

Term structure estimation without using latent factors

*Gregory R. Duffee**

Haas School of Business
University of California – Berkeley

This Draft: November 9, 2004

ABSTRACT

The term structure is modeled as a function of observed and unobserved (latent) factors. I use restrictions implied by no-arbitrage to extract information about the term structure from the observed factors, without specifying or estimating any of the parameters associated with the latent factors. In particular, both the number of latent factors and their correlations with future observed factors may be unknown. Estimation is straightforward because it reduces to fitting the moment conditions of a set of regressions, where no-arbitrage imposes cross-equation restrictions on the coefficients. I apply the methodology to the joint dynamics of inflation and the term structure. Outside of the disinflationary period of 1979 through 1983, short-term interest rates move one-for-one with expected inflation. In addition, bond risk premia are insensitive to inflation. Therefore the effect of changes in inflation on long-maturity bond yields is driven by changes in expected future short-term rates.

*Voice 510-642-1435, fax 510-643-1420, email duffee@haas.berkeley.edu. Address correspondence to 545 Student Services Building #1900, Berkeley, CA 94720. I thank Qiang Dai, George Pennacchi, Richard Stanton, and seminar participants at many universities for helpful comments. The most recent version of this paper is at <http://faculty.haas.berkeley.edu/duffee/>.

1 Introduction

Beginning with Vasicek (1977) and Cox, Ingersoll, and Ross (1985), researchers have built increasingly sophisticated no-arbitrage models of the term structure. These models specify the evolution of state variables under both the physical and equivalent martingale measures, and thus represent complete descriptions of the dynamic behavior of yields at all maturities. Much of this work focused on latent variable settings, where the evolution of yields is described in terms of yields themselves. This rather introspective view was broadened by the important work of Piazzesi (2003) and Ang and Piazzesi (2003), who included macroeconomic variables in the workhorse latent variable affine framework of Duffie and Kan (1996). This extension allows us to investigate questions at the boundaries of macroeconomics and finance. For example, what is the information in the output gap about the compensation investors demand to face interest rate risk? What does today's inflation rate say about the components of expected future real returns to nominal long-term bonds? These and related questions are the focus of intensive research using models that describe the entire term structure using a combination of macroeconomic and latent factors.¹

Yet many of these issues can be examined without attempting to estimate the complete dynamics of the term structure. In particular, only part of the dynamics are relevant for calculating expectations of future bond yields conditioned on observed macroeconomic variables. The logic follows the general asset-pricing approach introduced in Hansen and Singleton (1982), who noted that restrictions implied by no-arbitrage can be exploited without using (or knowing) the complete joint dynamics of asset prices and the pricing kernel. Recall that a zero-coupon bond's price is the expected value of the pricing kernel at the bond's maturity. If we condition this expectation on the information in a set of macroeconomic variables, combine it with the dynamics of the same macro variables, and add a couple of assumptions about risk compensation, we can exploit no-arbitrage restrictions without specifying the remainder of the term structure.

In this paper I describe how to estimate part of the dynamics of affine term structure models. The relation between observable macro variables and the term structure is parameterized without putting much structure on other features of the term structure. There are latent factors in the background, but they play no role in either parameter estimation or in statistical tests of the model's adequacy. The results from partial term structure estimation allow us to extract information from macroeconomic factors about the future evolution of the term structure.

¹Recent work includes Dewachter, Lyrio, and Maes (2002), Dewachter and Lyrio (2002), Hördaahl, Tristani, and Vestin (2002), Ang and Bekaert (2003), Ang, Piazzesi, and Wei (2003), and Rudebusch and Wu (2003).

The main advantage offered by partial term structure estimation is that the researcher need not specify features of the term structure model that are not of direct interest. This has two effects. First, estimation is simplified substantially. Second, misspecification of these features is less likely to contaminate estimates of the dynamics that are of interest.

