}$ are much higher than the others in general, which implies that correlations among the reduced VAR errors are small, but the contemporaneous effect of short rate shock $u_{r,t}$ on the term spread $s_t$ is too large to be ignored. The output growth shock has the largest volatility, and this is about three times as large as the second largest volatility, that for the inflation shock.

Table 3 reports the estimates for $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$. Seven out of 16 estimates of $\boldsymbol{\delta}$ are significantly different from zero at size of 5%. This result supports the time-variation of the market prices of risk, depending on economic variables. Among these significant parameters, the (1,1) and (1,2) elements of $\boldsymbol{\delta}$, $\delta_{11}$ and $\delta_{12}$ are most influential on the movement of term structure. The reason for this is as follows. Given the factors $\mathbf{X}_t$, the movement of term structure depends only on the factor loadings $\mathbf{b}^{(n)}$, which depend on $\tilde{\mathbf{\Phi}} - \tilde{\mathbf{\Sigma}}\tilde{\boldsymbol{\delta}}$ from (54). So the influence of $\boldsymbol{\delta}$ on the movement of term structure depends on $\tilde{\mathbf{\Sigma}}$ (i.e. $\mathbf{\Sigma}$). As we can see in Table 2, the (1,1) element of $\mathbf{\Sigma}$, the volatility of output growth shock, is much larger than the others. So the first row of $\boldsymbol{\delta}$ is most influential. Among the estimates in the first row, only $\delta_{11}$ and $\delta_{12}$ are significantly different from zero. In fact, as we will discuss in the next section, $\delta_{12}$ plays a key role in the predictabilities, while $\delta_{11}$ does not.

The positive sign of $\delta_{12}$ implies that, when inflation rate $\pi_t$ is higher, $\lambda_{g,t}$ is higher and bond holders are willing to pay a higher premium for output growth risk hedge, which results in a lower term premium. Why do they pay the higher premium during the higher inflation regime? A possible explanation can be obtained in the framework of the C-CAPM with MIU function discussed in subsection 3.4. Although this C-CAPM has only output and inflation shocks, we can generalize this model to be consistent with the VAR-ATSM by adding monetary policy and spread shocks into (37) and letting the time-varying weights on shocks depend on all four VAR variables. From (40), $\delta_{12} = -\theta\omega_{g\pi}$. So since $\theta > 0$, $\delta_{12} > 0$ implies $\omega_{g\pi} < 0$. This means that, when the inflation $\pi_t$ is high, the weight on output growth shock $w_{g,t}$ is small and the Fed is less accommodating toward the output growth shock. This result makes sense if the Fed considers that the output growth shock during high inflation regime tends to cause serious future inflation. According to this consideration, when inflation is high, the Fed tends to suppress the real money growth rate in responses to the output growth shock. This less accommodating response of the Fed reduces the correlation between output growth shock $u_{g,t+1}$ and the real money shock $\varepsilon_{\mu,t+1}$. This reduced correlation causes future marginal utility,

$$u(C_{t+1}, m_{t+1}) = C_{t+1}{}^{-\rho} m_{t+1}{}^{\theta}, \tag{59}$$

to be more sensitive to the output growth shock, that is, bonds are more valuable for consumption hedge. Therefore, consumers are willing to pay more premium for holding bonds during the higher inflation regime. We can also discuss the positive sign of $\delta_{11}$ in a similar way.

Finally, the J-test supports our estimates with high p-value of 1.0000[13]. For more evaluation of the estimation results, let's compare the model-implied discount rates $\hat{r}_t^{(n)} = a^{(n)} + \mathbf{b}^{(n)}\mathbf{X}_t$ and the sample rates $r_t^{(n)}$. Table 4 reports means and standard deviations of $r_t^{(n)}$ and $\hat{r}_t^{(n)}$, and correlations between them for $n$ = 2, 4, 8, 16, 20. Since they have very similar values for means and standard deviations and the correlations are close to unity, we can conclude that $\hat{r}_t^{(n)}$ approximates $r_t^{(n)}$ very well.

## 6. Impulse Response Functions and the Predictabilities of Term Spreads

In the previous section, we obtained estimates for our VAR-ATSM with great efficiency gains from the no-arbitrage assumption. Let's use this model to examine the predictabilities of term spreads.

From the VAR-ATSM, we can calculate the optimal forecasts conditional on 16 state variables in $\mathbf{X}_t$. However, our main interest is not the forecasts conditional on these large numbers of variables, but on a term spread alone as the regressions (4)-(6) use. For our purpose, in subsection 6.1, we first consider the relationship between impulse response functions of variables in regressions (4)-(6) and the $R^2$'s. Since both regressands and regressors of the regressions can be represented as affine functions of $\mathbf{X}_t$, we can calculate the impulse response functions and the $R^2$'s from parameters in the VAR-ATSM. The considerations for the relationship between the impulse response functions and the $R^2$'s will be used for clarifying the

[13] The p-value is calculated from the J-stat (3.2313) and the degree of freedom (23 = 122 - 98 - 1). Note that since the one of the restrictions on eigenvalues binds, 1 should be subtracted form the degree of freedom.

source of predictabilities in subsection 6.2.

*6.1. Impulse Response Functions and Model-Implied $R^2$'s*

Since $\mathbf{x}_t = (g_t, \pi_t, r_t^{(1)}, s_t)'$ obeys the VAR in (42), we can calculate their impulse response functions, and represent the system in MA($\infty$) form with identified exogenous shocks. For example, $g_t$ can be represented as

$$g_t = \bar{g} + \sum_{j=0}^{\infty} \psi_{gg,j} u_{g,t-j} + \sum_{j=0}^{\infty} \psi_{g\pi,j} u_{\pi,t-j} + \sum_{j=0}^{\infty} \psi_{gr,j} u_{r,t-j} + \sum_{j=0}^{\infty} \psi_{gs,j} u_{s,t-j} , \qquad (60)$$

where $\bar{g}$ is the unconditional mean of $g_t$, and impulse response functions $\psi_{gg,j}$, $\psi_{g\pi,j}$, $\psi_{gr,j}$, and $\psi_{gs,j}$ are functions of $\mathbf{\Phi}$ and $\mathbf{\Sigma}$. So the future output growth $g_{t+h}$ can be represented as

$$g_{t+h} = \hat{g}_{t+h|t} + \sum_{j=0}^{h-1} \psi_{gg,j} u_{g,t+h-j} + \sum_{j=0}^{h-1} \psi_{g\pi,j} u_{\pi,t+h-j} + \sum_{j=0}^{h-1} \psi_{gr,j} u_{r,t+h-j} + \sum_{j=0}^{h-1} \psi_{gs,j} u_{s,t+h-j} \qquad (61)$$

where

$$\hat{g}_{t+h|t} = \bar{g} + \sum_{j=h}^{\infty} \psi_{gg,j} u_{g,t+h-j} + \sum_{j=h}^{\infty} \psi_{g\pi,j} u_{\pi,t+h-j} + \sum_{j=h}^{\infty} \psi_{gr,j} u_{r,t+h-j} + \sum_{j=h}^{\infty} \psi_{gs,j} u_{s,t+h-j} \qquad (62)$$

is the optimal forecast of $g_{t+h}$ conditional on $\mathbf{X}_t$.

Since discount rates $r_t^{(n)} = a^{(n)} + \mathbf{b}^{(n)}{}'\mathbf{X}_t$ and term spreads $r_t^{(n)} - r_t^{(m)}$ are affine functions of $\mathbf{X}_t = (\mathbf{x}_t', \mathbf{x}_{t-1}', \mathbf{x}_{t-2}', \mathbf{x}_{t-3}')'$, we can also calculate their impulse response functions, and represent them in MA($\infty$) form. For example, $r_t^{(n)} - r_t^{(m)}$ can be represented as

$$r_t^{(n)} - r_t^{(m)} = \overline{r^{(n)} - r^{(m)}} + \sum_{j=0}^{\infty} \kappa_{g,j}^{(n,m)} u_{g,t-j} + \sum_{j=0}^{\infty} \kappa_{\pi,j}^{(n,m)} u_{\pi,t-j} + \sum_{j=0}^{\infty} \kappa_{r,j}^{(n,m)} u_{r,t-j} + \sum_{j=0}^{\infty} \kappa_{s,j}^{(n,m)} u_{s,t-j}, \tag{63}$$

where $\overline{r^{(n)} - r^{(m)}}$ is the unconditional mean of $r_t^{(n)} - r_t^{(m)}$, and impulse response functions $\kappa_{g,j}^{(n,m)}$, $\kappa_{\pi,j}^{(n,m)}$, $\kappa_{r,j}^{(n,m)}$ and $\kappa_{s,j}^{(n,m)}$ are functions of $\mathbf{\Phi}$, $\mathbf{\Sigma}$, and $\boldsymbol{\delta}$.

