No-Arbitrage Taylor Rules*

Andrew Ang†
Columbia University and NBER

Sen Dong‡
Columbia University

Monika Piazzesi§
University of Chicago, NBER and CEPR

June 2005

JEL Classification: C13, E43, E52, G12

Keywords: affine term structure model,
monetary policy, interest rate risk

*We thank Ruslan Bikbov, Sebastien Blais, Dave Chapman, Mike Chernov, John Cochrane, Michael Johannes, Andy Levin, David Marshall, Thomas Philippon, Tom Sargent, Martin Schneider, and George Tauchen for helpful discussions and we especially thank Bob Hodrick for providing detailed comments. We also thank seminar participants at the American Finance Association, SFBSF Conference on Fiscal and Monetary Policy, Bank of Canada, Columbia University, European Central Bank, Federal Reserve Board of Governors, and University of Southern California for comments. Andrew Ang and Monika Piazzesi both acknowledge financial support from the National Science Foundation.

†Columbia Business School, 3022 Broadway 805 Uris, New York, NY 10027; ph: (212) 854-9154; fax: (212) 662-8474; email: aa610@columbia.edu; WWW: http://www.columbia.edu/∼aa610

‡Columbia Business School, 3022 Broadway 311 Uris, New York, NY 10027; email: sd2068@columbia.edu; WWW: http://www.columbia.edu/∼sd2068

§University of Chicago, Graduate School of Business, 5807 S. Woodlawn, Chicago, IL 60637; ph: (773) 834-3199; email: monika.piazzesi@gsb.uchicago.edu; WWW: http://gsbwww.uchicago.edu/fac/monika.piazzesi/research/

Abstract

We estimate Taylor (1993) rules and identify monetary policy shocks using no-arbitrage pricing techniques. Long-term interest rates are risk-adjusted expected values of future short rates and thus provide strong over-identifying restrictions about the policy rule used by the Federal Reserve. The no-arbitrage framework also accommodates backward-looking and forward-looking Taylor rules. We find that inflation and GDP growth account for over half of the time-variation of time-varying excess bond returns and we attribute almost all of the movements in the term spread to inflation. Taylor rules estimated with no-arbitrage restrictions differ significantly from Taylor rules estimated by OLS, and monetary policy shocks identified with no-arbitrage techniques are less volatile than their OLS counterparts.

1 Introduction

Most central banks, including the U.S. Federal Reserve (Fed), conduct monetary policy to only influence the short end of the yield curve. However, the entire yield curve responds to the actions of the Fed because long interest rates are conditional expected values of future short rates, after adjusting for risk premia. These risk-adjusted expectations of long yields are formed based on a view of how the Fed conducts monetary policy. Thus, the whole yield curve reflects the monetary actions of the Fed, so the entire term structure of interest rates can be used to estimate monetary policy rules and extract estimates of monetary policy shocks.

According to the Taylor (1993) rule, the Fed sets short interest rates by reacting to CPI inflation and the deviation of GDP from its trend. To exploit the cross-equation restrictions on yield movements implied by the assumption of no arbitrage, we place the Taylor rule in a term structure model. The no-arbitrage assumption is reasonable in a world of large investment banks and active hedge funds, who take positions eliminating arbitrage opportunities arising in bond prices that are inconsistent with each other in either the cross-section or their expected movements over time. Moreover, the absence of arbitrage is a necessary condition for an equilibrium in most macroeconomic models. Imposing no arbitrage, therefore, can be viewed as a useful first step towards a structural model.

We describe expectations of future short rates by the Taylor rule and a Vector Autoregression (VAR) for macroeconomic variables. Following the approach taken in many papers in macro (see, for example, Fuhrer and Moore, 1995), we could infer the values of long yields from these expectations by imposing the Expectations Hypothesis (EH). However, there is strong empirical evidence against the EH (see, for example, Fama and Bliss, 1987; Campbell and Shiller, 1991; Bansal, Tauchen and Zhou, 2004; Cochrane and Piazzesi, 2005, among many others). Term structure models can account for deviations from the EH by explicitly incorporating time-varying risk premia (see, for example, Fisher, 1998; Duffee, 2002; Dai and Singleton, 2002).

We develop a methodology to embed Taylor rules in an affine term structure model with time-varying risk premia. The structure accommodates standard Taylor rules, backward-looking Taylor rules that allow multiple lags of inflation and GDP growth to influence the actions of the Fed, and forward-looking Taylor rules where the Fed responds to anticipated inflation and GDP growth. The framework also accommodates monetary policy shocks that are serially correlated but uncorrelated with macro factors. The model specifies standard VAR dynamics for the macro indicators, inflation and GDP growth, together with an additional latent factor that drives interest rates. This latent factor captures other movements in yields that may be correlated with inflation and GDP growth, including monetary policy shocks. Our framework also allows risk premia to depend on the state of the macroeconomy.

By combining no-arbitrage pricing with the Fed's policy rule, we extract information from the entire term structure about monetary policy, and vice versa, use our knowledge about monetary policy to model the term structure of interest rates. The model allows us to efficiently measure how different yields respond to systematic changes in monetary policy, and how they respond to unsystematic policy shocks. Interestingly, the model implies that a large amount of interest rate volatility is explained by systematic changes in policy that can be traced back to movements in macro variables. For example, 41% of the variance of the 1-quarter yield and 33% of the variance of the 5-year yield can be attributed to movements in inflation and GDP growth. Over 95% of the variance in the 5-year term spread is due to time-varying inflation and inflation risk. The estimated model also captures the counter-cyclical properties of time-varying expected excess returns on bonds.

To estimate the model, we use Bayesian techniques that allow us to estimate flexible dynamics and extract estimates of latent monetary policy shocks. Existing papers that incorporate macro variables into term structure models make strong – and often arbitrary – restrictions on the VAR dynamics, risk premia, and measurement errors. For example, Ang and Piazzesi (2003) assume that macro dynamics do not depend on interest rates. Dewachter and Lyrio (2004), and Rudebusch and Wu (2005), among others, set arbitrary risk premia parameters to zero. Hördahl, Tristani, and Vestin (2005), Rudebusch and Wu (2005), and Ang, Piazzesi, and Wei (2005) assume that only certain yields are measured with error, while others are observed without error. These restrictions are not motivated from economic theory, but are only made for reasons of econometric tractability. In contrast, we do not impose these restrictions and find that the added flexibility helps the performance of the model.

We estimate Taylor rules following the large macro literature that uses low frequencies (we use quarterly data) at which GDP and inflation are reported. Under the cross-equation restrictions for yields implied by the no-arbitrage model, we estimate a flexible specification where the macro and latent factors can Granger-cause each other as well as allow the macro and latent factors to be conditionally correlated. Other models in the macroeconomics literature are more restrictive. One standard approach is to assume that the Fed either ignores information from the bond market (Evans and Marshall 1998, 2001), or that bond markets ignore Fed announcements (Bagliano and Favero 1998). Both assumptions seem unappealing. Another approach take by Christiano, Eichenbaum, and Evans (1996), Evans and Marshall (1998, 2001), and many others, is to assume that macroeconomic variables these macroeconomic variables react only slowly – not within the same quarter – to monetary policy shocks. While this assumption seems reasonable, these authors must make additional arbitrary assumptions to handle bond yields within their VAR systems. Our strategy is to assume that the Fed adopts a Taylor rule and thus only cares about inflation and output movements. The Fed looks at current yield data, but only because current yields contain information about future values of these macro variables.[1]

Our paper is related to a growing literature on linking the dynamics of the term structure with macro factors. However, the other papers in this literature are less interested in estimating various Taylor rules, rather than embedding a particular form of a Taylor rule, sometimes pre-estimated, in a macroeconomic model. For example, Bekaert, Cho, and Moreno (2003), Gallmeyer, Hollifield, and Zin (2005), Hördahl, Tristani, and Vestin (2005), and Rudebusch and Wu (2005) estimate structural models with interest rates and macro variables. In contrast to these studies, we do not impose any structure in addition to the assumption of no arbitrage, which makes our approach more closely related to the identified VAR literature in macroeconomics (for a survey, see Christiano, Eichenbaum and Evans, 1999) and gives us additional flexibility in matching the dynamics of the term structure. Bernanke, Boivin and Eliasz (2004), and Diebold, Rudebusch, and Aruoba (2005) estimate latent factor models with macro variables, but they do not preclude no-arbitrage movements of bond yields. Dai and Philippon (2005) examine the effect of fiscal shocks on yields with a term structure model, whereas our focus is embedding monetary policy rules into a no-arbitrage model.

We do not claim that our new no-arbitrage techniques are superior to estimating monetary policy rules using structural models (see, among others, Bernanke and Mihov, 1998) or using real-time information sets like central bank forecasts to control for the endogenous effects of monetary policy taken in response to current economic conditions (see, for example, Romer and Romer, 2004). Rather, we believe that estimating policy rules using no-arbitrage restrictions are a useful addition to existing methods. Our framework enables the entire cross-section and time-series of yields to be modeled and provides a unifying framework to jointly estimate standard, backward-, and forward-looking Taylor rules in a single, consistent framework. Indeed, we show that many formulations of policy rules imply term structure dynamics that are observationally equivalent. Naturally, our methodology can be used in more structural approaches that effectively constrain the factor dynamics and risk premia, and we

[1] An alternative, high-frequency, identification approach is taken by Piazzesi (2005). By assuming that the Fed reacts to information available right before its policy decision, she identifies the unexpected change in the target as the monetary policy shock and the expected target as the policy rule.

can extend our set of instruments to include richer information sets. We intentionally focus on the most parsimonious set-up where Taylor rules can be identified in a no-arbitrage model.

The rest of the paper is organized as follows. Section 2 outlines the model and develops the methodology showing how Taylor rules can be identified with no-arbitrage conditions. We briefly discuss the estimation strategy in Section 3. In Section 4, we lay out the empirical results. After describing the parameter estimates, we attribute the time-variation of yields and expected excess holding period returns of long-term bonds to economic sources. We describe in detail the implied Taylor rule estimates from the model and contrast them with OLS estimates. We compare the no-arbitrage monetary policy shocks and impulse response functions with traditional VAR and other identification approaches. Section 5 concludes.

2 The Model

We describe the setup of the model in Section 2.1. Section 2.2 derives closed-form solutions for bond prices (yields) and expected returns. In Sections 2.3 to 2.8, we explain how to incorporate various Taylor rules in our setup.

2.1 General Set-up

We denote the 3×1 vector of state variables as

$$X_t = [g_t \;\; \pi_t \;\; f_t^u]^\top ,$$

where g_t is quarterly GDP growth from $t-1$ to t, π_t is the quarterly inflation rate from $t-1$ to t, and f_t^u is a latent term structure state variable. Both GDP growth and inflation are continuously compounded. We use one latent state variable because this is the most parsimonious set-up with Taylor rule residuals (as the next section makes clear). The latent factor, f_t^u, is a standard latent term structure factor in the tradition of the affine term structure literature. However, we show below that this factor can be interpreted as a transformation of policy actions taken by the Fed on the short rate.

We specify that X_t follows a VAR(1):

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

where $\varepsilon_t \sim$ IID $N(0, I)$. The short rate is given by:

$$r_t = \delta_0 + \delta_1^\top X_t, \tag{2}$$

for δ_0 a scalar and δ_1 a 3×1 vector. To complete the model, we specify the pricing kernel to take the standard form:

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1}\right), \tag{3}$$

with the time-varying prices of risk:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{4}$$

for the 3×1 vector λ_0 and the 3×3 matrix λ_1. The pricing kernel prices all assets in the economy, which are zero-coupon bonds, from the recursive relation:

$$P_t^{(n)} = \mathrm{E}_t[m_{t+1} P_{t+1}^{(n-1)}],$$

where $P_t^{(n)}$ 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:

$$P_t^{(n)} = \mathrm{E}_t^Q\left[\exp\left(-\sum_{i=0}^{n-1} r_{t+i}\right)\right],$$

where E_t^Q denotes the expectation under the risk-neutral probability measure, under which the dynamics of the state vector X_t are characterized by the risk-neutral constant and autocorrelation matrix:

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

If investors are risk-neutral, $\lambda_0 = 0$ and $\lambda_1 = 0$, and no risk adjustment is necessary.

This model belongs to the Duffie and Kan (1996) affine class of term structure models, but uses both latent and observable macro factors. The affine prices of risk specification in equation (4) has been used by, among others, Constantinides (1992), Fisher (1998), Duffee (2002), Dai and Singleton (2002), and Brandt and Chapman (2003) in continuous time and by Ang and Piazzesi (2003), Ang, Piazzesi, and Wei (2005), and Dai and Philippon (2005) in discrete time. As Dai and Singleton (2002) demonstrate, the flexible affine price of risk specification is able to capture patterns of expected holding period returns on bonds that we observe in the data.

2.2 Bond Prices and Expected Returns

Ang and Piazzesi (2003) show that the model (1)-(4) implies that bond yields take the form:

$$y_t^{(n)} = a_n + b_n^\top X_t, \tag{5}$$

where $y_t^{(n)}$ is the yield on an n-period zero coupon bond at time t that is implied by the model, which satisfies $P_t^{(n)} = \exp(-ny_t^{(n)})$.

The scalar a_n and the 3×1 vector b_n are given by $a_n = -A_n/n$ and $b_n = -B_n/n$, where A_n and B_n satisfy the recursive relations:

$$\begin{aligned} A_{n+1} &= A_n + B_n^\top(\mu - \Sigma\lambda_0) + \frac{1}{2}B_n^\top \Sigma\Sigma^\top B_n - \delta_0 \\ B_{n+1}^\top &= B_n^\top(\Phi - \Sigma\lambda_1) - \delta_1^\top, \end{aligned} \tag{6}$$

where $A_1 = -\delta_0$ and $B_1 = -\delta_1$. In terms of notation, the one-period yield $y_t^{(1)}$ is the same as the short rate r_t in equation (2).

Since yields take an affine form and the conditional mean of the state vector is affine, expected holding period returns on bonds are also affine in X_t. We define the one-period excess holding period return as:

$$\begin{aligned} rx_{t+1}^{(n)} &= \log\left(\frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}\right) - r_t \\ &= ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t. \end{aligned} \tag{7}$$

The conditional expected excess holding period return can be computed using:

$$\begin{aligned} \mathrm{E}_t[rx_{t+1}^{(n)}] &= -\frac{1}{2}B_{n-1}^\top \Sigma\Sigma^\top B_{n-1} + B_{n-1}^\top \Sigma\lambda_0 + B_{n-1}^\top \Sigma\lambda_1 X_t \\ &\equiv A_n^x + B_n^{x\top} X_t, \end{aligned} \tag{8}$$

which again takes an affine form for the scalar $A_n^x = -\frac{1}{2}B_{n-1}^\top \Sigma\Sigma^\top B_{n-1} + B_{n-1}^\top \Sigma\lambda_0$ and the 3×1 vector $B_n^x = \lambda_1 \Sigma^\top B_{n-1}$. From equation (8), we can see directly that the expected excess return comprises three terms: (i) a Jensen's inequality term, (ii) a constant risk premium, and (iii) a time-varying risk premium. The time variation is governed by the parameters in the matrix λ_1. Since both bond yields and the expected holding period returns of bonds are affine functions of X_t, we can easily compute variance decompositions following standard VAR methods.

2.3 The Benchmark Taylor Rule

The Taylor (1993) rule captures the notion that the Fed adjusts short-term interest rates in response to movements in inflation and real activity. The rule is consistent with a monetary authority that minimizes a quadratic loss function that tries to stabilize inflation and output around a long-run inflation target and the natural output rate (see, for example, Svensson 1997). Following Taylor's original specification, we define the benchmark Taylor rule to be:

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi} \pi_t + \varepsilon_t^{MP,T}, \tag{9}$$

where the short rate is set by the Federal Reserve in response to current output and inflation. The basic Taylor rule (9) can be interpreted as the short rate equation (2) in a standard affine term structure model, where the unobserved monetary policy shock $\varepsilon_t^{MP,T}$ corresponds to a latent term structure factor, $\varepsilon_t^{MP,T} = \gamma_{1,u} f_t^u$. This corresponds to the short rate equation (2) in the term structure model with $[\delta_{1,g} \;\; \delta_{1,\pi} \;\; \delta_{1,u}] = [\gamma_{1,g} \;\; \gamma_{1,\pi} \;\; \gamma_{1,u}]$.

The Taylor rule (9) can be estimated consistently using OLS under the assumption that $\varepsilon_t^{MP,T}$, or f_t^u, is contemporaneously uncorrelated with GDP growth and inflation. This assumption is satisfied, if GDP and inflation only react slowly to policy shocks. However, there are several advantages to estimating the policy coefficients, $\gamma_{1,g}$ and $\gamma_{1,\pi}$, and extracting the monetary policy shock, $\varepsilon_t^{MP,T}$, using no-arbitrage restrictions rather than simply running OLS on equation (9). First, no-arbitrage restrictions can free up the contemporaneous correlation between the macro and latent factors. Second, even if the macro and latent factors are conditionally uncorrelated, OLS is consistent but not efficient. By imposing no arbitrage, we impose additional restrictions that produce more efficient estimates by exploiting information contained in the whole term structure in the estimation of the Taylor rule coefficients, while OLS only uses data on the short rate. Third, the term structure model provides estimates of the effect of a policy or macro shock on any segment of the yield curve, which an OLS estimation of equation (9) cannot provide. Finally, our term structure model allows us to trace the predictability of risk premia in bond yields to macroeconomic or monetary policy sources.

