The Role of Monetary Policy Uncertainty in the Term Structure of Interest Rates

Junko Koeda* Ryo Kato†‡

July 2010

Abstract

We examine the effect of uncertainty arising from policy-shock volatility on yield-curve dynamics. Contrary to the assumption of many macro-finance models, policy-shock processes appear to be time varying and persistent. We allow for this key feature by constructing a no-arbitrage GARCH affine term structure model, in which policy shock volatility is modeled as the conditional volatility of the error term in a Taylor rule. We find that an increase in monetary policy uncertainty raises the medium- and longer-term spreads in a model that incorporates macroeconomic dynamics.

JEL Classification: C13, C32, E43, E44, E52

Keywords: GARCH, Estimation, Term Structure of Interest Rates, Financial markets and the macroeconomy, Monetary policy

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

†Senior economist, Bank of Japan.

‡We are grateful to Kazuo Ueda, Takatoshi Ito, Tack Yun, Shin-ichi Fukuda, Tatsuyoshi Okimoto, Akihiko Takahashi, and seminar participants at the Federal Reserve Board of Governors, the Bank of Japan, the International Monetary Fund, Hitotsubashi University, National Graduate Institute for Policy Studies, and the University of Tokyo for their useful discussion and comments. We thank Yoshifumi Hisata and Masaki Higurashi for the role they played at an early stage of our work. Junko Koeda acknowledges funding from Nomura Foundation for Academic Promotion. All remaining errors are our own.

1 Introduction

The time-varying volatility of factors that explain yield-curve dynamics may have important macroeconomic implications. For example, if the short-term interest rate follows a monetary policy rule such as a Taylor rule, then its conditional volatility captures monetary policy uncertainty that can affect the amount of interest rate risk perceived by market participants. In line with this widely acknowledged idea, some authors (Rudebusch 2002, Rudebusch, Swanson, and Wu, 2006) have suggested investigating the role of uncertainty factors in explaining yield curve dynamics. However, little formal analysis has followed, and most macro-finance no-arbitrage affine term structure models (ATSMs) remain to be homoskedastic.[1] In order to fill this gap, this paper examines the role of uncertainty arising from the heteroskedastic policy shock process in accounting for yield curve dynamics.

In general, policy uncertainty may at times be large and long-lived, while at other times relatively small and short-lived. At a time of unusual distress—for example the Volcker shock in the early 1980s, the Black Monday in 1987, 9/11 in 2001, and the Lehman shock in 2008—the Fed undertook extraordinary action deviating from any known simple policy rule. As a result, uncertainty in the federal funds (FF) and other financial markets has increased. On the other hand, there are indications that FF market volatility has declined since the Federal Open Market Committee (FOMC) began publicly announcing the target FF rate in 1995 (Favero and Mosca, 2001). In a somewhat similar vein, in 2004, the FOMC explicitly signaled that its future course of monetary policy would be less volatile and more predictable for market participants.[2]

On these grounds, it may be more reasonable to assume that the policy shock process consists of large occasional shocks. Once it largely deviates from the policy rule, the increased uncertainty in financial markets cannot easily be eliminated. One way to accommodate this type of shock process is to apply a generalized autoregressive conditional heteroskedasticity (GARCH) process that allows for serial correlation in the con-

[1] A body of empirical evidence, however, indicates that homoskedasticity is disputable (e.g., Brenner, Harjes, and Kroner 1996).

[2] For example, the FOMC made explicit policy commitments with statements such as, "Policy accommodation can be maintained for a considerable period" (August 2003) and "Accommodative monetary policy stance will be removed at a measured pace" (June 2004).

ditional volatility.[3] [4] To this end we construct a discrete-time macro-finance GARCH term structure model. Specifically, we extend Heston and Nandi's (2003) multivariate GARCH "ATSM"[5] with a richer macro structure. The main difference between Heston and Nandi's (2003) model and other GARCH term-structure models is that the yield equation in their model can be written as an affine function of state variables. This allows for greater tractability and generates a closed-form solution for term rates with any maturity as well as option pricing.

With the existing macro-finance ATSMs having performed broadly successfully,[6] we take Ang and Piazzesi (2003) as a point of departure and generalize their model in three directions. First, we allow the short-term interest rate to follow a GARCH-type process with the conditional volatility of the error term following an autoregressive moving average process. Second, we allow the dynamics of macro variables[7] to depend on the lagged short term interest rate as well as their own lagged variables, in a spirit similar to Ang, Piazzesi, and Wei (2006) and Hördahl, Tristani, and Vestin (2006). Thus, the policy interest rate can directly influence future macro variables, and vice versa. Third, to enhance the link between financial econometrics and macroeconomics, we include no latent variables, which are commonly used in many term structure models

[3]Previously developed "pure finance" ATSMs (e.g., Dai and Singleton 2000) are compatible with stochastic volatility, and they typically assume a square-root process for factor heterosckedasticity—for example, in a single-factor ATSM, where the short rate is the only factor explaining yield curves, the factor variance is the level of short rate itself. However, the square-root models tend to overstate the sensitivity of volatility to levels (Brenner *et al.*, 1996), and to date no consensus has been reached on how one should model the short-rate volatility.

[4]Evidence of time-varying conditional volatility can be provided by single-equation GARCH estimation. A regression of the FF rate on a constant, its first lag, 12-month inflation, 12-month change in unemployment (in percent), where the conditional variance of the FF rate follows the autoregressive moving average process, generates statistically significant GARCH and ARCH terms.

[5]"ATSM" in the sense that model-implied yields can be expressed as an affine function of state variables. Because the continuous version of the GARCH equation reduces to an ordinary differential equation rather than an affine diffusion process, our model lies outside the continuous ATSM framework formally defined by Dai and Singleton (2000).

[6]For example, Ang and Piazzesi (2003), using a discrete-time version of the affine class introduced by Duffie and Kan (1996), found that macro factors explain up to 85 perceent of movements in the short and middle parts of yield curves, and around 40 percent at the long end.

[7]In the baseline model, we assume homoskedasity for the dynamics of inflation and real activity. We can extend our model to allow heteroskedasticity for the macro dynamics, though such heteroskedasticity is less evidently confirmed when the sample period is short.

to improve empirical performance, because they alone cannot outfit any macroeconomic interpretations. We show that the inclusion of economically interpretable conditional volatility can significantly improve the empirical fit of the ATSMs, effectively replacing uninterpretable latent factors.

The model-implied conditional volatility is significantly time varying and persistent—it soared in the early 1980s and tapered off during the period of the "Great Moderation." The gradual decline halted in the early 2000s, when the Fed undertook expansionary policy deviating from the Taylor rule (Taylor 2009), but resumed its decline after the FOMC began making explicit policy announcements. Then it increased again during the global financial crisis of late 2000s.

Our model-estimated results indicate that the conditional volatility of the short term interest rate—monetary policy uncertainty—plays a significant role in determining the shape of yield curves in the presence of the Taylor rule and endogenous macro dynamics. An increase in the uncertainty factor in the short-rate dynamics raises term spreads by lifting the middle and longer-end parts of the yield curves. In addition, we focus on a new aspect of policy shock process—policy shock volatility—in explaining yield curves, whereas the existing literature focuses on the policy shock itself, assuming that policy shocks are i.i.d. normal, presumably for tractability. For example, Evans and Marshall (2001), using VARs with yields of various maturities and macro variables, find that positive monetary policy shocks would bear-flatten a yield curve.

To exemplify how our model performs on real data, we set forth a case study, highlighting the so-called Greenspan conundrum period of 2004-06, on the grounds that monetary policy uncertainty declined during this period (for example, see Figure 1). Our model with the estimated parameters successfully generate the continued bear-flattening of yield curves.[8] It also suggests that the greater predictability in monetary policy in this period reined in the risk premium. Meanwhile, it offsets the upward pressures from the rising short term interest rate and the expanding economic activity.

The paper is organized as follows. The next section describes our macro-finance GARCH term-structure model. Section 3 sets out our estimation strategy, and Section 4 discusses estimated results and a case study on the conundrum period of 2004-06

[8] In the run-up to the 2008 global financial crisis, US yield curves continued to bearflatten, despite the consecutive hikes in the FF rate and the expanding economic activity.

during which monetary policy uncertainty declined. Section 5 concludes.

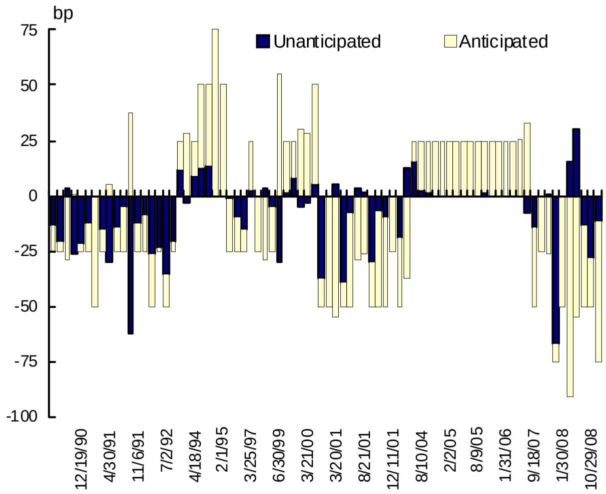

Figure 1. Monetary policy uncertainty (in basis points). Following the methodology of Kuttner (2001), this figure reports recent developments in monetary policy uncertainty; unanticipated policy changes are estimated by differences between the spot-month futures rates before and after each FOMC meeting; anticipated changes are the actual minus the estimated unanticipated changes). During the tightening period of 2004—2007, as can be seen from the figure, the interest rate hikes were mostly well anticipated by investors.

