No-Arbitrage Taylor Rules*

Andrew Ang†
Columbia University and NBER

Sen Dong
Lehman Brothers

Monika Piazzesi‡
University of Chicago, FRB Minneapolis, NBER and CEPR

September 2007

*We thank Ruslan Bikbov, Sebastien Blais, Dave Chapman, Mike Chernov, John Cochrane, Charlie Evans, Michael Johannes, Andy Levin, David Marshall, Thomas Philippon, Tom Sargent, Martin Schneider, George Tauchen, and John Taylor for helpful discussions. We especially thank Bob Hodrick for providing detailed comments. We also thank seminar participants at the American Economics Association, American Finance Association, a CEPR Financial Economics meeting, the CEPR Summer Institute, the European Central Bank Conference on Macro-Finance, the Federal Reserve Bank of San Francisco Conference on Fiscal and Monetary Policy, an NBER Monetary Economics meeting, the Society of Economic Dynamics, the Western Finance Association, the World Congress of the Econometric Society, Bank of Canada, Carnegie Mellon University, Columbia University, European Central Bank, Federal Reserve Board of Governors, Lehman Brothers, Morgan Stanley, PIMCO, and the University of Southern California for comments. Andrew Ang and Monika Piazzesi both acknowledge financial support from the National Science Foundation. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

†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

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

No-Arbitrage Taylor Rules

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 output gap account for over half of the variation of time-varying excess bond returns and most of the movements in the term spread. Taylor rules estimated with no-arbitrage restrictions differ from Taylor rules estimated by OLS, and the resulting monetary policy shocks are somewhat 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-term 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 the short-term interest rate by reacting to CPI inflation and the output gap. To exploit the cross-equation restrictions on yield movements implied by the assumption of no arbitrage, we place the Taylor rule into 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 fully specified general equilibrium model.

We describe expectations of future short rates by versions of 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, Campbell and Shiller, 1991; 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, Dai and Singleton, 2002) .

We present a setup that embeds 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 output gap to influence the actions of the Fed (for example, Eichenbaum and Evans, 1995; Christiano, Eichenbaum and Evans, 1996), and forward-looking Taylor rules where the Fed responds to anticipated inflation and output gap (Clarida, Galí and Gertler, 2000). 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 output gap, 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 output gap, 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, 74% of the variance of the 1-quarter yield and 66% of the variance of the 5-year yield can be attributed to movements in inflation and the output gap. Over 78% of the variance of the 5-year term spread is due to time-varying output gap and output gap risk. The estimated model also captures the counter-cyclical properties of time-varying expected excess returns on bonds.

We estimate Taylor rules following the large macro literature that uses low frequencies (we use quarterly data) at which the output gap and inflation are reported. Under the cross-equation restrictions for yields implied by the no-arbitrage model, we estimate a flexible specification for the macro and latent factors. This setup offers a natural solution to the usual identification problem in VAR dynamics that contain financial data, such as bond yields (for example, Evans and Marshall 1998, 2001; Piazzesi 2005). The Fed's endogenous policy reactions are described by the Taylor rule as movements in the short rate which can be traced to movements in the macro variables that enter the rule: inflation and output. While the Fed may take current yield data into account, it does so only because current yields contain information about future values of these macro variables.

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), Rudebusch and Wu (2005), and Hördahl, Tristani, and Vestin (2006) estimate structural models with interest rates and macro variables. In contrast to these studies, we do not impose any structural restrictions other than the absence of arbitrage. This makes our approach more closely related to the identified VAR literature in macroeconomics (for a survey, see Christiano, Eichenbaum and Evans, 1999) and this provides us additional flexibility in matching the dynamics of the term structure. Other non-structural term structure models with macro factors, like Ang and Piazzesi (2003), and Dewachter and Lyrio (2006), among many others, also do not investigate how no-arbitrage restrictions can help estimate different policy rules.

We do not claim that no-arbitrage techniques are superior to estimating monetary policy rules using structural models. 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.

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. 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 various Taylor rules can be identified in the no-arbitrage model.

2.1 General Set-up

Our state variables are the output gap at quarter t, g_t; the continuously compounded year-on-year inflation rate from quarter $t-4$ to t, π_t; and a latent term structure state variable, f^u_t. We measure year-on-year inflation using the GDP deflator. Our system uses four lags of the output gap and year-on-year inflation variables but parsimoniously captures the dynamics of the latent factor with only one lag. This specification is flexible enough to match the autocorrelogram of year-on-year inflation and the output gap at a quarterly frequency. We assume that in the full state vector, X_{t-1}, potentially up to four lags of the output gap and inflation Granger-cause g_t and π_t, but only the first lag of the variables, g_{t-1}, π_{t-1}, f^u_{t-1}, Granger-cause the latent factor f^u_t. Below we show that this assumption is not restrictive (for example, in the sense of matching impulse responses.) Thus, we can write the dynamics of the state variables as:

$$\begin{aligned} f^o_t &= \mu_1 + \begin{pmatrix} \Phi_{11} & \Phi_{12} \end{pmatrix} \begin{pmatrix} f^o_{t-1} \\ f^u_{t-1} \end{pmatrix} + \begin{pmatrix} \Phi_{13} & \Phi_{14} & \Phi_{15} \end{pmatrix} \begin{pmatrix} f^o_{t-2} \\ f^o_{t-3} \\ f^o_{t-4} \end{pmatrix} + v^o_t \\ f^u_t &= \mu_2 + \begin{pmatrix} \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} f^o_{t-1} \\ f^u_{t-1} \end{pmatrix} + v^u_t, \end{aligned} \tag{1}$$

for $f^o_t = [g_t \; \pi_t]^\top$ the vector of observable macro variables, the output gap and inflation, f^u_t the latent factor, and

$$v_t = \begin{pmatrix} v^o_t \\ v^u_t \end{pmatrix} \sim \text{ IID } N(0, \Sigma_v \Sigma_v^\top).$$

For ease of notation, we collect the four lags of all the state variables in a vector of $K = 12$ elements:

$$X_t = \begin{bmatrix} g_t & \pi_t & f^u_t & \dots & g_{t-3} & \pi_{t-3} & f^u_{t-3} \end{bmatrix}^\top,$$

and write the VAR in equation (1) in companion form as:

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

where

$$\varepsilon_t = \begin{pmatrix} v_t \\ 0_{9\times 1} \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \Sigma_v & 0_{3\times 9} \\ 0_{9\times 3} & 0_{9\times 9} \end{pmatrix}$$

and μ and Φ collect the appropriate conditional means and autocorrelation matrices of the VAR in equation (1), respectively.

We use only one latent state variable because this is the most parsimonious set-up with Taylor rule residuals (as the next section makes clear). This latent factor, f^u_t, is a standard latent factor in the tradition of the term-structure literature. Our focus is to show how this factor is related to monetary policy and how the no-arbitrage restrictions can identify various policy rules.

We specify the short rate equation to be:

$$r_t = \delta_0 + \delta_1^\top X_t, \tag{3}$$

for δ_0 a scalar and δ_1 a $K \times 1$ vector. To keep the model tractable, our baseline system has only contemporaneous values of g_t, π_t and f_t^u and no lags of these factors determining r_t, so only the first three elements of δ_1 are non-zero.

To complete the model, we specify the pricing kernel to take the standard form:

$$m_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t^\top \lambda_t - \lambda_t^\top \varepsilon_{t+1}\right), \tag{4}$$

with the prices of risk:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{5}$$

for the $K \times 1$ vector λ_0 and the $K \times K$ matrix λ_1. To keep the number of parameters down, we only allow the rows of λ_t that correspond to current variables to differ from zero. We specify

$$\lambda_0 = \begin{pmatrix} \bar{\lambda}_0 \\ 0_{9\times 1} \end{pmatrix},$$

where $\bar{\lambda}_0$ is a 3×1 vector. Likewise, we specify that the time-varying components of the prices of risk λ_1, depends on current and past values of macro variables, but only the contemporaneous value of the latent factor: $[g_t \;\; \pi_t \;\; f_t^u \;\; g_{t-1} \;\; \pi_{t-1} \;\; g_{t-2} \;\; \pi_{t-2} \;\; g_{t-3} \;\; \pi_{t-3}]^\top$. That is, we can write:

$$\lambda_1 = \begin{pmatrix} \bar{\lambda}_1 \\ 0_{9\times 12} \end{pmatrix},$$

where $\bar{\lambda}_1$ is a 3×12 matrix with zero columns corresponding to f_{t-1}^u, f_{t-2}^u and f_{t-3}^u.

The pricing kernel determines the prices of zero-coupon bonds in the economy from the recursive relation:

$$P_t^{(n)} = \mathrm{E}_t[m_{t+1} P_{t+1}^{(n-1)}], \tag{6}$$

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.