For concreteness, consider the relation between aggregate output and the term structure. We know that output growth forecasts yields. In addition, yields forecast output growth—they contain information about future output that is not contained in the history of output. To capture these dynamics in a complete term structure model such as Ang et al. (2003), we must specify the number of latent factors (i.e., factors that can be expressed in terms of yields themselves) and functional forms for their dynamics. For example, do they follow moving-average or autoregressive processes? Are the latent factors Gaussian or do they exhibit stochastic volatility? Is the information in the latent factors about future output primarily information about near-term output growth (e.g., today’s one-quarter-ahead forecast of output depends on today’s realization of shocks to latent factors) or more distant output growth (e.g., today’s one-quarter-ahead forecast depends on lagged shocks to latent factors)?

If our research goal is to model the complete term structure, we cannot avoid taking a stand on the entire functional form of the term structure. But if our goal is simply to use the information in the history of output to forecast current and future bond yields and risk premia, the latent factors are nuisance features of the model. The estimation procedure proposed here puts little structure on these factors. Neither the number of factors nor their functional relation with macro factors needs to be specified. Intuitively, the procedure can be viewed as the joint estimation of two sets of regressions. The first set consists of regressions of changes in bond yields on changes in the macro factors. These are estimated with instrumental variables, where the instruments are lagged macro factors. The second set are the regressions comprising a vector autoregression for the macro factors. No-arbitrage imposes cross-equation restrictions on the parameters.

I use this estimation framework to study the relation between inflation and the nominal term structure. I address two questions. First, how sensitive are short-term interest rates to inflation? Second, how sensitive are bond risk premia to inflation? I focus my attention on two periods. The first, from 1960 through the second quarter of 1979, is the “pre-Volcker” sample. The second, from 1984 through 2003, is the “post-disinflation” sample. I find that over both periods, short-term rates move approximately one-for-one with changes in expected future inflation, where the expectations are conditioned on the history of inflation. This result might appear to contradict the existing Taylor rule literature which concludes that the Fed has reacted more aggressively to inflation in the disinflationary period than

in the pre-Volcker period. However, the discrepancy is largely driven by the behavior of inflation and interest rates during 2002 and 2003.

Surprisingly, I find that bond risk premia are fairly insensitive to inflation in both periods. Risk premia are somewhat lower when inflation is high, but the contribution of inflation to variation in risk premia is economically small. The effect is strongest in the early period, where the standard deviation of excess quarterly returns to a five-year bond conditioned on inflation is about thirteen basis points. Put differently, the effect of changes in inflation on the shape of the term structure is determined almost entirely by their effect on expected future short rates, not by their effect on risk premia.

In the next section I describe the modeling framework and the estimation methodology. Section 3 applies the methodology to the relation between inflation and the term structure. Section 4 concludes.

2 The model and estimation technique

Underlying the dynamics of bond yields is some structural model that explains these dynamics in terms of the state of the macroeconomy, central bank policy, and investors' willingness to bear interest-rate risk. Although the model here includes observable macro variables, it is not a structural model. It is closer in spirit to a reduced-form model linking bond yields to macro variables. The formal structure is closely related to the model of Ang and Piazzesi (2003).

Time is indexed by discrete periods t . The length of a period is η years. There are n_0 macro variables realized at time t and stacked in a vector f_t^0 . The term "macro" is fairly arbitrary. In principle these factors can include any observed variable that we are interested in relating to bond yields. We work with an expanded vector of macro factors f_t , which includes lags zero through $p - 1$ of f_t^0 :

$$f_t \equiv \left(f_t^{0'} \quad f_{t-1}^{0'} \quad \cdots \quad f_{t-(p-1)}^{0'} \right)' \quad (1)$$

The length of f_t is $n_f = pn_0$. The choice of p is discussed at various places in this section. For the moment, it is sufficient to note that lags are important both in forming forecasts of future realizations of f_t^0 and in capturing variations in short-term interest rates that are not associated with f_t^0 . In a term structure setting it is important to distinguish between contemporaneous macro variables f_t^0 and the entire state vector f_t . Bond prices depend on compensation investors require to face one-step-ahead uncertainty in the state vector. In (1), only f_t^0 is stochastic given investors' information at $t - 1$.