The Taylor rule in equation (9) does not depend on the past level of the short rate. Therefore, OLS regressions typically find that the implied series of monetary policy shocks from the benchmark Taylor rule, $\varepsilon_t^{MP,T}$, is highly persistent (see, for example, Rudebusch and Svensson, 1999). The statistical reason for this finding is that the short rate is highly autocorrelated, and its movements are not well explained by the right-hand side variables in equation (9). This causes the implied shock to inherit the dynamics of the level of the persistent short rate. In affine term structure models, this finding is reflected by the properties of the implied latent variables. In particular, $\varepsilon_t^{MP,T}$ corresponds to $\delta_{1,u} f_t^u$, which is the scaled latent term structure variable. Ang and Piazzesi (2003) show that the first latent factor implied by an affine model with both latent factors and observable macro factors closely corresponds to the traditional first, highly persistent, latent factor in term structure models with only unobservable factors. This latent variable also corresponds closely to the first principal component of yields, or the average level of the yield curve, which is highly autocorrelated.

2.4 Backward-Looking Taylor Rules

Eichenbaum and Evans (1995), Christiano, Eichenbaum, and Evans (1996), Clarida, Galí, and Gertler (1998), among others, consider modified Taylor rules that include current as well as lagged values of macro variables and the previous short rate:

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi}\pi_t + \gamma_{2,g} g_{t-1} + \gamma_{2,\pi}\pi_{t-1} + \gamma_{2,r} r_{t-1} + \varepsilon_t^{MP,B}, \tag{10}$$

where $\varepsilon_t^{MP,B}$ is the implied monetary policy shock from the backward-looking Taylor rule. This formulation has the statistical advantage that we compute monetary policy shocks recognizing that the short rate is a highly persistent process. The economic mechanism behind such a backward-looking rule may be that the objective of the central bank is to smooth interest rates (see Goodfriend, 1991).

In the setting of our model, we can modify the short rate equation (2) to take the same form as equation (10). Collecting the macro factors g_t and π_t into a vector of observable variables $f_t^o = [g_t \;\; \pi_t]^\top$, we can rewrite the short rate dynamics (2) as:

$$r_t = \delta_0 + \delta_{1,o}^\top f_t^o + \delta_{1,u} f_t^u, \tag{11}$$

where we decompose the vector δ_1 into $\delta_1 = \left[\delta_{1,o}^\top \;\; \delta_{1,u}\right]^\top = [\delta_{1,g} \;\; \delta_{1,\pi} \;\; \delta_{1,u}]^\top$. We also rewrite the dynamics of $X_t = \left[f_t^{o\top} \;\; f_t^u\right]^\top$ in equation (1) as:

$$\begin{pmatrix} f_t^o \\ f_t^u \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} f_{t-1}^o \\ f_{t-1}^u \end{pmatrix} + \begin{pmatrix} v_t^o \\ v_t^u \end{pmatrix}, \tag{12}$$

where $v_t = (v_t^{o\top} \; v_t^u)^\top \sim$ IID $N(0, \Sigma\Sigma^\top)$. Equation (12) is equivalent to equation (1), but the notation in equation (12) separates the dynamics of the macro variables, f_t^o, from the dynamics of the latent factor, f_t^u.

Using equation (12), we can substitute for f_t^u in equation (11) to obtain:

$$r_t = (1 - \Phi_{22})\delta_0 + \delta_{1,u}\mu_2 + \delta_{1,o}^\top f_t^o + (\delta_{1,u}\Phi_{21}^\top - \delta_{1,o}\Phi_{22}^\top)^\top f_{t-1}^o + \Phi_{22} r_{t-1} + \varepsilon_t^{MP,B}, \tag{13}$$

where we substitute for the dynamics of f_t^u and define the backward-looking monetary policy shock to be $\varepsilon_t^{MP,B} \equiv \delta_{1,u} v_t^u$. Equation (13) expresses the short rate as a function of current and lagged macro factors, f_t^o and f_{t-1}^o, the lagged short rate, r_{t-1}, and a monetary policy shock $\varepsilon_t^{MP,B}$.

Interestingly, the response to contemporaneous GDP and inflation captured by the $\delta_{1,o}$ coefficient on f_t^o in the backward-looking Taylor rule (13) is identical to the response in the benchmark Taylor rule (9), because the $\delta_{1,o}$ coefficient is unchanged. The intuition behind this result is that the short rate equation (2) describes the response of the short rate to current macro factors. The latent factor, however, contains a predictable component that depends on past values of the short rate and the macro factors. The backward-looking Taylor rule makes this dependence explicit. Importantly, the backward-looking Taylor rule in equation (13) and the benchmark Taylor rule (9) lead to observationally equivalent reduced-form dynamics for interest rates and macro variables.

The implied monetary policy shocks from the backward-looking Taylor rule, $\varepsilon_t^{MP,B}$, are potentially very different from the benchmark shocks, $\varepsilon_t^{MP,T}$. In the no-arbitrage model, the backward-looking monetary policy shock $\varepsilon_t^{MP,B}$ is identified as the scaled shock to the latent term structure factor, $\delta_{1,u}v_t^u$. In the set-up of the factor dynamics in equation (1), or equivalently equation (12), the v_t^u shocks are IID. In comparison, the shocks in the standard Taylor rule (9), $\varepsilon_t^{MP,T}$ are highly autocorrelated. Note that the coefficients on lagged macro variables in the extended Taylor rule (13) are equal to zero only if $\delta_{1,u}\Phi_{21}^\top = \delta_{1,o}\Phi_{22}^\top$. Under this restriction, the combined movements of the past macro factors must exactly offset the movements in the lagged term structure latent factor so that the short rate is affected only by unpredictable shocks.

Once our model is estimated, we can easily back out the implied extended Taylor rule (10) from the estimated coefficients. This is done by using the implied dynamics of f_t^u in the factor dynamics (12):

$$v_t^u = f_t^u - \mu_2 - \Phi_{21}f_{t-1}^o - \Phi_{22}f_{t-1}^u.$$

Again, if $\varepsilon_t^{MP,B} = \delta_{1,u}v_t^u$ is unconditionally correlated with the shocks to the macro factors f_t^o, then OLS does not provide efficient estimates of the monetary policy rule, and may provide biased estimates of the Taylor rule in small samples.

2.5 Taylor Rules with Serially Correlated Policy Shocks

Backward-looking Taylor rules are observationally equivalent to a policy rule where the Fed reacts to the entire history of macro variables, but with serially correlated errors. To see this, we recursively substitute for r_{t-j}, $j \geq 1$, in equation (13) and obtain

$$r_t = c_t + \Psi_t(L)f_t^o + \varepsilon_t^{MP,AR}, \tag{14}$$

where c_t is constant, $\Psi_t(L)$ is a polynomial of lag operators, and $\varepsilon_t^{MP,AR}$ is a serially correlated shock. The variables c_t, $\Psi_t(L)$, and $\varepsilon_t^{MP,AR}$ are given by:

$$\begin{aligned} c_t &= \delta_0 + \delta_{1,u} \sum_{i=0}^{t-2} \Phi_{22}^i \mu, \\ \Psi_t(L) &= \delta_{1,0} + \delta_{1,u} \sum_{i=0}^{t-2} \Phi_{22}^i \Phi_{21} L^{i+1}, \\ \varepsilon_t^{MP,AR} &= \sum_{i=0}^{t-1} \Phi_{22}^i \delta_{1,u} v_{t-i}^u, \end{aligned}$$

where v_t^u are the innovations to the latent factor in the VAR (12). The shock $\varepsilon_t^{MP,AR}$ is orthogonal to the macro variables, f^o, and follows an AR(1) process: $\varepsilon_t^{MP,AR} = \Phi_{22}\varepsilon_{t-1}^{MP,AR} + \delta_{1,u} v_t^u$. Whereas in the backward-looking Taylor rule (13), the policy shocks are scaled innovations of the latent factor, $\varepsilon_t^{MP,B} = \delta_{1,u} v_t^u$, the autocorrelated policy errors $\varepsilon_t^{MP,AR}$ are linear combinations of current and past latent factor innovations in equation (14).[2]

2.6 Forward-Looking Taylor Rules

Finite-Horizon, Forward-Looking Taylor Rules

Clarida and Gertler (1997) and Clarida, Galí and Gertler (2000), among others, propose a forward-looking Taylor rule, where the Fed sets interest rates based on expected future GDP growth and expected future inflation over the next few quarters. For example, a forward-looking Taylor rule using expected GDP growth and inflation over the next quarter takes the form:

$$r_t = \gamma_0 + \gamma_{1,g} \mathrm{E}_t(g_{t+1}) + \gamma_{1,\pi} \mathrm{E}_t(\pi_{t+1}) + \varepsilon_t^{MP,F}, \tag{15}$$

where we define $\varepsilon_t^{MP,F}$ to be the forward-looking Taylor rule monetary policy shock.

We can map the forward-looking Taylor rule (15) into the framework of an affine term structure model as follows. The conditional expectations of future GDP growth and inflation are simply a function of current X_t that can be computed from the state dynamics (1):

$$\mathrm{E}_t(X_{t+1}) = \mu + \Phi X_t.$$

[2] Bikbov and Chernov (2005) use a projection procedure to also decompose latent factors into a macro-related component and an innovation component with different statistical properties that can apply to models with more than one latent factor.

Denoting e_i as a vector of zeros with a one in the ith position, we can write equation (15) as:

$$r_t = \gamma_0 + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)^\top \mu + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)^\top \Phi X_t + \varepsilon_t^{MP,F}, \tag{16}$$

as g_t and π_t are ordered as the first and second elements in X_t.

Equation (16) is an affine short rate equation where the short rate coefficients are a function of the parameters of the dynamics of X_t:

$$r_t = \bar{\delta}_0 + \bar{\delta}_1^\top X_t, \tag{17}$$

where

$$\begin{aligned}\bar{\delta}_0 &= \gamma_0 + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)^\top \mu \\ \bar{\delta}_1^\top &= \Phi^\top (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2) + \gamma_{1,u}e_3^\top,\end{aligned}$$

and $\varepsilon_t^{MP,F} \equiv \gamma_{1,u} f_t^u$ with $\gamma_{1,u} = \delta_{1,u}$. Hence, we can identify a forward-looking Taylor rule by redefining the bond price recursions in equation (6) in terms of the new $\bar{\delta}_0$ and $\bar{\delta}_1$ coefficients. The complete term structure model is defined by the same set-up as equations (1)-(4), except we use the new short rate equation (17) that embodies the forward-looking structure in place of the basic short rate equation (2). The relations involving γ_0, γ_1, μ, and Φ in equation (17) show that the forward-looking Taylor rule imposes restrictions on the parameters of an affine term structure model.

The new no-arbitrage bond recursions using the restricted coefficients $\bar{\delta}_0$ and $\bar{\delta}_1$ reflect the conditional expectations of GDP and inflation that enter in the short rate equation (17). Furthermore, the conditional expectations $\mathrm{E}_t(g_{t+1})$ and $\mathrm{E}_t(\pi_{t+1})$ are those implied by the underlying dynamics of g_t and π_t in the VAR process (1). Other approaches, like Rudebusch and Wu (2005), specify the future expectations of macro variables entering the short rate equation in a manner not necessarily consistent with the underlying dynamics of the macro variables. The monetary policy shocks in the forward-looking Taylor rule (15) or (16), $\varepsilon_t^{MP,F}$, can be consistently estimated by OLS only if f_t^u is orthogonal to the dynamics of g_t and π_t.

Since k-period ahead conditional expectations of GDP and inflation remain affine functions of the current state variables X_t, we can also specify a more general forward-looking Taylor rule based on expected GDP or inflation over the next k quarters:

$$r_t = \gamma_0 + \gamma_{1,g}\mathrm{E}_t(g_{t+k,k}) + \gamma_{1,\pi}\mathrm{E}_t(\pi_{t+k,k}) + \varepsilon_t^{MP,F}, \tag{18}$$

where $g_{t+k,k}$ and $\pi_{t+k,k}$ represent GDP growth and inflation over the next k periods:

$$g_{t+k,k} = \frac{1}{k}\sum_{i=1}^{k} g_{t+i} \quad \text{and} \quad \pi_{t+k,k} = \frac{1}{k}\sum_{i=1}^{k} \pi_{t+i}.$$

The forward-looking Taylor rule monetary policy shock $\varepsilon_t^{MP,F}$ is the scaled latent term structure factor, $\varepsilon_t^{MP,F} = \gamma_{1,u} f_t^u$. As Clarida, Galí and Gertler (2000) note, the general case (18) also nests the benchmark Taylor rule (9) as a special case by setting $k = 0$. Appendix A details the appropriate transformations required to map equation (18) into an affine term structure model and discusses the estimation procedure for a forward-looking Taylor rule based on a k-quarter horizon.

Infinite-Horizon, Forward-Looking Taylor Rules

An alternative approach to fixing some forecasting horizon k is to view the Fed as discounting the entire expected path of future economic conditions. For simplicity, we assume the Fed discounts both expected future GDP growth and expected future inflation at the same discount rate, β. In this formulation, the forward-looking Taylor rule takes the form:

$$r_t = \gamma_0 + \gamma_{1,\hat{g}} \hat{g}_t + \gamma_{1,\hat{\pi}} \hat{\pi}_t + \varepsilon_t^{MP,F}, \tag{19}$$

where $\hat{g}_t$ and $\hat{\pi}_t$ are infinite sums of expected future GDP growth and inflation, respectively, both discounted at rate β per period. Many papers have set β at one, or very close to one, sometimes motivated by calibrating it to an average real interest rate (see Salemi, 1995; Rudebusch and Svenson, 1999; Favero and Rovelli, 2003; Collins and Siklos, 2004).

We can estimate the discount rate β as part of a standard term structure model by using the dynamics of X_t in equation (1) to write $\hat{g}_t$ as:

$$\begin{aligned}
\hat{g}_t &= \sum_{i=0}^{\infty} \beta^i e_1^\top \mathrm{E}_t(X_{t+i}) \\
&= e_1^\top (X_t + \beta\mu + \beta\Phi X_t + \beta^2 (I+\Phi)\mu + \beta^2 \Phi^2 X_t + \cdots) \\
&= e_1^\top (\mu\beta + (I+\Phi)\mu\beta^2 + \cdots) + e_1^\top (I + \Phi\beta + \Phi^2\beta^2 + \cdots) X_t \\
&= \frac{\beta}{(1-\beta)} e_1^\top (I - \Phi\beta)^{-1}\mu + e_1^\top (I-\Phi\beta)^{-1} X_t.
\end{aligned}$$

Similarly, we can also write discounted future inflation, $\hat{\pi}_t$, in a similar fashion as:

$$\hat{\pi}_t = \frac{\beta}{(1-\beta)} e_2^\top (I-\Phi\beta)^{-1}\mu + e_2^\top (I-\Phi\beta)^{-1} X_t.$$

To place the forward-looking rule with discounting in a term structure model, we re-write the short rate equation (2) as:

$$r_t = \hat{\delta}_0 + \hat{\delta}_1^\top X_t, \tag{20}$$

where

$$\hat{\delta}_0 = \gamma_0 + [\gamma_{1,g} e_1 \;\; \gamma_{1,\pi} e_2]^\top \left(\frac{\beta}{(1-\beta)} (I - \Phi\beta)^{-1} \mu \right),$$
$$\hat{\delta}_1^\top = [\gamma_{1,g} e_1 \;\; \gamma_{1,\pi} e_2]^\top (I - \Phi\beta)^{-1} + \gamma_{1,u}\, e_3^\top.$$

Similarly, we modify the bond price recursions for the standard affine model in equation (6) by using the new $\hat{\delta}_0$ and $\hat{\delta}_1$ coefficients that embody restrictions on β, γ_0, γ_1, μ, and Φ.

2.7 Forward- and Backward-Looking Taylor Rules

As a final case, we combine the forward- and backward-looking Taylor rules, so that the monetary policy rule is computed taking into account forward-looking expectations of macro variables, lagged realizations of macro variables, while also controlling for lagged short rates. We illustrate the rule considering expectations for inflation and GDP over the next quarter ($k = 1$), but similar rules apply for other horizons.

We start with the standard forward-looking Taylor rule in equation (15):

$$r_t = \gamma_0 + \gamma_{1,o}^\top \mathrm{E}_t(f_{t+1}^o) + \varepsilon_t^{MP,F},$$

where $\mathrm{E}_t(f_{t+1}^o) = [\mathrm{E}_t(g_{t+1}) \; \mathrm{E}_t(\pi_{t+1})]^\top$ and $\varepsilon_t^{MP,F} = \gamma_{1,u} f_t^u$. We substitute for f_t^u using equation (12) to obtain:

$$\begin{aligned} r_t &= \gamma_0 + \gamma_{1,u}\mu_2 - \frac{\gamma_{1,u}\Phi_{22}\delta_0}{\delta_{1,u}} + \gamma_{1,o}^\top \mathrm{E}_t(f_{t+1}^o) \\ &+ \left(\gamma_{1,u}\Phi_{21}^\top - \frac{\gamma_{1,u}\Phi_{22}\delta_{1,o}}{\delta_{1,u}} \right)^\top f_{t-1}^o + \frac{\gamma_{1,u}\Phi_{22}}{\delta_{1,u}} r_{t-1} + \varepsilon_t^{MP,FB}. \end{aligned} \tag{21}$$

Equation (21) expresses the short rate as a function of both expected future macro factors and lagged macro factors, the lagged short rate, r_{t-1}, and a forward- and backward-looking monetary policy shock, $\varepsilon_t^{MP,FB} = \gamma_{1,u} v_t^u$. The forward- and backward-looking Taylor rule (21) is an equivalent representation of the forward-looking Taylor rule in (15). Hence, similar to how the coefficients on contemporaneous macro variables in the backward-looking Taylor rule (13) are identical to the coefficients in the benchmark Taylor rule (9), the coefficients $\delta_{1,o}$ on future expected macro variables are exactly the same as the coefficients in the forward-looking Taylor rule.

