Monetary Policy Shifts and the Term Structure*

Andrew Ang†
Columbia University and NBER

Jean Boivin‡
HEC Montréal, CIRPÉE, CIRANO and NBER

Sen Dong§
Lehman Brothers

This Version: 28 January 2008

JEL Classification: C13, C50, E43, E52, G12

Keywords: Quadratic term structure model; Monetary policy; Interest rate risk; time-varying parameter model.

*The paper has benefited from discussions with Michael Johannes and Monika Piazzesi. We thank seminar participants at UT Austin, the Capital Group, and the Federal Reserve Board. We thank Rudy Loo-Kung for excellent research assistance. Andrew Ang and Jean Boivin acknowledge support from the NSF (SES-0137145 and SES-0518770).

†Columbia Business School, 3022 Broadway 413 Uris, New York NY 10027, ph: (212) 854-9154; email: aa610@columbia.edu; WWW: http://www.columbia.edu/∼aa610.

‡HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, Québec, Canada H3T 2A7, Canada; ph: (514) 340-7002; email: jean.boivin@hec.ca; WWW: http://www.hec.ca/en/profs/jean.boivin.html

§Lehman Brothers Inc., 745 7th Ave, New York NY 10019; email: sen.dong@lehman.com

Abstract

We estimate the effect of shifts in monetary policy using the term structure of interest rates. In our no-arbitrage model, the short rate follows a version of the Taylor (1993) rule where the coefficients on inflation and output can vary over time. We find that monetary policy loadings on inflation, but not output, changed substantially over the last 50 years. Agents tend to assign a risk discount to monetary policy shifts and are willing to pay to be exposed to activist monetary policy. Over 1952-2006, if agents had assigned no value to active monetary policy, the slope of the yield curve would have been approximately 50 basis points higher, and up to twice as volatile, than what actually occurred in data.

1 Introduction

A large body of narrative and empirical evidence suggests that the conduct of monetary policy in the U.S. has changed in substantial ways over the last 50 years. The Volcker disinflation in the early 1980's is a well-known example of a drastic change in the way monetary policy is set in response to economic developments. The possibility of shifts in monetary policy has spurred a large body of empirical research attempting to document and quantify the importance of these changes.[1] One important reason to be interested in these changes in monetary policy is that they can serve as "monetary policy experiments" that could help us better identify and measure the effect of systematic monetary policy on the economy. So far, the focus in the literature has been mainly on the impact of monetary policy on real activity and inflation. In particular, a lot of attention has recently been devoted to determining the role played by monetary policy in explaining the "Great Moderation," referring to the fact that the volatility of real activity and inflation, and other macro series, has decreased since the mid-1980's.[2]

One aspect that has received little attention so far is the implications of these monetary policy changes for financial markets, in particular for the term structure of interest rates.[3] A growing number of studies that have employed macro factors in term structure models have found that macroeconomic fluctuations are an important source of uncertainty affecting bond risk premia (see, among others, Ang and Piazzesi, 2003; Ang, Bekaert and Wei, 2007). An unanswered question is what are the effects of the Fed's changing monetary policy stances vis-à-vis output or inflation on the term structure. Monetary policy changes affect the entire term structure because the actions of the Fed at the short end of the yield curve influence the dynamics of the long end of the yield curve through no-arbitrage restrictions. Consequently, the term structure of yields also provides valuable information in estimating monetary policy shifts.

It is unclear how changing monetary policy affects long-term yields and the slope of the yield curve. If monetary policy is entirely neutral, then agents would assign the same risk premia to long-term bonds in a world where monetary policy changed over time and in a world where monetary policy was constant. The slope of the yield curve would be the same in both

[1] See, among others, Clarida, Galí and Gertler (2000), Orphanides (2001), Cogley and Sargent (2001, 2005), Sims and Zha (2006), and Boivin (2006).

[2] See, for instance, Stock and Watson (2003), Boivin and Giannoni (2006), Sims and Zha (2006), and Justiniano and Primiceri (2006).

[3] One exception is Bikbov (2006), who examines the effect of discrete regime shifts in monetary policy rules on the term structure. Our monetary policy shifts are continuous and we also estimate the price of risk of monetary policy changes, which Bikbov does not do.

worlds. If monetary policy risk is priced and agents dislike the uncertainty of policy changes, then part of the risk premium for holding long-term bonds is needed to compensate investors for their risk aversion to monetary policy shifts. On the other hand, if monetary policy is valued by investors and active monetary policy is priced at a risk discount, then the slope of the yield curve in the real world would be lower than the slope of the yield curve in an economy where investors assigned no value to Fed policy changes. Similarly, it is not clear whether changes in monetary policy increase or decrease long-term yields and how these yield changes may differ across different parts of the yield curve.

The goal of this paper is to determine the influence of monetary policy on the term structure of interest rates. The existence of historical shifts in monetary policy, or different monetary policy "experiments," provides an opportunity to statistically estimate the effects of changes in the policy rule on the term structure. To do so we estimate a quadratic term structure model, where the dynamics of the short rate follow a version of Taylor's (1993) rule. Our no-arbitrage model allows for the Fed response to inflation and output to potentially vary over time. In contrast to most existing empirical models, we do not impose that the time variation in the policy parameters is exogenous. For instance, we entertain the possibility that a high response to inflation today might be due to the fact that inflation was high in the recent past.

An additional advantage of estimating the Taylor rule with policy shifts jointly with a term structure model is that by including more information, it can potentially yield sharper estimates of the changes in the policy rule. This is potentially important since conflicting evidence has been reported in the literature on the importance of monetary policy shifts, and the evidence based on the estimation of single equation have been subject to considerable statistical uncertainty. Exploiting term structure information in this context could thus lead to more conclusive evidence on how monetary has evolved.

Using the estimated model, we perform a series of exercises. We first document the importance of the historical changes in monetary policy. We then investigate the effect of these changes on the term structure of interest rates by computing the impulse responses of yields at various maturities to shocks to the monetary policy stance. Our key findings can be summarized as follows. First, our estimates suggest that monetary policy changed substantially over the last 50 years. The Fed's sensitivity to inflation has changed markedly over time and our estimates are largely consistent with the evidence reported in Clarida, Galí and Gertler (2000), Cogley and Sargent (2005) and Boivin (2006). The use of term structure information in the estimation of the policy rule leads to sharper parameter estimates, which statistically allow us to reject the

hypothesis that the Taylor principle was satisfied throughout the 1970's. More recently, the coefficient on inflation was below one during the early 2000's. We also find that the Fed's inflation response is more aggressive and flexible, on average, compared to estimates from random-walk coefficient models similar to Cogley and Sargent (2001), Cogley (2005), Boivin (2006), and Justiano and Primiceri (2006).

Second, we find that shifts of monetary policy stances with regards to output shocks exhibit very small variation. Our model estimates imply that most of the discretion in monetary policy has resulted from changing the response of the Fed to inflation shocks rather than to output shocks. The finding of very small variation in the output loadings is the opposite conclusion to estimates from models using random walks to capture the time variation of monetary policy coefficients. In these models estimated without yield curve information, the policy loadings on output shocks exhibit large time variation.

Third, changes in monetary policy have a quantitatively important influence on the shape of the term structure. A shock to the Fed response to inflation fluctuations, ceteris paribus, raises short term rates and shrinks the term spread. We find that in the real world, past monetary policy stances have little effect on future output and inflation. However, under the dynamics of the risk-neutral measure implied by bond prices, tighter monetary policy measured by a higher inflation coefficient has a perceived ability to lower future inflation. These real world versus risk neutral differences suggest that the actions of the Fed are valued by investors.

Finally, we find that Fed policy shifts are priced at a risk discount. Investors are generally willing to pay, rather than requiring to be paid, to become exposed to monetary policy changes, especially if the Fed responses to output and inflation shocks are sufficiently large. We find that if agents do not value activist monetary policy, the slope of the yield curve would have been, on average, 50 basis points higher than what occurred over the 1952-2006 sample and up to two times as volatile. Thus, monetary policy partly helps to determine the low risk premia on long-term bonds.

The rest of the paper is organized as follows. Section 2 describes the modelling framework. It first describes the short rate equation, specified as a time-varying policy reaction function, and then derives bond prices based on a quadratic, arbitrage-free, term structure model. Section 3 discusses the empirical results and describes the estimated time series of the policy coefficients, how policy changes affect the yield curve, and how policy risk is priced. Section 4 concludes. The details of the bond pricing derivations and the Bayesian estimation technique can be found in the Appendix.

2 Model

We assume the dynamics of the short end of the yield curve (the one-quarter short rate) follows a version of Taylor's (1993) rule where the monetary authority sets the short rate as a function of inflation and the output gap. Unlike a standard Taylor rule, we let the policy responses on output and inflation vary over time:

$$r_t = \delta_0 + (\bar{a} + a_t)g_t + (\bar{b} + b_t)\pi_t, \tag{1}$$

where r_t is the 1-quarter yield, g_t is the output gap, and π_t is inflation. In estimating equation (1) in our model, we also include a small orthogonal error term. The coefficient on the output gap, $(\bar{a} + a_t)$, measures how much the monetary authority adjusts the short rate to output gap shocks and consists of a base level, $\bar{a}$, and a zero-mean time-varying component, a_t. Similarly, the policy response to inflation consists of an average response, $\bar{b}$, and a zero-mean deviation around this mean level, b_t. If there has been no change to the Fed's policy reaction function, then $a_t = b_t = 0$, otherwise time variation in a_t and b_t represent policy shifts in the relative importance of output gap or inflation shocks in setting short-term interest rates.

We collect the macro and policy variables in the state vector $X_t = [g_t \ \pi_t \ a_t \ b_t]^\top$, which follows the stationary VAR:

$$X_t = \mu + \Phi X_{t-1} + \Sigma\varepsilon_t, \tag{2}$$

where $\varepsilon_t \sim$ IID $N(0, I)$. We order the macro variables first in the VAR. We set Σ to be the Cholesky decomposition of $\Sigma\Sigma^\top$ and so allow all factor shocks to be correlated. We parameterize Φ as

$$\Phi = \begin{pmatrix} \Phi_{gg} & \Phi_{g\pi} & \Phi_{ga} & 0 \\ \Phi_{\pi g} & \Phi_{\pi\pi} & 0 & \Phi_{\pi b} \\ \Phi_{ag} & 0 & \Phi_{aa} & 0 \\ 0 & \Phi_{b\pi} & 0 & \Phi_{bb} \end{pmatrix}. \tag{3}$$

Without the time-varying policy coefficients a_t and b_t, the upper 2×2 matrix of Φ represents a regular VAR of output and inflation. The coefficients Φ_{ga} and $\Phi_{\pi b}$ allow the policy coefficients to influence the future path of output and inflation. A negative coefficient Φ_{ga} means that a more aggressive response to the output gap would reduce the output gap next period. Similarly, if $\Phi_{\pi b} < 0$, then future inflation reduces as the Fed tightens monetary policy.[4]

[4] We do not allow the Fed's response to inflation to influence the future output gap or the Fed's output gap sensitivity to influence future inflation ($\Phi_{gb} = \Phi_{\pi a} = 0$). Estimates with non-zero Φ_{gb} and $\Phi_{\pi a}$ are hard to identify and resulted in VAR estimates that were non-stationary.

We treat the policy variables a_t and b_t as latent factors and are especially interested in their variation through the sample. In systems with latent factors, the same reduced-form model may often be produced by arbitrarily scaling or shifting the coefficients governing the dynamics of a_t and b_t in Φ or Σ. To identify a_t and b_t, we allow their shocks to be correlated in Σ, but do not allow any feedback between a_t and b_t in Φ. We allow the macro variables to potentially influence future policy stances. If Φ_{ag} or $\Phi_{\pi b}$ are positive, then the Fed responds to an environment with an increasing output gap or inflation by raising the policy responses to the output gap or inflation. Similarly, we also allow the policy variables to potentially influence the future path of output and inflation by specifying Φ_{ga} and $\Phi_{\pi b}$ to be non-zero.

The time-varying policy rule in equation (1) can be written in the form of a standard time-invariant Taylor (1993) rule with a "policy shock," η_t, that depends explicitly on the level of the output gap and inflation, combined with a time-varying policy stance. For ease of exposition, we ignore the orthogonal error term that is included in estimating equation (1) and write equation (1) as:[5]

$$\begin{aligned} r_t &= \delta_0 + \bar{a} g_t + \bar{b} \pi_t + (a_t g_t + b_t \pi_t) \\ &= \delta_0 + \bar{a} g_t + \bar{b} \pi_t + \eta_t, \end{aligned} \tag{4}$$

where $\eta_t = (a_t g_t + b_t \pi_t)$. In the special case that a_t and b_t are uncorrelated with the output gap and inflation (or $\Phi_{ga} = \Phi_{\pi b} = \Phi_{ag} = \Phi_{b\pi} = 0$ in equation (3)), then the average policy responses $\bar{a}$ and $\bar{b}$ in equation (4) can be consistently estimated by OLS. However, if a_t or b_t are correlated with the output gap or inflation, then conventional estimates of a linear policy rule like equation (4) will produce biased estimates of the policy responses to macro variable shocks.

If policy shifts do occur over time in the form of equation (1), then we can interpret the traditional policy shock, η_t, in the linear Taylor setting (4) as comprising two components: policy reaction components a_t and b_t, and macro variable components, g_t and π_t. The residual η_t would

[5] We also investigated an alternative policy shift model, where the short rate took the form:

$$\hat{r}_t = \delta_0 + (\bar{a} + a_t) g_t + (\bar{b} + b_t) \pi_t + f_t,$$

where f_t was an additional IID latent factor. This latent factor differs the measurement error u_t^1 put on the short rate in the estimation because the observation errors are yield specific, whereas the latent factor f_t is also priced by all other yields. This model resulted in extremely small estimates of f_t with almost zero improvement in model fit (observation error standard deviations). A Bayes factor test also makes this model extremely unlikely compared to the benchmark quadratic model.

also exhibit conditional heteroskedasticity. We can decompose a traditional linear policy shock into policy shifts by the Fed (a_t and b_t terms) and separate shocks to output gap and inflation components ($a_t g_t$ and $b_t \pi_t$, respectively). Previous research in affine models have found that a linear latent factor, η_t, is related to movements in macro variables and can represent a monetary policy shock (see, among others, Ang, Dong and Piazzesi, 2006; Bikbov and Chernov, 2006). In our set up, we can quantify the variation in short rates directly emanating from policy shifts versus shocks to macro variables.

We assume that the time variation in the policy coefficients is a covariance stationary process, that is all the eigenvalues of Φ lie inside the unit circle. This is in contrast to previous approaches which model time variation in policy parameters using random walks (see, among others, Cooley and Prescott, 1976; Cogley and Sargent, 2001, 2005; Cogley, 2005; Boivin, 2006; Justiniano and Primiceri, 2006). While the random walk is a convenient framework to account for permanent changes in coefficients, inferring how the term structure reacts to policy shifts is better done with a stationary process for several reasons. First, since yields are intertemporal marginal rates of substitution, they should be stationary in well-defined exchange models with representative agents having utility over consumption. Second, random walk models cannot be used to attribute the variance of long-term yields to policy shift components and shocks to macro factors, as the unconditional variance is infinite in a random walk process. Similarly, since there is no well-defined long-run mean in a random walk system, it is hard to define the long-run effects of policy shifts on yields or macro factors.

The time-varying Taylor rule (1) is an example of a regression model with stochastically varying coefficients. Using only macro data and short rates, the system is asymptotically identified (see Pagan, 1980). However, it is hard to use only one observable variable, short rates, to identify two latent processes in small samples. Fortunately, it is not only the short rate that responds to policy shifts – we identify the variation in a_t and b_t by using information from the entire yield curve. A further advantage of using the entire term structure is that we can identify the prices of risk that agents assign to the policy authority's time-varying policy rules. Thus, we can infer the effect on long-term yields of a policy shift by the Fed on its inflation stance, as well as the traditional analysis of tracing through the effect of an inflation shock on the term structure. In contrast, the previous literature on macro-finance term structure models assumes that the policy coefficients are time invariant.

2.1 Bond Prices

To derive bond prices from the policy shift model of equation (1), we write the short rate as a quadratic function of the factors $X_t = [g_t \ \pi_t \ a_t \ b_t]^\top$:

$$\hat{r}_t = \delta_0 + \delta_1^\top X_t + X_t^\top \Omega X_t \tag{5}$$

where δ_0 is a scalar and $\delta_1 = [\bar{a} \ \bar{b} \ 0 \ 0]^\top$. We use the carrot notation for yields which are direct functions of the model to distinguish the model-implied short rate, $\hat{r}_t$, from the short rate in data, r_t. In the quadratic term $X_t^\top \Omega X_t$ in equation (5), Ω is specified as

$$\Omega = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \end{pmatrix}. \tag{6}$$

The short rate is linear in the observable macro variables and the quadratic form results from the interaction of the stochastic policy coefficients with the macro factors. If there are no policy shifts, then $\Omega = 0$, and the model simplifies to a standard affine term structure model.