2 The Model

The basic setup of our model essentially builds on the prevailing discrete macro-finance no-arbitrage term structure model, where the stochastic process of the short-term interest rate is driven by a Taylor-type (1993) monetary policy rule. With no-arbitrage bond pricing restrictions, term rates for any maturity can be expressed as an affine function of factors such as the short term interest rate and macro variables.

2.1 Short-term interest rate and macro-variable dynamics

We employ a few variants of the standard Taylor rule that includes the lagged short-term interest rate and expected inflation rate (rather than the concurrent inflation rate). This specification including the expected inflation may be labeled a forward-looking version of the Taylor rule as proposed by Clarida *et al.* (2000). The baseline dynamics of short-term and macro variables are given by

$$r_{t+1} = \underset{1\times1}{\mu_0} + \underset{1\times1}{\mu_1} r_t + \underset{1\times2}{\boldsymbol{\mu}_2} X_{t+1} + \sqrt{h_{t+1}} z_{t+1} \tag{1}$$

$$X_{t+1} = \underset{2\times1}{\boldsymbol{\delta}_0} + \underset{2\times1}{\boldsymbol{\delta}_1} r_t + \underset{2\times2}{\Phi} X_t + \underset{2\times2}{\Sigma} \varepsilon_{t+1} \tag{2}$$

$$h_{t+1} = \underset{1\times1}{\beta_0} + \underset{1\times1}{\beta_1} h_t + \underset{1\times1}{\alpha} z_t^2 \tag{3}$$

$$X_t = [\pi_t \ y_t]', \tag{4}$$

where r_t denotes the short-term interest rate (FF rate). X_t is a 2×1 macro-variable vector of inflation (π), and real activity (y) measures following an autoregressive (AR) process. Σ is an upper triangular matrix, while h_t is the conditional variance of the short-term interest rate. A scalar random shock z and a 2×1 random shock vector ε are assumed to be independent and jointly normal.

We take Ang and Piazzesi (2003) as a point of departure and generalize their model in two directions. First, we allow the short-term interest rate to follow a GARCH-type process with the conditional volatility of the error term following an autoregressive moving average process given by equation (3). Note that h_{t+1} is included in the information set in period t by (3). The $\sqrt{h_{t+1}} z_{t+1}$ term in the short-rate equation (1) could be interpreted as discretionary changes in the FF rate deviated from the Taylor rule. In some preceding macro-finance models as well as in broader monetary policy-related works, the "policy shock" is broadly assumed to be a random shock following i.i.d. normal distribution on account of tractability rather than empirical plausibility. As discussed in the previous section, empirical evidence supports that the policy shock has time-varying (conditional) variance as opposed to the homoskedasticity frequently assumed in most of the early macro-finance studies.

Second, we allow the dynamics of macro variables to depend on the lagged short term interest rate as well as their own lagged variables, in a spirit similar to Ang, Piazzesi, and Wei (2006) and Hördahl, Tristani, and Vestin (2006). Thus, the policy interest rate can directly influence future macro variables. In the next model-estimation section, we will explain that the inclusion of the lagged short term interest rate requires us to modify the Ang and Piazzesi-type specification of the system of equations.

Third, our model has no latent variables, which are commonly used in term structure models to explain the yield curve dynamics, because they alone cannot provide any macroeconomic interpretations. Instead, we treat the conditional volatility of the short term interest rate as an additional factor that explains the yield curves. We then jointly estimate this unobservable variable via maximum likelihood estimation.

Substituting (2) into (1), we obtain

$$\begin{aligned} r_{t+1} &= \mu_0 + \mu_1 r_t + \boldsymbol{\mu}_2 X_{t+1} + \sqrt{h_{t+1}} z_{t+1} \\ &= \mu_0 + \mu_1 r_t + \boldsymbol{\mu}_2 \left(\boldsymbol{\delta}_0 + \boldsymbol{\delta}_1 r_t + \Phi X_t + \boldsymbol{\Sigma}\varepsilon_{t+1}\right) + \sqrt{h_{t+1}} z_{t+1} \\ &= \underbrace{(\mu_0 + \boldsymbol{\mu}_2\boldsymbol{\delta}_0)}_{\overline{\mu}_0} + \underbrace{(\mu_1 + \boldsymbol{\mu}_2\boldsymbol{\delta}_1)}_{\overline{\mu}_1} r_t + \underbrace{(\boldsymbol{\mu}_2\Phi)}_{\overline{\mu}_2} X_t + \sqrt{h_{t+1}} z_{t+1} + \mu_2 \boldsymbol{\Sigma}\varepsilon_{t+1} \end{aligned} \tag{5}$$

$$\text{where } \bar{\mu}_0 = \mu_0 + \boldsymbol{\mu}_2\boldsymbol{\delta}_0,\ \bar{\mu}_1 = \mu_1 + \boldsymbol{\mu}_2\boldsymbol{\delta}_1,\ \bar{\boldsymbol{\mu}}_2 = \boldsymbol{\mu}_2\Phi. \tag{6}$$

The above short-term interest rate and macro-variable dynamics can be rewritten in a more concise form:

$$\begin{pmatrix} r_{t+1} \\ X_{t+1} \end{pmatrix} = \begin{pmatrix} \overline{\mu}_0 \\ \delta_0 \end{pmatrix} + \begin{pmatrix} \overline{\mu}_1 \\ \delta_1 \end{pmatrix} r_t + \begin{pmatrix} \overline{\mu}_2 \\ \Phi \end{pmatrix} X_t + \underbrace{\begin{pmatrix} \sqrt{h_{t+1}} & \boldsymbol{\mu}_2\Sigma \\ 0 & \Sigma \end{pmatrix}}_{\equiv \Sigma_{t+1}} \underbrace{\begin{pmatrix} z_{t+1} \\ \varepsilon_{t+1} \end{pmatrix}}_{\equiv e_{t+1}},$$

$$h_{t+1} = \beta_0 + \beta_1 h_t + \alpha z_t^2.$$

2.2 Pricing kernel and the price of risk

We define a time-dependent 1×3 price of risk vector Ω_t and assume that the price of risk takes a certain affine form in state variables, as handled in many existing affine term structure models.

$$\begin{aligned} \Omega_t' &\equiv (\omega_{r,t}\ \omega_{\pi,t}\ \omega_{y,t}) \\ & \underbrace{\begin{bmatrix} \omega_{0r} \\ \omega_{0\pi} \\ \omega_{0y} \end{bmatrix}}_{\Omega_0} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix}}_{\Omega_1} \begin{bmatrix} r_t \\ \pi_t \\ y_t \end{bmatrix}. \end{aligned} \tag{7}$$

$$\omega_0 \equiv [\omega_{0\pi}\ \omega_{0y}]'\ \omega_1 \equiv [\omega_{21}\ \omega_{31}]'\ \ \tilde{\Omega}_1 \equiv \begin{bmatrix} \omega_{22} & \omega_{23} \\ \omega_{32} & \omega_{33} \end{bmatrix},$$

where Ω_0 is a 3×1 constant vector, and Ω_1 is a 3×3 constant matrix where we impose some zero restrictions.[9] Note that with the zero restriction, $\omega_{r,t} = \omega_{0r}$.

[9] The first row in Ω_js must be zero, as this is a critical condition to ensure that the model lies within the affine framework (in the sense that yield equations can be written as a linear function of factors).