2.8 Summary of Taylor Rules

The no-arbitrage framework is able to estimate several structural Taylor rule specifications from the same reduced-form term structure model. Table 1 summarizes the various specifications. The benchmark, backward-looking Taylor rules, and the Taylor rule with serially correlated shocks are different structural rules that give rise to the same term structure dynamics. Similarly, the forward-looking and the backward- and forward-looking Taylor rules produce observationally equivalent term structure models. In all cases, the monetary policy shocks are transformations of either levels or innovations of the latent term structure variable. Finally, the last column of Table 1 reports if the no-arbitrage model requires additional restrictions. Both the forward-looking specifications require parameter restrictions in the short rate equation to ensure that we compute the expectations of the macro variables consistent with the dynamics of the VAR.

3 Data and Econometric Methodology

The objective of this section is to briefly discuss the data and the econometric methodology used to estimate the model. We relegate all technical issues to Appendix C.

3.1 Data

To estimate the model, we use continuously compounded yields of maturities 1, 4, 8, 12, 16, and 20 quarters, at a quarterly frequency. The bond yields of one year maturity and longer are from the CRSP Fama-Bliss discount bond files, while the short rate (one-quarter maturity) is taken from the CRSP Fama risk-free rate file. The sample period is June 1952 to December 2002. The consumer price index and real GDP numbers are taken from the Federal Reserve Database (FRED) at Saint Louis.

3.2 Estimation and Identification

We estimate the term structure model using Markov Chain Monte Carlo (MCMC) and Gibbs sampling methods. There are two main reasons why we choose to use a Bayesian estimation approach. First, we can avoid the usual assumption that some (arbitrary) yields are observed without any measurement error, while other yields are observed with error (going back to

Chen and Scott, 1993). This assumption is clearly ad hoc. Instead, we assume that all yields are observed with error, so that the equation for each yield is:

$$\hat{y}_t^{(n)} = y_t^{(n)} + \eta_t^{(n)}, \tag{22}$$

where $y_t^{(n)}$ is the model-implied yield from equation (5) and $\eta_t^{(n)}$ is the zero-mean observation error is IID across time and yields. We specify $\eta_t^{(n)}$ to be normally distributed and denote the standard deviation of the error term as $\sigma_\eta^{(n)}$.

Second, the Bayesian estimation method provides a posterior distribution of the time-series path of f_t^u and monetary policy shocks. That is, we can compute the mean of the posterior distribution of the time-series of f_t^u through the sample, and, consequently, we can obtain a best estimate of implied monetary policy shocks. Importantly, by not assigning one arbitrary yield to have zero observation error (and the other yields to have non-zero observation error), we do not bias our estimated monetary policy shocks to have undue influence from only one particular yield. Instead, the extracted latent factor reflects the dynamics of the entire cross-section of yields.

The third advantage of our estimation method is tractability. Although the likelihood function of yields and related variables can be written down (see Ang and Piazzesi, 2003), the model is high dimensional and non-linear in the parameters. This may cause the likelihood function to have multiple local optima, some of which may lie in unreasonable or implausible regions of the parameter space. In a Baysian estimation setting, we can specify priors on reasonable regions of the parameter space that effectively rule out parameter values that are economically implausible. In our estimation, the only informative prior we impose is that we constrain our state-space system to be stationary. Moreover, the maximum likelihood estimator involves a difficult optimization problem, whereas the Bayesian algorithm is based on a series of simulations that are computationally much more tractable.

Finally, in models with latent factors, the latent factors can be arbitrarily shifted and scaled to yield an observationally equivalent model. Dai and Singleton (2000) and Collin-Dufresne, Goldstein and Jones (2003) discuss some identification issues for affine models with latent factors. We discuss our identification strategy in Appendix B.

4 Empirical Results

Section 4.1 discusses the parameter estimates and the fit of the model to data. Section 4.2 investigates the driving determinants of the yield curve. We compare benchmark, backward-looking and forward-looking Taylor rules in Section 4.3. Sections 4.4 and 4.5 discuss the implied no-arbitrage monetary policy shocks and impulse responses, respectively.

4.1 Parameter Estimates

Table 2 presents the parameter estimates of the unconstrained term structure model (1)-(4). The first row of the companion form Φ shows that GDP growth can be forecasted by lagged inflation and lagged GDP growth. The parameter estimates of the second row of Φ shows that term structure information helps to forecast inflation. The large coefficient on lagged inflation reflects the fact that quarterly inflation is persistent. The third row of Φ shows that both inflation and GDP help forecast the latent term structure factor. This is consistent with results in Ang and Piazzesi (2003), who show that adding macro variables improves out-of-sample forecasts of interest rates. The large coefficient on the lagged latent factor indicates the f_t^u series is more persistent that inflation.

Interestingly, the estimated covariance matrix $\Sigma\Sigma^\top$ shows that innovations to inflation and GDP growth are negatively correlated, whereas high inflation Granger-causes low GDP growth in the conditional mean. In contrast, the conditional covariances between the latent factor and the macro factors are not significant. This implies that the common recursive identification strategy in low-frequency VARs (see, for example, Christiano, Eichenbaum, and Evans, 1996) – macro factors do not respond contemporaneously to policy shocks – is automatically satisfied and therefore not restrictive at our parameter estimates. For identification, we impose the constraint of zero conditional covariances between latent and macro factors when we estimate the forward-looking Taylor rules.

The short rate coefficients in δ_1 are all positive, so higher inflation and GDP growth lead to increases in the short rate, which is consistent with the basic Taylor-rule intuition. In particular, a 1% increase in contemporaneous inflation leads to a 28 basis point (bp) increase in the short rate, while the effect of a 1% increase in GDP growth is small at 5.6bp. Below, we compare these magnitudes with OLS estimates of the Taylor rule.

The risk premia parameters in λ_1 indicate that expected excess returns vary significantly over time. The diagonal elements of λ_1 are all statistically significant. The off-diagonal elements of λ_1 in the third row corresponding to GDP growth and inflation are also statistically significant. Hence, GDP growth, inflation and the latent factor are all going to drive time-varying expected excess returns. Below, we will come back to this issue.

The standard deviations of the observation errors are fairly large. For example, the observation error standard deviation of the one-quarter yield (20-quarter yield) is 19bp (6bp) per quarter. For the one-quarter yield, the measurement errors are comparable to, and slightly smaller than, other estimations containing latent and macro factors (see, for example, Dai and Phillipon, 2005). This is not surprising, because we only have one latent factor to fit the entire yield curve. Piazzesi (2005) shows that traditional affine models often produce large observation errors of the short end of the yield curve relative to other maturities. Indeed, the largest observation error variance occurs at the short end of the yield curve, which indicates that treating the short rate as an observable factor may lead to large discrepancies between the true latent factor and the short rate.[3]

Latent Factor Dynamics

The monetary policy shocks identified by no arbitrage depend crucially on the behavior of the latent factor, f_t^u. Figure 1 plots the latent factor together with the OLS Taylor rule residual and the demeaned short rate. We plot the time-series of the latent factor posterior mean produced from the Gibbs sampler. The plot illustrates the strong relationship between these three series. The correlation of the time-series of the posterior mean of the latent factor with GDP growth (inflation) is -0.122 (0.530). The correlations implied by the model point estimates is -0.128 (0.522), very similar to the correlations computed using posterior mean of the latent factor. These strong correlations suggest that simple OLS estimates of the Taylor rule (9) may be biased in small samples, which we investigate below. The correlations between f_t^u and the yields range between 93% (the short rate) and 99% (the 16-quarter yield). Hence, f_t^u can be interpreted as level factor, similar to the findings of Ang and Piazzesi (2003). In comparison, the correlation between f_t^u and term spreads is below 20%.

[3] We have also estimated the model using only the short rate as an observation equation to draw the latent factor. The estimation results (not reported) show that we can indeed marginally fit the short rate better, but at the expense of the other yields. The gain is limited, as the measurement error for the short rate drops slightly to 15bp, compared to 19bp for our benchmark model, while the measurement errors for the other yields deteriorate significantly. For example, the measurement errors for the 8-quarter yield (20-quarter yield) is 19bp (24bp), compared to 6bp (6bp) in our benchmark estimates. If we invert the latent factor directly from the short rate and so assume that the short rate contains zero measurement error, then the measurement errors for the other yields are even larger.

Matching Moments of Yields and Macro Variables

Table 3 reports the first and second unconditional moments of yields and macro variables computed from data and implied by the model. We compute standard errors of the data estimates using GMM. To test if the model estimates match the data, it is most appropriate to use standard errors from data. This is because large standard errors of parameters may result when the data provide little information about the model, while very efficient estimates produce small standard errors. Nevertheless, we also report posterior standard deviations of the model-implied moments. The moments computed from the model are well within two standard deviations from their counterparts in data for macro variables (Panel A), yields (Panel B), and correlations (Panel C). Panel A shows that the model provides an almost exact match with the unconditional moments of inflation and GDP.

Panel B shows that the autocorrelations in data increase from 0.925 for the short rate to 0.959 for the 5-year yield. In comparison, the model-implied autocorrelations exhibit a smaller range in point estimates from 0.953 for the short rate to 0.963 for the 5-year yield. However, the model-implied estimates are well within two standard deviations of the data point estimates. The smaller range of yield autocorrelations implied by the model is due to having only one latent factor. Since inflation and GDP have lower autocorrelations than yields, the autocorrelations of the yields are primarily driven by the single latent factor f_t^u.

Panel C shows that the model is able to match the correlation of the short rate with GDP and inflation present in the data. The correlation of the short rate with f_t^u implied by the model is 0.962. This implies that using the short rate to identify monetary policy shocks may potentially lead to different estimates than the no-arbitrage shocks identified through f_t^u.

4.2 What Drives the Dynamics of the Yield Curve?

From the yield equation (5), the variables in X_t explain all yield dynamics in our model. To understand the role of each state variable in X_t, we compute variance decompositions from the model and the data. These decompositions are based on Cholesky decompositions of the innovation variance in the following order: $[g_t \, \pi_t \, f_t^u]$, which is consistent with the Christiano, Eichenbaum, and Evans (1996) recursive scheme. We ignore observation error in the yields when computing variance decompositions.

Yield Levels

Panel A of Table 4 reports unconditional variance decompositions of yield levels for various forecasting horizons. The columns under the heading "Risk Premia" report the proportion of the forecast variance attributable to time-varying risk premia. The remainder is the proportion of the variance implied by the predictability embedded in the VAR dynamics without risk premia, under the EH.

To compute the variance of yields due to risk premia, we partition the bond coefficient b_n on X_t in equation (5) into an EH term and into a risk-premia term:

$$b_n = b_n^{EH} + b_n^{RP},$$

where we compute the b_n^{EH} bond pricing coefficient by setting the prices of risk $\lambda_1 = 0$. We let $\Omega^{F,h}$ represent the forecast variance of the factors X_t at horizon h, where $\Omega^{F,h} = \text{var}(X_{t+h} - \mathrm{E}_t(X_{t+h}))$. Since yields are given by $y_t^{(n)} = b_n + b_n^\top X_t$, the forecast variance of the n-maturity yield at horizon h is given by $b_n^\top \Omega^{F,h} b_n$. We compute the unconditional forecast variance using a horizon of $h = 100$ quarters.

We decompose the forecast variance of yields as follows:

$$\text{Risk Premia Proportion} = \frac{b_n^{RP^\top} \Omega^{F,h} b_n^{RP}}{b_n^\top \Omega^{F,h} b_n}.$$

Note that this risk premia proportion reports only the pure risk premia term and ignores any covariances of the risk premia with the state variables. Panel A of Table 4 shows that risk premia play important roles in explaining the level of yields. Unconditionally, the pure risk premia proportion of the 20-quarter yield is 18%. As the maturity increases, the importance of the risk premia increases. Panel B shows that risk premia matter even more for yield spreads. Roughly 1/2 of the variance of yield spreads is due to time-varying risk premia.

The numbers under the line "Variance Decompositions" report the variance decompositions for the total forecast variance, $b_n^\top \Omega^{F,h} b_n$ and the pure risk premia variance, $b_n^{RP^\top} \Omega^{F,h} b_n^{RP}$, respectively. The total variance decompositions reveal that, unconditionally, shocks to macro variables explain about 30-40% of the total variance of yield levels. Shocks to inflation are about twice as important as shocks to GDP in explaining the forecast variance of yield levels. In the pure risk premia term, the proportion of variance attributable to GDP and inflation is also around 30%.

Yield Spreads

Panel B of Table 4 reports variance decompositions of yield spreads of maturity n quarters in excess of the one-quarter yield, $y_t^{(n)} - y_t^{(1)}$. The variance decompositions in Panel B document that shocks to inflation are the main driving force of yield spreads. Shocks to inflation explain more than 87% of the variance of yield spreads. Inflation shocks are even more important for long maturity yield spreads. For example, movements in inflation account for 96% of the unconditional variance of the 5-year spread. These results are consistent with Mishkin (1992) and Ang and Bekaert (2004), who find that inflation accounts for a large proportion of term spread changes.

Expected Excess Holding Period Returns

Panel C of Table 4 examines variance decompositions of expected excess holding period returns. By definition, time-varying expected excess returns must be due only to time-varying risk premia, which is why the total and pure risk premia variance decompositions are identical. Panel C shows that the proportion of the expected excess return variance explained by macro variables significantly increases as the yield maturity increases. At a 4-quarter (20-quarter) maturity, macro factors account for up to 37% (64%). Inflation is more important for explaining time-varying excess returns than GDP, with the proportion for inflation reaching close to 50% for the 20-quarter bond. Thus, inflation and inflation risk impressively account for almost half of the dynamics of expected excess returns. At a one-quarter forecasting horizon (not reported), GDP growth and inflation account for even larger proportions of holding period return variance, at 22% and 64%, respectively.

Table 5 further characterizes conditional expected excess returns. Panel A reports the means and standard deviations of the approximate excess returns computed from data and implied by the model. To compute the one-quarter excess returns on holding the, say, 20-quarter bond from t to $t+1$, we would need data on the price of the 19-quarter bond at $t+1$. Because of data availability, we implement the approximation by Campbell and Shiller (1991):

$$arx_{t+1} = \log \frac{P_{t+1}^{(n)}}{P_t^{(n)}} - r_t. \tag{23}$$

Panel A shows that the moments of excess returns computed from the model are nearly identical to their (approximate) counterparts in data. Hence, our model matches unconditional excess returns almost exactly.

Panel B reports regressions of (approximate) excess returns onto macro factors and yield variables both in data and implied by the model. We choose the 20-quarter yield to be representative of a level factor. The predictability of one-quarter excess returns is fairly weak, compared to the results for longer holding periods reported by Cochrane and Piazzesi (2005). Nevertheless, comparing the model-implied coefficients with the data reveals that the model is able to closely match the predictability patterns in the data. In particular, for the excess returns of longer maturity bonds, the significantly negative (positive) coefficients on inflation (the 20-quarter yield) are well within one standard deviation of their counterparts in data. The point estimates of the loadings on GDP and inflation both increase in magnitude with maturity, indicating that long bond excess returns are more affected by macro factor variation.

Panel C reports the coefficients of the conditional (exact) expected excess holding period return $\mathrm{E}_t(rx_{t+1}^{(n)}) = A_n^x + B_n^{x\top} X_t$ defined in equation (8). The B_n^x coefficients on GDP and inflation are negative, indicating that conditional expected excess returns are strongly counter-cyclical. High GDP growth and high inflation rates are more likely to occur during the peaks of economic expansions, so excess returns of holding bonds are lowest during the peaks of economic expansions. The exposure to this counter-cyclical risk premium also increases as the maturity of the bond increases.

Figure 2 plots the time-series of one-period expected excess holding period returns for the 4-quarter and 20-quarter bond. We compute the expected excess returns using the posterior mean of the latent factors through the sample. Expected excess returns are much more volatile for the long maturity bond, reaching a high of over 16% per year during the 1982 recession and drop below -4% during 1953, 1973 and, 1978. In contrast, expected excess returns for the 4-quarter bond lie in a more narrow range between -0.3% and 2.9% per year. During every recession, expected excess returns increase. In particular, the increase in expected excess returns for the 20-quarter bond at the onset of the 1981 recession is dramatic, rising from 4.5% per year in September 1981 to 16.5% per year in March 1982.[4]

[4] At 1982:Q1, the 16.5% expected excess return for the 20-quarter bond can be broken down into the various proportions: 35% to the constant term premium, 13% to GDP, -14% to inflation, and 65% to the latent factor. Note that there is a large exposure, in absolute values, to macro factors. Although the exposure to the latent factor is large at this date, the implied monetary policy shock is much smaller, as it is the scaled latent factor, $\delta_{1,u} f_t^u$. We discuss this below in further detail.

4.3 A Comparison of Taylor Rules

We now compare the benchmark, backward-looking, and forward-looking Taylor rules estimated by no-arbitrage techniques. We first discuss the estimates of each Taylor rule in turn, and then compare the monetary policy shocks computed from each specification.