The period t price of a zero-coupon bond that pays a dollar at period $t + \tau$ is $P_{t,\tau}$. The continuously-compounded annualized yield is $y_{t,\tau}$. The short-term interest rate, which is equivalent to the yield on a one-period bond, is r_t . Macro factors are related to the term structure, but they are insufficient to explain the complete dynamics of the term structure. Latent factors pick up all other variation in bond yields. There are n_x latent factors stacked in a vector x_t . The relation between the factors and the short rate is affine:

$$r_t = \delta_0 + \delta'_f f_t + \delta'_x x_t. \quad (2)$$

Bond prices satisfy the law of one price

$$P_{t,\tau} = E_t(M_{t+1}P_{t+1,\tau-1}) \quad (3)$$

where M_{t+1} is the pricing kernel. The term structure of bond yields depends on the joint dynamics of the pricing kernel, the macro factors, and the latent factors. To motivate the method for estimating the relation between macro factors and the term structure, it is easiest to start with the special case in which the macro factors are independent of the latent factors. The estimation technique in the more general case of correlated factors requires only a slight (but vital) modification to the method that is appropriate for independence.

2.1 Case 1: Independence between macro and latent factors

The contemporaneous macro variables f_t^0 are assumed to follow a vector autoregressive process (VAR) with at most p lags. (Recall that p is the number of lags of f_t^0 included in the vector f_t .) We can always embed a VAR with fewer than p lags into a VAR(p). Thus, since the mathematics of affine term structure models are usually expressed in terms of first-order dynamics, it is convenient to express the macro dynamics as a VAR(1) model for f_t :

$$f_{t+1} - f_t = \mu_f - K_{ff}f_t + \Sigma_f \epsilon_{f,t+1}. \quad (4)$$

The components on the right of (4) are

$$\mu_f = \begin{pmatrix} \mu_0 \\ 0 \end{pmatrix}, \quad K_{ff} = \begin{pmatrix} K_0 \\ C \end{pmatrix}, \quad \Sigma_f = \begin{pmatrix} \Sigma_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \epsilon_{f,t+1} = \begin{pmatrix} \epsilon_{0,t+1} \\ 0 \end{pmatrix}. \quad (5)$$

The vector μ_0 has length n_0 , the matrix K_0 is $n_0 \times n_f$, and the matrix Σ_0 is $n_0 \times n_0$. The vectors of zeros in μ_f and $\epsilon_{f,t+1}$ have length $n_f - n_0$, and the submatrices of zeros in the square matrix Σ_f have the natural dimensions. The elements of $\epsilon_{0,t+1}$ are independent

standard normal innovations. The companion matrix C has the form

$$C = \begin{pmatrix} -I & I & 0 & \dots & 0 & 0 \\ & & & \dots & & \\ 0 & 0 & 0 & \dots & -I & I \end{pmatrix}. \quad (6)$$

The square submatrices in C all have n_0 rows. The double subscript on K_{ff} is used for consistency with the model of correlated factors presented in Section 2.3.

The dynamics of the latent factors have the general affine representation:

$$x_{t+1} - x_t = \mu_x - K_{xx}x_t + \Sigma_x S_{xt} \epsilon_{x,t+1}, \quad (7)$$

where S_{xt} is a diagonal matrix with elements

$$S_{xt(ii)} = \sqrt{\alpha_{xi} + \beta'_{xi}x_t}. \quad (8)$$

The elements of $\epsilon_{x,t+1}$ are independent standard normal innovations. No additional detail about latent factor dynamics is either necessary or useful.

The pricing kernel has the standard log-linear form

$$\log M_{t+1} = -\eta r_t - \Lambda'_{ft} \epsilon_{f,t+1} - \Lambda'_{xt} \epsilon_{x,t+1} - (1/2)(\Lambda'_{ft} \Lambda_{ft} + \Lambda'_{xt} \Lambda_{xt}). \quad (9)$$

The vectors Λ_{ft} and Λ_{xt} are the prices of $\epsilon_{f,t+1}$ risk and $\epsilon_{x,t+1}$ risk respectively. Since f_{t+1}^0 is the only component of f_{t+1} that is unknown at t , without loss of generality the former price of risk can be expressed

$$\Lambda_{ft} = \begin{pmatrix} \Lambda_{0t} \\ 0 \end{pmatrix}. \quad (10)$$

The n_0 -vector Λ_{0t} is the price of risk associated with innovations to f_{t+1}^0 and vector of zeros has length $n_f - n_0$. The price of macro risk, which is the product of macro-factor volatility and the compensation for exposure to $\epsilon_{f,t+1}$, may depend on the macro and latent factors:

$$\Sigma_f \Lambda_{ft} \equiv \begin{pmatrix} \Sigma_0 \Lambda_{0t} \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_f + (\lambda_{ff} \ \lambda_{fx}) \begin{pmatrix} f_t \\ x_t \end{pmatrix} \\ 0 \end{pmatrix}. \quad (11)$$