2.2 Bond Prices and Expected Returns

The model (2)-(5) belongs to the Duffie and Kan (1996) affine class of term structure models, but incorporates both latent and observable macro factors. The model implies that bond yields take the form:

$$y_t^{(n)} = a_n + b_n^\top X_t, \tag{7}$$

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 $K \times 1$ 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:

$$\begin{aligned} A_{n+1} &= A_n + B_n^\top(\mu - \Sigma\lambda_0) + \frac{1}{2} B_n^\top \Sigma\Sigma^\top B_n - \delta_0 \\ B_{n+1}^\top &= B_n^\top(\Phi - \Sigma\lambda_1) - \delta_1^\top, \end{aligned} \tag{8}$$

where $A_1 = -\delta_0$ and $B_1 = -\delta_1$. The recursions (8) are derived by Ang and Piazzesi (2003). In terms of notation, the one-period yield $y_t^{(1)}$ is the same as the short rate r_t in equation (3).

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:

$$\begin{aligned} rx_{t+1}^{(n)} &= \log\left(\frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}\right) - r_t \\ &= ny_t^{(n)} - (n-1)y_{t+1}^{(n-1)} - r_t. \end{aligned} \tag{9}$$

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

$$\begin{aligned} \mathrm{E}_t[rx_{t+1}^{(n)}] &= -\frac{1}{2} B_{n-1}^\top \Sigma\Sigma^\top B_{n-1} + B_{n-1}^\top \Sigma\lambda_0 + B_{n-1}^\top \Sigma\lambda_1 X_t \\ &\equiv A_n^x + B_n^{x\top} X_t, \end{aligned} \tag{10}$$

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 $K \times 1$ vector $B_n^x = \lambda_1 \Sigma^\top B_{n-1}$. From equation (10), 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 describes the Fed as adjusting 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:

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi} \pi_t + \varepsilon_t^{MP,T}, \tag{11}$$

where the short rate is set by the Federal Reserve in response to current output and inflation. The basic Taylor rule (11) can be interpreted as the short rate equation (3) 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 (3) in the term structure model with

$$\delta_1 \equiv \begin{pmatrix} \delta_{1,g} \\ \delta_{1,\pi} \\ \delta_{1,u} \\ 0_{9\times1} \end{pmatrix} = \begin{pmatrix} \gamma_{1,g} \\ \gamma_{1,\pi} \\ \gamma_{1,u} \\ 0_{9\times1} \end{pmatrix},$$

which has zeros for all coefficients on lagged g and π.

The Taylor rule (11) can be estimated consistently using OLS under the assumption that $\varepsilon_t^{MP,T}$, or f_t^u, is contemporaneously uncorrelated with the output gap and inflation. This assumption is satisfied if the output gap 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 (11). 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 use cross-equation 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 (11) 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 (11) 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 (11). 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. For example, 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:

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi} \pi_t + \gamma_{2,g} g_{t-1} + \gamma_{2,\pi} \pi_{t-1} + \gamma_{2,r} r_{t-1} + \varepsilon_t^{MP,B}, \tag{12}$$

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 (3) to take the same form as equation (12). Using the notation f_t^o and f_t^u to refer to the observable macro and latent factors, respectively, we can rewrite the short rate dynamics (3) as:

$$r_t = \delta_0 + \delta_{1,o}^\top f_t^o + \delta_{1,u} f_t^u, \tag{13}$$

where

$$\delta_1 \equiv \begin{pmatrix} \delta_{1,o} \\ \delta_{1,u} \\ 0_{9\times 1} \end{pmatrix}.$$

Using equation (1), we can substitute for f_t^u in equation (13) to obtain:

$$r_t = (1 - \Phi_{22})\delta_0 + \delta_{1,u}\mu_2 + \delta_{1,o}^\top f_t^o + (\delta_{1,u}\Phi_{21}^\top - \delta_{1,o}\Phi_{22}^\top)^\top f_{t-1}^o + \Phi_{22} r_{t-1} + \varepsilon_t^{MP,B}, \quad (14)$$

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 (14) 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 $\varepsilon_t^{MP,B}$. Equating the coefficients in equations (12) and (14) allows us to identify the structural coefficients as:

$$\begin{aligned}
\gamma_0 &= (1 - \Phi_{22})\delta_0 + \delta_{1,u}\mu_2 \\
\begin{pmatrix} \gamma_{1,g} \\ \gamma_{1,\pi} \end{pmatrix} &= \delta_{1,o} \\
\begin{pmatrix} \gamma_{2,g} \\ \gamma_{2,\pi} \end{pmatrix} &= (\delta_{1,u}\Phi_{21}^\top - \delta_{1,o}\Phi_{22}^\top) \\
\gamma_{2,r} &= \Phi_{22}.
\end{aligned} \quad (15)$$

Interestingly, the response to contemporaneous output gap and inflation captured by the $\delta_{1,o}$ coefficient on f_t^o in the backward-looking Taylor rule (14) is identical to the response in the benchmark Taylor rule (11), because the $\delta_{1,o}$ coefficient is unchanged. The intuition behind this result is that the short rate equation (3) 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 (14) and the benchmark Taylor rule (11) 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), the v_t^u shocks are IID. In comparison, the shocks in the standard Taylor rule (11), $\varepsilon_t^{MP,T}$ are highly autocorrelated. Note that the coefficients on lagged macro variables in the extended Taylor rule (14) 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 (12) from the estimated coefficients. This is done by using the implied dynamics of f_t^u in the factor dynamics (1):

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

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}, for $j \geq 1$, in equation (14) to obtain:

$$r_t = c_t + \Psi_t(L) f_t^o + \varepsilon_t^{MP,AR}, \tag{16}$$

where c_t is a scalar, $\Psi_t(L)$ 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 in equation (1). 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 (14), 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 (16).[1]

[1] 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.

2.6 Forward-Looking Taylor Rules

Finite-Horizon, Forward-Looking Taylor Rules

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

$$r_t = \gamma_0 + \gamma_{1,g} \mathrm{E}_t(g_{t+1}) + \gamma_{1,\pi} \mathrm{E}_t(\pi_{t+1}) + \varepsilon_t^{MP,F}, \tag{17}$$

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 (17) into the framework of an affine term structure model as follows. The conditional expectations of future output gap and inflation are simply a function of current X_t that can be computed from the state dynamics (2):

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

Denoting e_i as a vector of zeros with a one in the ith position, we can write equation (17) as:

$$r_t = \gamma_0 + (\gamma_{1,g} e_1 + \gamma_{1,\pi} e_2)^\top \mu + (\gamma_{1,g} e_1 + \gamma_{1,\pi} e_2)^\top \Phi X_t + \varepsilon_t^{MP,F}, \tag{18}$$

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

Equation (18) is an affine short rate equation where the short rate coefficients are a function of the parameters of the dynamics of X_t:

$$r_t = \bar{\delta}_0 + \bar{\delta}_1^\top X_t, \tag{19}$$

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 (8) 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 (2)-(5), except we use the new short rate equation (19) that embodies the forward-looking structure in place of the basic short rate equation (3). To the extent that lagged values of the output gap and inflation help forecast their own future values, the vector $\bar{\delta}_1$ now has nonzero elements corresponding to the coefficients on lagged macro variables. The relations in equation (19) explicitly show that the forward-looking Taylor rule structural coefficients (γ_0, $\gamma_{1,g}$, $\gamma_{1,\pi}$) impose 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 output gap and inflation that enter in the short rate equation (19). 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 (2). The monetary policy shocks in the forward-looking Taylor rule (17) or (18), $\varepsilon_t^{MP,F}$, can only be consistently estimated by OLS if f_t^u is orthogonal to the dynamics of g_t and π_t.

Since k-period ahead conditional expectations of output gap 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 output gap or inflation over the next k quarters:

$$r_t = \gamma_0 + \gamma_{1,g}\mathrm{E}_t(g_{t+k,k}) + \gamma_{1,\pi}\mathrm{E}_t(\pi_{t+k,k}) + \varepsilon_t^{MP,F}, \tag{20}$$

where $g_{t+k,k}$ and $\pi_{t+k,k}$ represent output gap 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 (20) also nests the benchmark Taylor rule (11) as a special case by setting $k = 0$. Appendix A details the appropriate transformations required to map equation (20) 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 that the Fed discounts both expected future output gap and expected future inflation at the same discount rate, β. In this formulation, the forward-looking Taylor rule takes the form:

$$r_t = \gamma_0 + \gamma_{1,\hat{g}}\hat{g}_t + \gamma_{1,\hat{\pi}}\hat{\pi}_t + \varepsilon_t^{MP,F}, \tag{21}$$

where $\hat{g}_t$ and $\hat{\pi}_t$ are infinite sums of expected future output gap 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, among others, Rudebusch and Svenson, 1999).

We can estimate the discount rate β as part of a standard term structure model by using the dynamics of X_t in equation (2) to write $\hat{g}_t$ as:

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

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

$$\hat{\pi}_t \equiv \sum_{i=0}^{\infty} \beta^i e_2^\top \mathrm{E}_t(X_{t+i}) = \frac{\beta}{(1-\beta)} e_2^\top (I - \Phi\beta)^{-1}\mu + e_2^\top (I - \Phi\beta)^{-1} X_t.$$

To place the forward-looking rule with discounting in a term structure model, we re-write the short rate equation (3) as:

$$r_t = \hat{\delta}_0 + \hat{\delta}_1^\top X_t, \tag{22}$$

where

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

Similarly, we modify the bond price recursions for the standard affine model in equation (8) by using the new $\hat{\delta}_0$ and $\hat{\delta}_1$ coefficients that embody restrictions on β, γ_0, $\gamma_{1,\hat{g}}$, $\gamma_{1,\hat{\pi}}$, 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 output gap over the next quarter ($k = 1$), but similar rules apply for other horizons.

