TERM STRUCTURE DYNAMICS WITH MACROECONOMIC FACTORS

A Dissertation

by

HA-IL PARK

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2009

Major Subject: Economics

TERM STRUCTURE DYNAMICS WITH MACROECONOMIC FACTORS

A Dissertation

by

HA-IL PARK

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Hwagyun Kim
Committee Members,	Joon Y. Park
	Yoosoon Chang
	Neal E. Galpin
Head of Department,	Larry Oliver

December 2009

Major Subject: Economics

ABSTRACT

Term Structure Dynamics with Macroeconomic Factors. (December 2009)

Ha-Il Park, B.A., Yonsei University;

M.S., Korea Advanced Institute of Science and Technology

Chair of Advisory Committee: Dr. Hwagyun Kim

Affine term structure models (ATSMs) are known to have a trade-off in predicting future Treasury yields and fitting the time-varying volatility of interest rates. First, I empirically study the role of macroeconomic variables in simultaneously achieving these two goals under affine models. To this end, I incorporate a liquidity demand theory via a measure of the velocity of money into affine models. I find that this considerably reduces the statistical tension between matching the first and second moments of interest rates. In terms of forecasting yields, the models with the velocity of money outperform among the ATSMs examined, including those with inflation and real activity. My result is robust across maturities, forecasting horizons, risk price specifications, and the number of latent factors. Next, I incorporate latent macro factors and the spread factor between the short-term Treasury yield and the federal funds rate into an affine term structure model by imposing cross-equation restrictions from no-arbitrage using daily data. In doing so, I identify the high-frequency monetary policy rule that describes the central bank's reaction to expected inflation and real activity at daily frequency. I find that my affine model with macro factors and the spread factor shows better forecasting performance.

To my family

ACKNOWLEDGMENTS

I would like to thank my advisor, Hwagyun Kim, for his encouragement, insightful guidance, and involvement in progressing my dissertation. For valuable comments I thank Joon Y. Park, Yoosoon Chang, Neal E. Galpin, and Michael F. Gallmeyer. I am grateful to my colleagues for helpful suggestions at seminars. I thank Courtney Collins for help in editing the manuscript. I would like to express my gratitude to my mother, Misoon Yang, and my mother-in-law, Boksoon Lee. They have always supported me. I extend special thanks to my wife, Sehyun Song. I could not have completed my study without her love and encouragement. Lastly thanks to my lovely sons, Jaeseok Park and Jaeyoung Park.

This dissertation is dedicated to my deceased father, Gyeonggeun Park.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

TABLE Page

TABLE Page

TABLE Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Affine term structure models (ATSMs) appeal to both practitioners and academic researchers for tractability in econometric implementation and a sufficient degree of freedom in specifying how bond market compensates investors for taking systematic risk. Beginning with the pioneering studies by Vasicek (1977) and Cox, Ingersoll, and Ross (1985), researchers extended ATSMs to explain several important aspects of Treasury yields data.

Recently, Dai and Singleton (2000), Duffee (2002), and Duarte (2004) reported that there exists a trade-off between improving forecast ability on future bond yields and matching interest rate volatility in affine models. Since the market price of risk setup can be modeled leaving the affine form of term structure intact as shown by Duffie and Kan (1996), a sufficiently flexible setup for the market price of risk can be one way to resolve these issues. Along this line, Duffee (2002) and Duarte (2004) use alternative parameterizations of the market price of risk and report some success in increasing the predictive power of ATSMs, but they still have difficulty in generating the third fact unless they dispense with stochastic volatility.

In chapter II, I attempt to reduce this statistical tension by incorporating some observable macroeconomic variables into ATSMs. Specifically, I use a measure of liquidity demand, the velocity of money. For comparison, I also examine measures of inflation and output gap which are now popular in macroeconomic term structure studies. It is well known that inflation, output, and/or the velocity of money are closely related to interest rates both empirically and theoretically. Obviously, the first

---

This dissertation follows the style of *Econometrica*.

two variables are motivated from either the Fisherian theory or the recent monetary policy rule literature, and the third variable is inferred by a typical liquidity demand theory.

The influential work by Ang and Piazzesi (2003) shows that incorporating monetary policy behaviors via inflation and output measure can help forecast bond yields in an affine model with a conditionally homoscedastic setup. I emphasize the importance of macro factors in modeling term structure dynamics as well, but we depart from their work in two directions. First, I explicitly model stochastic volatility to tackle the estimation and prediction issues mentioned above. Second, I focus on liquidity or money demand theory to impose an economic restriction to both bond yields and key macroeconomic variables and compare my results with other affine models.

In this vein, I ask two research questions. First, do macroeconomic variables help reduce the trade-off between matching the first and second moments of bond yields? Second, which affine model, either in a latent or a (macro-latent) hybrid factor structure, predicts future bond yields better? Regarding the first question, I find that the statistical "tension" is considerably relaxed in my model. In addition, consistent with this result, the affine model with the velocity of money predict future yields better than all other ATSMs examined, even compared with the models including inflation and real activity both in sample and out of sample.

In chapter III, I construct a tractable model at daily frequency with both the typical latent factors and latent macro factors by imposing cross-equation restrictions on yield movements from no-arbitrage while most term structure models incorporate observable macroeconomic variables in monthly or quarterly frequency. In term structure models using low-frequency macro variables, it is hard to examine the role of macro variables in explaining term structure dynamics in continuous time. Additionally, I add the spread factor between the short-term Treasury yield and the federal funds rate into an affine term structure model to identify the high-frequency monetary policy rule that describes the central bank's reaction to expected inflation and real activity at daily frequency. The benchmark and backward-looking high-frequency policy rules are identified without difficulties. Although many other researchers place Taylor rules incorporating inflation and the output gap in an affine model, those macro variables are observable in monthly or quarterly frequency. In my model, different bond yields such as the real yield, nominal yield, and defaultable yield are used and latent macro factors and the spread factor are extracted from yield relationships by using cross-equation restrictions. Accordingly, I do not need to worry about the discrepancy of data frequency between yields and macro variables. Thus, I do not lose information available in matching high frequency yields data and low frequency macro variables data.

When I assess my model in terms of out-of-sample forecasting, the term structure model with macro factors and the spread factor shows better performance. Moreover, I show that the spread between the 3-month Treasury yield and the federal funds rate has strong predictive power for predicting excess bond returns and future changes in yields from the results of two different regressions. Finally, I find that short-maturity yields tend to rise and long-maturity yields tend to fall when the yield spreads widen, which is inconsistent with the expectations hypothesis.

CHAPTER II

YIELD FORECASTS AND STOCHASTIC VOLATILITY IN AFFINE MODELS WITH MACRO FACTORS

A. Introduction

Affine term structure models (ATSMs) appeal to both practitioners and academic researchers for tractability in econometric implementation and a sufficient degree of freedom in specifying how bond market compensates investors for taking systematic risk. Beginning with the pioneering studies by Vasicek (1977) and Cox, Ingersoll, and Ross (1985), researchers extended ATSMs to explain several important aspects of Treasury yields data. Some of them include:

- Treasury yields are persistent and move similarly over time. Most of yield variations are well explained by three to five factors.
- Treasury yields have time-varying volatilities which are high (low) when the levels of yields are high (low).
- Expected excess returns vary over time and the slope of yield curve has very good predictive power for expected excess returns.

Litterman and Scheinkman (1991) conduct an exploratory factor analysis of yields to show the first stylized fact. In response to this empirical finding, most affine models are now estimated in their multivariate forms. The strong evidence of time-varying volatility and volatility clustering is reported in various econometric works on the models of conditional heteroscedasticity. Term structure models incorporating stochastic volatility such as Cox, Ingersoll, and Ross (1985) can generate this feature. Lastly, Fama and Bliss (1987) and Campbell and Shiller (1991) are the first papers establishing the third stylized fact; their results imply that the expectations hypothesis of interest rates is rejected unless a relevant risk adjustment is made. Putting together, a multi-factor term structure model with time-varying market price of risk and stochastic volatility appears to explain all the stylized facts.

Recently, Dai and Singleton (2000), Duffee (2002), and Duarte (2004) reported that there exists a trade-off between improving forecast ability on future bond yields and matching interest rate volatility in affine models. Related, the models with stochastic volatility fail to account for the third stylized fact. Since the market price of risk setup can be modelled leaving the affine form of term structure intact as shown by Duffie and Kan (1996), a sufficiently flexible setup for the market price of risk can be one way to resolve these issues. Along this line, Duffee (2002) and Duarte (2004) use alternative parameterizations of the market price of risk and report some success in increasing the predictive power of ATSMs, but they still have difficulty in generating the third fact unless they dispense with stochastic volatility.

I attempt to reduce this statistical tension by incorporating some observable macroeconomic variables into ATSMs. Specifically, I use a measure of liquidity demand, the velocity of money. For comparison, I also examine measures of inflation and output gap which are now popular in macroeconomic term structure studies. It is well known that inflation, output, and/or the velocity of money are closely related to interest rates both empirically and theoretically. Obviously, the first two variables are motivated from either the Fisherian theory or the recent monetary policy rule literature, and the third variable is inferred by a typical liquidity demand theory.

The influential work by Ang and Piazzesi (2003) shows that incorporating monetary policy behaviors via inflation and output measure can help forecast bond yields in an affine model with a conditionally homoscedastic setup. I emphasize the importance of macro factors in modelling term structure dynamics as well, but I depart from their work in two directions. First, I explicitly model stochastic volatility to tackle the estimation and prediction issues mentioned above. Second, I focus on liquidity or money demand theory to impose an economic restriction to both bond yields and key macroeconomic variables and compare my results with other affine models. As Duffee (2002) points out, "[i]mposing these (economic) restrictions should allow us to explain more of the information in the current term structure, and thus forecasts" Furthermore, given that alternative economic restrictions are available, I believe that it is important to compare those in light of forecasting term structure dynamics. In so doing, it is also critical to verify if the models do not challenge the stylized facts on the conditional moments of bond yields.

In this vein, I ask two research questions. First, do macroeconomic variables help reduce the trade-off between matching the first and second moments of bond yields? Second, which affine model, either in a latent or a (macro-latent) hybrid factor structure, predicts future bond yields better? Regarding the first question, I find that the affine models with the velocity of money and stochastic volatility can explain all three stylized facts, while the purely latent factor models cannot. That is, the statistical "tension" is considerably relaxed in my model. In addition, consistent with this result, the affine model with the velocity of money predict future yields better than all other ATSMs examined, even compared with the models including inflation and real activity both in sample and out of sample. This finding is robust across different maturities, forecasting horizons, price of risk specifications, and the number of latent factors.

Chapter II is organized as follows. The next section shows some figures to motivate my study. Then I present my affine term structure model with latent and macroeconomic variables. Then, I explain my estimation method, followed by presenting estimation results. In so doing, I spell out econometric specifications. Then

I compare in-sample and out-of-sample forecasts, and check whether or not the improvement of forecasting performance helps to resolve the trade-off mentioned above. After some further discussions on my results, I conclude.

B. Interest Rates and Macroeconomic Variables

1. Data Description

My data set consists of two groups, bond yield data and macroeconomic data. Regarding bond yields, I use monthly yield series of the U.S. Treasuries with the maturities of 1, 3, 6, 12, 36 and 60 months, from June 1964 through December 2006, taken from the Fama-Bliss data file in the Center for Research in Security Prices (CRSP) data set. All bond yields are continuously compounded. Figure 2-1 plots monthly yields of maturity 1 month, 12 months and 60 months and Table 2-1 shows summary statistics for bond yields data. As stated in the introduction, the Treasury yields of different maturities move persistently and similarly. There exists a common variation of yields which is often called the level factor, followed by fluctuation related to the difference between long-term and short-term yields (the slope). In addition, one can notice that the differences between long-term and mid-term and between mid-term and short-term yields vary in a heterogeneous fashion over time. This is so called the curvature or the twist factor according to Litterman and Scheinkman (1991).

Figure 2-2 displays the time-varying volatility of the Treasury yields. The upper panel displays the band-pass filtered part of the Treasury yields with the maturities of 1, 12, and 60 months and the lower panel computes the realized volatility of those yields. It is clear that high volatilities of yields are matched with high levels of yields.

Three macro variables are used in this study. Those are the velocity of money, inflation, and output growth. All of the data come from the St. Louis Fed (FRED). For

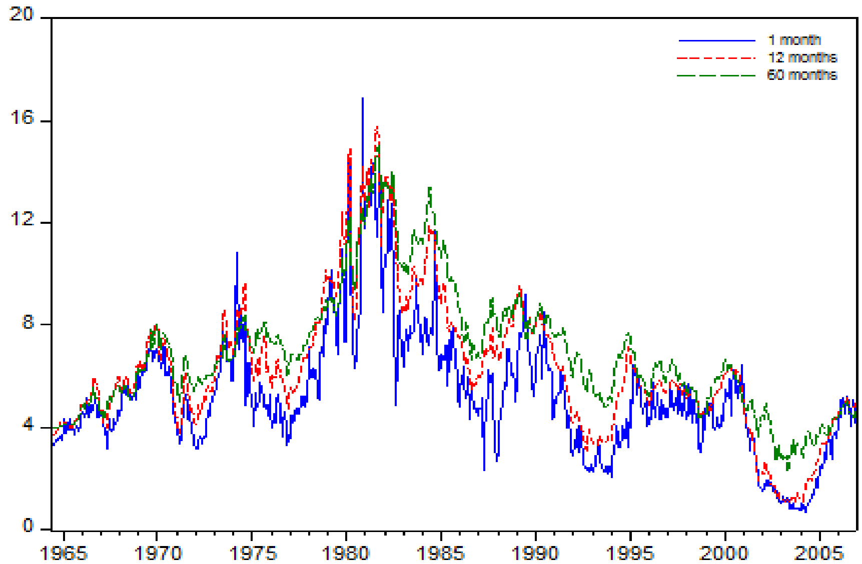

Fig. 2-1. US Interest Rates

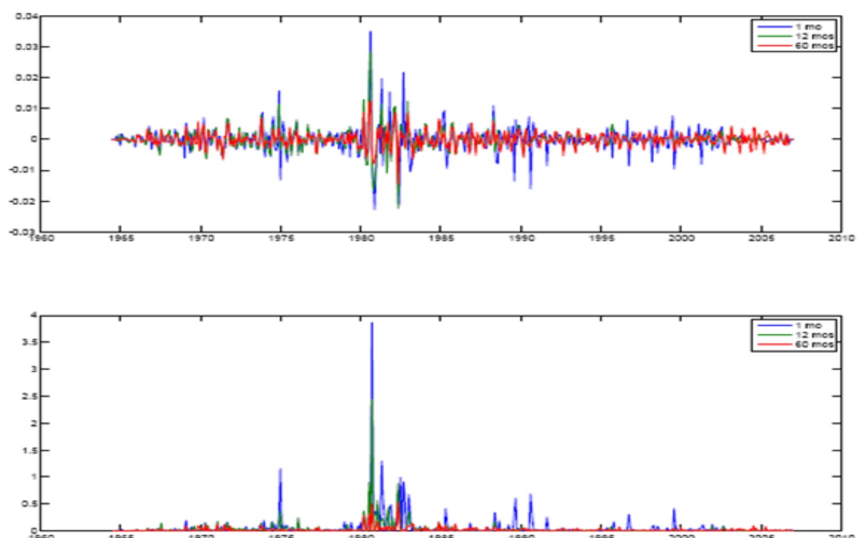

Fig. 2-2. Time Varying Volatility of Treasury Yields

the velocity of money, I use the money definition of M2 minus. This money aggregate is obtained from M2 subtracting small time deposits. Thus, M2 minus entails currency, demand/checkable deposits, savings accounts, money market deposit accounts (MMDA), and retail money market fund (RMMF). Most of the assets do pay some interest for holding, but they are close to zero. More importantly, these assets allow check writing with no or very small transactions fee, and therefore can be regarded as good substitutes for cash. Emergence of these monetary assets is mainly due to financial innovations and deregulation since the late 1960s. Thus, without considering this, measuring money demand using traditional definitions such as monetary base or M1 will be misleading in light of capturing transactions motive.[1] I calculate the inflation measure using $\log(\mathrm{P}_t/\mathrm{P}_{t-12})$ where P_t is the consumer price index (CPI). The output growth is calculated as annual growth of industrial production.[2]

Figure 2-3 shows the three macro variables representing the velocity of money, inflation, and output growth and Table 2-1 presents summary statistics for the data. In Figure 2-3, I also mark recession periods recorded by the National Bureau of Economic Research (NBER) as shaded areas to indicate that macroeconomic variables have considerable co-movements around recession periods. In the next section, I explain economic theories that I will use to analyze affine term structure models.

[1]Instead of using a broader concept of money, one can alternatively use a more narrow definition of money such as M1 with an additional term or a function describing the financial innovations. Since I focus on liquidity demand side rather than money supply side, I simply adopted the former method.

[2]I also experimented using other macroeconomic variables reflecting real activities. They show very little differences.

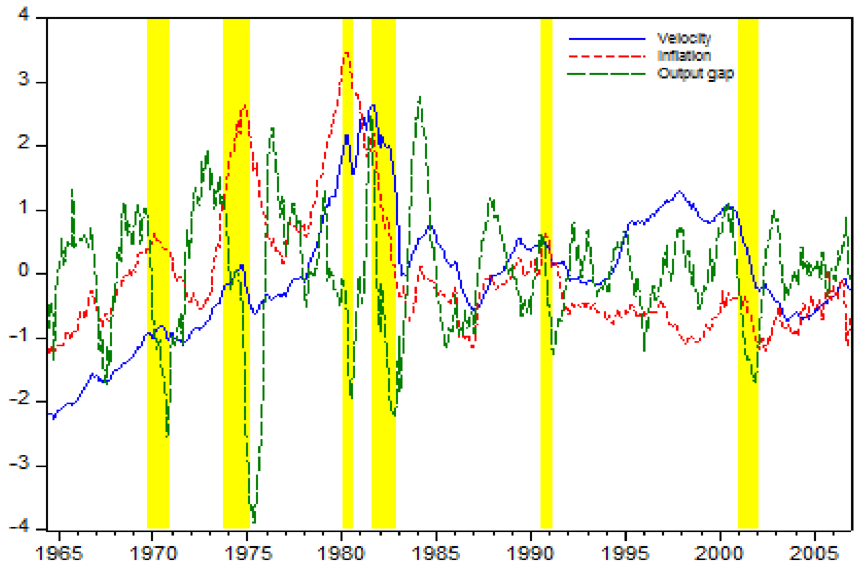

Fig. 2-3. Macroeconomic Variables and Business Cycles

Table 2-1. Summary Statistic of Data

	1 mth	3 mth	6 mth	12 mth	36 mth	60 mth	velo	inf	output
Mean	5.295	5.835	6.086	6.333	6.724	6.940	4.037	4.547	0.017
Std	2.483	2.714	2.776	2.758	2.582	2.481	0.675	2.898	3.111
Skew	1.122	1.015	0.935	0.804	0.846	0.888	0.132	1.484	-0.578
Kurt	5.497	4.832	4.487	4.067	3.808	3.598	3.065	4.696	4.742
Auto(1)	0.916	0.975	0.981	0.980	0.984	0.985	0.993	0.991	0.946
Auto(12)	0.703	0.758	0.775	0.790	0.820	0.835	0.826	0.746	-0.304
Correlation									
	1 mth	3 mth	6 mth	12 mth	36 mth	60 mth	velo	inf	output
1 mth	1.000								
3 mth	0.954	1.000							
6 mth	0.947	0.994	1.000						
12 mth	0.933	0.984	0.994	1.000					
36 mth	0.880	0.941	0.956	0.977	1.000				
60 mth	0.841	0.906	0.923	0.949	0.993	1.000			
velocity	0.549	0.594	0.591	0.591	0.611	0.613	1.000		
inflation	0.655	0.686	0.696	0.672	0.618	0.591	0.408	1.000	
output gap	0.051	0.051	0.053	0.056	0.030	0.023	0.003	-0.133	1.000

2. Liquidity Demand and Monetary Policy

A recent trend in studying term structure dynamics is to include inflation and real activity variables as observable macro factors. This can be understood as an attempt to incorporate monetary policy behavior into term structure models. According to Taylor (1993), rule-like behaviors appear to approximate actual monetary policy changes. Especially, Taylor suggests the following form

$$r_t = \theta_0 + \pi_t + \theta_\pi(\pi_t - \pi_t^*) + \theta_g g_t, \tag{2.1}$$

where r_t is the short-term interest rate that the central bank can control, π_t is a measure of inflation, and g_t is a measure of output gap. A caveat related to the monetary policy literature is that interpretations of any results based on this approach hinge upon how legitimate this type of linear policy rules is in light of describing actual policy behaviors. If this rule appeals mainly to the normative side of monetary policy, this may not be suitable for explaining and forecasting yield dynamics. In addition, Fisher hypothesis tells that nominal rate is approximately the sum of (expected) inflation and real interest rate. Thus, the identification of a policy rule is somewhat ambiguous. Another issue with this policy rule is that expected variables, not the current ones are the relevant target variables. Of course, there are several versions of forward-looking rules available in the literature. However, since expected inflation and expected output are not observable, using those forward looking rules to study the term structure of interest rates is not a simple task. In this sense, estimated parameters such as θ_π and θ_g using current variables may entail information about conditional expectation on future economy as well as target behaviors of the central bank. This can cloud the issue of monetary policy.

Nevertheless, the main advantage of this approach is that I can see how yields are

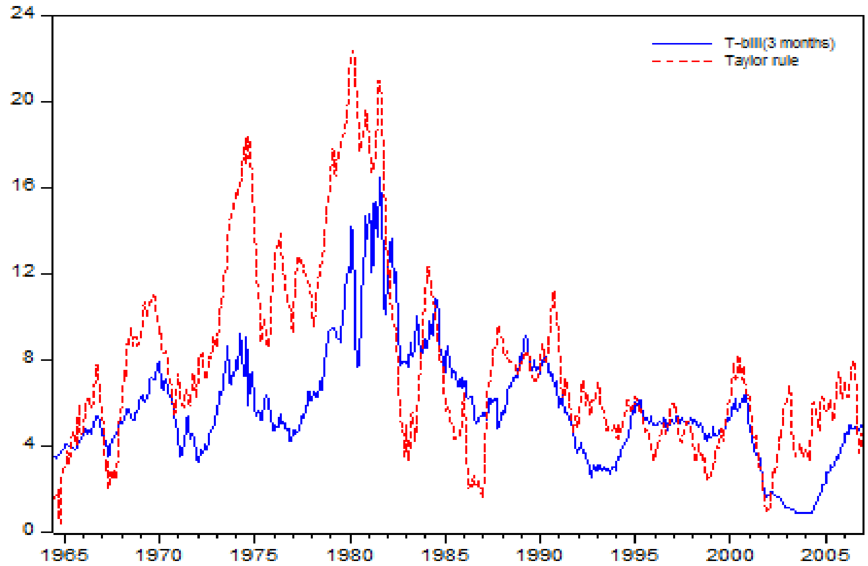

Fig. 2-4. Treasury Bill (3 Months) and Taylor Rule

linked to some of the fundamental macroeconomic variables such as output growth and inflation. In Figure 2-4 I plot the policy rule suggested by Taylor (1993) in comparison with Treasury yield of three months maturity. The policy rule seems to delineate the long-term trend of the bond yield well, though it is clearly more volatile than the movement of actual yields.