The vector λ_f has length n_0 , the matrix λ_{ff} is $n_0 \times n_f$, and the matrix λ_{fx} is $n_0 \times n_x$. This is the Gaussian special case of the essentially affine price of risk introduced in Duffee (2002).

The price of risk associated with latent factor shocks has the similar form

$$\Sigma_x S_{xt} \Lambda_{xt} = \lambda_x + (\lambda_{xf} \ \lambda_{xx}) \begin{pmatrix} f_t \\ x_t \end{pmatrix}. \quad (12)$$

Conditions under which this form satisfies no-arbitrage (in the continuous-time limit) are discussed in Kimmel, Cheridito, and Filipovic (2004). As written, (12) allows the price of latent factor risk to depend on both macro and latent factors. I use this general form to highlight the role of a key restriction that will soon be placed on (12).

Bond pricing with this affine setup is standard. Campbell, Lo, and MacKinlay (1997) provide a textbook treatment. I nonetheless go through a few of the steps here for future reference. Guess that log bond prices are affine in the factors:

$$\log P_{t,\tau} = A_\tau + B'_{f,\tau} f_t + B'_{x,\tau} x_t. \quad (13)$$

The recursion implied by law of one price (3), combined with the normally-distributed shocks to f_t and x_t and independence between f_t and x_t , produces

$$\begin{aligned} A_\tau + B'_{f,\tau} f_t + B'_{x,\tau} x_t &= -\eta r_t + A_{\tau-1} \\ &+ B'_{f,\tau-1} E_t(f_{t+1}) + B'_{x,\tau-1} E_t(x_{t+1}) \\ &+ \frac{1}{2} (B'_{f,\tau-1} \Sigma_f \Sigma'_f B'_{f,\tau-1} + B'_{x,\tau-1} \Sigma_x S_{xt}^2 \Sigma'_x B'_{x,\tau-1}) \\ &- B'_{f,\tau-1} \begin{pmatrix} \lambda_f + (\lambda_{ff} \ \lambda_{fx}) \begin{pmatrix} f_t \\ x_t \end{pmatrix} \\ 0 \end{pmatrix} \\ &- B'_{x,\tau-1} \begin{pmatrix} \lambda_x + (\lambda_{xf} \ \lambda_{xx}) \begin{pmatrix} f_t \\ x_t \end{pmatrix} \end{pmatrix}. \end{aligned} \quad (14)$$

The factor loadings $B_{f,\tau}$ and $B_{x,\tau}$ are determined by this recursion. Substitute into (14) the short rate equation (2) and the conditional expectation of f_{t+1} from (4), then match coefficients in f_t to determine one part of this recursion:

$$B'_{f,\tau} = -\eta \delta'_f + B'_{f,\tau-1} (I - K_{ff}^q) - B'_{x,\tau-1} \lambda_{xf}. \quad (15)$$

The matrix K_{ff}^q in (15) is the counterpart to K_{ff} under the equivalent martingale measure:

$$K_{ff}^q = \begin{pmatrix} K_0 + \lambda_{fff} \\ C \end{pmatrix}. \quad (16)$$

Matching coefficients in x_t produces another recursion that, combined with (15), allows for the joint calculation of the loadings $B_{f,\tau}$ and $B_{x,\tau}$. Yet another recursion produces the constant terms A_τ . These other recursions are not relevant here.

The combination of the macro factor dynamics (4), the latent factor dynamics (7), and the coefficients of log bond prices in (13) completely characterize the behavior of bond prices. For example, both the unconditional expectation of $\log P_{t,\tau}$ and its expectation conditioned on time $t-1$ factor values can be calculated. This characterization allows us to use observed bond-price dynamics to estimate the model's parameters. To date, researchers using no-arbitrage models to study term structure dynamics have estimated these complete term structure models. In other words, each parameter's value is either fixed by the researcher or estimated. The motivation behind this methodology is simple: our ultimate goal is to understand all of the dynamic patterns in the term structure.