The Benchmark Taylor Rule

Panel A of Table 6 contrasts the OLS and model-implied estimates of the benchmark Taylor rule in equation (9). Over the full sample, the OLS estimate of the output coefficient is small at 0.036, and is not significant. The model-implied coefficient is similar in magnitude at 0.056. In contrast, the OLS estimate of the inflation coefficient is 0.643 and strongly significant. The model-implied coefficient on π_t of 0.281 is much smaller. Hence, OLS over-estimates the response of the Fed on the short rate by approximately half compared to the model-implied estimate. This indicates that the endogenous fluctuations in inflation and output are important in estimating the Taylor rule (also see comments by Woodford, 2000).

Moreover, the model estimation extracts information about the policy rule from the entire panel of yield data and not only the time series of the short rate. This approach increases efficiency, which we can see from the number in brackets reported below the model and OLS estimates in Table 6. While these numbers are not directly comparable — OLS regressions produce classical standard errors, while Bayesian estimations produce posterior standard deviations — we can still see that the model estimation produces tighter posterior standard deviations.

To further understand the difference between OLS and model estimates, we compute the OLS coefficients and the OLS R^2 of the benchmark Taylor rule implied by the model, i.e., the model-implied OLS coefficients on g_t and π_t while omitting the latent factor f_t^u from the equation. These coefficients are 0.007 for the constant, 0.041 (0.648) for growth (inflation) – almost identical to the OLS regression coefficients. Moreover, the model-implied OLS regression R^2 is 54%, very similar to the OLS R^2 of 49%. These results suggest that the larger magnitude of the OLS regression estimate of the inflation coefficient in the benchmark Taylor rule compared to the model-implied coefficient is due to an omitted variable that is correlated with GDP growth and inflation.

Our results are based on quarterly growth rates of GDP and the CPI in the Taylor rule. The advantage of our specification is that we would have to deal with moving average errors in the quarterly factor dynamics, if we had used annual growth rates. Moreover, our estimation gives similar magnitudes for the policy rule coefficients. Specifically, we can re-write the Taylor rule (9) to use GDP growth and inflation over the past year:

$$r_t = \gamma_0 + \gamma_{1,g}\frac{1}{4}(g_t + g_{t-1} + g_{t-2} + g_{t-3}) + \gamma_{1,\pi}\frac{1}{4}(\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}) + \varepsilon_t^{MP,T}, \quad (24)$$

where $\varepsilon_t^{MP,T} \equiv \gamma_{1,u} f_t^u$. In this formulation, bond yields now become affine functions of X_t, X_{t-1}, X_{t-2}, and X_{t-3}. Using annual GDP growth and inflation, the posterior mean of the coefficient on GDP growth (inflation) is 0.148 (0.260), with a posterior standard deviation of 0.018 (0.022). These values are almost identical to the estimates from Table 6 based on quarterly growth and inflation.

By estimating the model over the full sample, we follow Christiano, Eichenbaum, and Evans (1996), Cochrane (1998) and others and assume that the Taylor rule relationships are stable. However, our results are surprisingly stable when we estimate the model across subsamples. Panel A of Table 6 reports estimates of both OLS Taylor rules and the benchmark Taylor rule estimated by no-arbitrage restrictions over the pre-1982 and post-1983 monetary policy regimes.[5] For example, the model (OLS) coefficient on inflation is 0.272 (0.677) over the pre-1982 sample and 0.241 (0.605) over the post-1983 sample, compared with 0.281 (0.643) over the whole sample. The model coefficients on GDP are also fairly stable, at 0.051 (0.032) over the pre-1982 (post-1982) period. In contrast, the OLS coefficient on GDP differs widely across the samples, ranging from 0.004 in the pre-1982 sample to 0.238 in the post-1982 sample. Hence, the OLS coefficients of GDP are much more dissimilar across the pre-1982 and post-1983 samples compared to the no-arbitrage Taylor rule estimation.

The Backward-Looking Taylor Rule

Panel B of Table 6 reports the estimation results for the backward-looking Taylor rule. Consistent with equation (13), the model coefficients on g_t and π_t are unchanged from the benchmark Taylor rule in Panel A at 0.056 and 0.281, respectively. The corresponding OLS estimates of the backward-looking Taylor rule coefficients on GDP and inflation are 0.074 and 0.182, respectively. Here, the model-implied rule predicts that the Fed reacts more to inflation than the OLS estimates suggest. As expected, the coefficients on the lagged short rate in both

[5] Several recent studies have emphasized that the linear coefficients γ_1 potentially vary over time (see, among others, Clarida, Galí, and Gertler, 2000; Cogley and Sargent, 2001; Boivin, 2004). However, other authors like Bernanke and Mihov (1998), Sims (1999 and 2001), Sims and Zha (2002), and Primiceri (2003) find either little or no evidence for time-varying policy rules, or negligible effect on the impulse responses of macro variables from time-varying policy rules.

the OLS estimates and the model-implied estimates are similar to the autocorrelation of the short rate (0.925 in Table 3). The large and significant coefficient on the lagged short rate is often interpreted as reflecting the interest rate smoothing behavior of the Fed. We can rewrite the backward-looking Taylor rule into the partial adjustment format used in the macro literature as follows:

$$r_t = (1-0.922)(-0.001+0.714g_t+3.607\pi_t-0.164g_{t-1}-2.522\pi_{t-1})+0.922r_{t-1}+\varepsilon_t^{MP,B}.$$

Hence, our model implies a long-run response to inflation of $3.607-2.522=1.085$. This is consistent with the Taylor principle that the coefficient on inflation should be larger than one.

The Taylor Rule with Serially Correlated Shocks

Figure 3 plots the monetary policy shocks of the Taylor rule with serially correlated errors (see equation (14)) as well as the OLS Taylor rule residual for comparison. Not surprisingly, the serially correlated shocks are much smoother. As a measure of how much predictable variation is contained in the short rate as it responds to contemporaneous and lagged macro variables, we plot the fitted short rate implied from the serially correlated Taylor rule in the bottom panel. From equation (14), we can construct a fitted short rate, r_t^{AR}, where

$$r_t^{AR} = c_t + \Psi_t(L)f_t^o$$

is the predictable variation in the short rate from the entire past history of macro factors. The fitted short rate bears a very high resemblance to the level of the short rate in data, and the R^2 of regressing the short rate in data onto r_t^{AR} is over 80%. Thus, although short rates do not resemble contemporaneous GDP growth or inflation, in a serially correlated Taylor rule, the entire past history of GDP growth and inflation contains a lot of information about the level of the short rate.

The Finite-Horizon, Forward-Looking Taylor Rule

In Panel C of Table 6, we list the estimates of the forward-looking Taylor rule coefficients $\gamma_{1,g}$ and $\gamma_{1,\pi}$ in equation (18) for various horizons k. For each k, we re-estimate the whole term structure model, but only report the forward-looking Taylor rule coefficients for comparison. For a one-quarter ahead forward-looking Taylor rule, the coefficient on expected GDP growth (inflation) is 0.061 (0.401). These are larger than the contemporaneous responses for GDP growth and inflation over the past quarter in the benchmark Taylor rule, which are 0.056 and

0.281, respectively. For a one-year ($k = 4$) horizon, the short interest rate responds quite aggressively to inflation expectations, with $\gamma_{1,\pi} = 0.68$. In all cases, the response of the Fed to future inflation expectations is large, whereas the response of the Fed to future GDP expectations is comparatively small.[6]

As k increases beyond one year, the coefficients on GDP and inflation expectations differ widely and the posterior standard deviations become very large. This is due to two reasons. First, as k becomes large, the conditional expectations approach their unconditional expectations, or $\mathrm{E}_t(g_{t+k,k}) \rightarrow \mathrm{E}(g_t)$ and $\mathrm{E}_t(\pi_{t+k,k}) \rightarrow \mathrm{E}(\pi_t)$. Econometrically, this makes $\gamma_{1,g}$ and $\gamma_{1,\pi}$ hard to identify for large k, and unidentified in the limit as $k \rightarrow \infty$. The intuition behind this result is that as $k \rightarrow \infty$, the only variable driving the dynamics of the short rate in equation (18) is the latent monetary policy shock:

$$r_t = \gamma_0 + \gamma_{1,g}\mathrm{E}(g_t) + \gamma_{1,\pi}\mathrm{E}(\pi_t) + \varepsilon_t^{MP,F},$$

and it is impossible to differentiate the (scaled) effect of GDP or inflation expectations from γ_0. Hence, for large k, identification issues cause the coefficients $\gamma_{1,g}$ and $\gamma_{1,\pi}$ to be poorly estimated.

The second reason is that each estimation for different k tries to capture the same contemporaneous relation between g_t, π_t, and r_t. Panel C also reports the estimates $\delta_1 = [\delta_{1,g}\ \delta_{1,\pi}\ \delta_{1,u}]^\top$ of the short rate equation (17) implied by the forward-looking Taylor rules. These coefficients are very similar across horizons. In particular, the inflation coefficient $\delta_{1,\pi}$ is almost unchanged at around 0.25 for all k. The coefficients on g_t and π_t are also very similar to the coefficients in the benchmark Taylor rule in Panel A. The forward-looking Taylor rule transforms the same contemporaneous response of the short rate to GDP and inflation as the benchmark Taylor rule into the loadings on conditional expectations of future macro factors.

The Infinite-Horizon, Forward-Looking Taylor Rule

We report the estimates of the infinite-horizon, forward-looking Taylor rule (21) in Panel D of Table 6. The coefficient on future discounted GDP growth (inflation) is 0.02 (0.10). The discount rate $\beta = 0.931$, which implies an effective horizon of $1/(1-0.939)$ quarters, or 4.1 years.[7] This estimate is much lower than the discount rates above 0.97 used in the literature

[6] We compute a Bayes factor test among all the forward-looking rules using the harmonic mean of likelihood values proposed by Newton and Raftery (1994), and find strong evidence in favor of the rule with a 4-quarter horizon.

[7] We can also allow different discount rates on future expected GDP growth and future expected inflation. The

(see Salemi, 1995; Rudebusch and Svenson, 1999; Favero and Rovelli, 2003), but still much higher than the estimate of 0.76 calibrated by Collins and Siklos (2004). The effective horizon of approximately four years is consistent with transcripts of FOMC meetings, which indicate that the Fed usually considers forecasts and policy scenarios of up to three to five years ahead.

The Forward- and Backward-Looking Taylor Rule

Finally, Panel E of Table 6, reports the estimates of the forward- and backward-looking Taylor rule in equation (21) for horizons of $k = 1$ and $k = 4$ quarters. These are the same restricted estimations as the forward-looking Taylor rules in Panel C for the corresponding horizons and, hence, have the same coefficients on $\mathrm{E}_t(g_{t+k,k})$ and $\mathrm{E}_t(\pi_{t+k,k})$. Naturally, the lagged short rate continues to play a large role. The estimates show that after taking into account the effect of forward-looking components of GDP and inflation, the response of the Fed to lagged macro variables is negligible.

4.4 Monetary Policy Shocks

The no-arbitrage monetary policy shocks are transformations of either levels or innovations of the latent factor. There are different no-arbitrage policy shocks depending on the chosen structural specification, like benchmark, forward-looking, or backward-looking Taylor rules. Note that the implied policy shock is a choice of a particular structural rule, but the same no-arbitrage model produces several versions of monetary policy shocks (see Table 1).

As an example, we graph the model-implied monetary policy shocks based on the backward-looking Taylor rule in Figure 4 and contrast them with OLS estimates of the backward Taylor rule. We plot the OLS estimate in the top panel and the model-implied shocks, $\varepsilon_t^{MP,B}$, from equation (13) in the bottom panel. We compute $\varepsilon_t^{MP,B}$ using the posterior mean estimates of the latent factor through time. Figure 4 shows that the model-implied shocks are much smaller than the shocks estimated by OLS. In particular, during the early 1980s, the OLS shocks range from below -6% to above 4%. In contrast, the model-implied shocks lie between -3% and 2% during this period. This indicates that according to the no-arbitrage estimates, the Volcker-experience was not as big a surprise as suggested by OLS. These results are consistent with our findings that the pre-1982 and post-1983 estimates of the Taylor rule using no-arbitrage identification techniques are very similar.

discount rate for the future expected GDP growth is estimated as 0.870 but with a larger standard error (0.050), while the estimate for the discount rate for future expected inflation is 0.921 with a standard error of 0.005.

Table 7 characterizes the various model-implied monetary policy shocks in more detail and explicitly compares them with OLS estimates. We list model-implied estimates of the no-arbitrage benchmark Taylor rule shock, $\varepsilon_t^{MP,T}$, which is the scaled latent factor, f_t^u, in equation (9); the backward-looking Taylor rule shocks $\varepsilon_t^{MP,B}$ from equation (13); the autocorrelated shocks $\varepsilon^{MP,AR}$ from equation (14); the forward-looking Taylor rule shocks $\varepsilon_t^{MP,F}$ over a horizon of $k = 4$ quarters from equation (18); and the no-arbitrage forward- and backward-looking Taylor policy shock $\varepsilon^{MP,FB}$ from equation (21), also with a $k = 4$-quarter horizon.

First, the only difference between the no-arbitrage benchmark shock, $\varepsilon_t^{MP,T}$, and the forward-looking shock, $\varepsilon_t^{MP,F}$, are the restrictions imposed on the estimation to take the VAR-implied expectations of future GDP growth and inflation. These estimations are near identical (see Panels A and C of Table 6), so $\varepsilon_t^{MP,T}$ and $\varepsilon_t^{MP,F}$ have a correlation of over 99%. Similarly, both the no-arbitrage backward-looking shock, $\varepsilon_t^{MP,B}$ and the forward- and backward-looking shock, $\varepsilon_t^{MP,FB}$, are both scaled innovations of the latent factor, with the only difference being the restrictions to take the expectations of future macro factors in the short rate equation. Again, these estimations are very similar, producing a correlation between $\varepsilon_t^{MP,B}$ and $\varepsilon_t^{MP,FB}$ over 99%.

We also compare the no-arbitrage shocks with the Romer and Romer (2004) policy shocks that are computed using the Fed's internal forecasts of macro variables and intended changes of the federal funds rate. The OLS residual from the backward Taylor rule has the highest correlation with the Romer-Romer shock, at 72%. In comparison, the no-arbitrage backward-looking shocks have only a 54% correlation with the Romer-Romer series. However, the volatility of the OLS backward-looking residuals are more volatile, at 0.9% per annum than either the no-arbitrage $\varepsilon_t^{MP,B}$ shocks, which have a volatility of 0.6% per annum. Figure 4 clearly illustrates this. The volatility and range of the no-arbitrage $\varepsilon_t^{MP,B}$ shocks are closer to the volatility and range of the Romer-Romer shocks. The OLS backward-looking Taylor residuals are also more negatively autocorrelated (-0.267) than the $\varepsilon_t^{MP,B}$ series, which has an autocorrelation of -0.185. This is very similar to the -0.183 autocorrelation of the Romer-Romer series.

The last two columns of Table 7 report statistics on the one-quarter short rate, r, and the change in the short rate, Δr. The OLS backward-looking Taylor rule shocks are more highly correlated with r, at 32%, than $\varepsilon_t^{MP,B}$, which has a correlation of only 26% with r and only 68% with Δr. Hence, using the short rate as an instrument to estimate monetary policy shocks produces dissimilar estimates to extracting a no-arbitrage estimate of the Taylor rule shocks using the whole yield curve.

4.5 Impulse Responses of Yields

To gauge the effect of the various shocks on the yield curve, we compute impulse response functions. We obtain the posterior distribution of the impulse responses by computing the implied impulse response functions in each iteration of the Gibbs sampler. In the plots, we show the posterior mean of the impulse response functions. These responses are based on Cholesky decompositions that use the same ordering as the variance decompositions: $[g_t \ \pi_t \ f_t^u]$.[8]

Figure 5 plots the responses of yields and yield spreads to GDP shocks, inflation shocks, and a short rate shock. A 1% inflation shock produces persistent effects on all yields. The initial response is highest for the short rate, at 32bp per annum, while the initial response of the long, 20-quarter yield is approximately 16bp per annum. Hence, the term spread narrows from an unexpected inflation shock. Shocks to GDP also increase yields, but the effect from a GDP shock is much smaller. The initial response from a 1% GDP shock is almost the same across the yield curve, at approximately 10bp. The 1% shock to the short rate is constructed by shocking all of the state variables in proportion to their Cholesky decomposition so that the sum of the shock adds up to 1%. This allows us to trace the effect of a change in the short rate across the yield curve. As expected, the initial shock to a 1% increase in the short rate dies out gradually across the yield curve. At a five-year maturity, the response reaches approximately 82bp per annum after two quarters.

Figure 6 plots the responses of yields and yield spreads to -1% expected GDP shocks, and 1% expected inflation shocks over the next year. We construct the -1% shock to $\mathrm{E}_t(g_{t+4,4})$ by noting that the conditional expectation can be written as a linear combination of the factors X_t. Similar to the 1% shock to the short rate in Figure 5, we construct the -1% shock to $\mathrm{E}_t(g_{t+4,4})$ by shocking each variable by X_t in proportion to their Cholesky decomposition so that the sum of the shocks adds up to -1%. The 1% shock to $\mathrm{E}_t(\pi_{t+4,4})$ is constructed the same way. Note that because $\mathrm{E}_t(g_{t+4,4})$ and $\mathrm{E}_t(\pi_{t+4,4})$ are not orthogonal, the effect of the shocks differs from the initial response of the forward-looking Taylor rule coefficients reported in Panel C of Table 6.

[8] We note that, in common with standard macro VARs, our model has a "price puzzle," where after a shock to the short rate, the response of inflation initially increases. However, the no-arbitrage restrictions mitigate the price puzzle. To fully eliminate the prize puzzle, we would need to add certain state variables to our system, such as commodity prices (see comments by Sims, 1992; Christiano, Eichenbaum and Evans, 1996, among others). This is an interesting avenue for future research; the goal of our paper is to illustrate how Taylor rules can be estimated using no-arbitrage techniques, and so we keep the system as low-dimensional as possible.