An alternative way to impose a restriction on macro variables and interest rates is using a money demand relationship

$$\frac{M_t}{p_t y_t} = L(r_t, z_t), \tag{2.2}$$

where M/p is the real money balances, y_t is the output, and z_t is a vector of other variables affecting the money demand. If properly defined, monetary variables can reflect demand for transactions services. Specifically, transactions services can be well represented by monetary assets providing low and stable interest rates close to zero. For this purpose, I use M2 minus small time deposits, so called 'M2 minus' to

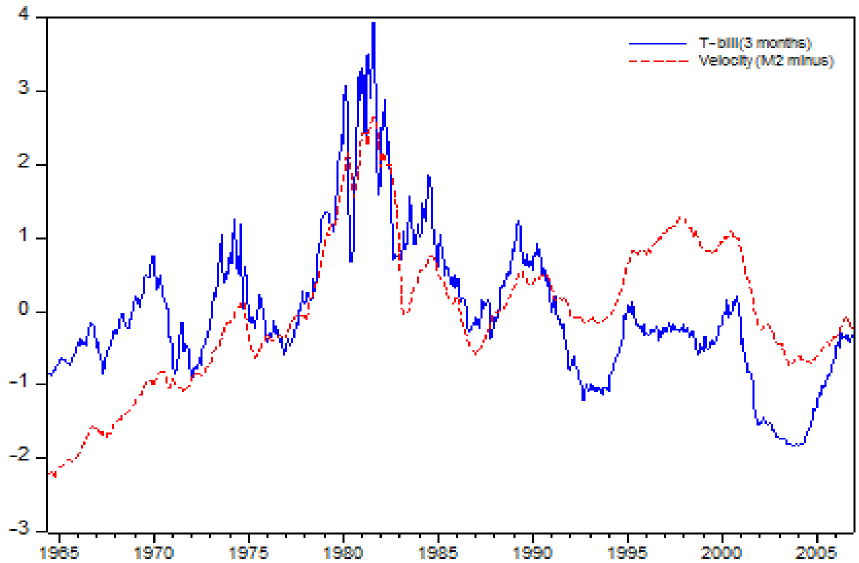

Fig. 2-5. Treasury Bill (3 Months) and Velocity of Money (M2 Minus)

measure the liquidity demand of an economy. Figure 2-5 indicates that the velocity of money ($p_t y_y / M_t$) and the short-term interest rate move closely together over the post-war period. As long as the opportunity cost of holding these liquid assets varies closely with policy rate changes, I expect monetary variables to be inversely related to monetary policy changes. That is, under a stable money demand relationship, changes in the money demand will have close link to monetary policy behaviors as well. In this sense, including monetary variables in estimating term structure dynamics can be regarded as an attempt to account for yield movements resulting from changes in monetary policy behaviors via the lens of liquidity demand.

Both relationships offer strong theoretical links between interest rates and macroeconomic variables. Therefore, it is worthwhile to examine these two alternative macroeconomic restrictions in terms of explaining the stylized facts of bond yields. Especially, I are interested in the role of macro factors in resolving the statistical trade-off between enhancing forecasting performance and fitting the time-varying volatility of bond yields.

C. Model

1. A Generic Affine Setup with Macro Variables

I present an affine term structure model with latent factors and observable macroeconomic variables. I denote a state variable vector by $X_t = (x_{1t}, \cdots, x_{kt}, x_{k+1t}, \ldots, x_{nt})'$, where the first k factors are unobservable and the remaining $(n-k)$ factors are macroeconomic variables. Suppose that X_t follows an Ito process

$$dX_t = K[\Theta - X_t]dt + \sqrt{S_t}dW_t, \tag{2.3}$$

where K is an $n \times n$ matrix, Θ is an $n \times 1$ vector. S_t is an $n \times n$ diagonal matrix and the ith diagonal element is given as $\alpha_i + \beta_i' X_t$, where α_i is a constant and β_i is an $n \times 1$ vector. W_t is an $n \times 1$ vector of independent Brownian motions under risk natural measure. Sans arbitrage opportunities, the price at t of a zero coupon bond maturing at τ denoted as $P_t(\tau)$ is expressed under the risk-neutral measure Q as

$$P_t(\tau) = \mathbb{E}_t^Q \left[e^{-\int_t^\tau r_u du} \right], \tag{2.4}$$

where r_t is the instantaneous short-term interest rate process. An affine term structure model implies that r_t is an affine function of X_t and the evolution of X_t under Q measure follows another affine diffusion,

$$dX_t = \tilde{K}[\tilde{\Theta} - X_t]dt + \sqrt{S_t}d\tilde{W}_t, \tag{2.5}$$

where $\tilde{W}_t$ is an $n \times 1$ vector of independent Brownian motions under the risk neutral Q measure. Changes between two measures are possible via Girsanov transformation under Novikov conditions. In this way, I can define a market price of risk Λ_t using the relationship $dW_t = d\tilde{W}_t - \Lambda_t dt$. While maintaining the affine diffusion property, I use the market price of risk specification proposed by Duffee (2002)

$$\Lambda_t = \sqrt{S_t}\lambda_1 + Z_t \lambda_2 X_t \tag{2.6}$$

where λ_1 is an $n \times 1$ vector, λ_2 is an $n \times n$ matrix, and Z_t is a diagonal $n \times n$ matrix with

$$Z_{t(ii)} = \begin{cases} \frac{1}{\sqrt{\alpha_i + \beta_i' X_t}} & \text{if } \inf(\alpha_i + \beta_i' X_t) > 0 \\ 0 & \text{otherwise} \end{cases}.$$

This setup is more flexible than the traditional one with λ_2 being null matrix and thus can better capture the time-variability of term premia. Dai and Singleton (2000) and Duffee (2002) provide an admissible class of this model by showing sufficient conditions for the existence of X_t for alternative specifications which restrict the parameter matrices and vectors. Under some additional technical conditions, Duffie and Kan (1996) show that the bond prices are exponentially affine as

$$P_t(\tau) = e^{A(\tau) - B(\tau)' X_t}, \tag{2.7}$$

$$r_t = \sum_{i=1}^{n} x_{it}, \tag{2.8}$$

where $A(\tau)$ and $B(\tau)$ satisfy the ordinary differential equation (ODE) system of

$$\frac{\partial B(\tau)}{\partial \tau} = 1_{n\times 1} - \tilde{K}' B(\tau) - \frac{1}{2} \sum_{i=1}^{n} [B(\tau)]_i^2 \, \beta_i \tag{2.9}$$

$$\frac{\partial A(\tau)}{\partial \tau} = -\tilde{\Theta}' \tilde{K}' B(\tau) - \frac{1}{2} \sum_{i=1}^{n} [B(\tau)]_i^2 \, \alpha_i. \tag{2.10}$$

This can be easily solved through a numerical method using the initial condition of $A(0) = 0$ and $B(0) = 0_{n\times 1}$.

D. Estimation Method

I use the Kalman filter approach to estimate the term structure models I proposed in the previous section. This method in term structure estimation was used in Pennacchi (1991), Lund (1997), and Duan and Simonato (1999).[3] In non-Gaussian settings where the exact form of conditional density for the state vector is not known, an approximate linear filtering can be used and Monte Carlo results have shown that this method performs well (Duffee and Stanton (2004)). Implementation of this filtering relies on the availability of the first two conditional moments of the state variables and hence works well with most of the affine term structure models. In the below, I explain my estimation procedure more in detail.

The affine term structure model is re-casted in a state space setup as the measurement equations and transition equations

$$\begin{pmatrix} y_t^{\tau_1} \\ y_t^{\tau_2} \\ y_t^{\tau_3} \\ y_t^{\tau_4} \\ \vdots \\ y_t^{\tau_N} \end{pmatrix} = \begin{pmatrix} -(1/\tau_1)A(\tau_1) \\ -(1/\tau_2)A(\tau_2) \\ -(1/\tau_3)A(\tau_3) \\ -(1/\tau_4)A(\tau_4) \\ \vdots \\ -(1/\tau_N)A(\tau_N) \end{pmatrix} + \begin{pmatrix} (1/\tau_1)B(\tau_1) \\ (1/\tau_2)B(\tau_2) \\ (1/\tau_3)B(\tau_3) \\ (1/\tau_4)B(\tau_4) \\ \vdots \\ (1/\tau_N)B(\tau_N) \end{pmatrix} X_t + \begin{pmatrix} \varepsilon_{t,1} \\ \varepsilon_{t,2} \\ \varepsilon_{t,3} \\ \varepsilon_{t,4} \\ \vdots \\ \varepsilon_{t,N} \end{pmatrix}, \tag{2.11}$$

$$dx_{it} = \kappa_i(\theta_i - x_{it})dt + \sqrt{\alpha_i + \beta_i x_{it}}dW_{it},\ i = 1, 2, \cdots, n. \tag{2.12}$$

[3] There is another approach developed by Chen and Scott (1993), which basically use the same idea of transforming yields into latent factors. In comparision, the Kalman approach allows measurement errors for all the yields observed, while the former method assumes that k yields are observed without an error for a k-factor model.

In order to use the discrete-time state space model technique, I write down the transition equation as

$$x_{it} = a_i(\psi, \Delta) + b_i(\psi, \Delta)x_{it-\Delta} + \eta_t \qquad \eta_t \sim N(0, \Phi_i(x_{it-\Delta}\ ; \psi, \Delta)), \tag{2.13}$$

where I know that

$$E(x_{it}|x_{it-\Delta}) = a_i(\psi, \Delta) + b_i(\psi, \Delta)x_{it-\Delta}, \tag{2.14}$$

$$Var(x_{it}|x_{it-\Delta}) = \Phi_i(x_{it-\Delta}; \psi, \Delta). \tag{2.15}$$

For affine diffusion models, I can compute the closed form solutions for $E(x_{it}|x_{it-\Delta})$ and $Var(x_{it}|x_{it-\Delta})$ as affine functions according to Fisher and Gilles (1996) and Duan and Simonato (1999). For my case, I compute those as

$$E(x_{it}|x_{it-\Delta}) = (1 - e^{-\kappa_i \Delta})\theta_i + e^{-\kappa_i \Delta}x_{it-\Delta} \tag{2.16}$$

$$Var(x_{it}|x_{it-\Delta}) = \frac{\alpha_i}{2\kappa_i}(1 - e^{-2\kappa_i \Delta}) + \theta_i \frac{\beta_i}{2\kappa_i}(1 - e^{-2\kappa_i \Delta}) + (x_{it-\Delta} - \theta_i)\frac{\beta_i}{\kappa_i}(e^{-\kappa_i \Delta} - e^{-2\kappa_i \Delta}) \tag{2.17}$$

This method, however, cannot be used in case of more flexible market price of setup such as Duarte (2004) because the drift of the state vector process under physical measure may be non-linear. Although my model is essentially affine as in Duffee (2002) such that the above method is still applicable, I suggest another method using Milstein approximation of diffusion process. The basic idea of the Milstein approximation can be explained as follows. Given the discrete nature of available data, I oftentimes need to discretize a diffusion process for estimation. The simplest method would be Euler scheme. Specifically, suppose I have

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t \tag{2.18}$$

Then, I can approximate this process as

$$X_t - X_{t-\Delta} \approx \mu(X_{t-\Delta})\Delta + \sigma(X_{t-\Delta})(W_t - W_{t-\Delta}) + O(\Delta) \tag{2.19}$$

for a small time interval Δ. Then I have the first two conditional moments as

$$E(X_t|X_{t-\Delta}) = \mu(X_{t-\Delta})\Delta + X_{t-}\Delta \tag{2.20}$$

$$Var(X_t|X_{t-\Delta}) = E([X_t - X_{t-\Delta} - \mu(X_{t-\Delta})\Delta]^2) = \sigma^2(X_{t-\Delta})\Delta \tag{2.21}$$

The Euler approximation may be reasonable if I have data sampled at a sufficiently high frequency, but otherwise it is biased. Instead, I can use a Milstein approximation in which the diffusion process is approximated by

(2.22)

$$X_t - X_{t-\Delta} \approx \mu(X_{t-\Delta})\Delta + \sigma(X_{t-\Delta})(W_t - W_{t-\Delta}) + \frac{1}{2}\sigma\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\left[(W_t - W_{t-\Delta})^2 - \Delta\right] + o(\Delta)$$

In this case, I can obtain the first two conditional moments as[4]

$$E(X_t|X_{t-\Delta}) = \mu(X_{t-\Delta})\Delta + X_{t-\Delta} \tag{2.23}$$

$$Var(X_t|X_{t-\Delta}) = \sigma^2(X_{t-\Delta})\Delta + \frac{1}{2}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2 \tag{2.24}$$

Using these I can calculate the conditional mean and variance of state variables for my model as

$$E(x_{it}|x_{it-\Delta}) = \theta_i\kappa_i\Delta + (1 - \kappa_i\Delta)x_{it-\Delta} \tag{2.25}$$

$$Var(x_{it}|x_{it-\Delta}) = (\alpha_i + \beta_i x_{it-\Delta})\Delta + \frac{1}{8}\beta_i^2\Delta^2 \tag{2.26}$$

This implies that the transition equation over a discrete time interval can be written

[4]Derivations are in Appendix D.

as

$$x_{it} = a_i(\psi,\Delta) + b_i(\psi,\Delta)x_{it-\Delta} + \eta_t \qquad \eta_t \sim N(0,\Phi_i(x_{it-\Delta};\psi,\Delta)), \tag{2.27}$$

where

$$a_i(\psi,\Delta) = \theta_i\kappa_i\Delta \tag{2.28}$$

$$b_i(\psi,\Delta) = 1-\kappa_i\Delta \tag{2.29}$$

$$\Phi_i(x_{it};\psi,\Delta) = (\alpha_i+\beta_i x_{it})\Delta + \frac{1}{8}\beta_i^2\Delta^2 \tag{2.30}$$

Given the linearity of $a_i(\psi,\Delta)+b_i(\psi,\Delta)x_{it-\Delta}$ and $\Phi_i(x_{it}\,;\psi,\Delta)$, I can use the Kalman filter recursion to obtain a prediction error decomposition in expressing and evaluating the log of the quasi-likelihood function recursively.

For notational simplicity, this model in the state space form can be expressed as follows:

$$Y_t = A + BX_t + \varepsilon_t \tag{2.31}$$

$$X_t = a + bX_{t-1} + \eta_t \tag{2.32}$$

$$E(\varepsilon_t)=0,\;\; E(\eta_t)=0,\;\; E(\varepsilon_t\varepsilon_t')=R,\;\; E(\eta_t\eta_t')=Q_t \tag{2.33}$$

Let $\hat{X}$ and $\hat{Y}$ be the estimates of the n state variables and bond yields with N different maturities. In addition, let P_t and V_t be the covariance matrices of the estimation errors respectively. Then, given $\hat{X}_{t-1}$ and P_{t-1}, I can compute the one-period-ahead prediction using $\hat{X}_{t/t-1} = a+b\hat{X}_{t-1/t-1}$ and $\hat{Y}_{t/t-1} = A+B\hat{X}_{t/t-1}$ and the one-period-ahead covariance matrices, $P_{t/t-1} = bP_{t-1/t-1}b' + Q_t$ and $V_{t/t-1} = BP_{t/t-1}B' + R$. When Y_t is observable, I update the prediction, $\hat{X}_{t/t} = \hat{X}_{t/t-1} + P_{t/t-1}B'V_{t/t-1}^{-1}(Y_t - \hat{Y}_{t/t-1})$ and the covariance matrix, $P_{t/t} = P_{t/t-1} - P_{t/t-1}B'V_{t/t-1}^{-1}BP_{t/t-1}$. With $e_t =$

$Y_t - \hat{Y}_{t/t-1}$, the log of the quasi-likelihood function is given by

$$\log L(Y;\,\psi) = \sum_{t=1}^{T} -\frac{1}{2}[N\log(2\pi) + \log(\det V_{t/t-1}) + e_t' V_{t/t-1}^{-1} e_t] \tag{2.34}$$

The quasi-likelihood function can be evaluated through the usual Kalman filtering recursive procedure. The estimated parameters are those which maximize the log of the quasi-likelihood function.

E. Empirical Results

1. Estimated Models

Since the main objective of the paper is to evaluate empirical performance of affine term structure model with the velocity of money, I will estimate various versions of the affine model for comparison. Broadly, I have three groups of the model: i) yield factor models, ii) hybrid factor models including the velocity of money, and iii) hybrid factor models with inflation and output growth. Within each group, I have different model specifications depending on the number of factors, the number of independent volatility drivers, and the setup for market price of risk.

Regarding the number of latent factors, I have two to three unobservable factors with one to two macro factors. The biggest model I estimate is a five-factor version which includes three latent factors and two macro factors. Another criterion for classifying the estimated models is the number of independent state variables determining the variance-covariance of the state vector X_t. As mentioned in the introduction, there exists a tension between matching conditional mean and conditional volatility of yields in that smaller number of instantaneous volatility drivers enhances the forecast ability of the model. Thus, I examine whether or not the inclusion of macroeconomic variables, especially the velocity of money helps relax this tension.

At the same time, the models need to be checked if they are able to explain the term premium variability.

For expositional purposes, I label the estimated models using notation similar to Dai and Singleton (2000). $\mathbb{A}_M(n, j; I)$ refers to an affine model with the total number n of factors, j macroeconomic variables, M state variables affecting the instantaneous volatility, and the market price of risk specification of Duffee (2002). The last element I in the model $\mathbb{A}_M(n, j, I)$ displays the macroeconomic variables included in the estimated model. This entails the velocity of money (v), inflation (π), and output growth (g). To distinguish the models in terms of risk price specifications, I also use the notation $C\mathbb{A}_M(n, j, I)$ referring to a completely affine model which employs the risk price specification à la Dai-Singleton (i.e. $\lambda_2 = 0$). $\mathbb{A}_M(M, 0)$ models are basically multi-factor Cox-Ingersoll-Ross model with Duffee's price of risk setup. Specifically I estimate the following cases:

- Three-factor latent model: $\mathbb{A}_1(3, 0)$, $C\mathbb{A}_1(3, 0)$, $\mathbb{A}_3(3, 0)$
- Three-factor hybrid model: $\mathbb{A}_1(3, 1; \{v\})$, $C\mathbb{A}_1(3, 1; \{v\})$, $\mathbb{A}_3(3, 1; \{v\})$
- Four-factor hybrid model: $\mathbb{A}_1(4, 1; \{v\})$, $C\mathbb{A}_1(4, 1; \{v\})$, $\mathbb{A}_2(4, 1; \{v\})$, $C\mathbb{A}_2(4, 1; \{v\})$, $\mathbb{A}_4(4, 1; \{v\})$, $\mathbb{A}_4(4, 2; \{\pi, g\})$, $\mathbb{A}_4(4, 2; \{v, g\})$
- Five-factor hybrid model: $\mathbb{A}_5(5, 2; \{\pi, g\})$, $\mathbb{A}_5(5, 2; \{v, g\})$

I impose restrictions (mostly zero restrictions) on the parameters of the models above. Some of those come from the representation of the affine model as in Dai and Singleton (2000), Duffee (2002), Ang and Piazzesi (2003), Duarte (2004), and Aït-Sahalia and Kimmel (2008). In addition, I make additional assumptions on the parameters to make the model parsimonious. Four-factor hybrid versions are regarded as my main model and I compare those with the three-factor latent models $\mathbb{A}_M(3, 0)$

and five-factor hybrid models $\mathbb{A}_M(5,2;\{\pi,g\})$ which are popular choices among the affine models.

2. Parameter Estimates

This section reports parameter estimates of the several affine models I estimated. Table A-1 $\sim$ Table A-3 in Appendix A display the estimates of the selected three-, four-, and five-factor models together with standard errors given in parentheses. Table A-1 entails two latent models: $\mathbb{A}_1(3,0)$, $\mathbb{A}_3(3,0)$, and two macro hybrid models: $\mathbb{A}_1(3,1;\{v\})$, $\mathbb{A}_3(3,1;\{v\})$. Most parameters in all cases are estimated reliably. A major difference between the latent models and the macro hybrid models in case of three-factor model is that the latter has two persistent parameters (κ_1 and κ_v), while the latent, three-factor models have one persistent factor (κ_1). I follow Duffee (2002) for zero restrictions on the market price of risk in case of the essentially affine setups and other restrictions come from Dai and Singleton (2000). Four and five factor models are estimated in an identical manner except the number of macro factors I incorporated. Table A-2 shows the estimation results from the four-factor models. For conserving space, I do not provide all the results. Two of those have three latent factors and one macro factor (money velocity), while the third one displayed in the table has two macro factors, money velocity and output growth.[5] Similar to the three-factor case, four-factor models also have two relatively persistent factors, one of which is the velocity of money. For five-factor models, I report only one case with inflation and output gap as macro factors. Regardless of the number of factors, macro factors seem to play an important role in capturing the time-variability of both bond risk premia (λ_1 and λ_2) and conditional volatility of yields (β). In the next section, I

[5] I do not include the result with inflation and the output gap in case of four factor model due to its relative weak performance.

verify if this is indeed the case. In addition, I evaluate the accuracy of yield forecasts these models generate both in and out of sample.

3. Yield Forecast, Term Premium, and Stochastic Volatility

My data set covers the periods between June 1964 and December 2006. To produce out-of-sample forecasts, I estimate my models using the in-sample period of June 1964 through December 2003. Then, I forecast future bond yields during the out-of-sample period of January 2004 through December 2006. I compute the root mean squared errors (RMSE) from both the in-sample and the out-of-sample forecasts and display in Table 2-2 $\sim$ Table 2-6 respectively. Lower RMSE values (%) indicate better forecasts. I report the results for bonds with maturities of 1, 3, 6 months, 1 year, 3 years, and 5 years. For out-of-sample forecast horizons, 1, 3, 6, and 12 months are used.

In Table 2-2 $\sim$ Table 2-6, the most significant result would be that the models showing the lowest RMSE for each maturity are uniformly the ones with the velocity of money. Although there is not a single specification dominating others, most of four-factor models perform very well. Still, the forecasts from the essentially affine model ($\mathbb{A}_1(3,0)$) are comparable to those of the four-factor hybrid models. A notable exception is the forecast horizon of twelve months for which a three factor model with the velocity of money ($\mathbb{A}_3(3,1;\{v\})$) shows the best performance. Its error reduction over a three-factor essentially affine model ($\mathbb{A}_1(3,0)$) is up to 47 basis point! These results are maintained even if I use a different accuracy measure such as mean absolute deviation (MAD). Macro models with inflation and output gap also show decent, often better performance compared to latent factor models. But the models with the money velocity still outperform in all horizons. In addition, the results show that the number of factors is not a major ingredient boosting up forecast ability.

