No-Arbitrage Taylor Rules*

Andrew Ang†
Columbia University, USC and NBER

Sen Dong‡
Columbia University

Monika Piazzesi§
University of Chicago and NBER

Preliminary Version: 15 November 2004

JEL Classification: C13, E43, E52, G12

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

*We thank Ruslan Bikbov, Mike Chernov, John Cochrane, Bob Hodrick, Michael Johannes, and George Tauchen for helpful discussions and seminar participants at Columbia University and USC for useful comments. Andrew Ang and Monika Piazzesi both thank the NSF for financial support.

†Marshall School of Business at USC, 701 Exposition Blvd, Rm 701, Los Angeles, CA 90089-1427; ph: (213) 740-5615; fax: (213) 740-6650; email: aa610@columbia.edu; WWW: http://www.columbia.edu/∼aa610

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

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

Abstract

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

1 Introduction

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

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

We describe expectations of future short rates by the Taylor rule and a Vector Autoregression (VAR) for macroeconomic variables. Following the approach taken in many papers in macroeconomics (see, for example, Fuhrer and Moore, 1995; Cogley, 2003), we could infer the values of long yields from these expectations by imposing the Expectations Hypothesis (EH). However, there is strong empirical evidence against the EH (see, for example, Fama and Bliss, 1987; Campbell and Shiller, 1991; Cochrane and Piazzesi, 2004). Term structure models can account for deviations from the EH by explicitly incorporating time-varying risk premia (see, for example, Fisher, 1998; Dai and Singleton, 2002).

We develop a methodology to embed Taylor rules in an affine term structure model with time-varying risk premia. The structure accommodates standard Taylor rules, backward-looking Taylor rules that allow multiple lags of inflation and GDP growth to influence the actions of the Fed, and forward-looking Taylor rules where the Fed responds to anticipated inflation and GDP growth. The model specifies standard VAR dynamics for the macro indicators, inflation and GDP growth, and an additional latent factor that drives interest rates and is related to 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 interest rates. In particular, we use information from the whole yield curve to obtain more efficient estimates of how monetary policy shocks affect the future path of macro aggregates. Similarly, the term structure model allows us to measure how a yield of any maturity responds to monetary policy or macro shocks. Interestingly, the model implies that a large amount of the interest rate volatility is explained by movements in macro variables. For example, 65% of the variance of the 1-quarter yield and 61% of the variance of the 5-year yield can be attributed to movements in inflation and GDP growth. Over 95% of the variance in the 5-year term spread is due to time-varying inflation and inflation risk. The estimated model also captures the counter-cyclical properties of time-varying expected excess returns on bonds.

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

Our paper is related to a growing literature on linking the dynamics of the term structure with macro factors. Piazzesi (2004) develops down a term structure model, where the Federal Reserve targets the short rate and reacts to information contained in the yield curve. To identify monetary policy shocks, Piazzesi uses data measured at high-frequencies and by assuming that the Fed reacts to information available right before its policy decision, she identifies the unexpected change in the target as the monetary policy shock, and identifies the expected target as the policy rule. In contrast, we estimate Taylor rules following the large macro literature that uses the standard low frequencies (we use quarterly data) at which GDP and inflation are reported. At low frequencies, the Piazzesi identification scheme does not make sense because we would have to assume that the Fed uses only lagged one-quarter bond market information and ignores more recent data.

In contrast, we assume that the Fed follows the Taylor rule, and thus reacts to contemporaneous output and inflation numbers. This identification strategy relies on the reasonable assumption that these macroeconomic variables react only slowly – not within the same quarter – to monetary policy shocks. This popular identification strategy has also been used by Christiano, Eichenbaum, and Evans (1996), Evans and Marshall (1998, 2001) and many others. By using this strategy, we are not implicitly assuming that the Fed completely ignores current and lagged information from the bond market (or other financial markets). To the contrary, yields in our model depend on the current values of output and inflation. Thus, we are implicitly assuming that the Fed cares about yield data, but only to the extent that they provide information about these macro variables.

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), Hördahl, Tristani, and Vestin (2003), and Rudebusch and Wu (2003) estimate structural term structure models with macroeconomic variables. In contrast to these studies, we do not impose any structure in addition to the assumption of no arbitrage, which makes our approach more closely related to the identified VAR literature in macroeconomics (for a survey, see Christiano, Eichenbaum and Evans, 1999). This gives us additional flexibility in matching the dynamics of the term structure. While Bagliano and Favero (1998), and Evans and Marshall (1998, 2001), among others, estimate VARs with many yields and macroeconomic variables, they do not impose no-arbitrage restrictions. Bernanke, Boivin and Eliasz (2004), and Diebold, Rudebusch, and Aruoba (2004) estimate latent factor models with macro variables, but also do not impose no arbitrage on bond yields.

We do not claim that our new no-arbitrage identification techniques are superior to estimating monetary policy rules using structural models (see, among others, Bernanke and Mihov, 1998) or using real-time information sets like central bank forecasts to control for the endogenous effects of monetary policy taken in response to current economic conditions (see, for example, Romer and Romer, 2004). Rather, we believe that identifying monetary policy shocks 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. Naturally, our methodology can be used in more structural approaches that effectively constrain the factor dynamics and risk premia and we can extend our set of instruments to include richer information sets. We intentionally focus on the most parsimonious set-up where Taylor rules can be identified in a no-arbitrage model.

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

2 The Model

We detail the set-up of the model in Section 2.1. Section 2.2 shows how the model implies closed-form solutions for bond prices (yields) and expected returns. In Sections 2.3 to 2.5, we detail how Taylor rules can be identified using the over-identifying restrictions imposed on bond prices through no-arbitrage. We briefly discuss some estimation issues in Section 2.6.

2.1 General Set-up

We denote the 3×1 vector of state variables as

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

where g_t is quarterly GDP growth from $t-1$ to t, π_t is the quarterly inflation rate from $t-1$ to t, and f_t^u is a latent term structure state variable. Both GDP growth and inflation are continuously compounded. We use one latent state variable because this is the most parsimonious set-up where the Taylor rule residuals can be identified (as the next section makes clear) using no-arbitrage restrictions.

We specify that X_t follows a VAR(1):

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

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

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

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

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

with the time-varying prices of risk:

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

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

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

where $P_t^{(n)}$ is the price of a zero-coupon bond of maturity n quarters at time t.

Equivalently, we can solve the price of a zero-coupon bond as

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

where E_t^Q denote 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

$$\mu^Q = \mu - \Sigma\lambda_0$$
$$\Phi^Q = \Phi - \Sigma\lambda_1.$$

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

This model falls into the Duffie and Kan (1996) affine class of term structure models, but uses both latent and observable macro factors. The affine prices of risk specification in equation (4) has been used by, among others, Constantinides (1992), Fisher (1998), Dai and Singleton (2002), Duffee (2002), and Brandt and Chapman (2003) in continuous time and by Ang and Piazzesi (2003) and Dai and Philippon (2004) in discrete time. This flexible specification is able to capture patterns of expected holding period returns on bonds that we observe in the data.

2.2 Bond Prices and Expected Returns

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \mathrm{E}_t(rx_{t+1}^{(n)}) &= \mathrm{E}_t(A_{n-1} + B_{n-1}^\top X_{t+1}) + \frac{1}{2}\mathrm{var}_t(B_{n-1}^\top X_{t+1}) - (A_n + B_n^\top X_t) - (\delta_0 + \delta_1^\top X_t) \\ &= A_{n-1} + B_{n-1}^\top(\mu + \phi X_t) + \frac{1}{2}B_{n-1}^\top \Sigma\Sigma^\top B_{n-1} - (A_n + B_n^\top X_t) - (\delta_0 + \delta_1^\top X_t) \\ &= B_{n-1}^\top \Sigma\lambda_0 + B_{n-1}^\top \Sigma\lambda_1 X_t \\ &= B_{n-1}^\top \Sigma\lambda_t. \end{aligned}$$

From this expression, we can see directly that if λ_1 is zero, expected excess returns do not vary over time. Since both bond yields and the expected holding period returns of bonds are affine functions of X_t, we can also easily compute variance decompositions following standard methods for a VAR.

2.3 The Benchmark Taylor Rule

We can interpret the short rate equation (2) of the term structure model as a Taylor rule of monetary policy. Following Taylor (1993), we define the benchmark Taylor rule as:

$$r_t = \delta_0 + \delta_{1,g} g_t + \delta_{1,\pi}\pi_t + \varepsilon_t^{MP,T}, \tag{8}$$

where the short rate is set by the Federal Reserve to be a function of current output and inflation. The basic Taylor rule (8) can be interpreted as the short rate equation (2) in a standard affine term structure model, where the unobserved monetary policy shock $\varepsilon_t^{MP,T}$ corresponds to a latent term structure factor, $\varepsilon_t^{MP,T} = \delta_{f^u} f_t^u$. This corresponds to the short rate equation (2) in the term structure model with $\delta_1 = (\delta_{1,g}\,\delta_{1,\pi}\,\delta_{1,f^u})^\top$.

The Taylor rule (8) can be estimated consistently using OLS under the assumption that $\varepsilon_t^{MP,T}$, or f_t^u, is contemporaneously uncorrelated with GDP growth and inflation. If monetary policy is effective, policy actions by the Federal Reserve today predict the future path of GDP and inflation, causing an unconditional correlation between monetary policy actions and macro factors. In this case, running OLS on equation (8) may not provide efficient estimates of the Taylor rule. In our setting, we allow $\varepsilon_t^{MP,T}$ to be unconditionally correlated with GDP or inflation and thus our estimates should be more efficient, under the null of no-arbitrage, than OLS. In our model, the coefficients $\delta_{1,g}$ and $\delta_{1,\pi}$ in equation (8) are simply the coefficients on g_t and π_t in the vector δ_1 in the short rate equation (2).

The Taylor rule in equation (8) does not depend on the past level of the short rate. Therefore, empirical studies typically find that the implied monetary policy shock, $\varepsilon_t^{MP,T}$, is highly persistent (see Rudebusch and Svensson, 1999). The reason is that the short rate is highly autocorrelated and its movements are not well explained by the right-hand side variables in equation (8). This makes the implied shock, $\varepsilon_t^{MP,T}$, inherit the dynamics of the level of the short rate and, thus, is highly persistent. 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 $\delta_{1,f^u} f_t^u$, the latent term structure variable. Ang and Piazzesi (2003) show that the first latent factor implied by an affine model with both latent factors and observable macro factors closely corresponds to the traditional first, highly persistent, latent factor in term structure models with only unobservable factors. This latent variable also corresponds closely to the first principal component of yields, or the average level of the yield curve, which is highly autocorrelated.

2.4 Backward-Looking Taylor Rules

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

$$r_t = \delta_0 + \delta_{1,g} g_t + \delta_{1,\pi}\pi_t + \delta_{2,g} g_{t-1} + \delta_{2,\pi}\pi_{t-1} + \delta_{2,r} r_{t-1} + \varepsilon_t^{MP,B}, \tag{9}$$

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 equation (9) may be that the objective of the central bank is to smooth interest rates (see Goodfriend, 1991).

In the setting of our model, we can modify the short rate equation (2) to take the same form as equation (9). Collecting the macro factors g_t and π_t into a vector of observable variables $f_t^o = (g_t\ \pi_t)^\top$, we can rewrite the short rate dynamics in equation (2) as:

$$r_t = \delta_0 + \delta_{11}^\top f_t^o + \delta_{12} f_t^u, \tag{10}$$

where we decompose the vector δ_1 into $\delta_1 = (\delta_{11}^\top\ \delta_{12})^\top$. We also rewrite the dynamics of

$X_t = (f_t^{o\top}\ f_t^u)^\top$ in equation (1) as:

$$\begin{pmatrix} f_t^o \\ f_t^u \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{pmatrix} \begin{pmatrix} f_{t-1}^o \\ f_{t-1}^u \end{pmatrix} + \begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix}, \tag{11}$$

where $u_t = (u_t^{1\top}\ u_t^2)^\top \sim$ IID $N(0, \Sigma\Sigma^\top)$. Equation (11) is equivalent to equation (1), but the notation in equation (11) separates the dynamics of the macro variables, f_t^o, from the dynamics of the latent factor, f_t^u.

Using equation (11), we can substitute for f_t^u in equation (10) to obtain:

$$\begin{aligned} r_t &= \delta_0 + \delta_{11}^\top f_t^o + \delta_{12}(\mu_2 + \Phi_{21} f_{t-1}^o + \Phi_{22} f_{t-1}^u + u_t^2) \\ &= (1 - \Phi_{22})\delta_0 + \delta_{12}\mu_2 + \delta_{11} f_t^o + (\delta_{12}\Phi_{21}^\top - \Phi_{22}\delta_{11})^\top f_{t-1}^o + \Phi_{22} r_{t-1} + \varepsilon_t^{MP,B}, \end{aligned} \tag{12}$$

where we substitute for the dynamics of f_t^u in the first line and where we define the backward-looking monetary policy shock to be $\varepsilon_t^{MP,B} \equiv \delta_{12} u_t^2$ in the second line. Equation (12) 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 $\epsilon_t^{MP,B}$.