Since $\mathbf{u}_t \sim N(\mathbf{0}, \mathbf{I})$ i.i.d., we can calculate the unconditional variances of VAR variables, the optimal forecasts of them, and term spreads. From (60), (62) and (63),

$$\begin{aligned} \sigma_g{}^2 &\equiv \mathrm{var}(g_t) \\ &= \sum_{j=0}^{\infty} \psi_{gg,j}{}^2 + \sum_{j=0}^{\infty} \psi_{g\pi,j}{}^2 + \sum_{j=0}^{\infty} \psi_{gr,j}{}^2 + \sum_{j=0}^{\infty} \psi_{gs,j}{}^2, \end{aligned} \tag{64}$$

$$\begin{aligned} \sigma_{\hat{g},h}{}^2 &\equiv \mathrm{var}(\hat{g}_{t+h|t}) \\ &= \sum_{j=h}^{\infty} \psi_{gg,j}{}^2 + \sum_{j=h}^{\infty} \psi_{g\pi,j}{}^2 + \sum_{j=h}^{\infty} \psi_{gr,j}{}^2 + \sum_{j=h}^{\infty} \psi_{gs,j}{}^2, \end{aligned} \tag{65}$$

$$\begin{aligned} (\sigma^{(n,m)})^2 &\equiv \mathrm{var}(r_t^{(n)} - r_t^{(m)}) \\ &= \sum_{j=0}^{\infty} \kappa_{g,j}^{(n,m)2} + \sum_{j=0}^{\infty} \kappa_{\pi,j}^{(n,m)2} + \sum_{j=0}^{\infty} \kappa_{r,j}^{(n,m)2} + \sum_{j=0}^{\infty} \kappa_{s,j}^{(n,m)2}. \end{aligned} \tag{66}$$

Similarly we can calculate the correlations among these variables. The correlation between future output growth $g_{t+h}$ and the current term spread $r_t^{(n)} - r_t^{(m)}$ can be represented as

$$\begin{aligned}\operatorname{corr}(g_{t+h}, r_t^{(n)} - r_t^{(m)}) &= \frac{\operatorname{cov}(g_{t+h}, r_t^{(n)} - r_t^{(m)})}{\sigma_g \sigma^{(n,m)}} \\ &= \sum_{j=0}^{\infty} \frac{\psi_{gg,j+h} \kappa_{g,j}^{(n,m)}}{\sigma_g \sigma^{(n,m)}} + \sum_{j=0}^{\infty} \frac{\psi_{g\pi,j+h} \kappa_{\pi,j}^{(n,m)}}{\sigma_g \sigma^{(n,m)}} + \sum_{j=0}^{\infty} \frac{\psi_{gr,j+h} \kappa_{r,j}^{(n,m)}}{\sigma_g \sigma^{(n,m)}} + \sum_{j=0}^{\infty} \frac{\psi_{gs,j+h} \kappa_{s,j}^{(n,m)}}{\sigma_g \sigma^{(n,m)}} .\end{aligned} \tag{67}$$

Since the forecasting error of the optimal forecast $g_{t+h} - \hat{g}_{t+h|t}$ is unpredictable by any variable known at time *t* such as $r_t^{(n)} - r_t^{(m)}$,

$$\operatorname{corr}(g_{t+h}, r_t^{(n)} - r_t^{(m)}) = \operatorname{corr}(\hat{g}_{t+h|t}, r_t^{(n)} - r_t^{(m)}) . \tag{68}$$

By squaring the correlation, we can obtain the $R^2$. For example, the $R^2$ of the regression (4) can be represented as

$$R^{2(n,m)}_{g,h} = \operatorname{corr}(\hat{g}_{t+h|t}, r_t^{(n)} - r_t^{(m)})^2 . \tag{69}$$

Since the $R^2$'s are functions of parameters in our VAR-ATSM, we can calculate the $R^2$'s with the estimates of the parameters. We call them the model-implied $R^2$'s. Equation (69) implies that if $r_t^{(n)} - r_t^{(m)}$ is a good predictor for future output growth $g_{t+h}$, $r_t^{(n)} - r_t^{(m)}$ should have similar responses to exogenous shocks as $\hat{g}_{t+h|t}$ has. We investigate this by looking at the variance decomposition of $\hat{g}_{t+h|t}$ in the next subsection. Finally, as we can see from (67)-(69), the $R^2$'s depend on the sum of products of impulse response functions for regressands and regressors. Note that, in (67), indexes for $\psi$'s start from $t+h$, not $t$, because future shocks $\mathbf{u}_{t+1},\ldots,\mathbf{u}_{t+h}$ are unpredictable. This implies that since $\psi$'s typically decay with the horizon $j$, $r_t^{(n)}-r_t^{(m)}$ is a good predictor if this responds to recent shocks well, i.e. $\kappa$'s are large for smaller $j$.

*6.2. Why do term spreads help predict?*

Figure 3 displays the model-implied $R^2$'s of the regressions (4)-(6) for three selected term spreads, and is the model-calculated analog of Figure 2. The results show that the model-implied $R^2$'s replicate three properties of the sample $R^2$'s in Figure 2 very well. First, the 12Q-8Q spread is better than the 20Q-1Q spread except for output growth predictions at shorter horizons. Second, the 2Q-1Q spread is almost useless. Finally, it is difficult to predict output growth at 1Q ahead. Therefore it is reasonable to try to explain the sample $R^2$'s in Figure 2 in terms of the factors that determine the model-implied $R^2$'s in Figure 3. Since the model-implied $R^2$'s are functions of parameters in our VAR-ATSM, we can analyze how these parameters affect the $R^2$'s.

Figure 4 shows impulse response functions of VAR variables $g_t$, $\pi_t$, $r_t^{(1)}$, and $s_t$ to one unit exogenous shocks. These are based on the estimates from the restricted GMM estimation of the VAR-ATSM. In general, these results are consistent with those in the VAR literature. For example, (4-a) and (4-b) show that the short rate, the instrument of the monetary policy authority, responds positively to output growth and inflation shocks. Panel (4-c) demonstrates that the estimated monetary policy shock sharply reduces output growth. This shock also suppresses inflation rates in the long run. These reasonable results imply reasonable estimates of the monetary policy shock. Further support is provided by Panel (4-d). As we discussed in Section 3, the most questionable part of our identification strategy may come from the contamination between the monetary policy shock and the spread shock. Panel (4-d) indicates that the estimated spread shock raises output growth and suppresses inflation. Since the output growth and inflation should respond to a monetary policy shock in the same direction, the results in (4-d) suggest that the spread shock is not measuring a change in monetary policy.

Figure 5 shows variance decompositions of the optimal forecasts, where the variances of the forecasts such as (65) are normalized to unity. As discussed in the previous subsection, this indicates which exogenous shocks should be useful for prediction. (5-a) shows that the output growth shock dominates output growth predictability at one quarter ahead. Then around 2-4 quarter ahead, the monetary policy shock is the most important. The importance of the inflation shock increases with the forecasting horizon, and at last this shock is most influential at 12 quarters ahead. These results are consistent with the impulse response functions in Figure 4. The output growth shock causes a sharp jump of output growth only in the short run. The monetary policy shock has negative effects on output growth with 2-4 quarter lags. But in the long run, the inflation shock raises the short rate persistently, which continues to suppress output growth. Panels (5-b) and (5-c) show that the inflation shock is most important for predicting inflation and short rates at most horizons. Accordingly, how the term spreads respond to the inflation shock is most important for specifying the sources of the predictabilities especially at longer horizons. Note that, as Figure 4 implies, the effects of exogenous shocks decay with the horizon. So we can also say that good predictors should respond to recent shocks rather than old shocks.

Figure 6 shows impulse response functions of selected discount rates. There are three notable features. First, the inflation shock has very persistent effects on levels of discount rates. That is, the discount rates do not return to zero even after 40 quarters. Since good predictors should respond to recent shocks, this is an important reason why levels of yield curves do not have great predictive power.

Second, discount rates with different maturities display different responses to recent shocks, while they respond to old shocks in similar ways. This implies that most movements in term spreads are due to recent shocks, because old shocks shift the yield curve almost in parallel. In fact, the upper graphs of Figure 7 display that both the 20Q-1Q and 12Q-8Q spreads depend much on recent shocks. This is a reason why the term spreads have predictive powers.

Why do the discount rates respond in such ways? We find that the time-varying market price of risk plays important roles as follows. As discussed in Section 5, the parameters corresponding to the effects of the output growth and inflation rates on the market price of output growth risk $\delta_{11}$ and $\delta_{12}$ are most influential on the movement of long rates. In fact, only $\delta_{12}$ plays a supportive role in the predictability. As shown in Figure 5, the inflation shock is most important for the predictability, and the positive $\delta_{12}$ causes the market price of output growth risk to respond positively to the shock well. On the other hand, the positive $\delta_{11}$ makes the predictability even worse. As shown in (4-b), the positive inflation shock causes a decrease in the output growth rate, which has negative effects on the market price of output growth risk. Since the effect though $\delta_{12}$ dominates the effect though $\delta_{11}$, the market price of output growth risk responds positively and so term premium responses negatively to the inflation shock.

For evaluating the influence of $\delta_{12}$, we calculated the impulse response functions of discount rates when $\delta_{12} = 0$ and the other parameters are unchanged in Figure 8. The main change in the impulse response functions appears in (8-b), which is totally different from (6-b). In (6-b), the responses of longer rates are smaller than the short rate, and the difference between the long and short rates almost disappear around 20 quarters ahead. On the other hand, in (8-b), the responses of longer rates are stronger than the short rate, and the difference does not disappear even around 40 quarter ahead. Why are they so different? The expectations hypothesis says that the long rate is the average of expected short rates plus a constant term premium. From (4-b), the inflation shock continues to raise the short rate up to around 20 quarters ahead. So according to the hypothesis, the initial responses of long rates with maturities up to 20 quarters should be stronger than the response of the short rate, as displayed in (8-b). But since in fact $\delta_{12}$ is positive, the inflation shock raises the market price of output growth risk, and so reduces the term premium. This is why long rates respond less strongly than the short rate in (6-b). The difference of responses in (6-b) and (8-b) has large effects on the predictabilities. Figure 9 shows model-implied $R^2$'s corresponding to the case of $\delta_{12} = 0$. Surprisingly, the $R^2$'s almost disappear. So now we can conclude that the positive $\delta_{12}$, which can be interpreted that consumers are willing to pay a higher premium for output growth risk hedge during the higher inflation regime, is a key explanation for the predictabilities.