To price long-term bonds, we specify the pricing kernel to take the standard form:

$$m_{t+1} = \exp\left(-\hat{r}_t - \frac{1}{2}\lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1}\right), \tag{7}$$

with the time-varying prices of risk:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{8}$$

for the 4×1 vector λ_0 and the 4×4 matrix λ_1. The prices of risk control the response of long-term yields to macro and policy shocks, and cause the expected holding period returns of long-term bonds to vary over time (see Dai and Singleton, 2002). Of particular interest are the risk premia parameters on the policy shift variables a_t and b_t. These have not been examined before because the prices of risk in equation (8) have almost exclusively been employed in traditional affine macro-term structure models where the policy coefficients are constant (see, for example, Ang and Piazzesi, 2003).

The pricing kernel prices zero-coupon bonds from the recursive relation:

$$\hat{P}_t^n = \mathrm{E}_t[m_{t+1} \hat{P}_{t+1}^{n-1}],$$

where $\hat{P}_t^n$ is the price of a zero-coupon bond of maturity n quarters at time t. Equivalently, we can solve the price of a zero-coupon bond as:

$$\hat{P}_t^n = \mathrm{E}_t^{\mathbb{Q}} \left[\exp\left(-\sum_{i=0}^{n-1} \hat{r}_{t+i} \right) \right], \tag{9}$$

where $\mathrm{E}_t^{\mathbb{Q}}$ denotes the expectation under the risk-neutral probability measure $\mathbb{Q}$, under which the dynamics of the state vector X_t are characterized by the risk-neutral constant and autocorrelation matrix:

$$\begin{aligned} \mu^Q &= \mu - \Sigma\lambda_0 \\ \Phi^Q &= \Phi - \Sigma\lambda_1, \end{aligned} \tag{10}$$

where X_t follows the process

$$X_t = \mu^Q + \Phi^Q X_{t-1} + \Sigma\varepsilon_t$$

under $\mathbb{Q}$. In our estimation, we impose Φ^Q to take the same restrictions as the companion form under the real measure, Φ, given in equation (3).

The relevant dynamics for bond prices are given by the risk-neutral parameters μ^Q and Φ^Q. The actual dynamics of macro variables and policy risk factors in the real world may differ from the dynamics of the factors under the risk-neutral measure. These differences precisely highlight how macro and monetary policy changes are priced by agents. For example, agents may have different perceptions of how monetary policy changes impact the economy than what actually happens in reality. Similarly, the yield curve may reflect beliefs on how past economic growth and inflation influence future monetary policy actions that is different to how the Fed actually conducts policy in the real world. We will explicitly contrast the different implied dynamics of monetary policy factors under the risk-neutral and real measures.

The quadratic short rate (1) or (5), combined with the linear VAR in equation (2), and the pricing kernel (7) gives rise to a quadratic term structure model. We can write the bond price for maturity n implied by the model as:

$$\hat{P}_t^n = \exp(A_n + B_n^\top X_t + X_t^\top C_n X_t), \tag{11}$$

where the terms A_n, B_n, and C_n are given in Appendix A. Hence, if we denote the yield on a zero-coupon bond with maturity n quarters as $\hat{y}_t^n = -1/n \log \hat{P}_t^n$, yields are quadratic functions of X_t:

$$\hat{y}_t^n = a_n + b_n^\top X_t + X_t^\top c_n X_t, \tag{12}$$

where $a_n = -A_n/n$, $b_n = -B_n/n$, and $c_n = -C_n/n$. This analytical form enables the estimation of the model and allows us to investigate how the entire term structure responds to policy changes and macro shocks.

Since the yields are quadratic functions of the state variables, the model belongs to the class of quadratic term structure models developed by Longstaff (1989), Beaglehole and Tenney (1992), Constantinides (1992), Leippold and Wu (2002, 2003), and Ahn, Dittmar and Gallant (2002).[6] None of these authors incorporate observable macro factors or investigate policy shifts. Ahn, Dittmar and Gallant (2002) and Brandt and Chapman (2003) demonstrate that quadratic models have several advantages over the affine class in adding more flexibility to better match yield dynamics, particularly conditional moments. The non-linearity of yields also aids in identifying prices of risk because there is an additional source of identification, through the non-linear mapping of state variables to yields, that is absent in an affine setting.

To estimate the model, we assume that all yields, including the short rate, are measured with error. Specifically, we assume:

$$y_t^n = \hat{y}_t^n + u_t^n, \tag{13}$$

where $\hat{y}_t^n$ is the model-implied yield in equation (12), y_t^n is the yield observed in data, and u_t^n IID $\sim N(0, \sigma_n^2)$, are additive measurement errors for all yields n. The quadratic form of the yields implies that there is not a one-to-one correspondence between certain yields assumed to be observed without error and latent state variables. Thus, standard filtering techniques for estimating affine models cannot be used to estimate our quadratic term structure model. We employ a Bayesian filtering algorithm that requires no approximation to estimate the model, which we detail in Appendix B.

3 Empirical Results

In Section 3.1, we describe the construction of the output gap and inflation and how the model matches macro variables and yields in data. Section 3.2 discusses the parameter estimates. Section 3.3 documents how the Fed reaction to output gap and inflation shocks have changed

[6] These quadratic models are members of the broader affine class of term structure models of Duffie and Kan (1996) as they have linear representations of yields involving factors X_t and second moments of factors, $vech(X_t X_t')$. The quadratic term itself follows an affine process, as shown by Filipovic and Teichmann (2002) and Gourieroux and Sufana (2003). Buraschi, Cieslak and Trojani (2007) show that the quadratic short rate process can be supported in a Cox, Ingersoll and Ross (1985) production economy with a representative agent.

over time. In Section 3.4, we discuss how the yield curve reacts to changes in the Fed's policy parameters. Section 3.5 explores the price of risk of Fed policy shifts.

3.1 Data

All our data is at a quarterly frequency and the sample period is from June 1952 to December 2006. The output gap is constructed following Rudebusch and Svensson (2002) and is given by

$$g_t = \frac{1}{4}\frac{Q_t - Q_t^*}{Q_t^*}, \tag{14}$$

where Q_t is real GDP and Q_t^* is potential GDP. We obtain real GDP from the Bureau of Economic Analysis (BEA), which is produced using chained 2000 dollars. We use the measure of potential output published by the Congressional Budget Office (CBO) in the Budget and Economic Outlook using chained 1996 dollars. To make the BEA series comparable to the CBO series, we translate real GDP to 1996 dollars. Finally, we demean the output gap and divide the output gap by four to correspond to quarterly units. Since we will be using per quarter short rates, this allows us to read the coefficient on the output gap as an annualized number. Our series for inflation is the year-on-year GDP deflator expressed as a continuously compounded growth rate. This is also divided by four to be in per quarter units. In addition to the one-quarter short rate, our term structure of interest rates comprises take zero-coupon bond yields from CRSP of maturities 4, 8, 12, 16, and 20 quarters. These are all expressed as continuously compounded yields per quarter.

Figure 1 plots the output gap, inflation, and the short rate over our sample in annualized terms. The output gap decreases during all the NBER recessions and reaches a low of -7.1% during the 1981:Q3 to 1983:Q4 recession. The output gap strongly trends upwards during the expansions of the 1960's, the mid-1980's, and the 1990's. Inflation is slightly negatively correlated with the output gap at -0.245. Inflation rises to near 10% during the mid-1970's and early 1980's, but otherwise remains below 5%. In the data, the correlation between the output gap and the short rate is -0.155 and the correlation between inflation and the short rate is 0.698. These correlations are matched closely by the model, with implied correlations of g_t and π_t with the short rate of -0.128 and 0.741, respectively.

As a benchmark, we report OLS estimates of simple Taylor (1993) rules where the short rate is a linear combination of macro factors and lagged inflation:

$$\begin{array}{ccccccccc} r_t & = & 0.001 & + & 0.020\ g_t & + & 0.904\ \pi_t & + & \varepsilon_t, \\ & & (0.001) & & (0.060) & & (0.064) & & \end{array} \tag{15}$$

where standard errors are reported in parentheses. Adding lagged short rates we obtain:

$$\begin{array}{ccccccccccc} r_t & = & 0.001 & + & 0.071\ g_t & + & 0.141\ \pi_t & + & 0.873\ r_{t-1} & + & \varepsilon_t, \\ & & (0.001) & & (0.028) & & (0.041) & & (0.031) & & \end{array} \tag{16}$$

which can be written in partial adjustment format as:

$$r_t = 0.001 + 0.873\, r_{t-1} + (1 - 0.873)(0.560\, g_t + 1.105\, \pi_t) + \varepsilon_t.$$

These estimates are very similar to those reported in the literature. In our model, the coefficients on g_t and π_t from these simple OLS estimates do not correspond to the policy coefficients by the monetary authority on the output gap and inflation. While the short rate equation (4) also specifies the short rate as a linear combination of g_t and π_t, the OLS shock ε_t is not orthogonal to the macro factors.

In Table 1, we report summary statistics of the factors in data and implied by the estimated model. The factors and yields are expressed in percentage terms at a quarterly frequency. The model provides an excellent match to the data, with model-implied unconditional means and standard deviations very close to the moments in data. In Panel A, the unconditional moments of the output gap and inflation implied by the model are well within 95% confidence bounds of the data estimates. Panel B of Table 1 compares the yields in data with the model-implied yields. All yields are expressed in percentage terms per quarter. We construct the posterior moments of the model-implied yields by using the generated latent factors in each iteration from the Gibbs sampler estimation. The tight posterior standard deviations indicate that the draws of the latent a_t and b_t factors in the estimation result in yields that very closely lie around the data yields. All of the model-implied estimates are almost identical to the data. Note that because we match the short rate exactly in the estimation, the mean of the short rate aligns exactly with the data mean by construction.

3.2 Parameter Estimates

We report the estimates of the model parameters in Table 2. We report posterior means, with posterior standard deviations in parentheses, of the model parameters. The first panel of Table 2 reports the long-run responses of the Fed to the output gap and inflation. Unconditionally, the long-run response to the output gap is small, at 0.223, with a posterior standard deviation of 0.045. The posterior mean of the long-run response to inflation is well above one, at 1.442, with a posterior standard deviation of 0.100. These are larger than the simple Taylor rule estimates of

0.020 and 0.904 in equation (15) suggesting that the time variation of a_t and b_t play an important role in determining the short rate and OLS estimates contain some bias.

In the conditional volatility matrix, the conditional shocks of all factors are lowly correlated with each other. The Fed sensitivity to output gap shocks has small variation, with the conditional volatility of a_t only 0.005×10^{-3}, whereas the Fed's response to inflation exhibits much larger variation, with the conditional volatility of b_t 38.65×10^{-3}. The posterior standard deviations of a_t and b_t across all sample paths in the estimation are 0.008 and 0.587, respectively, so the Fed's stance to output gap shocks is very stable while b_t has changed substantially over time. Below, we further investigate the time-series variation of a_t and b_t.

Not surprisingly, Table 2 shows that in the companion form Φ, all the factors are highly autocorrelated, with the coefficients of Φ lying very near one along the diagonal. High inflation Granger-causes low economic activity next quarter ($\Phi_{g\pi} = -0.098$), but this effect is statistically weak with a posterior standard deviation of 0.058. On the other hand, high economic growth today suggests that next-period inflation will accelerate ($\Phi_{\pi g} = 0.064$) and this effect is statistically stronger with a posterior standard deviation of 0.011. These effects have been noted before in standard VAR macro models like Christiano, Eichenbaum and Evans (1996, 1999).

Table 2 shows that the feedback coefficients between $(g_t\, \pi_t)$ and $(a_t\, b_t)$ are estimated with considerable error and 95% confidence intervals of the posterior estimates encompass zero. The parameter $\Phi_{ga} = -0.008$ is negative indicating that changes in the Fed's response to output gap shocks induce a small, but insignificant, influence over the path of next-period future economic activity. The coefficient $\Phi_{b\pi}$ is also estimated to be effectively zero. The effect of past output gaps on a_t is small with $\Phi_{ag} = 0.004$, whereas past high inflation Granger-causes monetary policy to tighten and become more sensitive to inflation shocks with $\Phi_{b\pi} = 1.183$. However, while this coefficient is large in magnitude, it is estimated with considerable error with a posterior standard deviation of 2.518.

The lack of evidence that past macro factors influence future monetary policy movements and vice versa implies that the specification used by previous studies estimating drifting policy coefficients, such as Cogley (2005) and Boivin (2006), among others, may be econometrically sufficient. These studies use independent random walk specifications which are not affected by past macro shocks to model changing monetary policy. However, these studies did not consider the pricing of changing monetary policy risk by the yield curve, which our model captures by the price of risk parameters λ_0 and λ_1.[7]

[7] The zero entry in the λ_1 matrix results from the companion form Φ taking the form of equation (3) under both

Table 2 also reports the risk-neutral companion form Φ^Q which differs from Φ in two important ways. Under the risk-neutral measure, macro shocks are noticeably less persistent than their real measure counterparts ($\Phi^Q_{gg} = 0.637 < \Phi_{gg} = 0.913$ and $\Phi^Q_{\pi\pi} = 0.838 < \Phi_{\pi\pi} = 0.988$ while the persistence of a_t and b_t is largely unchanged. Thus monetary policy changes have more persistent effects on bond prices than macro factors. Second, there is significant evidence that the yield curve prices in Granger-causality of past macro variables affecting monetary policy stances and that monetary policy endogenously responds to past g_t and π_t levels.

Under $\mathbb{Q}$, the market perceives the Fed to significantly affect the future path of the economy, with $\Phi^Q_{ga} = -0.594$ and $\Phi^Q_{\pi b} = -0.001$. This implies that the market prices such that a higher output sensitivity decreases future output and a more aggressive stance towards inflation decreases future inflation. The negative coefficient $\Phi^Q_{ag} = -0.007$ indicates that the yield curve prices in a counter-cyclical response to output gap shocks. When g_t is high during expansions, a_t is small and the short rate response to output gap shocks is small. When g_t is negative during recessions, a_t is large and the Fed moves to reduce the short rate aggressively to bad output shocks more than if the same shocks occur during expansions when g_t is high. These feedback effects are much stronger in Φ^Q than what actually occur in Φ.

Of all the parameters in Φ^Q, only $\Phi^Q_{b\pi} = -0.099$ has a posterior 95% confidence bound that includes zero. Interestingly, this coefficient has the opposite sign to $\Phi_{b\pi}$, but both $\Phi^Q_{b\pi}$ and $\Phi_{b\pi}$ are insignificantly different from zero. Thus, also under the risk-neutral measure, there is little evidence that the Fed's monetary policy stance with respect to inflation responds significantly to past inflation shocks.

3.3 Policy Shifts in Output and Inflation Responses

Figure 2 displays the policy parameters, $\bar{a}+a_t$ and $\bar{b}+b_t$, over the sample. We plot the mean posterior estimates at each point in time of the Fed's response to output and inflation produced by the Gibbs sampler, along with two posterior standard deviations. These estimates lend support to the conjecture that the changes in monetary policy during this period were substantial.

The Fed's response to output gap shocks is centered around $\bar{a} = 0.223$ and is generally above 0.20 through the sample. From this low, the loading on the output gap reaches a low of 0.199 in 1981:Q3. The output gap loading rose in the early post-Volcker era back to around 0.22 and has has been fairly stable, with the exception of the most recent 2001 recession, which saw

the risk neutral and the real measure. The risk prices are inferred from equation (8). See the Appendix for further details.

a decrease in the output gap sensitivity. The most notable feature of the output gap response is that its range is relatively narrow, with a minimum of 0.199 and a maximum of 0.234. Thus, the Fed has exhibited little change in its responsiveness to economic growth.

The bottom plot of Figure 2 graphs the Fed's response to inflation. In contrast to the severe smoothness of the Fed's sensitivity to output gap shocks, the Fed's response to inflation has changed markedly over time. The Fed's loading on inflation takes on a minimum of 0.297 in 2003:Q3 and a maximum of 2.981 in 2008:Q1. Overall, the time-series pattern of the inflation coefficient is roughly consistent with the evidence reported in Clarida, Galì and Gertler (2000), Cogley and Sargent (2005), and Boivin (2006).