Table 2-2. In-sample Fit Performance Using RMSE

Model \ Maturity	1 month	3 months	6 months	12 months	36 months	60 months
$\mathbb{A}_1(3,0)$	0.6386	0.1682	0.1931	0.1159	0.2150	0.0517
$C\mathbb{A}_1(3,0)$	0.5732	0.1773	0.1749	0.1677	0.2057	0.0961
$\mathbb{A}_3(3,0)$	0.7016	0.1569	0.2164	0.1282	0.2039	0.2321
$\mathbb{A}_1(3,1;\{v\})$	0.7246	0.2491	0.2532	0.1731	0.2084	0.2042
$C\mathbb{A}_1(3,1;\{v\})$	0.7612	0.2949	0.2588	0.2228	**0.0827**	0.1820
$\mathbb{A}_3(3,1;\{v\})$	0.8011	0.2388	0.2428	0.2076	0.1225	0.1770
$\mathbb{A}_1(4,1;\{v\})$	0.6015	0.1707	0.1762	0.1366	0.1665	0.1360
$C\mathbb{A}_1(4,1;\{v\})$	0.6620	**0.1332**	0.1722	0.0853	0.1095	0.1578
$\mathbb{A}_2(4,1;\{v\})$	0.5455	0.3063	**0.1661**	0.0751	0.1444	**0.0135**
$C\mathbb{A}_2(4,1;\{v\})$	**0.4801**	0.3316	0.1703	0.1053	0.1421	0.0268
$\mathbb{A}_4(4,1;\{v\})$	0.7087	0.1403	0.1863	**0.0643**	0.1283	0.0235
$\mathbb{A}_4(4,2;\{\pi,g\})$	0.7841	0.2987	0.2588	0.1978	0.2303	0.2550
$\mathbb{A}_4(4,2;\{v,g\})$	0.8358	0.2564	0.2464	0.2148	0.1714	0.2327
$\mathbb{A}_5(5,2;\{\pi,g\})$	0.7810	0.1358	0.2062	0.1193	0.1497	0.1371
$\mathbb{A}_5(5,2;\{v,g\})$	0.7352	0.1385	0.1930	0.1121	0.1293	0.0250

Ang and Piazzesi (2003) find that including observable variables especially measures of inflation and real activity helps forecasting in an affine model with conditional homoscedasticity. Specifically, they report that their hybrid five-factor models can beat random walk and unconstrained vector auto regression models. My results show that macro hybrid models with the velocity of money and stochastic volatility predict future yields better than all the other models I considered. Given my focus on matching the conditional mean and the conditional volatilities of yields, I would be interested in further investigating if the macro affine models with stochastic volatility are able to generate yield forecasts consistent with the third stylized fact.

Toward this end, I run the following regressions echoing Fama and Bliss (1987) and Campbell and Shiller (1991),

$$y^{\tau}_{t+k} - E_t[y^{\tau}_{t+k}] = \gamma_0 + \gamma_1 \times \text{term spread} + \text{error}, \tag{2.35}$$

where $E_t[y^{\tau}_{t+k}]$ is computed as a forecast out of sample, using a term structure model

Table 2-3. Out-of-sample Forecasts Performance Using RMSE (Forecast Horizon: 1 month)

Model\Maturity (months)	1	3	6	12	36	60
$\mathbb{A}_1(3,0)$	0.3638	0.1634	0.1879	0.1712	0.2620	0.2486
$C\mathbb{A}_1(3,0)$	0.3684	0.1989	0.2184	0.1795	0.2566	0.2561
$\mathbb{A}_3(3,0)$	0.4965	0.3027	0.2566	0.2606	0.2589	0.2633
$\mathbb{A}_1(3,1;\{v\})$	**0.3601**	0.1641	0.2503	0.2685	0.2468	0.2803
$C\mathbb{A}_1(3,1;\{v\})$	0.3755	0.1680	0.2799	0.3314	0.2523	0.2857
$\mathbb{A}_3(3,1;\{v\})$	0.4374	0.2020	0.1820	0.2372	0.2562	0.3032
$\mathbb{A}_1(4,1;\{v\})$	0.3641	0.1571	0.1530	**0.1689**	0.2550	0.2516
$C\mathbb{A}_1(4,1;\{v\})$	0.3659	0.1690	0.1759	0.1856	0.2690	0.2555
$\mathbb{A}_2(4,1;\{v\})$	0.4006	**0.1483**	0.1540	0.1801	**0.2353**	0.2496
$C\mathbb{A}_2(4,1;\{v\})$	0.3607	0.1534	0.1628	0.1716	0.2559	0.2426
$\mathbb{A}_4(4,1;\{v\})$	0.4211	0.2096	**0.1412**	0.1692	0.2510	0.2444
$\mathbb{A}_4(4,2;\{\pi,g\})$	0.4802	0.2427	0.2316	0.2598	0.2574	0.2808
$\mathbb{A}_4(4,2;\{v,g\})$	0.5096	0.2711	0.2460	0.2859	0.3046	0.3224
$\mathbb{A}_5(5,2;\{\pi,g\})$	0.5281	0.3052	0.2157	0.1995	0.2545	0.2676
$\mathbb{A}_5(5,2;\{v,g\})$	0.4766	0.2774	0.2193	0.1989	0.2469	**0.2426**

Table 2-4. Out-of-sample Forecasts Performance Using RMSE (Forecast Horizon: 3 months)

Model\Maturity (months)	1	3	6	12	36	60
$\mathbb{A}_1(3,0)$	0.3736	0.2468	0.3291	0.3358	0.4068	0.4236
$C\mathbb{A}_1(3,0)$	0.4656	0.3808	0.4425	0.3887	0.4058	0.4354
$\mathbb{A}_3(3,0)$	0.8463	0.6559	0.5326	0.4900	0.4318	0.4106
$\mathbb{A}_1(3,1;\{v\})$	0.4055	0.3647	0.4716	0.4798	0.4289	0.4338
$C\mathbb{A}_1(3,1;\{v\})$	0.4466	0.4418	0.5652	0.5752	0.4441	0.4299
$\mathbb{A}_3(3,1;\{v\})$	0.5239	0.3379	0.2941	0.3444	**0.3836**	0.4160
$\mathbb{A}_1(4,1;\{v\})$	0.3691	**0.2299**	0.3156	0.3566	0.4307	0.4204
$C\mathbb{A}_1(4,1;\{v\})$	0.4128	0.3235	0.4144	0.4385	0.4707	0.4255
$\mathbb{A}_2(4,1;\{v\})$	**0.3636**	0.2386	0.3251	0.3652	0.4174	0.4112
$C\mathbb{A}_2(4,1;\{v\})$	0.3936	0.2843	0.3621	0.3805	0.4336	0.4093
$\mathbb{A}_4(4,1;\{v\})$	0.5666	0.3701	**0.2714**	**0.3103**	0.4124	0.4081
$\mathbb{A}_4(4,2;\{\pi,g\})$	0.6241	0.5165	0.4884	0.5184	0.5534	0.5520
$\mathbb{A}_4(4,2;\{v,g\})$	0.6280	0.4554	0.4207	0.4250	0.3867	**0.3925**
$\mathbb{A}_5(5,2;\{\pi,g\})$	0.8054	0.6400	0.5084	0.4951	0.5602	0.5630
$\mathbb{A}_5(5,2;\{v,g\})$	0.7038	0.5068	0.3967	0.3606	0.4122	0.4067

Table 2-5. Out-of-sample Forecasts Performance Using RMSE (Forecast Horizon: 6 months)

Model\Maturity (months)	1	3	6	12	36	60
$\mathbb{A}_1(3,0)$	0.4999	0.4246	0.4997	0.4715	0.4475	0.4850
$C\mathbb{A}_1(3,0)$	0.7171	0.6819	0.7221	0.6091	0.4747	0.5081
$\mathbb{A}_3(3,0)$	1.1567	0.8873	0.6920	0.5969	0.4407	0.3834
$\mathbb{A}_1(3,1;\{v\})$	0.7160	0.6774	0.7623	0.7145	0.4793	0.4133
$C\mathbb{A}_1(3,1;\{v\})$	0.8552	0.8397	0.9233	0.8581	0.5359	0.4200
$\mathbb{A}_3(3,1;\{v\})$	0.6904	0.4494	0.3548	0.3827	**0.3501**	**0.3551**
$\mathbb{A}_1(4,1;\{v\})$	0.5208	0.4657	0.5655	0.5692	0.5054	0.4562
$C\mathbb{A}_1(4,1;\{v\})$	0.7222	0.7139	0.8134	0.7844	0.6113	0.4737
$\mathbb{A}_2(4,1;\{v\})$	**0.4844**	**0.4136**	0.5140	0.5274	0.4876	0.4398
$C\mathbb{A}_2(4,1;\{v\})$	0.6234	0.5929	0.6773	0.6414	0.5234	0.4401
$\mathbb{A}_4(4,1;\{v\})$	0.7043	0.4512	**0.3259**	**0.3638**	0.4410	0.4298
$\mathbb{A}_4(4,2;\{\pi,g\})$	0.7450	0.5837	0.5694	0.5892	0.5465	0.5254
$\mathbb{A}_4(4,2;\{v,g\})$	0.7168	0.5188	0.5089	0.5303	0.4138	0.3764
$\mathbb{A}_5(5,2;\{\pi,g\})$	0.9205	0.6852	0.5389	0.5285	0.5621	0.5695
$\mathbb{A}_5(5,2;\{v,g\})$	0.6947	0.4424	0.3825	0.4160	0.4544	0.4249

Table 2-6. Out-of-sample Forecasts Performance Using RMSE (Forecast Horizon: 12 months)

Model\Maturity (months)	1	3	6	12	36	60
$\mathbb{A}_1(3,0)$	0.8184	0.8123	0.9039	0.8590	0.7095	0.7508
$C\mathbb{A}_1(3,0)$	1.2498	1.2204	1.2537	1.1114	0.8055	0.7991
$\mathbb{A}_3(3,0)$	1.0623	0.8269	0.6700	0.6081	0.5470	0.4812
$\mathbb{A}_1(3,1;\{v\})$	1.2215	1.2481	1.3248	1.2137	0.7659	0.5755
$C\mathbb{A}_1(3,1;\{v\})$	1.4867	1.5035	1.5609	1.4175	0.8744	0.5955
$\mathbb{A}_3(3,1;\{v\})$	0.6138	**0.4339**	**0.4346**	**0.4925**	**0.4641**	0.3996
$\mathbb{A}_1(4,1;\{v\})$	1.0169	1.0484	1.1519	1.0971	0.7792	0.6315
$C\mathbb{A}_1(4,1;\{v\})$	1.5392	1.5626	1.6341	1.5170	1.0032	0.6926
$\mathbb{A}_2(4,1;\{v\})$	0.8966	0.9214	1.0407	1.0103	0.7575	0.6197
$C\mathbb{A}_2(4,1;\{v\})$	1.2673	1.2696	1.3373	1.2309	0.8271	0.6140
$\mathbb{A}_4(4,1;\{v\})$	0.6418	0.4891	0.5200	0.5897	0.5748	0.5161
$\mathbb{A}_4(4,2;\{\pi,g\})$	0.5810	0.5504	0.6473	0.6648	0.5712	0.5131
$\mathbb{A}_4(4,2;\{v,g\})$	0.6270	0.6024	0.6886	0.7008	0.4905	**0.3802**
$\mathbb{A}_5(5,2;\{\pi,g\})$	0.7276	0.5526	0.4937	0.5194	0.5418	0.5555
$\mathbb{A}_5(5,2;\{v,g\})$	**0.5627**	0.5483	0.6960	0.7616	0.6326	0.5148

estimated in sample. The spirit of this regression is simple: If a term structure model captures the time-variability of the term premium, γ_1 should not be significantly different from zero. Previous affine models with flexible market price of risk have difficulty in generating this result unless the models assume a conditional homoscedastic volatility. Since my macro models produce better forecasts than latent models despite the stochastic volatility, it is interesting to see how my macro models perform. I report the results in Table 2-7.

As clearly seen, the p-values of the latent factor models ($\mathbb{A}_3(3,0)$, $\mathbb{A}_1(3,0)$) imply that the affine models have trouble matching the term premium variability. Adjusted R^2s also indicate that term spreads still have considerable explanatory power for forecasts errors. Meanwhile, the p-values of the hybrid models with the velocity of money imply that most of the term premium variability is captured by the velocity augmented affine models. This result is robust across different model specifications. In conjunction with the best forecasts performance of those models shown in Table 2-3 $\sim$ Table 2-6, it is argued that the tension between improving yield forecasts and explaining the yield stochastic volatility disappears to a substantial degree. Note that the model with inflation and output gap has similar p-values to those from latent factor models, which implies that the model with inflation and output does not sufficiently capture the term premium variability when stochastic volatility resides.

From a perspective of estimation, matching the first and the second moments is more burdensome than explaining only the first moment. However, as I put in the introduction, stochastic volatility and its clustering behavior is prevalent in the yield data, hence it is necessary to deal with this statistical tension rather than ignore the second moment property. Related, a good term structure model should be able to explain the variety of stylized facts about the Treasury yield data. My finding suggests that observable macro variables, especially those describing the liquidity demand of an

Table 2-7. Regression of the Forecasting Residuals on the Term Spread (Out-of-sample 6-month Ahead Forecasting)

Model\Maturity	6 months	12 months	36 months	60 months
$\mathbb{A}_1(3,0)$				
Adj. R^2	0.0685	0.1381	0.1502	0.1083
γ_1	-0.0993	-0.1556	-0.2062	-0.1781
	(0.0493)	(0.0540)	(0.0778)	(0.0862)
p-value	0.0539	0.0075	0.0131	0.0482
$\mathbb{A}_3(3,0)$				
Adj. R^2	0.7185	0.6035	0.2488	0.0640
γ_1	-0.5907	-0.5077	-0.2646	-0.1351
	(0.0872)	(0.0807)	(0.0711)	(0.0715)
p-value	0.0000	0.0000	0.0009	0.0693
$\mathbb{A}_3(3,1;\{v\})$				
Adj. R^2	0.2369	0.1760	-0.0296	-0.0014
γ_1	-0.1937	-0.1884	-0.0306	0.0714
	(0.0774)	(0.0779)	(0.0721)	(0.0717)
p-value	0.0185	0.0224	0.6748	0.3279
$\mathbb{A}_1(4,1;\{v\})$				
Adj. R^2	0.0427	0.1093	0.0980	0.0434
γ_1	-0.0913	-0.1524	-0.1846	-0.1353
	(0.0551)	(0.0598)	(0.0811)	(0.0861)
p-value	0.1088	0.0166	0.0307	0.1275
$\mathbb{A}_2(4,1;\{v\})$				
Adj. R^2	-0.0153	0.0561	0.1237	0.0998
γ_1	-0.0420	-0.1068	-0.1894	-0.1751
	(0.0511)	(0.0549)	(0.0800)	(0.0871)
p-value	0.4182	0.0617	0.0250	0.0542
$\mathbb{A}_4(4,2;\{v,g\})$				
Adj. R^2	0.1234	0.1420	0.0094	-0.0357
γ_1	-0.1984	-0.2196	-0.0981	0.0003
	(0.1151)	(0.1121)	(0.0933)	(0.0778)
p-value	0.0958	0.0601	0.3019	0.9972
$\mathbb{A}_5(5,2;\{\pi,g\})$				
Adj. R^2	0.0020	0.0463	0.0736	0.0608
γ_1	-0.1088	-0.1671	-0.2088	-0.1904
	(0.0567)	(0.0512)	(0.0659)	(0.0762)
p-value	0.0656	0.0029	0.0037	0.0186

economy are very useful to achieve this goal. Macroeconomic variables are known to capture the variations of business conditions. Since expected excess returns of assets are often negatively correlated with business conditions, the use of macro factors may help capture expected bond return behaviors, or more specifically the market price of risk. Furthermore, I emphasize the role of variables related to liquidity demand such as the velocity of money among many observable macro factors. When properly defined to reflect the correct nature of transactions services, the velocity of money can explain the money demand of an economy which is in turn closely and contemporaneously related to nominal interest rates. Thus, I believe that the velocity of money plays an instrumental role in explaining both the level and the slope features of bond yields, thereby providing the affine models with more room for fitting the conditional variances of yields.

F. Conclusion

I report that incorporating observable macroeconomic variables not only helps the affine term structure models better predict future yields but also considerably reduces the tension between matching the first and the second conditional moments. Especially, the affine models with the velocity of money measured by M2 minus small time deposits can capture all three major stylized facts in Treasury yields. For each maturity, I estimate affine models with different risk price specifications, the number of factors, the number of independent volatility factors, and different combinations of macro factors. Although there is no clear winner that dominates across maturities, I find that three to four factor models with the velocity of money, stochastic volatility, and flexible market price of risk perform better than others in terms of out-of-sample forecasts. These models can also match the term premium variability observed in the data.

My results imply that for the purpose of yield forecasts, macro-latent affine models can provide better results than simple forecasting methods such as random walk or unconstrained vector auto regressions. It is well known that economic restrictions such as the money demand relationship or monetary policy play a key role in understanding interest rates. My results suggest that those conditions are important for the empirical evaluation of the term structure of interest rates as well.

One final point I would mention is that the results of my paper indicate that affine models with stochastic volatility can be useful in studying macro term structure models in general equilibrium frameworks, provided that liquidity demand is properly modeled.

CHAPTER III

ESTIMATING TERM STRUCTURE MODELS WITH MACRO FACTORS USING HIGH FREQUENCY DATA

A. Introduction

The term structure of interest rate reflects the expectations and influences from macroeconomic variables. Hence, the yield curve conveys information about the future economy. Many other recent papers have modeled term structure dynamics including observable macroeconomic variables such as inflation and the output gap. Ang and Piazzesi (2003) propose a macro-finance model incorporating macroeconomic variables as observable state variables other than the typical latent factors. They present a VAR model of the yield curve with inflation and real activity along with latent factors. Ang and Piazzesi (2003) report that macro variables account for a substantial portion of the variation in the short and middle of the yield curve and incorporating macro variables into affine models helps to improve forecasts. Moreover, imposing the cross-equation restriction from no arbitrage improves the performance of out-of-sample forecasting. Dewachter and Lyrio (2006) find that inflation expectations are crucial for long-term bond yields and that both inflation and the real interest rate are especially important for the short-end of the term structure.

Diebold, Rudebusch, and Aruoba (2006) estimate a term structure model with latent factors and observable macro variables (inflation, real activity, and the monetary policy instrument) to examine the interactions between the macroeconomic variables and bond yields. They find evidences of macroeconomic effects on the future yield curve and yield curve effects on the future macro economy as well. However, Diebold, Rudebusch, and Aruoba (2006) do not impose no-arbitrage restrictions, which is unusual in the term structure models, arguing that if the restrictions hold for the data and the yield curve has a good fit, then these restrictions are approximately met. Hördahl, Tristani, and Vestin (2006) construct a joint model of macroeconomic and term structure dynamics. Their model performs very well in forecasting future bond yields while yields do not appear to help to improve the performance in forecasting macroeconomic variables. Duffee (2006) estimates a term structure model without latent factors. He finds that there is a positive relationship between the short-term interest rate and inflation and that short-term interest rates move approximately one-for-one with changes in expected inflation.

However, macroeconomic variables which are used in those papers are measured at monthly or quarterly frequency; thus, these types of macro-finance models cannot match the higher frequency of the interest rates and the lower frequency of macro variables. In this sense, it is hard to examine the role of macro variables in explaining term structure dynamics in continuous time. For example, in the macro-finance model using monthly frequency data, the intra-month information can be missed. To deal with this problem, I propose a term structure model employing daily data. It is known that using high frequency data improves the accuracy of the conditional volatility in the same sample size. Particularly, with a shorter sample, employing daily data is crucial for analyzing term structure dynamics.

Therefore, I construct a tractable model at daily frequency with both the typical latent factors and latent macro factors by imposing cross-equation restrictions on yield movements from no-arbitrage while most macro-finance models incorporate observable macroeconomic variables in monthly or quarterly frequency. Additionally, I add the spread factor between the short-term Treasury yield and the federal funds rate into an affine term structure model to identify the high-frequency monetary policy rule that describes the central bank's reaction to expected inflation and real activity at daily frequency. The benchmark and backward-looking high-frequency policy rules are identified without difficulties. Although many other researchers place Taylor rules incorporating inflation and the output gap in an affine model, those macro variables are observable in monthly or quarterly frequency. When I assess my model in terms of out-of-sample forecasting, my term structure model with macro factors and the spread factor shows better performance.

Finally, I show that the spread between the 3-month Treasury yield and the federal funds rate has strong predictive power for predicting excess bond returns and future changes in yields from the results of two different regressions. Also, I find that short-maturity yields tend to rise and long-maturity yields tend to fall when the yield spreads widen. These results are inconsistent with the expectations hypothesis.

In my model, different bond yields such as the real yield, nominal yield, and defaultable yield are used and latent macro factors and the spread factor are extracted from yield relationships by using cross-equation restrictions. Accordingly, I do not need to worry about the discrepancy of data frequency between yields and macro variables. Thus, I do not lose information available in matching high frequency yields data and low frequency macro variables data. Another benefit is that the no arbitrage assumption is enforced by imposing cross-equation restrictions.

Chapter III is organized as follows. The next section shows data used in this paper. Section C presents affine term structure models estimated. Then, I explain presenting estimation results. Finally, after some further discussions on my results, I conclude.

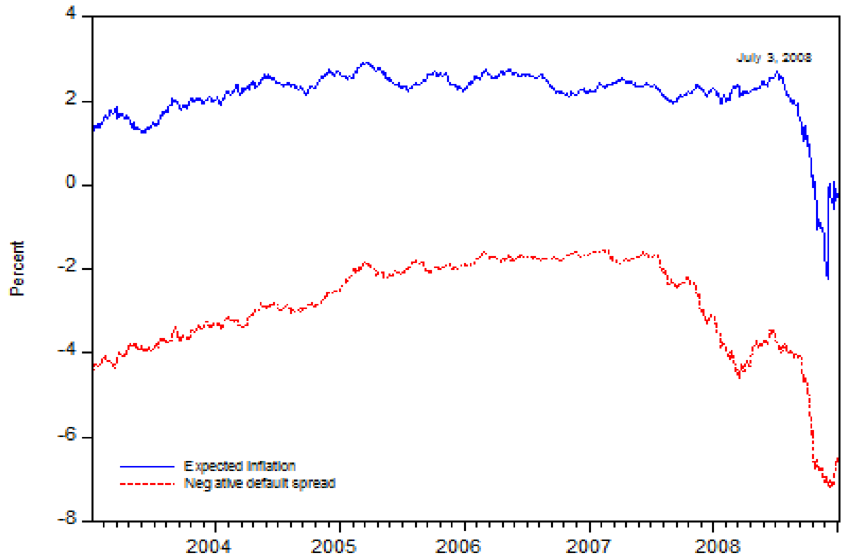

Fig. 3-1. US Daily Macroeconomic Variables

B. Data

As I mentioned, macro variables are not measured at daily frequency. Thus I should find proxies for macro variables on a daily basis. From the Fisher Hypothesis, the difference between the nominal and real yield can be regarded as expected inflation. Regarding real activity or output, if business conditions become worse, investors require a higher expected return for the extra risk; accordingly, the default spread increases. In other words, the default spread is inversely related to business conditions. Therefore, the difference between the nominal and real yield, which represents expected inflation, and the difference between the Treasury yield and corporate bond yield, which represents the negative default spread, are considered to be the macro variables in a term structure model at daily frequency.