The last notable feature of Figure 6 is the lagged responses of 1Q rate (the monetary policy authority) to output growth and inflation shocks. Panel (6-a) shows that the immediate response of 1Q rate to output growth shock is smallest among discount rates, although the response of 1Q rate is largest at several quarters ahead. Panel (6-b) shows that the immediate response of 1Q rate to inflation shock is smaller than 2Q rate, and almost coincides with 8Q rate.

These results are consistent with the monetary policy authority's inertial behavior empirically shown in the literature such as Clarida, Gali, and Gertler (2000). The lower graphs in Figure 7 show the impulse response functions of 20Q-1Q and 12Q-8Q spreads to output growth and inflation shocks. The near responses of 20Q-1Q spread are much weaker than 12Q-8Q spread because of the slow responses of 1Q rate. Since recent shocks are very important for predictions, we can conclude that this is the reason that 20Q-1Q spread is worse than 12Q-8Q spread. That is, the monetary authority's inertial behavior disturbs the responses of term spreads using the short end of the yield curve to the output growth and inflation shocks.

Further support for this view is provided by the correlations between future predicted variables and current term spreads. Since model-implied $R^2$'s are squares of these model-implied correlations, we can use the correlations for analyzing why we found the $R^2$'s shown in Figure 3 or 2. Equation (67) has four summed terms, and each of them can be interpreted as the contribution of a given exogenous shock to the predictability. Figure 10 shows the contributions of exogenous shocks to the absolute values of correlations for 20Q-1Q and 12Q-8Q spreads. The inflation and output growth shocks contribute to the correlations with 12Q-8Q spread rather than 20Q-1Q spread. These differences are the reason for the usefulness of the 12Q-8Q spread for prediction. This result is consistent with our discussion about the results in lower graphs in Figure 7.

Another notable property in Figure 10 is the hump-shapes of the contributions of monetary policy shocks to output growth predictability. So we can conclude that the hump-shape of $R^2$'s for output growth predictions is attributable to the monetary policy shock. That is, the monetary policy shock affects output growth with a lag, while the term structure responds to the shock immediately. This difference in timing makes it harder for term spreads to help forecast output growth at short horizons.

Finally, Figure 11 shows the contributions in the case of $\delta_{12} = 0$. Obviously the sharp drops of $R^2$'s are attributable to the different sign of contribution of the inflation shock, which are caused by strong long rate responses to the shock.

## 7. Conclusion

Why do term spreads predict output growth, inflation, and short rates? For answering this question, we used the VAR-ATSM model with four lags and four variables, which is less restricted than those in the literature of affine term structure models with observable factors. And we succeeded in estimating this model by using an efficient method.

We have three main findings. First, the time-varying market price of output growth risk, which is sensitive to the level of inflation, plays a key role in explaining why the term spread helps forecast output growth, inflation, and interest rates. This finding can be interpreted as follows. When the inflation rate is higher, consumers are willing to pay a higher premium for output growth risk hedge, possibly because the marginal utility is more sensitive to the output growth shock due to less accommodating response of the Fed. This causes term spreads to react to recent inflation shocks, which also prove useful for forming longer-run forecasts. Second, we also found that term spreads using the short end of yield curve have less predictive power than many spreads between longer rates. This fact is attributable to the inertial character of monetary policy. Finally, it is hard to predict output growth with term spreads at short horizons, because monetary policy shock affects output growth with a lag while the term structure responses to the shock immediately.

## References


Ang, A. and M. Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, Journal of Monetary Economics 50, 745-787.

Ang, A., M. Piazzesi and M. Wei, 2003, What does the yield curve tell us about GDP growth?, working paper, Andersen school, UCLA.

Backus, D., S. Foresi and C. Telmer, 1998, Discrete-time models of bond pricing, working paper, NBER 6736.

Balduzzi, P., G. Bertola and S. Foresi, 1997, A model of target changes and the term structure of interest rates, Journal of Monetary Economics 39, 223-249.

Bekaert, G., R. J. Hodrick and D. A. Marshall, 2001, Peso problem explanations for term structure anomalies, Journal of Monetary Economics 48, 241-270.

Bernanke, B. S., M. Gertler and M. Watson, 1997, Systematic monetary policy and the effects of oil price shocks, Brookings Papers on Economic Activity 1:1997, 91-142.

Bernard, H. and S. Gerlach, 1998, Does the term structure predict recessions? The international evidence, International Journal of Finance and Economics 3, 195-215.

Bonser-Neal, C., and T. R. Morley, 1997, Does the yield spread predict real economic activity? A multicountry analysis, Economic Review 82, 3rd quarter, 37-53.

Breeden, D. T., 1986, Consumption, production, inflation, and interest rates, Journal of Financial Economics, 16, 3-39.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, 1997, The econometrics of financial markets (Princeton University Press).

Campbell, J. Y., and R. J. Shiller, 1991, yield spread and interest rate movements: a bird's eye view, Review of Economic Studies 58, 495-514.

Christiano, L. J., M. Eichenbaum and C. L. Evans, 1999, Monetary policy shocks: what have we learned and to what end?, in: J. B. Taylor and M. Woodford, eds., Handbook of macroeconomics, Vol. 1.

Clarida, R., J. Gali and M. Gertler, 2000, Monetary policy rules and macroeconomic stability: evidence and some theory, The Quarterly Journal of Economics 115, 147-180.

Cliff, M. T., 2000, GMM and MINZ program libraries for MATLAB, Krannet Graduate School of Management, Purdue University.

Cochrane, J. H., 2001, Asset pricing (Princeton University Press).

Cochrane, J. H., and M. Piazzesi, 2002, Bond risk premia, working paper, Andersen school, UCLA.

Dai, Q., and K. J. Singleton, 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, Journal of Financial Economics 63, 3, 415-441.

Dewachter, H., and M. Lyrio, 2002, Macro factors and the term structure of interest rates, working paper, Center for economic studies, Catholic University of Leuven.

Dotsey, M., 1998, The predictive content of the interest rate term spread for future economic growth, Economic Quarterly, Federal Reserve Bank of Richmond 84, 3, 31-51.

Dotsey, M. and C. Otrok, 1995, The rational expectations hypothesis of the term structure, monetary policy, and time-varying term premia, Economic Quarterly, Federal Reserve Bank of Richmond 81, 1, 65-81.

Dueker, M., 1997, Strengthening the case for the yield curve as a predictor of U.S. recessions, Review, Federal Reserve Bank of St. Louis 79, March/April, 41-51.

Duffee, G. R., 2002, term premia and interest rate forecasts, Journal of Finance 57, 1, 405-443.

Duffie, J. D. and R. Kan, 1996, A yield-factor model of interest rates, Mathematical Finance 6, 379-406.

Engle, R. F., and D. M. Lilien, and Russell P. Robins, Estimating time varying risk premia in the term structure: the Arch-M model, Econometrica 55, 2, 391-407.

Engle, R. F., and V. K. Ng, 1993, Time-varying volatility and the dynamic behavior of the term structure, Journal of Money, Credit, and Banking 25, 3, 336-349.

Estrella, A., and G. Hardouvelis, 1991, The term structure as a predictor of real economic activity, Journal of Finance 46, 2, 555-576.

Estrella, A., and F. S. Mishkin, 1997, The predictive power of the term structure of interest rates in Europe and the United States: implications for the European Central Bank, European Economics Review 41, July, 1375-1401.

Evans, C. L., and D. A. Marshall, 1998, Monetary policy and the term structure of nominal interest rates, evidence and theory, Carnegie-Rochester Conference Series on Public Policy 49, 53-111.

Fama, E. F., 1975, Short term interest rates as predictors of inflation, American Economic Review 65, 269-282.

Fama, E. F., 1990, Term-structure forecasts of interest rates, inflation, and real returns, Journal of Monetary Economics 25, 1, 59-76.

Fama, E. F., and R. R. Bliss, 1987, The information in long-maturity forward rates, American Economic Review 77, 4, 680-692.

Gordon, D. B. and E. M. Leeper, 1994, The dynamic impacts of monetary policy: an exercise in tentative identification, Journal of Political Economy 102, 6, 1228-1247.

Hamilton, J. D., 1994, Time series analysis (Princeton University Press).

Hamilton, J. D. and D. H. Kim, 2002, A re-examination of the predictability of the yield spread for real economic activity, Journal of Money, Credit, and Banking 34, 2, 340-360.

Harrison, J. M., and D. M. Kreps, 1979, Martingales and arbitrage in multiperiod securities markets, Journal of Economic Theory 20, 381-408.

Hansen, L. P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.

Hansen, L. P., and K. Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269-1286.

Hansen, L. P., and R. Jagannathan, 1991, Assessing specification errors in stochastic discount factor models, Journal of Finance 52, 557-590.

Harvey, C., 1988, The real term structure and consumption growth, Journal of Financial Economics 22, 305-334.

Hardouvelis, G., 1994, The term structure spread and future changes in long and short rates in the G7 countries. Is there a puzzle?, Journal of Monetary Economics 33, 255-283.

Haubrich, J. G., and A. M. Dombrosky, 1996, Predicting real growth using the yield curve, Economic Review, Federal Reserve Bank of Cleveland 32, 1st quarter, 26-35.

Hordahl, P., O. Tristani, and D. Vestin, 2002, A joint econometric model of macroeconomic and term structure dynamics, working paper, European Central Bank.

Jorion, P., and F. S. Mishkin, 1991, A multicountry comparison of term-structure forecasts at long horizons, Journal of Financial Economics 29, 1, 59-80.

Kozicki, S., 1997, Predicting real growth and inflation of monetary policy, Economic Review, Federal Reserve Bank of Kansas City 82, $4^{th}$ quarter, 39-57.

Longstaff, F. A., and E. S. Schwartz, 1992, Interest rate volatility and the term structure: a two-factor general equilibrium model, Journal of Finance 47, Sep., 1259-1282.