The response to inflation during the 1950's was below one sharply increasing to well above one during the late 1950's. In 1959:Q3 $\bar{b}+b_t$ reached a temporary high of 2.595. From this high, the Fed's inflation coefficient started to decrease during the 1960's but generally remained above one during this time. The response to inflation was generally below one throughout the 1970's. An appealing feature of these estimates is that, consistent with the narrative evidence (see, for example, Meltzer, 2005), it clearly shows that the response to inflation started to increase in 1979. Interestingly, and as in Boivin (2006), the sharpest increase in the inflation response was not in late 1979, as is often assumed because of the appointment of Volcker in July 1979, but after 1981. The rapid increase in $\bar{b}+b_t$ over the early 1980's reached a temporary high of 2.800 in 1984:Q2.

The estimated increase in b_t from the 1970's to the 1980's is sizeable. The inflation loading starts from a level less than one in the 1970's, where the Taylor principle is not satisfied. That means that during most of the 1970's a unit increase in inflation translated into a less than unit increase in the nominal policy rate, which represents a decline in the real rate, and hence implies an easing of monetary policy. Whenever the Taylor principle is not satisfied it is possible for inflation expectations, and thus economic fluctuations, to be driven by non-fundamental sunspot shocks. A failure to rule out the presense of such shocks could thus have been responsible for the greater economic volatlity of the 1970's (see the discussion by Taylor, 1999; and Clarida, Galí and Gertler, 2000). The importance and the direction of these shifts are overall consistent with the view that the conduct of monetary policy was not stable during the 1970's and has evolved under Volcker toward a more stabilizing conduct. Moreover, the timing of these changes are broadly consistent with the general decline in the volatility of the US economy, suggesting that monetary policy could have been in part responsible for the Great Moderation post-1985 (see comments by Stock and Watson, 2003).

Recently, the response to inflation dipped well below one during the 2001 recession and the aftermath of the September 2001 terrorism acts, where the short rate declined from 4.25% in 2001:Q1 to 0.90% in 2003:Q5. The Fed's response to inflation shocks sank below one in 2001:Q3 to 0.686 reaching a low of 0.297 in 2003:Q3. From this low to the end of our sample in 2006, the Fed response to inflation increases sharply, rising above one in 2005:Q2 and ending at 1.613 in 2006:Q4. Thus, we find that the last few years of monetary policy under Greenspan also did not satisfy the Taylor principle.

It is an interesting question to see what the yield curve would have looked like had the Fed not changed its inflation loading over the post-2001 period. Inflation during this time was low, possibly even below an implicit target (see Figure 1), so interest rates may have declined over this period even with unchanged policy coefficients. In Figure 3 we report the results of a counter-factual experiment where we hold the Fed weight on inflation at the average weight of $\bar{b} + b_t$ over 2000 and trace the effects on the yields post-2001. We allow the other factors to be take their sample values. Figure 3 plots the path of the short rate and term spread if the Fed had maintained the same inflation stance as in 2000 in the dashed lines and overlays the actual short rate and term spread in the solid lines.

The top panel of Figure 3 shows that had the Fed maintained the same inflation stance in 2000, short rates would indeed have been considerably higher post-2001 than in data. With the same inflation tolerance in 2000, the short rate in 2003:Q4 would have been well over 4% compared to 0.92% in data. In the bottom panel of Figure 3 we plot the term spread. Interestingly, we find that there is little difference in the slope of the yield curve over 2001-2004 comparing actual data and the counter-factual exercise where the Fed did not take a more dovish stance.

3.3.1 Comparisons with Estimations Using No Term Structure Information

Our model uses the entire term structure to identify the time series of output and inflation policy responses. We now demonstrate that this leads to sharper estimates of the Taylor rule coefficients than models that omit term structure information. Intuitively this is because no-arbitrage restrictions, through the bond prices in equation (11), link policy actions on the short rate together with movements in long-term bonds. The omission of term structure information not only increases, sometimes substantially, the estimation error of the policy responses – it also results in estimates of policy paths that are different from the full model which incorporates long-term bond information.

The top panel of Figure 4 plots the output gap coefficient from our full estimation and is the same figure as the top panel of Figure 2, except on a different scale. Clearly, the output gap coefficient does not vary much around $\bar{a} = 0.223$ and is precisely estimated. We compare these estimates with the output gap response from two other models in the two lower panels of Figure 4.[8] In the middle panel, we draw the latent factors a_t and b_t from the same model as equation (2) except only information on g_t, π_t, and r_t is used, that is no term structure information is employed. In this exercise, we hold all the parameters of the VAR constant at their posterior mean estimates in Table 2 to isolate the effect of the term structure information on the latent factor distributions. This is a conservative exercise because the additional sampling error from the VAR and price of risk parameters increases the posterior distributions of a_t and b_t in the benchmark estimation relative to the model in the middle panel of Figure 4.

Comparing the middle panel of Figure 4 to the first panel, we observe that the standard error bands of a_t with the term structure information are an order of magnitude smaller than the standard error bands of a_t omitting the information from the yield curve. While the path of a_t in the full model (top panel) exhibits small variation, it does trend downwards to reach a low in 1981:Q3 and then rises gradually towards the end of the sample. In comparison, the best estimates of a_t without the yield curve information (middle panel) stay very evenly around 0.223, indicating that the yield curve does provide auxiliary information to the short rate alone.

In the last panel of Figure 4, we estimate a model close to the models in the literature where a_t and b_t follow a random walk and are orthogonal to the macro variables. In this case, we change the companion form, Φ, and the square-root of the conditional covariance matrix, Σ, to:

$$\Phi = \begin{pmatrix} \Phi_{gg} & \Phi_{g\pi} & 0 & 0 \\ \Phi_{\pi g} & \Phi_{\pi\pi} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{gg} & 0 & 0 & 0 \\ \Sigma_{\pi g} & \Sigma_{\pi\pi} & 0 & 0 \\ 0 & 0 & \Sigma_{aa} & 0 \\ 0 & 0 & \Sigma_{ab} & \Sigma_{bb} \end{pmatrix}. \tag{17}$$

To be comparable to the literature, we estimate all the parameters of this model using only the time series of r_t, g_t, and π_t.

The bottom graph of Figure 4 shows that the posterior standard deviation bands of a_t are also very large for the random walk model compared to the benchmark model in the top panel. The unit root estimates of a_t also exhibit much larger time variation. In the benchmark model, the

[8] Technically, these models can be estimated using a methodology similar to the full model described in the Appendix, except that no accept/reject draw for the latent factors a_t and b_t is required. There are also no accept/reject draws needed for the term structure likelihood for the VAR or short rate parameters.

best estimates of a_t have a standard deviation of 0.008. In comparison, the standard deviation of the posterior mean of a_t from the unit-root model is 0.204. The unit-root model's estimates of a_t are lowly correlated, at 0.207, with the benchmark model's estimates. Thus, the inference of the Fed's sensitivity to output growth is very much affected by whether a unit-root specification for the policy coefficients and long-term bond information are used.

In Figure 5, we plot the inflation policy responses from the full model (top panel), the full model estimated without yield curve information (middle panel), and the unit root model (bottom panel). (The top panel repeats the same curve as the bottom graph of Figure 2 except on a different scale.) Not surprisingly, the posterior standard error bands increase moving from the benchmark model in the top panel to the middle panel that omits term structure information. In the bottom panel, the standard error bands of the unit-root model are similar in magnitude to the the benchmark model estimates without the yield curve in the middle panel.

In contrast to the output gap responses, the time series of $\bar{b} + b_t$ from the full model and the unit-root model are highly correlated at 0.829. The unit-root specification also allows us to reject the hypothesis that monetary policy satisfied the Taylor principle in the 1970's and was also substantially less than one in the early 2000's. While the general comovement of the inflation response is similar across the full and unit-root models, there are important differences. First, the general level of the Fed's inflation sensitivity coefficient is generally lower for the unit-root model estimated without term structure information. In the full model, the unconditional inflation response $\bar{b} = 1.442$ compared to $\bar{b} = 0.857$ in the unit root model. Second, the range of the inflation responses is higher in the benchmark specification. The lowest (highest) inflation response is 0.304 (2.9794) in the benchmark model compared to -0.988 (2.082) in the last panel. This implies that using term structure information, the estimate of the inflation response is more aggressive, on average, and that the Fed exhibits a more flexible, active response compared to a unit-root estimate of inflation policy that ignores the yield curve.

3.3.2 Factor Impulse Responses

Figure 6 reports the impulse responses of the factors on each other. In the first two rows, we consider standard VAR responses of a 1% shock from g_t or π_t on each other. These results are consistent with many other studies. A 1% shock to output causes future inflation to increase reaching approximately 40 basis points after 15 quarters. The effect of an inflation shock to output is stronger, with a 1% shock to inflation causing future output to contract approximately 70 basis points after 12 quarters.

In the bottom two rows of Figure 6, we plot the impulse responses of b_t from shocks to macro factors and the effect of shocks to $b - t$ to the macro factors. We denote the responses in the real measure by the red solid line and we overlay the responses in the real measure in dashed green lines. In the third row, we consider the effect of a 1% shock to g_t and π_t on b_t. In the last row, we shock b_t by 1.00 and trace the effects on g_t and π_t. We focus on the effect of b_t because there is little variation in the response of the Fed to output fluctuations.

The third line of Figure 6 shows that in the real world, a 1% shock to g_t and π_t causes the Fed to tighten monetary policy. The effect is quantitatively small, with a 1% shock to g_t causing an increase in b_t of less than 0.05 and a 1% shock to π_t raising b_t by less than 0.1. Under the risk-neutral measure, the responses implied by the market prices of bonds show that b_t are even more unresponsive with b_t barely budging to g_t or π_t shocks. While the actual policy responses suggest that the Fed modestly tightens monetary policy in economic expansions (when g_t and π_t shocks are positive), economic participants price bonds believing that there is little automatic endogenous response of monetary policy to past macro shocks and agents believe that monetary policy is more discretionary.

In the last row of Figure 6, we plot the effect of a unit move in b_t on the macro factors g_t and π_t. The responses of the macro variables under the real measure reveal little reaction of g_t and π_t to changes in the Fed response to inflation. In contrast, under $\mathbb{Q}$, agents believe that the macro environment does respond modestly to Fed policy changes. In particular, raising b_t by one unit produces an increase in g_t of 20 basis points and a decrease in π_t of 50 basis points after 10 quarters. This response of inflation under the risk-neutral measure is economically large because the variation in b_t has exhibited wide range over the sample. (For example, Figure 2 shows that the change in b_t from the mid-1970's to the early 1980's is around 3.5.) Hence, the yield curve prices in a belief that the Fed can engineer successful disinflation by increasing the policy loading on inflation shocks. However, there is little evidence this actually occurs in the real world.

3.4 How Policy Shifts Affect the Yield Curve

3.4.1 Short Rate Components

In our model, short rates move due to movements in the output gap component, $(\bar{a} + a_t)g_t$, and the inflation component, $(\bar{b} + b_t)\pi_t$. Figure 7 highlights these two components of the short rate. The policy factors are evaluated at the best estimates of a_t and b_t through the sample, together

with the short rate in per quarter units. The correlation between the actual short rate and the fitted components is 0.985, indicating that movements in the macro variables and policy rule account for almost all of the variation in the short rate. The bottom panel of Figure 7 shows that most of the variation in the short rate comes from inflation components. Although the variation in the output gap and inflation are similar (see Table 1), the small policy response on output shocks and the relatively large response on inflation cause the inflation component to dominate in the short rate variance.

3.4.2 Impulse Responses of the Term Structure

In Figure 8 we plot the response of the yield curve to macro shocks and inflation policy shifts. Since the yields are non-linear functions of macro and policy variables, we compute the impulse responses numerically, which we detail in Appendix C. We graph in columns the response of a 1% shock to g_t, a 1% shock to π_t, and a 1.00 shock to b_t on the short rate, r_t, the 20-quarter long rate, y_t^{20}, and the yield spread, $y_t^{20} - r_t$, which are presented in rows.

The first column of Figure 8 shows that positive output shocks increase short rates and decrease spreads. A 1% shock to g_t increases the short rate by 65 basis points after 15 quarters. The same shock causes the long rate to also increase, initially to 10 basis points, reaching a peak of 15 basis points around 10 quarters. Thus, the lower left-hand graph shows that after a 1% shock to g_t the term spread initially shrinks by approximately 15 basis points reaching -50 basis points by around 15 quarters. The second column of Figure 8 shows similar patterns for a 1% shock to π_t. Initially, the short rate increases by 1.4% and the yield spread shrinks by approximately 1%. The impulse response of the yield spread falls below 25 basis points after 20 quarters. These results are similar to those reported by Ang and Piazzesi (2003), among many others, who show that the macro shocks influence mostly the short end of the yield curve, which responds more to macro shocks than the long end of the yield curve.

In the third column we plot the response of a 1.00 change in b_t. We focus on b_t because the movements in a_t over the sample have been quantitatively very small. A 1.00 change in b_t causes the short rate to initially increase by 80 basis points, which slowly dies down to fall below 25 basis points after 15 quarters. The 1.00 shock to b_t causes the long rate to barely move and the response by y_t^{20} remains below 10 basis points over all horizons. Thus, the term spread initially decreases by 75 basis points when b_t increases by 1.00. Thus, changing monetary policy stances also mostly impact the short end of the yield curve.

In the two lower figures in the third column, we overlay the responses of the long rate and

term spread under the risk-neutral measure. These risk-neutral responses are the responses of yields under the Expectations Hypothesis (EH) and are produced by setting the time-varying prices of risk $\lambda_1 = 0$.[9] Clearly, there are differences in the response of yields under risk neutrality compared to the case with risk premia. Under the EH, long rates would move much more than under the real measure, with an initial response of 40 basis points to a 1.00 move in b_t, compared to the actual initial response of less than 10 basis points. Thus, under the EH, the yield spread only initially falls by 40 basis points compared to 75 basis points in the full model. That is, shifting monetary policy risk is priced so that long rates are more stable and move less when b_t moves. In our model, this is due to agents' beliefs under $\mathbb{Q}$ that Fed actions are stabilizing, which causes the yield curve not to amplify the effects of b_t shocks as much as the case where policy risk is not priced.

3.5 The Fed Policy Risk Premium

The different response of the yield curve in Figure 8 under the case where all factor risk is priced compared to the case under the EH indicate that investors value the role of active monetary policy. In this section, we characterize the risk premium assigned by investors to the risk of policy shifts by the Fed in several ways.

3.5.1 Interpreting Price of Risk Parameters

To directly interpret the λ_0 and λ_1 price of risk coefficients, consider first a standard CRRA representative agent economy with the pricing kernel

$$m_{t+1} = \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma} = \exp(-\gamma(\mu_c + \sigma_c \varepsilon^c_{t+1})),$$

where C_t is aggregate consumption, γ is the representative agent's risk aversion, μ_c and σ_c are the mean and volatility of log consumption growth, respectively, and $\varepsilon^c_{t+1} \sim N(0,1)$ is the shock to consumption growth. In this economy, the price of a security with the same payoff as the unit consumption shock is given by:

$$\begin{aligned} P_t &= \mathrm{E}_t[m_{t+1}\varepsilon^c_{t+1}] = \mathrm{E}_t[e^{-\gamma(\mu_c+\sigma_c\varepsilon^c_{t+1})}\varepsilon^c_{t+1}] \\ &= -e^{-r_t}\gamma, \end{aligned} \tag{18}$$

[9] In an affine model the term spread would not move under the EH, except for some small Jensen's inequality terms.

where the risk-free rate $r_t = \gamma\mu_c - \frac{1}{2}\gamma^2\sigma_c^2$. Equation (18) shows that if the agent is risk averse, $\gamma > 0$, then the price of the consumption shock is less than the risk-neutral price of zero, $P_t < 0$. Thus, risk-averse agents must be paid to bear consumption risk and the price of a unit risk is equal to aggregate risk aversion multiplied by a bond.

In the term structure model there is no direct correspondence to representative risk aversion because there are multiple shocks, the prices of risk vary over time, and the prices of risk of a_t and b_t also depend on the correlated movements of g_t and π_t as well as each other. Nevertheless, we can use the difference between the actual price and risk-neutral price of claims to the factor shocks to provide economic intuition for the policy shift risk priced by the yield curve. The prices of unit shock payoffs are given by:

$$\begin{aligned} \mathrm{E}_t[m_{t+1}\varepsilon_{t+1}] &= \mathrm{E}_t\left[\exp\left(-r_t - \frac{1}{2}\lambda_t^\top\lambda_t - \lambda_t^\top\varepsilon_{t+1}\right)\varepsilon_{t+1}\right] \\ &= e^{-r_t}\lambda_t = e^{-r_t}(\lambda_0 + \lambda_1 X_t) \end{aligned} \tag{19}$$

where we use the definition of the pricing kernel in equation (7) and the short rate $r_t = \delta_0 + \delta_1^\top X_t + X_t^\top \Omega X_t$ is also a function of X_t.