Now suppose that the pricing kernel (m)[10] is given by

$$m_{t+1} \equiv \exp(-r_t + \Omega_t \Sigma_{t+1} e_{t+1} - \frac{1}{2}\Omega_t \Sigma_{t+1} \Sigma_{t+1}' \Omega_t').$$

Then the log price of n-period bond follows the following affine form (see Appendix B for the derivation):

$$\begin{aligned}
p_t^n &= \exp(\bar{A}_n + \bar{B}_n r_t + \bar{C}_n h_{t+1} + \bar{\mathbf{D}}_n X_t), \\
\text{where } \bar{A}_{n+1} &= \bar{A}_n + \bar{B}_n \overline{\mu}_0 + \bar{C}_n \beta_0 + \bar{\mathbf{D}}_n \delta_0 + \frac{1}{2} H_n \Sigma \Sigma' H_n' \qquad (8) \\
&\quad - \frac{1}{2}\log(1 - 2\bar{C}_n \alpha) + \omega_{0r} \boldsymbol{\mu}_2 \Sigma \Sigma' H_n' + H_n \Sigma \Sigma' \omega_0 \\
\bar{B}_{n+1} &= \bar{B}_n \mu_1 + \bar{\mathbf{D}}_n \boldsymbol{\delta}_1 + H_n \Sigma \Sigma' \omega_1 - 1 \qquad (9) \\
\bar{C}_{n+1} &= \bar{C}_n \beta_1 + \bar{B}_n \omega_{0r} + \frac{1}{2} \bar{B}_n^2 \qquad (10) \\
\bar{\mathbf{D}}_{n+1} &= \bar{\mathbf{D}}_n \Phi + \bar{B}_n \overline{\mu}_2 + H_n \Sigma \Sigma' \tilde{\Omega}_1 \qquad (11) \\
\overline{\mu}_1 &= \mu_1 + \boldsymbol{\mu}_2 \delta_1 \,,\ \overline{\boldsymbol{\mu}}_2 = \boldsymbol{\mu}_2 \Phi,\ H_n = \bar{B}_n \boldsymbol{\mu}_2 + \mathbf{D}_n.
\end{aligned}$$

Note that according to basic asset pricing theory, the n-period bond yield is given by

$$r_t^n = A_n + B_n r_t + C_n h_{t+1} + \mathbf{D}_n X_t,$$

where $A_n = -\bar{A}_n/n,\ B_n = -\bar{B}_n/n,\ C_n = -\bar{C}_n/n,\ \mathbf{D}_n = -\bar{\mathbf{D}}_n/n.$

3 Model Estimation

For our estimation, we use monthly data on interest rates and macro variables that capture inflation and real activity from July 1954 to December 2009.[11] We assume that the policy reaction function remains fully stable throughout the period.[12] The summary statistics and data sources are provided in Appendix C.

[10] For the pricing kernel expressed in terms of risk-neutral probabilities, see Appendix A.

[11] Our sample period starts from July 1954 because the FF rate data are available from that month.

[12] Thus we subscribe the view expressed by Sims and Zha (2006) that the monetary policy rule was stable.

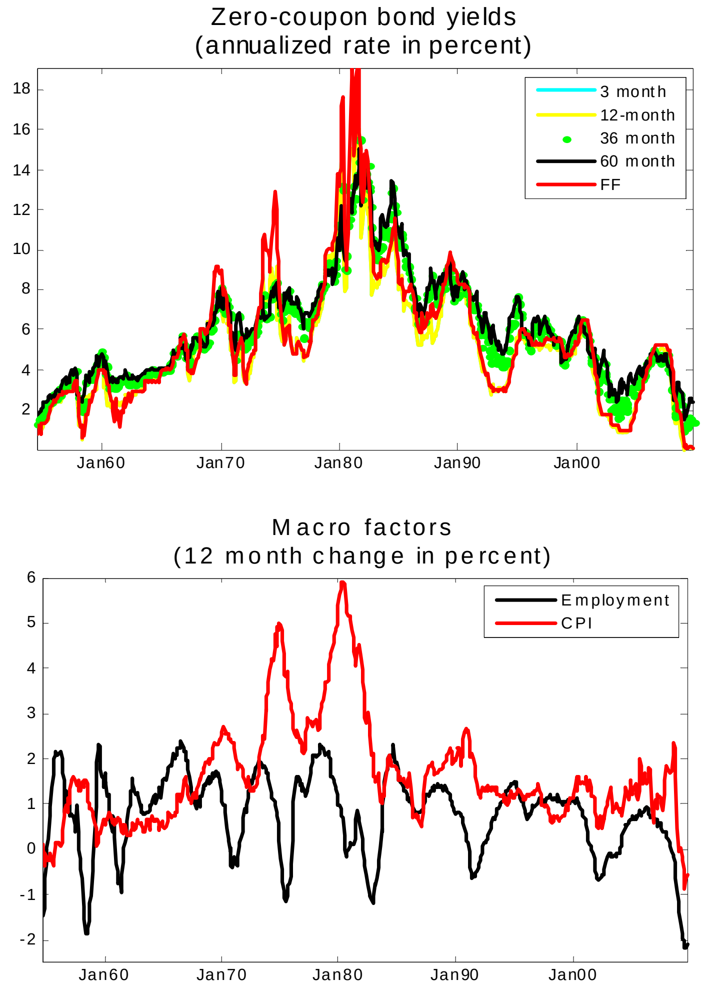

Figure 2. Bond yields and macro principal components. The top panel plots the monthly FF rate and zero-coupon bond yields of maturity 3 months, 12 months, 36 months, and 60 months at an annualized rate in percent. The bottom panel plots employment and CPI in year-on-year percentage change, representing real activity and inflation, respectively. The sample period is July 1954 to December 2009.

We use the FF rates for the short term interest rate and zero-coupon bond yields of 3-, 12-, 36-, and 60-month maturities (Figure 2, top panel); the FF rates are obtained from the Fed. The bond yields are from the CRSP US Treasury Database (the Fama-Bliss Discount Bond Files for 12-, 36-, and 60-month data and from the Risk-Free Rate Files for 3-month data). All bond yields are continuously compounded and expressed at annualized rates in percentages. Regarding inflation and real activity measures, we use the consumer price index (CPI) and employment data (Figure 2, bottom panel). These macro variables are expressed in the year-on-year difference in logs of the original series (see appendix C for data description).

As explained in the previous section, our model dynamics consist of macro dynamics and static yield equations. The macro dynamics are summarized by equation (2) and the static yield equations are given by

$$R_t = A + Br_t + Ch_{t+1} + \mathbf{D}X_t + \Sigma^m \varepsilon^m_{t+1},$$

where $R_t = \left[r_t^3, r_t^{12}, r_t^{36}, r_t^{60}\right]'$ is a 4×1 vector of bond yields with maturities corresponding to the superscript numbers (in months). The yield dynamics are an affine function of the state variables with the coefficient vectors of A, B, C, and $\mathbf{D}$ corresponding to (i) the constant term, (ii) the short-rate term, (iii) the conditional variance term, and (iv) the macro-variable term, respectively. These vectors are time-invariant 4×1 vectors with maturities corresponding to the subscript numbers (i.e., $A = [A_3, A_{12}, A_{36}, A_{60}]'$ $B = [B_3, B_{12}, B_{36}, B_{60}]'$ $C = [C_3, C_{12}, C_{36}, C_{60}]'$ $\mathbf{D} = [D_3, D_{12}, D_{36}, D_{60}]'$). Their elements are derived from the recursive equations; in other words, the model implicitly imposes cross-equation restrictions reducing the number of parameters to be estimated. Measurement errors ε^m are assumed to have constant variance and Σ^m is a diagonal matrix.

We can summarize the system of equations to be estimated as follows:

$$\underbrace{\begin{pmatrix} r_{t+1} \\ X_{t+1} \\ R_t \end{pmatrix}}_{\equiv Y_{t+1}} = \begin{pmatrix} \overline{\mu}_0 \\ \boldsymbol{\delta}_0 \\ A \end{pmatrix} + \begin{pmatrix} \overline{\mu}_1 \\ \boldsymbol{\delta}_1 \\ B \end{pmatrix} r_t + \begin{pmatrix} 0 \\ 0 \\ C \end{pmatrix} h_{t+1} \tag{12}$$

$$+ \begin{pmatrix} \overline{\boldsymbol{\mu}}_2 \\ \Phi \\ \mathbf{D} \end{pmatrix} X_t + \begin{pmatrix} \sqrt{h_{t+1}} & \boldsymbol{\mu}_2 \Sigma & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma^m \end{pmatrix} \begin{pmatrix} z_{t+1} \\ \varepsilon_{t+1} \\ \varepsilon^m_{t+1} \end{pmatrix},$$

$$h_{t+1} = \lambda + \beta_1(h_t - \lambda) + \alpha(z_t^2 - 1), \tag{13}$$

where z_t, ε_t, and ε_t^m are jointly normal and independent to each other and over time. Thus, the observation equation linking R_t to the state (r_t, h_{t+1}, X_t) is appended to the garch-VAR equations describing the state dynamics. We set the lag of X_t and h_t at one.[13]

[13] We tried other lag lengths, but the corresponding coefficients were insignificant.

λ is the unconditional variance of the short term interest rate given by $(\alpha+\beta_0)/(1-\beta_1)$. We estimate this system using the maximum likelihood method in which the likelihood is the joint density of the sample $(y_1, y_2, ..., y_T)$ (for details, see Appendix D).[14] A cursory glance at the model-implied yields (Figure 3) indicates a good fit to the data.