However, we might make more progress toward this ultimate goal if we are less ambitious in our modeling efforts. Instead of estimating all of the parameters of a term structure model that is unavoidably misspecified, we can estimate particular components of the model while leaving the rest unspecified. This is the point of the estimation procedure described in the next subsection. I estimate the relation between the macro factors and the term structure without characterizing the part of the term structure that is unrelated to the macro factors. I do not have to estimate parameters associated with the latent factors. In fact, I need not even specify the number of latent factors driving the term structure.

In order to proceed I need to add an additional assumption: The price of risk of innovations in the latent factors does not depend on the level of the macro factors. The general form of the price of risk in (12) is restricted by

$$\lambda_{xf} = 0. \tag{17}$$

The role of this assumption is highlighted in the methodology that I now describe.

2.2 Partial term structure estimation with independent factors

The parameters that are identified and estimated by this procedure are δ_f in (2), μ_0 and K_0 in (5), and λ_{ff} in (11). There are three key results that guide the econometric methodology. The first is that the macro factor loadings $B_{f,\tau}$ depend only on these parameters and not on any parameters associated with the latent factors. With assumption (17), the loading on the latent factors drops out of (15). We can solve explicitly the resulting recursion for macro

factor loadings without reference to the parameters of the latent factor dynamics:

$$B_{f,\tau} = - (K_{ff}^q)'^{-1} (I - (I - K_{ff}^q)^\tau) \eta \delta_f. \quad (18)$$

The second key result is that the expectation of differenced log bond yields conditioned on macro variables depends only on information about the macro variables. To understand this result, first-difference the general bond-pricing equation (13), divide by the negative of the bond's maturity (in years) $\eta\tau$ to express it in terms of annualized yields instead of log prices, and rearrange terms, denoting first differences with Δ :

$$\Delta y_{t,\tau} - \left(\frac{-B'_{f,\tau}}{\eta\tau} \right) \Delta f_t = \left(\frac{-B'_{x,\tau}}{\eta\tau} \right) \Delta x_t. \quad (19)$$

The purpose of the first differencing is to remove any information about both A_τ and the unconditional mean of the latent factors. Next, remove any other information about the latent factors by taking the expectation of (19) conditioned on Δf_t . Because f_t and x_t are independent, the conditional expectation of the right side of (19) is zero:

$$E \left(\Delta y_{t,\tau} - \left(\frac{-B'_{f,\tau}}{\eta\tau} \right) \Delta f_t \mid \Delta f_t \right) = 0. \quad (20)$$

The conditional expectation depends only on $B_{f,\tau}$ and Δf_t .

The third key result is that conditional expectations of the macro factors identify the physical dynamics of f_t , and thus identify the parameters of these dynamics. The expectation of Δf_t conditioned on f_{t-1} is:

$$E(\Delta f_t \mid f_{t-1}) - (\mu_f - K_{ff} f_{t-1}) = 0. \quad (21)$$

The parameters that link the macro factors to bond yields can be estimated with Generalized Method of Moments (GMM) using the bond-pricing formula (18) and the moment conditions (20) and (21). At each date $t = 1, \dots, T$ we observe both the current macro factors \tilde{f}_t^0 and the yields \tilde{y}_{t,τ_i} of L zero-coupon bonds with maturities τ_1 through τ_L . (A tilde represents a realized value.) Denote a candidate parameter vector as

$$\hat{\Phi} = \left(\hat{\mu}_0 \quad \hat{\delta}'_f \quad \text{vec}(\hat{K}_0)' \quad \text{vec}(\hat{\lambda}_{ff})' \right)'. \quad (22)$$

There are $n_0 + n_f + 2n_0n_f$ parameters in Φ ; n_0 in μ_0 , n_f in δ_f , and n_0n_f in each of K_0 and λ_{ff} . Denote the true parameter vector by Φ_0 .