In Panel A of Figure 9, we plot the price of a unit shock to a_t and b_t as a function of the monetary policy coefficient loadings $(\bar{a} + a_t)$ and $(\bar{b} + b_t)$, respectively. We denote with vertical lines the steady-state values of $\bar{a} = 0.223$ and $\bar{b} = 1.442$. The figure shows that for values of $(\bar{a} + a_t) < 0.226$ the price of a unit a_t shock is negative. For these values monetary policy risk commands a risk premium. When $(\bar{a} + a_t) > 0.226$ the price of the claim on a unit a_t shock is positive – in this case agents are willing to pay for activist monetary policy. That is, when the Fed is very responsive to output gap shocks, agents value monetary policy shifts and monetary policy has a risk discount.

We now turn to the price of a unit shock on b_t. The second graph in Panel A of Figure 9 shows that the price of this claim is negative for values of $(\bar{b} + b_t) < 1.62$. Thus, when monetary policy is loose and unresponsive to inflation shocks, agents demand a monetary policy risk premium. When monetary policy responds sufficiently aggressively to inflation shocks, $(\bar{b} + b_t) > 1.62$, agents assign a risk discount to activist monetary policy and are willing to pay to be subject to this risk. Note that for both the long-run values of the monetary policy responses $\bar{a}$ and $\bar{b}$, monetary policy has a very slight risk premium, but it is close to zero.

3.5.2 The Price of Policy Shifts Under Full Risk and Risk-Neutral Specifications

An alternative way to characterize monetary policy risk is to compare the price of a security in the full model, where a_t and b_t are priced, to the price of a security with the same payoff in a model where a_t and b_t carry zero risk premia. Consider the price of a security paying off $(1 + X_{t+1})$:

$$\begin{aligned} P_t &= \mathrm{E}_t[m_{t+1}(1 + X_{t+1})] = \mathrm{E}_t^Q[e^{-r_t}(1 + X_{t+1})] \\ &= e^{-r_t}(1 + \mu^Q + \Phi^Q X_t) \end{aligned} \tag{20}$$

The risk-neutral parameters μ^Q and Φ^Q include risk prices arising from both macro factors as well as policy risk. To isolate the effects of the risk prices on a_t and b_t, we consider risk-neutral parameters μ^{Q*} and Φ^{Q*} where all parameters corresponding to the rows and columns of a_t and b_t in μ^{Q*} and Φ^{Q*} are set equal to their corresponding values in μ and Φ. Thus, in μ^{Q*} and Φ^{Q*} there is no risk for a_t and b_t, but we allow for macro risk.

Panel B of Figure 9 considers a security paying off $(1 + b_{t+1})$, which is the Fed's response to inflation. The first plot graphs the price of this security under the benchmark with full risk and under the specification where a_t and b_t risk is zero (which we compute using μ^{Q*} and Φ^{Q*}). We refer to this restricted risk-neutral price as the "Risk-Neutral a_t and b_t Price," which denotes that only the risk with respect to a_t and b_t has been turned off. The first plot of Panel B graphs the full-risk price and the risk-neutral a_t and b_t price. The difference between the two is graphed in the second plot. We plot values of the policy coefficient $(\bar{b} + b_t)$ on the x-axis.

Panel B shows that the actual price with full risk is always above the risk-neutral a_t and b_t price. Normally if a factor carries a risk premium, actual prices are lower than risk-neutral prices. The fact that the risk-neutral a_t and b_t price lies below the actual price means that activist monetary policy is valued by investors, who assign a risk discount to policy shift risk. The perceived value by investors to monetary policy shifts increases as monetary policy becomes increasing aggressive in combating inflation shocks.

The final graph in Panel B shows the gross returns of holding the security paying off $(1 + b_{t+1})$ under the two scenarios for values of $(\bar{b} + b_t)$ greater than 1.50. In the full risk case the gross return is negative, so investors are willing to pay to be exposed to Fed policy changes. When there is no a_t and b_t risk and only macro risk is priced, the expected gross return of the security paying off $(1 + b_{t+1})$ is approximately 1.02.

3.5.3 Policy Risk and the Yield Curve

Evaluating the effect of policy risk on the yield curve is more difficult because the posterior distribution of the policy latent factors, and hence the posterior distribution of implied yields, depend on the risk parameters. Computing yields requires both the best estimates of the latent factors as well as the prices of risk. In comparison, the previous exercises of Figure 9 required no estimates of the latent factors. Because of this, we compute the yield curve through the sample under several different risk scenarios and report the results in Table 3.

The first row of Table 3 reports the means and standard deviations of short rates and term spreads in the data (Case 0). The benchmark model estimates reported in Case 1 are almost identical. In Case 2, we compute yields under a model where all risk to a_t and b_t is turned off but we use the same sample path of the best estimates of a_t and b_t as the benchmark case. In Case 2, we set the price of risk for a_t and b_t to zero by specifying that all parameters in μ^Q and Φ^Q involving a_t or b_t are set to their real-measure counterparts in μ and Φ. The estimates of the short rates under Case 2 are identical to Case 1 by construction, but the different risk prices affect the term spreads. With no a_t or b_t risk, implied term spreads through the sample are 0.0099, slightly higher than the sample mean of 0.0091. The term spreads in Case 2 are significantly more volatile at 0.0197, compared to 0.0101 in data. Thus, if monetary policy were not priced by agents, term spreads would have been almost twice as volatile.

Case 2 suffers from the shortcoming that the best estimates for a_t and b_t are jointly determined with the specification of the prices of risk. In Case 3, we redraw the posterior distributions of a_t and b_t so that they are consistently estimated with the no-risk a_t and b_t specification in Case 2. Under Case 3 means and standard deviations of short rates are also near-identical to the data. But, the mean term spread is 0.0139, which is almost 50 basis points higher than the data. Term spreads are also more volatile, at 0.0131, than the data estimate of 0.0101. This implies that if agents did not value monetary policy, term spreads would have been higher and more volatile than what actually occurred. Stated differently, part of the reason why term spreads are low in our sample is due to the risk discount assigned by the market to monetary policy shifts.

Finally, Case 4 in the last row of Table 3 reports a case where there were no monetary policy shifts, $a_t = b_t = 0$, and short rates are set according to the Taylor rule $r_t = \bar{a} g_t + \bar{b} \pi_t$. The mean and standard deviation of the short rate are close to the data estimates, which has been noted by many authors investigating the fitted short rates implied by Taylor rules (see, for example, Taylor, 1993). These studies usually have not focused on the implied fit at the long end of the yield curve. With no discretionary policy shifts, term spreads would have been approximately

10 basis points higher than the data, but almost twice as volatile (0.0232 compared to 0.0101 in the data). Thus, activist monetary policy has contributed to a less volatile long-term yield.

4 Conclusion

Existing results suggest that monetary policy has changed in substantial ways over the last 50 years. While the implications of these changes for the business cycles dynamics has received considerable attention, little is known about the implications of these changes for financial markets, in particular for the term structure of interest rates.

In this paper we propose a quadratic term structure model where the coefficients of the short rate equation – which describe the behavior of monetary policy – can change over time. By exploiting term structure information, we are able to obtain sharper estimates of the changes in monetary policy compared to previous studies that do not use term structure information. Contrary to what is typically assumed, our empirical model does not impose that the shifts in monetary policy are exogenous. We find that the endogenous response of inflation to past changes in inflation loadings is an important component of how bond prices reflect monetary policy risk under the risk-neutral measure. An appealing feature is that our framework provides an estimate of the price of risk that financial market participants attribute to policy shifts.

Our empirical results show that monetary policy has changed in important ways and the shifts we estimate line up largely with narrative accounts of monetary policy and with some existing empirical estimates. We find that monetary policy shifts in inflation loadings show that monetary policy is more aggressive and flexible, on average, than estimations without long-term bonds employing unit-root changing coefficient models. In contrast, we find that policy shifts in output gap loadings exhibit little time series variation, so almost all changes in monetary policy stances have been done with respect to inflation.

A central contribution of the paper is to show that monetary policy shifts are priced by investors. We find that market participants assign an important value to activist monetary policy and agents are generally willing to pay to become exposed to monetary policy changes. If investors assigned no value to monetary policy shifts, then the slope of the yield curve would have been, on average, up to 50 basis points higher than the data. The term spread would have also been significantly more volatile without activist monetary policy that is priced by investors. This valuable contribution of monetary policy discretion is due to the risk discount assigned by investors for monetary policy shifts.

Appendix

A Bond Pricing

The price of a one-period zero-coupon bond is given by:

$$\begin{aligned}\hat{P}_t^1 &= \exp(-\hat{r}_t) = \exp(-\delta_0 - \delta_1^\top X_t - X_t^\top \Omega X_t) \\ &= \exp(A_1 + B_1^\top X_t + X_t^\top C_1 X_t),\end{aligned} \tag{A-1}$$

where $A_1 = -\delta_0$, $B_1 = -\delta_1 = -[\bar{a}\ \bar{b}\ 0\ 0]^\top$, and $C_1 = -\Omega$, with Ω given in equation (6).

Under measure $\mathbb{Q}$, the price of a n-period zero-coupon bond, $\hat{P}_t^n$, is:

$$\begin{aligned}\hat{P}_t^n &= \mathrm{E}_t^{\mathbb{Q}}(\exp(-\hat{r}_t)P_{t+1}^{n-1}) \\ &= \mathrm{E}_t^{\mathbb{Q}}(\exp\left(-\hat{r}_t + A_{n-1} + B_{n-1}^\top X_{t+1} + X_{t+1}^\top C_{n-1} X_{t+1}\right)) \\ &= \exp\left(-\hat{r}_t + A_{n-1} + B_{n-1}^\top(\mu^Q + \Phi^Q X_t) + (\mu^Q + \Phi^Q X_t)^\top C_{n-1}(\mu^Q + \Phi^Q X_t)\right) \\ &\quad \times \mathrm{E}_t^{\mathbb{Q}}(\exp\left((B_{n-1}^\top \Sigma + 2(\mu^Q + \Phi^Q X_t)^\top C_{n-1}\Sigma)\varepsilon_{t+1} + \varepsilon_{t+1}^\top \Sigma^\top C_{n-1}\Sigma\varepsilon_{t+1}\right)).\end{aligned} \tag{A-2}$$

To take the expectation, note that the expectation of the exponential of a quadratic Gaussian variable is given by:

$$\mathrm{E}[\exp(A^\top \epsilon + \epsilon^\top \Gamma \epsilon)] = \exp\left(-\frac{1}{2}\ln\det(I - 2\Psi\Gamma) + \frac{1}{2}A^\top(\Psi^{-1} - 2\Gamma)^{-1}A\right)$$

for $\epsilon \sim N(0, \Psi)$. This can be derived by general properties of Gaussian quadratic forms (see Mathai and Provost, 1992; Searle, 1997).

After taking the expectation and equating the terms with

$$\hat{P}_t^n = \exp(A_n + B_n^\top X_t + X_t^\top C_n X_t),$$

the coefficients A_n, B_n, and C_n are given by the recursions:

$$\begin{aligned}A_n &= -\delta_0 + A_{n-1} + B_{n-1}^\top \mu^Q + {\mu^Q}^\top C_{n-1}\mu^Q - \frac{1}{2}\ln\det(I - 2\Sigma^\top C_{n-1}\Sigma) \\ &\quad + \frac{1}{2}(\Sigma^\top B_{n-1} + 2\Sigma^\top C_{n-1}\mu^Q)^\top(I - 2\Sigma^\top C_{n-1}\Sigma)^{-1}(\Sigma^\top B_{n-1} + 2\Sigma^\top C_{n-1}\mu^Q) \\ B_n^\top &= -\delta_1^\top + B_{n-1}^\top \Phi^Q + 2{\mu^Q}^\top C_{n-1}\Phi^Q + 2(\Sigma^\top B_{n-1} + 2\Sigma^\top C_{n-1}\mu^Q)^\top(I - 2\Sigma^\top C_{n-1}\Sigma)^{-1}\Sigma^\top C_{n-1}\Phi^Q \\ C_n &= -\Omega + {\Phi^Q}^\top C_{n-1}\Phi^Q + 2(\Sigma^\top C_{n-1}\Phi^Q)^\top(I - 2\Sigma^\top C_{n-1}\Sigma)^{-1}(\Sigma^\top C_{n-1}\Phi^Q)\end{aligned} \tag{A-3}$$

If the model were specified in continuous time, then the recursions in equation (A-3) are versions of the ordinary differential equations derived by Ahn, Dittmar and Gallant (2002) on the bond pricing coefficients.

B Estimating the Model

The model is estimated using a Bayesian Gibbs sampling algorithm. While there are several examples of these types of estimations for affine models (see, among others, Lamoureux and Witte, 2002; Johannes and Polson, 2005; Ang, Dong and Piazzesi, 2006; Dong, 2006), these cannot be directly employed to estimate the quadratic model because in an affine setting, drawing the latent factors requires a Kalman filter. The Kalman filter assumes that yields are linear functions of state variables, whereas they are non-linear functions in the quadratic model. In this appendix, we develop an acceptance-rejection algorithm to draw the latent factors without approximation.

For ease of notation, we group the macro variables as $M_t = [g_t\ \pi_t]^\top$ and the latent factors as $L_t = [a_t\ b_t]^\top$ and rewrite the dynamics of $X_t = \left[M_t^\top\ L_t^\top\right]^\top$ in equation (2) as:

$$\begin{pmatrix} M_t \\ L_t \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} M_{t-1} \\ L_{t-1} \end{pmatrix} + \begin{pmatrix} \Sigma_{11} & 0 \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_{M,t} \\ \varepsilon_{L,t} \end{pmatrix}, \tag{B-1}$$

where $\varepsilon_t = (\varepsilon_{M,t}^\top \varepsilon_{L,t}^\top)^\top \sim$ IID $N(0, I)$ and Σ_{11} and Σ_{22} are lower triangular.

The parameters of the model are $\Theta = (\mu,\ \Phi,\ \Sigma,\ \delta_0,\ \delta_1,\ \Omega,\ \mu^Q,\ \Phi^Q,\ \sigma_u)$, where μ^Q and Φ^Q are parameters governing the state variable process under the risk neutral probability measure, and σ_u denotes the vector of observation error volatilities $\{\sigma_n\}$. We draw μ^Q and Φ^Q, but invert the prices of risk λ_0 and λ_1 using the relations:

$$\begin{aligned} \lambda_0 &= \Sigma^{-1}(\mu - \mu^Q) \\ \lambda_1 &= \Sigma^{-1}(\Phi - \Phi^Q). \end{aligned} \tag{B-2}$$

The latent factors $L_t = \{a_t\ b_t\}$ are generated in each iteration of the Gibbs sampler. Note that Ω is not a parameter, but is fixed from equation (6). We also do not draw δ_0, but set δ_0 in each iteration to match the sample mean of the short rate.

We now detail the procedure for drawing each of these variables. We denote the factors $X = \{X_t\}$ and the set of yields for all maturities in data as $Y = \{y_t^n\}$. Note that the model-implied yields $\hat{Y} = \{\hat{y}_t^n\}$ differ from the yields in data, Y, by observation error. By definition, $Y = \hat{Y} + u$, where $u = \{u_t^n\}$ is the set of all observation errors for all yields. This notation also implies that the short rate in data, r_t, is the same as y_t^1.

B.1 Drawing the Latent Factors

We use a single-move algorithm based on Jacquier, Polson and Rossi (1994, 2004) adapted to our model. We derive a draw from the distribution $P(L_t|Y, L_{-t}, M)$, where L_t is the t-th observation of the latent factors, L_{-t} denotes all the latent factors except the t-th observation, and Y and M are the complete time-series of yields and macro variables, respectively. We use the notation Y_t and M_t to denote the t-th observation of the set of yields and macro variables. We draw the latent factors L_t conditional on the macro factors, yields, and other parameters.