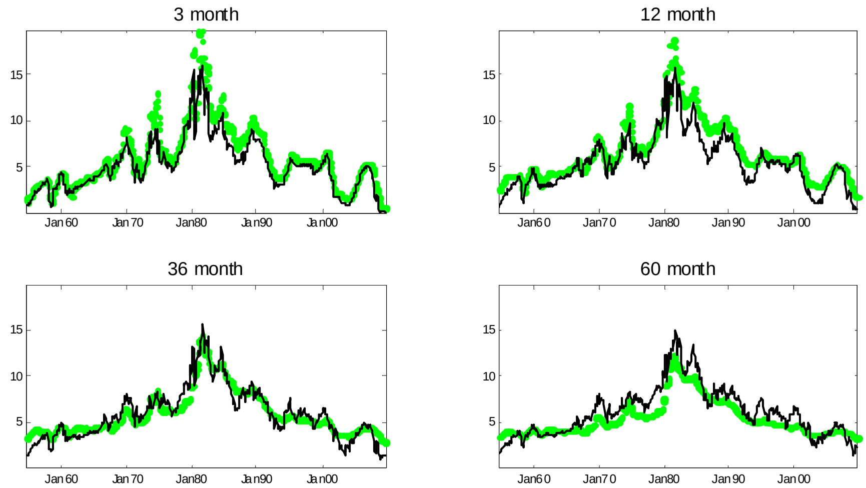

Figure 3: Model-implied yields (in annualized rate in percent). These figures plot model-implied yields for the indicated maturities in annualized rate in percent. The dotted-lines show one-period-ahead in-sample forecasting, and the solid lines show the actual data.

The parameter estimates of our model are reported in Table 1. The estimated conditional volatility of the short rate is highly persistent as the two coefficients in β_1 sum up to near one (0.989). The price of risk coefficients corresponding to inflation and real activity are significant, implying that the macro factors drive time-variation in risk premia. The Taylor rule coefficients $(\mu_0, \mu_1, \boldsymbol{\mu}_2)$ have the right signs with the implied long-run response to inflation (γ)[15] near 1 (0.96)—we cannot reject the null hypothesis that γ is equal to 1 and γ would increase with the exclusion of the recent global financial crisis period. The estimated parameters describing inflation and real-activity $(\delta_0, \delta_1, \Phi)$

[14] That is, the sample here is $(y_1, ..., y_T) = (r_1, X_1, R_0; r_2, X_2, R_1; ..., r_T, X_T, R_{T-1})$. It would be more natural to consider the sample $(r_1, X_1, R_1; r_2, X_2, R_2; ..., r_T, X_T, R_T)$, but the usual factorization argument can be more readily applied to the former. If the sample size T is large, the choice of the sample would not matter for the point estimation.

[15] The implied long-run response to inflation γ can be calculated by $\gamma = \mu_2[1, 0]'/(1-\mu_1)$.

appear reasonable. We will discuss robustness checks of these results in Section 4.2.

Short-rate dynamics

μ_0	μ_1	μ_2	
0.240	0.870	0.125	0.101
0.003	*0.001*	*0.006*	*0.009*

λ	α	β_1
0.063	0.001	0.989
0.005	*3.1E-05*	*0.001*

Dynamics of macro variables

δ_0	δ_1	Φ		Σ	
0.020	0.001	0.972	0.035	0.170	-0.008
0.013	*0.003*	*0.008*	*0.008*	*0.005*	*0.007*
0.121	0.006	-0.067	0.975	---	0.134
0.010	*0.002*	*0.006*	*0.006*	---	*0.004*

Prices of risk

	Ω_0	$\tilde{\Omega}_1$ Inflation	$\tilde{\Omega}_1$ Real activity
Inflation	---	-10.847	-0.479
	---	*0.038*	*0.032*
Real activity	---	4.737	-3.574
	---	0.032	0.036
FF	6.967		
	0.052		

Measurement error

3 months	12 months	36 months	60 months
1.610	2.364	0.891	1.410
0.032	*0.081*	*0.050*	*0.082*

Table 1. Estimated coefficients. This table reports estimated coefficients in our macro-finance GARCH term-structure model. Numbers in italic indicate standard errors. Insignificant prices of risk parameters are set to zero. The delta method is used to calculate the standard errors of μ_0 and μ_1.

4 Estimated Results

4.1 Estimation summary

The key results are as follows. First, our model-implied conditional volatility is considerably time-varying and persistent. Figure 4 reports the dynamics of conditional variance[16] and shows that the model-implied conditional standard deviation increased notably in the wake of the Volcker shock in the early 1980s (left panel) and tapered off during the "Great Moderation." The gradual decline halted in the early 2000s when the

[16] This GARCH process is stationary, as the absolute values of the corresponding polynomial roots are all greater than one.

Fed undertook expansionary policy deviating from the Taylor rule (Taylor 2009) but resumed its decline when the FOMC made explicit policy announcements with statements such as, "policy accommodation can be maintained for a considerable period" (August 2003) and "accommodative monetary policy stance will be removed at a measured pace" (June 2004) (right panel). Then, the conditional volatility increased again during the global financial crisis of late 2000s.

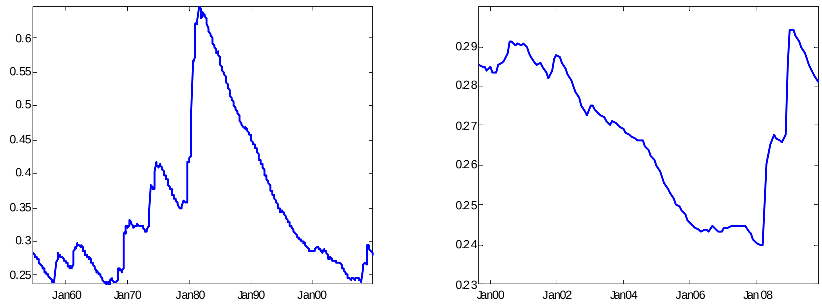

Figure 4: Model-implied conditional standard deviation of the short rate (at an annualized rate in percent). The left panel shows the dynamics of the conditional standard deviation of the short rate for the entire sample period. The right panel enlarges the dynamics in recent years.

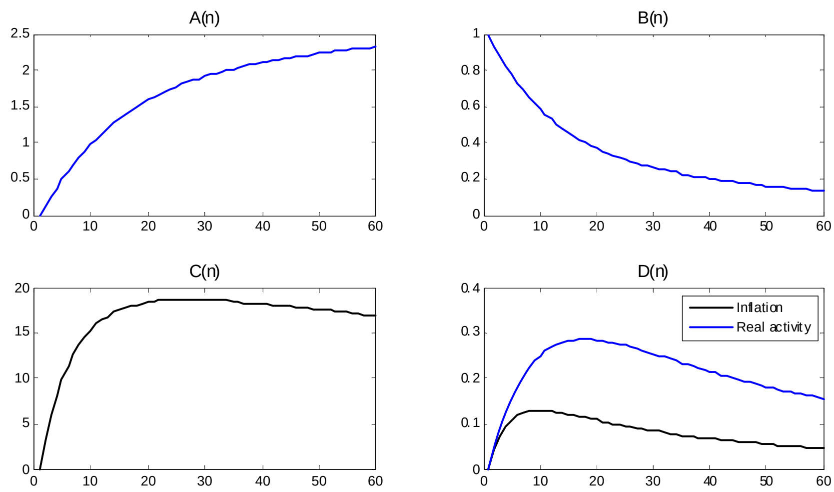

Figure 5. Factor weights against maturity. This figure plots the coefficients of the yield equation against maturity (in months). A(n), B(n), C(n), and D(n) correspond to the constant term, the short-rate term, the conditional-variance term, and macro-variable term, respectively.

Second, our results confirm that the conditional volatility of the short term interest rate plays a significant role in determining yield curves in the presence of endogenous macro dynamics. Figure 5 shows how the yield-equation coefficients change against maturity. The upward-sloping of A_n represents the shape of average yield curves, while the downward slope of B_n implies that an increase in the short term interest rate has a more positive impact on the shorter-end of yield curves, thereby reducing term spreads. The shape of C_n implies that the conditional volatility increases the term spreads by lifting the middle parts and longer-end of yield curves. It implies that one standard deviation increase in h (≈ 0.09) increases the five-year bond yield by more than 150 basis points. The curves of $\mathbf{D}_n$ appear similar to the corresponding dynamics in the existing macro-finance literature, and capture the positive impact of macro variables on yield curves.

Third, in the absence of heteroskedasticity, the model performance deteriorates considerably. Note that we can obtain the homoskedastic version of the model by simply setting the coefficients of the ARCH term (α) and GARCH term (β_1) in the GARCH equation equal to zero and re-maximizing the log-likelihood function. Clearly, this homoskedastic model with no other latent variables turns out to be overly inflexible to provide a reasonable fit to the data, notably at the longer-end of yield curves as shown in Figure 6.

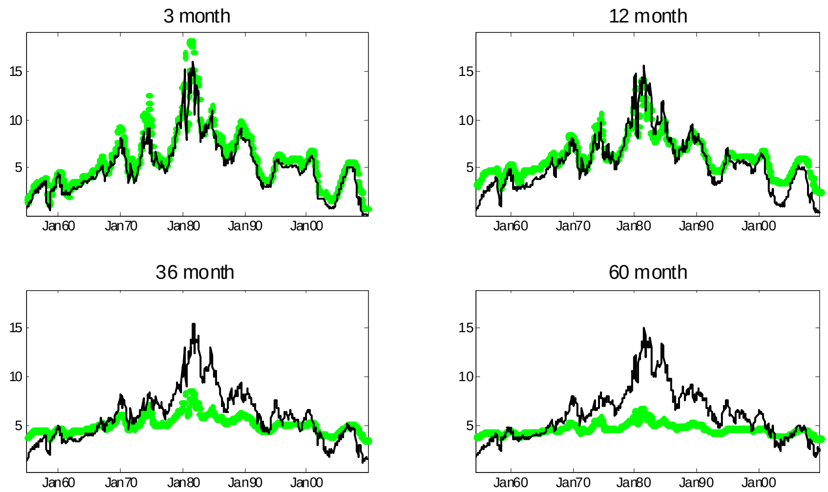

Figure 6: Model-implied yields without heteroskedasticity. With no other latent variables, the model has a poor and unreasonable fit to the data, notably at the longer-end of yield curves.