Figure 6 shows that a negative shock to $\mathrm{E}_t(g_{t+4,4})$ increases all yields. This is consistent with the empirical findings that higher interest rates predict lower future GDP growth (see, for example, Ang, Piazzesi, and Wei 2005). Note that the short rate increases more than longer term yields, which implies a negative response of term spreads to the -1% shock to $\mathrm{E}_t(g_{t+4,4})$. Thus, the model implies that we are more likely to see inverted yield curves before future periods of negative GDP growth, consistent with data. When we shock $\mathrm{E}_t(\pi_{t+4,4})$ by 1%, all yields increase, but the responses are not one-to-one. Thus, the simple real rate implied by the Fisher hypothesis (not adjusted for real rate risk premia), $y_t^{(4)} - \mathrm{E}_t(\pi_{t+4,4})$, decreases, so our model produces real rates that are negatively correlated with expected inflation. The negative correlation between real rates and expected inflation is termed the Mundell-Tobin effect, and our model implies a correlation between real rates and expected inflation of -30%.

5 Conclusion

We exploit information from the entire term structure to estimate monetary policy rules. The framework accommodates original Taylor (1993) rules that describe the Fed as reacting to current values of GDP growth and inflation; backward-looking Taylor rules where the Fed reacts to current and lagged macro variables and lagged policy rates; and forward-looking Taylor rules where the Fed takes into account conditional expectations of future real activity and inflation. The framework also accommodates Taylor rules with serially correlated policy shocks. An advantage of the no-arbitrage model is that all these types of Taylor rules are estimated jointly in a unified system that provides consistent expectations of future interest rates and macro factors.

Our methodology embeds the Taylor rules in a term structure model with time-varying risk premia that excludes arbitrage opportunities. The absence of arbitrage implies that long yields are expected values of future short rates after adjusting for risk. The tractability of the system is based on flexible VAR dynamics for the macro and latent state variables and by specifying risk premia that are also linear combinations of the VAR state variables. In our model, monetary policy shocks are transformations of either levels or innovations to the latent factor, depending on the Taylor rule specification. The cross-equation restrictions implied by no arbitrage help us to estimate this shock more efficiently.

We find that shocks to GDP growth and inflation account for over 60% of the time-variation of time-varying expected excess returns on long-term bonds, while inflation shocks are mostly responsible for driving yield spreads. Macro factors induce a counter-cyclical risk premium for holding long-term bonds. We find that monetary policy shocks identified by no-arbitrage are significantly less volatile than Taylor rule residuals estimated by OLS. Interesting extensions of our approach are to impose more structure on the VAR dynamics or to expand the state space to include other macro factors.

Appendix

A Forward-Looking Taylor Rules

In this appendix, we describe how to compute δ_0, δ_1 in equation (17) of a forward looking Taylor rule without discounting for a k-quarter horizon. From the dynamics of X_t in equation (1), the conditional expectation of k-quarter ahead GDP growth and inflation can be written as:

$$\begin{aligned} \mathrm{E}_t(g_{t+k,k}) &= \mathrm{E}_t\left(\frac{1}{k}\sum_{i=1}^{k} g_{t+i}\right) = \frac{1}{k}e_1^\top \left(\sum_{i=1}^{k}\tilde{\Phi}_i\mu + \tilde{\Phi}_k\Phi X_t\right) \\ \mathrm{E}_t(\pi_{t+k,k}) &= \mathrm{E}_t\left(\frac{1}{k}\sum_{i=1}^{k} \pi_{t+i}\right) = \frac{1}{k}e_2^\top \left(\sum_{i=1}^{k}\tilde{\Phi}_i\mu + \tilde{\Phi}_k\Phi X_t\right), \end{aligned} \tag{A-1}$$

where e_i is a vector of zeros with a 1 in the ith position, and $\tilde{\Phi}_i$ is given by:

$$\tilde{\Phi}_i = \sum_{j=0}^{i-1}\Phi^j = (I-\Phi)^{-1}(I-\Phi^i). \tag{A-2}$$

The bond price recursions for the standard affine model in equation (6) are thus based on the short rate equation $r_t = \bar{\delta}_0 + \bar{\delta}_1^\top X_t$, where:

$$\begin{aligned} \bar{\delta}_0 &= \gamma_0 + \frac{1}{k}[\gamma_{1,g}e_1 \;\; \gamma_{1,\pi}e_2]^\top \left(\sum_{i=1}^{k}\tilde{\Phi}_i\right)\mu, \\ \delta_1^\top &= \frac{1}{k}[\gamma_{1,g}e_1 \;\; \gamma_{1,\pi}e_2]^\top \tilde{\Phi}_k\Phi + \gamma_{1,u}e_3^\top. \end{aligned} \tag{A-3}$$

As $k \to \infty$, both $\mathrm{E}_t(g_{t+k,k})$ and $\mathrm{E}_t(\pi_{t+k,k})$ approach their unconditional means and there is no state-dependence. Hence, the limit of the short rate equation in equation (18) as $k \to \infty$ is:

$$r_t = \gamma_0 + [\gamma_{1,g}e_1 \;\; \gamma_{1,\pi}e_2]^\top (I-\Phi)^{-1}\mu + \gamma_{1,u}f_t^u, \tag{A-4}$$

which implies that when k is large, the short rate effectively becomes a function only of f_t^u, and g_t and π_t can only indirectly affect the term structure through the feedback in the VAR equation (1). In the limiting case $k=\infty$, the coefficients $\gamma_{1,g}$ and $\gamma_{1,\pi}$ are unidentified because they act exactly like the constant term γ_0.

B Econometric Identification

For our benchmark model, our identification strategy is to set the mean of f_t^u to be zero and to pin down $\delta_{1,u}$ while the conditional variance matrix $\Sigma\Sigma^\top$ is unconstrained. To ensure that f_t^u is mean zero, we parameterize $\mu = [\mu_g \;\mu_\pi \;\mu_f]^\top$ so that μ_f solves the equation:

$$e_3^\top (I-\Phi)^{-1}\mu = 0,$$

where e_3 is a vector of zeros with a one in the third position. We set $\delta_{1,u} = 1$. We find that fixing $\delta_{1,u}$ to other values does not change the estimates of $\delta_{1,o}$ because the latent factor can be arbitrarily scaled. We estimate the most fully flexible parameterization that is fully identified.

For the models with forward-looking Taylor rules, certain parameters associated with the horizon are unidentified. Note that the finite-horizon forward-looking rule nests the benchmark model for a horizon of $k=0$ in equation (18). Similarly, the infinite-horizon forward-looking rule nests the benchmark model for a discount rate of $\beta = 0$ in equation (20). Because we estimate the most flexible model as the benchmark case, the forward-looking rules cannot be identified without additional parameter restrictions.

For the forward-looking rules, we impose the identifying assumption that the conditional correlations between the latent and macro factors are zero in the $\Sigma\Sigma^\top$ matrix. This is the same restriction that Christiano, Eichenbaum, and Evans (1996), Evans and Marshall (1998, 2001), among others, employ in estimating policy rules in VARs. The small conditional correlations between the latent and macro factors in the full estimation (see Table 2) indicate that this is a reasonable restriction to impose. For the forward-looking rules, the identifying restriction is:

$$\Sigma\Sigma^\top = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{12} & \Sigma_{22} & 0 \\ 0 & 0 & c \end{pmatrix}. \tag{B-1}$$

Note that while the conditional shocks to the latent factor are uncorrelated with g_t and π_t, because the matrix Φ has full feed-back, all three variables are potentially unconditionally correlated with each other. We set $c = 0.233 \times 10^{-5}$, the estimated value in the benchmark model in which $\Sigma\Sigma^\top$ is a free parameter, so that the coefficient $\delta_{1,u}$ in the short rate equation (2) is of the same magnitude as the benchmark model in the forward-looking systems. Naturally, by fixing c, we draw $\delta_{1,u}$ as the latent factor can be arbitrarily scaled.

In the finite-horizon forward-looking rule, it is theoretically possible to estimate a full covariance matrix while holding the horizon k constant. We find the estimates are very close to those in the covariance matrix of our benchmark model, in which $\Sigma\Sigma^\top$ is unconstrained. However, since the discounting factor, β, is not identified in the full model, we choose to report all the forward-looking rules with the identifying restriction in equation (B-1) so that all the forward-looking rules are comparable with each other.

To match the mean of the short rate in the sample, we set δ_0 in each Gibbs iteration so that:

$$\delta_0 = \bar{r} - \delta_1^\top \bar{X}, \tag{B-2}$$

where $\bar{r}$ is the average short rate from data and $\bar{X}$ is the time-series average of the factors X_t, which change because f_t^u is drawn in each iteration. This means that δ_0 is not individually drawn as a separate parameter, but δ_0 changes its value in each Gibbs iteration because it is a function of δ_1 and the draws of the latent factor f_t^u.

C Estimating the Model

We estimate the model by MCMC with a Gibbs sampling algorithm. Lamoureux and Witte (2002), Mikkelsen (2002), Bester (2003), and Johannes and Polson (2005) develop similar Bayesian methods for estimating term structure models, but their settings do not have macro variables or time-varying prices of risk.

The parameters of the model are $\Theta = (\mu, \Phi, \Sigma, \delta_0, \delta_1, \mu^Q, \Phi^Q, \sigma_\eta)$, where μ^Q and Φ^Q are parameters governing the state variable process under the risk neutral probability measure, σ_η denotes the vector of observation error volatilities $\{\sigma_\eta^{(n)}\}$. We draw μ^Q and Φ^Q, but invert λ_0 and λ_1 using $\lambda_0 = \Sigma^{-1}(\mu - \mu^Q)$, and $\lambda_1 = \Sigma^{-1}(\Phi - \Phi^Q)$. The latent factor $f^u = \{f_t^u\}$ is also generated in each iteration of the Gibbs sampler. We simulate 50,000 iterations of the Gibbs sampler after an initial burn-in period of 10,000 observations.

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 $\hat{Y} = \{\hat{y}_t^{(n)}\}$. Note that the model-implied yields $Y = \{y_t^{(n)}\}$ differ from the yields in data, $\hat{Y}$ by observation error. Note that observing X is equivalent to observing the term structure Y through the bond recursions in equation (6).

Drawing the Latent Factor f^u

The factor dynamics (1), together with the yield equations (22), imply that the term structure model can be written as a state-space system. The state and observation equations for the system are linear in f_t^u, but also involve the macro variables g_t and π_t. To generate f^u, we use the Carter and Kohn (1994) forward-backward algorithm (see also Kim and Nelson, 1999). We first run the Kalman filter forward taking the macro variables (g_t, π_t) to be exogenous variables, and then sample f^u backwards. We use a Kalman filter algorithm that includes time-varying exogenous variables in the state equation following Harvey (1989). Since we specify the mean of f^u to be zero for identification, we set each generated draw of f^u to have a mean of zero.

Drawing μ and Φ

We follow Johannes and Polson (2005) and explicitly differentiate between $\{\mu, \Phi\}$ and $\{\mu^Q, \Phi^Q\}$. As X_t follows a VAR in equation (1), the draw of μ and Φ is standard Gibbs sampling with conjugate normal priors and posteriors. We note that the posterior of μ and Φ conditional on X, $\hat{Y}$ and the other parameters is:

$$\begin{aligned} P(\mu, \Phi \mid \Theta_-, X, \hat{Y}) &\propto P(\hat{Y} \mid \Theta, X) P(X \mid \mu, \Phi, \Sigma) P(\mu, \Phi) \\ &\propto P(\hat{Y} \mid \Sigma, \delta_0, \delta_1, \mu^Q, \Phi^Q, \sigma_\eta, X) P(X \mid \mu, \Phi, \Sigma) P(\mu, \Phi) \\ &\propto P(X \mid \mu, \Phi, \Sigma) P(\mu, \Phi), \end{aligned} \tag{C-1}$$

where Θ_- denotes the set of all parameters except μ and Φ. $P(X|\mu, \Phi, \Sigma)$ is the likelihood function, 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 (6): given μ^Q and Φ^Q, the bond price is independent of μ and Φ. We specify the prior $P(\mu, \Phi)$ to be $N(0, 1000)$, so, consequently, the posterior of (μ, Φ) is a natural conjugate normal distribution and the draw of μ and Φ is standard Gibbs sampling. We draw μ and Φ separately for each equation in the VAR system (1).

We impose the restriction that f_t^u is mean zero for identification. We set μ_3 to satisfy $e_3^\top (I - \Phi)^{-1} \mu = 0$ to ensure that the factor f_t^u has mean zero. Hence μ_3 is simply a function of the other parameters in the factor VAR in equation (1).

Drawing $\Sigma\Sigma^\top$

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

$$P(\Sigma\Sigma^\top \mid \Theta_-, X, \hat{Y}) \propto P(\hat{Y} \mid \Theta, X) P(X \mid \mu, \Phi, \Sigma) P(\Sigma\Sigma^\top), \tag{C-2}$$

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

$$\begin{aligned} \alpha &= \min \left\{ \frac{P((\Sigma\Sigma^\top)^{m+1} \mid \Theta_-, X, \hat{Y})}{P((\Sigma\Sigma^\top)^m \mid \Theta_-, X, \hat{Y})} \frac{q((\Sigma\Sigma^\top)^m)}{q((\Sigma\Sigma^\top)^{m+1})}, 1 \right\} \\ &= \min \left\{ \frac{P(\hat{Y} \mid (\Sigma\Sigma^\top)^{m+1}, \Theta_-, X)}{P(\hat{Y} \mid (\Sigma\Sigma^\top)^m, \Theta_-, X)}, 1 \right\}, \end{aligned} \tag{C-3}$$

where $P(\hat{Y}|\mu, \Phi, \Theta_-, X)$ is the likelihood function, which is normally distributed from the assumption of normality for the observation errors $\eta^{(n)}$. From equation (C-3), α is just the ratio of the likelihoods of the new draw of $\Sigma\Sigma^\top$ relative to the old draw.

Drawing δ_1

We draw δ_1 using a Random Walk Metropolis step:

$$\delta_1^{m+1} = \delta_1^m + \zeta_{\delta_1} v \tag{C-4}$$

where $v \sim N(0, 1)$ and ζ_{δ_1} is the scaling factor used to adjust the acceptance rate. The acceptance probability α for δ_1 is given by:

$$\begin{aligned} \alpha &= \min \left\{ \frac{P(\delta_1^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\delta_1^m \mid \Theta_-, X, \hat{Y})} \frac{q(\delta_1^m \mid \delta_1^{m+1})}{q(\delta_1^{m+1} \mid \delta_1^m)}, 1 \right\} \\ &= \min \left\{ \frac{P(\delta_1^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\delta_1^m \mid \Theta_-, X, \hat{Y})}, 1 \right\}, \end{aligned} \tag{C-5}$$

where the posterior $P(\delta_1|\Theta_-, X, \hat{Y})$ is given by:

$$P(\delta_1|\Theta_-, X, \hat{Y}) \propto P(\hat{Y}|\delta_1, \Theta_-, X) P(\delta_1).$$

Thus, in the case of the draw for δ_1, α is the posterior ratio of the new and old draws of δ_1. We set δ_0 to match the sample mean of the short rate.

To draw γ_1 in the forward-looking Taylor rule system, we rewrite the short rate in data as a regression:

$$\hat{y}_t^{(1)} = \gamma_0 + \gamma_1^\top \bar{X}_t + \eta_t^{(1)},$$

where $\bar{X}_t = [\mathrm{E}_t(g_{t+k,k}\ \mathrm{E}_t(\pi_{t+k,k})\ f_t^u]^\top$, and we can compute the conditional expectations for GDP growth and inflation implied from the VAR parameters at every date t. We generate a proposal draw from the regression for γ_1, and then accept/reject based on the likelihood of the bond yields. We first draw a proposal for the $(m+1)$th value of γ_1 from the proposal density:

$$q(\gamma_1) \propto P(\hat{y}_t^{(1)} \mid \gamma_0, \gamma_1, X, \eta^{(1)}) P(\gamma_1),$$

where we specify the prior $P(\gamma_1)$ to be normally distributed, so, consequently, $q(\gamma_1)$ is a natural conjugate normal distribution. The proposal draw γ_1^{m+1}, is then accepted with probability α, where

$$\begin{aligned} \alpha &= \min\left\{ \frac{P(\gamma_1^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\gamma_1^{m} \mid \Theta_-, X, \hat{Y})} \frac{q(\gamma_1^m)}{q(\gamma_1^{m+1})}, 1 \right\} \\ &= \min\left\{ \frac{P(\hat{Y}_- \mid \gamma_1^{m+1}, \Theta_-, X)}{P(\hat{Y}_- \mid \gamma_1^{m}, \Theta_-, X)}, 1 \right\}, \end{aligned} \tag{C-6}$$

where $P(\hat{Y_-} | \gamma_1, \Theta_-, X)$ is the likelihood function of yields other than the short rate $\hat{r}$, which is normally distributed from the assumption of normality for the observation errors $\eta^{(n)}$. We set γ_0 to match the sample mean of the short rate.

Drawing μ^Q and Φ^Q

We draw μ^Q and Φ^Q with a Random Walk Metropolis algorithm. We assume 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 similar to equation (C-5). 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, as it is easier to interpret market prices of risk than parameters under the risk-neutral measure.

Drawing σ_η

Drawing the variance of the observation errors, σ_η^2, is straightforward, because we can view the observation errors η as regression residuals from equation (22). 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).