From the Markov structure of the model, we can write:

$$P(L_t|L_{-t}, Y, M, \Theta) \propto P(L_t|L_{t-1}, M, \Theta)P(Y_t|L_t, M, \Theta)P(L_{t+1}|L_t, M, \Theta). \tag{B-3}$$

To keep the notation to a minimum, we write this as:

$$P(L_t|L_{-t}) \propto P(L_t|L_{t-1})P(Y_t|L_t)P(L_{t+1}|L_t).$$

Since M and Θ are treated as known, we can write the dynamics for L_t in equation (B-1) as:

$$\begin{aligned} L_t &= \mu_2 + \Sigma_{12}\varepsilon_{M,t} + \Phi_{21}M_{t-1} + \Phi_{22}L_{t-1} + \Sigma_{22}\varepsilon_{L,t} \\ &= \mu_{L,t} + \Phi_L L_{t-1} + \Sigma_L \varepsilon_{L,t}, \end{aligned} \tag{B-4}$$

where $\mu_{L,t} = \mu_2 + \Sigma_{12}\varepsilon_{M,t}$, $\Phi_L = \Phi_{22}$, and $\Sigma_L = \Sigma_{22}$. Since M is observable and we hold Θ as fixed, $\mu_{L,t}$ is known at time t.

Each conditional distribution of the RHS of equation (B-3) is known. From equation (B-4) we have

$$P(L_t|L_{t-1}) \propto \exp\left(-\frac{1}{2}(L_t - \mu_L - \Phi_L L_{t-1})^\top (\Sigma_L \Sigma_L^\top)^{-1}(L_t - \mu_L - \Phi_L L_{t-1})\right). \tag{B-5}$$

Similarly, from the VAR in equation (B-4) we can write:

$$P(L_{t+1}|L_t) \propto \exp\left(-\frac{1}{2}(L_{t+1} - \mu_L - \Phi_L L_t)^\top (\Sigma_L \Sigma_L^\top)^{-1}(L_{t+1} - \mu_L - \Phi_L L_t)\right). \tag{B-6}$$

Finally, the likelihood of bond yields, $P(Y_t|L_t)$, is given by:

$$P(Y_t|L_t) \propto \exp\left(-\frac{1}{2}\sum_n \left[\frac{(y_t^n - (a_n + b_n^\top X_t + X_t^\top c_n X_t))^2}{\sigma_n^2}\right]\right), \tag{B-7}$$

where $X_t = [L_t^\top\ M_t^\top]^\top]$ and the summation is taken over yield maturities n, which includes the short rate with $n = 1$. In the likelihood, the model-implied yield, $\hat{y}_t^n = a_n + b_n^\top X_t + X_t^\top c_n X_t$, is given in equation (12), and σ_n^2 is the observation error variance of the yield of maturity n.

We can combine equations (B-5)-(B-7) and complete the square to obtain:

$$\begin{aligned} P(L_t|L_{-t}) &\propto P(Y_t|L_t)\exp\Big(-\frac{1}{2}\Big[L_t^\top(\Phi_L'^\top(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+(\Sigma_L\Sigma_L^\top)^{-1})L_t \\ &\quad -2(L_{t+1}^\top(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+L_{t-1}^\top\Phi_L(\Sigma_L\Sigma_L^\top)^{-1}-\mu_L(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+\mu_L(\Sigma_L\Sigma_L^\top)^{-1})L_t\Big]\Big) \\ &\propto P(Y_t|L_t)\exp\left(-\frac{1}{2}(L_t-\mu_t^*)^\top(\Sigma_t^*)^{-1}(L_t-\mu_t^*)\right) \end{aligned} \tag{B-8}$$

where

$$\begin{aligned} \Sigma_t^* &= \left(\Phi_L'^\top(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+(\Sigma_L\Sigma_L^\top)^{-1}\right)^{-1} \\ \mu_t^* &= \Sigma_t^*(L_{t+1}^\top(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+L_{t-1}^\top\Phi_L(\Sigma_L\Sigma_L^\top)^{-1}-\mu_L(\Sigma_L\Sigma_L^\top)^{-1}\Phi_L+\mu_L(\Sigma_L\Sigma_L^\top)^{-1})^\top. \end{aligned}$$

Since this distribution is not recognizable, we use a Metropolis draw. We draw a proposal from the distribution $N(\mu_t^*,\Sigma_t^*)$ and then the acceptance probability is based on the likelihood of $P(Y_t|L_t)$. Since we specify the mean of L to be zero for identification, we set each generated draw of L to have a mean of zero.

To generate initial values for the very first draw, we use the Carter and Kohn (1994) forward-backward algorithm to first run a Kalman filter forward and then sample L_t backwards. The Kalman filter is constructed linearizing the yields at $L_{t|t-1}$. Note that this Kalman filter is only used to produce initial values for the draw; the steady-state distribution of L_t relies on the single-step accept/reject algorithm given above.

B.2 Drawing μ and Φ

We follow Johannes and Polson (2005) and explicitly differentiate between $\{\mu,\Phi\}$ under the real measure and $\{\mu^Q,\Phi^Q\}$ under the risk-neutral measure. As X_t follows a VAR in equation (2), we follow standard Gibbs sampling and use conjugate normal priors and posteriors for the draw of μ and Φ. We note that the posterior of μ and Φ conditional on X, Y and the other parameters is:

$$\begin{aligned} P(\mu,\Phi|\Theta_-,X,Y) &\propto P(Y|\Theta,X)P(X|\mu,\Phi,\Sigma)P(\mu,\Phi) \qquad \text{(B-9)} \\ &\propto P(Y|\Sigma,\delta_0,\delta_1,\mu^Q,\Phi^Q,\sigma_\eta,X)P(X|\mu,\Phi,\Sigma)P(\mu,\Phi) \\ &\propto P(X|\mu,\Phi,\Sigma)P(\mu,\Phi), \qquad \text{(B-10)} \end{aligned}$$

where Θ_- denotes the set of all parameters except μ and Φ, and $P(X|\mu,\Phi,\Sigma)$ is the likelihood function of the VAR, which is normally distributed from the assumption of normality for the errors in the VAR. The validity of going from the first line to the second line is ensured by the bond recursion in equation (A-3): given μ^Q and Φ^Q, the bond price is independent of μ and Φ. We specify the prior $P(\mu,\Phi)$ to be $N(0,1000)$, which effectively represents an uninformative prior. We draw μ and Φ separately for each equation in the VAR system (2).

B.3 Drawing $\Sigma\Sigma^\top$

To draw $\Sigma\Sigma^\top$, we note that the posterior of $\Sigma\Sigma^\top$ conditional on X, Y and the other parameters is:

$$P(\Sigma\Sigma^\top|\Theta_-,X,Y)\propto P(Y|\Theta,X)P(X|\mu,\Phi,\Sigma)P(\Sigma\Sigma^\top), \tag{B-11}$$

where Θ_- denotes the set of all parameters except Σ. This posterior suggests an Independence Metropolis draw. We draw $\Sigma\Sigma^\top$ from the proposal density

$$q(\Sigma\Sigma^\top)=P(X\mid\mu,\Phi,\Sigma)P(\Sigma\Sigma^\top),$$

which is an Inverse Wishart (IW) distribution if we specify the prior $P(\Sigma\Sigma^\top)$ to be IW, so that $q(\Sigma\Sigma^\top)$ is an IW natural conjugate. The proposal draw $(\Sigma\Sigma^\top)^{m+1}$ for the $(m+1)$th draw is then accepted with probability α, where

$$\begin{aligned} \alpha &= \min\left\{\frac{P((\Sigma\Sigma^\top)^{m+1}\mid\Theta_-,X,Y)}{P((\Sigma\Sigma^\top)^{m}\mid\Theta_-,X,Y)}\frac{q((\Sigma\Sigma^\top)^{m})}{q((\Sigma\Sigma^\top)^{m+1})},1\right\} \\ &= \min\left\{\frac{P(Y\mid(\Sigma\Sigma^\top)^{m+1},\Theta_-,X)}{P(Y\mid(\Sigma\Sigma^\top)^{m},\Theta_-,X)},1\right\}, \end{aligned} \tag{B-12}$$

where $P(Y|\mu, \Phi, \Theta_-, X)$ is the likelihood function of all yields, including the short rate, which is normally distributed from the assumption of normality for the observation errors. From equation (B-12), α is just the ratio of the likelihoods of the new draw of $\Sigma\Sigma^\top$ relative to the old draw.

B.4 Drawing $\bar{a}$ and $\bar{b}$

We draw $\bar{a}$ and $\bar{b}$ jointly with a Random Walk Metropolis algorithm. We assume a flat prior. The accept/reject probability for the draws of $\bar{a}$ and $\bar{b}$ is the ratio of the likelihood of bond yields based on candidate and last draw of $\bar{a}$ and $\bar{b}$:

$$\begin{aligned} \alpha &= \min\left\{ \frac{P((\bar{a},\bar{b})^{m+1} \mid \Theta_-, X, Y)}{P((\bar{a},\bar{b})^{m} \mid \Theta_-, X, Y)} \frac{q((\bar{a},\bar{b})^{m})}{q((\bar{a},\bar{b})^{m+1})}, 1 \right\} \\ &= \min\left\{ \frac{P(Y \mid (\bar{a},\bar{b})^{m+1}, \Theta_-, X)}{P(Y \mid (\bar{a},\bar{b})^{m}, \Theta_-, X)}, 1 \right\}. \end{aligned} \tag{B-13}$$

B.5 Drawing μ^Q and Φ^Q

We draw μ^Q and Φ^Q with a Random Walk Metropolis algorithm assuming a flat prior. We draw each parameter separately in μ^Q, and each row in Φ^Q. The accept/reject probability for the draws of μ^Q and Φ^Q is the ratio of the likelihood of bond yields based on candidate and last draw of μ^Q and Φ^Q:

$$\begin{aligned} \alpha &= \min\left\{ \frac{P((\mu^Q,\Phi^Q)^{m+1} \mid \Theta_-, X, Y)}{P((\mu^Q,\Phi^Q)^{m} \mid \Theta_-, X, Y)} \frac{q((\mu^Q,\Phi^Q)^{m})}{q((\mu^Q,\Phi^Q)^{m+1})}, 1 \right\} \\ &= \min\left\{ \frac{P(Y \mid (\mu^Q,\Phi^Q)^{m+1}, \Theta_-, X)}{P(Y \mid (\mu^Q,\Phi^Q)^{m}, \Theta_-, X)}, 1 \right\}, \end{aligned} \tag{B-14}$$

In each iteration, we invert λ_0 and λ_1 and report the estimates of the prices of risk instead of μ^Q and Φ^Q. We discard non-stationary draws of Φ^Q.

B.6 Drawing σ_u

Drawing the variance of the observation errors, σ_u^2, is straightforward, because we can view the observation errors η as regression residuals from equation (13). We draw the observation variance $(\sigma_\eta^n)^2$ separately from each yield. We specify a conjugate prior $IG(0, 0.00001)$, so that the posterior distribution of σ_η^2 is a natural conjugate Inverse Gamma distribution. The prior information roughly translates into a 30bp bid-ask spread in Treasury securities, which is consistent with studies on the liquidity of spot Treasury market yields (see, for example, Fleming, 2000).

C Impulse Responses

Since the yields are non-linear, we follow Gallant, Rossi and Tauchen (1993) and Potter (2000), among others, and compute the impulse response functions using simulation. We start with the sample series of data (g_t and π_t) and the posterior means of the latent factors (a_t and b_t) at each observation t. We term these points X_t^*. From the VAR in equation (2), we construct an orthogonalized error term ν_t by taking the Cholesky of $\Sigma\Sigma^\top$. To construct the impulse response for the jth variable of X_t, we first draw a shock v_t that represents a shock only to variable j from the error term distribution ν_t. From the points X_t^*, we construct a new series where each observation has been shocked by v_t, which we denote as $X_t^v = X_t^* + v_t$.

The impulse response functions are taken as the difference between the averaged response of the yields to the evolution of X_t^* without shocks to the evolution of the shocked X_t^v series:

$$\mathrm{E}(y_{t+k}^n | X_t^v) - \mathrm{E}(y_{t+k}^n | X_t^*).$$

Using the VAR in equation (2), we simulate out the value of X_{t+k}^v from X_t^v and the value of X_{t+k}^* from X_t^*. This is done at each observation t. Then, we construct the yields, y_{t+k}^n, from equation (12) corresponding to the state vectors X_{t+k}^v and X_{t+k}^*. We take values of $k = 1 \ldots 60$ quarters.

The impulse responses are computed at each observation by taking the average of the sample paths of y^n_{t+k} computed using X^v_{t+k} minus the average of the sample paths of y^n_{t+k} computed using X^*_{t+k}. We report the average of the impulse responses across all observations t. This procedure results in impulse responses that are identical to impulse responses computed for traditional VAR systems for large numbers of observations.

References

[1] Ahn, D. H., R. F. Dittmar, and A. R. Gallant, 2002, "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243-288.

[2] Ang, A., G. Bekaert, and M. Wei, 2007, "The Term Structure of Real Rates and Expected Inflation," forthcoming *Journal of Finance*.

[3] Ang, A., S. Dong, and M. Piazzesi, 2006, "No-Arbitrage Taylor Rules," working paper, Columbia University.

[4] 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.

[5] Beaglehold, D., and M. Tenney, 1992, "A Nonlinear Equilibrium Model of the Term Structure of Interest Rates: Corrections and Additions," *Journal of Financial Economics*, 12, 763-806.

[6] Bikbov, R., 2006, "Monetary Policy Regimes and the Term Structure of Interest Rates," unpublished dissertation, Columbia University.

[7] Bikbov, R., and M. Chernov, 2006, "No-Arbitrage Macroeconomic Determinants of the Yield Curve," working paper, Columbia University.

[8] Boivin, J., 2006, "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," *Journal of Money, Credit and Banking*, 38, 5, 1149-1174.

[9] Boivin, J. and M. Giannoni, 2006, "Has Monetary Policy Become More Effective?," forthcoming *The Review of Economics and Statistics*. 88, 3, 445-462.

[10] Brandt, M. W., and Chapman, D. A., 2003, "Comparing Multifactor Models of the Term Structure," working paper, Duke University.

[11] Buraschi, A., A. Cieslak, and F. Trojani, 2007, "Correlation Risk and the Term Structure of Interest Rates," working paper, Imperial College.

[12] Carter, C. K., and R. Kohn, 1994, "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541-553.

[13] Christiano, L. J., M. Eichenbaum, and C. Evans, 1996, "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics*, 78, 16-34.

[14] Christiano, L. J., M. Eichenbaum, M., and C. Evans, 1999, "Monetary Policy Shocks: What have we Learned and to what End?" in J. B. Taylor, and M. Woodford, ed., *Handbook of Macroeconomics*, Elsevier Science, North Holland.

[15] Clarida, R., Galí, J., and M. Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147-80.

[16] Cogley, T., 2005, "Changing Beliefs and the Term Structure of Interest Rates: Cross-Equation Restrictions with Drifting Parameters," *Review of Economic Dynamics*, 8, 420-451.

[17] Cogley, T., and T. J. Sargent, 2001, "Evolving Post World War II U.S. Inflation Dynamics," *NBER Macroeconomics Annual*, 16, 331-373.

[18] Cogley, T., and T. J. Sargent, 2005, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics*, 8, 262-302.

[19] Constantinides, G., 1992, "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5, 531-552.

[20] Cooley, T., and E. Prescott, 1976, "Estimation in the Presence of Stochastic Parameter Variation," *Econometrica*, 44, 167-184.

[21] Cox, J. C., J. E. Ingersoll, and S. A. Ross, 1985, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363-384.

[22] Dai, Q., and K. J. Singleton, 2002, "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure," *Journal of Financial Economics*, 63, 415-41.

[23] Dong, S., 2006, "Macro Variables Do Drive Exchange Rate Movements: Evidence from a No-Arbitrage Model," unpublished dissertation, Columbia University.

[24] Duffie, D., and R. Kan, 1996, "A Yield-Factor Model of Interest Rates," *Mathematical Finance*, 6, 379-406.

[25] Estrella, A., and F. S. Mishkin, 1998, "Predicting U.S. Recessions: Financial Variables as Leading Indicators," *Review of Economics and Statistics*, 1, 45-61.

[26] Filipovic, D., and J. Teichmann, 2002, "On Finite Dimensional Term Structure Models," working paper, Princeton University.

[27] Fleming, M. J., 2003, "Measuring Teasury Market Liquidity," *Federal Reserve Bank of New York Economic Policy Review*, 9, 83-108.

[28] Gallant, A. R., P. E. Rossi, and G. Tauchen, 1993, "Nonlinear Dynamic Structures," *Econometrica*, 61, 871-908.

[29] Gourieroux, C., and R. Sufana, 2003, "Wishart Quadratic Term Structure Models," working paper, University of Toronto.

[30] Jacquier, E., N. G. Polson, and P. E. Rossi, 1994, "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics*, 12, 371-417.