4.2 Robustness checks

We now turn to some robustness checks of these basic results. The estimated parameters describing inflation and real activity are robust to whether or not we include the term structure with no-arbitrage conditions: they are comparable to on those based on a multivariate GARCH model in the absense of the term structure (for description of multivariate GARCH model, see Appendix E). The GARCH and ARCH coefficients in the GARCH equation (3) are statistically significant as well.

We have considered additional measures of macro variables. Similar to Ang and Piazessi (2003), we sort macro variables in two groups and extract the first principal component of each group of variables separately. The first group consists of inflation measures: CPI-U and the Producer Price Index (PPI) for finished goods. The second group contains real activity measures: industrial production and employment. All variables are expressed in the year-on-year difference in logs of the original series. We use only the first principal components in our analysis because they can explain about 92 percent of variance of real-activity variables and 95 percent of variance of inflation variables. The main results were robust to the inclusion of additional macro variables (the corresponding estimated results are available upon request).

Lastly, we have estimated the model with a shorter sample period from January 1988 to December 2009, i.e., the period that covers Alan Greenspan's tenure as Fed chairman. The main results did not change, although the convergence of maximum likelihood estimators became less smooth (the corresponding estimated results are available upon request).

4.3 A case study: Around the time of the conundrum period

In the runup to the 2008 global financial crisis, US yield curves continued to bear-flatten, despite the consecutive hikes in the FF rate and expanding economic activity. This development, labeled a "conundrum" by then-Fed Chairman Alan Greenspan, poses a challenge to the existing macro-finance models, because they tend to perform poorly in explaining this period unless the term premiums fell beyond the range predicted by these models.

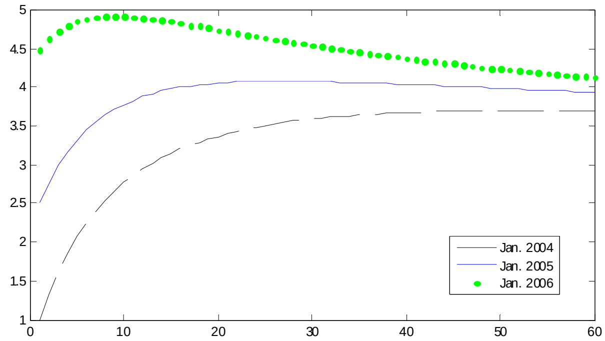

Figure 7. Model-implied yield curves (at an annualized rate in percent) The implied yield curves continued to bear-flatten during the low-yield period.

In our paper, on the other hand, the model-implied yield curves (Figure 7) successfully generate the continued bear-flattening of yield curves between 2004-06. To facilitate understanding of the mechanism behind this bear-flattening, Figure 8 reports factor dynamics around this period: they are characterized by a decline in conditional variance while the short term interest rate was rising and economic activity was expanding. Keeping in mind the factor weights discussed in the previous paragraph, we originally conjectured that it must have been the volatility channel that put downward pressure on the longer rate during this period. The contribution of each term to the model-implied yields, however, only *partially* confirms this conjecture (Figure 9, bottom left panel), as there was a significant decline in model residuals with respect to longer-maturity yield equations, particularly in 2002 (Figure 10). This suggests that there are still unexplained factors accounting for the conundrum. In particular, a demand shift caused by the increased demand for the long-maturity bonds by foreign central banks and institutions might be an important underlying factor.

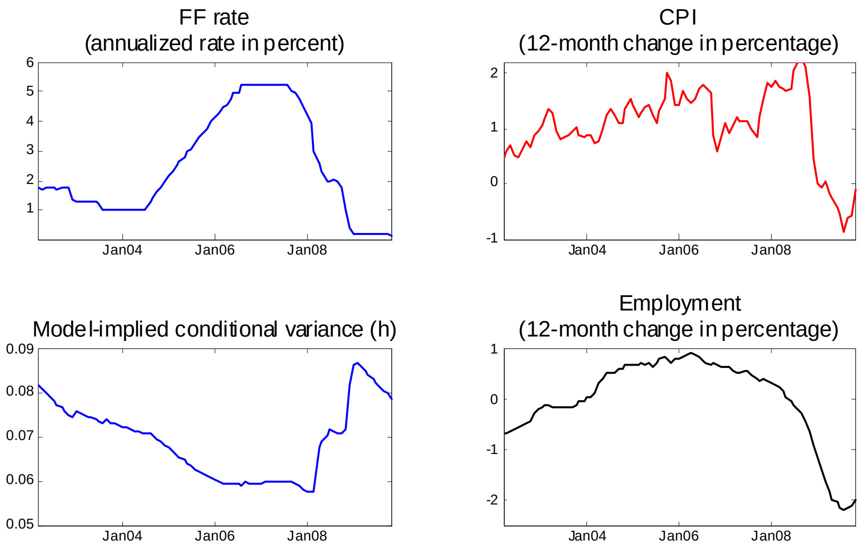

Figure 8. Factor dynamics around the conundrum. These figures plot the dynamics of state variables (i.e., the short rate, the conditional volatility of the short rate, and macro variables between January 2002 and December 2009.

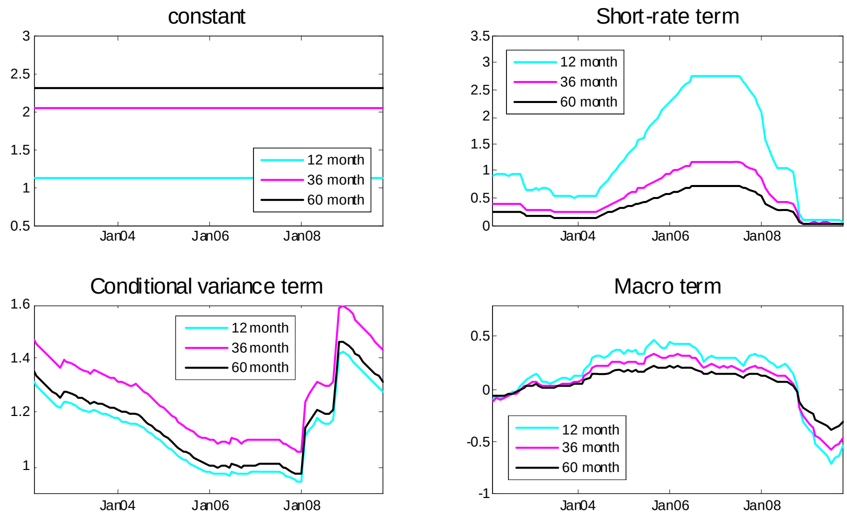

Figure 9. Contributions to the model-implied yields (in annualized rate in percent). These figures demonstrate the contribution to the model-implied yields by each term in the yield equation. Note that the sum of each factor contribution is equal to the model-implied yields.

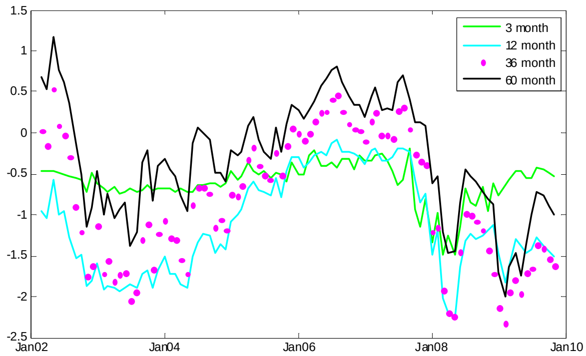

Figure 10. Model residuals for the yield equations (actual minus model-implied yields, at an annualized rate in percent). The model residuals of 3-month-maturity yield equation are more stable than those of longer-maturity yield equations during the conundrum period. The model residuals of 60-month yield equation dropped in 2002 and gradually turned from negative to positive values prior to the global financial crisis.

5 Conclusion

We analyzed a new aspect of monetary policy effects—the role of the policy shock volatility or policy uncertainty—rather than the policy shock itself (i.e., its level or the first moment, in contrast to our focus; the second moment), in accounting for yield curve dynamics. Our estimation results confirmed that the newly included uncertainty factor improved the empirical performance of our ATSM remarkably, greatly reducing the unexplained portion or residuals, particularly at the longer-end of the yield curves. Furthermore, the results indicated that the time-varying and persistent policy shocks increase term spreads as they lift the middle-part or longer-end of the yield curves.