Figure 3-1 and Figure 3-2 show macroeconomic variables on a daily basis. Figure 3-1 plots US daily macro variables representing 5-year TIPS-derived expected

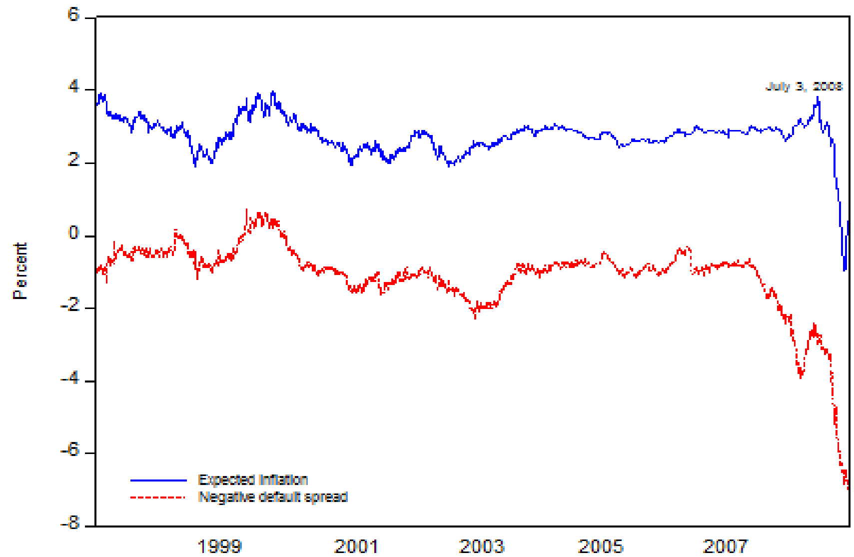

Fig. 3-2. UK Daily Macroeconomic Variables

inflation and negative default spread from January 2003 through December 2008. Figure 3-2 plots UK daily macro variables from March 1997 through December 2008. As clearly seen in Figure 3-1 and Figure 3-2, expected inflation is very stable and low, but expected inflation has drastically decreased since July 2008. The negative default spread has also decreased since July 2008. Thus, it is possible to guess that the economy might fall into a recession. Figure 3-3 plots the inflation measure using $\log(P_t/P_{t-12})$ where P_t is the consumer price index (CPI) and 5-year TIPS-derived expected inflation on a monthly basis in the US. Inflation is more volatile than expected inflation. Figure 3-4 plots the growth rate of industrial production (IP) measured as the log difference at time t and $t-12$ and the negative default spread (5-year Treasury bond yield - corporate bond yield(baa)) on a monthly basis in the US. The shaded areas show periods of recessions as defined by the NBER. As clearly seen, they are moving very similarly and are highly correlated. Thus, the negative default spread could be a good proxy for real activity or output.

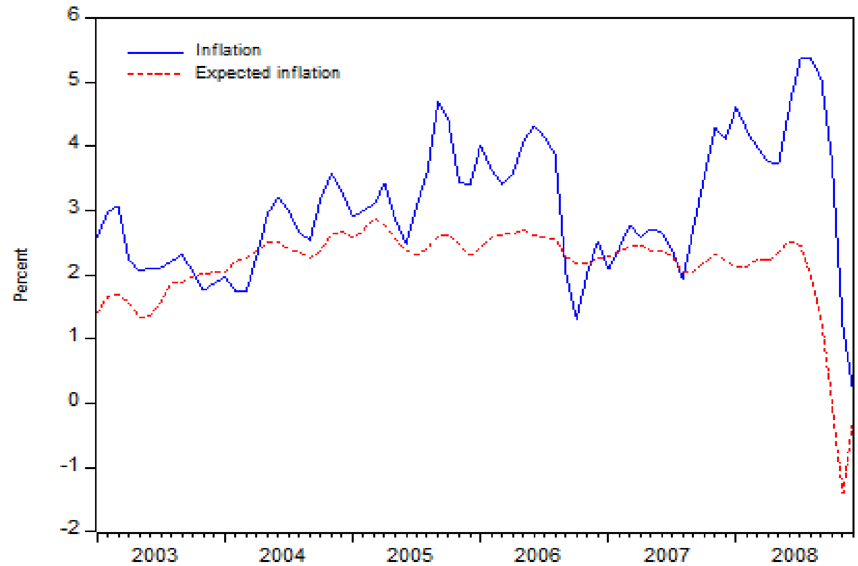

Fig. 3-3. US Inflation and Expected Inflation

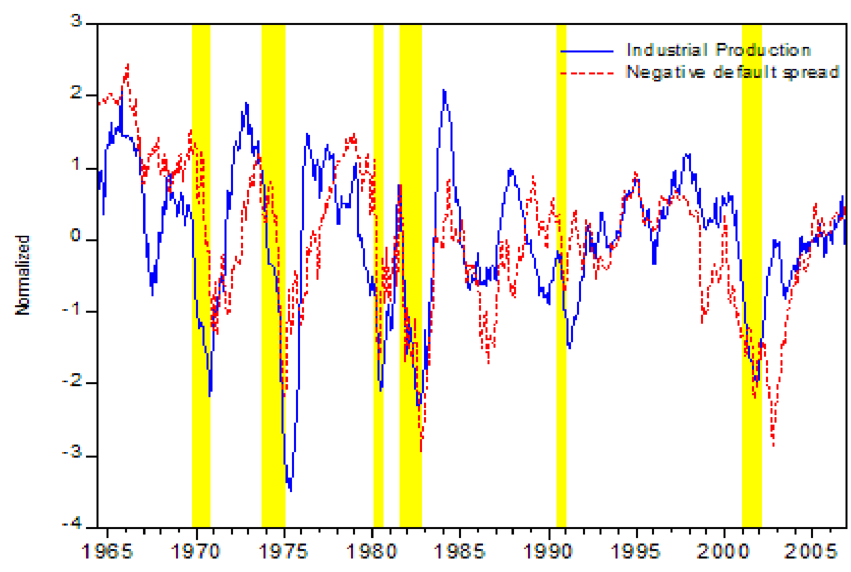

Fig. 3-4. US Real Activity and the Negative Default Spread

Regarding bond yields, I use daily yield series of US Treasury constant maturity bond yields with maturities of 3, 6, 12, 24, 36 and 60 months, from January 2003 through December 2008, taken from the the St. Louis Fed (FRED). In addition, UK government bond yields with the maturities of 3, 6, 12, 24, 36 and 60 months, from March 1997 through December 2008 are used, come from the Bank of England. All bond yields are continuously compounded.

C. Model

1. A Term Structure Model with Macro Factors

Many empirical papers studying term structure dynamics incorporate inflation and real activity as observable macro variables. This is because the central banks set the short-term interest rate by reacting to inflation and the output gap. Taylor (1993) suggests the following form

$$i_t = r^* + \pi_t + \theta_\pi(\pi_t - \pi_t^*) + \theta_g g_t, \tag{3.1}$$

where i_t is the short-term interest rate that the central banks can control, r^* is the equilibrium real rate, π_t^* is the central bank's inflation target, π_t is a measure of inflation, and g_t is a measure of output gap. In this Taylor rule, the short-term interest rate is regarded as the sum of $(r^* + \pi_t)$ and deviations from the policy goals. Other versions of the Taylor rule are forward-looking Taylor rules using expected inflation and expected output gap. Ang, Dong, Piazzesi (2007) place the Taylor rule into a term structure model under no-arbitrage framework. They define the bench mark Taylor rule as follows:

$$i_t = \gamma_0 + \gamma_{1,\pi}\pi_t + \gamma_{1,g} g_t + \varepsilon_t^{MP,T} \tag{3.2}$$

where $\varepsilon_t^{MP,T}$ is the unobserved monetary policy shock which corresponds to a latent term structure factor, f_t^u. Hence, two observable macro variables π_t, g_t, and the only one latent factor f_t^u are incorporated into a term structure model as state variables and various versions of Taylor rules are estimated.

Inspired by Ang, Dong, Piazzesi (2007), I set up a five-factor term structure model with two macro factors besides the typical latent factors using the cross-equation restrictions on yield movements from no arbitrage assumption. Now I know that the nominal interest rate is approximately the sum of the real interest rate and expected inflation, and the defaultable bond yield is the sum of the default-free bond yield and the default spread. Using these relationships, I denote different bond yield as follows:

$$\begin{aligned} y_t^{R,D} &= \text{real defaultable bond yield (Corporate bond yield} - \text{expected inflation)} \\ y_t^{\$,D} &= \text{nominal defaultable bond yield (Corporate bond yield)} \\ y_t^{\$} &= \text{nominal default-free bond yield (Treasury yield)} \end{aligned}$$

Instantaneous short-term interest rates are defined as below and $x_{1t}, x_{2t,}x_{3t}$ are considered as the usual latent factors (level, slope, curvature).

$$r_t^{R,D} = x_{1t} + x_{2t} + x_{3t} \tag{3.3}$$

$$r_t^{\$,D} = x_{1t} + x_{2t} + x_{3t} + \pi_t^e \tag{3.4}$$

$$r_t^{\$} = x_{1t} + x_{2t} + x_{3t} + \pi_t^e + d_t^- \quad (\text{where } d_t^- = -d_t) \tag{3.5}$$

In measurement equations, two more yields $y_t^{R,D}$, $y_t^{\$,D}$ are added to Treasury yields with 6 different maturities in order to extract π_t^e capturing expected inflation and d_t^- capturing the negative default spread. Contrary to other macro-finance models, macro variables in my model are latent factors by cross-equation restrictions on yield movements. The subscript of a matrix displays its dimensions.

$$(3.6)\qquad \begin{pmatrix} y^{R,D}_{t,1\times1} \\ y^{\$,D}_{t,1\times1} \\ y^{\$}_{t,6\times1} \end{pmatrix} = \begin{pmatrix} A^{R,D}_{1\times1} \\ A^{\$,D}_{1\times1} \\ A^{\$}_{6\times1} \end{pmatrix} + \begin{pmatrix} B^{R,D}_{1\times3} & 0 & 0 \\ & B^{\$,D}_{1\times4} & 0 \\ & B^{\$}_{6\times5} & \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ \pi^e_t \\ d^-_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t,8\times1} \end{pmatrix}$$

2. Identification of the High-frequency Monetary Policy Rule

To identify the high-frequency monetary policy rule, I set up a six-factor term structure model. I denote s_t as a latent factor capturing the spread between the short-term Treasury yield and the federal funds rate (repo rate in the UK). Thus, the federal funds rate, $y^{F,F}_t$, is added to the measurement equations in other to extract a latent factor s_t.

$$(3.7)\qquad r^{R,D}_t = x_{1t} + x_{2t} + x_{3t}$$

$$(3.8)\qquad r^{\$,D}_t = x_{1t} + x_{2t} + x_{3t} + \pi^e_t$$

$$(3.9)\qquad r^{F,F}_t = x_{1t} + x_{2t} + x_{3t} + \pi^e_t + d^-_t \quad (\text{where } d^-_t = -d_t)$$

$$(3.10)\qquad r^{\$}_t = x_{1t} + x_{2t} + x_{3t} + \pi^e_t + d^-_t + s_t$$

(3.11)

$$\begin{pmatrix} y^{R,D}_{t,1\times1} \\ y^{\$,D}_{t,1\times1} \\ y^{F,F}_{t,1\times1} \\ y^{\$}_{t,6\times1} \end{pmatrix} = \begin{pmatrix} A^{R,D}_{1\times1} \\ A^{\$,D}_{1\times1} \\ A^{F,F}_{1\times1} \\ A^{\$}_{6\times1} \end{pmatrix} + \begin{pmatrix} B^{R,D}_{1\times3} & 0 & 0 & 0 \\ & B^{\$,D}_{1\times4} & 0 & 0 \\ & & B^{F,F}_{1\times5} & 0 \\ & & & B^{\$}_{6\times6} \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \\ \pi^e_t \\ d^-_t \\ s_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t,9\times1} \end{pmatrix}$$

Identifying the high-frequency policy rule that describes the central bank's reaction to expected inflation and the negative default spread, the parameters are estimated by the Kalman filter procedure in the first step. Therefore, I obtain a measurement equation of the federal funds rate as follows:

$$y_t^{F,F} = A^{F,F} + B_1^{F,F} x_{1t} + B_2^{F,F} x_{2t} + B_3^{F,F} x_{3t} + B_4^{F,F} \pi_t^e + B_5^{F,F} d_t^- + \varepsilon_t \tag{3.12}$$

In the above equation, the parameters $B_4^{F,F}$ and $B_5^{F,F}$ are known, but π_t^e and d_t^- are the latent factors denoting expected inflation and the negative default spread, respectively. Thus, I run OLS regressions to obtain equations to see how unobservable macro factors (π_t^e, d_t^-) respond to changes in observable counterparts (Π_t^e , D_t^-) in the second step. This can be interpreted as the process of scale adjustments to compare my policy rules with other policy rules such as the standard Taylor rule.

(3.13)

$$\pi_t^e = \gamma_{0,\pi} + \gamma_{1,\pi} \Pi_t^e \quad (\ \Pi_t^e = \text{5-year TIPS-derived expected inflation} \ )$$

(3.14)

$$d_t^- = \gamma_{0,d} + \gamma_{1,d} D_t^- \quad (\ D_t^- = \text{5-year Treasury bond yield} \ - \text{corporate bond yield})$$

By plugging the equations in the second step into the measurement equation in the first step, I can obtain a equation representing monetary policy rule on a daily basis. This rule can be interpreted as the central bank's activity adjusting policy rate in response to movements in expected inflation and the negative default spread.

(3.15)

$$y_t^{F,F} = A_{new}^{F,F} + B_1^{F,F} x_{1t} + B_2^{F,F} x_{2t} + B_3^{F,F} x_{3t} + (B_4^{F,F} \times \gamma_{1,\pi}) \Pi_t^e + (B_5^{F,F} \times \gamma_{1,d}) D_t^- + \varepsilon_t^{new}$$

where $A_{new}^{F,F}$ and ε_t^{new} denote rearranged constant term and error term.

If I modify the above equation by including lagged macro variables in the second step, backward-looking policy rules can be identified easily. This could be one of the advantages in my two-step estimation method. The intuition behind a backward-looking policy rule is to smooth the federal funds rate by considering lagged macro variables as well as current ones.

$$\pi_t^e = \gamma_{0,\pi} + \gamma_{1,\pi}\Pi_t^e + \gamma_{2,\pi}\Pi_{t-1}^e \tag{3.16}$$

$$d_t^- = \gamma_{0,d} + \gamma_{1,d}D_t^- + \gamma_{2,d}D_{t-1}^- \tag{3.17}$$

$$\begin{aligned} y_t^{F,F} = {} & A_{new}^{F,F} + B_1^{F,F}x_{1t} + B_2^{F,F}x_{2t} + B_3^{F,F}x_{3t} + (B_4^{F,F} \times \gamma_{1,\pi})\Pi_t^e \\ & + (B_5^{F,F} \times \gamma_{1,d})D_t^- + (B_4^{F,F} \times \gamma_{2,\pi})\Pi_{t-1}^e + (B_5^{F,F} \times \gamma_{2,d})D_{t-1}^- + \varepsilon_t^{new} \end{aligned} \tag{3.18}$$

3. Estimated Models

Regarding the number of latent factors, I estimate the affine models with three typical latent factors which are referred to as three-factor yields-only models. Additionally, the affine models with five latent factors are estimated in comparison with a five-factor term structure model with macro factors. I refer to an affine model with two macro factors as a five-factor macro model. Also, a six-factor macro model is estimated. Another criterion for classifying the estimated models is the number of independent state variables determining the variance-covariance of the state vector X_t. It is well known that there exists a tension between matching conditional mean and conditional volatility of yields in that smaller number of instantaneous volatility drivers enhances the forecast ability of the model. Thus, I estimate three-factor and five-factor yields-only models with only one state variable determining the variance-covariance of the state vector X_t.

For expositional purposes, I label the estimated models using notation similar to Dai and Singleton (2000). $\mathbb{A}_M(n, j)$ refers to an affine model with the total number

n of factors, j macroeconomic factors, M state variables affecting the instantaneous volatility, and the market price of risk specification of Duffee (2002). Specifically I estimate the following cases:

- Three-factor Yields-Only Model: $\mathbb{A}_3(3,0)$, $\mathbb{A}_1(3,0)$
- Five-factor Yields-Only Model: $\mathbb{A}_5(5,0)$, $\mathbb{A}_1(5,0)$
- Five-factor Macro Model: $\mathrm{XR}\mathbb{A}_5(5,2)$
- Six-factor Macro Model: $\mathrm{XR}\mathbb{A}_6(6,2)$

I impose restrictions (mostly zero restrictions) on the parameters of the models above. Some of those come from the representation of the affine model as in Dai and Singleton (2000), Duffee (2002), Ang and Piazzesi (2003), Duarte (2004), and Aït-Sahalia and Kimmel (2008). Moreover, I make additional assumptions on the parameters to make the model parsimonious. A five-factor macro model and a six-factor macro model are regarded as my main models and I compare these with the three-factor yields-only models $\mathbb{A}_M(3,0)$ which are popular choices among the affine models.

This is the conventional model of term structure with the no-arbitrage restriction, $\mathbb{A}_3(3,0)$, $\mathbb{A}_1(3,0)$ depending on the number of independent factors describing instantaneous volatility of yields.

$$
\begin{aligned}
r_t &= \sum_{i=1}^{3} x_{it}, \\
d \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix} &= \begin{pmatrix} \kappa_1(\theta_1 - x_{1t}) \\ \kappa_2(\theta_2 - x_{2t}) \\ \kappa_3(\theta_3 - x_{3t}) \end{pmatrix} dt + \sqrt{S_t} dW_t, \\
[S_t]_i &= \alpha_i + \beta_i x_{it}, \\
\Lambda_t &= \sqrt{S_t} \begin{pmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \end{pmatrix} + Z_t \begin{bmatrix} \lambda_{2(11)} & 0 & 0 \\ 0 & \lambda_{2(22)} & 0 \\ 0 & 0 & \lambda_{2(33)} \end{bmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ x_{3t} \end{pmatrix}.
\end{aligned}
$$

For the $\mathbb{A}_1(3,0)$ model, $\alpha_1 = 0$, $\beta_2 = \beta_3 = 0$, $\theta_2 = \theta_3 = 0$ are additionally imposed. The admissibility condition by Dai and Singleton (2000) dictates that this model is over-identified, especially in terms of the drift component of the X_t process, but I find that relaxing this part does not change the likelihood nor the out-of-sample forecasting performance much. The model specifications of two versions of five-factor yields-only models, $\mathbb{A}_5(5,0)$, $\mathbb{A}_1(5,0)$, are identical to the three-factor yields-only models except that two additional state variables are incorporated. The term structure models are estimated through the Kalman filter approach. Given the first two conditional moments of the state variables under the Milstein approximation, I can use the usual Kalman filter procedure to evaluate the logarithm of the quasi-likelihood function recursively.

D. Empirical Results

1. Parameter Estimates and State Variables

This section reports parameter estimates of several affine models I estimated. Table A-4 $\sim$ Table A-7 display the estimates of the 3-factor yields-only models and 5-factor yields-only models together with standard errors given in parentheses. Most parameters in all cases are estimated reliably. I follow Duffee (2002) for zero restrictions on the market price of risk in case of the essentially affine setups and other restrictions come from Dai and Singleton (2000). The 5-factor macro model and 6-factor macro model are estimated in an identical manner. Table A-8 $\sim$ Table A-9 show the estimation results from the 5-factor macro model. The 5-factor yields-only models have only one relatively persistent factor, the level factor. The 4th and 5th factor have lower persistence. On the other hand, the 5-factor macro model has two relatively persistent factors, one of which is the level factor, the other is the negative default spread representing real activity.

By using the Kalman filter procedure to estimate the term structure models, I can extract the usual latent factors (level, slope, curvature), macro factors, and the spread factor. I compute correlation coefficients between the usual latent factors and corresponding empirical proxies in case of the 6-factor macro model in the US. A long-term yield (y_t^{60m}) is considered as an empirical proxy for level. The correlation between the level factor and the proxy is -0.68, which is caused by including the expected inflation factor, the negative default spread factor, and the spread factor in the 6-factor macro model. The slope factor displays high 0.86 correlation with its empirical counterpart ($y_t^{3m} - y_t^{60m}$) and the curvature factor also shows high 0.91 correlation with its empirical proxy, $(y_t^{3m} + y_t^{60m}) - 2y_t^{24m}$. Therefore, two macro factors and the spread factor affect the movement of the level factor while the slope factor and

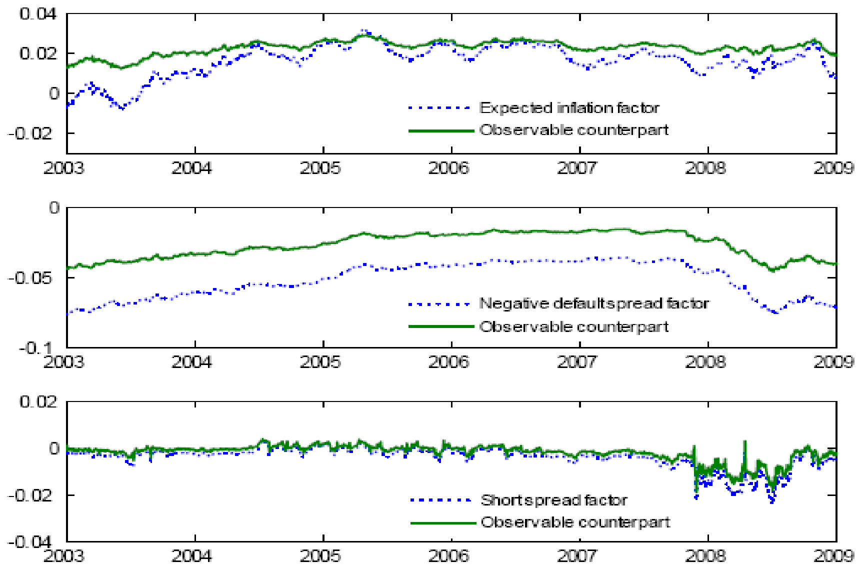

Fig. 3-5. Macro Factors and Observable Counterparts : US Six-factor Macro Model

curvature factor are not affected by those factors. Contrary to other macro-finance models incorporating observable macro variables, I extract latent macro factors by imposing cross-equation restrictions. In Figure 3-5, I plot the latent macro factors and the spread factor between the short-term Treasury yield and the federal funds rate together with their observable counterparts in case of the 6-factor macro model in the US. The factors are moving together with their related observable counterparts and their correlations are 1, 0.99, and 0.90, respectively. Thus, two latent macro factors and the spread factor are perfectly capturing expected inflation, negative default spread, and the spread between the 3-month yield and the federal funds rate.

2. High-frequency Monetary Policy Rule

In the first step, I obtain a measurement equation of the federal funds rate by the Kalman filter procedure in a state-space setup and x_{1t}, x_{2t}, x_{3t} are identified as level, slope, and curvature factor, respectively. The coefficients in a measurement equation are close to 1 in both the US and the UK.

$$FF_t^{US} = 0.0001 + 0.9995x_{1t} + 0.9967x_{2t} + 0.9999x_{3t} + 0.9992\pi_t^e + 0.9997d_t^-$$
$$RP_t^{UK} = 0.0005 + 0.9946x_{1t} + 0.9797x_{2t} + 0.9981x_{3t} + 0.9931\pi_t^e + 0.9992d_t^-$$

Since π_t^e and d_t^- are latent macro factors, I run the regression by ordinary least squares in the second step. From the standard errors in parentheses, most parameters are estimated significantly.