Given a parameter vector, the implied macro factor loadings \hat{B}_{f,τ_1} through \hat{B}_{f,τ_L} can be

calculated with (18). The moment vector for observation t is

$$h_t(\hat{\Phi}) = \begin{pmatrix} \left(\Delta \tilde{y}_{t,\tau_1} - \left(\frac{-\hat{B}'_{f,\tau_1}}{\eta\tau_1} \right) \Delta \tilde{f}_t \right) \otimes \Delta \tilde{f}_t \\ \dots \\ \left(\Delta \tilde{y}_{t,\tau_L} - \left(\frac{-\hat{B}'_{f,\tau_L}}{\eta\tau_L} \right) \Delta \tilde{f}_t \right) \otimes \Delta \tilde{f}_t \\ (\Delta \tilde{f}_t^0 - \hat{\mu}_0 - \hat{K}_0 \tilde{f}_{t-1}) \otimes \begin{pmatrix} 1 \\ \tilde{f}_{t-1} \end{pmatrix} \end{pmatrix}. \quad (23)$$

The unconditional expectation of h_t is zero when it is evaluated at the true parameter vector.

We can think of these moments as the moments associated with $L + n_0$ OLS regressions, modified by the requirement of no-arbitrage. To make this clear, consider the top expression in the moment vector, which represents n_f moments associated with the τ_1 -maturity bond. If no-arbitrage is not imposed, the vector B_{f,τ_1} is unrestricted. Then this set of moments corresponds to the moments of the OLS regression of differenced bond yields on differenced macro factors. (There is no constant term in the regression.) Without the requirement of no-arbitrage, the estimate of $-B_{f,\tau_1}/(\eta\tau_1)$ equals the coefficients produced by this regression. Similar OLS regressions are estimated for each of the L bonds. By imposing no-arbitrage, the coefficients from these regressions are required to satisfy cross-equation restrictions.

Now consider the bottom expression in the moment vector, which represents $n_0 \times (1 + n_f)$ moments. If no-arbitrage is not imposed, it corresponds to the moments of n_0 OLS regressions of the VAR(p) model of the macro factors. The estimate of K_0 is then determined by the VAR parameter estimates. If no-arbitrage is imposed but the feedback matrix K_0 under the physical measure has no parameters in common with the feedback matrix $K_0 + \lambda_{ff}$ under the equivalent martingale measure, the interpretation of these moments is unchanged. If any parameter restrictions are placed on λ_{ff} , cross-equation restrictions link the macro factor dynamics and the bond price dynamics.

The parameter estimates solve

$$\hat{\Phi}^* = \underset{\hat{\Phi}}{\operatorname{argmax}} g_T(\hat{\Phi})' W g_T(\hat{\Phi}) \quad (24)$$

where g_T is the mean moment vector

$$g_T(\hat{\Phi}) = \sum_{t=1}^T h_t(\hat{\Phi}) \quad (25)$$

and W is some weighting matrix. The moment vector has length $Ln_f + n_0(1 + n_f)$. If no

restrictions are placed on the model's parameters, the number of moments less the number of free parameters is $n_f(L - 1 - n_0)$. Thus all of the parameters are exactly identified when the number of bonds L is one greater than the number of variables in the contemporaneous macro vector f_t^0 . Additional bonds result in overidentifying restrictions that can be used to test the adequacy of the model.

2.3 Case 2: Correlated factors

A large literature documents that the term structure contains information about future realizations of some macro variables, such as output and inflation, that is not contained in the history of these macro variables.² Thus for at least some choices of macro variables, the assumption of independence between macro and latent factors is untenable. In this subsection I generalize the model to allow for correlations between macro and latent factors. Conveniently, the partial term structure estimation technique described in Section 2.2 requires little modification in order to incorporate the correlation structure introduced here.

If we allow unrestricted joint dynamics between macro factors and an arbitrary number of latent factors, the link between the term structure and macro factors is also arbitrary. To take an extreme case, simply include among the latent factors a vector that is identical to the vector of macro factors. When we estimate term structure models using standard techniques such as maximum likelihood, this indeterminacy does not pose any practical difficulties because the number of latent factors is fixed by the econometrician. The likelihood function is then maximized by allowing the latent factors to pick up variation in the term structure that is not closely related to the macro factors. Here I do not want to restrict the number of latent factors. Therefore I must impose a normalization to eliminate the indeterminacy. I use a normalization that maximizes the ability of current and lagged macro variables to explain variations in the short rate. The latent variables pick up variation in the short rate associated with news that investors receive about future macro variables.