We start with the standard forward-looking Taylor rule in equation (17):

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

where $\mathrm{E}_t(f^o_{t+1}) = [\mathrm{E}_t(g_{t+1})\ \mathrm{E}_t(\pi_{t+1})]^\top$ and $\varepsilon^{MP,F}_t = \gamma_{1,u} f^u_t$. We substitute for f^u_t using equation the implied short rate equation (19) that is implied by the forward-looking Taylor rule (17):

$$\begin{aligned} r_t &= \gamma_0 + \gamma_{1,u}\mu_2 - \frac{\gamma_{1,u}\Phi_{22}\bar{\delta}_0}{\bar{\delta}_{1,u}} + \gamma_{1,o}^\top \mathrm{E}_t(f^o_{t+1}) + \frac{\gamma_{1,u}\Phi_{22}}{\bar{\delta}_{1,u}} r_{t-1} \\ &+ \gamma_{1,u}\Phi_{21}^\top f^o_{t-1} + \frac{\gamma_{1,u}\Phi_{22}}{\bar{\delta}_{1,u}}(\bar{\delta}_1^\top X_{t-1} - \bar{\delta}_{1,u} f^u_{t-1}) + \varepsilon^{MP,FB}_t, \end{aligned} \tag{23}$$

where $\bar{\delta}_{1,u}$ is the coefficient on f^u_t in $\bar{\delta}_1$.

Equation (23) 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, $\varepsilon^{MP,FB}_t = \gamma_{1,u} v^u_t$. The forward- and backward-looking Taylor rule (23) is an equivalent representation of the forward-looking Taylor rule in (17). Similar to how the coefficients on contemporaneous macro variables in the backward-looking Taylor rule (14) are identical to the coefficients in the benchmark Taylor rule (11), the coefficients on future expected macro variables in the forward- and backward-looking Taylor rule are exactly the same as the corresponding coefficients in the forward-looking Taylor rule.

2.8 Summary of Taylor Rules

We can identify several structural policy rules 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. 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 B.

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 2004. The consumer price index and real GDP numbers are taken from the Federal Reserve Database (FRED) at Saint Louis. The output gap is computed by applying the Hodrick and Prescott (1997) filter on quarterly real GDP using a smoothing parameter of 1,600. When we estimate the model, we divide the Hodrick-Prescott output gap measure by 4 so that all the variables are expressed in per quarter units.

In Figure 1, we plot the output gap, inflation and the short rate (all expressed in annual units) over time and indicate recessions in solid bars defined by the NBER. As expected, each recession coincides with decreases in the output gap. Inflation and the short rate are strongly positively correlated, at 70%, with both inflation and the short rate peaking during the early and mid-1970s and the monetary targeting period from 1979-1983. In contrast, the short rate is weakly correlated with the output gap, at 19%. Unconditionally, the output gap and inflation are almost uncorrelated, with a correlation of 1%, but this does not capture the stronger lead-lag effects of output and inflation in the VAR, which we show below.

3.2 Estimation and Identification

The VAR dynamics for the state vector in equation (1) are homoskedastic, and since bond yields (7) in our model are linear in the state vector, they are also Gaussian. We deal with potential time variation in volatilities and other parameters such as policy-rule coefficients (as documented by Clarida, Galí and Gertler, 2000) by estimating the model over different subsamples. This approach assumes that bond investors form their expectations in equation (6) based on recent data. They do not take into account that the economy may return to a previously observed "regime." For example, investors during the high-inflation Volcker years did not anticipate that there would be a return to a low-inflation regime under Greenspan.

We estimate the term structure model using Markov Chain Monte Carlo (MCMC) and Gibbs sampling methods. We assume that all yields are observed with error, so that the equation for each yield is:

$$\hat{y}_t^{(n)} = y_t^{(n)} + \eta_t^{(n)}, \tag{24}$$

where $y_t^{(n)}$ is the model-implied yield from equation (7) 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 $\sigma_\eta^{(n)}$.

A major advantage of the Bayesian estimation method is that it 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 measurement error (and the other yields to have non-zero measurement 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.

Another advantage of our estimation method is tractability. Although the likelihood function of yields and related variables can be written down, the model has high dimension and is non-linear in the parameters. 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. In a Bayesian estimation setting, we can also 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 to constrain our state-space system to be stationary.

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. Section 4.4 discusses the implied no-arbitrage monetary policy shocks.

4.1 Parameter Estimates

Table 2 presents the parameter estimates of the term structure model (1)-(5). The first row of the companion form Φ shows that the output gap is significantly forecasted by the first lag of inflation. Similarly, a high lagged output gap significantly Granger-causes high current inflation. In the third row of Φ, both the lagged output gap and lagged inflation significantly predict the latent 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. Naturally, the diagonal coefficients on the first lag reveal that all the variables are highly autocorrelated.

With four lags of the output gap and inflation, many coefficients for the output gap and inflation corresponding to lags 3 to 4 are insignificant. Including the four lag structure is, however, necessary for the model to provide sufficient flexibility for the model to fit year-on-year inflation with a quarterly frequency model. For example, the effect of the relatively large negative coefficient on the second lag of inflation predicting current inflation can only be captured by adding complicated moving average error terms to a VAR system with only one lag.

In Table 2, the estimated covariance matrix $\Sigma_v \Sigma_v^\top$ shows that the innovations to inflation and the output gap are lowly correlated. 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) – where macro factors do not respond contemporaneously to policy shocks – is automatically satisfied, but not a priori imposed, and therefore not restrictive at our parameter estimates.

The short-rate coefficients in δ_1 are all positive, so higher inflation and output gap lead to increases in the short rate, which is consistent with the basic Taylor-rule intuition. In particular, a 1 percent increase in the output gap leads to an increase a 51 basis point (bp) increase in the short rate, while the effect of a 1 percent increase in contemporaneous inflation leads to a 24bp increase of the short rate. Below, we compare these magnitudes with OLS estimates of the Taylor rule.

The risk premia parameters $\bar{\lambda}_0$ and $\bar{\lambda}_1$ indicate that macro-factor risk is significantly priced by the yield curve. There are significant constant prices of risk for g and π in $\bar{\lambda}_0$. There are also many significant prices of time-varying risk in the $\bar{\lambda}_1$ matrix for all three factors. Hence, the output gap, inflation, and the latent factor will play important roles in driving time-varying expected excess returns, as we show below.

The standard deviations of the measurement errors are fairly large. For example, the measurement-error standard deviation of the one-quarter yield (20-quarter yield) is 18bp (6bp) per quarter. This is not surprising, because our system only has one latent factor. Interestingly, short rates have the largest measurement-error variance. This finding suggests that the standard approach of backing out latent factors from data on selected yields by constraining these yields to have zero measurement errors may lead to misspecification, especially at the short end of the yield curve. Indeed, Piazzesi (2005) documents evidence for such misspecification by showing that short rates implied by standard three-factor models are only weakly correlated with those in the data.

Finally, to summarize the dynamics of the VAR, we plot impulse responses of $(g_t\, \pi_t\, r_t)^\top$ implied by the model in the left column of Figure 2. Note that the model VAR is specified in terms of $(g_t\, \pi_t\, f_t^u)^\top$, so to compute the effects of a 1-percent shock on r_t, we invert the appropriate shock to f_t^u so that the shocks from $(g_t\, \pi_t\, f_t^u)^\top$ sum to 1-percent in the short rate equation (3) using a Cholesky decomposition that orders the variables as $(g_t\, \pi_t\, f_t^u)^\top$. For comparison, we contrast the model-implied impulse responses with the impulse responses computed from an empirical VAR(4) on $(g_t\, \pi_t\, r_t)^\top$ in the right column of Figure 2. The empirical VAR allows all lags of r_t to be non-zero, unlike the model-implied VAR, which constrains lags 2 to 4 of f_t^u to be zero (see equation (1)).

The impulse responses generated by our model and the empirical VAR are very similar. In both the model and the empirical VAR, inflation and the short rate increase after a positive shock to the output gap, while the short rate increases after an inflation shock. However, inflation dampens immediately after a 1% shock to r in the responses generated by our model, while the empirical VAR has a very weak price puzzle (see comments by Sims, 1992) as inflation initially slightly increases and then drops below zero about 10 quarters later. There is no price puzzle in the model-implied VAR dynamics. Overall, we conclude that limiting the model VAR to exclude lags of the latent factor as in equation (1) is inconsequential as it captures the same macro variable dynamics.

Latent Factor Dynamics

The monetary policy shocks identified by no arbitrage depend crucially on the behavior of the latent factor, f_t^u. Figure 3 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 output gap (inflation) is -0.10 (0.61). The corresponding correlation implied by the model posterior mean point estimates is -0.08 (0.61), which is very similar to the correlations computed using the posterior mean of the latent factor. These strong correlations suggest that simple OLS estimates of the Taylor rule (11) may be biased in small samples, which we investigate below. The correlations between f_t^u and the yields range between 94% (the short rate) and 98% (the 20-

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 near zero.