Drawing β

For the case of the forward-looking Taylor rule over an infinite horizon with discounting, we augment the parameter space to include the discount rate, β. To draw β, we use an Independence Metropolis-Hastings step. The candidate draw, β^{m+1}, is drawn from a proposal density, $q(\beta^{m+1} \mid \beta^m) = q(\beta^{m+1})$, which we specify to be a doubly truncated normal distribution, with mean 0.95 and standard deviation 0.03 but truncated at 0.8 from below and at 0.99 from above.

Assuming a flat prior, the acceptance probability α for β^{m+1} is given by:

$$\begin{aligned} \alpha &= \min\left\{ \frac{P(\beta^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\beta^{m} \mid \Theta_-, X, y)} \frac{q(\beta^m)}{q(\beta^{m+1})}, 1 \right\} \\ &= \min\left\{ \frac{P(\hat{Y} \mid \beta^{m+1}, \Theta_-, X)}{P(\hat{Y} \mid \beta^{m}, \Theta_-, X)} \frac{q(\beta^m)}{q(\beta^{m+1})}, 1 \right\}, \end{aligned} \tag{C-7}$$

where Θ_- represents all the parameters except the β parameter that is being drawn and $P(\hat{Y}|\beta, \Theta_-, X)$ is the likelihood function.

Scaling Factors and Accept Ratios

The table below lists the scaling factors and acceptance ratios used in the Random Walk Metropolis steps for the benchmark Taylor rule and backward-looking Taylor rule estimation.

Parameter	Scaling Factor	Acceptance Ratio	Parameter	Scaling Factor	Acceptance Ratio
δ_1	0.00100	0.150	μ_1^Q	0.00050	0.133
μ_2^Q	0.00005	0.379	μ_3^Q	0.00002	0.141
$\Phi_{1,\cdot}^Q$	0.00500	0.413	$\Phi_{2,\cdot}^Q$	0.00100	0.582
$\Phi_{3,\cdot}^Q$	0.00050	0.168			

where $\mu^Q = (\mu_1^Q \; \mu_2^Q \; \mu_3^Q)^\top$ and $\Phi_{i,\cdot}^Q$ denotes the element of Φ^Q in the ith row.

Checks for Convergence

To check the reliability of our estimation approach, we perform several exercises. First, we tried starting the chain from many different initial values on real data and we obtained almost exactly the same results for the posterior means and standard deviations of the parameters. We also check that the posterior distributions for the parameters Θ are unimodal.

Second, we compute the Raftery and Lewis (1992) minimum burn-in and the minimum number of runs required to estimate the 0.025 quantile to within ± 0.025 with probability 0.95, using every draw in the MCMC-Gibbs algorithm, which is conservative. For all the parameters (with one exception) and the complete time-series of the latent factors f^u, the minimum required burn-in is only several hundred and the minimum number of runs is several thousand. This is substantially below the burn-in sample (10,000) and the number of iterations (50,000) for our estimation.

The third, and probably most compelling check of the estimation method is that the MCMC-Gibbs sampler works very well on simulated data. We perform Monte Carlo simulations, similar to the experiments performed by Eraker, Johannes and Polson (2003). We take the posterior means of the parameters in Table 2 as the population values and simulate a small sample of 203 quarterly observations, which is the same length as our data. Applying our MCMC algorithm to the simulated small sample, we find that the draws of the VAR parameters (μ, Φ, Σ), the short rate parameters (δ_0, δ_1), the constant prices of risk (λ_0), and the observation error standard deviations $(\sigma_\eta^{(n)})$ converge extremely fast. After our estimation procedure, the posterior means for these parameters are all well within one posterior standard deviation of the population parameters. We find that a burn-in sample of only 1,000 observations is sufficient to start drawing values for these parameters that closely correspond to the population distributions. The time-varying prices of risk (λ_1) were estimated less precisely on the simulated data, but the posterior means of eight out of nine prices of risk were also within one posterior standard deviation of the population parameters. The algorithm is also successful in estimating the time-series of the latent factor f^u, where the true series of f^u in the simulated sample lies within one posterior standard deviation bound of the posterior mean of the generated f^u from the Gibbs sampler.

In summary, these results verify that we can reliably estimate the parameters of the term structure model given our sample size and, thus, we are very confident about the convergence of our algorithm.

References

[1] Ang, A., and G. Bekaert, 2004, "The Term Structure of Real Rates and Expected Inflation," working paper, Columbia University.

[2] 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.

[3] Ang, A., M. Piazzesi, and M. Wei, 2005, "What does the Yield Curve Tell us about GDP Growth?" forthcoming *Journal of Econometrics*.

[4] Bagliano, F. C., and C. A. Favero, 1998, "Measuring Monetary Policy with VAR Models: An Evaluation," European Economic Review, 42, 1069-112.

[5] Bansal, R., G. Tauchen, and H. Zhou, 2004, "Regime Shifts, Risk Premiums in the Term Structure, and the Business Cycle," *Journal of Business and Economic Statistics*, 22, 4, 396-409.

[6] Bekaert, G., S. Cho, and A. Moreno, 2003, "New-Keynesian Macroeconomics and the Term Structure," working paper, Columbia University.

[7] Bernanke, B. S., and I. Mihov, 1998, "Measuring Monetary Policy," *Quarterly Journal of Economics*, 113, 3, 869-902.

[8] Bester, A., 2003, "Random Fields and Affine Models of Interest Rates," working paper, Duke University.

[9] Bikbov, R., and M. Chernov, 2005, "No-Arbitrage Macroeconomic Determinants of the Yield Curve," working paper, Columbia University.

[10] Boivin, J., 2004, "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," working paper, Columbia University.

[11] Boivin, J., B. Bernanke, and P. Eliasz, 2004, "Measuring Monetary Policy: A Factor Augmented Vector Autoregressive (FAVAR) Approach," forthcoming *Quarterly Journal of Economics*.

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

[13] Campbell, J. Y., and R. J. Shiller, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495-514.

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

[15] Chen, R. R., and L. Scott, 1993, "Maximum Likelihood Estimation for a Multi-factor Equilibrium Model of the Term Structure of Interest Rates," *Journal of Fixed Income*, 3, 14-31.

[16] 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.

[17] 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.

[18] Clarida, R., and M. Gertler, 1997, "How the Bundesbank Conducts Monetary Policy," in *Reducing Inflation: Motivation and Strategy*, Romer, C. D., and D. H. Romer, eds., University of Chicago Press, Chicago, 363-406.

[19] Clarida, R., Galí, J., and M. Gertler, 1998, "Monetary Policy Rules in Practice: Some International Evidence," *European Economic review*, 42, 6, 1033-1067.

[20] 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.

[21] Cochrane, J. H., 1998, "What Do The VARs Mean? Measuring The Output Effects Of Monetary Policy," *Journal of Monetary Economics*, 41, 2, 277-300.

[22] Cochrane, J. H., and M. Piazzesi, 2005, "Bond Risk Premia," *American Economic Review* 95, 1, pp. 138-160.

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

[24] Collin-Dufresne, P., R. S. Goldstein, and C. S. Jones, 2003, "Identification and Estimation of 'Maximal' Affine Term Structure Models: An Application to Stochastic Volatility," working paper, USC.

[25] Collins, S., and P. L. Siklos, 2004, " 'Discounting' Optimal Taylor Rules: Can the Policy Horizon Masquerade as Interest Rate Smoothing?" working paper, Investment Company Institute, Washington DC.

[26] Constantinides, G. M., 1992, "A Theory of the Nominal Term Structure of Interest Rates," *Review of Financial Studies*, 5, 531-553.

[27] Dai, Q., and T. Philippon, 2005, "Government Deficits and Interest Rates: A No-Arbitrage Structural VAR Approach," working paper, NYU.

[28] Dai, Q., and K. J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 5, 1943-1978.

[29] 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.

[30] Dewachter, H., and M. Lyrio, 2004, "Macro Factors and the Term Structure of Interest Rates," forthcoming *Journal of Money, Credit and Banking*.

[31] Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba, 2005, "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," forthcoming *Journal of Econometrics*.

[32] Duffee, G. R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, 57, 405443.

[33] Evans, C. L. and D. Marshall, 1998, "Monetary Policy and the Term Structure of Nominal Interest rates: Evidence and Theory," Carnegie-Rochester Conference Series on Public Policy, 49, 53-111.

[34] Evans, C. L. and D. Marshall, 2001, "Economic Determinants of the Term Structure of Interst Rates," working paper, Federal Reserve Bank of Chicago.

[35] Eichenbaum, M., and C. L. Evans, 1995, "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates," *Quarterly Journal of Economics*, 110, 4, 975-1009.

[36] Eraker, B., M. Johannes, and N. Polson, 2003, "The Impact of Jumps in Volatility and Returns," *Journal of Finance*, 58, 3, 1269-1300.

[37] Favero, C., and R. Rovelli, 2003, "Macroeconomic Stability and the Preferences of the Fed: A Formal Analysis," *Journal of Money, Credit and Banking*, 35, 4, 545-556.

[38] Fisher, M., 1998, "A Simple Model of the Failure of the Expectations Hypothesis, working paper, Federal Reserve Bank of Atlanta.

[39] Fleming, M. J., 2003, "Measuring Treasury Market Liquidity", *Federal Reserve Bank of New York Economic Policy Review*, 9, 3, 83-108.

[40] Fuhrer, J. C., and G. R. Moore, 1995, "Monetary Policy Trade-Offs and the Correlation between Nominal Interest Rates and Real Output," *American Economic Review*, 85, 1, 219-239.

[41] Gallmeyer, Michael F., Burton Hollifield, and Stanley E. Zin, 2005, "Taylor Rules, McCallum Rules, and the Term Structure of Interest rates," NBER working paper 11276.

[42] Goodfriend, M., 1991, "Interest-Rate Smoothing in the Conduct of Monetary Policy," *Carnegie-Rochester Conference Series on Public Policy*, 34, 7-30.

[43] Harvey, A. C., 1989, *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, U.K.

[44] Hördahl, P., O. Tristani, and D. Vestin, 2005, "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," forthcoming *Journal of Econometrics*.

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

[46] Kim, C., and C. R. Nelson, 1999, *State-Space Models with Regime Switching*, MIT Press, Cambridge, MA.

[47] 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.

[48] Mikkelsen, P., 2002, "MCMC Based Estimation of Term Structure Models," working paper, Aarhus School of Business, Denmark.

[49] Mishkin, F. S., 1992, "Is the Fisher Effect for Real?" *Journal of Monetary Economics*, 30, 195-215.

[50] Newton, M. A., and A. E. Raftey, 1994, "Approximate Bayesian Inference by the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society*, 56, 3-48.

[51] Piazzesi, M., 2005, "Bond Yields and the Federal Reserve," *Journal of Political Economy*, 113, 2, pp. 311-344.

[52] Primiceri, G. E., 2003, "Time Varying Structural Vector Autoregressions and Monetary Policy," working paper, Princeton University.

[53] Raftery, A. E., and S. Lewis, 1992, "How Many Iterations in the Gibbs Sampler?" in J. M. Bernardo, J. Berger, A. P. David, and A. F. M. Smith, ed., *Bayesian Statistics 4*, Oxford University Press, Oxford, pp. 763-773.

[54] Romer, C. D., and D. H. Romer, 2004, "A New Measure of Monetary Policy Shocks: Derivation and Implication," *American Economic Review* 94, 4, pp. 1055-1084.

[55] Rudebusch, G., and L. E. O. Svensson, 1999, "Policy Rules for Inflation Targeting," in Taylor, J. B., ed., *Monetary Policy Rules*, Chicago University Press, Chicago, IL, 203-246.

[56] Rudebusch, G., and T. Wu, 2005, "A Macro-Finance Model of the Term Strucutre, Monetary Policy, and the Economy," working paper, Federal Reserve Bank of San Francisco.

[57] Salemi, M., 1995, "Revealed Preferences of the Federal Reserve using Inverse Control Theory to Interpret the Policy Function of a Vector Autoregression," *Journal of Business and Economic Statistics*, 13, 419-433.

[58] Sims, C. A., 1992, "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, 36, 975-1000.

[59] Sims, C. A., 1999, "Drifts and Breaks in Monetary Policy," working paper, Yale University.

[60] Sims, C. A., 2001, "Comment on Sargent and Cogley's 'Evolving Post World War II U.S. Inflation Dynamics,' " *NBER Macroeconomics Annual*, 16, 373-379.

[61] Sims, C. A., and T. Zha, 2002, "Macroeconomic Switching," working paper, Princeton University.

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

[63] Svensson, L. E. O., 1997, "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets," *European Economic Review*, 41, 1111-1146.

[64] Woodford, M., 2000, "Pitfalls of Forward-Looking Monetary Policy," *American Economic Review*, 90, 2, 100-104.

Table 1: SUMMARY OF NO-ARBITRAGE TAYLOR RULES

Taylor Rule Specification	Equivalent to	Monetary Policy Shocks is a Transformation of the	Monetary Policy Shock Denoted as	Restrictions?
Benchmark	Backward-Looking	Level of f_t^u	$\varepsilon_t^{MP,T} = \delta_{1,u} f_t^u$	No
Backward-Looking	Benchmark	Current Innovation in f_t^u	$\varepsilon_t^{MP,B} = \delta_{1,u} v_t^u$	No
Serially Correlated Shocks	Benchmark	Current and Past Innovations in f_t^u	$\varepsilon_t^{MP,AR} = \sum_{i=0}^{t-1} \Phi_{22}^i \delta_{1,u} v_{t-i}^u$	No
Forward-Looking	Backward- and Forward-Looking	Level of f_t^u	$\varepsilon_t^{MP,F} = \delta_{1,u} f_t^u$	Yes
Backward- and Forward-Looking	Forward-Looking	Current Innovation in f_t^u	$\varepsilon_t^{MP,FB} = \delta_{1,u} v_t^u$	Yes

Note: The table summarizes the various Taylor Rule specifications identified using no-arbitrage conditions of Sections 2.3 to 2.8. The models in the equivalence column indicates that the Taylor rule specification is equivalent in the sense that one term structure model gives rise to both Taylor rule specifications. The variable f_t^u represents the latent term structure factor and the restrictions refer to parameter restrictions in the short rate equation imposed to take future expectations of the macro variables consistent with the VAR dynamics of the model (see Section 2.6).

Table 2: PARAMETER ESTIMATES

Factor Dynamics

		Companion Form Φ			$\Sigma\Sigma' \times 10000$		
	μ	g	π	f^u	g	π	f^u
g	0.008	0.301	-0.240	-0.019	0.828	-0.011	0.007
	(0.001)	(0.068)	(0.095)	(0.132)	(0.086)	(0.034)	(0.014)
π	0.002	0.048	0.700	0.198	-0.011	0.265	0.001
	(0.001)	(0.038)	(0.055)	(0.075)	(0.034)	(0.027)	(0.009)
f^u	-0.001	0.039	0.063	0.922	0.007	0.001	0.023
	(0.001)	(0.014)	(0.018)	(0.022)	(0.014)	(0.009)	(0.003)

Short Rate Equation

	δ_1		
δ_0	g	π	f^u
0.010	0.056	0.281	1.000
(0.001)	(0.013)	(0.029)	—

Risk Premia Parameters

		λ_1		
	λ_0	g	π	f^u
g	0.945	-45.674	-48.621	-32.501
	(0.551)	(16.913)	(34.717)	(23.072)
π	-0.240	2.404	-22.405	-1.023
	(0.193)	(10.360)	(13.034)	(17.444)
f^u	-0.669	22.390	48.335	-34.369
	(0.202)	(9.337)	(14.545)	(15.771)

Observation Error Standard Deviation

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
$\sigma_\eta^{(n)}$	0.189	0.128	0.063	0.035	0.046	0.064
	(0.010)	(0.007)	(0.004)	(0.002)	(0.003)	(0.004)

Note: The table lists parameter values for the model in equations (1)-(4) and observation error standard deviations in equation (22) for yields of maturity n quarters. We use 50,000 simulations after a burn-in sample of 10,000 for the Gibbs sampler. We report the posterior mean and posterior standard deviation (in parentheses) of each parameter. The sample period is June 1952 to December 2002 and the data frequency is quarterly.

Table 3: SUMMARY STATISTICS

PANEL A: MOMENTS OF MACRO FACTORS

	Means %		Standard Deviations %		Autocorrelations	
	Data	Model	Data	Model	Data	Model
g	0.803	0.809	0.964	1.007	0.340	0.356
	(0.088)	(0.204)	(0.067)	(0.073)	(0.068)	(0.077)
π	0.950	0.949	0.792	0.879	0.762	0.776
	(0.110)	(0.643)	(0.097)	(0.225)	(0.057)	(0.064)

PANEL B: MOMENTS OF YIELDS

	$n = 1$	$n = 4$	$n = 8$	$n = 12$	$n = 16$	$n = 20$
Means						
Data	1.334	1.438	1.488	1.528	1.558	1.576
	(0.107)	(0.108)	(0.107)	(0.105)	(0.105)	(0.104)
Model	1.334	1.419	1.488	1.530	1.557	1.575
	–	(0.005)	(0.003)	(0.002)	(0.002)	(0.004)
Standard Deviations						
Data	0.720	0.721	0.713	0.697	0.693	0.684
	(0.092)	(0.083)	(0.082)	(0.080)	(0.080)	(0.076)
Model	0.707	0.707	0.705	0.699	0.691	0.680
	(0.011)	(0.006)	(0.003)	(0.003)	(0.003)	(0.005)
Autocorrelations						
Data	0.925	0.932	0.943	0.949	0.955	0.959
	(0.032)	(0.031)	(0.029)	(0.028)	(0.028)	(0.026)
Model	0.953	0.958	0.961	0.963	0.963	0.963
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

PANEL C: SHORT RATE CORRELATIONS

	g	π	f^u
Data	-0.090	0.695	–
	(0.100)	(0.063)	
r	-0.103	0.735	0.962
	(0.125)	(0.087)	(0.021)

Note: Panel A lists moments of GDP and inflation in data and implied by the model.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. In Panel C, we report correlations of the short rate with various factors. For the model, we compute the posterior distribution of the correlations of the model-implied short rate r in equation (2). In all the panels, the data standard errors (in parentheses) are computed using GMM and all moments are computed at a quarterly frequency. For the model, we report posterior means and standard deviations (in parentheses) of each moment. The sample period is June 1952 to December 2002 and the data frequency is quarterly.