In equation (12), the response of the Fed to GDP and inflation captured by the δ_{11} coefficient on f_t^o is identical to the response of the Fed in the benchmark Taylor rule (8) because the δ_{11} coefficient is unchanged. The intuition behind this result is that the short rate equation (2) already embeds the full response of the short rate to current macro factors. The latent factor, however, represents the action of past short rates and past macro factors. We have rewritten the benchmark Taylor rule to equivalently represent the latent factor as lagged macro variables and lagged short rates.

The implied monetary policy shocks from the backwards-looking Taylor rule, $\varepsilon_t^{MP,B}$, are potentially very different from the benchmark Taylor rule 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, $\delta_{12} u_t^2$. In the set-up of the factor dynamics in equation (1) (or equivalently equation (11)), the u_t^2 shocks are IID. In comparison, the shocks in the standard Taylor rule (8), $\varepsilon_t^{MP,T}$ are highly autocorrelated. Note that the coefficients on lagged macro variables in the extended Taylor rule (12) are equal to zero only if $\delta_{12}\Phi'_{21} = \Phi_{22}\delta_{11}$. 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 changed only by unpredictable shocks.

Once our model is estimated, we can easily back out the implied extended Taylor rule (9) from the estimated coefficients. This is done by using the implied dynamics of f_t^u in the factor dynamics (11):

$$u_t^2 = f_t^u - \mu_2 - \Phi_{21} f_{t-1}^o - \Phi_{22} f_{t-1}^u.$$

Again, if $\varepsilon_t^{MP,B} = \delta_{12} u_t^2$ 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 Forward-Looking Taylor Rules

Finite Horizon, Without Discounting

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

$$r_t = \delta_0 + \delta_{1,g} \mathrm{E}_t(g_{t+1}) + \delta_{1,\pi} \mathrm{E}_t(\pi_{t+1}) + \varepsilon_t^{MP,F}, \quad (13)$$

where we define $\varepsilon_t^{MP,F}$ to be the forward-looking Taylor rule monetary policy shock. We now show how $\varepsilon_t^{MP,F}$ can be identified using no-arbitrage restrictions from a term structure model.

We can compute the conditional expectation of GDP growth and inflation from our model by noting that:

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

from the dynamics of X_t in equation (1). Since the conditional expectations of future GDP growth and inflation are simply a function of current X_t, we can map the forward-looking Taylor rule (13) into the framework of an affine term structure model. Denoting e_i as a vector of zeros with a one in the ith position, we can write equation (13) as:

$$r_t = \delta_0 + (\delta_{1,g} e_1 + \delta_{2,\pi} e_2)^\top \mu + (\delta_{1,g} e_1 + \delta_{2,\pi} e_2)^\top \Phi X_t + \varepsilon_t^{MP,F}, \quad (14)$$

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

Equation (14) is an affine short rate equation where the short rate coefficients are a function of the parameters of the dynamics of X_t:

$$r_t = \bar{\delta}_0 + \bar{\delta}_1^\top X_t, \quad (15)$$

where

$$\begin{aligned} \bar{\delta}_0 &= \delta_0 + (e_1 + e_2)^\top \mu \\ \bar{\delta}_1 &= (\delta_{1,g} e_1 + \delta_{2,\pi} e_2)^\top \Phi + \delta_{3,f^u} e_3, \end{aligned}$$

and $\varepsilon_t^{MP,F} \equiv \delta_{1,f^u} f_t^u$. Hence, we can identify a forward-looking Taylor rule by imposing no-arbitrage restrictions by redefining the bond price recursions in equation (6) using the new $\bar{\delta}_0$ and $\bar{\delta}_1$ coefficients in place of δ_0 and δ_1. Hence, a complete term structure model is defined by the same set-up as equations (1) to (4), except we use the new short rate equation (15) that embodies the forward-looking structure, in place of the basic short rate equation (2). The restrictions on δ_0, δ_1, μ and Φ in equation (15) imply that the forward-looking Taylor rule is effectively a constrained estimation of a general affine term structure model.

The new no-arbitrage bond recursions using $\bar{\delta}_0$ and $\bar{\delta}_1$ reflect the conditional expectations of GDP and inflation that enter in the short rate equation (15). Furthermore, the conditional expectations $\mathrm{E}_t(g_{t+1})$ and $\mathrm{E}_t(\pi_{t+1})$ are those implied by the underlying dynamics of g_t and π_t in the VAR process (1). Other approaches, like Rudebusch and Wu (2003), specify the future expectations of macro variables entering the short rate equation in a manner not necessarily consistent with the underlying dynamics of the macro variables. Similar to the monetary policy shocks, $\varepsilon_t^{MP,T}$, in the basic Taylor rule (8), the monetary policy shocks in the forward-looking Taylor rule (13) or (14), $\varepsilon_t^{MP,F}$, 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 GDP and inflation remain affine functions of the current state variables X_t, we can also specify a more general forward-looking Taylor rule based on expected GDP or inflation over the next k quarters:

$$r_t = \delta_0 + \delta_{1,g}\mathrm{E}_t(g_{t+k,k}) + \delta_{1,\pi}\mathrm{E}_t(\pi_{t+k,k}) + \varepsilon_t^{MP,F}, \tag{16}$$

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

$$g_{t+k,k} = \frac{1}{k}\sum_{i=1}^{k} g_{t+i} \quad \text{and} \quad \pi_{t+k,k} = \frac{1}{k}\sum_{i=1}^{k} \pi_{t+i}. \tag{17}$$

The forward-looking Taylor rule monetary policy shock $\varepsilon_t^{MP,F}$ is the scaled latent term structure factor, $\varepsilon_t^{MP,F} = \delta_{1,f^u} f_t^u$. As Clarida, Galí and Gertler (2000) note, the general case (16) also nests the benchmark Taylor rule (8) as a special case by setting $k = 0$. In Appendix A, we detail the appropriate transformations required to map equation (16) into an affine term structure model and discuss the estimation procedure for a forward-looking Taylor rule for a k-quarter horizon.

Infinite Horizon, With Discounting

An alternative approach to assigning a k-period horizon for which the Fed considers future GDP growth and inflation in its policy rule is that the Fed discounts the effect of future economic conditions. For simplicity, we assume the Fed discounts both expected future GDP growth and expected future inflation at the same discount rate, β. In this formulation, the forward-looking Taylor rule takes the form:

$$r_t = \delta_0 + \delta_{1,\hat{g}}\hat{g}_t + \delta_{1,\hat{\pi}}\hat{\pi}_t + \delta_{1,f^u}f_t^u \tag{18}$$

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

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

$$\begin{aligned}
\hat{g}_t &= \sum_{i=0}^{\infty} \beta^i e_1^\top \mathrm{E}_t(X_{t+i}) \\
&= e_1^\top (X_t + \beta\mu + \beta\Phi X_t + \beta^2(I+\Phi)\mu + \beta^2\Phi^2 X_t + \cdots) \\
&= e_1^\top (\mu\beta + (I+\Phi)\mu\beta^2 + \cdots) + e_1^\top (I + \Phi\beta + \Phi^2\beta^2 + \cdots)X_t \\
&= e_1^\top \frac{\beta}{(1-\beta)}(I-\Phi\beta)^{-1}\mu + e_1^\top (I-\Phi\beta)^{-1}X_t,
\end{aligned} \tag{19}$$

where e_1 is a vector or zeros with a one in the first position to pick out g_t, which is ordered first in X_t. We can also write discounted future inflation, $\hat{\pi}_t$, in a similar fashion:

$$\hat{\pi}_t = e_2^\top \frac{\beta}{(1-\beta)}(I-\Phi\beta)^{-1}\mu + e_2^\top (I-\Phi\beta)^{-1}X_t, \tag{20}$$

where e_2 is a vector of zeros with a one in the second position.

In a similar fashion to mapping the Clarida-Gali-Gertler forward-looking Taylor without discounting into a term structure model, we can accommodate a forward-looking Taylor rule with discounting by re-writing the short rate equation (2) as:

$$r_t = \hat{\delta}_0 + \hat{\delta}_1^\top X_t, \tag{21}$$

where

$$\begin{aligned}
\hat{\delta}_0 &= \delta_0 + [\delta_{1,g}e_1 \ \ \delta_{1,\pi}e_2]^\top \left(\frac{\beta}{(1-\beta)}(I-\Phi\beta)^{-1}\mu\right)\mu, \\
\hat{\delta}_1 &= [\delta_{1,g}e_1 \ \ \delta_{1,\pi}e_2]^\top (I-\Phi\beta)^{-1} + \delta_{1,f^u}\, e_3.
\end{aligned} \tag{22}$$

Similarly, we modify the bond price recursions for the standard affine model in equation (6) by using $\hat{\delta}_0$ and $\hat{\delta}_1$ in place of δ_0 and δ_1.

2.6 The Econometric Methodology

The objective of this section is to discuss the econometric method used to estimate the model outlined in Section 2.1. In particular, we motivate our estimation approach and discuss several econometric issues. We relegate all technical issues to Appendix B.

Data

To estimate the model, we use continuously compounded yields of maturities one, four, eight, twelve, sixteen, and twenty quarters, at a quarterly frequency. The longer bond yields (one, two, three, four, and five years) are from the Fama CRSP zero coupon files, while the short maturity rate (one quarter) is taken from the Fama CRSP Treasury Bill files. The sample period is June 1952 to December 2003. The consumer price index and real GDP numbers are taken from the Federal Reserve Database (FRED) at Saint Louis.

Estimation Method

We estimate the term structure model using Markov Chain Monte Carlo (MCMC) and Gibbs sampling methods. There are three main reasons why we choose to use a Bayesian estimation approach. First, the term structure factor, f_t^u, and the corresponding monetary policy shocks implied by f_t^u, are unobserved variables. In a Bayesian estimation strategy, we obtain a posterior distribution of the time-series path of f_t^u and monetary policy shocks. That is, the Bayesian algorithm provides a way to 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.

The second advantage of our estimation method is that, although the maximum likelihood function of the model can be written down (see Ang and Piazzesi, 2003), the model is high dimensional and extremely non-linear. This causes the maximum likelihood function to have many possible local optima, some of which may lie in unreasonable or implausible regions of the parameter space. In our Bayesian setting, using uninformative priors on reasonable regions of the parameter space effectively rules out parameter values that are implausible. A maximum likelihood estimator also involves a difficult optimization problem, whereas the Bayesian algorithm is based on a series of simulations that are computationally much more tractable.

Third, in a situation with only one yield and one latent factor, the maximum likelihood function has a point mass at zero for the set of parameter values that assign a one-to-one correspondence between the observed yield and the latent factor. In this set of parameter values, there is effectively zero effect of macro variables on the dynamics of interest rates, and the yield is driven entirely by the latent factor that takes on the same dynamics as the yield itself. Specifically, in the maximum likelihood function, the coefficients δ_{11} on the observable macro variables in equation (10) may tend to go to zero, and the feedback coefficients between the latent factor and the macro variables in the VAR equation (1) may also tend to go to zero. A similar problem occurs in our setting with a cross-section of yields and one latent factor, where a maximum likelihood estimator may assign almost all explanatory power to the latent factor and assign little role to the macro factors. Given that there must be some underlying economic relation between bond prices and macro variables, we have strong priors that this set of parameters is not a reasonable representation of the true joint dynamics of term structure and macro variables. A Bayesian estimation avoids this stochastic singularity by a suitable choice of priors.

Observation Error

An affine term structure model can only exactly price the same number of yields as the number of latent factors. In our case, the model in equations (1)-(4) can only price one yield exactly since we use only one latent factor, f^u. The usual estimation approach, following Chen and Scott (1993), is to specify some (arbitrary) yield maturities to be observed without error, and the remaining yields to have observation, or measurement, error. We do not arbitrarily impose observation error across certain yields. Instead, we assign an observation error to each yield, so that the equation for each yield is:

$$\hat{y}_t^{(n)} = y_t^{(n)} + \eta_t^{(n)}, \tag{23}$$

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

Importantly, by not assigning one arbitrary yield to have zero observation error (and the other yields to have non-zero observation error), we do not bias our estimated monetary policy shocks to have undue influence from only one yield. Instead, the extracted latent factor reflects the dynamics of the entire cross-section of yields. Below, we discuss the effect of choosing an arbitrary yield, like the short rate, to invert the latent factor.

Identification

Since the factor f_t^u is latent, f_t^u can be arbitrarily shifted and scaled to yield an observationally equivalent model. Dai and Singleton (2000) and Collin-Dufresne, Goldstein, and Jones (2003) discuss some identification issues for affine models with latent factors. Our identification strategy is to set the mean of f_t^u to be zero and to pin down the conditional variance of f_t^u. This allows δ_0 and δ_1 to be unconstrained parameters in the short rate equation (2). 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:

$$e_3^\top (I - \Phi)^{-1} \mu = 0, \tag{24}$$

where e_3 is a vector of zero's with a one in the third position.