- US

$$\underset{}{\pi_t^e} = \underset{(0.0000)}{-0.0373} + \underset{(0.0001)}{2.3574} \times \Pi_t^e$$

$$d_t^- = \underset{(0.0001)}{-0.0135} + \underset{(0.0045)}{1.4138} \times D_t^-$$

$$\pi_t^e = \underset{(0.0000)}{-0.0373} + \underset{(0.0008)}{2.3291} \times \Pi_t^e + \underset{(0.0008)}{0.0284} \times \Pi_{t-1}^e$$

$$d_t^- = \underset{(0.0001)}{-0.0135} + \underset{(0.1099)}{1.2990} \times D_t^- + \underset{(0.1098)}{0.1146} \times D_{t-1}^-$$

- UK

$$\pi_t^e = \underset{(0.0000)}{-0.0451} + \underset{(0.0001)}{2.1754} \times \Pi_t^e$$

$$d_t^- = \underset{(0.0000)}{-0.0115} + \underset{(0.0034)}{1.1217} \times D_t^-$$

$$\pi_t^e = \underset{(0.0000)}{-0.0451} + \underset{(0.0003)}{2.1529} \times \Pi_t^e + \underset{(0.0003)}{0.0225} \times \Pi_{t-1}^e$$

$$d_t^- = \underset{(0.0000)}{-0.0115} + \underset{(0.0373)}{1.1186} \times D_t^- + \underset{(0.0373)}{0.0024} \times D_{t-1}^-$$

Finally, the benchmark and backward-looking high-frequency policy rules are identified. In the benchmark high-frequency policy rule for the US, the expected inflation coefficient on the variable Π_t^e is 2.3555, which implies that the Fed reacts to expected inflation very aggressively. It is well known that the standard Taylor rule is $1 + 1.5 \times$ inflation $+ 0.5 \times$ output gap. As plotted in Figure 3-3, expected inflation is less volatile than inflation. Thus, the greater value of the expected inflation coefficient relative to the inflation coefficient in the original Taylor rule shows the Fed's relatively more aggressive response to expected inflation. Even though the negative default spread coefficient of 1.4134 is smaller than the expected inflation coefficient, it is greater than 1. Compared to the coefficients in the UK, since the coefficients of both Π_t^e and D_t^- are larger, the Fed appears to adjust the federal funds rate more actively. The backward-looking high-frequency policy rules including lagged macro variables show that the central banks consider the lagged macro information as well as the current macro variables to some extent. However, the sizes of the coefficients of the lagged macro variables are relatively small in both the US and the UK. Until now, I consider the policy rule for the federal funds rate, but I can also get the high-frequency policy rule for the 3-month Treasury yield. Additionally, the spread between the 3-month yield and the federal funds rate (repo rate) is included in the policy rule. I find that the spread between the 3-month yield and the federal funds rate affects the movements of the 3-month yield and that the coefficients of Π_t^e, D_t^-

are slightly smaller than those in the previous policy rules.

- Benchmark high-frequency policy rule for the federal funds rate

$$\begin{aligned} FF_t^{US} &= -0.0508 + 0.9995 \times Level_t + 0.9967 \times Slope_t + 0.9999 \times Curvature_t \\ &+ 2.3555 \times \Pi_t^e + 1.4134 \times D_t^- \\ RP_t^{UK} &= -0.0557 + 0.9946 \times Level_t + 0.9797 \times Slope_t + 0.9981 \times Curvature_t \\ &+ 2.1604 \times \Pi_t^e + 1.1208 \times D_t^- \end{aligned}$$

- Backward-looking high-frequency policy rule for the federal funds rate

$$\begin{aligned} FF_t^{US} &= -0.0508 + 0.9995 \times Level_t + 0.9967 \times Slope_t + 0.9999 \times Curvature_t \\ &+ 2.3272 \times \Pi_t^e + 1.2986 \times D_t^- + 0.0284 \times \Pi_{t-1}^e + 0.1145 \times D_{t-1}^- \\ RP_t^{UK} &= -0.0557 + 0.9946 \times Level_t + 0.9797 \times Slope_t + 0.9981 \times Curvature_t \\ &+ 2.1381 \times \Pi_t^e + 1.1178 \times D_t^- + 0.0224 \times \Pi_{t-1}^e + 0.0024 \times D_{t-1}^- \end{aligned}$$

- Benchmark high-frequency policy rule for the 3-month yield

$$\begin{aligned} 3m_t^{US} &= -0.0489 + 0.9951 \times Level_t + 0.9685 \times Slope_t + 0.8167 \times Curvature_t \\ &+ 2.2409 \times \Pi_t^e + 1.3891 \times D_t^- + 0.6145 \times S_t \\ 3m_t^{UK} &= -0.0532 + 0.9951 \times Level_t + 0.9685 \times Slope_t + 0.8167 \times Curvature_t \\ &+ 2.0801 \times \Pi_t^e + 1.1161 \times D_t^- + 0.5123 \times S_t \end{aligned}$$

- Backward-looking high-frequency policy rule for the 3-month yield

$$
\begin{aligned}
3m_t^{US} &= -0.0489 + 0.9951 \times Level_t + 0.9685 \times Slope_t + 0.8167 \times Curvature_t \\
&+ 2.2140 \times \Pi_t^e + 1.2763 \times D_t^- + 0.5244 \times S_t \\
&+ 0.0270 \times \Pi_{t-1}^e + 0.1126 \times D_{t-1}^- + 0.0939 \times S_{t-1} \\
3m_t^{UK} &= -0.0532 + 0.9951 \times Level_t + 0.9685 \times Slope_t + 0.8167 \times Curvature_t \\
&+ 2.0587 \times \Pi_t^e + 1.1131 \times D_t^- + 0.3409 \times S_t \\
&+ 0.0215 \times \Pi_{t-1}^e + 0.0024 \times D_{t-1}^- + 0.1816 \times S_{t-1}
\end{aligned}
$$

3. Out-of-sample Forecasts Performance

My data sets cover the periods between January 2003 and December 2008 in the US and between March 1997 and December 2008 in the UK. To produce out-of-sample forecasts, I first estimate my models using the in-sample period and forecast future bond yields during the out-of-sample period of September 2008 through December 2008. I compute the root mean squared errors (RMSE) from the out-of-sample forecasts. Lower RMSE values denote better forecasts with the best forecast highlighted in bold. I report the results for bonds with maturities of 3, 6 months, 1 year, 2 years, 3 years, and 5 years. For out-of-sample forecast horizons, 1, 2, and 3 months are used.

In Table 3-1 and Table 3-2, I show a comparison of out-of-sample forecasts for the several affine models. In the case of the US, the 6-factor macro model shows better performance except for maturities of 3 months and 6 months at a 1-month forecast horizon. At a 2-month and a 3-month forecast horizon, the 6-factor macro model shows the lowest RMSE for maturities of 3 months, 6 months, 1 year, 2 years, and 3 years while the 5-factor macro model performs best for maturity of 5 years. As I increase the forecast horizon, the 5-factor macro model and 6-factor macro model perform very well in terms of forecasting future bond yields. Interestingly, the

Table 3-1. Out-of-sample Yields Forecasts Performance Using RMSE (US)

Model \ Maturity	3 m	6 m	12 m	24 m	36 m	60 m
			(Forecast Horizon: 1 month)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	0.7056	0.4628	0.4108	0.4758	0.4851	0.5717
$\mathbb{A}_1(3,0)$	0.6179	**0.3656**	0.3154	0.4799	0.5788	0.5787
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	0.9869	0.8307	0.7447	0.6015	0.5385	0.6481
$\mathbb{A}_1(5,0)$	0.8128	0.6904	0.5819	0.4641	0.4763	0.5913
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	**0.4297**	0.4162	0.3254	0.3036	0.3375	0.3588
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	0.4299	0.3859	**0.2956**	**0.2717**	**0.3146**	**0.3564**
			(Forecast Horizon: 2 months)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	1.1605	0.9220	0.8355	0.8532	0.8424	0.9524
$\mathbb{A}_1(3,0)$	1.0034	0.7527	0.7089	0.8831	0.9653	0.9723
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	1.6650	1.4588	1.2668	1.0550	0.9668	1.0840
$\mathbb{A}_1(5,0)$	1.3494	1.1671	0.9644	0.8320	0.8501	0.9883
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	0.3399	0.3518	0.3528	0.4043	0.4467	**0.4236**
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	**0.3337**	**0.2936**	**0.2802**	**0.3558**	**0.4092**	0.4281
			(Forecast Horizon: 3 months)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	1.5903	1.4590	1.3522	1.2278	1.1642	1.2859
$\mathbb{A}_1(3,0)$	1.4342	1.3217	1.2886	1.3341	1.3588	1.3426
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	2.0759	1.9358	1.7450	1.4551	1.3316	1.4263
$\mathbb{A}_1(5,0)$	1.7517	1.6399	1.4512	1.2611	1.2478	1.3496
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	0.3971	0.3426	0.3161	0.4515	0.5023	**0.2068**
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	**0.3634**	**0.2772**	**0.2251**	**0.3848**	**0.4550**	0.2190

Table 3-2. Out-of-sample Yields Forecasts Performance Using RMSE (UK)

Model \ Maturity	3 m	6 m	12 m	24 m	36 m	60 m
			(Forecast Horizon: 1 month)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	1.0961	1.2355	1.1710	0.8314	0.6129	0.5380
$\mathbb{A}_1(3,0)$	0.9479	1.0996	1.0648	0.7568	0.5463	0.4720
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	1.0876	1.1472	1.0733	0.8582	0.7008	0.5242
$\mathbb{A}_1(5,0)$	0.9160	0.9673	0.9447	0.7619	0.5797	0.4643
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	0.8002	**0.8949**	**0.7920**	**0.3904**	**0.3062**	0.4184
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	**0.7327**	0.9743	0.9920	0.5467	0.3132	**0.3205**
			(Forecast Horizon: 2 months)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	2.2412	2.3619	2.2181	1.6280	1.2194	0.9362
$\mathbb{A}_1(3,0)$	1.9744	2.1259	2.0267	1.4866	1.0991	0.8317
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	2.2128	2.2861	2.1626	1.7038	1.3583	0.9657
$\mathbb{A}_1(5,0)$	1.8489	1.9592	1.9201	1.5044	1.1443	0.8343
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	**1.3058**	**1.4426**	**1.2982**	**0.6615**	**0.3511**	0.5143
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	1.3147	1.5494	1.4836	0.8105	0.3783	**0.3737**
			(Forecast Horizon: 3 months)			
3-factor Yields-Only						
$\mathbb{A}_3(3,0)$	3.1477	3.2574	3.0328	2.2051	1.6416	1.2015
$\mathbb{A}_1(3,0)$	2.8161	2.9625	2.7938	2.0329	1.4998	1.0669
5-factor Yields-Only						
$\mathbb{A}_5(5,0)$	3.2377	3.3173	3.0966	2.3495	1.8284	1.2792
$\mathbb{A}_1(5,0)$	2.6915	2.8332	2.7143	2.0530	1.5455	1.0747
5-factor Macro						
$XR\mathbb{A}_5(5,2)$	**1.4086**	**1.5788**	**1.4215**	**0.6386**	**0.2414**	0.2256
6-factor Macro						
$XR\mathbb{A}_6(6,2)$	1.4913	1.7026	1.5698	0.7410	0.2537	**0.2247**

5-factor yields-only models do not show better performance than the 3-factor yields-only models. This result implies that the 4th and 5th latent factors do not play a role of improving the forecasting performance. It is well known that the three factors account for almost all of the variation of yields. Ang, Piazzesi, and Wei (2006) find that higher factors other than the two factors (level, slope) account for less than 0.3% of the movements of yields at quarterly frequency. Moreover, studies like Dai and Singleton (2000), Duffee (2002), and Duarte (2004) report that there exists a trade-off between improving forecast ability on future bond yields and matching interest rate volatility in affine models. Accordingly, the 5-factor yields-only model with only one stochastic volatility, $\mathbb{A}_1(5,0)$ displays better forecasting performance than $\mathbb{A}_5(5,0)$. In the case of the UK, the macro models show better performance, which means that including macro factors by imposing the cross-equation restriction helps to improve the forecasting performance. However, the 5-factor macro model dominates the 6-factor macro model at almost all maturities, indicating that the yield spread between the 3-month yield and the repo rate is not helpful in terms of forecasting future bond yields.

Now consider the results of macro variables forecasting performance. Due to the cross-equation restrictions in the 5-factor macro model, I can easily obtain the forecast values of macro variables. In other words, the cross-equation restrictions from the no-arbitrage assumption allow the 5-factor macro model to endogenously generate the forecasts of expected inflation and the negative default spread. Hence, I can get the future forecast values of both the yields and macro variables at the same time. For example, the forecast values of the negative default spread can be derived from subtracting the forecast values of corporate bond yields from 5-year Treasury yields. My benchmark is the RMSE from the AR(1) model that is commonly used in forecasting macro variables. In addition, VAR(1) and ECM(1) models are

Table 3-3. Out-of-sample Macro Variables Forecasts Performance Using RMSE

	Expected inflation			Negative default spread		
	Forecast Horizon(month)			Forecast Horizon(month)		
	1	2	3	1	2	3
US						
AR(1)	1.6504	2.3073	2.0832	**1.1098**	**1.9204**	2.5157
VAR(1)	1.6512	2.3100	2.0929	1.1937	2.1062	2.5753
ECM(1)	1.5028	2.1779	**1.7661**	1.1348	2.0428	**2.3153**
5-factor Macro	**1.3132**	**2.0629**	2.0454	1.1713	2.0582	2.7389
UK						
AR(1)	1.3122	**2.2737**	3.0468	1.0605	2.0232	2.8879
VAR(1)	1.2992	2.2754	**3.0467**	1.0442	1.9984	**2.8448**
ECM(1)	1.3069	2.2844	3.1300	**1.0127**	**1.9490**	2.8545
5-factor Macro	**1.2250**	2.2935	3.1315	1.1408	2.1146	2.9686

used for comparison. Regarding expected inflation, the 5-factor macro model is the best performing model at shorter horizon in both the US and the UK. In contrast, with respect to the negative default spread, other benchmark models show better forecasting performance compared to the 5-factor macro model even though there is not a single dominating model at all forecast horizons. My results show that the forecast values of macro variables derived from the 5-factor macro model are comparable to those of other benchmark models.

4. Yield Spreads and the Expectations Hypothesis

The expectations hypothesis tells us that excess bond returns should not be predictable because long-maturity yield is an average of future short rates. However, Fama and Bliss (1987) show that the spread between forward rates and short-maturity yield predicts excess bond returns. Cochrane and Piazzesi (2005) find a single linear combination of forward rate predicts excess bond returns. Thus, I examine time varying excess bond returns via different yield spreads. The data are daily from January 1982 to August 2008 in the US and from March 1997 to August 2008 in the UK. The following regression model is estimated:

$$ehpr_{t,t+k} = \gamma_0 + \gamma_1 \times yield\ spread_t + \varepsilon_{t,t+k}, \qquad (k = 3,\ 6,\ 12,\ \text{and } 24 \text{ months}) \tag{3.19}$$

I denote $ehpr_{t,t+k}$ as the k-period excess bond return and all excess holding period returns are annualized. To the best of my knowledge, asset returns are volatile at high frequencies. This volatility may make the coefficients of the standard OLS regression meaningless. Thus, in addition to the classical OLS, I use the OLS with a time change in sampling to accommodate a stochastic volatility in excess bond returns. Table B-1 $\sim$ Table B-8 in Appendix B show the results from regressions of excess bond returns on different yield spreads. OLS-TC denotes the OLS with a time change method. The regression results show that the OLS with a time change makes the size of the coefficient smaller. This is because volatility is taken care of by using samples collected at random intervals with a time change method. In the US, the coefficients on the yield spread between the 3-month yield and the federal funds rate are significantly negative in both the 3-month and 6-month holding periods. In addition, the coefficients on the spreads between intermediate-maturity yield and short-maturity yield such as $y_t^{24m} - y_t^{3m}, y_t^{12m} - y_t^{3m}$ are significantly positive. On the other hand, the other spreads are statistically insignificant. Therefore, the spreads in the short and middle of the yield curve have predictive power for excess bond returns. In the UK, the some coefficients on the yield spread ($y_t^{60m} - y_t^{36m}, y_t^{60m} - y_t^{24m}, y_t^{24m} - y_t^{12m}$) are significantly negative, which implies that the spreads between long-maturity yield and intermediate-maturity yield have information about future excess bond returns.

I run additional OLS regressions that predict future changes in yields with a variety of yield spreads of the term structure. The regressions are

$$(3.20)\ \ y^{\tau}_{t+k} - y^{\tau}_{t} = \beta_0 + \beta_1 \times yield\ spread_t + \varepsilon_{t,t+k},\ (k = 1 \text{ month}, 2 \text{ months},\ 3 \text{ months})$$

The yield change after k-periods is regressed on different yield spreads at t. Campbell and Shiller (1991) find that a high yield spread between long-maturity yield and short-maturity yield implies that shorter-term yield rises over the long term and longer-term yield falls over the short term. These results are not consistent with the expectation hypothesis of the term structure. Table C-1 $\sim$ Table C-16 in Appendix C show the results from regressions of future changes in yields on the yield spreads using daily data from January 1982 to August 2008 in the US and from March 1997 to August 2008 in the UK. In the US, if the future changes in yields are regressed on the yield spreads such as $y^{24m}_t - y^{3m}_t$, $y^{12m}_t - y^{3m}_t$, the results tell that short-maturity yields (federal funds rate and 3-month Treasury yield) rise and longer-maturity yields(24-month and 60-month Treasury yields) tend to decline when the spreads are greater. These results can be explained by time-varying risk premia, which is equivalent to the expectations hypothesis puzzle. In the UK, when the yield spreads ($y^{60m}_t - y^{36m}_t, y^{24m}_t - y^{3m}_t$, $y^{12m}_t - y^{3m}_t$) widen, short-maturity yields (repo rate, 3-month government yield) rise and also long-maturity yield rises. These imply the weak violation of the expectations hypothesis of interest rates in the UK.

Interestingly, the coefficients of the yield spread between the 3-month yield and the federal funds rate (repo rate in the UK) are significantly positive at all maturities in the US and at short-term maturity yields in the UK. It is known that changes in monetary policy behavior induce changes in the federal funds rate because the federal funds rate is directly controlled by the Fed. Consequently, the Treasury bill rate moves through the monetary policy transmission mechanism. In turn, the short-term Treasury bill rate may reflect the future federal funds rate. In other words, movements in the Treasury bill rate could affect the movement in the federal funds rate. Regarding future changes in the federal funds rate, the coefficients on the yield spread, $y_t^{3m} - ff_t$, are significantly positive at the one percent level. Accordingly, the federal funds rate tends to rise when the spread between the 3-month Treasury yield and the federal funds rate widens. Even though the coefficients increase with longer k-period, they still remain less than one. In terms of future changes in the 3-month Treasury yield, the coefficients are significantly greater than zero, but less than those in case of the federal funds rate. Regarding the future changes of the repo rate in the UK, the coefficients of the 3-month yield-repo rate spread are significantly greater than one at the one percent level in the case of the two-month period and the three-month period. Considering future changes in the 3-month government yield, the coefficients are significantly positive, but less than those in case of the repo rate.

E. Performance of the Milstein Approximation

I compare the performance between conditional mean and variance under the Milstein approximation and the closed form solutions for conditional moments in case of the five-factor macro models.

Furthermore, I show the results of a Monte Carlo simulation to assess the performance of the estimator by the Kalman filter under the Milstein scheme. I generate 1,000 data series of 2,872 daily observations ($\Delta = 1/252$) of the state variables following the square root process with the Milstein approximation. This sample size is used because it has the same length as our sample period of UK daily data.

Table 3-5 displays the results of 1,000 Monte Carlo simulations for a five-factor macro model (XR$\mathbb{A}_5$(5,2)), comparing the distribution of the estimator $\hat{\Phi}^{MIL}$ through the Kalman filter procedure using the approximated conditional moments under the

Table 3-4. Out-of-sample Yields Forecasts Performance of Milstein using RMSE

Method \ Maturity	3 m	6 m	12 m	24 m	36 m	60 m
			(Horizon : 1 month)			
US						
Closed	0.4298	0.4163	**0.3252**	0.3252	**0.3369**	**0.3587**
Milstein	**0.4297**	**0.4162**	0.3254	**0.3036**	0.3375	0.3588
UK						
Closed	0.8006	0.8954	0.7925	0.3908	**0.3061**	0.4184
Milstein	**0.8002**	**0.8949**	**0.7920**	**0.3904**	0.3062	**0.4184**
			(Horizon : 2 months)			
US						
Closed	**0.3399**	**0.3518**	**0.3522**	**0.4032**	**0.4455**	**0.4234**
Milstein	0.3399	0.3518	0.3528	0.4043	0.4467	0.4236
UK						
Closed	1.3070	1.4438	1.2994	0.6626	0.3516	0.5146
Milstein	**1.3058**	**1.4426**	**1.2982**	**0.6615**	**0.3511**	**0.5143**
			(Horizon : 3 months)			
US						
Closed	0.3975	0.3427	**0.3155**	**0.4500**	**0.5006**	**0.2067**
Milstein	**0.3971**	**0.3426**	0.3161	0.4515	0.5023	0.2068
UK						
Closed	1.4107	1.5810	1.4237	0.6406	0.2419	**0.2255**
Milstein	**1.4086**	**1.5788**	**1.4215**	**0.6386**	**0.2414**	0.2256

Milstein scheme around the true value, to the distribution of the estimator $\hat{\Phi}^{CLO}$ through the Kalman filter procedure using the closed form of conditional moments around the true value. From the Table 3-5, I see that the distribution of the $\hat{\Phi}^{MIL}$ is close to that of the $\hat{\Phi}^{CLO}$. Neither estimator seems to dominate the other; that is, the bias and standard deviation of the estimator $\hat{\Phi}^{CLO}$ are larger than those of the estimator $\hat{\Phi}^{MIL}$ for some parameters, but smaller for others. Accordingly, the approximated conditional moments under the Milstein scheme can be used when I have trouble in calculating the closed form of conditional moments. From Table 3-6, MSE (Mean Squared Error) comparison tells that no one dominates the other.

Moreover, I compute the IQR (Inter Quantile Range) which is defined as $|q_{75} - q_{25}|$ where q_i is the i-th quantile of the empirical distribution. IQR is the comparison criteria when the outlier problem exits. As shown in Table 3-6, neither method dominates the other, thus it is very difficult to tell which method performs better. Figure 3-6 displays empirical distribution of the approximation error $(\hat{\Phi}^{CLO} - \hat{\Phi}^{MIL})$ in a XR$\mathbb{A}_5$(5,2) model. The mean of the distribution is very small compared to the mean difference between $(\hat{\Phi}^{CLO} - \hat{\Phi}^{TRUE})$.