[31] Jacquier, E., N. G. Polson, and P. E. Rossi, 2004, "Bayesian Analysis of Fat-Tailed Stochastic Volatility Models with Correlated Errors," *Journal of Econometrics*, 122, 185-212.

[32] Johannes, M., and N. Polson, 2005, "MCMC Methods for Financial Econometrics," working paper, Columbia University.

[33] Justiniano, A., and G. E. Primiceri, 2006, "The Time Varying Volatility of Macroeconomic Fluctuations," NBER working paper 12022.

[34] Koop, G., M. H. Pesaran, and S. M. Potter, 1996, "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74, 119-147.

[35] Lamoureux, C., and D. Witte, 2002, "Empirical Analysis of the Yield Curve: The Information in the Data Viewed through the Window of Cox, Ingersoll, and Ross," *Journal of Finance*, 57, 1479-1520.

[36] Leippold, M., and L. Wu, 2002, "Asset Pricing under the Quadratic Class," *Journal of Financial and Quantitative Analysis*, 37, 271-295.

[37] Leippold, M., and L. Wu, 2003, "Design and Estimation of Quadratic Term Structure Models," *European Finance Review*, 7, 47-73.

[38] Longstaff, F., 1989, "A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates," *Journal of Finance*, 23, 1259-1282.

[39] Mathai, A. M., and S. B. Provost, 1992, *Quadratic Forms in Random Variables: Theory and Applications*, Marcel Dekker, New York.

[40] Meltzer, A. H., 2005, "From Inflation to More Inflation, Disinflation, and Low Inflation," Keynote Address, Conference on Price Stability, Federal Reserve Bank of Chicago.

[41] Orphanides, A., 2001, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review*, 91, 964-85.

[42] Pagan, A., 1980, "Some Identification and Estimation Results for Regression Models with Stochastically Varying Coefficients," *Journal of Econometrics*, 13, 341-363.

[43] Rudebusch, G. D., and L. E. O. Svensson, 2002, "Eurosystem Monetary Targeting: Lessons from U.S. Data," *European Economic Review*, 46, 417-442.

[44] Searle, S. R., 1997, *Linear Models*, John Wiley & Sons.

[45] Sims, C. A., and T. Zha, 2006, "Macroeconomic Switching," *American Economic Review*, 96, 54-81.

[46] Stock, J., and M. Watson, 2003, "Has the Business Cycle Changed? Evidence and Explanations," working paper, Princeton University.

[47] Taylor, J. B., 1993, "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214.

[48] Taylor, J. B., 1999, "A Historical Analysis of Monetary Policy Rules," in *Monetary Policy Rules*, ed. J. B. Taylor, NBER Conference Report Series, University of Chicago Press, 319-41.

Table 1: Summary Statistics

Panel A: Moments of Macro Factors

	Means (%)		Standard Deviations (%)		Autocorrelations	
	Data	Model	Data	Model	Data	Model
g	0.000	-0.018	0.589	0.626	0.930	0.929
	(0.082)	(0.134)	(0.057)	(0.167)	(0.034)	(0.024)
π	0.856	0.830	0.554	0.592	0.982	0.983
	(0.082)	(0.382)	(0.066)	(0.221)	(0.026)	(0.007)

Panel B: Moments of Yields

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
Means (%)						
Data	1.275	1.363	1.411	1.452	1.482	1.501
	(0.103)	(0.102)	(0.101)	(0.098)	(0.098)	(0.103)
Model	1.275	1.351	1.413	1.453	1.481	1.502
	–	(0.005)	(0.003)	(0.002)	(0.002)	(0.002)
Standard Deviations (%)						
Data	0.710	0.701	0.691	0.674	0.667	0.657
	(0.087)	(0.077)	(0.076)	(0.074)	(0.074)	(0.075)
Model	0.716	0.697	0.686	0.677	0.667	0.654
	(0.011)	(0.005)	(0.002)	(0.001)	(0.001)	(0.002)
Autocorrelations						
Data	0.936	0.944	0.952	0.958	0.961	0.964
	(0.030)	(0.029)	(0.027)	(0.026)	(0.026)	(0.024)
Model	0.942	0.953	0.961	0.965	0.966	0.967
	(0.004)	(0.003)	(0.002)	(0.001)	(0.000)	(0.001)

The table lists various moments of the factors in data and implied by the model. All the factors and yields are expressed at a quarterly frequency in percentage terms. All standard errors are reported in parentheses. Panel A lists moments of the output gap and inflation. For the model, we construct the posterior distribution of unconditional moments by computing the unconditional moments implied from the parameters in each iteration of the Gibbs sampler. Panel B reports data and model unconditional moments of n-quarter maturity yields. We compute the posterior distribution of the model-implied yields using the generated latent factors in each iteration of the Gibbs sampler. In Panels A and B, the data standard errors are computed using GMM with robust standard errors. The sample period is June 1952 to December 2006.

Table 2: Parameter Estimates

Short Rate Parameters

δ_0	$\bar{a}$	$\bar{b}$
0.001	0.223	1.442
(0.001)	(0.045)	(0.100)

VAR Parameters

		Φ				Volatility $\times 1000$/Correlation Matrix			
	$\mu \times 1000$	g	π	a	b	g	π	a	b
g	0.805	0.913	-0.098	-0.008	0	0.005	0.009	-0.065	0.112
	(0.344)	(0.028)	(0.058)	(0.028)	–	(0.001)	(0.076)	(0.075)	(0.067)
π	0.107	0.064	0.988	0	-0.000	0.009	0.001	-0.227	-0.186
	(0.089)	(0.011)	(0.007)	–	(0.000)	(0.076)	(0.000)	(0.072)	(0.095)
a	-0.035	0.004	0	0.957	0	-0.065	-0.227	0.005	-0.117
	(0.157)	(0.034)	–	(0.022)	–	(0.075)	(0.072)	(0.002)	(0.245)
b	-7.137	0	1.183	0	0.942	0.112	-0.186	-0.117	38.65
	(25.60)	–	(2.518)	–	(0.024)	(0.067)	(0.095)	(0.245)	(13.31)

Risk Premia Parameters

		λ_1				Φ^Q			
	λ_0	g	π	a	b	g	π	a	b
g	2.330	130.7	36.82	277.4	0	0.637	-0.175	-0.594	0
	(0.505)	(23.16)	(26.50)	(40.17)	–	(0.036)	(0.042)	(0.065)	–
π	-2.941	22.24	164.1	-3.672	1.055	0.042	0.838	0	-0.001
	(0.422)	(15.03)	(11.50)	(21.08)	(0.119)	(0.001)	(0.002)	–	(0.000)
a	-0.394	18.46	41.03	12.35	0.250	-0.007	0	0.970	0
	(0.325)	(19.82)	(13.78)	(24.62)	(0.087)	(0.005)	–	(0.010)	–
b	-1.156	-9.947	43.76	-34.97	0.080	0	-0.099	0	0.974
	(0.553)	(11.00)	(27.12)	(20.77)	(0.172)	–	(0.065)	–	(0.010)

Observation Error Standard Deviation

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
σ_u^n	0.132	0.054	0.031	0.020	0.020	0.026
	(0.010)	(0.010)	(0.005)	(0.002)	(0.001)	(0.002)

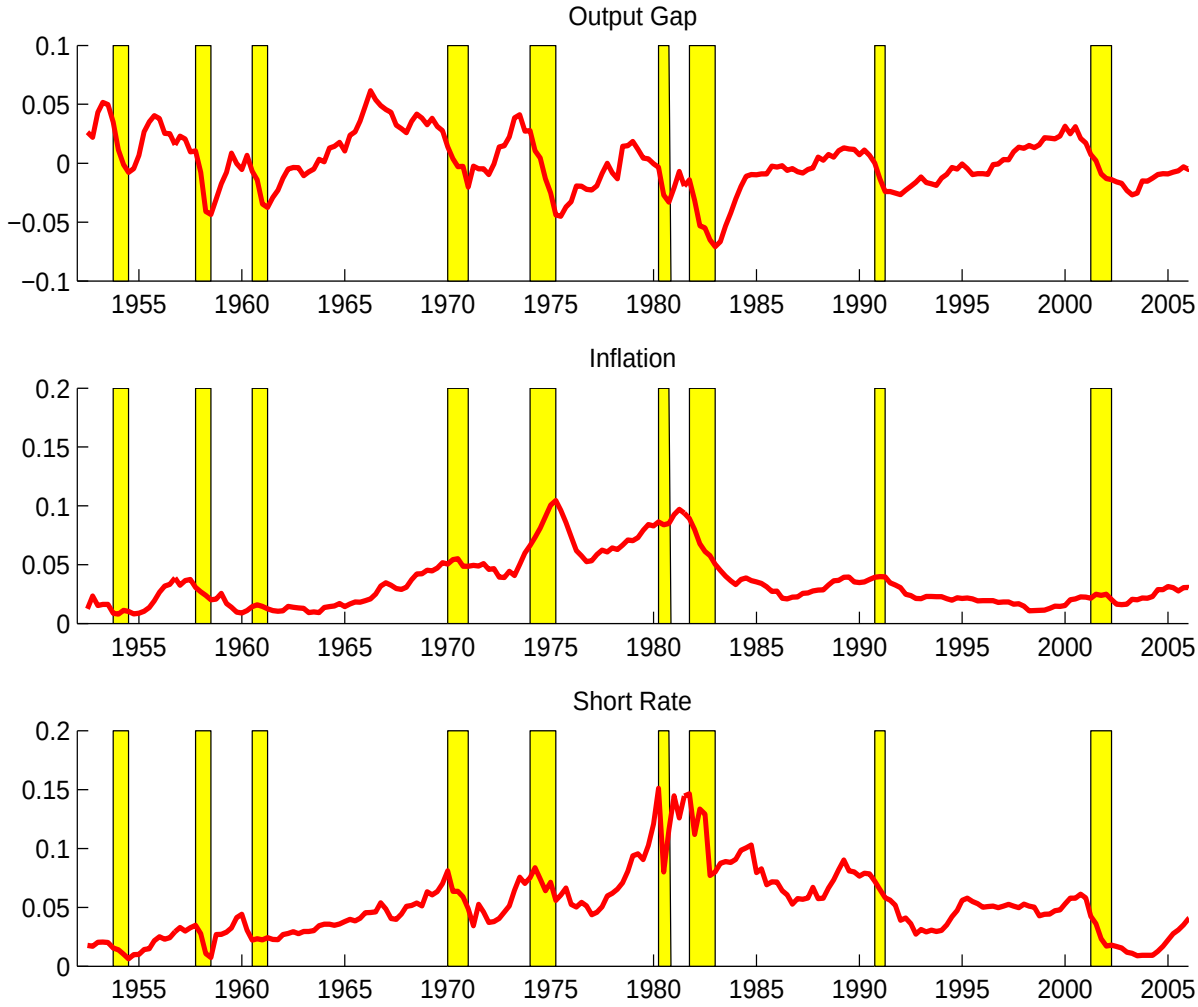
The table lists parameter values for the model in equations (2)-(8) and observation error standard deviations in equation (13) for yields of maturity n quarters. Any parameters set to zero without standard errors are not estimated. We also report the risk-neutral companion form Φ^Q given by equation (10). We estimate the model by Gibbs sampling using 50,000 simulations after a burn-in sample of 10,000. We report the posterior mean and posterior standard deviation (in parentheses) of each parameter. In the Volatility/Correlation matrix, we report standard deviations of each factor along the diagonal multiplied by 1000 and correlations between the factors on the off-diagonal elements. The sample period is June 1952 to December 2006 and the data frequency is quarterly.

Table 3: Yield Curve Under Different Risk Assumptions

		Short Rates r_t		Spreads $y_t^{20} - r_t$	
		Mean	Stdev	Mean	Stdev
0)	Data	0.0510	0.0286	0.0091	0.0101
1)	Full Risk Model	0.0510	0.0286	0.0091	0.0107
2)	No a_t, b_t Risk and Monetary Policy is the Same as (1)	0.0510	0.0286	0.0099	0.0197
3)	No a_t, b_t Risk	0.0510	0.0286	0.0139	0.0131
4)	Full Risk Model with $a_t = b_t = 0$	0.0547	0.0311	0.0102	0.0232

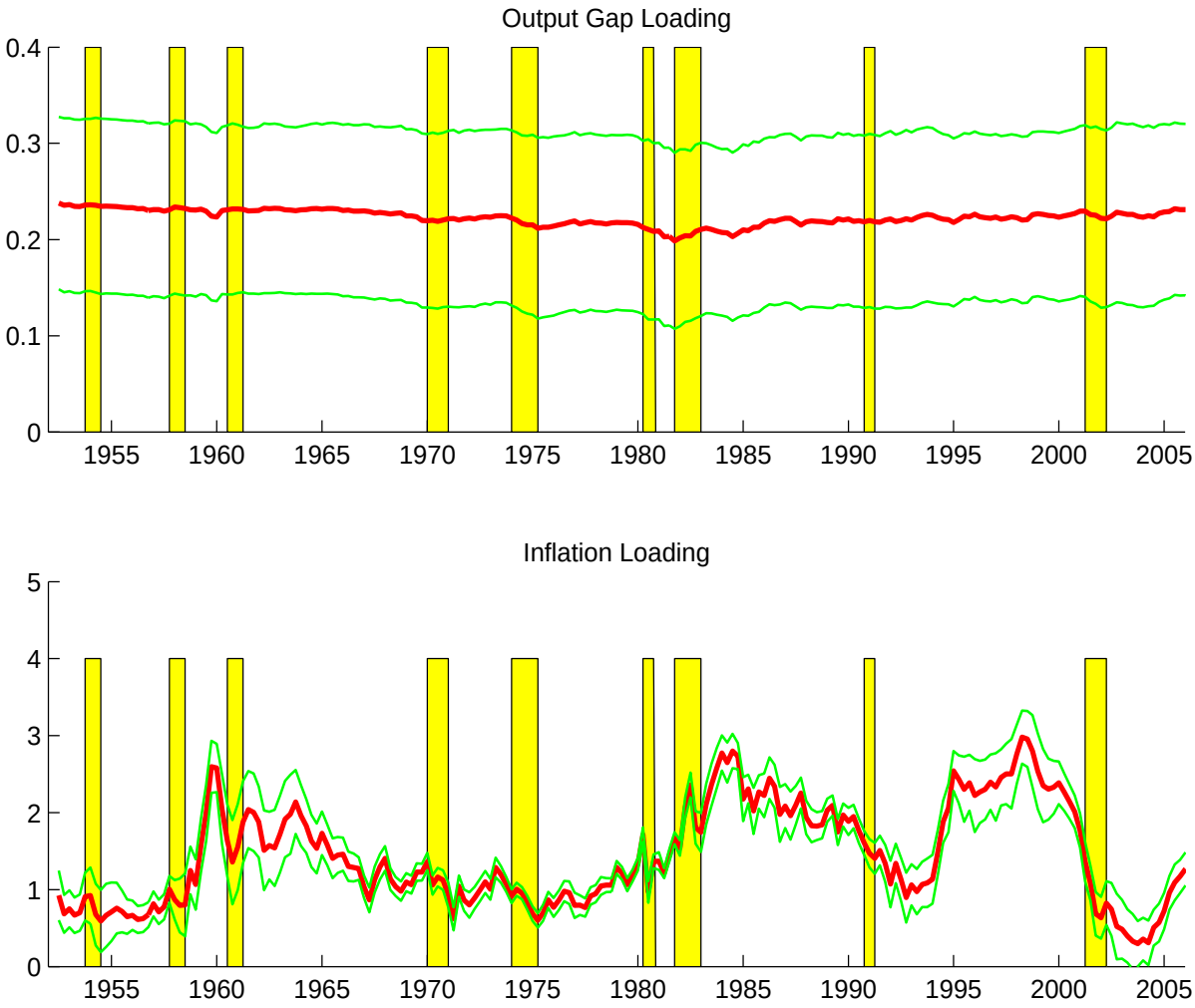
The table reports the annualized mean and standard deviation of short rates r_t and term spreads $y_t^{20} - r_t$ in the data (Case 0) and under different risk cases. Case 1 corresponds to the benchmark model and evaluates the short rates and spreads through the sample using the best-estimate posterior means of a_t and b_t. In Case 2, all parameters in the rows and columns of μ^Q and Φ^Q corresponding to a_t and b_t are set to their real-measure estimates μ and Φ and the best estimates of the policy factors a_t and b_t are the same as Case 1. Case 3 has the same parameter structure as Case 2 except a_t and b_t are redrawn optimally and the short rates and spreads are evaluated at these new optimal best estimates. In Case 4, we evaluate the short rates and spreads with $a_t = b_t = 0$ using the parameter values in Case 1. The sample period is June 1952 to December 2006 and the data frequency is quarterly.