We parameterize the conditional covariance matrix $\Sigma\Sigma^\top$ to take the form:

$$\Sigma\Sigma^\top = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{12} & \Sigma_{22} & 0 \\ 0 & 0 & c \end{pmatrix}, \tag{25}$$

which allows shocks to the macro factors to be conditionally correlated while the conditional shocks to the latent factor f_t^u are conditionally uncorrelated with g_t and π_t. The form for $\Sigma\Sigma^\top$ in equation (25) can be interpreted as the Taylor residual having no contemporaneous effect on current GDP growth or inflation, which is the same assumption made by Christiano, Eichenbaum and Evans (1999). However, because the companion form Φ allows full feedback, f_t^u is unconditionally correlated with both g_t and π_t. We set $c = 0.05$, which is chosen so that the coefficients δ_1 in the short rate equation (2) are all of the same magnitude.

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

$$\delta_0 = \bar{r} - \delta_1^\top \bar{X}, \tag{26}$$

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 becomes a function of δ_1 and the draws of the latent factor f_t^u.

3 Empirical Results

Section 3.1 discusses the parameter estimates, the behavior of the latent factor, and the fit of the model to data. Section 3.2 investigates what are the driving determinants of the yield curve. We compare benchmark, backward-, and forward-looking Taylor rules in Section 3.3. Sections 3.4 and 3.5 discuss the implied no-arbitrage monetary policy shocks and impulse responses, respectively.

3.1 Parameter Estimates

Table 1 presents the parameter estimates. The first row of the companion form Φ shows that GDP growth can be forecasted by lagged inflation and lagged GDP growth, while the latent term structure factor does not Granger-cause GDP growth. In particular, high inflation predicts lower future GDP, which is consistent with a Phillips curve. The parameter estimates of the second row of Φ shows that term structure information helps to forecast inflation. The large coefficient estimate on lagged inflation also reveals that inflation, even at the quarterly frequency, is highly persistent.

The third row of Φ shows that both inflation and GDP help forecast the latent term structure factor factor. This is consistent with results in Ang and Piazzesi (2003), who show that adding macro variables improves out-of-sample forecasts of interest rates. The large coefficient on the lagged latent factor indicates the f_t^u series is more persistent that inflation. Interestingly, the estimated covariance matrix $\Sigma\Sigma^\top$ indicates that innovations to inflation and GDP growth are positively correlated, whereas high inflation Granger-causes low GDP growth in the conditional mean equation.

The short rate coefficients in δ_1 are all positive, indicating that higher inflation and GDP growth lead to increases in the short rate, which is consistent with the basic Taylor-rule intuition. In particular, a 1% increase in inflation leads to a 32 basis point (bp) increase in the short rate, while the estimated effect of a 1% increase in GDP growth is small at 9bp and not significant. Below, we compare these magnitudes with OLS estimates of the Taylor rule.

The risk premia parameters in λ_1 indicate that risk premia vary significantly over time. Interestingly, we find that risk premia mostly depend on inflation and the latent factor. Although the estimates in λ_1 in the column corresponding to g are of the same order of magnitude, these parameters are insignificant. Hence, we expect inflation and the latent factor to drive time-varying expected excess returns with less of an effect from GDP growth.

The standard deviations of the observation errors are large. This is not surprising, because we only have one latent factor to fit the entire yield curve. Interestingly, the largest variance occurs at the short end of the yield curve, which indicates that treating the short rate as an observable factor may lead to large discrepancies between the true latent factor and the short rate. We further investigate this issue below.

Latent Factor Dynamics

The monetary policy shocks identified using no-arbitrage assumptions depend crucially on the behavior of the latent factor, f_t^u. Figure 1 plots the latent factor together with the OLS Taylor rule residual and the demeaned short rate. We plot the time-series of the latent factor posterior mean produced from the Gibbs sampler. The plot illustrates the strong relationship between these three series. However, note that the behavior of the OLS benchmark Taylor rule residual is more closely aligned with the short rate movements than the latent factor. This indicates that the behavior of monetary policy shocks based on f_t^u will look different to the estimates of Taylor rule residuals estimated by OLS.

To characterize the relation between f_t^u with macro factors and yields more formally, Table 2 reports correlations of the latent factor with various instruments. The table reports the correlations of time-series of the latent factor posterior mean with GDP, inflation, and yields. Table 2 shows that the latent factor is positively correlated with inflation at 49% and slightly negatively correlated with GDP growth at -17%. The correlation between f_t^u and the yield levels ranges between 91% and 98%. Hence, f_t^u can be interpreted as level factor, similar to the findings of Ang and Piazzesi (2003). In comparison, the correlation between f_t^u and term spreads is below 20%.

Interestingly, the correlation between the latent factor and any given yield data series is not perfect. This is because we are estimating the latent factor by extracting information from the entire yield curve, not just a particular yield. The estimation method could have led us to parameter values that minimize observation error on one particular yield and thereby maximize the correlation between f^u and this yield. However, the estimation results indicate that this is not optimal. This suggests that an estimation method based on an observable (arbitrarily chosen) yield like the short rate may give misleading results. Nevertheless, for estimations based on only observable yields, Table 2 gives useful advice. It suggests to pick the longest yield as measured without observation error to proxy for a single underlying latent factor.

Matching Moments

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

Panel B shows that the autocorrelations in data increase from 0.925 for the short rate to 0.959 for the 5-year yield. In comparison, the model-implied autocorrelations exhibit a smaller range in point estimates from 0.955 for the short rate to 0.965 for the 5-year yield. However, the model-implied estimates are well within two standard deviations of the data. The smaller range of yield autocorrelations implied by the model is due to only having one latent factor. Since inflation and GDP have lower autocorrelations than yields, the autocorrelations of the yields are primarily driven by f_t^u.

Panel C shows that the model is able to match the correlation of the short rate with GDP and inflation present in the data. The correlation of the short rate with f_t^u implied by the model is 0.947. 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.

3.2 What Drives the Dynamics of the Yield Curve?

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

Yields and Yield Spreads

Panels A and B of Table 4 report variance decompositions of yield levels and yield spreads, respectively. Panel A shows that shocks to macro variables explain more than 60% of the variance of yield levels. Shocks to GDP growth and inflation are about equally important; each of these shock series explains roughly 30% of the unconditional yield variance. Over shorter forecasting horizons, like one-quarter and four-quarter horizons, inflation shocks matter more for the short end of the yield curve, while GDP growth tends to be more important for longer yields.

Panel B documents that shocks to inflation are the main driving force behind the variance of yield spreads. Over any horizon, shocks to inflation explain more than 86% of the variance of yield spreads. Inflation shocks are even more important at longer horizons and for long maturity yield spreads. For example, movements in inflation account for 96% of the unconditional variance of the 5-year spread. These results are consistent with Mishkin (1992), who finds that inflation accounts for a large proportion of term spread changes. Ang and Bekaert (2004) also find that inflation accounts for a large amount of the movements of the term spreads in a term structure setting.

Expected Excess Holding Period Returns

We examine the variance decomposition of expected excess holding period returns in Panel C of Table 4. Time-varying expected excess returns are driven primarily by shocks to inflation and the latent factor. This is consistent with the variance decompositions to yield spreads in Panel B. The definition of excess holding period returns in equation (7) reveals that movements in excess returns are closely related to the movements in yield spreads. Therefore, we find that inflation is important for both yield spreads and excess returns.

Figure 2 shows the time-series of one-period expected excess holding period returns for the four-quarter and twenty-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 13% per quarter during the 1982 recession and drop below -4% during 1953 and 1978. In contrast, expected excess returns for the four-quarter bond lie in a more narrow range between -0.3% and 2.9% per quarter. Note that in 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 5.8% per quarter in September 1981 to 13.4% per quarter in March 1982.

To characterize how the macro variables affect expected excess returns, Table 5 reports regressions of expected excess returns onto macro factors and yield variables. Panel A reports the results from unrestricted OLS regressions, while Panel B reports the corresponding slope coefficients and R^2s computed from the model. Comparing the two panels reveals that the model is able to match the predictability patterns in the data well. Both the slope coefficients and the R^2 are of similar magnitudes across the two panels.

Interestingly, the point estimates of the loadings on both GDP and inflation are negative, so both high GDP and high inflation reduce the risk premia on long-term bonds. High GDP growth and high inflation rates are more likely to occur during the peaks of economic expansions, so bond risk premia are counter-cyclical. However, only the loading on inflation is significant. The remaining state variable in the model is latent, but we know from Table 2 that it is most highly correlated with the longest yield in our dataset. This is why we also included this yield in the regression. Its loading in Table 5 is positive and significant as well.

3.3 A Comparison of Taylor Rules

In this section, we provide a comparison of the benchmark, backward-looking, and forward-looking Taylor rules estimated by no-arbitrage techniques. We first discuss the estimates of each Taylor rule in turn, and then compare the monetary policy shocks computed from each different Taylor specification.

The Benchmark Taylor Rule

Panel A of Table 6 contrasts the OLS and model-implied estimates of the benchmark Taylor rule in equation (8). The OLS estimate of the output coefficient is small at 0.036, and is not significant. The model-implied coefficient is similar in magnitude at 0.091. In contrast, the OLS estimate of the inflation coefficient is 0.643 and strongly significant. The model-implied coefficient on π_t of 0.322 is much smaller. Hence, OLS over-estimates the response of the Fed on the short rate by approximately half compared to the model-implied estimate.

Although the estimation uses quarterly data, we obtain similar magnitudes for the δ_1 coefficients in equation (8) using annual GDP growth and inflation in the Taylor rule. Specifically, we re-write equation (8) to use GDP growth and inflation over the past year:

$$r_t = \delta_0 + \delta_{1,g}(g_t + g_{t-1} + g_{t-2} + g_{t-3}) + \delta_{1,\pi}(\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3}) + \varepsilon_t^{MP,T}, \qquad (27)$$

where $\varepsilon_t^{MP,T} \equiv \delta_{1,f^u} f_t^u$. In this formulation, bond yields now become affine functions of X_t, X_{t-1}, X_{t-2}, and X_{t-3}. Using annual GDP growth and inflation, the posterior mean of the coefficient on GDP growth (inflation) is 0.036 (0.334), with a posterior standard deviation of 0.023 (0.092). These values are almost identical to the estimates using the quarterly frequency data in Table 6.

An important question is whether the monetary policy rule coefficients in the short rate equation (2) are time-invariant. Several recent studies have emphasized that the linear coefficients δ_1 potentially vary over time (see, among others, Clarida, Galí and Gertler, 2000; Cogley and Sargent, 2001 and 2004; Boivin, 2004). However, other authors like Bernanke and Mihov (1998), Sims (1999 and 2001), Sims and Zha (2002), and Primiceri (2003) find either little or no evidence for time-varying policy rules, or negligible effect on the impulse responses of macro variables from time-varying policy rules. By estimating the model over the full sample, we follow Christiano, Eichenbaum, and Evans (1996), Cochrane (1998) and others and assume that the Taylor rule relationships are stable. We can address the potential time variation in these coefficients (and other parameters) by estimating our model over different subsamples, especially over the more recent post-1980's data corresponding to declining macroeconomic volatility (see Stock and Watson, 2003) and the post-Volcker era of leadership at the Federal Reserve.

If we estimate a benchmark Taylor rule using only data to the end of December 1982, the coefficients on GDP and inflation are very similar to the full sample estimates, at 0.067 and 0.352, respectively (compared to 0.091 and 0.322, respectively over the full sample). In the post-1982 period, estimating a benchmark Taylor rule with no-arbitrage restrictions over this period produces a slightly lower estimate of the weight on inflation, at 0.253, compared to the full sample estimate of 0.322. This is not surprising, because over the post-Volcker period, inflation is much lower, but it is surprising how close the inflation coefficient is across the two samples. In contrast, the post-1982 weight on GDP is 0.160, which is also close to the weight on GDP over the full sample, at 0.091.[1] Hence, the no-arbitrage estimates of the Taylor rule coefficients are fairly similar across the pre- and post-Volcker sample periods, and over the whole sample.

The Backward-Looking Taylor Rule

Panel B of Table 6 reports the estimation results for the backward-looking Taylor rule. Consistent with equation (12), the model coefficients on g_t and π_t are unchanged from the benchmark Taylor rule in Panel A at 0.091 and 0.332, respectively. The corresponding OLS estimates of the backward-looking Taylor rule coefficients on GDP and inflation are 0.074 and 0.182, respectively. Here, the model-implied rule predicts that the Fed reacts more to inflation than the OLS estimates suggest. The model also suggests that the Fed places a negative weight on past inflation; the coefficient on π_{t-1} is -0.217, but the sum of the coefficients on π_t and π_{t-1} is similar for both OLS and the model at approximately 0.10. 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 (0.925 in Table 3).