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. 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 output gap.

Panel B shows that the autocorrelations in data increase from 0.932 for the short rate to 0.962 for the 5-year yield. In comparison, the model-implied autocorrelations exhibit a smaller range in point estimates from 0.964 for the short rate to 0.962 for the 2-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.

Panel C shows that the model is able to match the correlation of the short rate with output gap and inflation present in the data. The correlation of the short rate with f_t^u implied by the model is 0.941. 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 (7), 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 order $[g_t \, \pi_t \, f_t^u]$.

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 (7) 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 $\Omega^{F,h} = \text{var}(X_{t+h} - \text{E}_t(X_{t+h}))$. 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 30%. As the maturity increases, the importance of the risk premia increases. Panel B shows that risk premia matter even more for yield spreads. Over one half 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, the shocks to macro variables explain about 65-75% of the total variance of yield levels. Shocks to inflation are slightly more important than shocks to output gap in explaining the forecast variance of yield levels. In the pure risk premia term, the proportion of variance attributable to output gap and inflation is also around 50%.

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 the macro variables are by far the main driving force of yield spreads, with the unexplained latent factor portion being generally less than 10%. In particular, shocks to output gap explain more than 62% of the variance of yield spreads and inflation shocks account for approximately 20% of the unconditional variance of the 5-year spread.

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 is about 50% for all maturities. Inflation is a little more important for explaining time-varying excess returns than output gap, with the proportion for inflation reaching close to 33% for the 20-quarter bond. Thus, inflation and inflation risk impressively account for over one half of the dynamics of expected excess returns.[2]

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, say, the 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):

$$arx_{t+1} = \log \frac{P_{t+1}^{(n)}}{P_t^{(n)}} - r_t. \tag{25}$$

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. The time-varying prices of risk are essential in this good fit. If λ_1 is set

[2] We also estimated a simpler system using quarter-on-quarter GDP growth, quarter-on-quarter inflation (measured using the GDP deflator), and a latent factor with only one lag in the VAR. This system produces similar variance decomposition attributions for yield levels and expected excess holding period returns, but assigns higher variance decompositions to inflation than Table 4. This is because the output gap is more persistent than GDP growth. Nevertheless, the proportion of the risk premia for yield levels, yield spreads, or excess returns, are very similar using either output gaps or GDP growth.

to zero, the model's ability to match excess returns deteriorates substantially, with the model implied mean (standard deviation) of the excess return on the 20-quarter bond changing to 0.63% (2.78%), compared to 0.23% (3.30%) in the data.

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 the output gap and inflation both increase in magnitude with maturity, indicating that long bond excess returns are more affected by macro factor variation. Note that if λ_1 were set to zero, the coefficients in the Panel B regressions would be zero, ignoring approximation and Jensen's inequality terms.

Panel C reports the coefficients of the conditional (exact) expected excess holding period return defined in equation (10). To make the coefficients easier to interpret, we report the summed coefficients on all lags of g and π. The B_n^x coefficients on the contemporaneous and lagged output gap and inflation are negative, indicating that conditional expected excess returns are counter-cyclical. High output gaps and high inflation rates are more likely to occur during the peaks of economic expansions, so excess returns of holding long-term 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 4 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 12% per year during the 1982 recession and drop below -5% during 1953, 1973 and, 1978. In contrast, expected excess returns for the 4-quarter bond lie in a more narrow range between -1.3% and 3.8% 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 4.0% per year in March 1981 to 12.9% per year in December 1981.[3]

[3] At 1981:Q4, the 12.9% expected excess return for the 20-quarter bond can be broken down into the various proportions: 174% to the constant term premium, 10% to the output gap, -65% to inflation, and -19% to the latent

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 various specifications.

The Benchmark Taylor Rule

Panel A of Table 6 contrasts the OLS and model-implied estimates of the benchmark Taylor rule in equation (11). Over the full sample, the OLS estimate of the output coefficient is 0.34, and this is highly significant. The model-implied coefficient is similar in magnitude at 0.51. The OLS estimate of the inflation coefficient is over twice as large at 0.90 and also strongly significant. In contrast, the model-implied coefficient on π_t of 0.24 is much smaller than the OLS estimate. Hence, OLS over-estimates the response of the Fed on the short rate by over two-thirds compared to the model-implied estimate.

There are two main reasons for the differences between the standard OLS estimate and the model-implied coefficients. First, the model accounts for the endogenous fluctuations in inflation and output, which are correlated with monetary policy shocks. This contemporaneous correlation causes OLS coefficients to be biased in small samples. Second, 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. These numbers must be carefully compared since OLS regressions produce classical standard errors, while Bayesian estimations produce posterior standard deviations, but we see that our estimation produces tighter posterior standard deviations than regular OLS standard errors.

To further investigate 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 0.005 for the constant, 0.37 (0.91) for the output gap (inflation) – almost identical to the OLS regression coefficients. Moreover, the model-implied OLS regression R^2 is 57%, very similar to the OLS R^2 of 52%. These results suggest that the larger magnitude of the OLS regression estimate of the inflation coefficient in the benchmark

factor. Note that there is a large exposure, in absolute values, to the 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.

Taylor rule compared to the model-implied coefficient is due to an omitted variable that is correlated with output gap and inflation.

By estimating the model over the full sample, we follow Christiano, Eichenbaum and Evans (1996), and others, who all assume that the Taylor rule relationships are stable. Interestingly, our results for the benchmark Taylor rule are indeed fairly stable when we estimate the model across different subsamples. (Below we show that this result does not apply to other specifications of the Taylor rule.) 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.[4] For example, the model (OLS) coefficient on inflation is 0.23 (0.98) over the pre-1982 sample and 0.52 (1.83) over the post-1983 sample, compared with 0.24 (0.90) over the whole sample. The model coefficients on output gap are also fairly stable, at 0.41 (0.46) over the pre-1982 (post-1982) period. In contrast, the OLS coefficient on output gap differs widely across the samples, ranging from 0.28 in the pre-1982 sample to 0.55 in the post-1982 sample. Hence, the OLS coefficients of output gap 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 (14), the model coefficients on g_t and π_t are unchanged from the benchmark Taylor rule in Panel A at 0.51 and 0.24, respectively. The corresponding OLS estimates of the backward-looking Taylor rule coefficients on output gap and inflation are 0.38 and 0.32, respectively. Here, OLS estimates of the backward-looking rule are closer to the model-implied estimates compared to the benchmark Taylor rule, particularly for contemporaneous and lagged inflation.

As expected, the coefficients on the lagged short rate in both the OLS estimates and the model-implied estimates are similar to the autocorrelation of the short rate (0.93 in Table 3). The large and significant coefficient on the lagged short rate reflects interest-rate smoothing. We can rewrite the backward-looking Taylor rule in partial-adjustment format as follows:

$$r_t = (1 - 0.911)(5.719 g_t + 2.674\pi_t - 4.180 g_{t-1} - 1.393\pi_{t-1}) + 0.911 r_{t-1} + \varepsilon_t^{MP,B}.$$

[4] Several recent studies have emphasized that the linear policy coefficients on the output gap and inflation potentially vary over time (see, among others, Cogley and Sargent, 2001). However, other authors like Sims and Zha (2006) find either little or no evidence for time-varying policy rules.

Hence, our model implies a long-run response to inflation of $2.674 - 1.393 = 1.281$. This is consistent with the Taylor principle that the coefficient on inflation should be larger than one (see comments by Taylor, 1999).

As already mentioned above, we find that the estimates of the backward-looking Taylor rule change across subsamples. Panel B in Table 6 shows that interest-rate smoothing is more important in the post-1983 sample than in the pre-1982 sample. The coefficient on the lagged interest rate goes up from 0.87 to 0.94 in the model estimation. We also find that the model estimation finds a higher long-run response to inflation in the more recent sample than in the earlier sample. These findings – more recently, both interest-rate smoothing and the inflation response have become stronger – are consistent with those by Clarida, Galí and Gertler (2000).

The Taylor Rule with Serially Correlated Shocks

Figure 5 plots the monetary policy shocks of the Taylor rule with serially correlated errors (see equation (16)) 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 (16), 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 71%. Thus, although short rates do not resemble contemporaneous output gap or inflation, in a serially correlated Taylor rule, the entire past history of output gap 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 (20) for various horizons k. For all the forward-looking Taylor rules, we re-estimate the term structure model holding the μ, Φ, and Σ in equation (2) fixed to the same values as the benchmark estimation, but only report the forward-looking Taylor rule coefficients for comparison. A different estimation is performed for each horizon k. By holding the VAR parameters constant across the specifications, we concentrate only on the effect of different horizons in the forward-looking Taylor rule, but allow best fits to the prices of risk across the term structure.

For a one-quarter ahead forward-looking Taylor rule, the coefficient on expected output gap (inflation) is 0.59 (0.29). These are larger than the contemporaneous responses for output gap and inflation over the past quarter in the benchmark Taylor rule, which are 0.51 and 0.24, respectively. For a one-year ($k = 4$) horizon, the short interest rate responds even more aggressively to output gap and inflation expectations, with $\gamma_{1,g} = 0.74$ and $\gamma_{1,\pi} = 0.37$. The response of the Fed to future inflation expectations increases quickly as the horizon increases.