Mishkin, F. S., 1988, The information in the term structure: some further results, Journal of Applied Econometrics 3, 307-314.

Mishkin, F. S., 1990a, What does the term structure tell us about future inflation?, Journal of Monetary Economics 25, January, 77-95.

Mishkin, F. S., 1990b, The information in the longer maturity term structure about future inflation, Quarterly Journal of Economics 105, 3, 815-828.

Mishkin, F. S., 1991, A multi-country study of the information in the shorter maturity term structure about future inflation, Journal of International Money and Finance 19, March, 2-22.

Mankiw, N. G. and J. A. Miron, 1986 The changing behavior of the term structure of interest rates, Quarterly Journal of Economics 101, 211-228.

Newey, W. K., and K. D. West, 1987, A simple positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, May, 703-778.

Pagan, A. R. and J. C. Robertson, 1995, Resolving the liquidity effect, Federal Reserve Bank of St. Louis Review 77, 3, 33-54.

Piazzesi, M., 2001, An econometric model of the yield curve with macroeconomic jump effects, working paper, Andersen school, UCLA.

Plosser, C. I., and K. G. Rouwenhost, 1994, International term structure and real economic growth, Journal of Monetary Economics 33, February, 135-155.

Roberds, W., and C. H. Whiteman, 1999, Endogenous term premia and anomalies in the term structure of interest rates: explaining the predictability smile, Journal of Monetary Economics 44, 555-580.

Rudebusch, G. D., 1995, Federal Reserve interest rate targeting, rational expectations, and the term structure, Journal of Monetary Economics 35, 2, 245-274.

Sims, C. A., 1980, Macroeconomics and reality, Econometrica 48, 1, 1-48.

Stock, J. H., and M. W. Watson, 1989, New index of coincident and leading indicators, in O. Blanchard and S. Fischer, eds., NBER Macroeconomic Annual, Vol. 4 (MIT Press).

Taylor, J., 1993, Discretion versus policy rules in practice, Carnegie-Rochester Conference on Public Policy 39, 195-214.

Vasicek, O., 1977, An equilibrium characterization of the term structure, Journal of Financial Economics 5, 177-188.

Walsh, C. E., 1998, Monetary theory and policy (The MIT Press).

Wu, T., 2002, Macro factors and the affine term structure of interest rates, working paper, Federal Reserve Bank of San Francisco.

Table 1: Estimated parameters of AR(1) model for the short rate factor

| $c_r$ | $\phi_r$ | $\sigma_r$ |
|---|---|---|
| 0.1499<br>(0.0621) | 0.9037<br>(0.0362) | 0.0057 |

The AR(1) model for the short rate (7) is estimated by OLS. Standard errors are in parentheses. The sample period is 1964:1Q-2001:4Q.

Table 2: Estimated parameters of $\mathbf{\Sigma}$

| | Shocks | | | |
|---|---|---|---|---|
| | $u_{g,t}$ | $u_{\pi,t}$ | $u_{r,t}$ | $u_{s,t}$ |
| $g_t$ | 0.0076 | 0 | 0 | 0 |
| $\pi_t$ | -0.0001 | 0.0025 | 0 | 0 |
| $r_t^{(1)}$ | 0.0007 | 0.0004 | 0.0023 | 0 |
| $s_t$ | -0.0002 | -0.0002 | -0.0014 | 0.0012 |

$\mathbf{\Sigma}$ is estimated by GMM introduced in section 5. The sample period is 1964:1Q-2001:4Q.

Table 3: Estimated parameters of $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$

| | $\boldsymbol{\gamma}$ | $\boldsymbol{\delta}$ | | | |
|---|---|---|---|---|---|
| | | $g_t$ | $\pi_t$ | $r_t^{(1)}$ | $s_t$ |
| $\lambda_{g,t}$ | -0.50<br>(0.43) | 140*<br>(22) | 78*<br>(32) | -26<br>(26) | -43<br>(63) |
| $\lambda_{\pi,t}$ | -0.89<br>(0.90) | -99*<br>(48) | 62<br>(57) | -60<br>(51) | -177<br>(114) |
| $\lambda_{r,t}$ | 0.25<br>(0.25) | -23*<br>(10) | -13<br>(11) | -30*<br>(12) | -45*<br>(17) |
| $\lambda_{s,t}$ | 0.67*<br>(0.31) | -46<br>(29) | 28*<br>(14) | -30<br>(16) | -114*<br>(26) |
| mean of factor | | 0.0080 | 0.0102 | 0.0159 | 0.0026 |
| s.d. of factor | | 0.0089 | 0.0061 | 0.0065 | 0.0032 |

$\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are estimated by GMM introduced in section 5. The estimates with * are significantly different from zero at 5%. Standard errors are in parentheses. Last two rows report means and standard deviations of $g_t$, $\pi_t$, $r_t^{(1)}$, and $s_t$. The sample period is 1964:1Q-2001:4Q.

Table 4: The comparison between model-implied rates and sample rates

| maturity ($n$) | | 2 | 4 | 8 | 12 | 16 | 20 |
|---|---|---|---|---|---|---|---|
| mean | $\hat{r}_t^{(n)}$ | 0.0166 | 0.0172 | 0.0177 | 0.0181 | 0.0184 | 0.0186 |
| | $r_t^{(n)}$ | 0.0166 | 0.0172 | 0.0177 | 0.0181 | 0.0184 | 0.0185 |
| s.d. | $\hat{r}_t^{(n)}$ | 0.0064 | 0.0062 | 0.0060 | 0.0059 | 0.0059 | 0.0058 |
| | $r_t^{(n)}$ | 0.0065 | 0.0063 | 0.0062 | 0.0060 | 0.0059 | 0.0058 |
| correlation | | 0.9927 | 0.9869 | 0.9885 | 0.9902 | 0.9926 | 0.9928 |

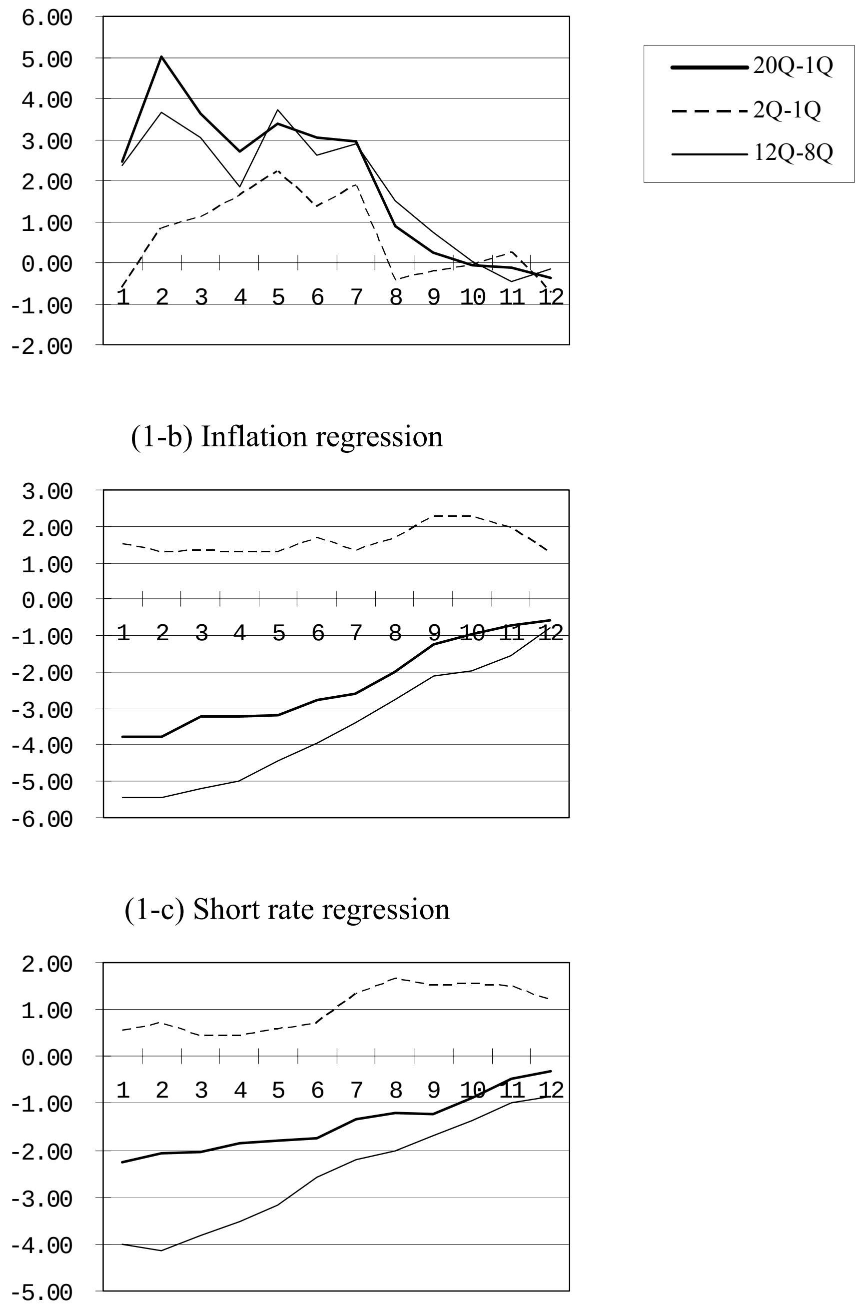
Means and standard deviations of $\hat{r}_t^{(n)}$ and $r_t^{(n)}$, and correlations between them are reported. The sample period is 1964:1Q-2001:4Q.