Table 3-5. Monte Carlo Simulations for a Five-factor Macro Model (Distribution)

Parameter	Φ^{TRUE}	$\hat{\Phi}^{CLO} - \Phi^{TRUE}$		$\hat{\Phi}^{MIL} - \Phi^{TRUE}$	
		Mean	Std. Dev.	Mean	Std. Dev.
θ_1	0.0030	**-0.0025**	0.0040	-0.0032	0.0040
θ_2	0.0250	-0.0304	0.0236	**-0.0281**	0.0199
θ_3	0.0100	**-0.0095**	0.0082	-0.0099	0.0140
θ_4	0.0300	**-0.0290**	0.0060	-0.0294	0.0054
θ_5	0.0100	**-0.0098**	0.0017	-0.0100	0.0020
κ_1	0.1000	**0.0477**	0.0269	0.0483	0.0323
κ_2	0.7500	-0.0145	0.0170	**-0.0127**	0.0156
κ_3	1.1500	-0.0073	0.0088	**-3.457e-05**	0.0062
κ_4	1.3200	0.0147	0.0122	**0.0085**	0.0100
κ_5	0.1000	**0.0189**	0.0190	0.0248	0.0171
α_1	0.0030	-0.0148	0.0166	**-0.0135**	0.0132
α_2	0.0020	0.0100	0.0199	**0.0099**	0.0136
α_3	0.0020	**-0.0038**	0.0213	-0.0049	0.0186
α_4	0.0040	0.0512	0.0152	**0.0510**	0.0129
α_5	0.0020	0.0090	0.0072	**0.0082**	0.0091
β_1	0.0040	**0.0005**	0.0323	-0.0111	0.0242
β_2	0.0050	-0.0571	0.0377	**-0.0414**	0.0330
β_3	0.0050	0.0016	0.0023	**0.0005**	0.0011
β_4	0.0050	0.0027	0.0030	**0.0011**	0.0020
β_5	0.0030	0.0020	0.0182	**0.0015**	0.0147

Table 3-6. Monte Carlo Simulations for a Five-factor Macro Model (MSE, IQR)

Parameter	$\hat{\Phi}^{CLO}$		$\hat{\Phi}^{MIL}$	
	MSE	IQR_{50}	MSE	IQR_{50}
θ_1	**2.250e-05**	**0.0006**	2.615e-05	0.0008
θ_2	0.0015	0.0303	**0.0012**	**0.0229**
θ_3	**0.0002**	0.0015	0.0003	**0.0013**
θ_4	**0.0009**	0.0028	0.0009	**0.0021**
θ_5	**0.0001**	**0.0002**	0.0001	0.0002
κ_1	**0.0030**	**0.0423**	0.0034	0.0439
κ_2	0.0005	**0.0240**	**0.0004**	0.0242
κ_3	0.0001	0.0109	**3.896E-05**	**0.0067**
κ_4	0.0004	0.0177	**0.0002**	**0.0146**
κ_5	**0.0007**	0.0198	0.0009	**0.0198**
α_1	0.0005	0.0222	**0.0004**	**0.0158**
α_2	0.0005	0.0275	**0.0003**	**0.0176**
α_3	0.0005	0.0264	**0.0004**	**0.0209**
α_4	0.0029	0.0156	**0.0028**	**0.0148**
α_5	**0.0001**	**0.0081**	0.0002	0.0098
β_1	0.0010	0.0393	**0.0007**	**0.0367**
β_2	0.0047	0.0726	**0.0028**	**0.0560**
β_3	7.782E-06	0.0026	**1.422e-06**	**0.0010**
β_4	1.611E-05	0.0037	**5.489e-06**	**0.0021**
β_5	0.0003	0.0215	**0.0002**	**0.0174**

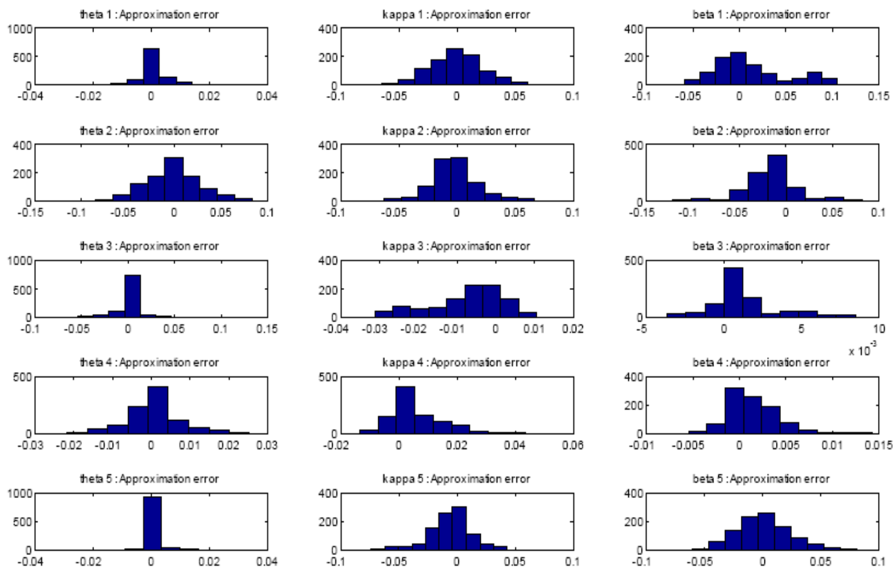

Fig. 3-6. Empirical Distribution of the Approximation Error ($\hat{\Phi}^{CLO} - \Phi^{MIL}$) in a Five–factor Macro Model

F. Conclusion

I propose an affine term structure model with both the typical latent factors and latent macro factors by imposing cross-equation restrictions on yield movements from no-arbitrage. Usually, macroeconomic variables are measured in monthly or quarterly frequency; thus, it is hard to match the higher frequency of the interest rates and the lower frequency of macro variables. To deal with this, instead of employing observable macro variables, more yields such as the real yield, nominal yield, and defaultable yield are used and latent macro factors are extracted from yield relationships by using cross-equation restrictions.

Additionally, I add the spread factor between the short-term Treasury yield and the federal funds rate into an affine term structure model to identify the high-frequency monetary policy rule that describes the central bank's reaction to expected inflation and real activity at daily frequency. By using this two-step method, the benchmark and backward-looking high-frequency monetary policy rules are identified easily. In these monetary policy rules, the sizes of the expected inflation coefficients are larger than 2 and the coefficients of the negative default spread are greater than 1, which implies that the central banks react to expected inflation very aggressively. In comparison with the Bank of England, the Fed appears to adjust the federal funds rate more actively. From the backward-looking high-frequency policy rules, I know that the central banks consider the lagged macro information as well as the current macro variables to some extent.

From the forecasting perspective, I find that macro factors and the spread factor can help the affine term structure models better predict future yields than usual latent models even though the spread factor is not helpful for improving out-of-sample forecasting performance in the UK. These results imply that for the purpose of yield forecasts, a term structure model with macro factors can provide better forecasting results. Moreover, my macro model can generate forecasts of future expected inflation and the negative default spread which are comparable to those by other benchmark models.

Finally, I show that the spread between the 3-month Treasury yield and the federal funds rate has strong predictive power for excess bond returns and future changes in yields. In addition, I find that short-maturity yields tend to rise and long-maturity yields tend to fall when the yield spreads are greater. These results are inconsistent with the expectations hypothesis.

CHAPTER IV

CONCLUSION

I show that incorporating observable macroeconomic variables not only helps the affine term structure models better predict future yields but also considerably reduces the tension between matching the first and the second conditional moments. Especially, the affine models with the velocity of money measured by M2 minus small time deposits can capture all major stylized facts in Treasury yields. For each maturity, I estimate affine models with different risk price specifications, the number of factors, the number of independent volatility factors, and different combinations of macro factors. Although there is no clear winner that dominates across maturities, I find that three to four factor models with the velocity of money, stochastic volatility, and flexible market price of risk perform better than others in terms of out-of-sample forecasts. These models can also match the term premium variability observed in the data.

The results imply that for the purpose of yield forecasts, macro-latent affine models can provide better results than simple forecasting methods such as random walk or unconstrained vector auto regressions. It is well known that economic restrictions such as the money demand relationship or monetary policy play a key role in understanding interest rates. My results suggest that those conditions are important for the empirical evaluation of the term structure of interest rates as well.

Usually, macroeconomic variables are measured in monthly or quarterly frequency; thus, it is hard to match the higher frequency of the interest rates and the lower frequency of macro variables. To deal with this, I propose an affine terms structure model with both the typical latent factors and latent macro factors at daily frequency by imposing cross-equation restrictions on yield movements from no-arbitrage. Additionally, I add the spread factor between the short-term Treasury yield and the federal funds rate into an affine term structure model to identify the high-frequency monetary policy rule. In the high-frequency monetary policy rules, the sizes of the expected inflation coefficients are larger than 2 and the coefficients of the negative default spread are greater than 1, which implies that the central banks react to expected inflation very aggressively. From the backward-looking high-frequency policy rules, I know that the central banks consider the lagged macro information as well as the current macro variables to some extent.

From the forecasting perspective, I find that macro factors and the spread factor can help the affine term structure models better predict future yields than the usual latent models even though the spread factor is not helpful for improving out-of-sample forecasting performance in the UK. I show that the spread between the 3-month Treasury yield and the federal funds rate has strong predictive power for excess bond returns and future changes in yields. In addition, I find that short-maturity yields tend to rise and long-maturity yields tend to fall when the yield spreads are greater. These results are inconsistent with the expectations hypothesis.

REFERENCES

AÏT-SAHALIA, Y., AND R. KIMMEL (2008): "Estimating Affine Multifactor Term Structure Models Using Closed-Form Likelihood Expansions," Working Paper, Department of Economics, Princeton University.

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 (2006): "What Does the Yield Curve Tell Us about GDP Growth," *Journal of Econometrics*, 131, 359–403.

ANG, A., S. DONG, AND M. PIAZZESI (2007): "No-Arbitrage Taylor Rules," Workng Paper, Graduate School of Business, University of Chicago.

CAMPBELL, J., AND R. SHILLER (1991): "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

CHEN, R., AND L. SCOTT (1993): "Maximum Likelihood Estimation for a Multi–factor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14–31.

COCHRANE, J., AND M. PIAZZESI (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

COX, J., J. INGERSOLL, AND S. ROSS (1985): "A Theory of the Nominal Term Structure of Interest Rates," *Econometrica*, 53, 363–384.

Dai, Q., and K. Singleton (2000): "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943–1978.

Dewachter, H., and M. Lyrio (2006): "Macro Factors and the Term Structure of Interest Rates," *Journal of Money, Credit and Banking*, 38, 119–140.

Diebold, F., G. Rudebusch, and S. Aruoba (2006): "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," *Journal of Econometrics*, 131, 309–338.

Duan, J., and J. Simonato (1999): "Estimating and Testing Exponential–Affine Term Sturucture Models by Kalman Filter," *Review of Quantitative Finance and Accounting*, 13, 111–135.

Duarte, J. (2004): "Evaluating an Alternative Risk Preference in Affine Term Structure Models," *Review of Financial Studies*, 17, 379–404.

Duffee, G. (2002): "Term Premia and Interest Rates Forecasts in Affine Models," *Journal of Finance*, 57, 405–443.

Duffee, G. (2006): "Term Structure Estimation without Using Latent Factors," *Journal of Financial Economics*, 79, 507–536.

Duffee, G., and R. Stanton (2004): "Estimation of Dynamic Term Structure Models," Working Paper, Hass School of Business, University of California at Berkeley.

Duffie, D., and R. Kan (1996): "A Yield–Factor Model of Interest Rates," *Mathematical Finance*, 6(4), 379–406.

Fama, E., and R. Bliss (1987): "The Information in Long–Maturity Forward Rates," *American Economic Review*, 77(4), 680–692.

Fisher, M., and C. Gilles (1996): "Estimating Exponential-Affine Models of the Term Structure," Working Paper, Board of Governors of the Federal Reserve System, Washington, DC.

Hördahl, P., O. Tristani, and D. Vestin (2006): "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," *Journal of Econometrics*, 131, 405–444.

Litterman, R., and J. Scheinkman (1991): "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, 1, 51–61.

Lund, J. (1997): "Non-linear Kalman Filtering Techniques for Term Structure Models," Working paper, Aarhus School of Business, University of Aarhus.

Pennacchi, G. (1991): "Identifying the Dynamics of Real Interest Rates and Inflation: Evidence Using Survey Data," *Review of Financial Studies* 4, 53–86.

Taylor, J. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

Vasicek, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

APPENDIX A

PARAMETER ESTIMATES

Table A-1. Parameter Estimates for Selected Three Factor Models

parameters	$\mathbb{A}_1(3,0)$	$\mathbb{A}_3(3,0)$	$\mathbb{A}_1(3,1;\{v\})$	$\mathbb{A}_3(3,1;\{v\})$
θ_1	0.0057 (0.0003)	0.0066 (0.0005)	0	0.0003 (0.0001)
θ_2	0	0.0321 (0.0058)	0	0.0322 (0.0028)
θ_3 (θ_v)	0	0.0172 (0.0014)	0.0296 (0.0006)	0.0311 (0.0007)
κ_1	0.0915 (0.0039)	0.0943 (0.0059)	0.0934 (0.0019)	0.0991 (0.0166)
κ_2	1.2741 (0.0575)	1.9312 (0.2832)	1.1022 (0.0507)	1.2101 (0.0639)
κ_3 (κ_v)	1.8624 (0.0725)	1.8219 (0.0314)	0.0869 (0.0018)	0.0874 (0.0019)
λ_{11}	-0.0827 (0.0037)	-0.2611 (0.1577)	0	-0.1321 (0.1464)
λ_{12}	0	-0.7652 (0.0142)	0	-0.7658 (0.1172)
λ_{13} (λ_{1v})	0	-0.2010 (0.0033)	-0.0998 (0.0049)	-0.0983 (0.0352)
$\lambda_{2(11)}$	0	-0.0182 (0.0062)	0.0696 (0.0053)	-0.0255 (0.0336)
$\lambda_{2(22)}$	-0.1372 (0.0539)	-0.5103 (0.2504)	-0.1585 (0.0464)	-0.0182 (0.0633)
$\lambda_{2(33)}$ ($\lambda_{2(vv)}$)	-0.1307 (0.0742)	-0.1678 (0.0805)	0	-0.0487 (0.0286)
α_1	0	0.0009 (0.0001)	0.0007 (0.0001)	0.0004 (0.0002)
α_2	0.0014 (0.0004)	0.0092 (0.0033)	0.0012 (0.0002)	0.0063 (0.0035)
α_3 (α_{3v})	0.0013 (0.0005)	0.0042 (0.0014)	0	0.0004 (0.0004)
β_1	0.0122 (0.0003)	0.0094 (0.0009)	0	0.0119 (0.0036)
β_2	0	0.0372 (0.0005)	0	0.0142 (0.0088)
β_3 (β_{3v})	0	0.0333 (0.0005)	0.0097 (0.0010)	0.0102 (0.0060)

Table A-2. Parameter Estimates for Selected Four Factor Models

$\mathbb{A}_2(4,1;\{v\})$ parameters	latent 1	latent 2	latent 3	velocity
θ	0.0057 (0.0004)	0	0	0.0296 (0.0001)
κ	0.0903 (0.0054)	1.2318 (0.0551)	1.8624 (0.0277)	0.0869 (0.0001)
λ_1	-0.0813 (0.0041)	0	0	-0.0998 (0.0045)
λ_2	0	-0.1391 (0.0544)	-0.1406 (0.0895)	0
α	0	0.0014 (0.0005)	0.0011 (0.0005)	0
β	0.0117 (0.0006)	0	0	0.0097 (0.0004)
$\mathbb{A}_4(4,1;\{v\})$ parameters	latent 1	latent 2	latent 3	velocity
θ	0.0063 (0.0004)	0.0319 (0.0042)	0.0104 (0.0012)	0.0311 (0.0008)
κ	0.0977 (0.0064)	1.2254 (0.0635)	1.7953 (0.0146)	0.0874 (0.0023)
λ_1	-0.2985 (0.2009)	-0.2103 (0.0676)	-0.7658 (0.0047)	-0.0983 (0.0690)
λ_2	-0.0901 (0.0009)	-0.3113 (0.0365)	-0.3018 (0.0211)	-0.0487 (0.0148)
α	0.0007 (0.0001)	0.0061 (0.0034)	0.0093 (0.0011)	0.0007 (0.0002)
β	0.0121 (0.0036)	0.0147 (0.0088)	0.0187 (0.0010)	0.0102 (0.0028)
$\mathbb{A}_4(4,2;\{v,g\})$ parameters	latent 1	latent 2	velocity	output gap
θ	0.0003 (0.0000)	0.0316 (0.0024)	0.0311 (0.0005)	0.0110 (0.0010)
κ	0.0997 (0.0019)	1.2553 (0.0424)	0.0874 (0.0016)	1.3750 (0.1235)
λ_1	-0.1001 (0.1471)	-0.2510 (0.0241)	-0.0983 (0.0225)	-0.1203 (0.0692)
λ_2	-0.0253 (0.0121)	-0.0010 (0.0416)	-0.0487 (0.0142)	-0.0057 (0.1220)
α	0.0004 (0.0002)	0.0059 (0.0028)	0.0004 (0.0001)	0.0096 (0.0018)
β	0.0122 (0.0055)	0.0146 (0.0128)	0.0102 (0.0022)	0.0192 (0.0213)

Table A-3. Parameter Estimates for a Five Factor Model

parameters	latent 1	latent 2	latent 3	inflation	output gap
θ	0.0085	0.0372	0.0102	0.0352	0.0111
	(0.0006)	(0.0069)	(0.0006)	(0.0013)	(0.0003)
κ	0.0765	1.3569	1.8979	0.0982	1.3750
	(0.0049)	(0.0647)	(0.0131)	(0.0036)	(0.0401)
λ_1	-0.3711	-0.4176	-0.6208	-0.1078	-0.1203
	(0.2776)	(0.0149)	(0.0081)	(0.0942)	(0.0347)
λ_2	-0.0931	-0.2109	-0.3241	-0.0077	-0.0057
	(0.0207)	(0.0705)	(0.0261)	(0.0289)	(0.0497)
α	0.0008	0.0096	0.0012	0.0005	0.0096
	(0.0002)	(0.0036)	(0.0003)	(0.0004)	(0.0035)
β	0.0111	0.0182	0.0132	0.0107	0.0192
	(0.0105)	(0.0010)	(0.0002)	(0.0053)	(0.0357)

Table A-4. Parameter Estimates for Three-factor Yields-only Models (US)

$\mathbb{A}_3(3,0)$ parameters	latent 1	latent 2	latent 3
θ	0.0037 (0.0002)	0.0298 (0.0011)	0.0182 (0.0009)
κ	0.1044 (0.0040)	0.8329 (0.0209)	1.4019 (0.0510)
λ_1	-0.2105 (0.0260)	-0.3717 (0.0037)	-0.2531 (0.0258)
λ_2	-0.0104 (0.0017)	-0.2876 (0.0109)	-0.1407 (0.0436)
α	0.0037 (0.0001)	0.0034 (0.0001)	0.0033 (0.0019)
β	0.0058 (0.0002)	0.0072 (0.0022)	0.0046 (0.0190)
$\mathbb{A}_1(3,0)$ parameters	latent 1	latent 2	latent 3
θ	0.0025 (0.0001)	0	0
κ	0.1049 (0.0038)	0.6237 (0.0169)	1.5943 (0.0401)
λ_1	-0.1898 (0.0040)	0	0
λ_2	0	-0.1151 (0.0129)	-0.1204 (0.0225)
α	0	0.0008 (0.0001)	0.0005 (0.0001)
β	0.0034 (0.0002)	0	0

Table A-5. Parameter Estimates for Three-factor Yields-only Models (UK)

$\mathbb{A}_3(3,0)$ parameters	latent 1	latent 2	latent 3
θ	0.0044 (0.0001)	0.0258 (0.0006)	0.0112 (0.0003)
κ	0.1037 (0.0027)	0.7801 (0.0117)	1.7650 (0.0463)
λ_1	-0.2114 (0.0124)	-0.1236 (0.0238)	-0.2570 (0.0059)
λ_2	-0.0103 (0.0018)	-0.2403 (0.0119)	-0.1154 (0.0494)
α	0.0022 (0.0000)	0.0024 (0.0000)	0.0022 (0.0017)
β	0.0040 (0.0002)	0.0052 (0.0034)	0.0042 (0.0110)
$\mathbb{A}_1(3,0)$ parameters	latent 1	latent 2	latent 3
θ	0.0022 (0.0001)	0	0
κ	0.1025 (0.0027)	0.7612 (0.0014)	1.7904 (0.0259)
λ_1	-0.2190 (0.0021)	0	0
λ_2	0	-0.1531 (0.0061)	-0.1324 (0.0301)
α	0	0.0006 (0.0000)	0.0003 (0.0000)
β	0.0028 (0.0001)	0	0

Table A-6. Parameter Estimates for Five-factor Yields-only Models (US)

$\mathbb{A}_5(5,0)$ parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0030	0.0283	0.0165	0.0009	0.0012
	(0.0001)	(0.0014)	(0.0009)	(0.0003)	(0.0003)
κ	0.0724	0.9332	1.3299	2.5322	2.6852
	(0.0027)	(0.0395)	(0.0493)	(0.3452)	(0.5917)
λ_1	-0.2107	-0.3998	-0.2375	-0.1389	-0.2015
	(0.0498)	(0.0129)	(0.0315)	(0.0558)	(0.1164)
λ_2	-0.0107	-0.2795	-0.1404	-0.1631	-0.3742
	(0.0032)	(0.0404)	(0.0485)	(0.0075)	(0.6336)
α	0.0028	0.0038	0.0028	0.0032	0.0012
	(0.0000)	(0.0005)	(0.0009)	(0.0024)	(0.0049)
β	0.0026	0.0072	0.0056	0.0012	0.0017
	(0.0009)	(0.0073)	(0.0067)	(0.1273)	(0.1899)
$\mathbb{A}_1(5,0)$ parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0019	0	0	0	0
	(0.0001)				
κ	0.0749	0.6237	1.4624	2.5241	2.6181
	(0.0028)	(0.0169)	(0.0536)	(0.1102)	(0.1325)
λ_1	-0.1912	0	0	0	0
	(0.0032)				
λ_2	0	-0.1151	-0.1311	-0.1141	-0.0915
		(0.0129)	(0.0596)	(0.0813)	(0.0861)
α	0	0.0009	0.0005	0.0001	0.0002
		(0.0001)	(0.0001)	(0.0001)	(0.0001)
β	0.0026	0	0	0	0
	(0.0002)				

Table A-7. Parameter Estimates for Five-factor Yields-only Models (UK)