As k increases, the posterior standard deviations increase so that the $\gamma_{1,g}$ and $\gamma_{1,\pi}$ coefficients become less precisely estimated. As k becomes large, the conditional expectations approach their unconditional expectations, or $\mathrm{E}_t(g_{t+k,k}) \to \mathrm{E}(g_t)$ and $\mathrm{E}_t(\pi_{t+k,k}) \to \mathrm{E}(\pi_t)$. Econometrically, this makes $\gamma_{1,g}$ and $\gamma_{1,\pi}$ hard to identify for large k, and unidentified in the limit as $k \to \infty$. The intuition behind this result is that as $k \to \infty$, the only variable driving the dynamics of the short rate in equation (20) 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 in the limit as $k \to \infty$, it is impossible to differentiate the (scaled) effect of output gap or inflation expectations from γ_0.

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 output gap (inflation) is 0.17 (0.05). The discount rate $\beta = 0.857$, which implies an effective horizon of $1/(1 - 0.857)$ quarters, or 1.75 years. This estimate is much lower than the discount rates close to 0.99 used in the literature (see, for example, 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 two years is consistent with transcripts of FOMC meetings, which indicate that the Fed often weighs forecasts and policy scenarios of up to three to five years ahead.

The Forward- and Backward-Looking Taylor Rule

For completeness, Panel E of Table 6 reports the estimates of the forward- and backward-looking Taylor rule 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 $\mathrm{E}_t(\pi_{t+k,k})$ as explained in Section 2.7. Naturally, the lagged short rate continues to play a large role. The summed coefficients on the lagged output gap and inflation variables cannot be as easily interpreted as the coefficients on the first lag of macro variables in the backward-looking rule. Nevertheless, the relatively large magnitude of these coefficients suggest that for traditional forward-looking Taylor rule estimates which use instruments for $\mathrm{E}_t(g_{t+k,k})$ and $\mathrm{E}_t(\pi_{t+k,k})$, long lags of macro variables would still be necessary to capture the endogenous correlated effects of latent monetary policy shocks.

4.4 An Example of No-Arbitrage Monetary Policy Shocks

The monetary policy shocks identified by no arbitrage 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, and the same reduced-form no-arbitrage model can produce 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 6 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 (14) in the bottom panel. We compute $\varepsilon_t^{MP,B}$ using the posterior mean estimates of the latent factor through time. Figure 6 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 quite similar.

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 output gap 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 this approach is that all these types of Taylor rules can be 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 output gap and inflation shocks account for over half of the time-variation of time-varying excess bond returns and almost all of the movements in the term spread. Macro factors induce a counter-cyclical risk premium for holding long-term bonds. We find that monetary policy rules identified by no-arbitrage are more stable over time than classical estimates of Taylor (1993) rule coefficients. Monetary policy shocks implied by backward-looking policy rules estimated with no-arbitrage restrictions are less volatile than their counterparts estimated by OLS. Interesting extensions of our approach are to impose more structure on the VAR dynamics, to expand the state space to include other macro factors, or to embed the no-arbitrage identification techniques in more structural models.

Appendix

A Forward-Looking Taylor Rules

In this appendix, we describe how to compute $\bar{\delta}_0$, $\bar{\delta}_1$ in equation (19) of a forward-looking Taylor rule for a k-quarter horizon. From the dynamics of X_t in equation (2), 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 (8) are thus based on the short rate equation $r_t = \bar{\delta}_0 + \bar{\delta}_1^\top X_t$, where:

$$\begin{aligned} \bar{\delta}_0 &= \gamma_0 + \frac{1}{k}[\gamma_{1,g}e_1 \;\; \gamma_{1,\pi}e_2]^\top\left(\sum_{i=1}^{k}\tilde{\Phi}_i\right)\mu, \\ \bar{\delta}_1^\top &= \frac{1}{k}[\gamma_{1,g}e_1 \;\; \gamma_{1,\pi}e_2]^\top\tilde{\Phi}_k\Phi + \gamma_{1,u}e_3^\top. \end{aligned} \tag{A-3}$$

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 (20) 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 (2). 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 Estimating the Model

We estimate the model by MCMC with a Gibbs sampling algorithm. 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 (8).

Drawing the Latent Factor f^u

The factor dynamics (2), together with the yield equations (24), 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. 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. 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 (2), 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:

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

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 (8): 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 (2).

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:

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

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

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

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 (B-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:

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

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:

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

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

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

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 (B-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 (24). 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 $\beta^{m+1}$is given by:

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

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	0.00500	0.296	μ_1^Q	0.000015	0.408
μ_2^Q	0.00002	0.324	μ_3^Q	0.000007	0.368
$\Phi_{1,\cdot}^Q$	0.00070	0.389	$\Phi_{2,\cdot}^Q$	0.001000	0.257
$\Phi_{3,\cdot}^Q$	0.00030	0.362			

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 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.

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.

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

$$\delta_0 = \bar{r} - \delta_1^\top \bar{X}, \tag{B-8}$$

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.

References

Ang, A., and M. Piazzesi, 2003, "A No-Arbitrage Vector Autoregression of Term Structure Dynamics with Macroeconomic and Latent Variables," *Journal of Monetary Economics*, 50, 4, 745-787.

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

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

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.

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

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.

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.

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.

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

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.

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

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

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.

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.

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

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

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.

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

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.

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.

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

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.

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.

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

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

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

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

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.

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.

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.

Rudebusch, G., and T. Wu, 2005, "A Macro-Finance Model of the Term Strucutre, Monetary Policy, and the Economy," forthcoming *Economic Journal*.

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

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

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

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

Taylor, J. B., 1999, "A Historical Analysis of Monetary Policy Rules," in Taylor, J. B., ed., *Monetary Policy Rules*, University of Chicago Press, Chicago, 319-340.

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								
	μ	g_{t-1}	π_{t-1}	f^u_{t-1}	g_{t-2}	π_{t-2}	g_{t-3}	π_{t-3}	g_{t-4}	π_{t-4}
g_t	0.022	1.037	0.333	0.011	-0.141	-0.453	-0.153	0.298	-0.069	-0.206
	(0.032)	(0.072)	(0.189)	(0.031)	(0.103)	(0.326)	(0.102)	(0.319)	(0.072)	(0.183)
π_t	0.014	0.060	1.443	-0.012	0.005	-0.449	-0.007	-0.048	0.015	0.040
	(0.000)	(0.028)	(0.073)	(0.012)	(0.040)	(0.127)	(0.039)	(0.124)	(0.028)	(0.071)
f^u_t	-0.081	0.091	0.093	0.911	0	0	0	0	0	0
	(0.025)	(0.032)	(0.026)	(0.024)	–	–	–	–	–	–

	$\Sigma_v\Sigma_v^\top \times 100,000$		
	g	π	f^u
g	0.395	0.013	-0.035
	(0.041)	(0.011)	(0.035)
π	0.013	0.057	-0.007
	(0.011)	(0.006)	(0.012)
f^u	-0.035	-0.007	0.268
	(0.035)	(0.012)	(0.030)

Short Rate Equation

		δ_1		
	δ_0	g_t	π_t	f^u_t
	0.011	0.509	0.238	1.000
	(0.001)	(0.044)	(0.079)	–

Risk Premia Parameters

		$\bar{\lambda}_1$								
	$\bar{\lambda}_0$	g_t	π_t	f^u_t	g_{t-1}	π_{t-1}	g_{t-2}	π_{t-2}	g_{t-3}	π_{t-4}
g	-1.39	84.3	59.7	-74.5	81.6	20.9	-59.6	-42.2	-116	-8.32
	(0.38)	(47.7)	(98.8)	(21.0)	(55.3)	(164)	(53.2)	(161)	(68.6)	(92.8)
π	1.29	25.7	-20.0	-53.5	-143	205	-92.4	-18.4	179	-99.8
	(0.38)	(81.1)	(108)	(66.1)	(73.4)	(190)	(70.5)	(174.9)	(61.0)	(126)
f^u	-0.31	-14.3	0.76	-24.0	12.7	30.2	-4.00	4.21	37.9	-0.26
	(0.22)	(29.5)	(26.1)	(18.1)	(18.0)	(36.5)	(18.6)	(34.5)	(20.7)	(20.6)

Observation Error Standard Deviation

	$n = 1$	$n = 4$	$n = 8$	$n = 12$	$n = 16$	$n = 20$
$\sigma_\eta^{(n)}$	0.177	0.111	0.056	0.034	0.046	0.064
	(0.009)	(0.006)	(0.004)	(0.002)	(0.003)	(0.004)

Note: The table lists parameter values for the model in equations (2)-(5) and observation error standard deviations in equation (24) 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 2004 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	0.000	-0.010	0.400	0.428	0.841	0.847
	(0.051)	(0.078)	(0.033)	(0.047)	(0.041)	(0.024)
π	0.866	0.875	0.567	0.675	0.984	0.986
	(0.086)	(0.509)	(0.066)	(0.256)	(0.026)	(0.005)

PANEL B: MOMENTS OF YIELDS

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
Means						
Data	1.307	1.412	1.464	1.507	1.540	1.560
	(0.108)	(0.109)	(0.108)	(0.105)	(0.104)	(0.103)
Model	1.307	1.402	1.462	1.509	1.540	1.559
	–	(0.006)	(0.003)	(0.002)	(0.002)	(0.004)
Standard Deviations						
Data	0.731	0.733	0.722	0.702	0.694	0.682
	(0.091)	(0.082)	(0.081)	(0.079)	(0.079)	(0.076)
Model	0.721	0.718	0.714	0.706	0.692	0.675
	(0.009)	(0.006)	(0.003)	(0.003)	(0.004)	(0.005)
Autocorrelations						
Data	0.932	0.940	0.949	0.955	0.959	0.962
	(0.031)	(0.030)	(0.028)	(0.027)	(0.027)	(0.025)
Model	0.964	0.963	0.962	0.963	0.963	0.964
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

PANEL C: SHORT RATE CORRELATIONS

	g	π	f^u
Data	0.188	0.696	–
	(0.138)	(0.075)	
r	0.228	0.714	0.941
	(0.124)	(0.134)	(0.031)

Note: Panel A lists moments of the output gap 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 (3). 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 2004 and the data frequency is quarterly.