The Forward-Looking Taylor Rule

Finite Horizon, Without Discounting

In Panel C, Table 6, we list the estimates of the forward-looking Taylor rule coefficients $\delta_{1,g}$ and $\delta_{1,\pi}$ without discounting in equation (16) for various horizons k. For each k, we re-estimate the whole term structure model, but only report the forward-looking Taylor rule coefficients for comparison across k.

[1] To obtain convergence, we specify the post-1982 estimation to have a diagonal λ_1 matrix in equation (4).

For a one-quarter ahead Taylor rule, the coefficient on expected GDP growth (inflation) is 0.151 (0.509). These are larger than the contemporaneous responses for GDP growth and inflation over the past quarter in the benchmark Taylor rule, which are 0.091 and 0.322, respectively. The $k = 1$ coefficients for $\delta_{1,g}$ and $\delta_{1,\pi}$ also correspond roughly to a weighted average of the pre-Volcker and Volcker-Greenspan coefficients reported by Clarida, Galí, and Gertler (2000). For a one-year ($k = 4$) horizon, the short interest rate responds by just over half the level of a shock to GDP expectations, with $\delta_{1,g} = 0.502$, and almost one-to-one with inflation expectations, with $\delta_{1,\pi} = 0.998$. Thus, in the no-arbitrage forward-looking Taylor rule, the Fed responds very aggressively to changes in inflation expectations over a one-year horizon.

As k increases beyond one year, the coefficients on GDP and inflation expectations differ widely and the posterior standard deviations become very large. This is due to two reasons. First, as k becomes large, the conditional expectations approach their unconditional expectations, or $\mathrm{E}_t(g_{t+k,k}) \to \mathrm{E}(g_t)$ and $\mathrm{E}_t(\pi_{t+k,k}) \to \mathrm{E}(\pi_t)$. Econometrically, this makes $\delta_{1,g}$ and $\delta_{1,\pi}$ 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 (16) is the latent monetary policy shock:

$$r_t = \delta_0 + \delta_{1,g}\mathrm{E}(g_t) + \delta_{1,\pi}\mathrm{E}(\pi_t) + \varepsilon_t^{MP,F},$$

and it is impossible to differentiate the (scaled) effect of GDP or inflation expectations from δ_0. Hence, for large k, identification issues cause the coefficients $\delta_{1,g}$ and $\delta_{1,\pi}$ to be poorly estimated.

The second reason is that each estimation for different k is trying to capture the same contemporaneous relation between g_t, π_t, and r_t. Panel C also reports the estimates $\bar{\delta}_0$ and $\bar{\delta}_1 = (\bar{\delta}_{1,g}\ \bar{\delta}_{1,\pi}\ \bar{\delta}_{1,f^u})'$ of the short rate equation (15) implied by the forward-looking Taylor rules. These coefficients are very similar across horizons. In particular, the inflation coefficient $\bar{\delta}_{1,\pi}$ is almost unchanged at around 0.32 for all k. The coefficients on g_t and π_t are also very similar to the coefficients in the benchmark Taylor rule in Panel A. Both the benchmark estimation and the forward-looking estimation are trying to capture the same response of the short rate to macro factors, but the forward-looking Taylor rule transforms the contemporaneous response into the loadings on conditional expectations of future macro factors implied by the factor dynamics.

As k increases, even though $\delta_{1,g}$ and $\delta_{1,\pi}$ are theoretically identified, the state-dependence of r_t on g_t and π_t is diminished. This means that as k increases, the coefficients $\delta_{1,g}$ and $\delta_{1,\pi}$ must be large in order for there to be any contemporaneous effect of g_t and π_t on the short rate. Since the data exhibits a strong contemporaneous relation between g_t, π_t and r_t (see Table 3) and the forward-looking restrictions for large k shrink the contemporaneous effect of g_t and π_t on r_t, the estimation compensates by increasing the values of $\delta_{1,g}$ and $\delta_{1,\pi}$ to match the same short rate dynamics.

Infinite Horizon, With Discounting

We report the forward-looking Taylor rule where the Fed discounts future expected GDP growth and inflation in Panel D of Table 6. The coefficient on future discounted GDP growth (inflation) is 0.27 (0.62), which is between the $k = 1$ and $k = 4$ horizons in Panel C for the forward-looking Taylor rule without discounting. The discount rate $\beta = 0.93$, which implies an effective horizon of $1/(1 - 0.93)$ quarters, or 3.7 years. This estimate is much lower than the discount rates above 0.97 used in the literature (see Salemi, 1995; Rudebusch and Svenson, 1999; Favero and Rovelli, 2003), but still much higher than the estimate of 0.76 calibrated by Collins and Siklos (2004). The effective horizon of approximately four years is consistent with transcripts of FOMC meetings, which indicate that the Fed usually considers forecasts and policy scenarios of up to three to five years ahead.

3.4 Monetary Policy Shocks

Figure 3 plots model-implied monetary policy shock series based on the backward Taylor rule. We plot the OLS estimate in the top panel and the model-implied shocks, $\varepsilon_t^{MP,B}$, from equation (12) in the bottom panel. We compute $\varepsilon_t^{MP,B}$ using the posterior mean estimates of the latent factor through time. Figure 3 shows that the model-implied shocks are much smaller than the shocks estimated by OLS. In particular, during the early 1980's, the OLS shocks range from below -6% to above 4%. In contrast, the model-implied shocks lie between -2% and 2% during this period. This indicates that the Volcker-experience according to the no-arbitrage estimates was not as big a surprise as suggested by OLS. These results are consistent with our findings that the pre- and post-Volcker estimates of the Taylor rule using no-arbitrage identification techniques are similar.

Table 7 characterizes the model-implied monetary policy shocks in more detail and explicitly compares the OLS estimates with the model-implied estimates. We list model-implied estimates of the latent factor, f_t^u; the backward-looking Taylor rule shocks, $\varepsilon_t^{MP,B}$; and the forward-looking Taylor rule shocks (without discounting), $\varepsilon_t^{MP,F}$, over a $k = 4$ quarter horizon. We also include the Romer and Romer (2004) policy shocks that are computed using the Fed's internal forecasts of macro variables and intended changes of the federal funds rate. To obtain a quarterly series, we sum the monthly Romer-Romer series over the quarter. The last column of Table 7 reports statistics on the short rate.

The correlation of $\varepsilon_t^{MP,B}$ with the OLS estimates of the backward Taylor rule is quite high, at 72%, but the volatility of $\varepsilon_t^{MP,B}$, at 0.006 per quarter, is about half of the volatility of its OLS counterpart, at 0.009 per quarter. The OLS backward-looking Taylor residuals are also more negatively autocorrelated (-0.267) than the $\varepsilon_t^{MP,B}$ series, which has an autocorrelation of -0.195. This is very similar to the -0.183 autocorrelation of the Romer-Romer series. The OLS backward-looking Taylor rule shocks are more highly correlated with the short rate, at 32%, than $\varepsilon_t^{MP,B}$, which has a correlation of 15% with the short rate. Hence, using the short rate as an instrument to estimate monetary policy shocks results in dissimilar estimates to extracting a no-arbitrage estimate of the backward-looking Taylor rule shock using the whole yield curve. Whereas the correlation of $\varepsilon_t^{MP,B}$ with the Romer-Romer measure is 47% compared to the 71% correlation of the OLS backward-looking residual with the Romer-Romer shock, the range of $\varepsilon_t^{MP,B}$ overlaps much closer with the Romer-Romer shocks.

Table 7 also shows that the model-implied forward Taylor rule shocks are almost equivalent to the latent factor, f_t^u, having a correlation of over 99%. This is because the benchmark Taylor rule estimates on GDP and inflation being very similar to the implied short rate coefficients on the macro factors from the forward-looking Taylor rule (see Panels A and C of Table 6). In both the benchmark Taylor rule and the forward-looking Taylor rule, f_t^u represents the (scaled) monetary policy shock.

3.5 Impulse Responses

In order to gauge the effect of the various shocks on the yield curve and on macro variables, we compute impulse response functions. We obtain the posterior distribution of the impulse responses by computing the implied impulse response functions in each iteration of the Gibbs sampler. In the plots, we show the posterior mean of the impulse response functions. These responses are based on Cholesky decompositions that use the same ordering as the variance decompositions: (g_t, π_t, f_t^u).

Impulse Responses of Factors to Factor Shocks

Figure 4 plots impulse responses from our model (left-hand column) and the corresponding impulse responses from an unrestricted VAR (right-hand column). The VAR contains GDP growth, inflation, and the short rate. We compute impulse responses from this VAR by Cholesky decomposing the variance in the order (g_t, π_t, r_t) to take account of the contemporaneous covariances.

The top row of Figure 4 shows responses to a 1% shock to GDP. In the model on the LHS, the GDP shock has a short-lived effect on GDP itself, and its effect on the short rate is modest, at just approximately 10bp per quarter, and very persistent. By comparison, the response of the short rate in the unconstrained VAR is even smaller. In contrast, there is almost no effect of the GDP shock on inflation in either the model or an unconstrained VAR.

In the middle row of Figure 4, we plot the impulse responses corresponding to a 1% increase in inflation. In both the model and the VAR, the inflation shock has a persistent effect on inflation itself. The effect of the inflation shock in the model is to contemporaneously increase the short rate by around 30bp per quarter, which slowly dies out. In the VAR, the response of the short rate is weaker, around 20bp per quarter, and the short rate more rapidly mean-reverts to zero. The response of GDP to the inflation shock is more pronounced in the model than in the unrestricted VAR; GDP declines by more than -30bp (-20bp) per quarter in the model (VAR) at horizons of two and three quarters.

The last row shows the responses to a 1% shock to the latent factor for the model and a 1% short rate shock in the left- and right-hand columns, respectively. The response of the short rate to the 1% shock to f_t^u is not one for one because macro variables also enter the short rate equation in the model. As expected, GDP falls in response to a contractionary policy shock. The responses of inflation reveal a "price puzzle," because inflation increases after a shock to f_t^u or to the short rate. To fully eliminate the prize puzzle, we would need to add certain state variables to our system, such as commodity prices (see comments by Sims, 1992; Christiano, Eichenbaum and Evans, 1996, among others). This is an interesting avenue for future research; the goal of our paper is to illustrate how Taylor rules can be estimated using no-arbitrage techniques, and so we keep the system as low-dimensional as possible.

Impulse Responses of Yields

Figure 5 plots the responses of yields and yield spreads to GDP shocks, inflation shocks, and a short rate shock. A 1% inflation shock produces persistent effects on all yields. The initial response is highest for the short rate, at 32 bp per annum, while the initial response of the long, twenty-quarter yield is approximately 16 bp per annum. Hence, the term spread narrows from an unexpected inflation shock. Shocks to GDP also increase yields, but the effect from a GDP shock is much less. The initial response from a 1% GDP shock is almost the same across the yield curve, at approximately 10 bp.

The 1% shock to the short rate is constructed by shocking all of the state variables in proportion to their Cholesky decomposition so that the sum of the shock adds up to 1%. This allows us to trace the effect of a change in the short rate across the yield curve. As expected, the initial shock to a 1% increase in the short rate dies out gradually across the yield curve. At a five-year maturity, the response is approximately 82 bp per annum.

In Figure 6, we plot the impulse responses for GDP growth and inflation to a 1% shock in the latent factor f_t^u for different forward-looking Taylor rules (without discounting). The top panel shows that in comparison to the benchmark Taylor rule ($k = 0$) shown in discrete squares, GDP growth declines slightly more (less) for a one-quarter (four-quarter) ahead forward-looking Taylor rule. For example, from an initial 1% shock to f_t^u, GDP growth declines by 11 bp (8 bp) per annum after one year for the Taylor rule that incorporates expectations of macro variables over the one-quarter (four-quarter) horizon, compared to a decline of 10 bp for the basic Taylor rule after six quarters.

The impulse response for inflation in the bottom panel shows that a forward-looking Taylor rule does not ameliorate the price puzzle. For all horizons k, inflation increases after a 1% latent factor shock. In fact, the price puzzle is exacerbated for the one-quarter horizon, where inflation rises by 26bp per annum after one year, compared to 19bp per annum for the benchmark Taylor rule. For the forward-looking Taylor rule over $k = 20$ quarters, inflation increases to approximately 30bp per annum after five quarters. The structure of the forward-looking Taylor rule cannot remove the price puzzle because the underlying short rate dynamics driving the term structure are largely unchanged (see Panel C of Table 6) and the forward-looking Taylor rule only indirectly imposes restrictions on the companion form Φ in the VAR dynamics in equation (1), which governs the impulse responses of the macro variables.