Table 4: VARIANCE DECOMPOSITIONS

		Variance Decompositions					
		Total			Risk Premia		
Maturity (qtrs)	Risk Premia	g	π	f^u	g	π	f^u
PANEL A: YIELD LEVELS $y_t^{(n)}$							
1	0.0	12.5	28.7	58.8	–	–	–
4	0.7	12.9	25.1	62.0	12.7	17.7	69.6
8	3.4	13.0	22.6	64.4	11.4	17.6	71.0
12	7.6	13.0	21.2	65.8	11.3	17.4	71.3
16	12.5	13.0	20.3	66.7	11.3	17.2	71.5
20	17.8	13.0	19.8	67.2	11.4	17.0	71.6
PANEL B: YIELD SPREADS $y_t^{(n)} - y_t^{(1)}$							
4	49.4	0.5	87.3	12.2	12.7	17.7	69.6
8	49.9	0.3	89.7	10.0	11.4	17.6	71.0
12	48.7	0.2	92.4	7.4	11.3	17.4	71.3
16	46.9	0.3	94.7	5.0	11.3	17.2	71.5
20	44.8	0.6	96.0	3.4	11.4	17.0	71.6
PANEL C: EXPECTED EXCESS HOLDING PERIOD RETURNS $\mathrm{E}_t(rx_{t+1}^{(n)})$							
4	100	16.2	20.7	63.1	16.3	20.7	63.0
8	100	15.8	30.4	53.8	15.8	30.4	53.8
12	100	15.6	38.9	45.5	15.6	38.9	45.5
16	100	15.5	44.9	39.6	15.5	44.9	39.6
20	100	15.5	49.0	35.5	15.5	49.0	35.5

Note: The table reports unconditional variance decompositions of forecast variance (in percentages) for yield levels $y_t^{(n)}$ in Panel A; yield spreads $y_t^{(n)} - y_t^{(1)}$ in Panel B; and unconditional expected excess holding period returns $\mathrm{E}(rx_{t+1}^{(n)}) = \mathrm{E}(ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t)$ in Panel C. In each panel, we also examine the variance decomposition due to time-varying risk premia. By definition, the variance decompositions of time-varying expected excess holding period returns must be due only to time-varying risk premia. All maturities are in quarters. We ignore observation error for computing variance decompositions for yield levels and yield spreads. All the variance decompositions are computed using the posterior mean of the parameters listed in Table 2.

Table 5: CHARACTERIZING EXCESS RETURNS

PANEL A: MOMENTS OF EXCESS RETURN

	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
Means					
Data	0.107	0.157	0.193	0.213	0.222
	(0.048)	(0.104)	(0.152)	(0.200)	(0.234)
Model	0.085	0.154	0.196	0.224	0.242
	(0.018)	(0.028)	(0.032)	(0.036)	(0.043)
Standard Deviations					
Data	0.755	1.558	2.204	2.796	3.310
	(0.129)	(0.222)	(0.262)	(0.295)	(0.339)
Model	0.629	1.410	2.167	2.904	3.619
	(0.034)	(0.075)	(0.115)	(0.154)	(0.194)

PANEL B: PREDICTABILITY REGRESSIONS

	Data Estimates				Model-Implied Estimates			
	g	π	$y^{(20)}$	R^2	g	π	$y^{(20)}$	R^2
$n=4$	-0.072	-0.078	0.223	0.036	-0.044	-0.039	0.164	0.040
	(0.064)	(0.090)	(0.096)		(0.047)	(0.069)	(0.077)	
$n=12$	-0.193	-0.461	0.752	0.040	-0.194	-0.449	0.743	0.053
	(0.184)	(0.240)	(0.296)		(0.159)	(0.236)	(0.262)	
$n=20$	-0.237	-0.719	1.129	0.039	-0.362	-0.957	1.331	0.061
	(0.266)	(0.366)	(0.450)		(0.265)	(0.393)	(0.434)	

PANEL C: FACTOR COEFFICIENTS

		Maturity (qtrs)				
		4	8	12	16	20
A_n^x		0.002	0.006	0.009	0.012	0.015
B_n^x	g	-0.039	-0.102	-0.176	-0.255	-0.334
	π	-0.066	-0.239	-0.462	-0.703	-0.946
	f^u	0.228	0.549	0.868	1.176	1.470

Note: Panel A lists moments of one-quarter approximate excess holding period returns, $arx_{t+1}^{(n)}$, in the data and implied by the model (see equation (7)). 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 regresses one-quarter approximate excess holding period returns for an n-period bond, $arx_{t+1}^{(n)}$ onto GDP, inflation, and a bond yield. The standard errors for the OLS estimates from data (in parentheses) are computed using robust standard errors. We compute the model-implied coefficients and R^2 as follows. We construct the posterior distributions of the model-implied estimates by computing the implied coefficients from the model parameters in each iteration of the Gibbs sampler. We report posterior means and standard deviations (in parentheses) of each coefficient. Panel C reports the coefficients of the conditional expected excess holding period return $\mathrm{E}_t(rx_{t+1}^{(n)}) = A_n^x + B_n^{x\top} X_t$ defined in equation (8) on the factors $X_t = [g_t \ \pi_t \ f_t^u]^\top$. The data frequency is quarterly and the sample period is June 1952 to December 2002.

Table 6: TAYLOR RULES

PANEL A: BENCHMARK TAYLOR RULE

	Full Sample		Pre-82:Q4		Post-83:Q1	
	OLS	Model	OLS	Model	OLS	Model
const	0.007	0.010	0.006	0.010	0.007	0.012
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
g_t	0.036	0.056	0.004	0.051	0.238	0.032
	(0.071)	(0.012)	(0.080)	(0.017)	(0.104)	(0.035)
π_t	0.643	0.281	0.677	0.272	0.605	0.241
	(0.080)	(0.029)	(0.082)	(0.031)	(0.130)	(0.052)

PANEL B: BACKWARD-LOOKING TAYLOR RULE

	const	g_t	π_t	g_{t-1}	π_{t-1}	r_{t-1}	R^2
OLS	0.000	0.074	0.182	-0.005	-0.077	0.879	0.895
	(0.000)	(0.027)	(0.046)	(0.029)	(0.041)	(0.035)	
Model	-0.000	0.056	0.281	-0.013	-0.197	0.922	0.955
	(0.001)	(0.013)	(0.029)	(0.015)	(0.028)	(0.022)	

PANEL C: FINITE-HORIZON, FORWARD-LOOKING TAYLOR RULE

	Forward-Looking Taylor Rule			Implied Short Rate		
	const	$\mathrm{E}_t(g_{t+k,k})$	$\mathrm{E}_t(\pi_{t+k,k})$	$\delta_{1,g}$	$\delta_{1,\pi}$	$\delta_{1,u}$
$k=1$	0.009	0.061	0.401	0.036	0.259	0.897
	(0.001)	(0.081)	(0.154)	(0.017)	(0.019)	(0.054)
$k=4$	0.006	0.157	0.676	0.044	0.265	0.918
	(0.003)	(0.221)	(0.138)	(0.014)	(0.018)	(0.029)
$k=8$	0.005	0.042	0.883	0.035	0.235	0.868
	(0.005)	(0.385)	(0.204)	(0.001)	(0.019)	(0.026)
$k=20$	0.005	-0.546	1.385	0.021	0.248	0.863
	(0.010)	(0.843)	(0.459)	(0.003)	(0.015)	(0.032)

Table 6 Continued

PANEL D: INFINITE-HORIZON, FORWARD-LOOKING TAYLOR RULE

	Forward-Looking Taylor Rule Coefficients				Implied Short Rate Coefficients		
	const	$\hat{g}_t$	$\hat{\pi}_t$	β	$\delta_{1,g}$	$\delta_{1,\pi}$	$\delta_{1,u}$
$k=\infty$	-0.003	0.020	0.096	0.939	0.038	0.275	1.078
	(0.002)	(0.010)	(0.008)	(0.007)	(0.015)	(0.021)	(0.116)

PANEL E: FORWARD- AND BACKWARD-LOOKING TAYLOR RULE

	const	$\mathrm{E}_t(g_{t+k,k})$	$\mathrm{E}_t(\pi_{t+k,k})$	g_{t-1}	π_{t-1}	r_{t-1}	R^2
$k=1$	-0.001	0.061	0.401	0.001	-0.054	0.846	0.962
	(0.001)	(0.081)	(0.051)	(0.016)	(0.051)	(0.039)	
$k=4$	-0.003	0.157	0.676	-0.008	-0.073	0.726	0.971
	(0.003)	(0.221)	(0.138)	(0.012)	(0.045)	(0.091)	

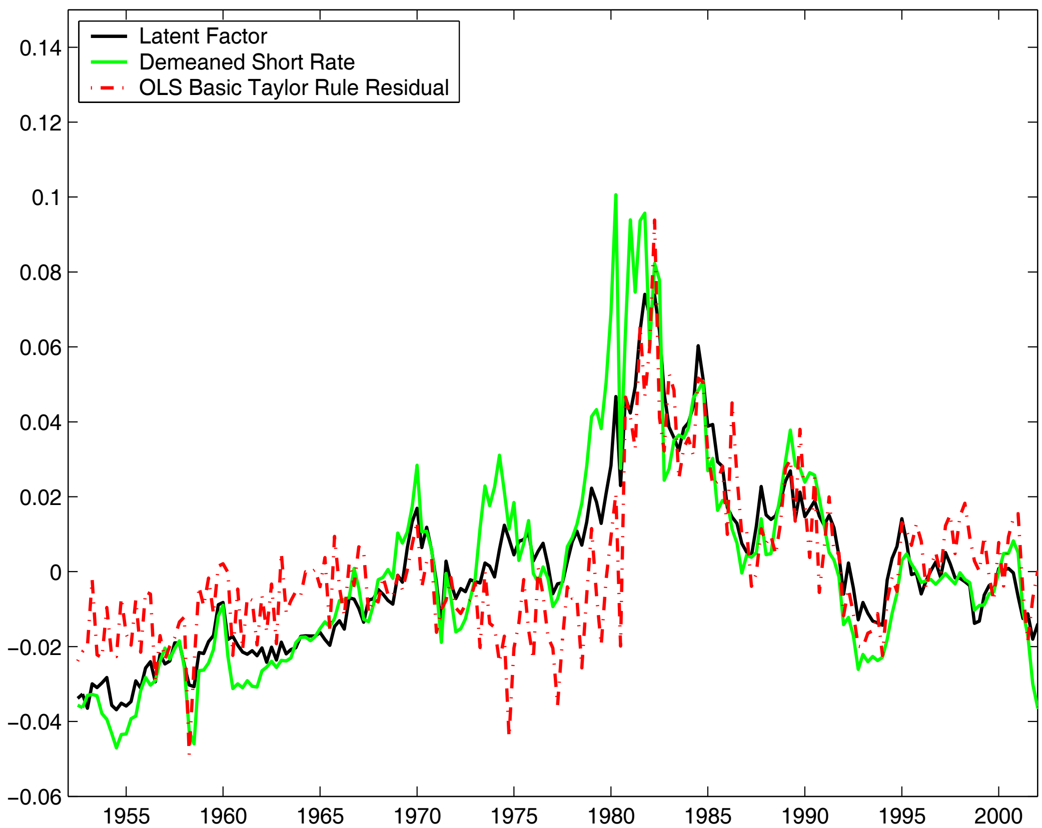
Note: Panel A reports the OLS and model-implied estimates of the benchmark Taylor (1993) rule in equation (9) over the full sample and over subperiods; Panel B reports the backward-looking Taylor rule (10); Panel C reports the finite-horizon, forward-looking Taylor rule without discounting in equation (18); Panel D reports the infinite-horizon, forward-looking Taylor rule with discounting in equation (19); and Panel E reports the forward- and backward-looking Taylor rule in equation (21). Panels C and D also report the implied short rate coefficients corresponding to the forward-looking Taylor rules without discounting in equation (17) for each horizon k and equation (20) for the forward-looking Taylor rule with discounting. For the model-implied coefficients, we construct the posterior distribution of Taylor rule coefficients by computing the implied coefficients from the model parameters in each iteration of the Gibbs sampler. We report posterior means and standard deviations (in parentheses) of each coefficient. The standard errors for the OLS estimates (in parentheses) are computed using robust standard errors. In each panel, the data frequency is quarterly and the full sample period is from June 1952 to December 2002.

Table 7: CHARACTERIZING DIFFERENT MONETARY POLICY SHOCKS

	OLS Estimates		Model-Implied Estimates							
	Bmk	Bwd	Bmk	Bwd	AR	Fwd	Fwd-Bwd	Romer-Romer	r	Δr
Correlations										
OLS Benchmark Rule	1.000	0.454	0.789	0.345	0.786	0.776	0.312	0.276	0.713	0.140
OLS Backward Rule		1.000	0.251	0.752	0.313	0.246	0.742	0.715	0.325	0.886
Model Benchmark Rule			1.000	0.264	0.725	0.999	0.260	0.176	0.934	0.047
Model Backward Rule				1.000	0.371	0.260	0.993	0.538	0.257	0.683
Model AR Shocks					1.000	0.714	0.353	0.260	0.590	0.100
Model Forward Rule						1.000	0.260	0.180	0.941	0.055
Model Fwd-Bwd Rule							1.000	0.554	0.273	0.720
Romer-Romer Shock								1.000	0.290	0.726
Short Rate r									1.000	0.187
Change of Short Rate										1.000
Autocorrelations	0.748	-0.267	0.962	-0.185	0.885	0.964	-0.169	-0.183	0.932	-0.184
Standard Deviation	0.021	0.009	0.023	0.006	0.013	0.018	0.004	0.007	0.029	0.011
Minimum	-0.049	-0.058	-0.037	-0.027	-0.033	-0.029	-0.022	-0.040	0.006	-0.073
Maximum	0.094	0.043	0.074	0.025	0.039	0.058	0.018	0.025	0.154	0.037

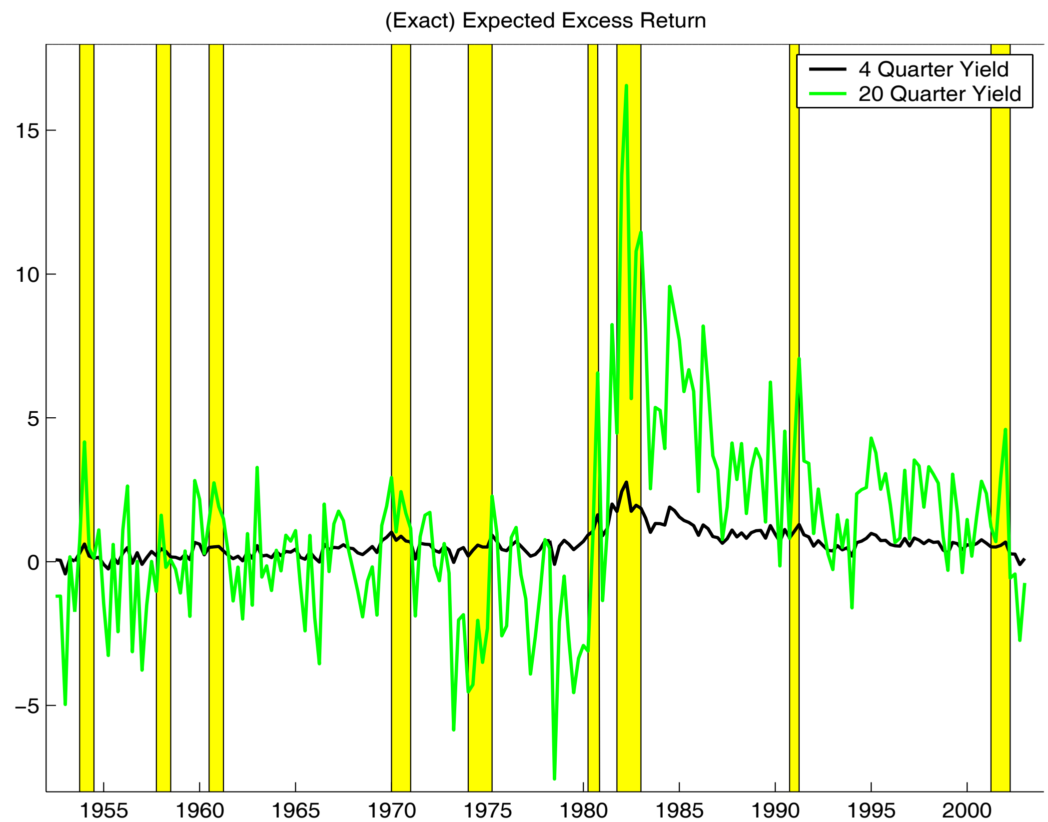
Note: The table reports summary statistics of OLS estimates and model-implied estimates of various Taylor rule specifications. We denote the benchmark Taylor rule as "Bmk" the forward-looking rule as "Fwd," the backward-looking Taylor rule as "Bwd," the Taylor rule with serially correlated shocks as "AR," and the forward- and backward-looking Taylor rule as "Fwd-Bwd" The OLS estimates of benchmark and backward-looking Taylor rules are simply the residuals from equations (9) and (10), respectively. The no-arbitrage benchmark shock is the scaled latent factor, $\varepsilon^{MP,T}$ in equation (9). The no-arbitrage backward-looking Taylor policy shocks are the $\varepsilon^{MP,B}$ terms in (13). The no-arbitrage serially correlated policy shocks are the $\varepsilon^{MP,AR}$ terms in (14). The no-arbitrage forward-looking Taylor rule shocks, $\varepsilon^{MP,F}$, are computed from equation (18) for a horizon of $k = 4$ quarters. The no-arbitrage forward- and backward-looking Taylor rule residuals represent $\varepsilon^{MP,FB}$ from equation (21), also with a $k = 4$-quarter horizon. We construct the Romer and Romer (2004) measure of policy shocks converted to a quarterly sequence by summing the monthly Romer-Romer shocks in Table 2 of Romer and Romer (2004). The last two columns report the correlations of the 1-quarter short rate, r, and the change in the 1-quarter short rate, Δr. All monetary policy shocks and the short rate are annualized. The no-arbitrage shocks are computed using the posterior mean of the latent factors. The sample period is June 1952 to December 2002 and the data frequency is quarterly.