Figure 1: The *t*-stats of OLS regressions

(1-a) Output growth regression

(1-b) Inflation regression

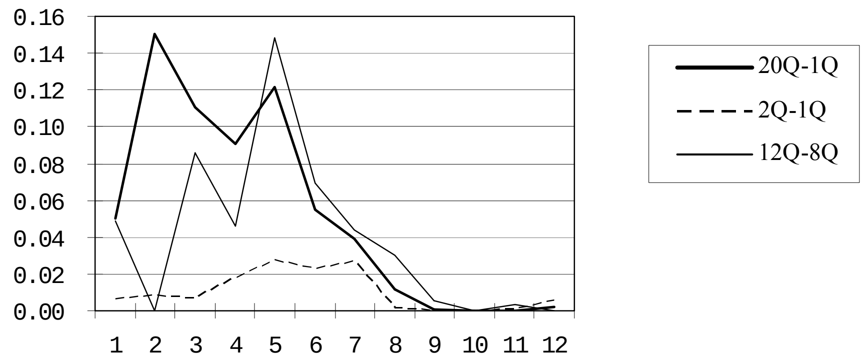
(1-c) Short rate regression



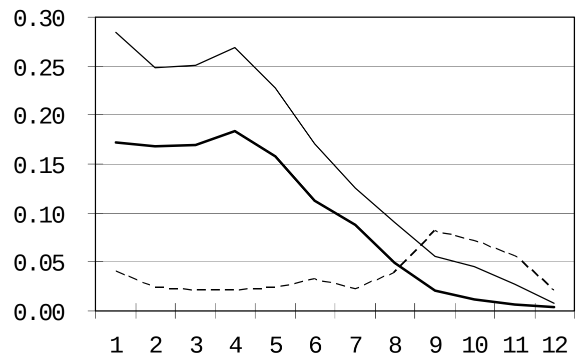

The *t*-stats of OLS regressions (4)-(6) are reported. The horizontal axes correspond to forecasting horizons (quarters). Thick, broken, and thin lines correspond to 20Q-1Q, 2Q-1Q, and 12Q-8Q term spreads respectively. The sample period is 1964:1Q-2001:4Q.

Figure 2: The sample $R^2$'s of OLS regressions

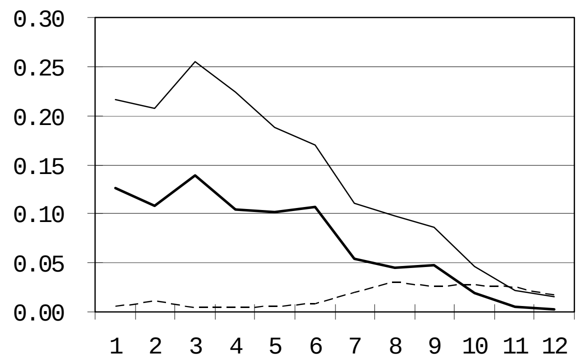
(2-a) Output growth regression




(2-b) Inflation regression



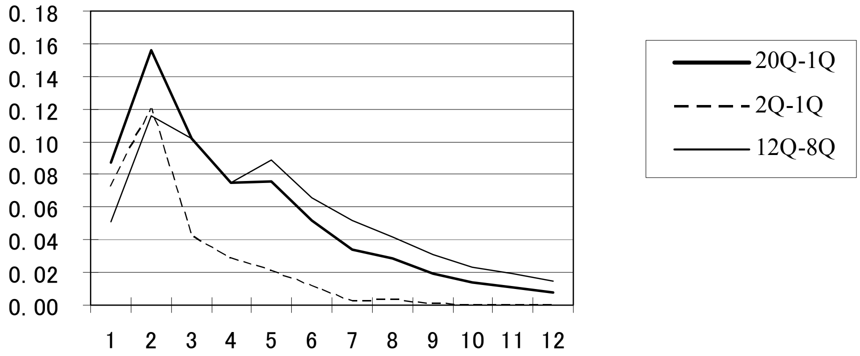

(2-c) Short rate regression



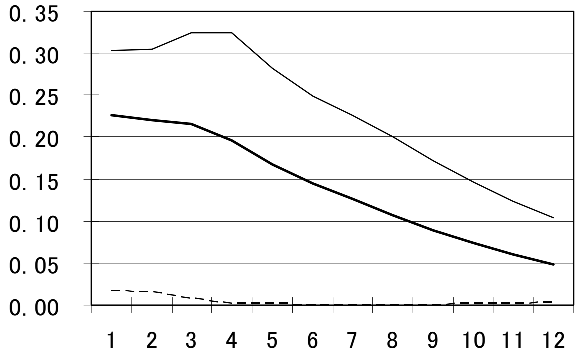

The sample $R^2$'s of OLS regressions (4)-(6) are reported. The horizontal axes correspond to forecasting horizons (quarters). Thick, broken, and thin lines correspond to 20Q-1Q, 2Q-1Q, and 12Q-8Q term spreads respectively. The sample period is 1964:1Q-2001:4Q.