There may be, however, other factors not yet included that could further reveal the unexplained portion of term premium dynamics or model residuals. For example, at a time of unusual distress, if the Fed were to undertake extraordinary policy actions, investors might lose their risk appetite, collectively switching to treasury bonds or other risk-free assets. This sort of "flight to quality" driven by a demand shift could fully offset the upward pressure on the interest rates arising from the elevated uncertainty as discussed in this paper. Looking ahead, the impact of demand-side shifts (i.e., investors' preference) on yield curves could be stressed more in the future research, particularly focusing on the crisis experience.

References

[1] 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 (4), 745-787.

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

[3] Brenner, R.J., R.H. Harjes, and K.F. Kroner. (1996). "Another look at models of the short-term interest rate," *Journal of Financial and Quantitative Analysis*, Vol. 31, No. 1.

[4] Clarida, R., J. Gali, and M. Gertler. (1999). "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, American Economic Association, vol. 37(4), 1661-1707, December 1999.

[5] Dai, Q. and K. J. Singleton. (2000). "Specification analysis of affine term structure models," *Journal of Finance* 55, 1943-1978.

[6] Duffie, D. and R. Kan. (1996). "A yield factor model of interest rates," *Mathematical Finance* 6, 379-406.

[7] Evans, C.L. and D.A. Marshall. (2001). "Economic Determinants of the Nominal Treasury Yield Curve," WP2001-16, Federal Reserve Bank of Chicago.

[8] Favero, C. A. and F. Mosca. (2001). "Uncertainty on monetary policy and the expectations model of the term structure of interest rates," *Economics Letters* 71, 369-375.

[9] Greenspan, A. (2005). Federal Reserve, "Semiannual Monetary Policy Report to the Congress," Board of Governors.

[10] Heston, S. and S. Nandi. (2003). "A Two-Factor Term Structure Model under GARCH Volatility," *The Journal of Fixed Income,* 13, 87-95.

[11] Hördahl, P, O. Tristani, and D. Vestin. (2006). “A joint econometric model of macroeconomic and term structure dynamics,” *Journal of Econometrics* 131, 405-444.

[12] Kuttner, K. N. (2001). “Monetary policy surprises and interest rates: Evidence from the Fed funds futures market,” *Journal of Monetary Economics* 47, 523-544.

[13] Rudebusch, G. D. and T. Wu. (2004). “A macro-finance model of the term structure, monetary policy, and the economy,” Federal Reserve Bank of San Francisco Proceedings, March 2004.

[14] Rudebusch, G. D. and T. Wu. (2007). “Accounting for a Shift in Term Structure Behavior with No-Arbitrage and Macro-Finance Models,” *Journal of Money, Credit and Banking*, Blackwell Publishing, vol. 39(2-3), 395-422.

[15] Rudebusch, G. D., Eric T. Swanson, and Tao Wu. (2006). “The Bond Yield "Conundrum" from a Macro-Finance Perspective,” *Monetary and Economic Studies*, Institute for Monetary and Economic Studies, Bank of Japan, vol. 24(S1), 83-109, December 2006.

[16] Rudebusch, Glenn D. (2002). “Term structure evidence on interest rate smoothing and monetary policy inertia,” *Journal of Monetary Economics*, Elsevier, vol. 49(6), 1161-1187, September 2002.

[17] Sims, C. and T. Zha. (2006). “Were There Regime Switches in U.S. Monetary Policy?,” *The American Economic Review*, Vol. 96, No. 1., 54-81, Mar., 2006.

[18] Taylor, J.B. (1993). “Discretion versus policy rules in practice,” Carnegie-Rochester Conference Series on Public Policy 39, 195–214.

A Pricing kernel and the Risk-Neutral Measure

Assume the existence of an equivalent martingale measure (or risk-neutral measure) Q, such that the price of any asset p_t with no dividends at time $t+1$ satisfies

$$p_t = E_t^Q \left(\exp(-r_t) p_{t+1}\right) \simeq E_t^Q \left(\frac{p_{t+1}}{1+r_t}\right),$$

where expectation is taken under the measure Q and $-log(1+r_t) = log(1+r_t)^{-1} \simeq -r_t$. Let the Radon-Nikodym derivative, which converts the risk-neutral measure to the data-generating measure exploiting the Girsanov theorem, be denoted by ζ_{t+1}. Then, for any random variable Z_{t+1}, we have

$$E_t^Q Z_{t+1} = E_t\left(\frac{\zeta_{t+1}}{\zeta_t} Z_{t+1}\right). \tag{14}$$

Condition 1 *Assume ζ_{t+1} follows the process described as,*

$$\begin{aligned}
\zeta_{t+1} &= \zeta_t \exp\left(\Omega\Sigma_{t+1}e_{t+1} - \frac{1}{2}\Omega\Sigma_{t+1}\Sigma'_{t+1}\Omega'\right) \\
E_t\Sigma_{t+1} &= \Sigma_{t+1},
\end{aligned}$$

where e_t is a vector of random variables that jointly follows $N(0,1)$ distribution and Σ_{t+1} denotes a lower or upper triangular standard deviation matrix. Σ_{t+1} can vary depending on t while it needs to be known at period t.

Under the condition, we define the pricing kernel m_{t+1} as,

$$m_{t+1} \equiv \exp(-r_t) \times \frac{\zeta_{t+1}}{\zeta_t}.$$

Using the kernel, the price of an asset without any dividend can be written as,

$$\begin{aligned}
p_t &= E_t\left(m_{t+1}p_{t+1}\right) \\
&= E_t\left[\exp(-r_t) \times \left(\frac{\zeta_{t+1}}{\zeta_t}\right) \times p_{t+1}\right] = \exp(-r_t)E^Q\left(p_{t+1}\right).
\end{aligned}$$

This clarifies the relationship between the pricing kernel and the risk-neutral measure. As shown here, the pricing kernel effectively adjusts the measure in addition to the discount effect arising from $\exp(-r_t)$.

B Recursive Bond Prices

We can confirm that the n-period bond pricing formula in

$$\begin{aligned}
p_t^{n+1} &= E_t\left(m_{t+1}p_{t+1}^n\right) \\
&= E_t\left[\begin{array}{c} \exp(-r_t + \Omega_t\Sigma_{t+1}e_{t+1} - \frac{1}{2}\Omega_t\Sigma_{t+1}\Sigma'_{t+1}\Omega'_t) \\ \times\exp(\bar{A}_n + \bar{B}_n r_{t+1} + \bar{C}_n h_{t+2} + \bar{\mathbf{D}}_n X_{t+1}) \end{array}\right] \\
&= \exp(-r_t + \bar{A}_n - \frac{1}{2}\Omega_t\Sigma_{t+1}\Sigma'_{t+1}\Omega'_t) \\
&\quad \times E_t\left[\exp(\Omega_t\Sigma_{t+1}e_{t+1} + \bar{B}_n r_{t+1} + \bar{C}_n h_{t+2} + \bar{\mathbf{D}}_n X_{t+1})\right].
\end{aligned}$$

Plugging in the dynamics of X_{t+1}, r_{t+1}, and h_{t+2} into the above gives

$$
\begin{aligned}
p_t^{n+1} &= \exp(-r_t + \bar{A}_n - \frac{1}{2}\Omega_t \Sigma_{t+1} \Sigma_{t+1}' \Omega_t') \\
&\quad \times E_t \left[\exp \begin{pmatrix} \Omega_t \Sigma_{t+1} e_{t+1} + \bar{B}_n \left(\overline{\mu}_0 + \overline{\mu}_1 r_t + \overline{\boldsymbol{\mu}}_2 X_t + \sqrt{h_{t+1}} z_{t+1} + \boldsymbol{\mu}_2 \boldsymbol{\Sigma} \varepsilon_{t+1} \right) \\ + \bar{C}_n h_{t+2} + \bar{\mathbf{D}}_n \left(\boldsymbol{\delta}_0 + \boldsymbol{\delta}_1 r_t + \Phi X_t + \Sigma \varepsilon_{t+1} \right) \end{pmatrix} \right] \\
&= \exp \begin{pmatrix} -r_t + \bar{A}_n - \frac{1}{2}\Omega_t \Sigma_{t+1} \Sigma_{t+1}' \Omega_t' + \bar{B}_n \left(\overline{\mu}_0 + \overline{\mu}_1 r_t + \overline{\boldsymbol{\mu}}_2 X_t \right) \\ + \bar{\mathbf{D}}_n \left(\boldsymbol{\delta}_0 + \boldsymbol{\delta}_1 r_t + \Phi X_t \right) \end{pmatrix} \\
&\quad \times E_t \left[\exp \begin{pmatrix} \Omega_t \Sigma_{t+1} e_{t+1} + \bar{B}_n \left(\sqrt{h_{t+1}} z_{t+1} + \boldsymbol{\mu}_2 \boldsymbol{\Sigma} \varepsilon_{t+1} \right) \\ + \bar{C}_n h_{t+2} + \bar{\mathbf{D}}_n \left(\Sigma \varepsilon_{t+1} \right) \end{pmatrix} \right].
\end{aligned}
$$