Figure 1: Output Gap, Inflation, and the Short Rate



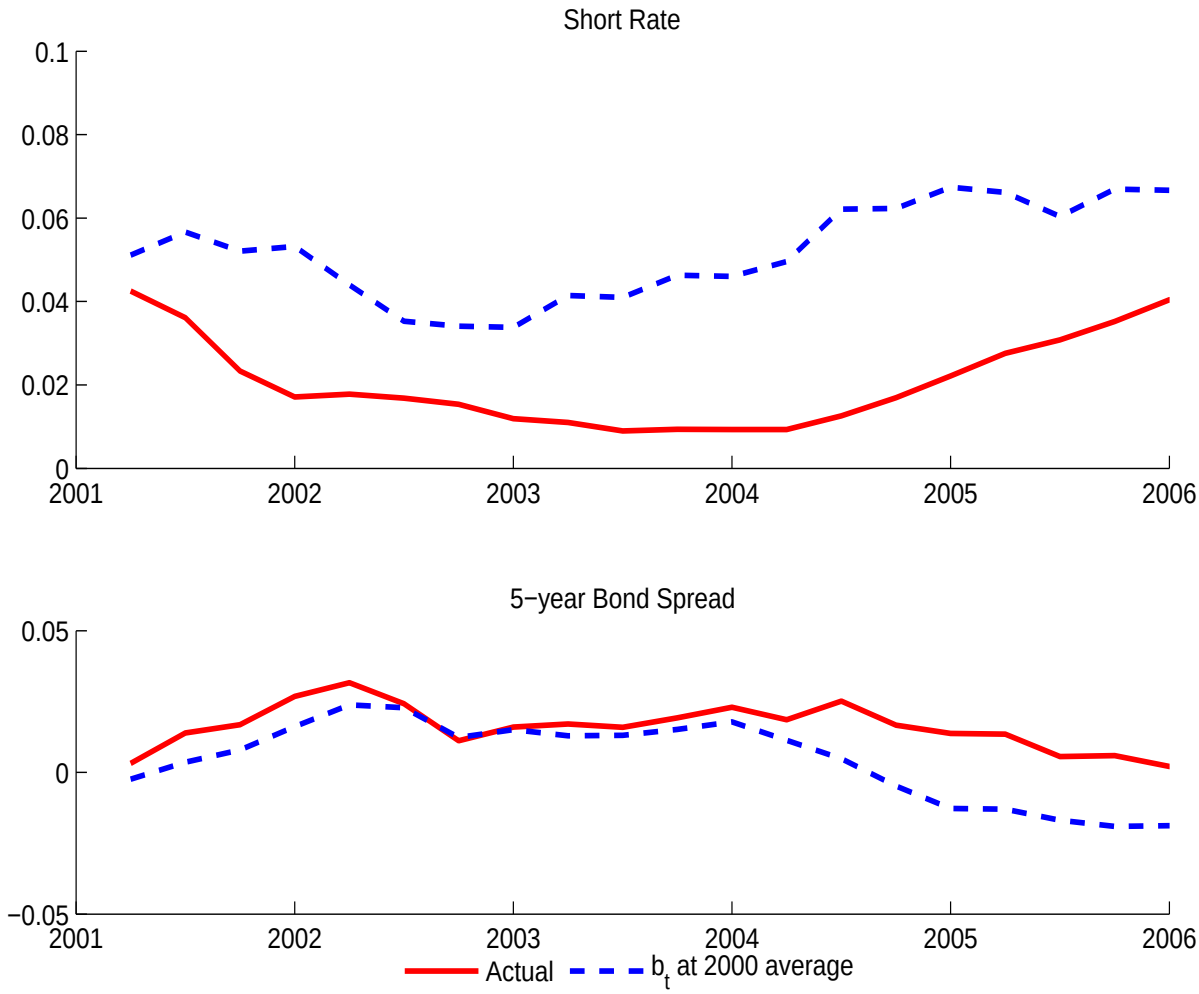
We plot the output gap, inflation, and the short rate. The output gap is defined as the proportional difference between actual and potential real GDP. Inflation is the year-on-year GDP deflator. The short rate is the 3-month T-bill yield. We overlay the NBER recession periods in shaded bars. The sample period is from June 1952 to December 2006 and the data frequency is quarterly. All data is annualized.

Figure 2: Time-Varying Policy Coefficients



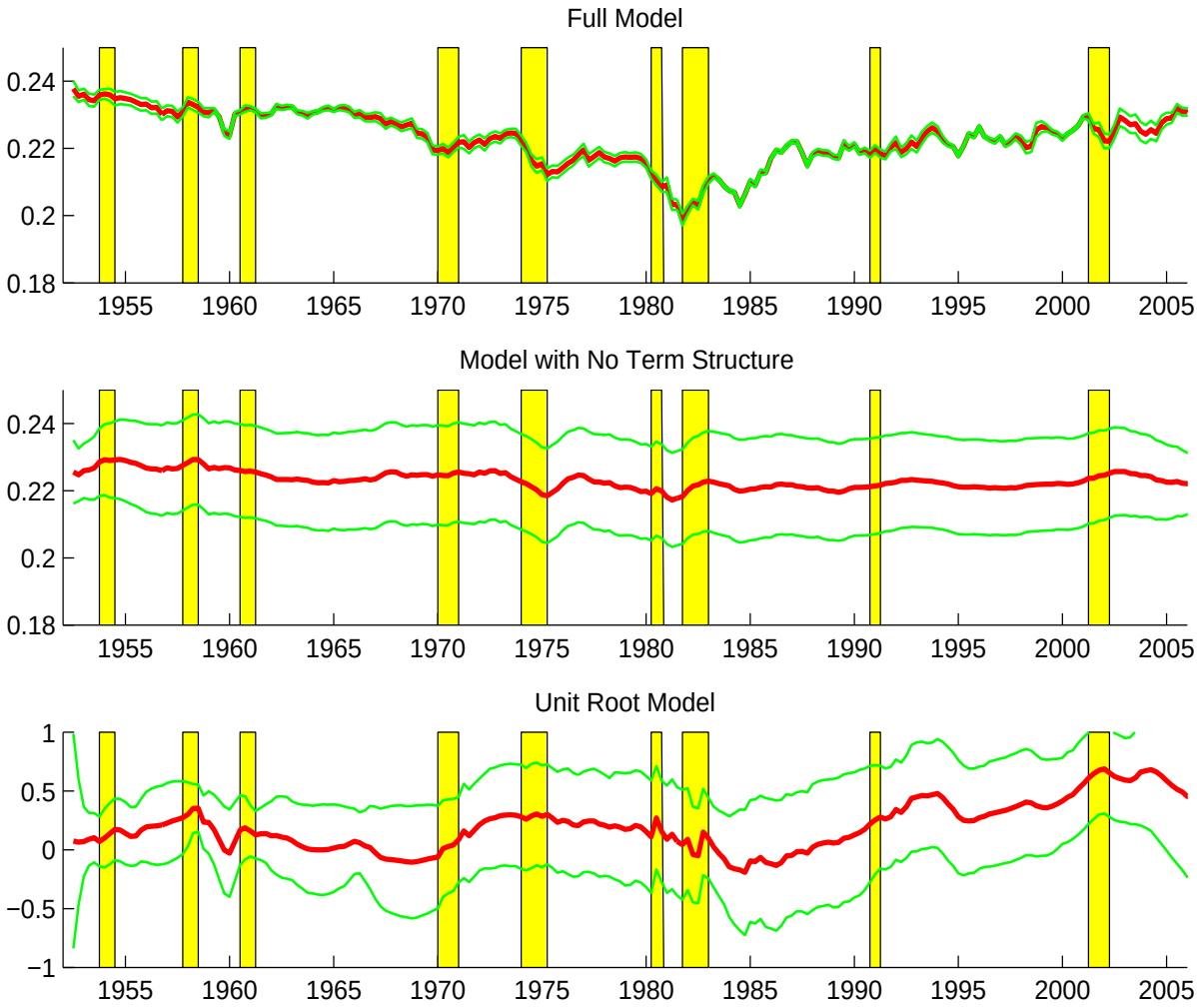
We plot the posterior mean of the time-varying coefficient $\bar{a} + a_t$ and $\bar{b} + b_t$ in the thick lines together with two posterior standard deviation bands in thin lines. We overlay the NBER recession periods in shaded bars. The sample period is from June 1952 to December 2006 and the data frequency is quarterly.

Figure 3: Counter-Factual Experiment



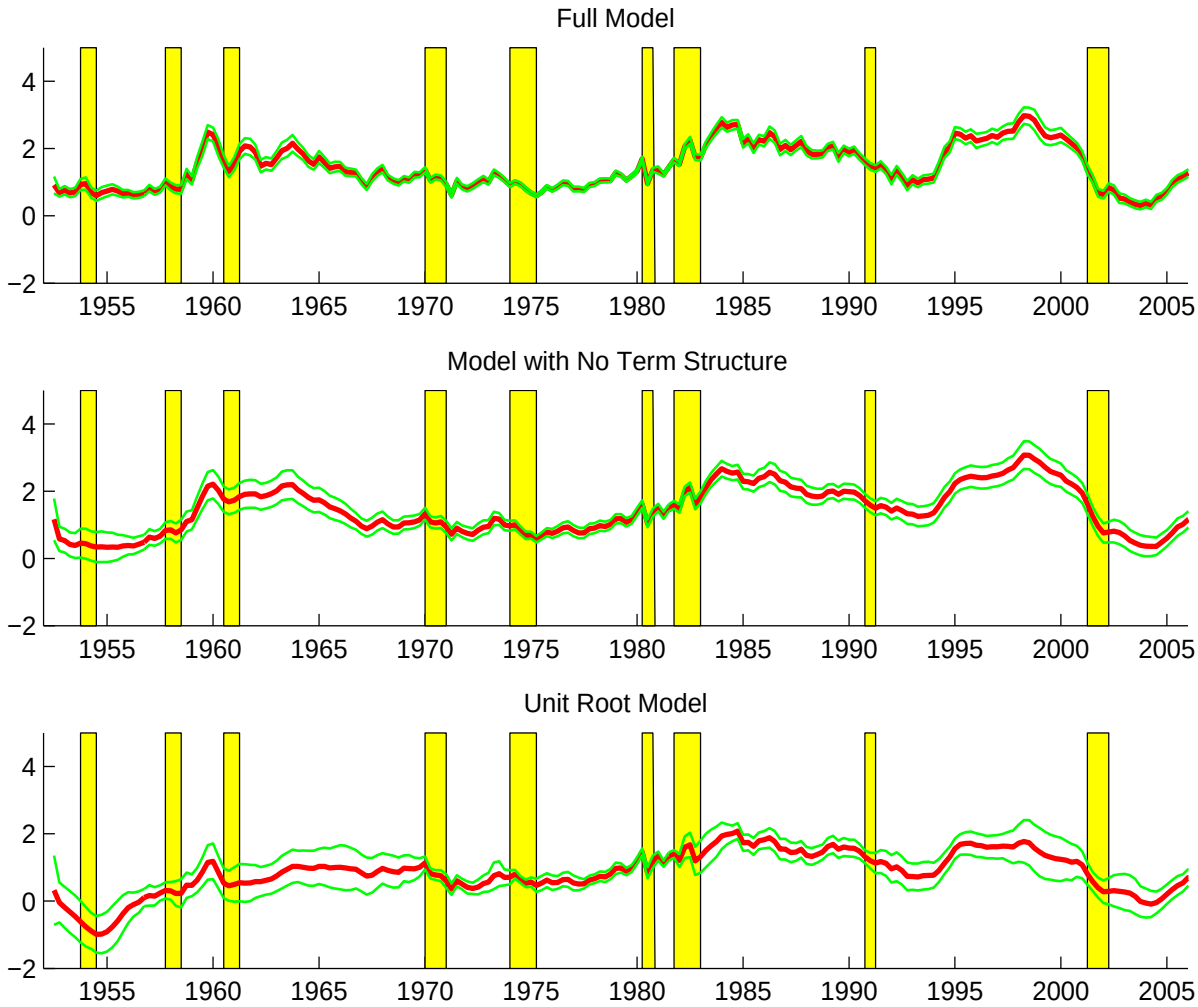
The figure plots the short rate (top panel) and the 5-year term spread (bottom panel), which is the 5-year yield minus the 3-month T-bill, from the results of a counter-factual experiment. We hold the Fed weight on inflation constant at its average level over 2000 and allow all other factors to take their sample values. We assume the posterior mean values for a_t. The figure plots the effect on the yield curve post-2001 in the dashed lines along with the actual paths of the yield curve in the solid lines. Units on the y-axis are annualized.

Figure 4: Output Gap Coefficients from Various Models



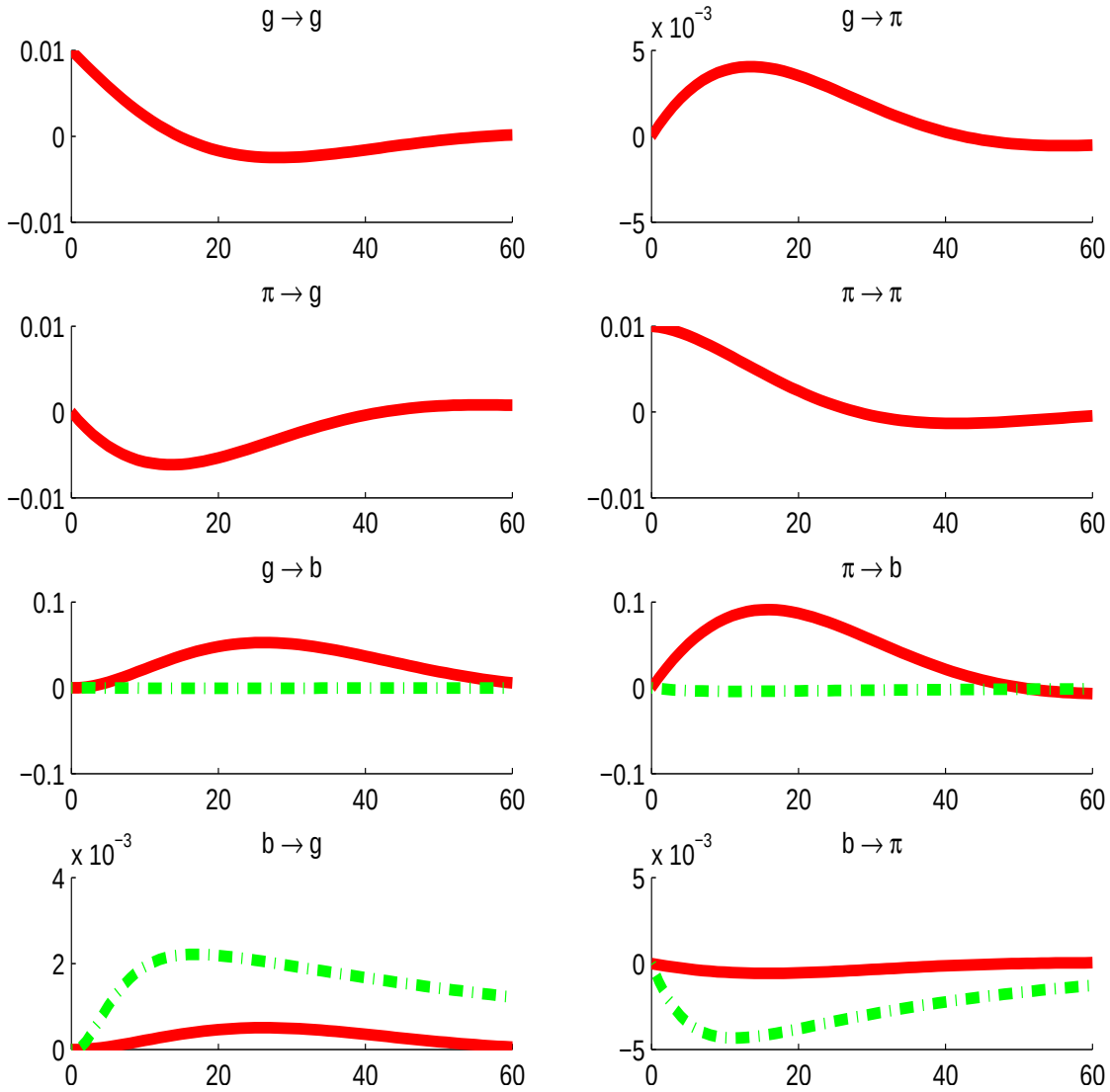
The figure plots the posterior mean of the time-varying output gap policy coefficient $\bar{a} + a_t$ implied by the full model (top panel), the model estimated without any yield curve information (middle panel), and a model where the policy coefficients follow random walks (bottom panel). In the middle panel, we use the full model holding the VAR coefficients constant at their posterior means in Table 2. In the bottom panel, we estimate the model where a_t and b_t follow random walks following equation (17). Two posterior standard deviation bands are also drawn in thin lines. We overlay the NBER recession periods in shaded bars. The sample period is from June 1952 to December 2006 and the data frequency is quarterly.