Table 4: VARIANCE DECOMPOSITIONS

Maturity (qtrs)	Risk Premia	Variance Decompositions: Total g	Total π	Total f^u	Risk Premia g	Risk Premia π	Risk Premia f^u
PANEL A: YIELD LEVELS $y_t^{(n)}$							
1	0.0	35.4	38.4	26.2	–	–	–
4	2.0	32.9	38.4	28.8	22.4	28.9	48.7
8	8.4	30.4	37.8	31.8	11.4	30.9	57.7
12	16.0	28.9	37.7	33.4	11.3	29.0	59.7
16	23.1	28.2	37.8	34.0	12.4	27.8	59.7
20	29.5	27.8	38.0	34.2	13.2	27.1	59.7
PANEL B: YIELD SPREADS $y_t^{(n)} - y_t^{(1)}$							
4	52.7	62.2	28.3	9.5	22.4	28.9	48.7
8	55.7	68.2	21.1	10.8	11.4	30.9	57.7
12	55.4	73.1	19.9	7.0	11.3	29.0	59.7
16	53.2	76.6	20.1	3.4	12.4	27.8	59.7
20	50.2	78.0	20.8	1.1	13.2	27.1	59.7
PANEL C: EXPECTED EXCESS HOLDING PERIOD RETURNS $\mathrm{E}_t(rx_{t+1}^{(n)})$							
4	100	24.3	28.1	47.6	24.3	28.1	47.6
8	100	19.8	30.5	49.7	19.8	30.5	49.7
12	100	18.6	31.6	49.8	18.6	31.6	49.8
16	100	18.2	32.4	49.3	18.2	32.4	49.3
20	100	18.3	32.9	48.8	18.3	32.9	48.8

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	0.102	0.149	0.188	0.212	0.226
	(0.047)	(0.102)	(0.149)	(0.196)	(0.229)
Model	0.095	0.157	0.204	0.236	0.254
	(0.014)	(0.021)	(0.027)	(0.038)	(0.055)
Standard Deviations					
Data	0.748	1.547	2.194	2.783	3.297
	(0.128)	(0.219)	(0.257)	(0.288)	(0.332)
Model	0.612	1.434	2.213	2.936	3.609
	(0.034)	(0.079)	(0.122)	(0.160)	(0.198)

PANEL B: PREDICTABILITY REGRESSIONS

	Data Estimates				Model-Implied Estimates			
	g	π	$y^{(20)}$	R^2	g	π	$y^{(20)}$	R^2
$n = 4$	-0.151	-0.284	0.338	0.060	-0.176	-0.272	0.304	0.093
	(0.108)	(0.145)	(0.117)		(0.107)	(0.099)	(0.083)	
$n = 12$	-0.520	-1.125	1.077	0.077	-0.531	-1.155	1.131	0.096
	(0.320)	(0.430)	(0.359)		(0.385)	(0.361)	(0.305)	
$n = 20$	-0.683	-1.853	1.662	0.082	-0.844	-1.943	1.811	0.096
	(0.485)	(0.650)	(0.536)		(0.631)	(0.597)	(0.503)	

PANEL C: FACTOR COEFFICIENTS

		Maturity (qtrs)				
		4	8	12	16	20
A_n^x		0.004	0.009	0.012	0.015	0.018
B_n^x	g	-0.103	-0.262	-0.444	-0.618	-0.777
	π	-0.313	-0.733	-1.090	-1.411	-1.706
	f^u	0.357	0.791	1.125	1.410	1.669

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 (25)). 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 the output gap, inflation, and the 20-quarter 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, summing the values over all lags of each factor, of the conditional expected excess holding period return defined in equation (10) on the factors. The data frequency is quarterly and the sample period is June 1952 to December 2004.

Table 6: TAYLOR RULES

PANEL A: BENCHMARK TAYLOR RULE

	Full Sample		Pre-82:Q4		Post-83:Q1	
	OLS	Model	OLS	Model	OLS	Model
const	0.005	0.011	0.003	0.011	0.002	0.010
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
g_t	0.338	0.509	0.283	0.503	0.548	0.457
	(0.095)	(0.044)	(0.282)	(0.103)	(0.182)	(0.103)
π_t	0.900	0.238	0.979	0.229	1.825	0.572
	(0.135)	(0.079)	(0.417)	(0.079)	(0.165)	(0.167)

PANEL B: BACKWARD-LOOKING TAYLOR RULE

		const	g_t	π_t	g_{t-1}	π_{t-1}	r_{t-1}	R^2
Full Sample	OLS	0.000	0.380	0.322	-0.268	-0.181	0.866	0.897
		(0.000)	(0.107)	(0.248)	(0.117)	(0.237)	(0.043)	
	Model	0.000	0.509	0.238	-0.372	-0.124	0.911	0.950
		(0.000)	(0.044)	(0.079)	(0.049)	(0.083)	(0.024)	
Pre-82:Q4	OLS	0.001	0.345	0.241	-0.183	-0.011	0.774	0.881
	Model	0.000	0.503	0.229	-0.323	-0.068	0.870	0.938
Post-83:Q1	OLS	0.000	0.516	0.791	-0.461	-0.761	0.960	0.956
	Model	0.000	0.457	0.519	-0.441	-0.403	0.941	0.912

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

	const	$\mathrm{E}_t(g_{t+k,k})$	$\mathrm{E}_t(\pi_{t+k,k})$
$k = 1$	0.011	0.590	0.292
	(0.001)	(0.030)	(0.016)
$k = 4$	0.010	0.741	0.365
	(0.001)	(0.039)	(0.016)
$k = 8$	0.009	0.975	0.486
	(0.001)	(0.078)	(0.016)
$k = 20$	0.007	0.903	0.708
	(0.001)	(0.163)	(0.027)

Table 6 Continued

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

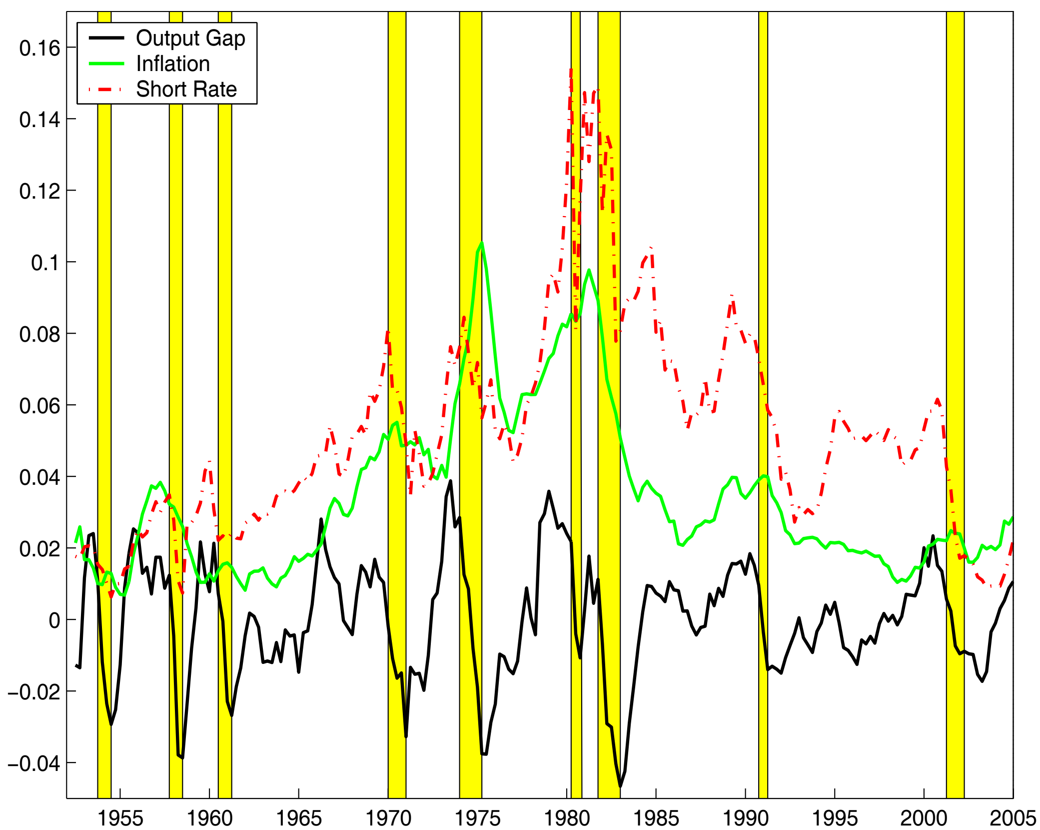
	const	$\hat{g}_t$	$\hat{\pi}_t$	β
$k=\infty$	0.010	0.168	0.054	0.857
	(0.001)	(0.008)	(0.002)	(0.007)