Figure 3: The model-implied $R^2$'s

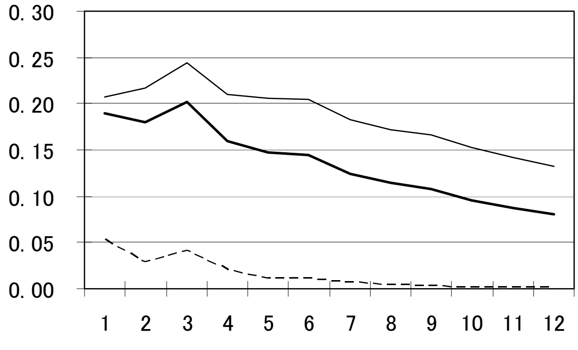
(3-a) Output growth regression




(3-b) Inflation regression

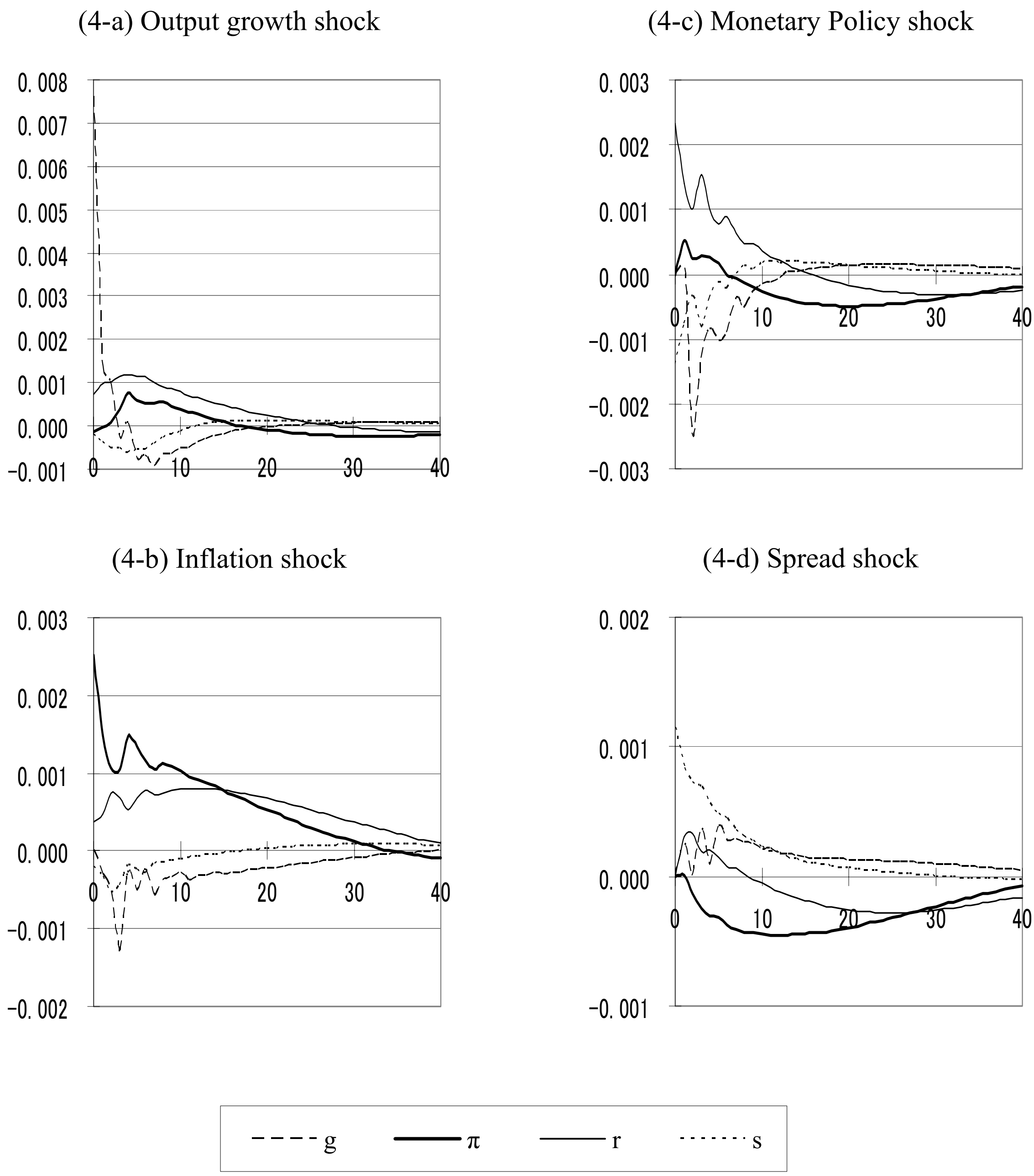



(3-c) Short rate regression



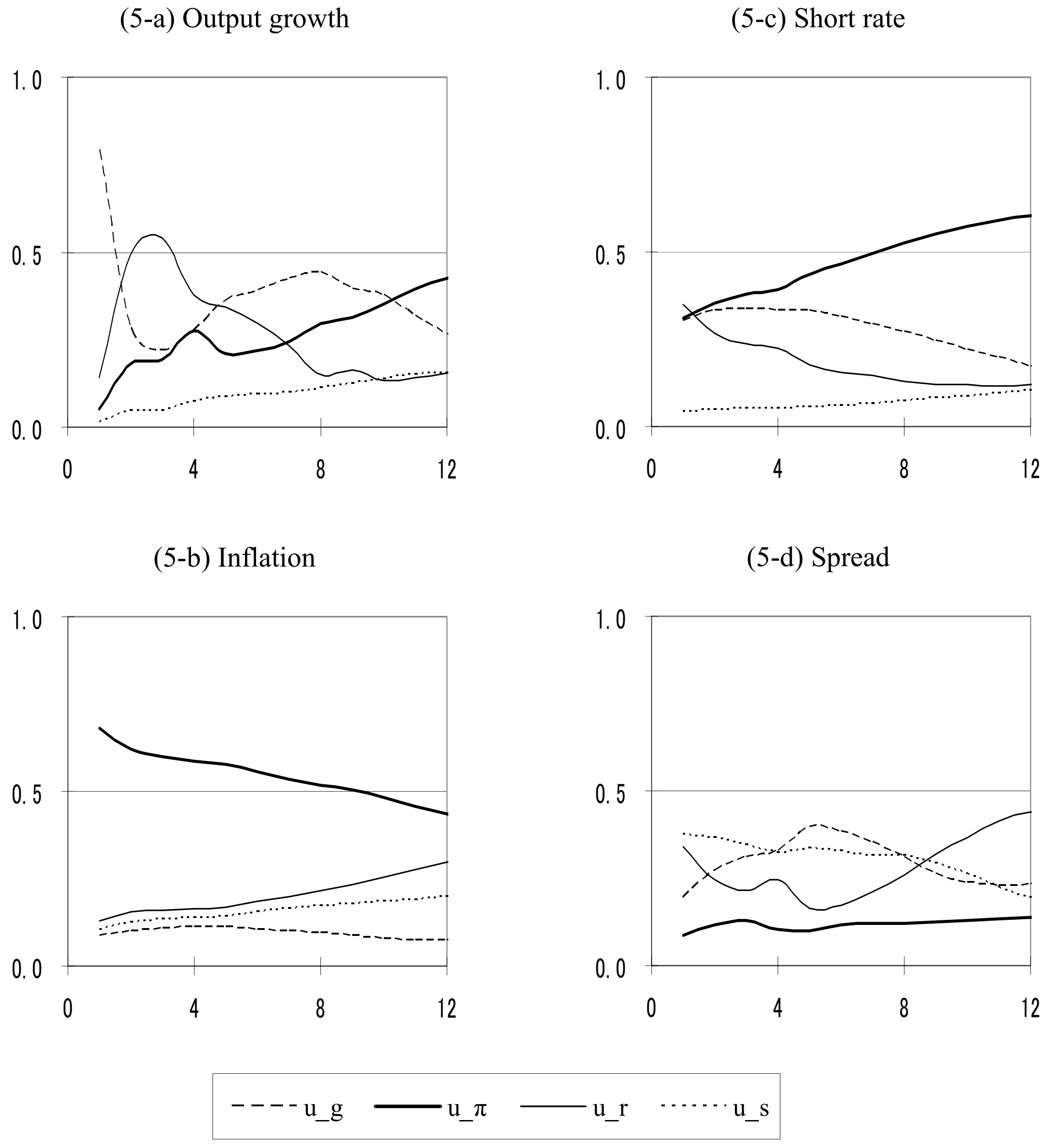

The model-implied $R^2$'s of regressions (4)-(6) are reported. The horizontal axes correspond to forecasting horizons (quarters). Thick, broken, and thin lines correspond to 20Q-1Q, 2Q-1Q, and 12Q-8Q term spreads respectively. The sample period is 1964:1Q-2001:4Q.

Figure 4: The impulse response functions of VAR variables



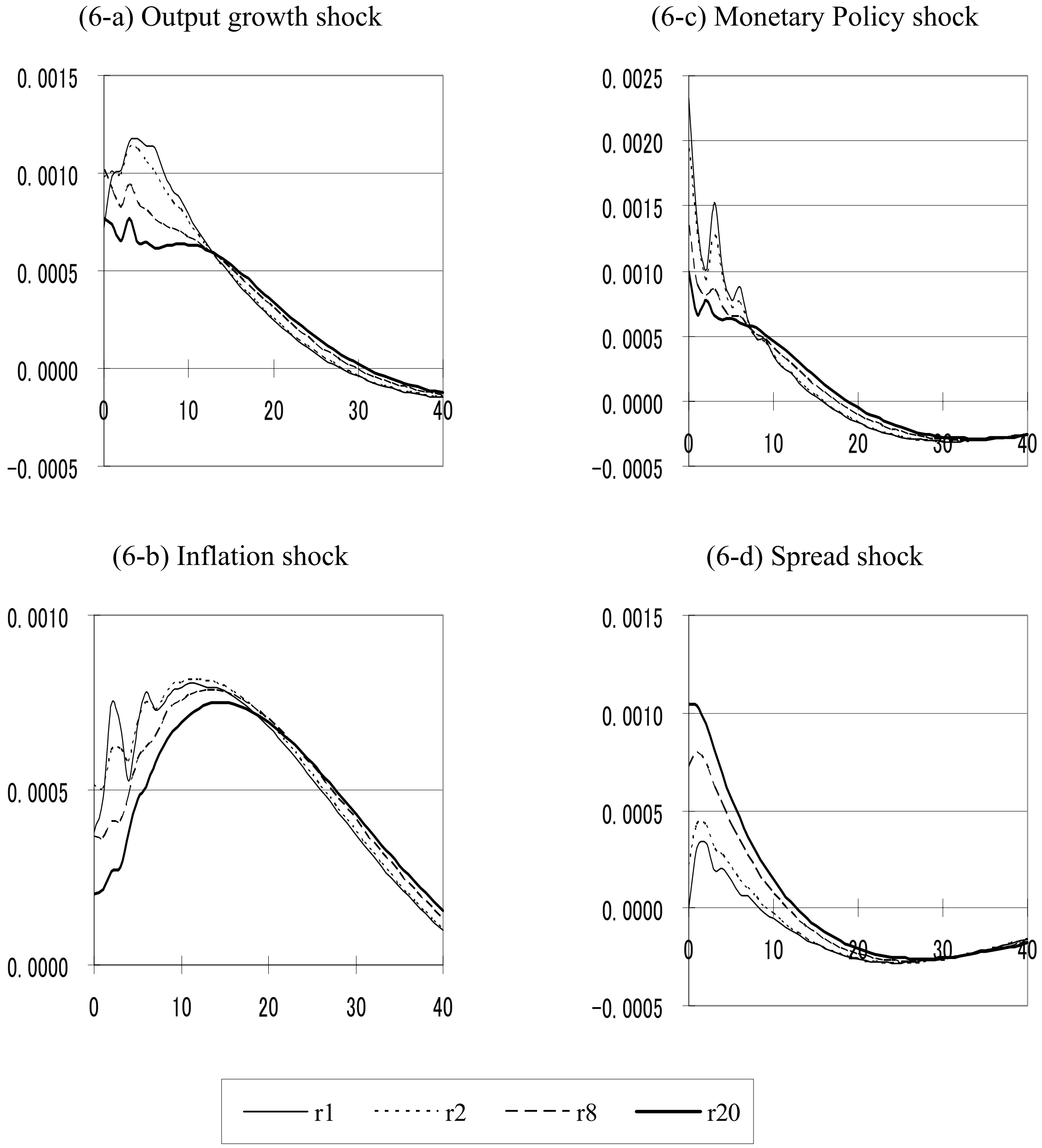

The impulse responses of VAR variables to one-unit exogenous shocks are reported. Broken, thick, thin, and dotted lines correspond to responses of output growth, inflation, short rates, and term spread respectively. The horizontal axes correspond to horizons (quarters). The sample period is 1964:1Q-2001:4Q.

Figure 5: The variance decompositions of the optimal forecasts for VAR variables




The variance decompositions of the optimal forecasts of VAR variables, in which the variances of the forecasts are normalized to unity, are reported. Broken, thick, thin, and dotted lines correspond to output growth, inflation, monetary policy, and spread shocks respectively. The horizontal axes correspond to forecasting horizons (quarters). The sample period is 1964:1Q-2001:4Q.