$\mathbb{A}_5(5,0)$ parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0031	0.0234	0.0130	0.0012	0.0010
	(0.0001)	(0.0006)	(0.0005)	(0.0001)	(0.0001)
κ	0.1104	0.8341	1.6134	2.7269	2.8618
	(0.0028)	(0.0152)	(0.0292)	(0.1210)	(0.1063)
λ_1	-0.2364	-0.1500	-0.2492	-0.1194	-0.2635
	(0.0856)	(0.0160)	(0.0137)	(0.0147)	(0.0220)
λ_2	-0.0133	-0.2295	-0.1366	-0.1817	-0.2421
	(0.0054)	(0.0222)	(0.0317)	(0.0510)	(0.0462)
α	0.0005	0.0032	0.0028	0.0023	0.0026
	(0.0000)	(0.0005)	(0.0016)	(0.0011)	(0.0003)
β	0.0060	0.0062	0.0052	0.0028	0.0039
	(0.0009)	(0.0057)	(0.0125)	(0.0849)	(0.0337)
$\mathbb{A}_1(5,0)$ parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0002	0	0	0	0
	(0.0000)				
κ	0.0912	0.8762	1.6624	2.4411	2.7631
	(0.0021)	(0.0228)	(0.1260)	(0.0259)	(0.0469)
λ_1	-0.2010(0.0029)	0	0	0	0
λ_2	0	-0.1311	-0.1297	-0.1231	-0.0751
		(0.0197)	(0.0924)	(0.0376)	(0.0315)
α	0	0.0006	0.0004	0.0001	0.0001
		(0.0000)	(0.0001)	(0.0000)	(0.0000)
β	0.0027	0	0	0	0
	(0.0001)				

Table A-8. Parameter Estimates for a Five-factor Macro Model (US)

parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0037	0.0289	0.0131	0.0301	0.0125
	(0.0011)	(0.0003)	(0.0004)	(0.0009)	(0.0005)
κ	0.0503	0.6653	1.8203	0.5366	0.1575
	(0.0104)	(0.0216)	(0.1117)	(0.0154)	(0.0052)
λ_1	-0.2104	-0.6160	-0.2575	-0.1014	-0.1235
	(0.1509)	(0.0905)	(0.0044)	(0.0075)	(0.0209)
λ_2	-0.0106	-0.4040	-0.1414	-0.1082	-0.0161
	(0.0117)	(0.0214)	(0.1161)	(0.0159)	(0.0029)
α	0.0038	0.0033	0.0028	0.0041	0.0027
	(0.0001)	(0.0000)	(0.0018)	(0.0001)	(0.0000)
β	0.0053	0.0072	0.0030	0.0028	0.0074
	(0.0005)	(0.0009)	(0.0019)	(0.0009)	(0.0006)

Table A-9. Parameter Estimates for a Five-factor Macro Model (UK)

parameters	latent 1	latent 2	latent 3	latent 4	latent 5
θ	0.0029	0.0233	0.0122	0.0282	0.0121
	(0.0001)	(0.0001)	(0.0003)	(0.0002)	(0.0003)
κ	0.1127	0.6714	1.0671	0.7322	0.0575
	(0.0036)	(0.0020)	(0.0263)	(0.0023)	(0.0014)
λ_1	-0.2388	-0.5821	-0.2745	-0.1133	-0.1489
	(0.0113)	(0.0096)	(0.0103)	(0.0016)	(0.0070)
λ_2	-0.0113	-0.3879	-0.1374	-0.3611	-0.0175
	(0.0040)	(0.0036)	(0.0266)	(0.0046)	(0.0009)
α	0.0025	0.0022	0.0022	0.0039	0.0025
	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0000)
β	0.0038	0.0052	0.0045	0.0045	0.0033
	(0.0002)	(0.0001)	(0.0014)	(0.0004)	(0.0003)

Table A-10. Parameter Estimates for a Six-factor Macro Model (US)

parameters	latent 1	latent 2	latent 3	latent 4	latent 5	latent 6
θ	0.0037	0.0278	0.0115	0.0287	0.0109	0.0021
	(0.0008)	(0.0005)	(0.0005)	(0.0008)	(0.0004)	(0.0001)
κ	0.0506	0.6654	1.8203	0.5367	0.1576	2.5923
	(0.0065)	(0.0066)	(0.0930)	(0.0144)	(0.0053)	(0.0972)
λ_1	-0.2104	-0.6166	-0.2576	-0.1014	-0.1235	-0.2033
	(0.0738)	(0.0558)	(0.0029)	(0.0073)	(0.0212)	(0.0040)
λ_2	-0.0106	-0.4039	-0.1414	-0.1278	-0.0161	-0.0311
	(0.0068)	(0.0043)	(0.0935)	(0.0149)	(0.0049)	(0.1034)
α	0.0038	0.0033	0.0028	0.0041	0.0028	0.0011
	(0.0000)	(0.0000)	(0.0005)	(0.0000)	(0.0000)	(0.0018)
β	0.0053	0.0072	0.0030	0.0028	0.0074	0.0015
	(0.0006)	(0.0012)	(0.0013)	(0.0007)	(0.0005)	(0.0957)

Table A-11. Parameter Estimates for a Six-factor Macro Model (UK)

parameters	latent 1	latent 2	latent 3	latent 4	latent 5	latent 6
θ	0.0032	0.0254	0.0115	0.0283	0.0125	0.0020
	(0.0001)	(0.0001)	(0.0003)	(0.0002)	(0.0003)	(0.0001)
κ	0.1132	0.6711	1.2141	0.7322	0.0575	3.2592
	(0.0043)	(0.0088)	(0.0306)	(0.0029)	(0.0014)	(0.1529)
λ_1	-0.2388	-0.5823	-0.2742	-0.1133	-0.1484	-0.2033
	(0.0143)	(0.0106)	(0.0101)	(0.0017)	(0.0174)	(0.0308)
λ_2	-0.0113	-0.3876	-0.1374	-0.3711	-0.0175	-0.0311
	(0.0047)	(0.0074)	(0.0308)	(0.0054)	(0.0008)	(0.1515)
α	0.0025	0.0023	0.0022	0.0039	0.0023	0.0010
	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0000)	(0.0040)
β	0.0038	0.0052	0.0045	0.0045	0.0034	0.0013
	(0.0002)	(0.0001)	(0.0040)	(0.0004)	(0.0003)	(0.6893)

APPENDIX B

REGRESSIONS OF EXCESS BOND RETURNS ON THE YIELD SPREAD

Table B-1. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{3m} - ff_t$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	**-1.0563***	**-1.4365***	-1.3381	-0.4276
	(0.4161)	(0.5736)	(0.7265)	(0.3679)
p-value	0.0116	0.0128	0.0665	0.2461
OLS-TC				
γ_1	**-0.4779****	-0.3383	-0.7314	-0.4142
	(0.1420)	(0.2548)	(0.3925)	(0.2769)
p-value	0.0009	0.1855	0.0635	0.1359
UK				
OLS				
γ_1	-0.2781	-0.0404	1.4918	**1.1630***
	(0.3359)	(0.4903)	(1.1012)	(0.5139)
p-value	0.4092	0.9345	0.1780	0.0255
OLS-TC				
γ_1	-0.1069	-0.3829	0.5026	**0.5760****
	(0.0992)	(0.3294)	(0.6111)	(0.1981)
p-value	0.2835	0.2475	0.4127	0.0045

Notes: * denotes significance at the 5 percent level and ** denotes significance at the 1 percent level.

Table B-2. Regressions of Excess Bond Returns on the Yield Spread $(y_t^{60m} - y_t^{3m})$

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	0.1325	0.2124	0.3088	0.0545
	(0.1026)	(0.1491)	(0.2602)	(0.1998)
p-value	0.1977	0.1553	0.2362	0.7854
OLS-TC				
γ_1	-0.0081	0.0199	-0.1937	0.0181
	(0.0298)	(0.0445)	(0.1449)	(0.1057)
p-value	0.7870	0.6559	0.1822	0.8638
UK				
OLS				
γ_1	-0.0895	-0.0538	0.1373	**0.3261****
	(0.0955)	(0.1937)	(0.3183)	(0.1119)
p-value	0.3506	0.7818	0.6670	0.0043
OLS-TC				
γ_1	-0.0475	0.0191	0.0528	**0.0808***
	(0.0349)	(0.0705)	(0.1083)	(0.0336)
p-value	0.1761	0.7866	0.6273	0.0182

Table B-3. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{60m} - y_t^{24m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	-0.3271	-0.2307	-0.2278	-0.5278
	(0.2185)	(0.3386)	(0.5445)	(0.3374)
p-value	0.1355	0.4961	0.6760	0.1188
OLS-TC				
γ_1	-0.0615	0.0477	**-0.4119***	-0.1346
	(0.0364)	(0.1006)	(0.2023)	(0.1938)
p-value	0.0926	0.6355	0.0427	0.4881
UK				
OLS				
γ_1	**-0.4039***	-0.6451	-0.5899	-0.0192
	(0.1777)	(0.3952)	(0.6706)	(0.3734)
p-value	0.0246	0.1050	0.3807	0.9591
OLS-TC				
γ_1	-0.1871	-0.0478	-0.3323	0.0389
	(0.0962)	(0.1909)	(0.3323)	(0.0656)
p-value	0.0541	0.8027	0.3195	0.5545

Table B-4. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{24m} - y_t^{3m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	**0.4764***	**0.6082***	**0.8427***	0.4403
	(0.2126)	(0.2726)	(0.4153)	(0.3151)
p-value	0.0257	0.0264	0.0433	0.1634
OLS-TC				
γ_1	0.0912	0.0216	-0.1173	0.1822
	(0.0733)	(0.0673)	(0.1935)	(0.1529)
p-value	0.2147	0.7484	0.5448	0.2345
UK				
OLS				
γ_1	-0.0384	0.1242	0.5563	**0.7242****
	(0.1377)	(0.2617)	(0.4241)	(0.1649)
p-value	0.7810	0.6359	0.1920	0.0000
OLS-TC				
γ_1	-0.0298	0.0546	0.2327	**0.2393****
	(0.0428)	(0.0882)	(0.1373)	(0.0860)
p-value	0.4872	0.5368	0.0929	0.0064

Table B-5. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{60m} - y_t^{36m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	-0.6088	**-1.4365***	-0.5440	-0.9333
	(0.3672)	(0.5736)	(0.8638)	(0.5173)
p-value	0.0984	0.0128	0.5293	0.0722
OLS-TC				
γ_1	-0.1089	-0.3383	**-0.6475***	-0.2572
	(0.0576)	(0.2548)	(0.3009)	(0.3043)
p-value	0.0597	0.1855	0.0323	0.3988
UK				
OLS				
γ_1	**-0.7421****	-0.0404	-1.4985	-0.4331
	(0.2697)	(0.4903)	(1.0359)	(0.6678)
p-value	0.0063	0.9345	0.1491	0.5171
OLS-TC				
γ_1	**-0.3423***	-0.3829	-0.7797	-0.0009
	(0.1594)	(0.3294)	(0.5647)	(0.1244)
p-value	0.0338	0.2475	0.1701	0.9943

Table B-6. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{36m} - y_t^{24m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	-0.5711	-0.2101	-0.0837	-1.0636
	(0.5019)	(0.8133)	(1.3853)	(0.9099)
p-value	0.2560	0.7963	0.9518	0.2434
OLS-TC				
γ_1	-0.1290	0.1530	-1.0978	-0.2318
	(0.0996)	(0.2476)	(0.6035)	(0.5183)
p-value	0.1966	0.5372	0.0700	0.6550
UK				
OLS				
γ_1	-0.0384	0.1242	0.5563	**0.7242****
	(0.1377)	(0.2617)	(0.4241)	(0.1649)
p-value	0.7807	0.6355	0.1906	0.0000
OLS-TC				
γ_1	-0.0293	0.0546	0.2327	**0.2393****
	(0.0428)	(0.0882)	(0.1373)	(0.0860)
p-value	0.4951	0.5368	0.0929	0.0064

Table B-7. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{24m} - y_t^{12m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	0.3558	0.6219	1.4227	0.6724
	(0.2749)	(0.4277)	(0.7986)	(0.6270)
p-value	0.1965	0.1469	0.0758	0.2844
OLS-TC				
γ_1	0.1121	0.1659	-0.3500	0.2040
	(0.1099)	(0.1297)	(0.4246)	(0.3467)
p-value	0.3086	0.2019	0.4105	0.5568
UK				
OLS				
γ_1	**-0.7421****	**-1.3004***	-1.4985	-0.4331
	(0.2697)	(0.6271)	(1.0359)	(0.6678)
p-value	0.0063	0.0389	0.1491	0.5171
OLS-TC				
γ_1	**-0.3423***	-0.1461	-0.7797	-0.0009
	(0.1594)	(0.3273)	(0.5647)	(0.1244)
p-value	0.0338	0.6562	0.1701	0.9943

Table B-8. Regressions of Excess Bond Returns on the Yield Spread ($y_t^{12m} - y_t^{3m}$)

	Holding Period			
	k= 3 months	k= 6 months	k= 12 months	k= 24 months
US				
OLS				
γ_1	**1.0406****	**1.2224***	1.3319	0.7137
	(0.3769)	(0.4787)	(0.6791)	(0.4844)
p-value	0.0061	0.0111	0.0508	0.1417
OLS-TC				
γ_1	0.1852	-0.1068	-0.1073	0.3782
	(0.1334)	(0.1419)	(0.3083)	(0.2241)
p-value	0.1664	0.4521	0.7280	0.0926
UK				
OLS				
γ_1	-0.7799	-1.0755	-0.4939	0.6311
	(0.4527)	(0.9645)	(1.6199)	(0.7412)
p-value	0.0859	0.2657	0.7607	0.3952
OLS-TC				
γ_1	-0.3586	0.0071	-0.3710	0.2017
	(0.2126)	(0.4214)	(0.7229)	(0.1504)
p-value	0.0943	0.9865	0.6088	0.1828

APPENDIX C

REGRESSIONS OF FUTURE CHANGES IN YIELDS ON THE YIELD SPREAD

Table C-1. Regressions of Future Changes in Policy Rate on the Yield Spread ($y_t^{3m} - ff_t$, $y_t^{60m} - y_t^{3m}$)

policy rate	$y_t^{3m} - ff_t$			$y_t^{60m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	**0.2819****	**0.6116****	**0.9209****	0.0104	0.0481	0.0895
	(0.0544)	(0.1283)	(0.2203)	(0.0262)	(0.0497)	(0.0704)
p-value	0.0000	0.0000	0.0000	0.6925	0.3336	0.2047
$\bar{R}^2$	0.1213	0.1949	0.2425	-0.0025	0.0032	0.0109
OLS-TC						
β_1	0.1199	**0.4132****	**0.5929****	0.0435	0.0270	0.0553
	(0.0698)	(0.1239)	(0.1703)	(0.0222)	(0.0567)	(0.0812)
p-value	0.0877	0.0010	0.0006	0.0518	0.6338	0.4963
$\bar{R}^2$	-0.0017	0.1734	0.1938	-0.0116	-0.0331	-0.0215
UK						
OLS						
β_1	**0.6982****	**1.2665****	**1.7252****	**0.0811***	**0.1579***	**0.2518***
	(0.0795)	(0.1444)	(0.2034)	(0.0327)	(0.0673)	(0.0980)
p-value	0.0000	0.0000	0.0000	0.0144	0.0206	0.0114
$\bar{R}^2$	0.5026	0.5682	0.5481	0.1449	0.1779	0.2370
OLS-TC						
β_1	**0.3926****	**0.5202***	0.5608	**0.0361***	0.0318	0.0169
	(0.0848)	(0.2398)	(0.4313)	(0.0163)	(0.0298)	(0.0327)
p-value	0.0000	0.0320	0.1961	0.0291	0.2887	0.6055
$\bar{R}^2$	0.3668	0.2415	0.1394	0.0948	0.0438	0.0087

Notes: * denotes significance at the 5 percent level and ** denotes significance at the 1 percent level.

Table C-2. Regressions of Future Changes in Policy Rate on the Yield Spread ($y_t^{60m} - y_t^{24m}$, $y_t^{24m} - y_t^{3m}$)

policy rate	$y_t^{60m} - y_t^{24m}$			$y_t^{24m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	-0.0126	0.0284	0.0829	0.0294	0.0902	0.1542
	(0.0441)	(0.0958)	(0.1438)	(0.0614)	(0.1121)	(0.1563)
p-value	0.7761	0.7668	0.5649	0.6328	0.4217	0.3248
$\bar{R}^2$	-0.0032	-0.0036	-0.0004	0.0002	0.0078	0.0157
OLS-TC						
β_1	-0.0343	-0.1317	-0.1133	**0.1273****	0.1072	0.1791
	(0.0310)	(0.1146)	(0.1825)	(0.0332)	(0.0784)	(0.1093)
p-value	0.2704	0.2516	0.5352	0.0002	0.1732	0.1025
$\bar{R}^2$	-0.0446	-0.0178	-0.0228	0.0994	-0.0033	0.0155
UK						
OLS						
β_1	0.0629	0.1280	0.2827	**0.1347****	**0.2635****	**0.3939****
	(0.0821)	(0.1603)	(0.2434)	(0.0420)	(0.0871)	(0.1290)
p-value	0.4441	0.4252	0.2463	0.0015	0.0027	0.0025
$\bar{R}^2$	0.0179	0.0188	0.0482	0.2079	0.2560	0.3016
OLS-TC						
β_1	0.0257	0.0183	-0.0347	**0.0652****	0.0684	0.0518
	(0.0301)	(0.0450)	(0.0464)	(0.0240)	(0.0526)	(0.0552)
p-value	0.3947	0.6851	0.4565	0.0076	0.1953	0.3502
$\bar{R}^2$	0.0030	0.0095	0.0077	0.1528	0.0774	0.0294

Table C-3. Regressions of Future Changes in Policy Rate on the Yield Spread ($y_t^{60m} - y_t^{36m}$, $y_t^{36m} - y_t^{24m}$)

policy rate	$y_t^{60m} - y_t^{36m}$			$y_t^{36m} - y_t^{24m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	-0.0326	0.0109	0.0794	0.0032	0.1618	0.3447
	(0.0801)	(0.1624)	(0.2373)	(0.1028)	(0.2329)	(0.3572)
p-value	0.6845	0.9464	0.7380	0.9753	0.4879	0.3354
$\bar{R}^2$	-0.0026	-0.0042	-0.0023	-0.0035	-0.0012	0.0042
OLS-TC						
β_1	-0.0726	-0.2332	-0.2262	-0.0222	-0.2704	-0.1749
	(0.0464)	(0.1840)	(0.2950)	(0.0907)	(0.2889)	(0.4546)
p-value	0.1198	0.2066	0.4438	0.8072	0.3504	0.7008
$\bar{R}^2$	-0.0367	-0.0138	-0.0183	-0.0526	-0.0245	-0.0284
UK						
OLS						
β_1	0.0463	0.1086	0.3253	0.2368	0.4636	0.8689
	(0.1317)	(0.2528)	(0.3930)	(0.1939)	(0.3890)	(0.5750)
p-value	0.7253	0.6678	0.4085	0.2228	0.2342	0.1317
$\bar{R}^2$	0.0064	0.0033	0.0199	0.0438	0.0511	0.0934
OLS-TC						
β_1	0.0221	0.0151	-0.0789	0.0939	0.0679	-0.0416
	(0.0559)	(0.0799)	(0.0807)	(0.0698)	(0.1087)	(0.1140)
p-value	0.6929	0.8509	0.3304	0.1811	0.5333	0.7154
$\bar{R}^2$	-0.0039	0.0077	0.0116	0.0178	0.0135	0.0039

Table C-4. Regressions of Future Changes in Policy Rate on the Yield Spread ($y_t^{24m} - y_t^{12m}$, $y_t^{12m} - y_t^{3m}$)

policy rate	$y_t^{24m} - y_t^{12m}$			$y_t^{12m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0314	0.1542	0.2732	0.0570	0.1431	0.2365
	(0.0814)	(0.1573)	(0.2217)	(0.1270)	(0.2286)	(0.3197)
p-value	0.7001	0.3279	0.2187	0.6537	0.5320	0.4601
$\bar{R}^2$	-0.0025	0.0037	0.0105	0.0020	0.0075	0.0139
OLS-TC						
β_1	**0.1771***	0.0546	0.1008	**0.2344****	**0.2556***	**0.4065***
	(0.0795)	(0.1611)	(0.2600)	(0.0592)	(0.1282)	(0.1812)
p-value	0.0272	0.7351	0.6985	0.0001	0.0475	0.0258
$\bar{R}^2$	0.0190	-0.0356	-0.0280	0.1323	0.0345	0.0602
UK						
OLS						
β_1	**0.2212***	**0.4345***	**0.6915***	**0.2563****	**0.4988****	**0.7267****
	(0.1000)	(0.2104)	(0.3083)	(0.0667)	(0.1380)	(0.2067)
p-value	0.0278	0.0397	0.0256	0.0001	0.0004	0.0005
$\bar{R}^2$	0.1119	0.1368	0.1833	0.2669	0.3272	0.3632
OLS-TC						
β_1	**0.1029***	0.0946	0.0492	**0.1283****	0.1486	0.1236
	(0.0499)	(0.0945)	(0.1007)	(0.0418)	(0.0975)	(0.0964)
p-value	0.0415	0.3192	0.6262	0.0026	0.1301	0.2020
$\bar{R}^2$	0.0778	0.0378	0.0077	0.2021	0.1116	0.0535

Table C-5. Regressions of Future Changes in 3-month Yield on the Yield Spread ($y_t^{3m} - ff_t$, $y_t^{60m} - y_t^{3m}$)

3-month yield	$y_t^{3m} - ff_t$			$y_t^{60m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	**0.1337***	**0.2854***	**0.5263***	0.0193	0.0598	0.0997
	(0.0528)	(0.1202)	(0.2116)	(0.0261)	(0.0447)	(0.0625)
p-value	0.0117	0.0181	0.0134	0.4594	0.1822	0.1116
$\bar{R}^2$	0.0249	0.0427	0.0842	0.0005	0.0100	0.0164
OLS-TC						
β_1	0.0784	0.0681	0.2096	0.0067	0.0085	0.0009
	(0.0483)	(0.0763)	(0.1079)	(0.0127)	(0.0170)	(0.0259)
p-value	0.1056	0.3731	0.0532	0.6002	0.6184	0.9715
$\bar{R}^2$	0.0122	-0.0327	0.0248	-0.0227	-0.0425	-0.0332
UK						
OLS						
β_1	**0.6110****	**1.0107****	**1.3533****	**0.1043****	**0.1754****	**0.2695****
	(0.0788)	(0.1686)	(0.2449)	(0.0327)	(0.0652)	(0.0964)
p-value	0.0000	0.0000	0.0000	0.0018	0.0081	0.0060
$\bar{R}^2$	0.4444	0.3631	0.3310	0.2551	0.2223	0.2654
OLS-TC						
β_1	**0.4014****	**0.4094****	**0.4974***	**0.0435***	0.0404	**0.1154***
	(0.0733)	(0.1389)	(0.2151)	(0.0187)	(0.0319)	(0.0480)
p-value	0.0000	0.0038	0.0224	0.0219	0.2072	0.0177
$\bar{R}^2$	0.3523	0.1458	0.1318	0.1596	0.0480	0.1218

Table C-6. Regressions of Future Changes in 3-month Yield on the Yield Spread ($y_t^{60m} - y_t^{24m}$, $y_t^{24m} - y_t^{3m}$)