PANEL E: FORWARD- AND BACKWARD-LOOKING TAYLOR RULE

	const	$\mathrm{E}_t(g_{t+k,k})$	$\mathrm{E}_t(\pi_{t+k,k})$	Lags of g	Lags of π	r_{t-1}	R^2
$k=1$	0.000	0.590	0.292	-0.672	-0.785	0.908	0.951
	(0.000)	(0.030)	(0.016)	(0.037)	(0.028)	(0.001)	
$k=4$	-0.001	0.741	0.365	-0.867	-1.318	0.911	0.950
	(0.000)	(0.039)	(0.016)	(0.047)	(0.053)	(0.001)	

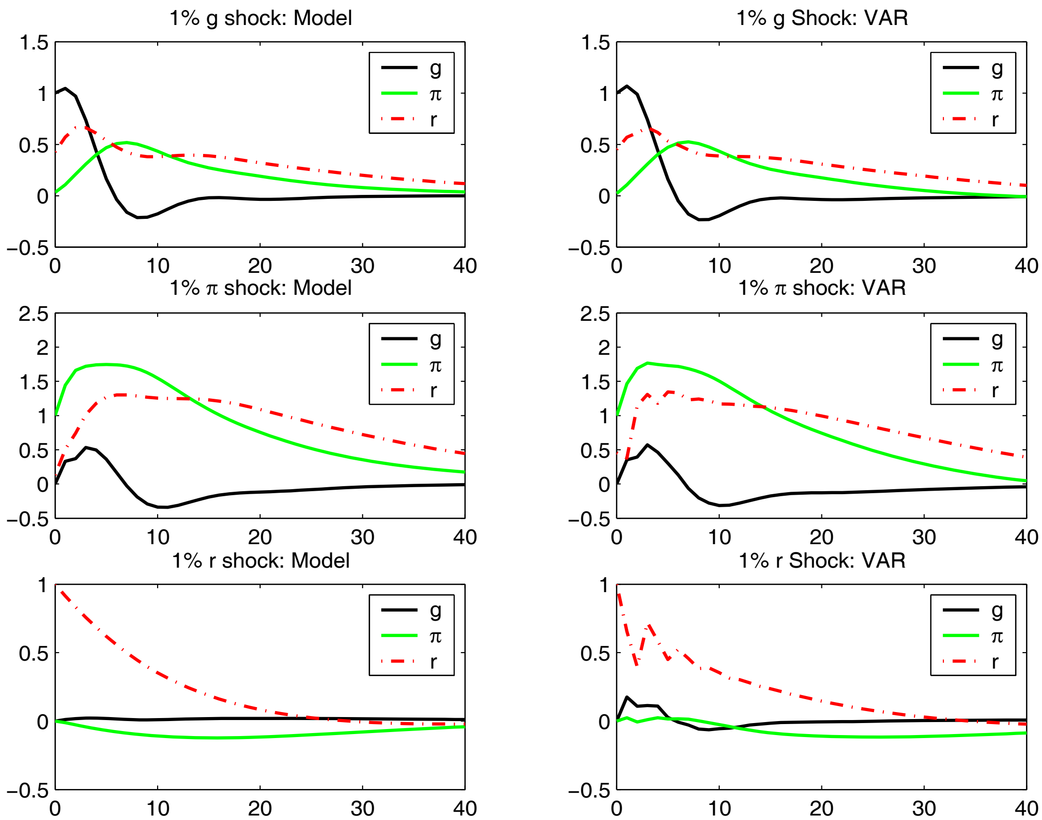
Note: Panel A reports the OLS and model-implied estimates of the benchmark Taylor (1993) rule in equation (11) over the full sample and over subperiods; Panel B reports the backward-looking Taylor rule (12); Panel C reports the finite-horizon, forward-looking Taylor rule without discounting in equation (20); Panel D reports the infinite-horizon, forward-looking Taylor rule with discounting in equation (21); and Panel E reports estimates of the forward- and backward-looking Taylor rule in Section 2.7. In Panel E, we report the sums of the coefficients of the output gap and inflation over all four lags. For the forward-looking Taylor rules in Panels C-E, we hold the estimates of the VAR parameters in equation (2) and re-estimate the forward-looking Taylor rule coefficients together with the prices of risk. 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 2004.

Figure 1: OUTPUT GAP, INFLATION, AND SHORT RATE



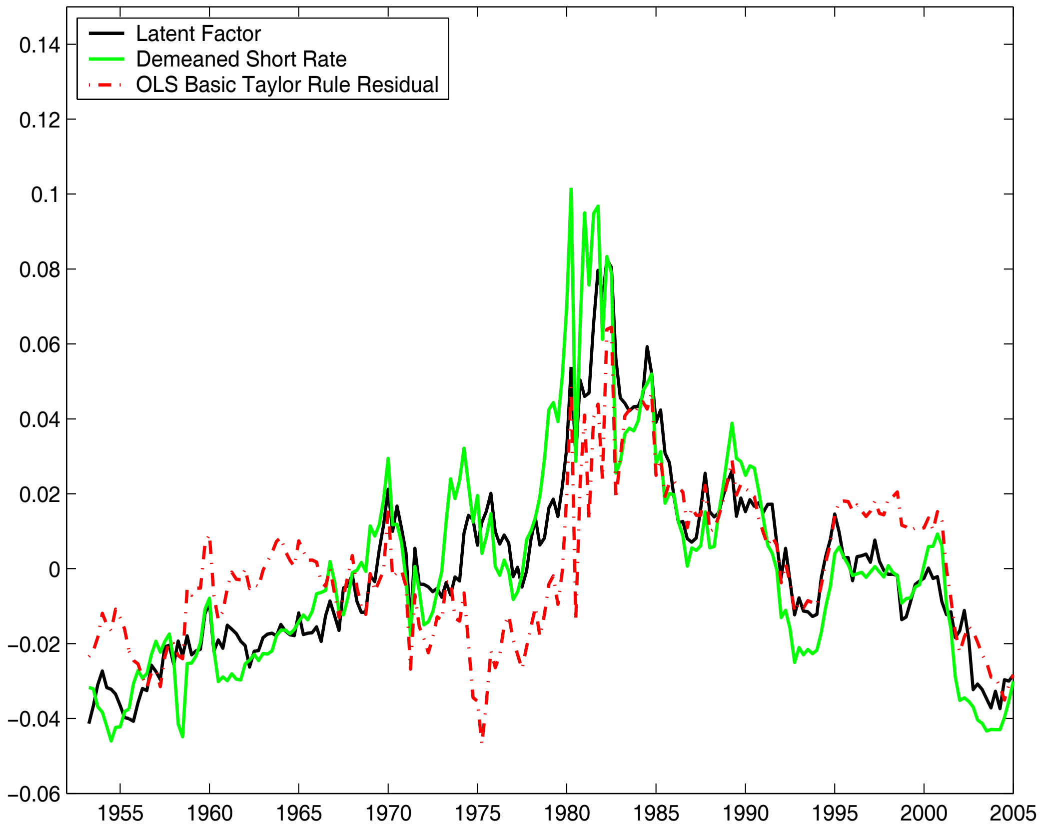
We plot the output gap, year-on-year inflation measured by the GDP deflator and the 1-quarter maturity short rate.