4 Conclusion

We exploit information from the entire term structure to estimate Federal Reserve (Fed) policy rules. The framework accommodates original Taylor (1993) rules that describe the Fed as reacting to current values of GDP growth and inflation; backward-looking Taylor rules where the Fed reacts to current and lagged macro variables and lagged policy rates; and forward-looking Taylor rules where the Fed takes into account conditional expectations of future real activity and inflation. An advantage of the model is that all these types of Taylor rules are estimated jointly in a unified system that provides consistent expectations of future macro factors.

The 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 Vector Autoregression (VAR) dynamics for the macro and latent state variables and by specifying risk premia that are also linear combinations of the VAR state variables. The key identifying assumption is that the scaled latent factor can be defined as monetary policy shocks from a Taylor rule and is identified by the no-arbitrage over-identifying restrictions on the cross-section of yields.

We find that shocks to GDP growth and inflation account for over 60% of the time-variation of yields, while inflation shocks are mostly responsible for driving yield spreads. Macro factors also account for a significant fraction of time-varying expected excess returns of long-term bonds. We find that monetary policy shocks identified by no-arbitrage are much less volatile that Taylor rule residuals estimated by OLS and than the no-arbitrage restrictions also produce larger responses of interest rates from macro shocks than unrestricted VARs. Interesting extensions of this no-arbitrage methodology are to estimate general Taylor rules in more structural models of the economy or to expand the state space.

Appendix

A Forward-Looking Taylor Rules

In this appendix, we describe how to compute $\bar{\delta}_0$, $\bar{\delta}_1$ in equation (15) of a forward looking Taylor rule without discounting for a k-quarter horizon. From the dynamics of X_t in equation (1), the conditional expectation of k-quarter ahead GDP growth and inflation can be written as:

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

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

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

The bond price recursions for the standard affine model in equation (6) are modified by using $\bar{\delta}_0$ and $\bar{\delta}_1$ in place of δ_0 and δ_1, where

$$\begin{aligned} \bar{\delta}_0 &= \delta_0 + \frac{1}{k}[\delta_{1,g}e_1 \;\; \delta_{1,\pi}e_2]^\top\left(\sum_{i=1}^{k}\tilde{\Phi}_i\right)\mu, \\ \bar{\delta}_1 &= \frac{1}{k}[\delta_{1,g}e_1 \;\; \delta_{1,\pi}e_2]^\top\tilde{\Phi}_k\Phi + \delta_{1,f^u}e_3. \end{aligned} \tag{A-3}$$

Using this notation also enables us to write the short rate observation equation as:

$$\begin{aligned} \hat{r}_t &= \delta_0 + \delta_1^\top\bar{X}_t + \eta_t^{(1)} \\ &= \bar{\delta}_0 + \bar{\delta}_1^\top X_t + \eta_t^{(1)}. \end{aligned} \tag{A-4}$$

where $\bar{X}_t = [\mathrm{E}_t(g_{t+k,k}) \; \mathrm{E}_t(\pi_{t+k,k}) \; f_t^u]$ and $r_t = \delta_0 + \delta_1^\top\bar{X}_t$ is the forward-looking Taylor rule in equation (16) with $\delta_1 = (\delta_{1,g}\,\delta_{1,\pi}\,\delta_{1,f^u})^\top$.

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 as $k \to \infty$ is:

$$r_t = \delta_0 + [\delta_{1,g}e_1 \;\; \delta_{1,\pi}e_2]^\top(I-\Phi)^{-1}\mu + \delta_{1,f^u}f_t^u, \tag{A-5}$$

which implies that when k is large, the short rate effectively becomes a function only of f_t^u, and g_t and π_t can only indirectly affect the term structure through the feedback in the VAR equation (1). In the limiting case $k = \infty$, the coefficients $\delta_{1,g}$ and $\delta_{1,\pi}$ are unidentified because they act exactly like the constant term δ_0.

B Estimating the Model

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

The parameters of the model are $\Theta = (\mu,\, \Phi,\, \Sigma,\, \delta_0,\, \delta_1,\, \lambda_0,\, \lambda_1,\, \sigma_\eta)$, where σ_η denotes the vector of observation error volatilities $\{\sigma_\eta^{(n)}\}$. 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 (6).

Drawing the Latent Factor f^u

The factor dynamics (1), together with the yield equations (23), imply that the term structure model can be written as a state-space system. The state and observation equations for the system are linear in f_t^u, but also involve the macro variables g_t and π_t. To generate f^u, we use the Carter and Kohn (1994) forward-backward algorithm (see also Kim and Nelson, 1999). We first run the Kalman filter forward taking the macro variables $(g_t\; \pi_t)$ to be exogenous variables, and then sample f^u backwards. We use a Kalman filter algorithm that includes time-varying exogenous variables in the state equation following Harvey (1989). Since we specify the mean of f_t^u to be zero for identification, we set each generated draw of f^u to have a mean of zero.

Drawing μ and Φ

Since X_t follows a VAR in equation (1), it is tempting to just draw μ and Φ using a conjugate normal posterior distribution given the factors X_t. However, as Johannes and Polson (2003) comment, this procedure ignores the information contained in the yields $\hat{Y}$. Updating μ and Φ requires Metropolis algorithms as the conditional posteriors are not standard distributions.

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

$$P(\mu, \Phi \mid \Theta_-, X, \hat{Y}) \propto P(\hat{Y} \mid \Theta, X) P(X \mid \mu, \Phi, \Sigma) P(\mu, \Phi), \tag{B-6}$$

where Θ_- denotes the set of all parameters except μ and Φ. This posterior suggests an Independence Metropolis draw, where we first draw a proposal for the $(m+1)$th value of μ and Φ (μ^{m+1} and Φ^{m+1}, respectively) from the proposal density:

$$q(\mu, \Phi) \propto P(X \mid \mu, \Phi, \Sigma) P(\mu, \Phi),$$

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

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

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-7), α is just the ratio of the likelihoods of the new draw of μ and Φ relative to the old draw. For the values of μ and Φ corresponding to the equations for g_t and π_t, we draw μ and Φ separately for each equation in the VAR system (1), but we perform the accept/reject step for each individual parameter.

For the values of μ and Φ corresponding to the f_t^u equation, we modify the algorithm. For ϕ_{33}, the element in the third row and third column of Φ, the Independence Metropolis algorithm yields very low accept rates because the likelihood is very sensitive to the value of ϕ_{33}. To increase the acceptance rate, we use a Random Walk Metropolis draw for ϕ_{33}:

$$\phi_{33}^{m+1} = \phi_{33}^m + \zeta_{\phi_{33}} v \tag{B-8}$$

where $v \sim N(0,1)$ and $\zeta_{\phi_{33}}$ is the scaling factor used to adjust the acceptance rate. The accept probability for ϕ_{33} is given by:

$$\begin{aligned} \alpha &= \min\left\{ \frac{P(\phi_{33}^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\phi_{33}^{m} \mid \Theta_-, X, \hat{Y})} \frac{q(\phi_{33}^m \mid \phi_{33}^{m+1})}{q(\phi_{33}^{m+1} \mid \phi_{33}^{m})}, 1 \right\} \\ &= \min\left\{ \frac{P(\phi_{33}^{m+1} \mid \Theta_-, X, \hat{Y})}{P(\phi_{33}^{m} \mid \Theta_-, X, \hat{Y})}, 1 \right\}, \end{aligned} \tag{B-9}$$

where the posterior $P(\phi_{33}|\Theta_-, X, \hat{Y})$ is given by:

$$P(\phi_{33}|\Theta_-, X, \hat{Y}) \propto P(\hat{Y}|\phi_{33}, \Theta_-, X) P(X|\mu, \phi_{33}, \Sigma).$$

Thus, in the case of the draw for ϕ_{33}, α is the posterior ratio of new and old draws of ϕ_{33}.

We also impose the restriction that f_t^u is mean zero for identification. We draw the parameters ϕ_{31} and ϕ_{32} (which are the elements in Φ in the third row and first column, and the third row and second column, respectively) separately from μ_3, the third element of the vector μ. We draw ϕ_{31} and ϕ_{32} jointly, but we set μ_3 to satisfy $e_3^\top(I - \Phi)^{-1}\mu = 0$ to ensure that the factor f_t^u has mean zero. Hence μ_3 is simply a function of the other parameters in the factor VAR in equation (1).

Drawing Σ

We draw Σ from the proposal density $q(\Sigma) = P(X \mid \mu, \Phi, \Sigma)P(\Sigma)$, which is an Inverse Wishart (IW) distribution if we specify the prior $P(\Sigma)$ to be IW, so that $q(\Sigma)$ is an IW natural conjugate. Similar to the independence metropolis case of μ and Φ, the accept/reject probability for the draws of Σ from the proposal density is the likelihood ratio of the new draw relative to old draw of Σ (see equation (B-7)).

Drawing δ_1

To draw δ_1, we exploit the observation short rate equation:

$$\hat{r}_t = \delta_0 + \delta_1' X_t + \eta_t^{(1)},$$

which is simply a regression of $\hat{r} = \hat{y}_t^{(1)}$ (data) on factors X, with observation error $\eta_t^{(1)}$. We draw δ_1 as a conjugate normal. Note that in drawing δ_1, since $\hat{r}_t$ is observable, and X_t is treated as observable for the purpose of drawing δ_1, we base the Taylor rule estimation only on short rate data. We do not draw the constant δ_0, but instead set δ_0 to match the sample mean of the short rate.

To draw δ_1 in the forward-looking Taylor rule system entails only a very simple modification. Since the observation equation for the short rate is a standard regression, we can draw δ_1 using a conjugate normal. Again, we set δ_0 to match the sample mean of the short rate.

Drawing λ_0 and λ_1

We draw λ_0 and λ_1 with a Random Walk Metropolis algorithm. We assume a flat prior, and for each parameter separately in $\lambda = (\lambda_0,\, \lambda_1)$, we draw the $(m+1)$th price of risk as:

$$\lambda^{m+1} = \lambda^m + \zeta_\lambda v \qquad \text{(B-10)}$$

where $v \sim N(0, 1)$ and ζ_λ is the scaling factor used to adjust the accept rate. Equation (B-10) represents the draw for each individual parameter in λ. The acceptance probability α is given by:

$$\begin{aligned}\alpha &= \min\left\{\frac{P(\lambda^{m+1} \mid \Theta_{-}, X, \hat{Y})}{P(\lambda^{m} \mid \Theta_{-}, X, y)}\frac{q(\lambda^{m} \mid \lambda^{m+1})}{q(\lambda^{m+1} \mid \lambda^{m})}, 1\right\} \\ &= \min\left\{\frac{P(\lambda^{m+1} \mid \Theta_{-}, X, \hat{Y})}{P(\lambda^{m} \mid \Theta_{-}, X, \hat{Y})}, 1\right\} \\ &= \min\left\{\frac{P(\hat{Y} \mid \lambda^{m+1}, \Theta_{-}, X)}{P(\hat{Y} \mid \lambda^{m}, \Theta_{-}, X)}, 1\right\}, \end{aligned} \qquad \text{(B-11)}$$

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

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 (23). We draw the observation variance $(\sigma_\eta^{(n)})^2$ separately from each yield using an Inverse Gamma posterior distribution.

Drawing β

For the case of the forward-looking Taylor rule with an infinite horizon and 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 $\beta^{m+}$is given by:

$$\begin{aligned}\alpha &= \min\left\{\frac{P(\beta^{m+1} \mid \Theta_{-}, X, \hat{Y})}{P(\beta^{m} \mid \Theta_{-}, X, y)}\frac{q(\beta^{m})}{q(\beta^{m+1})}, 1\right\} \\ &= \min\left\{\frac{P(\hat{Y} \mid \beta^{m+1}, \Theta_{-}, X)}{P(\hat{Y} \mid \beta^{m}, \Theta_{-}, X)}\frac{q(\beta^{m})}{q(\beta^{m+1})}, 1\right\}, \end{aligned} \qquad \text{(B-12)}$$

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
ϕ_{33}	0.001	0.458	$\lambda_{1,21}$	0.100	0.716
$\lambda_{0,1}$	0.010	0.484	$\lambda_{1,22}$	0.100	0.669
$\lambda_{0,2}$	0.010	0.465	$\lambda_{1,23}$	0.100	0.777
$\lambda_{0,3}$	0.010	0.409	$\lambda_{1,31}$	0.100	0.684
$\lambda_{1,11}$	0.100	0.720	$\lambda_{1,32}$	0.100	0.636
$\lambda_{1,12}$	0.100	0.675	$\lambda_{1,33}$	0.100	0.742
$\lambda_{1,13}$	0.100	0.786			

where $\lambda_0 = (\lambda_{0,1}\, \lambda_{0,2}\, \lambda_{0,3})^\top$ and $\lambda_{1,ij}$ denotes the element of λ_1 in the ith row and jth column. The acceptance rate for the β parameter when estimating the forward-looking Taylor rule with discounting is 0.606.