Figure 5: Inflation Coefficients from Various Models



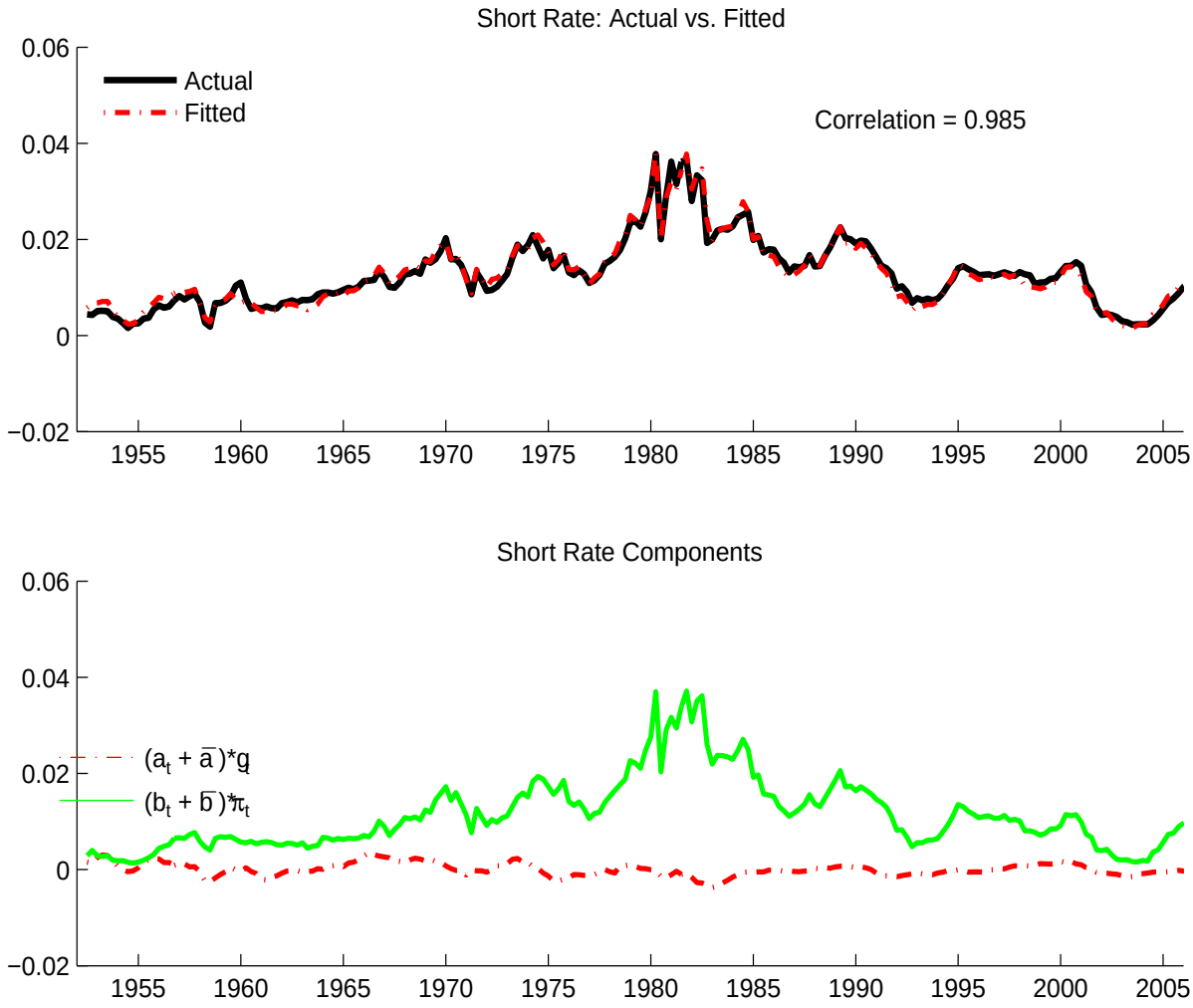
The figure plots the posterior mean of the time-varying output gap policy coefficient $\bar{b} + b_t$ implied by the full model (top panel), the model estimated without any yield curve information (middle panel), and a model where the policy coefficients follow random walks (bottom panel). In the middle panel, we use the full model holding the VAR coefficients constant at their posterior means in Table 2. In the bottom panel, we estimate the model where a_t and b_t follow random walks following equation (17). Two posterior standard deviation bands are also drawn in thin lines. We overlay the NBER recession periods in shaded bars. The sample period is from June 1952 to December 2006 and the data frequency is quarterly.

Figure 6: Impulse Responses of Factors



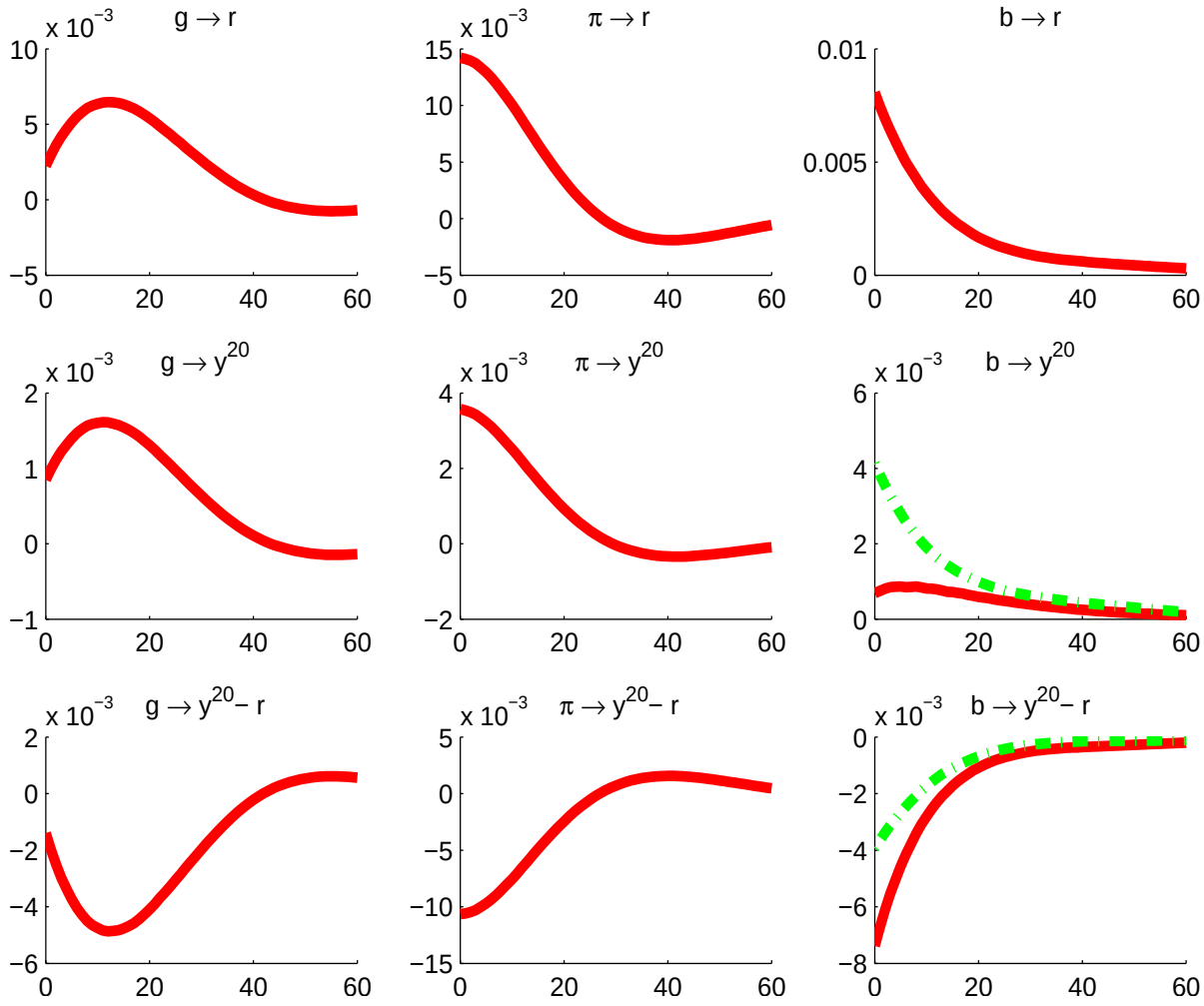
We plot selected impulse responses of the factors to each other. In the top two rows, we plot the impulse response of g_t and π_t to a 1% shock to g_t and the impulse response of g_t and π_t to a 1% shock to π_t, respectively. In the bottom two rows we graph the effect of shocks to macro factors on b_t and the effect of shocks to b_t to the macro factors. In the bottom two rows, the responses in the real measure are denoted by the red solid line and we overlay the responses under the risk-neutral measure in dashed green lines. In the third row, we consider the effect of a 1% shock to g_t and π_t on b_t. In the last row, we shock b_t by 1.00 and trace the effects on g_t and π_t. The x-axis units are quarters.

Figure 7: Components of the Short Rate



The top panel plots the short rate together with the fitted components $(\bar{a} + a_t)g_t + (\bar{b} + b_t)\pi_t$, where the policy factors a_t and b_t are evaluated at their posterior means at each observation from the Gibbs sampler. All variables are in per quarter units. The bottom panel plots each short rate component separately. The sample period is from June 1952 to December 2006 and the data frequency is quarterly.

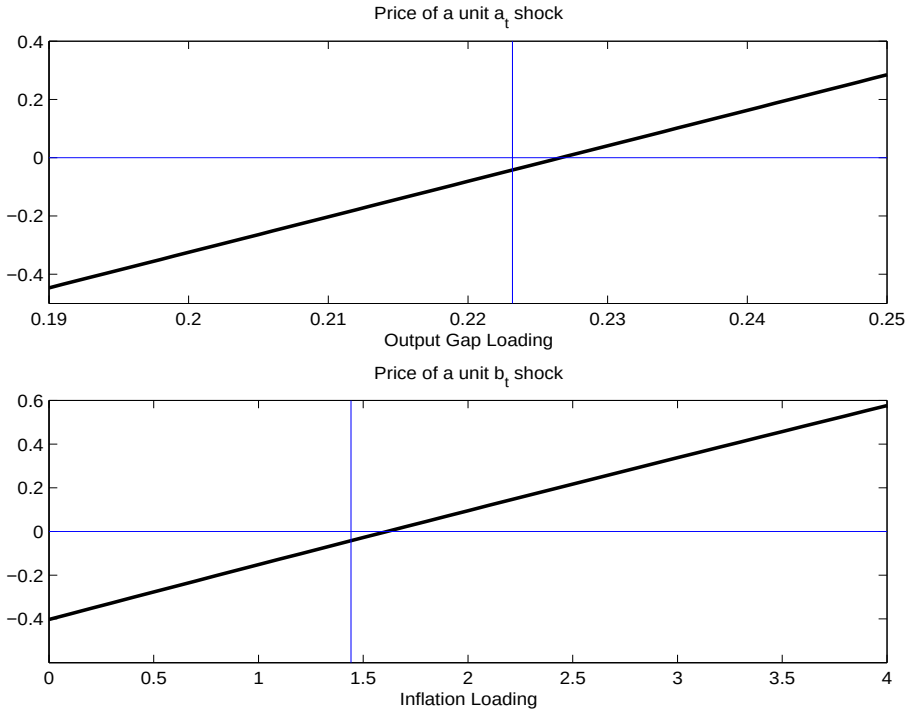
Figure 8: Yield Curve Impulse Responses to Factor Shocks



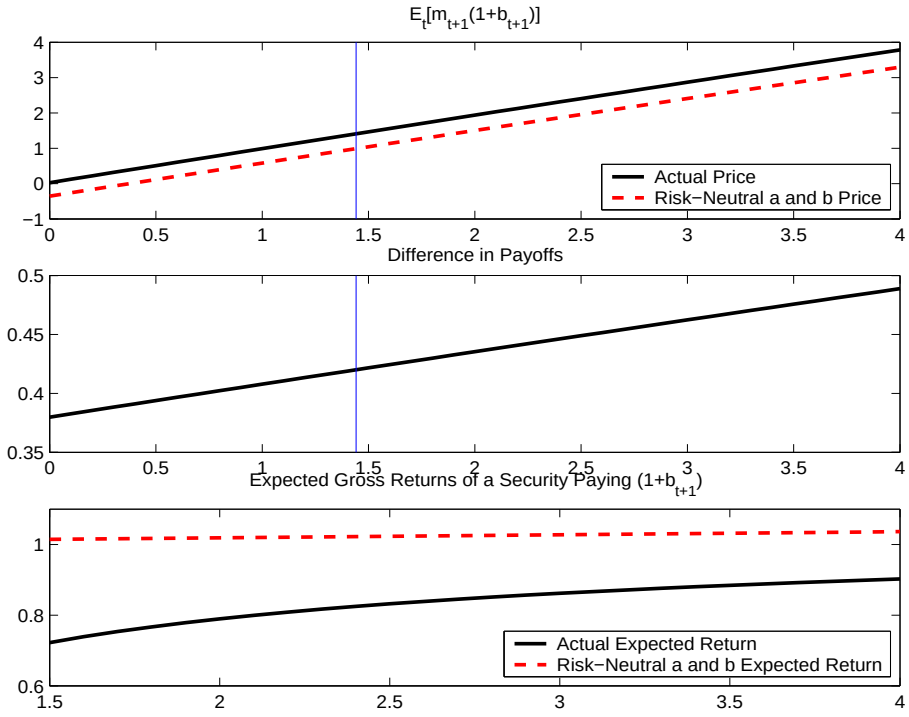
We plot the impulse responses of the short rate, r_t, the 20-quarter yield, y_t^{20}, and the yield spread, $y_t^{20} - r_t$, to a 1% shock in the output gap and inflation (g and π respectively) in the first two columns and to a 1.00 shock to b_t in the last column. For the impulse responses to y_t^{20} and $y_t^{20} - r$, we also overlay the risk-neutral impulse responses in dashed lines. We compute impulse responses following the method in Appendix C. Units on the x-axis are in quarters and the responses of yields on the y-axis are annualized.

Figure 9: The Price of Risk of Monetary Policy Shifts

Panel A: Price of Unit Shocks to a_t and b_t



Panel B: Price of a Security Paying $(1 + b_{t+1})$ at time t



Note to Figure 9
Panel A plots the price of a unit shock to a_t and b_t as a function of the monetary policy coefficients $(\bar{a} + a_t)$ and $(\bar{b} + b_t)$, respectively given in equation (19). For the top figure we hold g_t, π_t, and b_t at their population means and alter only a_t. Similarly, for the bottom figure, we hold g_t, π_t, and a_t at their population means and alter only b_t. We denote with vertical lines the steady-state value of $\bar{a} = 0.223$ and $\bar{b} = 1.442$.
In Panel B, we consider a security paying off $(1 + b_{t+1})$, which is the Fed's response to inflation. The first plot graphs the price of this security under the full model, where all risk is priced, and under the specification where a_t and b_t risk is zero. We refer to this as the "Risk-Neutral a_t and b_t Price," which is produced using risk-neutral parameters μ^{Q*} and Φ^{Q*}, where all rows and columns corresponding to a_t and b_t are set equal to their real measure counterparts in μ and Φ and the other parameters are set equal to their counterparts in μ^Q and Φ^Q. We graph the difference between the full-risk price and the risk-neutral a_t and b_t price in the second plot. In the last graph, we plot the gross expected return of the security, $\mathrm{E}_t(1 + b_{t+1})/P_t$, where P_t is the full risk price $P_t = e^{-r_t}(1 + e_4^\top(\mu^Q + \Phi^Q X_t))$ or the risk-neutral a_t and b_t price $P_t = e^{-r_t}(1 + e_4^\top(\mu^{Q*} + \Phi^{Q*} X_t))$ with $e_4 = (0\,0\,0\,1)^\top$. On each of the three figures in Panel B, we plot $(\bar{b} + b_t)$ on the x-axis. The vertical lines denote the steady-state value of $\bar{b} = 1.442$.