3-month yield	$y_t^{60m} - y_t^{24m}$			$y_t^{24m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0113	0.0701	0.1353	0.0373	0.0960	0.1490
	(0.0382)	(0.0841)	(0.1377)	(0.0596)	(0.0991)	(0.1279)
p-value	0.7681	0.4048	0.3266	0.5319	0.3335	0.2449
$\bar{R}^2$	-0.0028	0.0019	0.0060	0.0028	0.0118	0.0164
OLS-TC						
β_1	-0.0191	-0.0282	-0.0406	0.0368	0.0542	0.0370
	(0.0199)	(0.0287)	(0.0452)	(0.0246)	(0.0326)	(0.0485)
p-value	0.3362	0.3261	0.3691	0.1354	0.0973	0.4459
$\bar{R}^2$	-0.0189	-0.0379	-0.0259	0.0058	-0.0170	-0.0260
UK						
OLS						
β_1	0.1154	0.2165	0.4138	**0.1621****	**0.2676****	**0.3840****
	(0.0853)	(0.1588)	(0.2369)	(0.0430)	(0.0876)	(0.1317)
p-value	0.1772	0.1739	0.0816	0.0002	0.0024	0.0038
$\bar{R}^2$	0.0506	0.0564	0.1048	0.3230	0.2683	0.2800
OLS-TC						
β_1	0.0482	0.0664	0.1273	**0.0700****	0.0574	**0.1339***
	(0.0348)	(0.0554)	(0.0731)	(0.0165)	(0.0389)	(0.0598)
p-value	0.1690	0.2327	0.0837	0.0000	0.1423	0.0269
$\bar{R}^2$	0.0390	0.0215	0.0309	0.1928	0.0445	0.1009

Table C-7. Regressions of Future Changes in 3-month Yield on the Yield Spread ($y_t^{60m} - y_t^{36m}$, $y_t^{36m} - y_t^{24m}$)

3-month yield	$y_t^{60m} - y_t^{36m}$			$y_t^{36m} - y_t^{24m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0076	0.0881	0.1815	0.0558	0.2362	0.4266
	(0.0659)	(0.1395)	(0.2220)	(0.0925)	(0.2062)	(0.3466)
p-value	0.9086	0.5282	0.4142	0.5470	0.2530	0.2193
$\bar{R}^2$	-0.0030	0.0003	0.0036	-0.0020	0.0046	0.0098
OLS-TC						
β_1	-0.0345	-0.0519	-0.0675	-0.0398	-0.0562	-0.0991
	(0.0324)	(0.0467)	(0.0729)	(0.0512)	(0.0745)	(0.1178)
p-value	0.2867	0.2670	0.3559	0.4382	0.4508	0.4011
$\bar{R}^2$	-0.0172	-0.0360	-0.0253	-0.0213	-0.0407	-0.0271
UK						
OLS						
β_1	0.1290	0.2651	0.5728	0.3614	0.6413	**1.1139***
	(0.1385)	(0.2559)	(0.3870)	(0.2025)	(0.3793)	(0.5591)
p-value	0.3525	0.3010	0.1398	0.0754	0.0919	0.0472
$\bar{R}^2$	0.0197	0.0275	0.0666	0.1015	0.0997	0.1520
OLS-TC						
β_1	0.0705	0.1086	0.1656	0.1289	0.1547	0.3973
	(0.0551)	(0.0842)	(0.1079)	(0.0892)	(0.1440)	(0.2196)
p-value	0.2029	0.1996	0.1274	0.1511	0.2848	0.0728
$\bar{R}^2$	0.0266	0.0187	0.0184	0.0543	0.0222	0.0497

Table C-8. Regressions of Future Changes in 3-month Yield on the Yield Spread ($y_t^{24m} - y_t^{12m}$, $y_t^{12m} - y_t^{3m}$)

3-month yield	$y_t^{24m} - y_t^{12m}$			$y_t^{12m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0714	0.2011	0.2952	0.0542	0.1299	0.2107
	(0.0734)	(0.1413)	(0.2062)	(0.1263)	(0.2056)	(0.2573)
p-value	0.3314	0.1558	0.1533	0.6683	0.5280	0.4136
$\bar{R}^2$	0.0020	0.0121	0.0147	0.0017	0.0075	0.0119
OLS-TC						
β_1	0.0046	0.0234	-0.0300	**0.1282****	**0.1876****	**0.1944***
	(0.0383)	(0.0518)	(0.0689)	(0.0348)	(0.0588)	(0.0849)
p-value	0.9041	0.6522	0.6641	0.0003	0.0016	0.0229
$\bar{R}^2$	-0.0258	-0.0430	-0.0316	0.0803	0.0371	0.0155
UK						
OLS						
β_1	**0.2917****	**0.4958***	**0.7567***	**0.2927****	**0.4767****	**0.6593****
	(0.1067)	(0.2084)	(0.3124)	(0.0693)	(0.1400)	(0.2097)
p-value	0.0066	0.0180	0.0160	0.0000	0.0007	0.0018
$\bar{R}^2$	0.2071	0.1820	0.2146	0.3752	0.3026	0.2934
OLS-TC						
β_1	**0.1110***	0.1026	**0.3044***	**0.1302****	0.1052	0.1995
	(0.0522)	(0.0913)	(0.1263)	(0.0289)	(0.0617)	(0.1094)
p-value	0.0353	0.2633	0.0174	0.0000	0.0905	0.0704
$\bar{R}^2$	0.1075	0.0278	0.0842	0.2297	0.0507	0.0936

Table C-9. Regressions of Future Changes in 24-month Yield on the Yield Spread ($y_t^{3m} - ff_t$, $y_t^{60m} - y_t^{3m}$)

24-month yield	$y_t^{3m} - ff_t$			$y_t^{60m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	**0.1864****	**0.3235***	**0.4722***	-0.0017	-0.0163	-0.0245
	(0.0680)	(0.1298)	(0.1881)	(0.0328)	(0.0601)	(0.0801)
p-value	0.0065	0.0132	0.0125	0.9596	0.7866	0.7601
$\bar{R}^2$	0.0473	0.0515	0.0632	-0.0032	-0.0013	-0.0012
OLS-TC						
β_1	**0.1092***	**0.1911***	**0.2900****	0.0070	0.0110	0.0311
	(0.0467)	(0.0852)	(0.1047)	(0.0142)	(0.0232)	(0.0259)
p-value	0.0200	0.0257	0.0060	0.6224	0.6359	0.2307
$\bar{R}^2$	0.0319	0.0251	0.0349	-0.0126	-0.0251	-0.0327
UK						
OLS						
β_1	0.1973	0.2087	0.0870	**0.0704***	0.1093	0.1498
	(0.1176)	(0.2212)	(0.3179)	(0.0350)	(0.0715)	(0.1036)
p-value	0.0958	0.3472	0.7848	0.0461	0.1286	0.1504
$\bar{R}^2$	0.0247	0.0110	0.0020	0.0609	0.0504	0.0548
OLS-TC						
β_1	0.0951	0.0711	0.0023	0.0289	0.0334	0.0175
	(0.0804)	(0.1649)	(0.1612)	(0.0189)	(0.0239)	(0.0354)
p-value	0.2389	0.6671	0.9885	0.1296	0.1651	0.6211
$\bar{R}^2$	0.0006	-0.0173	-0.0116	0.0100	-0.0071	-0.0090

Table C-10. Regressions of Future Changes in 24-month Yield on the Yield Spread ($y_t^{60m} - y_t^{24m}$, $y_t^{24m} - y_t^{3m}$)

24-month yield	$y_t^{60m} - y_t^{24m}$			$y_t^{24m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0406	0.0810	0.1295	-0.0262	-0.0819	-0.1269
	(0.0379)	(0.0767)	(0.1171)	(0.0656)	(0.1177)	(0.1542)
p-value	0.2847	0.2917	0.2698	0.6899	0.4873	0.4111
$\bar{R}^2$	0.0004	0.0030	0.0052	-0.0004	0.0073	0.0108
OLS-TC						
β_1	-0.0104	-0.0090	0.0356	0.0287	0.0410	0.0601
	(0.0239)	(0.0406)	(0.0467)	(0.0240)	(0.0402)	(0.0432)
p-value	0.6635	0.8245	0.4464	0.2327	0.3076	0.1652
$\bar{R}^2$	-0.0133	-0.0265	-0.0382	-0.0035	-0.0188	-0.0303
UK						
OLS						
β_1	0.1488	**0.3152***	**0.5086***	0.0868	0.1050	0.1192
	(0.0821)	(0.1560)	(0.2286)	(0.0497)	(0.1045)	(0.1480)
p-value	0.0708	0.0442	0.0268	0.0812	0.3160	0.4212
$\bar{R}^2$	0.0475	0.0733	0.1088	0.0479	0.0249	0.0186
OLS-TC						
β_1	0.0238	0.0695	0.0336	**0.0492***	0.0385	0.0214
	(0.0479)	(0.0812)	(0.1387)	(0.0243)	(0.0301)	(0.0301)
p-value	0.6201	0.3939	0.8089	0.0452	0.2026	0.4799
$\bar{R}^2$	-0.0118	-0.0090	-0.0099	0.0211	-0.0115	-0.0096

Table C-11. Regressions of Future Changes in 24-month Yield on the Yield Spread ($y_t^{60m} - y_t^{36m}$, $y_t^{36m} - y_t^{24m}$)

24-month yield	$y_t^{60m} - y_t^{36m}$			$y_t^{36m} - y_t^{24m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0718	0.1499	0.2483	0.0809	0.1434	0.2065
	(0.0606)	(0.1205)	(0.1859)	(0.1041)	(0.2072)	(0.3097)
p-value	0.2372	0.2143	0.1826	0.4378	0.4895	0.5053
$\bar{R}^2$	0.0013	0.0049	0.0087	-0.0011	0.0002	0.0006
OLS-TC						
β_1	-0.0196	-0.0158	0.0541	-0.0182	-0.0189	0.0991
	(0.0371)	(0.0646)	(0.0735)	(0.0647)	(0.1061)	(0.1255)
p-value	0.5970	0.8065	0.4623	0.7791	0.8585	0.4305
$\bar{R}^2$	-0.0128	-0.0264	-0.0385	-0.0140	-0.0266	-0.0379
UK						
OLS						
β_1	0.2280	**0.5232***	**0.8725***	0.3647	0.7074	**1.0890***
	(0.1358)	(0.2585)	(0.3795)	(0.1955)	(0.3739)	(0.5417)
p-value	0.0941	0.0438	0.0221	0.0631	0.0594	0.0453
$\bar{R}^2$	0.0377	0.0681	0.1076	0.0563	0.0728	0.0986
OLS-TC						
β_1	0.0197	0.1107	0.0601	0.0897	0.1682	0.0678
	(0.0838)	(0.1413)	(0.2379)	(0.1094)	(0.1782)	(0.3002)
p-value	0.8150	0.4349	0.8009	0.4141	0.3469	0.8217
$\bar{R}^2$	-0.0143	-0.0102	-0.0097	-0.0065	-0.0083	-0.0103

Table C-12. Regressions of Future Changes in 24-month Yield on the Yield Spread ($y_t^{24m} - y_t^{12m}$, $y_t^{12m} - y_t^{3m}$)

24-month yield	$y_t^{24m} - y_t^{12m}$			$y_t^{12m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0178	-0.0083	-0.0443	-0.0802	-0.2105	-0.3089
	(0.0864)	(0.1606)	(0.2224)	(0.1216)	(0.2125)	(0.2751)
p-value	0.8373	0.9587	0.8421	0.5101	0.3227	0.2624
$\bar{R}^2$	-0.0029	-0.0021	-0.0019	0.0065	0.0214	0.0269
OLS-TC						
β_1	0.0130	0.0184	0.0520	**0.0838***	0.1207	0.1596
	(0.0424)	(0.0642)	(0.0717)	(0.0392)	(0.0772)	(0.0898)
p-value	0.7605	0.7746	0.4686	0.0334	0.1190	0.0767
$\bar{R}^2$	-0.0139	-0.0263	-0.0391	0.0133	-0.0063	-0.0195
UK						
OLS						
β_1	0.2212	0.3403	0.4555	0.1200	0.1025	0.0797
	(0.1184)	(0.2432)	(0.3488)	(0.0769)	(0.1608)	(0.2242)
p-value	0.0628	0.1629	0.1925	0.1197	0.5242	0.7223
$\bar{R}^2$	0.0617	0.0502	0.0521	0.0330	0.0099	0.0039
OLS-TC						
β_1	0.1048	0.1174	0.0559	**0.0775***	0.0436	0.0291
	(0.0643)	(0.0856)	(0.1206)	(0.0374)	(0.0549)	(0.0499)
p-value	0.1054	0.1724	0.6437	0.0404	0.4287	0.5615
$\bar{R}^2$	0.0163	-0.0051	-0.0091	0.0188	-0.0161	-0.0101

Table C-13. Regressions of Future Changes in 60-month Yield on the Yield Spread ($y_t^{3m} - ff_t$, $y_t^{60m} - y_t^{3m}$)

60-month yield	$y_t^{3m} - ff_t$			$y_t^{60m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	**0.1451***	**0.2423***	**0.3390***	-0.0199	-0.0577	-0.0839
	(0.0650)	(0.1208)	(0.1718)	(0.0290)	(0.0540)	(0.0719)
p-value	0.0263	0.0457	0.0493	0.4928	0.2859	0.2442
$\bar{R}^2$	0.0309	0.0312	0.0363	0.001	0.0096	0.0125
OLS-TC						
β_1	**0.0711***	**0.1318***	0.1153	0.0010	0.0070	0.0113
	(0.0296)	(0.0590)	(0.0759)	(0.0123)	(0.0221)	(0.0258)
p-value	0.0170	0.0262	0.1297	0.9348	0.7506	0.6619
$\bar{R}^2$	0.0116	0.0045	-0.0228	-0.0100	-0.0257	-0.0381
UK						
OLS						
β_1	0.1131	-0.0004	-0.2311	0.0510	0.0669	0.0845
	(0.1231)	(0.2242)	(0.3413)	(0.0298)	(0.0586)	(0.0887)
p-value	0.3598	0.9986	0.4995	0.0890	0.2558	0.3427
$\bar{R}^2$	0.0094	0.0038	0.0110	0.0348	0.0254	0.0244
OLS-TC						
β_1	0.0796	0.0443	0.0430	0.0249	0.0270	0.0289
	(0.0787)	(0.1843)	(0.2399)	(0.0136)	(0.0253)	(0.0359)
p-value	0.3139	0.8103	0.8581	0.0703	0.2889	0.4224
$\bar{R}^2$	0.0140	-0.0110	-0.0123	0.0251	-0.0026	-0.0062

Table C-14. Regressions of Future Changes in 60-month Yield on the Yield Spread ($y_t^{60m} - y_t^{24m}$, $y_t^{24m} - y_t^{3m}$)

60-month yield	$y_t^{60m} - y_t^{24m}$			$y_t^{24m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0181	0.0258	0.0454	-0.0551	-0.1446	-0.2146
	(0.0342)	(0.0674)	(0.1015)	(0.0556)	(0.1003)	(0.1318)
p-value	0.5971	0.7018	0.6551	0.3218	0.1502	0.1044
$\bar{R}^2$	-0.0021	-0.0016	-0.0012	0.0105	0.0306	0.0404
OLS-TC						
β_1	-0.0100	-0.0069	0.0029	0.0099	0.0248	0.0287
	(0.0220)	(0.0384)	(0.0483)	(0.0198)	(0.0357)	(0.0418)
p-value	0.6508	0.8570	0.9524	0.6152	0.4883	0.4925
$\bar{R}^2$	-0.0091	-0.0263	-0.0393	-0.0087	-0.0229	-0.0363
UK						
OLS						
β_1	0.1163	0.2297	0.3746	0.0593	0.0523	0.0375
	(0.0722)	(0.1355)	(0.2043)	(0.0412)	(0.0820)	(0.1195)
p-value	0.1084	0.0911	0.0676	0.1511	0.5240	0.7538
$\bar{R}^2$	0.0319	0.0481	0.0752	0.0247	0.0106	0.0056
OLS-TC						
β_1	0.0372	0.0488	0.0561	**0.0357***	0.0354	0.0338
	(0.0383)	(0.0791)	(0.1213)	(0.0177)	(0.0310)	(0.0436)
p-value	0.3329	0.5382	0.6444	0.0463	0.2561	0.4402
$\bar{R}^2$	0.0129	-0.0063	-0.0077	0.0261	-0.0039	-0.0083

Table C-15. Regressions of Future Changes in 60-month Yield on the Yield Spread ($y_t^{60m} - y_t^{36m}$, $y_t^{36m} - y_t^{24m}$)

60-month yield	$y_t^{60m} - y_t^{36m}$			$y_t^{36m} - y_t^{24m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	0.0436	0.0816	0.1381	0.0051	-0.0451	-0.0652
	(0.0548)	(0.1064)	(0.1613)	(0.0927)	(0.1820)	(0.2697)
p-value	0.4271	0.4438	0.3926	0.9564	0.8046	0.8091
$\bar{R}^2$	-0.0011	0.0002	0.0016	-0.0029	-0.0019	-0.0019
OLS-TC						
β_1	-0.0139	-0.0111	0.0049	-0.0304	-0.0175	0.0064
	(0.0348)	(0.0601)	(0.0767)	(0.0579)	(0.1036)	(0.1277)
p-value	0.6887	0.8534	0.9486	0.5996	0.8660	0.9599
$\bar{R}^2$	-0.0093	-0.0263	-0.0393	-0.0087	-0.0263	-0.0393
UK						
OLS						
β_1	0.1788	0.3881	0.6646	0.2838	0.5020	0.7639
	(0.1205)	(0.2277)	(0.3405)	(0.1681)	(0.3144)	(0.4742)
p-value	0.1390	0.0893	0.0518	0.0924	0.1114	0.1082
$\bar{R}^2$	0.0256	0.0463	0.0796	0.0373	0.0456	0.0623
OLS-TC						
β_1	0.0483	0.0760	0.0950	0.1088	0.1209	0.1250
	(0.0658)	(0.1378)	(0.2057)	(0.0886)	(0.1750)	(0.2729)
p-value	0.4646	0.5823	0.6450	0.2219	0.4909	0.6477
$\bar{R}^2$	0.0095	-0.0073	-0.0076	0.0180	-0.0054	-0.0083

Table C-16. Regressions of Future Changes in 60-month Yield on the Yield Spread ($y_t^{24m} - y_t^{12m}$, $y_t^{12m} - y_t^{3m}$)

60-month yield	$y_t^{24m} - y_t^{12m}$			$y_t^{12m} - y_t^{3m}$		
	k = 1m	k = 2m	k = 3m	k = 1m	k = 2m	k = 3m
US						
OLS						
β_1	-0.0310	-0.1196	-0.2023	-0.1267	-0.3083	**-0.4421***
	(0.0741)	(0.1415)	(0.1977)	(0.0991)	(0.1720)	(0.2216)
p-value	0.6762	0.3987	0.3069	0.2019	0.0740	0.0469
$\bar{R}^2$	-0.0019	0.0031	0.0066	0.0239	0.0540	0.0661
OLS-TC						
β_1	0.0030	0.0160	0.0095	0.0305	0.0667	0.0865
	(0.0355)	(0.0620)	(0.0724)	(0.0332)	(0.0628)	(0.0751)
p-value	0.9338	0.7967	0.8962	0.3599	0.2887	0.2502
$\bar{R}^2$	-0.0100	-0.0260	-0.0393	-0.0062	-0.0186	-0.0310
UK						
OLS						
β_1	0.1131	-0.0004	-0.2311	0.0708	0.0201	-0.0477
	(0.1231)	(0.2242)	(0.3413)	(0.0641)	(0.1254)	(0.1781)
p-value	0.3598	0.9986	0.4995	0.2702	0.8725	0.7890
$\bar{R}^2$	0.0094	0.0038	0.0110	0.0131	0.0042	0.0047
OLS-TC						
β_1	0.1704	0.2259	0.2739	0.0518	0.0442	0.0433
	(0.0989)	(0.1937)	(0.2871)	(0.0291)	(0.0520)	(0.0805)
p-value	0.0859	0.2443	0.3408	0.0778	0.3968	0.5914
$\bar{R}^2$	0.0401	0.0291	0.0260	0.0213	-0.0073	-0.0100

APPENDIX D

CONDITIONAL MOMENTS WITH MILSTEIN APPROXIMATION

Using the Milstein approximation, I can approximate the first two conditional moments as follows. For the conditional mean,

$$\begin{aligned} E(X_t|X_{t-\Delta}) &= E(X_t - X_{t-\Delta}|X_{t-\Delta}) + X_{t-\Delta} \\ &= \mu(X_{t-\Delta})\Delta + X_{t-\Delta} \end{aligned}$$

Regarding the conditional variance, I have

$$\begin{aligned} & Var(X_t|X_{t-\Delta}) \\ & = E([X_t - X_{t-\Delta} - \mu(X_{t-\Delta})\Delta]^2) \\ & = E([\sigma(X_{t-\Delta})(W_t - W_{t-\Delta}) + \frac{1}{2}\sigma\frac{\partial\sigma}{\partial X}(X_{t-\Delta})((W_t - W_{t-\Delta})^2 - \Delta)]^2) \\ & = E([\sigma(X_{t-\Delta})Y + \frac{1}{2}\sigma\frac{\partial\sigma}{\partial X}(X_{t-\Delta})(Y^2 - \Delta)]^2) \quad (\ let\ Y = W_t - W_{t-\Delta}, Y \sim N(0,\Delta)\ ) \\ & = \int_{-\infty}^{\infty}[\sigma(X_{t-\Delta})Y + \frac{1}{2}\sigma\frac{\partial\sigma}{\partial X}(X_{t-\Delta})(Y^2 - \Delta)]^2 p(Y)dY \quad (\ p(Y) : \text{normal density}\ ) \\ & = \sigma^2(X_{t-\Delta})\int_{-\infty}^{\infty}Y^2p(Y)dY + \sigma^2\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\int_{-\infty}^{\infty}Y^3p(Y)dY \\ & - \sigma^2\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\Delta\int_{-\infty}^{\infty}Yp(Y)dY + \frac{1}{4}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\int_{-\infty}^{\infty}Y^4p(Y)dY \\ & - \frac{1}{2}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta\int_{-\infty}^{\infty}Y^2p(Y)dY + \frac{1}{4}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2\int_{-\infty}^{\infty}p(Y)dY \end{aligned}$$

$(\because E(Y^2) = \Delta,\ E(Y^3) = 0,\ E(Y^4) = 3\Delta^2\ )$

$$= \sigma^2(X_{t-\Delta})\Delta + \frac{3}{4}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2 - \frac{1}{2}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2 + \frac{1}{4}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2$$

Therefore, I can obtain the following.

$$Var(X_t|X_{t-\Delta}) = \sigma^2(X_{t-\Delta})\Delta + \frac{1}{2}\sigma^2\left(\frac{\partial\sigma}{\partial X}(X_{t-\Delta})\right)^2\Delta^2$$

VITA

Ha-Il Park received his Bachelor of Arts degree in Business Administration from Yonsei University in February of 1997. He received his Master of Science degree in Management Engineering from Korea Advanced Institute of Science and Technology in February of 1999. He received his Ph.D. in Economics from Texas A&M University in December of 2009. His fields of specialization are Financial Economics, Financial Econometrics, and Macro Economics. Ha-Il Park can be reached at the Bank of Korea, 110, 3-Ga, Namdaemunno, Jung-Gu, Seoul 100-794, Republic of Korea, or by email at bluechip@bok.or.kr