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 $\pm$0.005 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 only exception is the estimates in the companion form Φ corresponding to the latent factor equation f_t^u, which require a Raftery-Lewis minimum burn-in and number of iterations of the same order of magnitude that we perform in the estimation. This is not surprising, because f_t^u is a very persistent series and mean-reversion parameters of persistent series are notoriously difficult to estimate in small samples.

The third, and probably most compelling check of the estimation method is that the MCMC-Gibbs sampler works very well on simulated data. We perform Monte Carlo simulations, similar to the experiments performed by Eraker, Johannes and Polson (2003).

We take the posterior means of the parameters in Table 1 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 that we are very confident about the convergence of our algorithm.

References

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

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

[3] Ang, A., M. Piazzesi, and M. Wei, 2004, "What does the Yield Curve Tell us about GDP Growth?" forthcoming *Journal of Econometrics*.

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

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

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

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

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

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

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

[11] 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.

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

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

[14] 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.

[15] 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.

[16] 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.

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

[18] 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.

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

[20] Cochrane, J. H., and M. Piazzesi, 2004, "Bond Risk Premia," forthcoming *American Economic Review*.

[21] Cogley, T., 2003, "An Exploration of Evolving Term-Structure Relations," working paper, Arizona State University.

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

[23] Cogley, T., and T. J. Sargent, 2004, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," working paper, UC Davis.

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

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

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

[27] Dai, Q., and T. Philippon, 2004, "Government Deficits and Interest Rates: A No-Arbitrage Structural VAR Approach," working paper, NYU.

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

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

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

[31] Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba, 2004, "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach," working paper, Univeristy of Pennsylvania.

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

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

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

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

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

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

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

[39] 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.

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

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

[42] Hördahl, P. O. Tristani, and D. Vestin, 2003, "A Joint Econometric Model of Macroeconomic and Term Structure Dynamics," working paper, European Central Bank.

[43] Johannes, M., and N. Polson, 2003, "MCMC Methods for Financial Econometrics," working paper, Columbia University.

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

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

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

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

[48] Piazzesi, M., 2004, "Bond Yields and the Federal Reserve," forthcoming *Journal of Political Economy*.

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

[50] 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, 763-773.

[51] Romer, C. D., and D. H. Romer, 2004, "A New Measure of Monetary Policy Shocks: Derivation and Implication," working paper, UC Berkeley.

[52] 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.

[53] Rudebusch, G., and T. Wu, 2003, “A Macro-Finance Model of the Term Strucutre, Monetary Policy, and the Economy,” working paper, Federal Reserve Bank of San Francisco.

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

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

[56] Sims, C. A., 1999, “Drifts and Breaks in Monetary Policy,” working paper, Yale University.

[57] Sims, C. A., 2001, “Comment on Sargent and Cogley’s ‘Evolving Post World War II U.S. Inflation Dynamics,’ ” *NBER Macroeconomics Annual*, 16, 373-379.

[58] Sims, C. A., and T. Zha, 2002, “Macroeconomic Switching,” working paper, Princeton University.

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

Table 1: PARAMETER ESTIMATES

Factor Dynamics

		Companion Form Φ			$\Sigma\Sigma' \times 10000$		
	μ	g	π	f^u	g	π	f^u
g	0.008	0.271	-0.299	-0.024	0.874	-0.008	0.000
	(0.002)	(0.064)	(0.010)	(0.069)	(0.112)	(0.043)	–
π	0.002	0.052	0.711	0.090	-0.008	0.270	0.000
	(0.001)	(0.038)	(0.060)	(0.044)	(0.043)	(0.032)	–
f^u	-0.002	0.133	0.162	0.931	0.000	0.000	0.050
	(0.001)	(0.056)	(0.066)	(0.032)	–	–	–

Short Rate Equation

	δ_1		
δ_0	g	π	f^u
0.009	0.091	0.322	0.499
(0.002)	(0.064)	(0.143)	(0.228)

Risk Premia Parameters

		λ_1		
	λ_0	g	π	f^u
g	-0.986	10.886	5.128	-18.749
	(0.487)	(7.426)	(6.292)	(5.300)
π	-0.402	-7.033	-10.055	-14.952
	(0.180)	(7.660)	(5.377)	(5.602)
f^u	0.583	4.378	23.867	-0.310
	(0.272)	(4.969)	(8.637)	(8.565)

Observation Error Standard Deviation

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
$\sigma_\eta^{(n)}$	0.188	0.130	0.066	0.032	0.046	0.072
	(0.011)	(0.013)	(0.020)	(0.026)	(0.036)	(0.064)

The table lists parameter values for the model in equations (1)-(4) and observation error standard deviations in equation (23) for yields of maturity n quarters. We use 50,000 simulations after a burn-in sample of 10,000. We report the posterior mean and posterior standard deviation (in parentheses) of each parameter. The sample period is June 1952 to December 2003 and the data frequency is quarterly.

Table 2: Latent Factor f^u Correlations

Correlations with Macro Factors

	g	π
Data	-0.173	0.489
Implied	-0.186	0.530

Correlations with Yields

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
Yield Levels $y_t^{(n)}$						
Data	0.912	0.941	0.965	0.975	0.977	0.977
Implied	0.950	0.966	0.975	0.979	0.981	0.982
Yield Spreads $y_t^{(n)} - y_t^{(1)}$						
Data		0.196	0.188	0.110	0.087	0.044
Implied		0.284	0.245	0.190	0.129	0.066

The table reports the model-implied correlations between the latent factor f^u and macro variables, yield levels, and yield spreads. The correlations are computed at the posterior mean of the model parameters.

Table 3: SUMMARY STATISTICS

PANEL A: MOMENTS OF MACRO FACTORS

	Means %		Standard Deviations %		Autocorrelations	
	Data	Model	Data	Model	Data	Model
g	0.803	0.666	0.964	1.061	0.340	0.361
	(0.088)	(0.301)	(0.067)	(0.132)	(0.068)	(0.097)
π	0.950	0.942	0.792	0.900	0.762	0.778
	(0.110)	(0.396)	(0.097)	(0.257)	(0.058)	(0.072)

PANEL B: MOMENTS OF YIELDS

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
Means						
Data	1.334	1.438	1.488	1.528	1.558	1.576
	(0.107)	(0.108)	(0.107)	(0.105)	(0.105)	(0.104)
Model	1.334	1.424	1.492	1.531	1.555	1.570
	–	(0.023)	(0.032)	(0.035)	(0.048)	(0.076)
Standard Deviations						
Data	0.720	0.721	0.713	0.697	0.693	0.684
	(0.092)	(0.083)	(0.082)	(0.080)	(0.080)	(0.076)
Model	0.697	0.702	0.703	0.699	0.692	0.685
	(0.048)	(0.048)	(0.047)	(0.046)	(0.044)	(0.045)
Autocorrelations						
Data	0.925	0.932	0.943	0.949	0.955	0.959
	(0.032)	(0.031)	(0.029)	(0.028)	(0.028)	(0.026)
Model	0.955	0.961	0.963	0.964	0.964	0.965
	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)

PANEL C: SHORT RATE CORRELATIONS

	g	π	f^u
Data	-0.090	0.695	–
	(0.100)	(0.063)	
r	-0.136	0.740	0.947
	(0.182)	(0.134)	(0.035)

Panel A lists moments of GDP and inflation in data and implied by the model. For the model, we construct the posterior distribution of unconditional moments by computing the unconditional moments implied from the parameters in each iteration of the Gibbs sampler. Panel B reports data and model unconditional moments of n-quarter maturity yields. We compute the posterior distribution of the model-implied yields $\hat{y}_t^{(n)}$ in equation (5) using the generated latent factors $\hat{f}^u$ in each iteration. In Panel C, we report correlations of the short rate with various factors. For the model, we compute the posterior distribution of the correlations of the model-implied short rate r in equation (2). In all the panels, the data standard errors (in parentheses) are computed using GMM and all moments are computed at a quarterly frequency. For the model, we report posterior means and standard deviations (in parentheses) of each moment. The sample period is June 1952 to December 2003 and the data frequency is quarterly.

Table 4: VARIANCE DECOMPOSITIONS

PANEL A: YIELD LEVELS

	$n=1$	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
One-Quarter Ahead						
g	14.5	21.7	25.0	26.3	27.0	27.4
π	59.2	42.5	31.2	25.8	22.9	21.1
f^u	26.3	35.9	43.8	47.9	50.1	51.5
Four-Quarters Ahead						
g	23.5	27.5	29.5	30.4	30.8	31.1
π	46.3	36.8	30.8	28.0	26.5	25.6
f^u	30.2	35.7	39.7	41.6	42.7	43.3
Unconditional						
g	29.9	30.8	31.2	31.4	31.5	31.5
π	33.7	31.7	30.4	29.9	29.6	29.4
f^u	36.4	37.6	38.4	38.8	39.0	39.1

PANEL B: YIELD SPREADS

	$n=4$	$n=8$	$n=12$	$n=16$	$n=20$
One-Quarter Ahead					
g	8.3	3.8	2.5	1.8	1.4
π	86.0	91.0	93.0	94.3	95.4
f^u	5.7	5.2	4.6	3.9	3.3
Four-Quarters Ahead					
g	4.9	2.6	1.6	1.1	0.7
π	87.7	90.9	92.8	94.3	95.6
f^u	7.3	6.6	5.6	4.6	3.7
Unconditional					
g	3.9	2.2	1.4	0.9	0.5
π	88.1	90.7	92.6	94.2	95.6
f^u	8.0	7.1	6.0	4.9	3.9

PANEL C: EXPECTED EXCESS RETURNS

	$n = 4$	$n = 8$	$n = 12$	$n = 16$	$n = 20$
One-Quarter Ahead					
g	11.9	20.2	23.2	24.6	25.4
π	49.1	56.9	59.1	60.1	60.5
f^u	39.0	22.9	17.7	15.4	14.1
Four-Quarters Ahead					
g	1.5	0.3	1.6	2.7	3.4
π	21.8	42.8	52.1	56.5	58.9
f^u	76.7	56.9	46.3	40.8	37.7
Unconditional					
g	8.7	1.9	0.4	0.0	0.0
π	11.2	30.6	41.7	47.5	50.9
f^u	80.2	67.5	58.0	52.4	49.1

The table reports variance decompositions (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. All maturities are in quarters. To compute variance decompositions for yield levels and yield spreads, we ignore observation error. The variance decompositions are computed using the posterior mean of the parameters listed in Table 1.

Table 5: PREDICTABILITY REGRESSIONS

PANEL A: DATA

	const	g_t	π_t	$y_t^{(20)}$	R^2
$n = 4$	-0.001	-0.072	-0.078	0.223	0.036
	(0.001)	(0.064)	(0.090)	(0.096)	
$n = 12$	-0.004	-0.193	-0.461	0.752	0.040
	(0.004)	(0.184)	(0.240)	(0.296)	
$n = 20$	-0.007	-0.237	-0.719	1.129	0.039
	(0.006)	(0.266)	(0.366)	(0.450)	

PANEL B: MODEL

	const	g_t	π_t	$y_t^{(20)}$	R^2
$n = 4$	0.001	-0.037	-0.051	0.163	0.034
	(0.002)	(0.049)	(0.087)	(0.065)	
$n = 12$	0.005	-0.171	-0.360	0.622	0.036
	(0.007)	(0.175)	(0.327)	(0.217)	
$n = 20$	0.009	-0.306	-0.678	1.040	0.037
	(0.013)	(0.317)	(0.598)	(0.356)	

We regress 1-quarter excess holding period returns for an n-period bond, $rx_{t+1}^{(n)}$ onto GDP, inflation, and a bond yield. Panel A reports the data regression results. The standard errors for the OLS estimates (in parentheses) are computed using robust standard errors. Panel B reports the model-implied coefficients and R^2. For the model, we construct the posterior distribution of those 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. In each panel, the data frequency is quarterly and the sample period is June 1952 to December 2003.