At this point, we can spell out the $\bar{\mathbf{C}}_n(.)$ and $\mathbf{h}_{t+2}(.)$ terms in the above as:

$$
\begin{aligned}
\bar{C}_n h_{t+2} &= \bar{C}_n \left[\beta_0 + \beta_1 h_{t+1} + \alpha z_{t+1}^2 \right] \\
&= \beta_0 \bar{C}_n + \bar{C}_n \beta_1 h_{t+1} + \bar{C}_n \alpha z_{t+1}^2,
\end{aligned}
$$

where s_1 and S are the selection vector and matrix, respectively. In the expectations operator, rearranging the terms leaves:

$$
\left[\exp \begin{pmatrix} \Omega_t \Sigma_{t+1} e_{t+1} \\ + \bar{B}_n \left(\sqrt{h_{t+1}} z_{t+1} + \boldsymbol{\mu}_2 \boldsymbol{\Sigma} \varepsilon_{t+1} \right) \\ + \bar{C}_n h_{t+2} \\ + \bar{\mathbf{D}}_n \left(\Sigma \varepsilon_{t+1} \right) \end{pmatrix} \right] = \left[\exp \begin{pmatrix} \Omega_t \Sigma_{t+1} e_{t+1} \\ + \bar{B}_n \sqrt{h_{t+1}} z_{t+1} \\ + \left(\bar{B}_n \boldsymbol{\mu}_2 + \bar{\mathbf{D}}_n \right) \boldsymbol{\Sigma} \varepsilon_{t+1} \\ + \beta_0 \bar{C}_n + \bar{C}_n \beta_1 h_{t+1} \\ + \bar{C}_n \alpha z_{t+1}^2 \end{pmatrix} \right].
$$

Putting this back into the bond pricing formula leaves

$$\begin{aligned}
p_t^{n+1} &= E_t\left(m_{t+1}p_{t+1}^n\right) \\
&= \exp\left(\begin{array}{c} -r_t + \bar{A}_n - \frac{1}{2}\Omega_t\Sigma_{t+1}\Sigma_{t+1}'\Omega_t' + \bar{B}_n\left(\overline{\mu}_0 + \overline{\mu}_1 r_t + \overline{\boldsymbol{\mu}}_2 X_t\right) \\ +\bar{\mathbf{D}}_n\left(\boldsymbol{\delta}_0 + \boldsymbol{\delta}_1 r_t + \Phi X_t\right) \\ +\beta_0\bar{C}_n + \bar{C}_n\beta_1 h_{t+1} \end{array}\right) \\
&\quad \times E_t\left(\exp\left(\begin{array}{c} \Omega_t\Sigma_{t+1}e_{t+1} \\ +\bar{B}_n\sqrt{h_{t+1}}z_{t+1} + \underbrace{\left(\bar{B}_n\boldsymbol{\mu}_2 + \bar{\mathbf{D}}_n\right)}_{\equiv \mathbf{H}_n}\boldsymbol{\Sigma}\varepsilon_{t+1} \\ +\bar{\mathbf{C}}_n\alpha z_{t+1}^2 \end{array}\right)\right)
\end{aligned}$$

$$\begin{aligned}
&= \exp\left(\begin{array}{c} -r_t + \bar{A}_n - \frac{1}{2}\Omega_t\Sigma_{t+1}\Sigma_{t+1}'\Omega_t' + \bar{B}_n\left(\overline{\mu}_0 + \overline{\mu}_1 r_t + \overline{\boldsymbol{\mu}}_2 X_t\right) \\ +\bar{\mathbf{D}}_n\left(\boldsymbol{\delta}_0 + \boldsymbol{\delta}_1 r_t + \Phi X_t\right) \\ +\beta_0\bar{C}_n + \bar{C}_n\beta_1 h_{t+1} \end{array}\right) \\
&\quad \times E_t\left(\exp\left(\left[\Omega_t\Sigma_{t+1} + \underbrace{\left(\bar{B}_n\sqrt{h_{t+1}}\ \mathbf{H}_n\Sigma\right)}_{\equiv J_n}\right]e_{t+1} + \bar{C}_n\alpha z_{t+1}^2\right)\right).
\end{aligned}$$

Now with the aid of proposition used in Heston and Nandi (2003), i.e., $E_t \exp\left(a z_{t+1}\right) = \exp(a^2 1/2)$, and $E_t \exp\left[k\left(z_{t+1}-a\right)^2\right] = \exp\left(\frac{ka^2}{1-2k} - \frac{1}{2}\log\left(1-2k\right)\right)$, where z is i.i.d standard normal, all $t+1$ variables $(z_{t+1}, \varepsilon_{t+1}, z_{t+1}^2)$ can be taken out from the expectations operators:

$$\begin{aligned}
E_t\left[\exp\left(\left[\Omega_t\Sigma_{t+1} + J_n\right]e_{t+1}\right)\right] &= \exp\left[\frac{1}{2}\left(\Omega_t\Sigma_{t+1}\Sigma_{t+1}'\Omega_t' + J_nJ_n' + 2\Omega_t\Sigma_{t+1}J_n'\right)\right] \\
&= \exp\left[\frac{1}{2}\left(\begin{array}{c} \Omega_t\Sigma_{t+1}\Sigma_{t+1}'\Omega_t' + \bar{B}_n^2 h_{t+1} + \mathbf{H}_n\Sigma\Sigma'\mathbf{H}_n' \\ +2\left(\begin{array}{c} \omega_{0r}\bar{B}_n h_{t+1} + \omega_{0r}\boldsymbol{\mu}_2\Sigma\Sigma'\mathbf{H}_n' \\ +\mathbf{H}_n\Sigma\Sigma'\omega_0 + \mathbf{H}_n\Sigma\Sigma'\omega_1 r_t \\ +\mathbf{H}_n\Sigma\Sigma'\tilde{\Omega}_1 X_t \end{array}\right) \end{array}\right)\right] \\
E_t\left[\bar{\mathbf{C}}_n\alpha z_{t+1}^2\right] &= \exp\left[-\frac{1}{2}\log\left(1-2\bar{C}_n\alpha\right)\right].
\end{aligned}$$

The bond price equation can finally be rewritten as

$$
\begin{aligned}
p_t^{n+1} &= E_t\left(m_{t+1}p_{t+1}^{n}\right)\\
&= \exp\begin{pmatrix}
+\bar{A}_n + \bar{B}_n\overline{\mu}_0 + \beta_0\bar{C}_n + \bar{\mathbf{D}}_n\boldsymbol{\delta}_0\\
+\frac{1}{2}\mathbf{H}_n\Sigma\Sigma'\mathbf{H}_n' - \frac{1}{2}\log\left(1-2\bar{C}_n\alpha\right)\\
+\omega_{0r}\mu_2\Sigma\Sigma'\mathbf{H}_n' + \mathbf{H}_n\Sigma\Sigma'\omega_0\\
\\
+\left(\bar{B}_n\overline{\mu}_1 + \bar{\mathbf{D}}_n\boldsymbol{\delta}_1 - 1 + \mathbf{H}_n\Sigma\Sigma'\omega_1\right)r_t\\
\\
+\left(\bar{C}_n\beta_1 + \frac{1}{2}\bar{B}_n^2 + \omega_{0r}\bar{B}_n\right)h_{t+1}\\
\\
+\left(\bar{B}_n\overline{\boldsymbol{\mu}}_2 + \bar{\mathbf{D}}_n\Phi + \mathbf{H}_n\Sigma\Sigma'\tilde{\Omega}_1\right)X_t
\end{pmatrix},
\end{aligned}
$$

corresponding to equations (8) - (11).

C Data

Table AC-1. Summary Statistics of the Data

	Central moments				Autocorrelations		
	Mean	Stdev	Skew	Kurt	Lag 1	Lag 2	Lag 3
FF rate	5.523	3.384	1.181	5.011	0.987	0.965	0.942
3-month	5.120	2.908	0.999	4.552	0.985	0.966	0.948
12-month	5.506	2.898	0.834	3.932	0.986	0.968	0.952
36-month	5.884	2.761	0.840	3.664	0.990	0.976	0.965
60-month	6.106	2.670	0.872	3.515	0.991	0.980	0.971
CPI	1.625	1.200	1.329	4.871	0.991	0.976	0.959
Employment	0.758	0.913	-0.804	3.491	0.985	0.958	0.915

Note: Normal distribution has skewness of zero and kurtosis of 3.

Table AC-2. Data sources

Variable	Source
Federal funds rate	Fed
Zero coupon bond yields (3, 12, 36, 60 month) 1/	CRSP US Treasury Database
CPI-U, all items, seasonally adjusted (1982-84=100)	Bureau of Labor Statistics
PPI for finished goods, seasonally adjusted (base year=1982)	Bureau of Labor Statistics
Nonfarm payroll employment	Establishment Survey Data, Bureau of Labor Statistics
Industrial production, major industry groups, seasonally adjusted (2000=100)	FRB

1/ CRSP currently does not provide zero-coupon bond yield data longer than five years.