Figure 6: The impulse response functions of discount rates

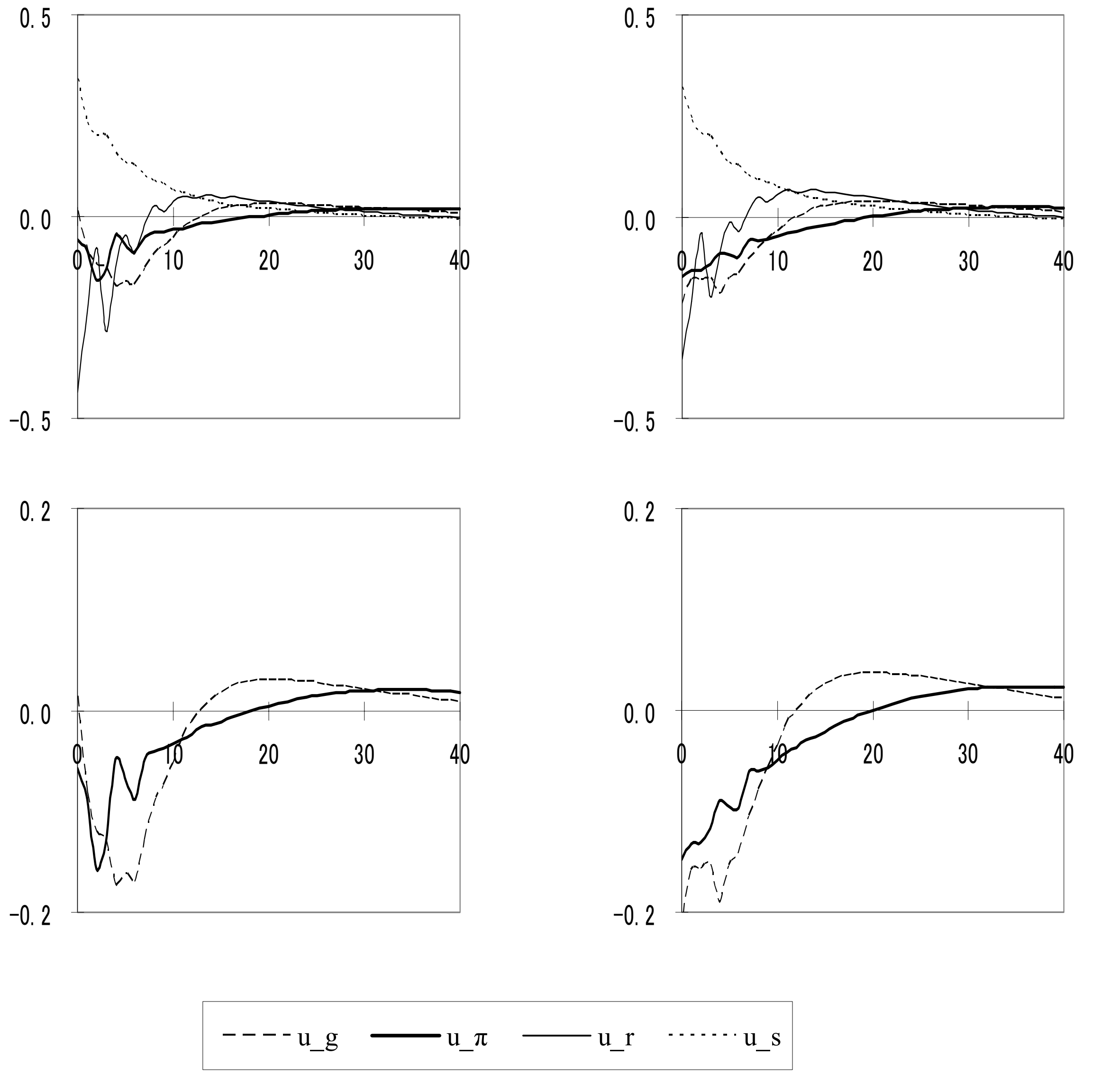



The impulse responses of discount rates to one-unit exogenous shocks are reported. Thin, dotted, broken, and thick lines correspond to 1Q, 2Q, 8Q, and 20Q rates respectively. The horizontal axes correspond to horizons (quarters). The sample period is 1964:1Q-2001:4Q.

Figure 7: The impulse response functions of term spreads

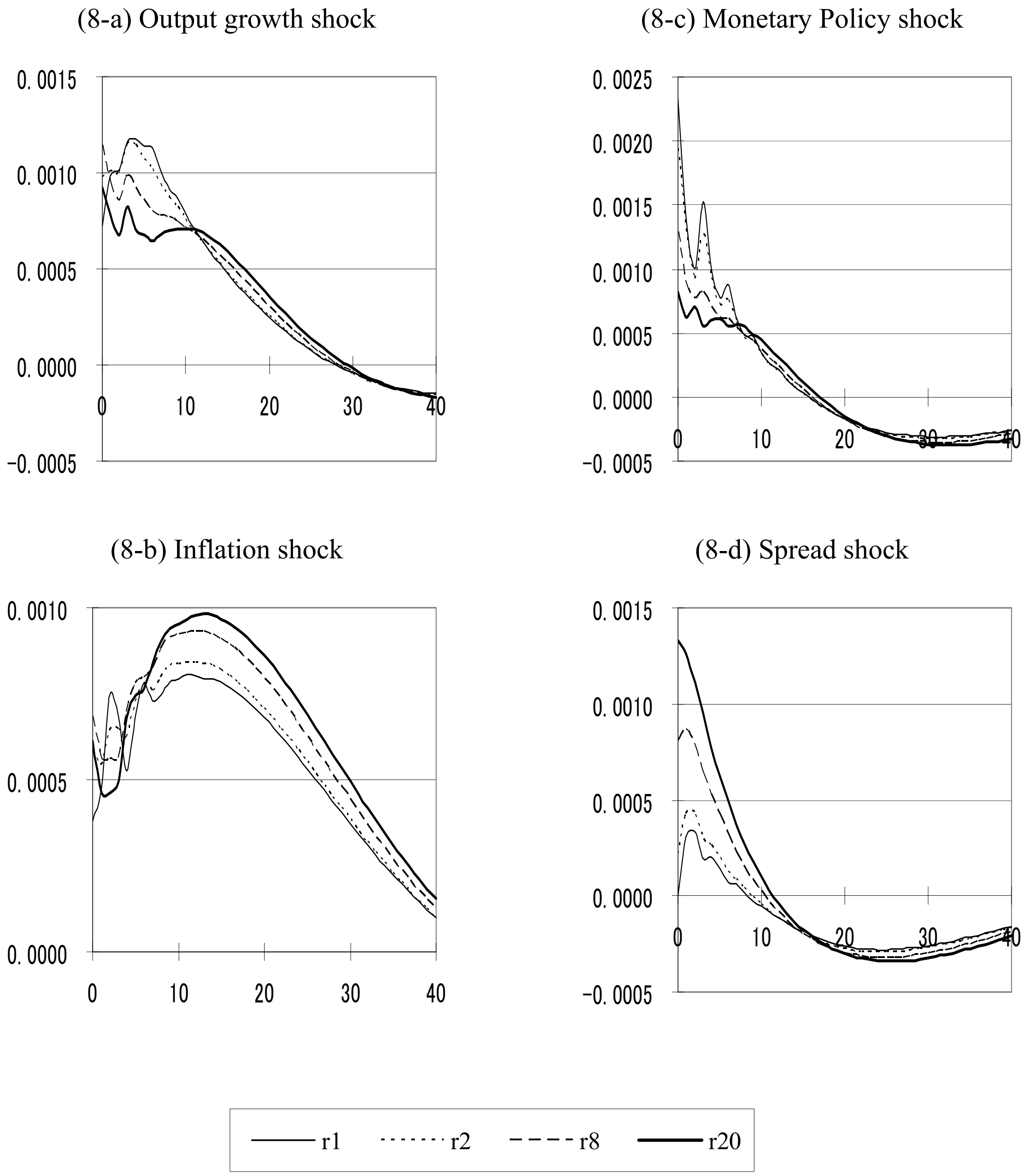
(7-a) The impulse response functions of the 20Q-1Q spread

(7-b) The impulse response functions of the 12Q-8Q spread




The impulse responses of 20Q-1Q and 12Q-8Q spreads to one-unit exogenous shocks are shown in upper graphs. The scales are normalized so that variances of spreads equal unity. Lower graphs show magnified impulse responses to output growth and inflation shocks. Broken, thick, thin, and dotted lines correspond to output growth, inflation, monetary policy, and spread shocks respectively. The horizontal axes correspond to horizons (quarters). The sample period is 1964:1Q-2001:4Q.

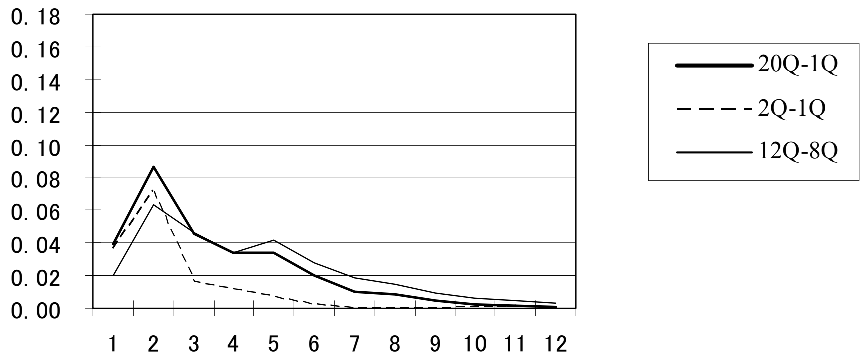
Figure 8: The impulse response functions of discount rates in the case of $\delta_{12} = 0$



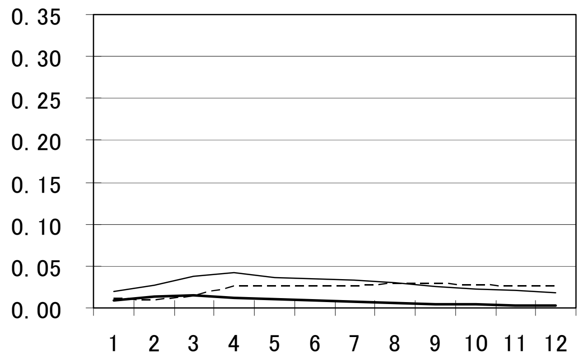

The impulse responses of discount rates to one-unit exogenous shocks are reported. Thin, dotted, broken, and thick lines correspond to 1Q, 2Q, 8Q, and 20Q rates respectively. The horizontal axes correspond to horizons (quarters). The sample period is 1964:1Q-2001:4Q.