Table 6: TAYLOR RULES

PANEL A: BENCHMARK TAYLOR RULE

	const	g_t	π_t	R^2
OLS	0.007	0.036	0.643	0.489
	(0.001)	(0.071)	(0.080)	
Model	0.009	0.091	0.322	0.137
	(0.002)	(0.064)	(0.143)	

PANEL B: BACKWARD-LOOKING TAYLOR RULE

	const	g_t	π_t	g_{t-1}	π_{t-1}	r_{t-1}	R^2
OLS	0.000	0.074	0.182	-0.005	-0.077	0.879	0.895
	(0.000)	(0.027)	(0.046)	(0.029)	(0.041)	(0.035)	
Model	-0.000	0.091	0.322	-0.019	-0.218	0.931	0.976
	(0.001)	(0.064)	(0.143)	(0.021)	(0.092)	(0.032)	

PANEL C: FORWARD-LOOKING TAYLOR RULE, FINITE HORIZON, WITHOUT DISCOUNTING

	Forward-Looking Taylor Rule Coefficients				Implied Short Rate Coefficients		
	const	$\mathrm{E}_t(g_{t+k,k})$	$\mathrm{E}_t(\pi_{t+k,k})$	R^2	$\bar{\delta}_{1,g}$	$\bar{\delta}_{1,\pi}$	$\bar{\delta}_{1,f^u}$
$k=1$	0.007	0.151	0.509	0.117	0.074	0.331	0.587
	(0.003)	(0.221)	(0.154)		(0.077)	(0.122)	(0.233)
$k=4$	0.001	0.502	0.998	0.139	0.099	0.331	0.492
	(0.006)	(0.417)	(0.286)		(0.052)	(0.081)	(0.146)
$k=8$	-0.008	0.751	1.277	0.132	0.072	0.306	0.514
	(0.096)	(5.096)	(1.658)		(0.083)	(0.146)	(0.284)
$k=20$	-0.010	0.003	2.066	0.142	0.016	0.311	0.600
	(0.172)	(6.878)	(4.854)		(0.089)	(0.127)	(0.224)

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

	Forward-Looking Taylor Rule Coefficients				Implied Short Rate Coefficients		
	const	$\hat{g}_t$	$\hat{\pi}_t$	β	$\hat{\delta}_{1,g}$	$\hat{\delta}_{1,\pi}$	$\hat{\delta}_{1,f^u}$
$k=\infty$	0.057	0.267	0.615	0.933	0.015	0.098	0.429
	(0.061)	(0.078)	(0.181)	(0.036)	(0.034)	(0.037)	(0.187)

NOTE TO TABLE 6

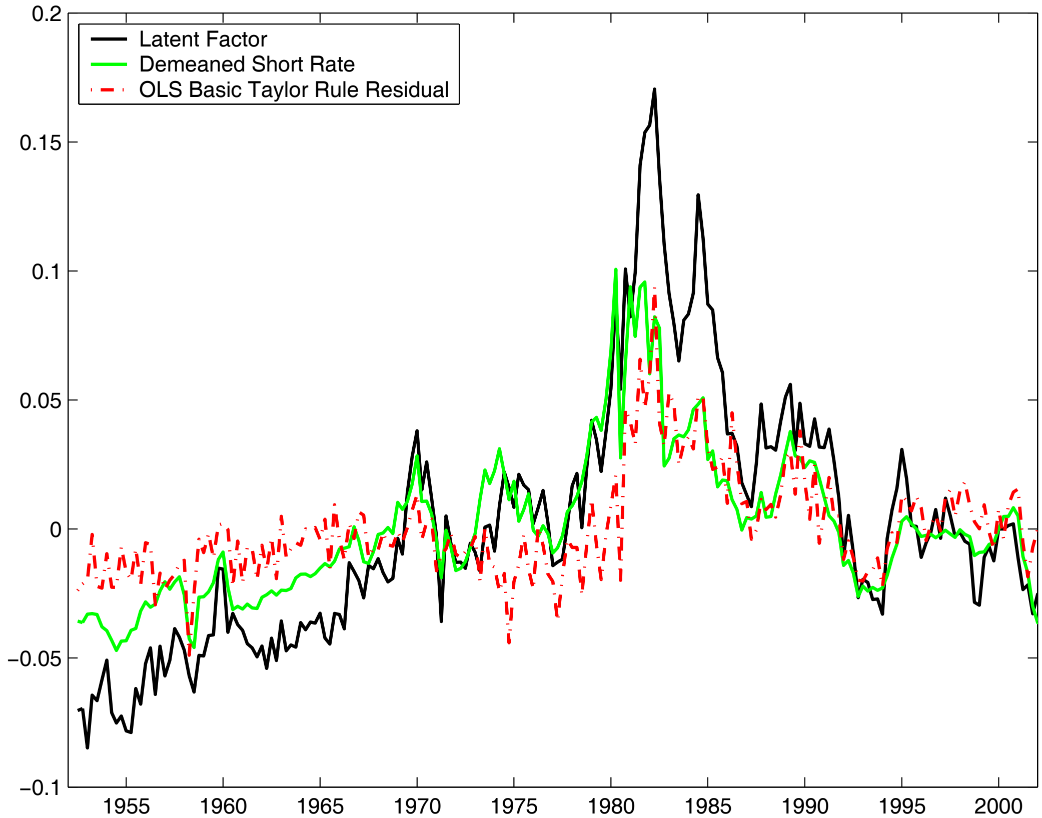
The table reports the OLS and model-implied estimates of a benchmark Taylor (1993) rule (8) in Panel A; the backward-looking Taylor rule (9) in Panel B; the finite-horizon, forward-looking Taylor rule without discounting in equation (16) in Panel C; and the infinite-horizon, forward-looking Taylor rule with discounting in equation (18) in Panel D. Panels C and D also report the implied short rate coefficients corresponding to the forward-looking Taylor rules without discounting in equation (15) for each horizon k and equation (21) for the forward-looking Taylor rule with discounting. For the model, 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 sample period is June 1952 to December 2003.

Table 7: MONETARY POLICY SHOCKS

	OLS Estimates		Model-Implied Estimates				
	Benchmark Taylor Rule	Backward Taylor Rule	Latent Factor	Backward Taylor Rule	Forward Taylor Rule	Romer-Romer	Short Rate
Correlations							
OLS Benchmark Taylor Rule	1.000	0.454	0.807	0.340	0.814	0.276	0.713
OLS Backward Taylor Rule		1.000	0.250	0.720	0.254	0.715	0.325
Latent Factor f^u			1.000	0.210	0.999	0.155	0.916
Model Backward Taylor Rule				1.000	0.223	0.465	0.147
Model Forward Taylor Rule					1.000	0.151	0.907
Romer-Romer Shock						1.000	0.290
Short Rate r							1.000
Autocorrelations	0.743	-0.267	0.948	-0.195	0.942	-0.183	0.924
Standard Deviation	0.021	0.009	0.049	0.006	0.022	0.007	0.029
Minimum	-0.049	-0.058	-0.085	-0.022	-0.040	-0.040	0.006
Maximum	0.094	0.043	0.171	0.024	0.079	0.025	0.154

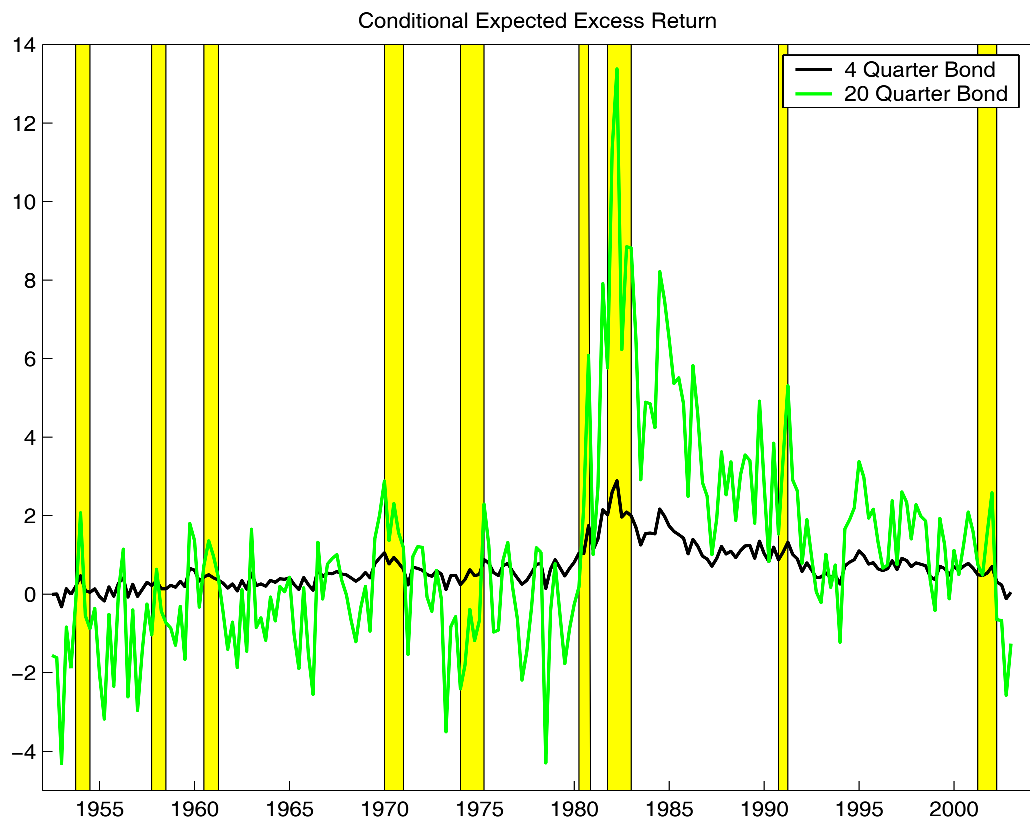
The table reports summary statistics of OLS estimates of residuals from the benchmark Taylor (1993) rule in equation (8); OLS estimates of residuals from the backward-looking Taylor rule in equation (9); posterior mean estimates of the latent factor f_t^u; the model-implied backward-looking Taylor rule residuals $\varepsilon^{MP,B}$ in equation (12); the model-implied forward-looking Taylor rule shocks (without discounting) from equation (16) for a $k = 4$ annual horizon; the Romer and Romer (2004) measure of policy shocks shocks (taken from Table 2 of Romer and Romer (2004)) converted into quarterly sequence by summing the monthly Romer-Romer shocks; and the one-quarter short rate. All monetary policy shocks and the short rate are annualized. The model-implied shocks are computed using the posterior mean of the latent factors. The sample period is June 1952 to December 2003 and the data frequency is quarterly.

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



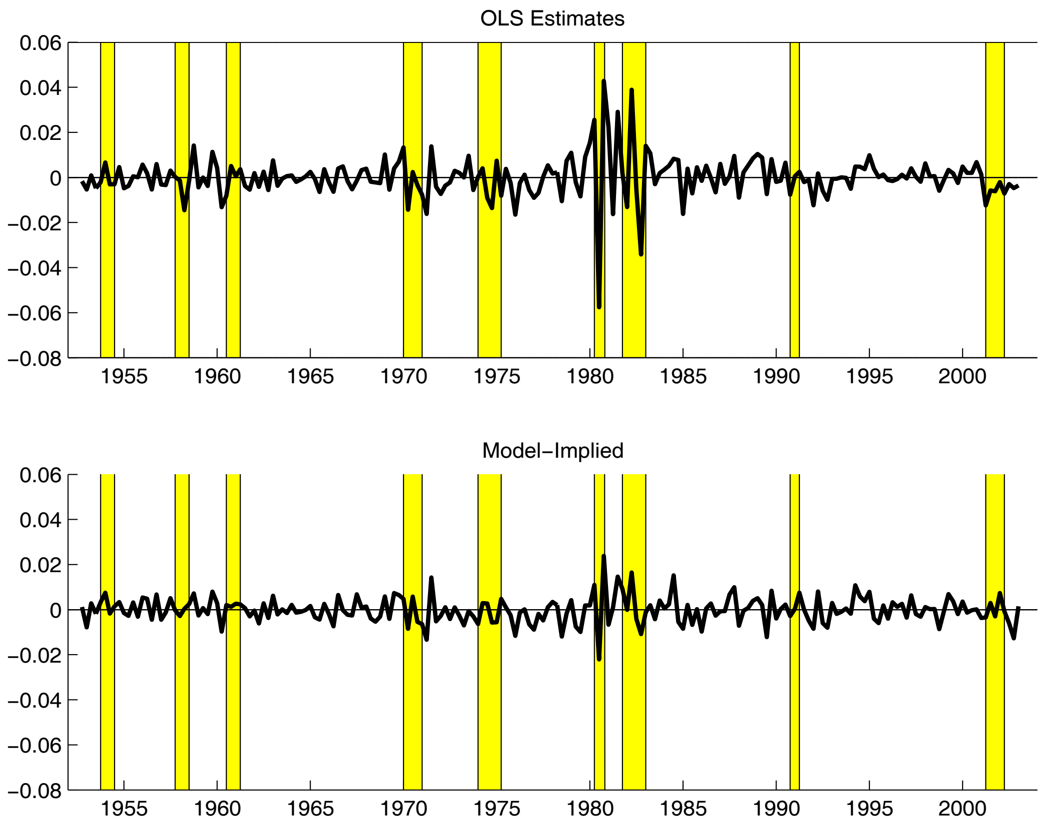
We plot the posterior mean of the latent factor f_t^u, the demeaned short rate from data, and the OLS estimate of the basic Taylor Rule, which is computed by running OLS on equation (8). The latent factor and short rate are annualized.