Figure 2: IMPULSE RESPONSE FUNCTIONS



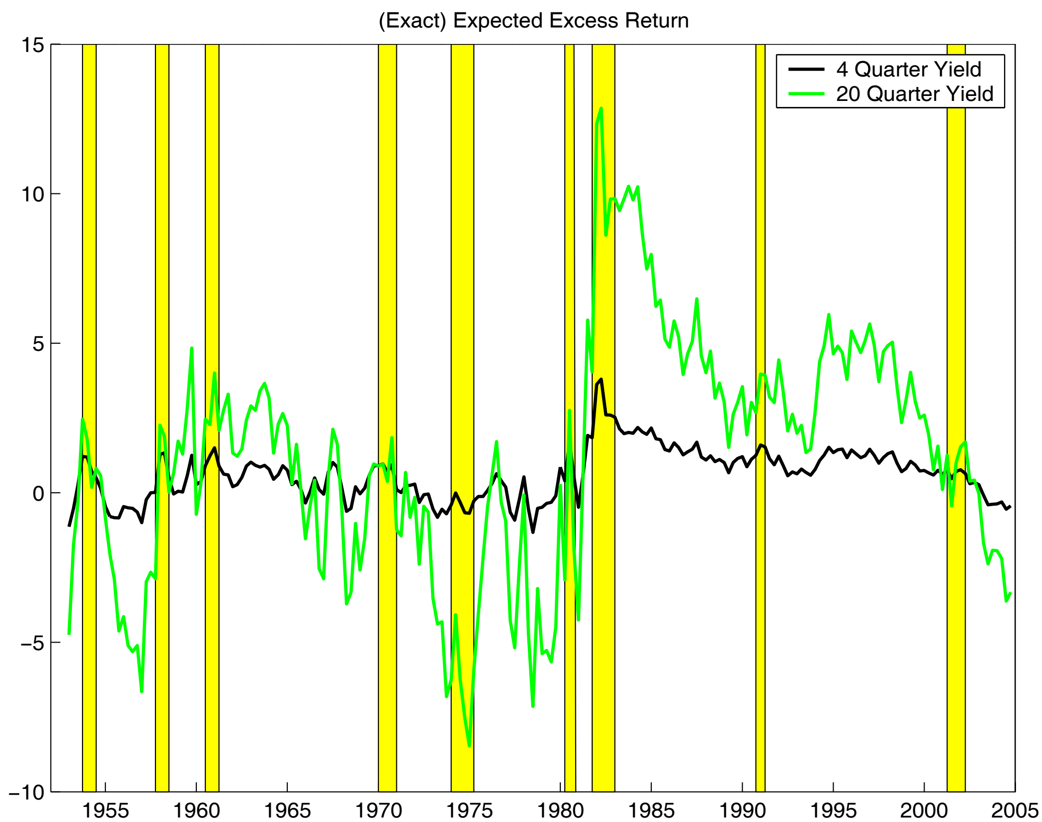
The panels show responses of the output gap, inflation, and the short rate to 1% shocks to the output gap g, inflation π, and the short rate r. The left column shows responses implied by the model, while the right column presents response functions computed from an unrestricted VAR(4) of $(g_t\, \pi_t\, r_t)^\top$. We compute the 1% r shock from the model by scaling an f shock so that the short rate changes by 1%. We show quarters on the x-axis. The impulse responses are computed using a Cholesky decomposition that orders the variables (g, π, f) for the model and (g, π, r) for the unrestricted VAR.

Figure 3: LATENT FACTOR, SHORT RATE, AND THE OLS BENCHMARK TAYLOR RULE



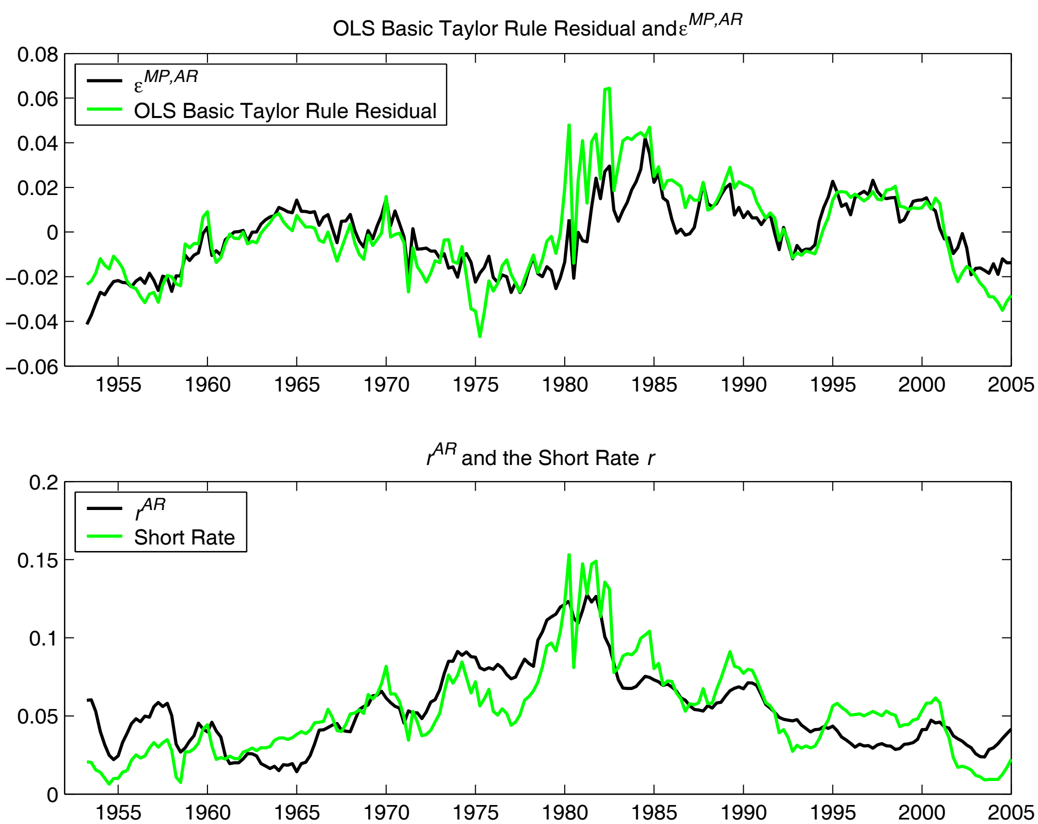
We plot the posterior mean of the latent factor f_t^u, the demeaned short rate from data, and the residuals from the OLS estimate of the basic Taylor Rule, which is computed by running OLS on equation (11). The latent factor, short rate, and OLS residuals are all annualized.

Figure 4: EXPECTED EXCESS BOND RETURNS



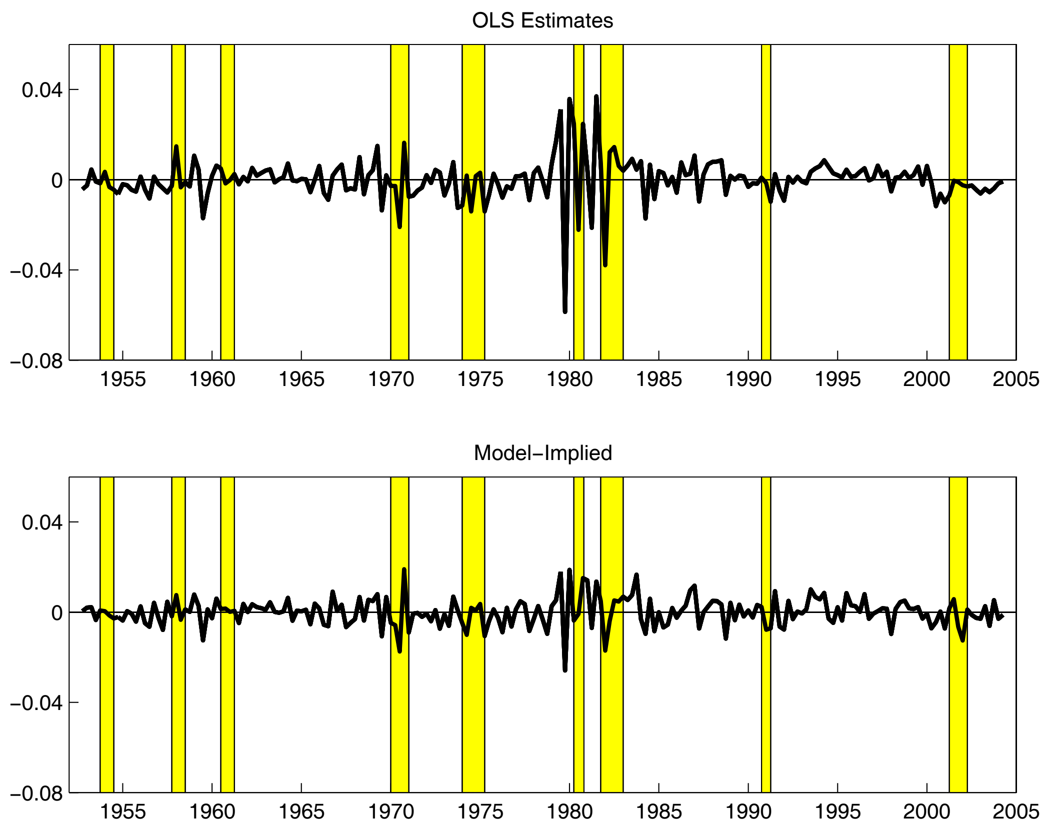
We plot the conditional expected excess holding period return $\mathrm{E}_t[rx_{t+1}^{(n)}]$ of a 4-quarter and 20-quarter bond implied by the posterior mean of the latent factors through time. The numbers on the y-axis are in percentage terms per annum.

Figure 5: 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 (11) and the posterior mean estimates of monetary policy shocks from a Taylor rule with serially correlated shocks ($\varepsilon_t^{MP,AR}$ in equation (16)). The bottom panel plots the short rate data and r^{AR}, which is the fitted short rate using equation (16), $r_t^{AR} = c_t + \Psi_t(L) f_t^o$. In both the top and bottom panels, we plot annualized numbers.

Figure 6: BACKWARD-LOOKING MONETARY POLICY SHOCKS



In the top panel, we plot the OLS estimates of the residuals of the backwards-looking Taylor rule (12). 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 (14). In both the top and bottom panels, we plot annualized monetary policy shocks. NBER recessions are shown as shaded bars.