Figure 1: LATENT FACTOR, SHORT RATE, AND THE OLS BENCHMARK TAYLOR RULE



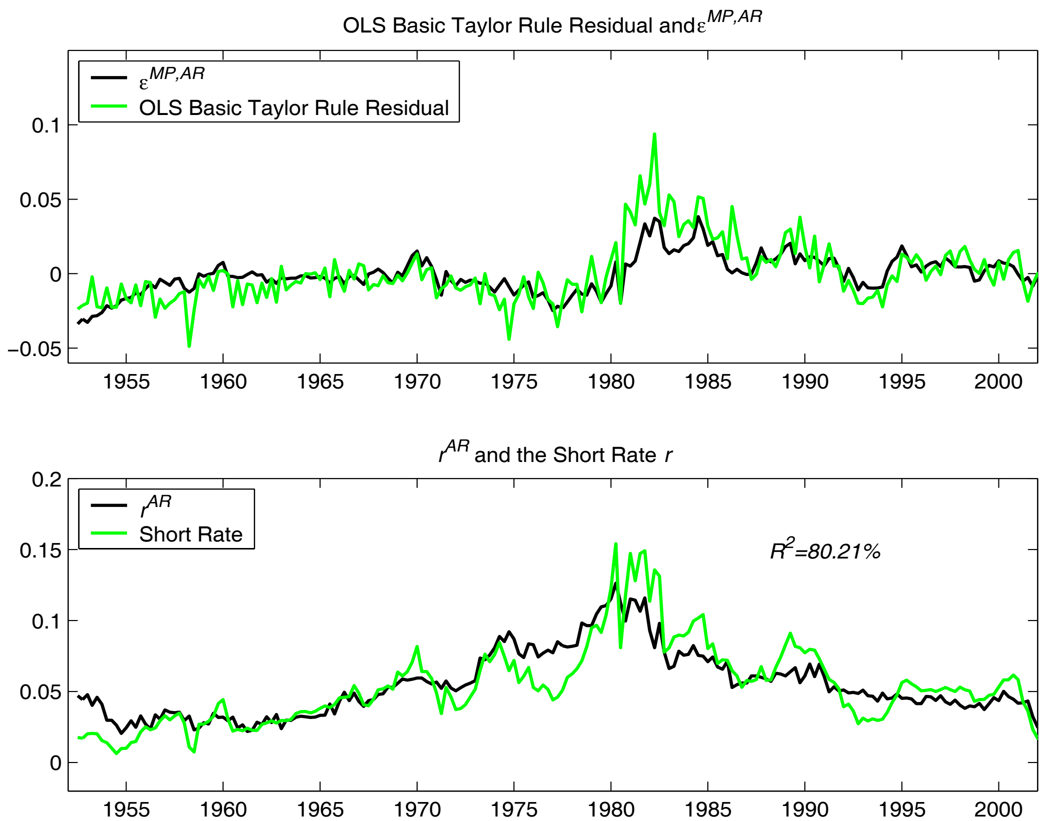
We plot the posterior mean of the latent factor f_t^u, the demeaned short rate from data, and the residuals from the OLS estimate of the basic Taylor Rule, which is computed by running OLS on equation (9). The latent factor, short rate, and OLS residuals are all annualized.

Figure 2: EXPECTED EXCESS BOND RETURNS



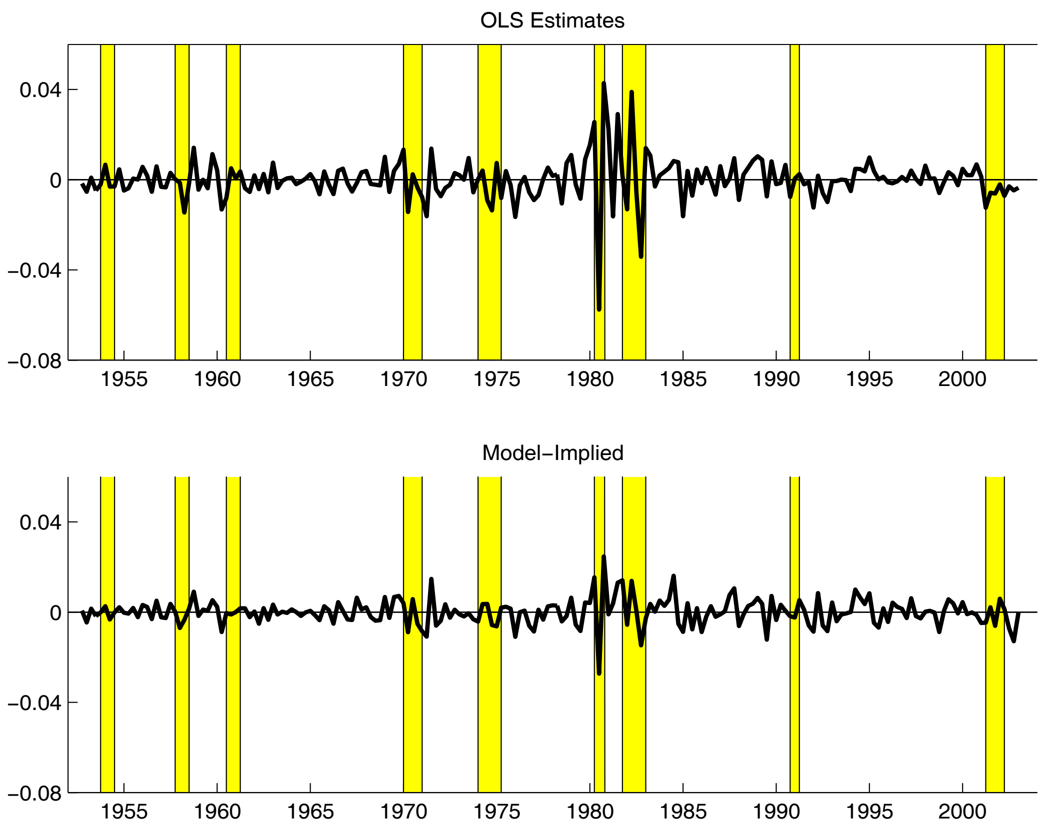
We plot the conditional expected excess holding period return $\mathrm{E}_t[rx_{t+1}^{(n)}]$ of a 4-quarter and 20-quarter bond implied by the posterior mean of the latent factors through time. The numbers on the y-axis are in percentage terms per annum.

Figure 3: SERIALLY CORRELATED MONETARY POLICY SHOCKS



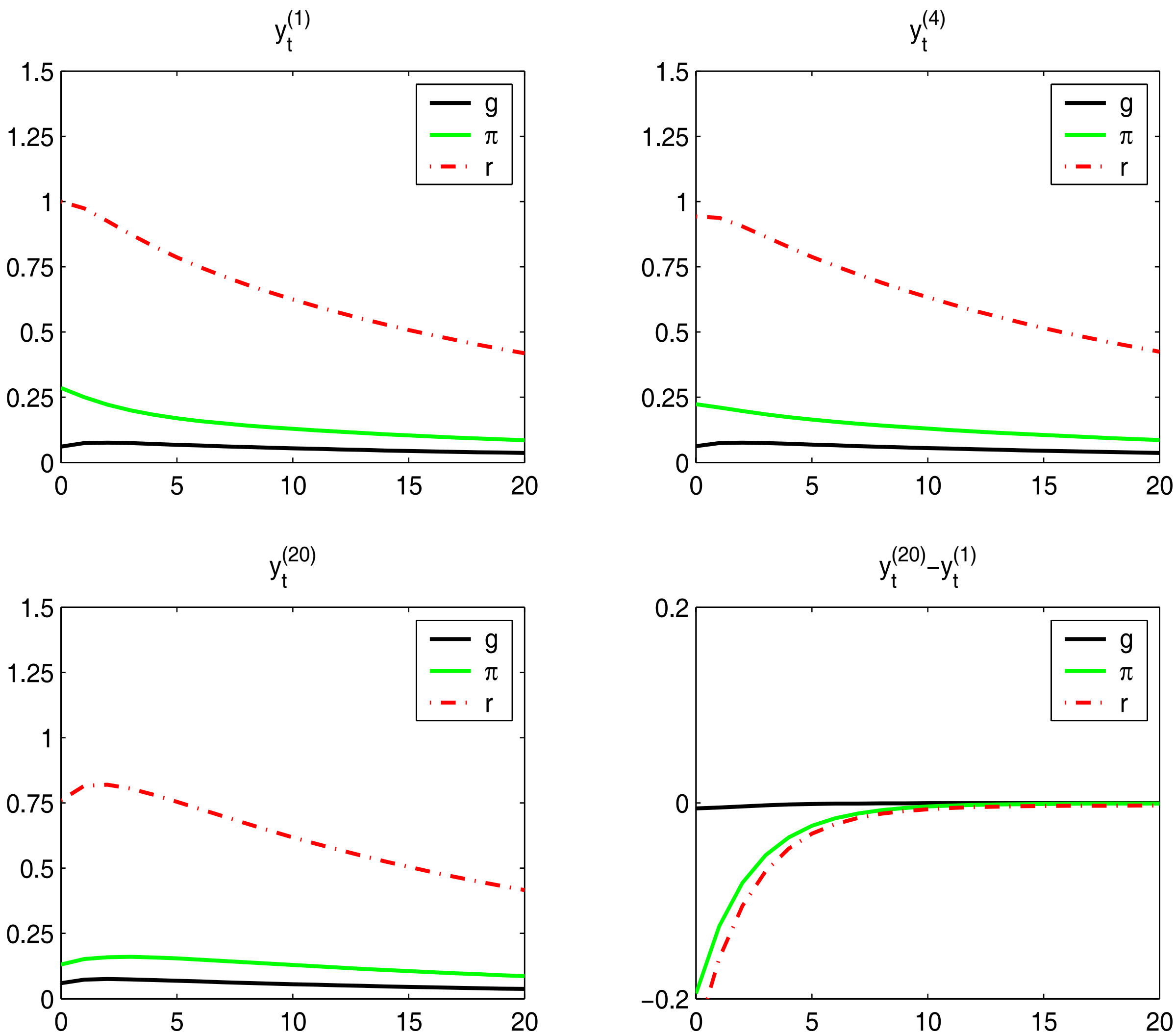
In the top panel, we plot the residuals from the OLS estimate of the basic Taylor Rule, which is computed by running OLS on equation (9) and the posterior mean estimates of monetary policy shocks from a Taylor rule with serially correlated shocks ($\varepsilon_t^{MP,AR}$ in equation (14)). The bottom panel plots the short rate data and r^{AR}, which is the fitted short rate using equation (14), $r_t^{AR} = c_t + \Psi_t(L) f_t^o$. In both the top and bottom panels, we plot annualized numbers.

Figure 4: BACKWARD-LOOKING MONETARY POLICY SHOCKS



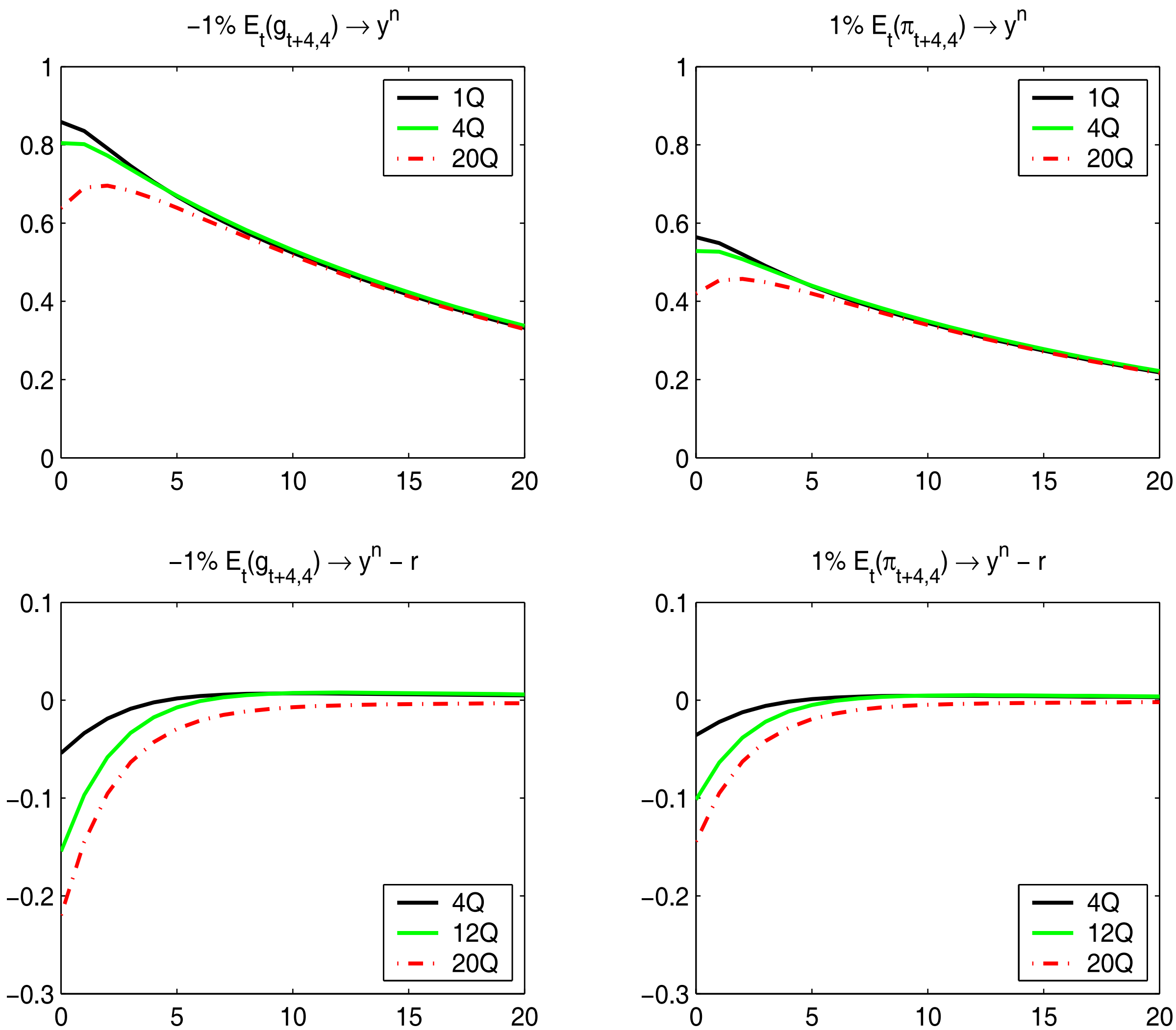
In the top panel, we plot the OLS estimates of the residuals of the backwards-looking Taylor rule (10). The bottom panel plots the corresponding model-implied monetary policy shocks, which are the posterior mean estimates of $\varepsilon_t^{MP,B} = \delta_{12} v_t^u$ from equation (13). In both the top and bottom panels, we plot annualized monetary policy shocks. NBER recessions are shown as shaded bars.

Figure 5: IMPULSE RESPONSES OF YIELDS I



The panels show responses of the one-, four- and twenty-quarter yield, and the term spread between the twenty- and one-quarter yields to 1% shocks to GDP growth g_t and inflation π_t. We also plot the response of a 1% shock in the short rate, which is computed by constructing a shock to the state variables proportional to their Cholesky decomposition that sums to a 1% short rate shock. Yields on the y-axis are annualized and we show quarters on the x-axis. The impulse responses are computed using a Cholesky decomposition that orders the variables (g_t, π_t, f_t^u).

Figure 6: IMPULSE RESPONSES OF YIELDS II



We plot the responses of yields of 1, 4, and 20 quarter maturities, and the corresponding term spreads to -1% shocks to expected GDP growth $E_t(g_{t+4,4})$ and inflation $E_t(\pi_{t,t+4})$ over the next year. We compute the 1% shocks to expected GDP growth or inflation by constructing a shock to the state variables proportional to their Cholesky decomposition that sums to a -1% $E_t(g_{t+4,4})$ or a 1% $E_t(\pi_{t+4,4})$ shock. Yields on the y-axis are annualized and we show quarters on the x-axis. We use a Cholesky decomposition that orders the variables (g_t, π_t, f_t^u).