Figure 2: Expected Excess Bond Returns



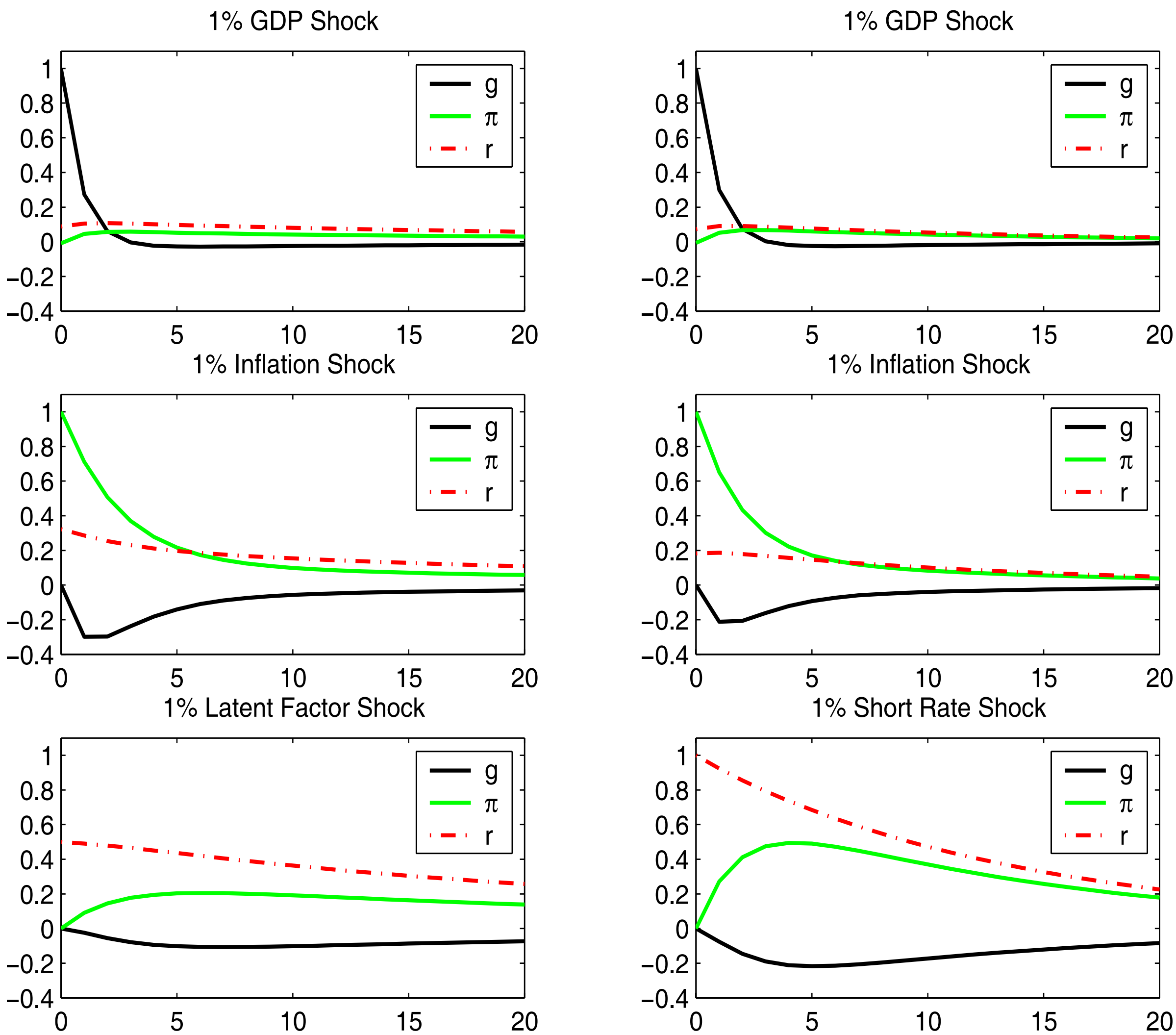
We plot the conditional expected excess holding period return $\mathrm{E}_t[rx_{t+1}^{(n)}]$ of a four-quarter and twenty-quarter bond implied by the posterior mean of the latent factors through time.

Figure 3: MONETARY POLICY SHOCKS



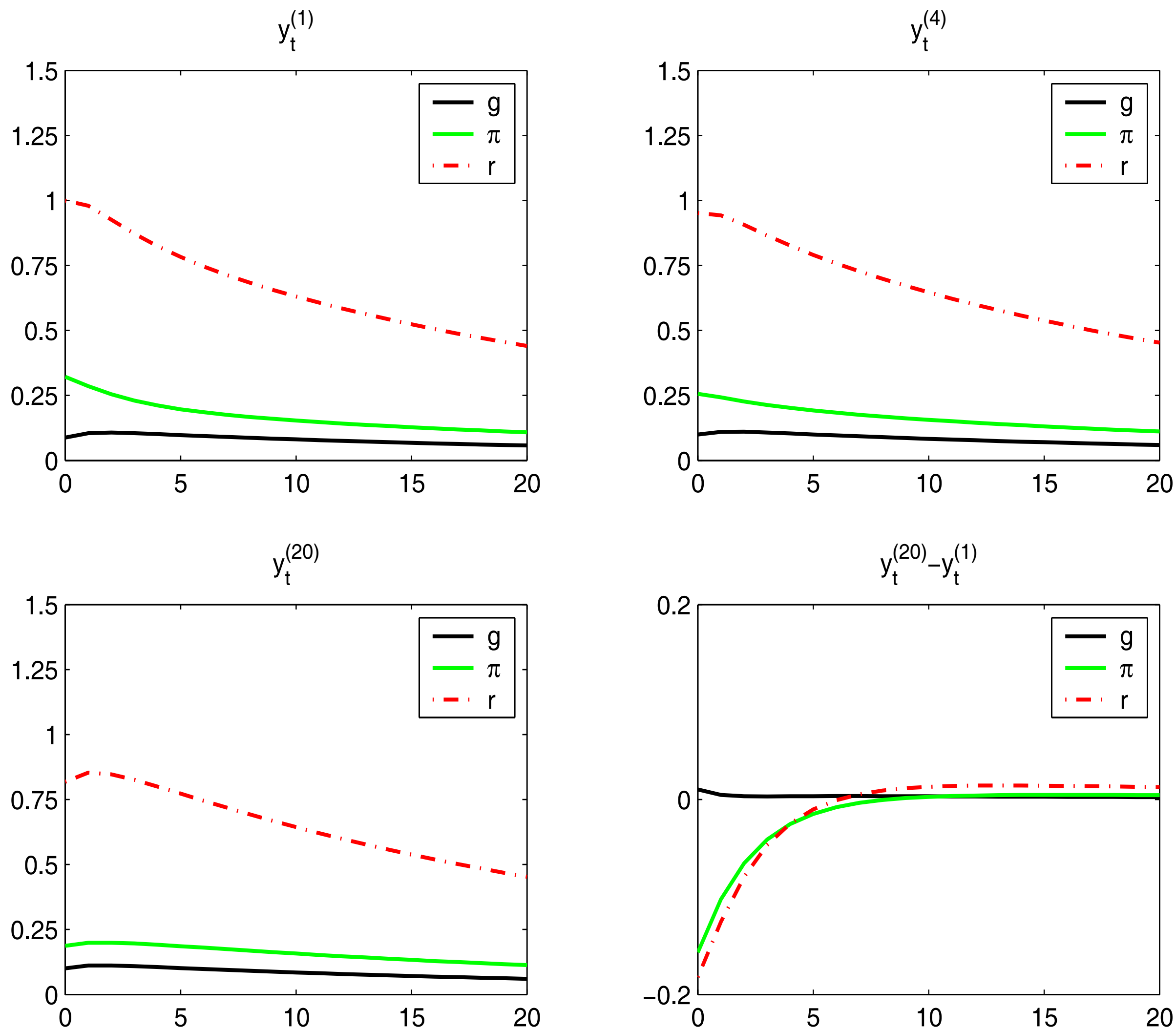
In the top panel, we plot the OLS estimates of the residuals of the backwards-looking Taylor rule (equation (9)). The bottom panel plots the corresponding model-implied monetary policy shocks, which are the posterior mean estimates of $\varepsilon_t^{MP,B} = \delta_{12} u_t^2$ from equation (12). In both the top and bottom panels, we plot annualized monetary policy shocks. NBER recessions are shown as shaded bars.

Figure 4: IMPULSE RESPONSES OF FACTORS



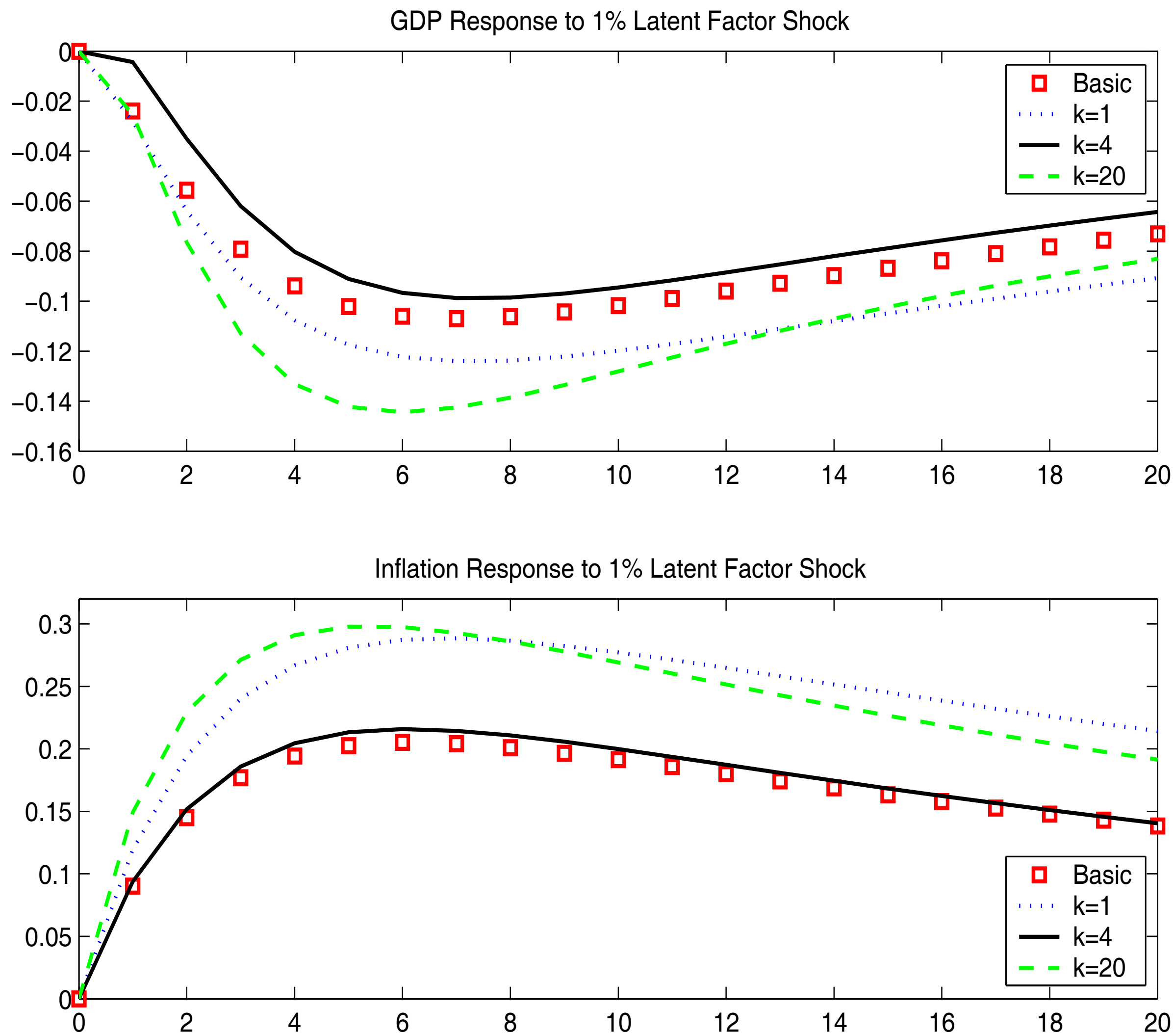
The left (right) column shows the impulse response for the model (an unrestricted VAR). The unrestricted VAR contains GDP growth, inflation, and the short rate. All y-axis responses are in percentages, and we show quarters on the x-axis. All initial shocks are 1% and are computed using a Cholesky decomposition that orders the variables (g_t, π_t, f_t^u) in the model and (g_t, π_t, r_t) in the unrestricted VAR.

Figure 5: IMPULSE RESPONSES OF YIELDS



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

Figure 6: IMPULSE RESPONSES FROM FORWARD-LOOKING TAYLOR RULES



We plot impulse responses to GDP (top panel) and inflation (bottom panel) from benchmark Taylor rules and from forward-looking Taylor rules (finite horizon, without discounting) to a 1% shock in the latent factor, f_t^u. Yields on the y-axis are annualized and we show quarters on the x-axis. The impulse responses are computed using a Cholesky decomposition that orders the variables (g_t, π_t, f_t^u).