Figure 9: The model-implied $R^2$'s in the case of $\delta_{12} = 0$

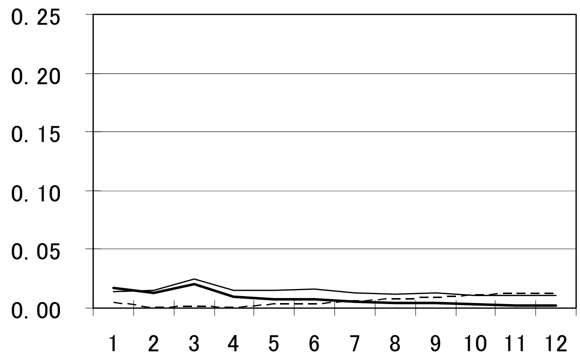
(9-a) Output growth regression




(9-b) Inflation regression

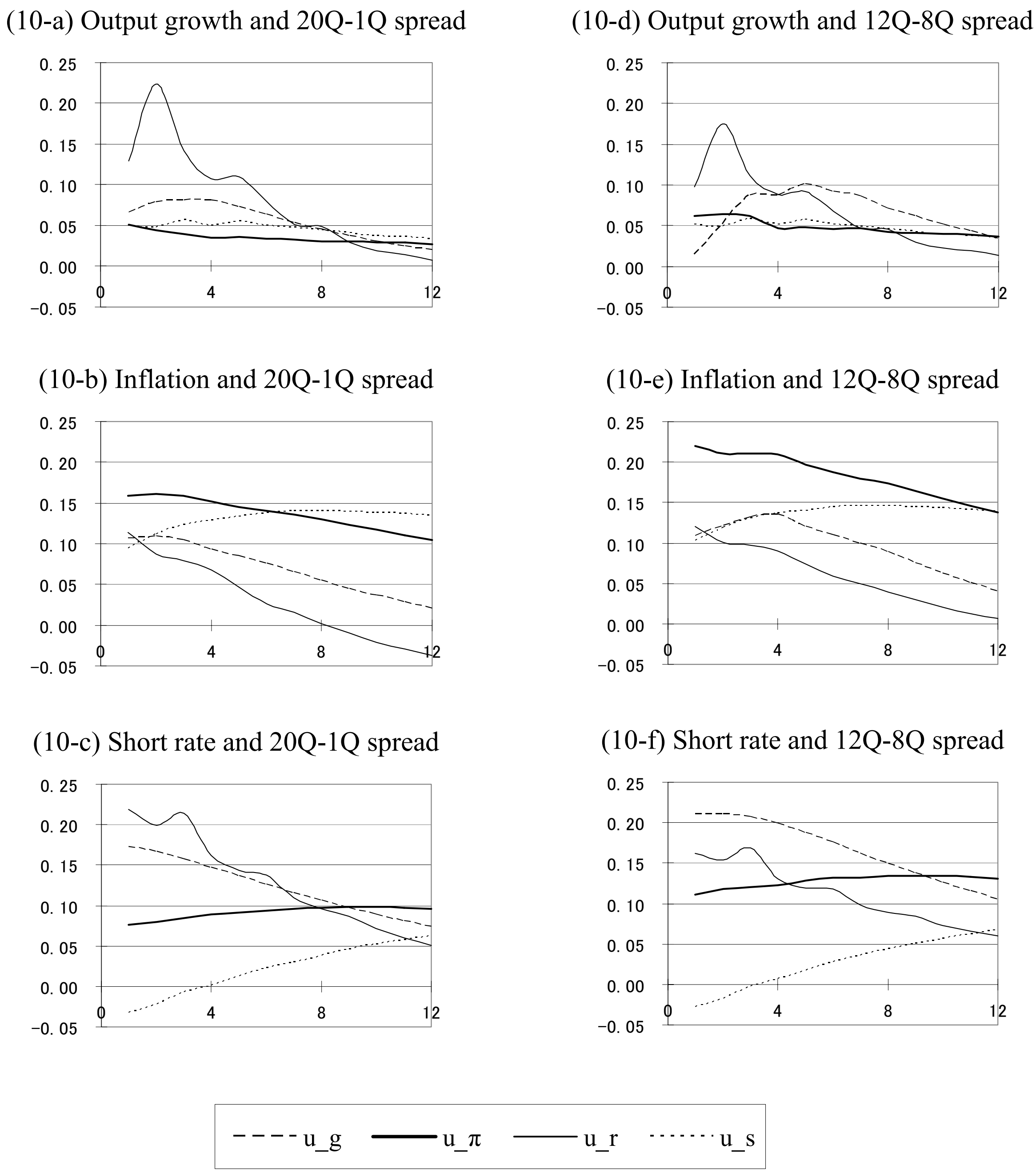



(9-c) Short rate regression



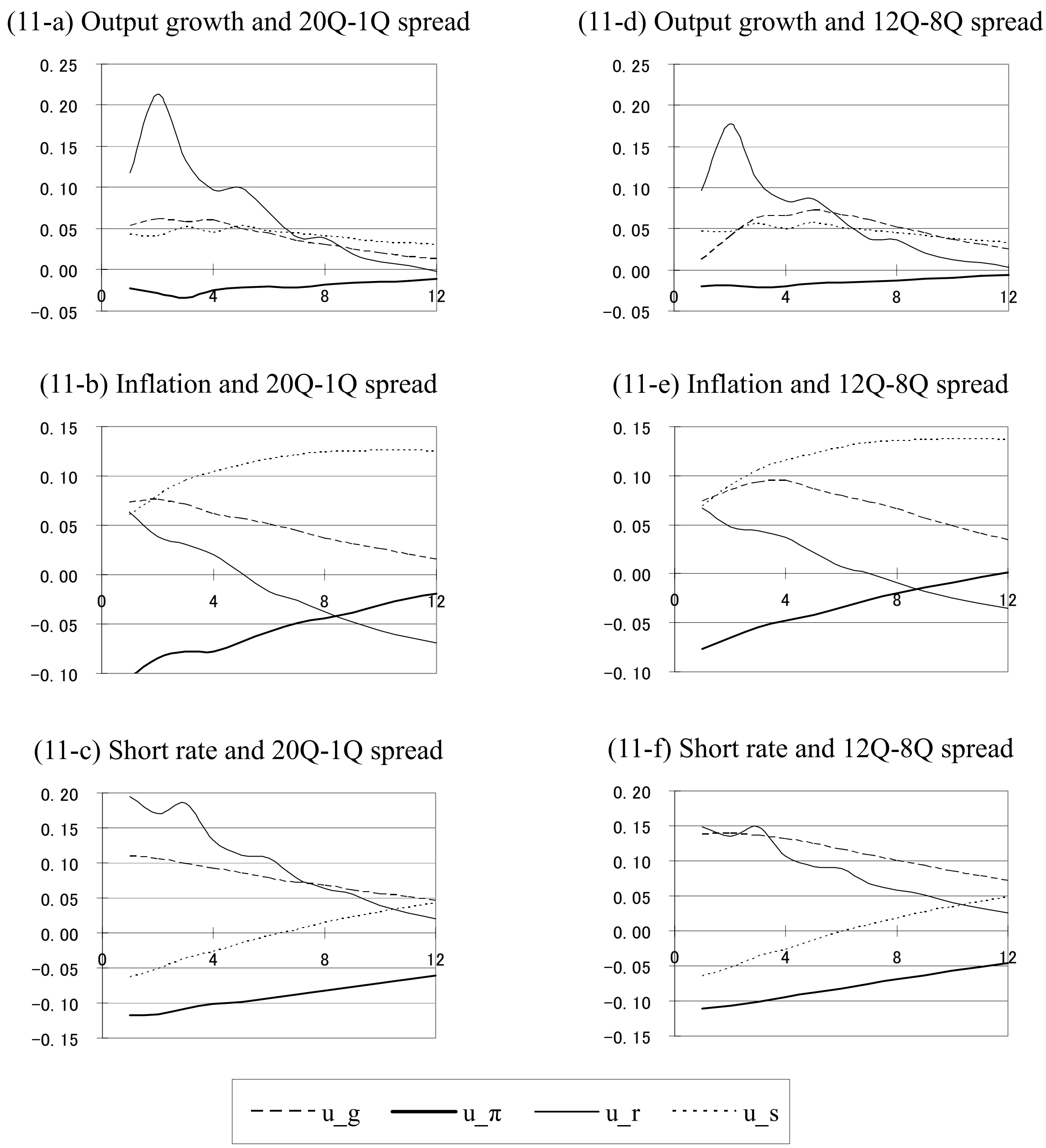

The model-implied $R^2$'s of regressions (4)-(6) are reported. The horizontal axes correspond to forecasting horizons (quarters). Thick, broken, and thin lines correspond to 20Q-1Q, 2Q-1Q, and 12Q-8Q term spreads respectively. The sample period is 1964:1Q-2001:4Q.

Figure 10: Decompositions of correlations between future VAR variables and term spreads




The contributions of shocks to the correlations between VAR variables and term spreads are shown. Since the correlations of term spreads with inflation and short rate are negative, the graphs are flipped for (10-b), (10-c), (10-e), and (10-f). Broken, thick, thin, and dotted lines correspond to output growth, inflation, monetary policy, and spread shocks respectively. The horizontal axes correspond to forecasting horizons (quarters). The sample period is 1964:1Q-2001:4Q.

Figure 11: Decompositions of correlations between future VAR variables and term spreads in the case of $\delta_{12} = 0$




The contributions of shocks to the correlations between VAR variables and term spreads are shown. Since the correlations of term spreads with inflation and short rate are negative, the graphs are flipped for (11-b), (11-c), (11-e), and (11-f). Broken, thick, thin, and dotted lines correspond to output growth, inflation, monetary policy, and spread shocks respectively. The horizontal axes correspond to forecasting horizons (quarters). The sample period is 1964:1Q-2001:4Q.