D The Log-Likelihood Function

In this appendix, we explain the derivation of the log likelihood function used in the paper. Our likelihood function is different from that for the standard multivariate garch model in the sense that the static yield equations are appended to the state dynamics. In preparation for the following discussion, we summarize the model as follows.

$$y_t = A_Y + \tilde{\Phi} y_{t-1} + C_Y h_t + \tilde{\Sigma}_t u_t, \tag{15}$$

$$h_t = \lambda + \beta_1 (h_{t-1} - \lambda) + \alpha (z_{t-1}^2 - 1), \tag{16}$$

where

$$y_t = (r_t, X_t, R_{t-1})', \; u_t = (z_t, \varepsilon_t, \varepsilon_t^m)',$$

$$\lambda = (\alpha + \beta_0)/(1 - \beta_1),$$

$$\tilde{\Phi} = \begin{bmatrix} \overline{\mu}_0 & \overline{\mu}_1 & \overline{\boldsymbol{\mu}}_2 & 0 \\ \boldsymbol{\delta}_0 & \boldsymbol{\delta}_1 & \Phi & 0 \\ A & B & \mathbf{D} & \mathbf{0} \end{bmatrix}, \; C_Y = \begin{pmatrix} 0 \\ 0 \\ C \end{pmatrix},$$

$$\tilde{\Sigma}_t = \begin{pmatrix} \sqrt{h_t} & \boldsymbol{\mu}_2 \Sigma & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma^m \end{pmatrix}, \; \overline{\boldsymbol{\mu}}_2 = \boldsymbol{\mu}_2 \Phi, \; \overline{\mu}_1 = \underbrace{1 - \boldsymbol{\mu}_2 [1, 0]'/\gamma}_{\mu_1} + \boldsymbol{\mu}_2 \boldsymbol{\delta}_1$$

and z_t, ε_t,and ε_t^m are jointly normal and independent to each other and over time. The elements in $A, B, \mathbf{C}$, and D are given recursively in equations $(8)-(11)$ in the text. We denote the vector of parameters to be estimated as θ,

$$\theta = \left[\boldsymbol{\delta}_0, \boldsymbol{\delta}_1, \Phi, \lambda, \beta_1, \alpha, \Sigma, \overline{\mu}_0, \gamma, \boldsymbol{\mu}_2, \Sigma^m, \Omega_0, \omega_1, \tilde{\Omega}_1\right].$$

We wish to describe the joint density of $(y_T, y_{T-1}, ..., y_1)$ given (y_0, y_{-1}, h_0). (The reason for conditioning the joint density with (y_0, y_{-1}, h_0) will be explained later in this appendix.) Note that the joint density of observations 1 through t conditioned on y_0, y_{-1},and h_0 satisfies

$$\begin{aligned} & f\left(y_t, y_{t-1}, ..., y_1 | y_0, y_{-1}, h_0; \theta\right) \\ & = f\left(y_{t-1}, ..., y_1 | y_0, y_{-1}, h_0; \theta\right) \\ & \times f\left(y_t | y_{t-1}, ..., y_0, y_{-1}, h_0; \theta\right), \end{aligned} \tag{17}$$

and through the usual sequential substitution, the joint density of $(y_T, y_{T-1}, ..., y_1)$ given (y_0, y_{-1}, h_0) satisfies

$$\begin{aligned} & f\left(y_T, y_{T-1}, ..., y_1 | y_0, y_{-1}, h_0; \theta\right) \\ & = \prod_{t=1}^{T} f\left(y_t | y_{t-1}, ..., y_0, y_{-1}, h_0; \theta\right). \end{aligned} \tag{18}$$

We are now ready to derive the conditional distribution in (18), i.e., $f\left(y_t | y_{t-1}, ..., y_0, y_{-1}, h_0; \theta\right)$. Since u_t is i.i.d. standard normal, the distribution of u_t conditioned on $(y_{t-1}, ..., y_0, y_{-1}, h_0)$ is

$$u_t | y_{t-1}, ..., y_0, y_{-1}, h_0 \sim N(0, I),$$

Note that by (15), u_0 is a function of (y_0, y_{-1}, h_0) and thus by (16), h_1 is also a function of (y_0, y_{-1}, h_0). In the following period, u_1 is a function of (y_1, y_0, y_{-1}, h_0) and h_2 is also a function of $(y_{2,} y_0, y_{-1}, h_0)$. It follows that u_t is a function of $(y_t, y_{t-1}, ..., y_0, y_{-1}, h_0)$ and h_t is a function of $(y_{t-1}, ..., y_0, y_{-1}, h_0)$. Since h_t is nonrandom given $(y_{t-1}, ..., y_0, y_{-1}, h_0)$, the distribution of y_t conditioned on $(y_{t-1}, ..., y_0, y_{-1}, h_0)$ is

$$\begin{aligned} & y_t | y_{t-1}, ..., y_0, y_{-1}, h_0 \\ & \sim N(A_Y + \tilde{\Phi} y_{t-1} + C_Y h_t, \tilde{\Sigma}_t \tilde{\Sigma}_t'). \end{aligned}$$

Therefore, the joint density of $(y_T, ..., y_1)$ conditioned on (y_0, y_{-1}, h_0) is given by

$$
\begin{aligned}
& f\left(y_T, y_{T-1}, ..., y_1 | y_0, y_{-1}, h_0; \theta\right) && (19) \\
& = (2\pi)^{-T/2} \left|(\tilde{\Sigma}_t \tilde{\Sigma}_t')^{-1}\right|^{1/2} \\
& \times \exp\left[(-1/2)\left(y_t - A_Y - \tilde{\Phi} y_{t-1} - C_Y h_t\right)' (\tilde{\Sigma}_t \tilde{\Sigma}_t')^{-1} \left(y_t - A_Y - \tilde{\Phi} y_{t-1} - C_Y h_t\right)\right].
\end{aligned}
$$

The conditional log likelihood is the log of the above expression, i.e.,

$$
\begin{aligned}
L\left(\theta\right) = const & - \frac{1}{2}\sum_{t=1}^{T} \log |\tilde{\Sigma}_t \tilde{\Sigma}_t'| \\
& - \frac{1}{2}\sum_{t=1}^{T} \left(y_t - A_Y - \tilde{\Phi} y_{t-1} - C_Y h_t\right)' (\tilde{\Sigma}_t \tilde{\Sigma}_t')^{-1} \left(y_t - A_Y - \tilde{\Phi} y_{t-1} - C_Y h_t\right)
\end{aligned}
$$

If the sample size T is large, the conditional maximum likelihood estimation would be asymptotically the same as the maximum likelihood estimation that maximizes the unconditional log likelihood, $\log f(y_T, ..., y_1)$.

Finally, we numerically maximize the conditional log-likelihood function, with the initial value of h given by the sum of squared residuals of the short-rate dynamics based on the low-inflation period of the 1950s.

E Estimating Macro Dynamics Without the Term Structure of Interest Rates

To see if our estimated parameters for macro dynamics lie within a reasonable range, we estimate the macro dynamics given by (2) and report the estimated results. The only difference between (2) and our macro-finance GARCH ATSM is that the former excludes the term structure.

The log-likelihood function is given by

$$
L(\tilde{\theta}) = -\frac{1}{2}\sum_{t=1}^{T} \log(\det(\tilde{H}_t)) - \frac{1}{2}\sum_{t=1}^{T} \varepsilon_t' \tilde{H}_t^{-1} \varepsilon_t,
$$

where θ is the vector of parameters to be estimated;

$$
\tilde{\theta} = \left[\boldsymbol{\delta}_0, \boldsymbol{\delta}_1, \Phi, \lambda, \beta_1, \alpha, \Sigma, \overline{\mu}_0, \overline{\mu}_1, \overline{\boldsymbol{\mu}}_2\right].
$$

H is the covariance-variance matrix

$$\tilde{H}_t = \begin{bmatrix} h_t + \boldsymbol{\mu}_2\Sigma(\boldsymbol{\mu}_2\Sigma)' & \boldsymbol{\mu}_2\Sigma\Sigma' \\ \Sigma(\boldsymbol{\mu}_2\Sigma)' & \Sigma\Sigma' \end{bmatrix},$$

and ε is the error term in the model defined by:

$$\varepsilon_{t+1} = \begin{pmatrix} r_{t+1} \\ X_{t+1} \end{pmatrix} - \begin{pmatrix} \overline{\mu}_0 \\ \boldsymbol{\delta}_0 \end{pmatrix} - \begin{pmatrix} \overline{\mu}_1 \\ \boldsymbol{\delta}_1 \end{pmatrix} r_t - \begin{pmatrix} \overline{\boldsymbol{\mu}}_2 \\ \Phi \end{pmatrix} X_t.$$

The estimation results are reported in Table AE. The delta method is used to calculate the standard errors of μ_0 and μ_1.

Table AE. Estimation Results: A Multivariate GARCH Model

Short-rate dynamics

μ_0	μ_1	μ_2	
0.049	0.964	0.038	0.097
0.009	*0.002*	*0.006*	*0.013*

λ	α	β_1
0.020	0.088	0.734
0.003	*0.002*	*0.024*

Dynamics of macro variables

δ_0	δ_1	Φ		Σ	
-0.012	0.001	0.986	0.040	0.154	0.007
0.010	*0.003*	*0.008*	*0.007*	*0.004*	*0.006*
0.053	9.00E-04	*-0.029*	*0.986*	---	0.151
0.010	*0.003*	*0.009*	*0.007*	---	*0.004*