Project the short-term interest rate on a constant and the vector of macro factors f_t :

$$r_t = \delta_0 + \delta_f' f_t + \omega_t, \quad \text{Cov}(f_t, \omega_t) = 0, \quad E(\omega_t) = 0. \quad (26)$$

This projection defines δ_f and the residual ω_t . Although this is simply a projection, and therefore unrestrictive, it takes us outside of commonly-used models of the joint dynamics of macro and latent factors. Consider, for example, a bivariate VAR(1) relation between a

²The literature is too large (and only indirectly related to this paper) to cite fully. See, e.g., Ang et al. (2003) and Diebold, Rudebusch, and Aruoba (2003) for discussions of this forecastability and references to the relevant literature.

scalar f_t and the residual ω_t :

$$\begin{pmatrix} f_{t+1} - f_t \\ \omega_{t+1} - \omega_t \end{pmatrix} = \mu - K \begin{pmatrix} f_t \\ \omega_t \end{pmatrix} + \begin{pmatrix} \epsilon_{f,t+1} \\ \epsilon_{\omega,t+1} \end{pmatrix}. \quad (27)$$

The dynamics (27) do not satisfy the covariance requirement of (26) for arbitrary K . Joint dynamics that satisfy this covariance requirement are discussed informally below and formally in the Appendix.

Assume for the moment that the vector of macro factors f_t includes only the contemporaneous realization f_t^0 . In this case, (26) says that f_t^0 is linked directly to today's short rate through the projection (26). It is also linked to expected future short rates through the VAR dynamics of the macro variables. Put differently, f_t^0 is correlated with future macro variables, and thus correlated with future short-term rates. In principle, there is another channel through which f_t^0 can affect expectations of future short rates: f_t^0 may be correlated with future residuals ω_{t+j} . In order to maximize the explanatory power of the macro factors for the short rate, I rule out this channel. To do so, I include as many lags of f_t^0 in f_t as is necessary in order to pick up any relation between f_t^0 and future short rates that does not work through the correlation between f_t^0 and future macro variables. Formally, I require that ω_t has zero covariance with lags of f_t as well as the contemporaneous f_t :

$$\text{Cov}(\omega_t, f_{t-j}) = 0, \quad j \geq 0. \quad (28)$$

If the covariance in (28) is nonzero for some f_{t-j} , augment f_t with additional lags of f_t^0 . Unlike (26), this is not simply a normalization. The restrictive content of (28) is implicit in the assumption that f_t includes a finite number of lags of f_t^0 . We could write down a model that violates this requirement, in the sense that (28) requires $f_t = (f_t^0 \ f_{t-1}^0 \ \dots \ f_{t-\infty}^0)$.

The residual ω_t is a linear function of a vector of latent variables:

$$\omega_t = \delta'_x x_t. \quad (29)$$

Because I work within an affine framework, I replace the covariance restriction (28) with the stronger restriction that the expectation of x_t conditioned on current and lagged f_t is zero.

$$E(x_t | f_{t-j}) = 0 \quad \forall j \geq 0. \quad (30)$$

This restriction implies that, although the macro and latent factors are correlated, the macro factors do not affect the dynamics of the latent factors. Hence the latent-factor dynamics fit into the general affine class (7) used in Section 2.1. The correlation is created

by the dynamics of the macro factors:

$$f_{t+1} - f_t = \mu_f - K_{ff}f_t - K_{fx}x_t + \Sigma_f \epsilon_{f,t+1}. \quad (31)$$

Consider the “own” dynamics of macro factors—the dynamics conditioned on the history of the macro factors. From (31) and (30),

$$f_{t+1} - f_t = \mu_f - K_{ff}f_t + \xi_{t+1}, \quad (32)$$

$$\xi_{t+1} = -K_{fx}x_t + \Sigma_f \epsilon_{f,t+1}, \quad E(\xi_{t+1} | f_t, \dots, f_{t-\infty}) = 0. \quad (33)$$

Equations (32) and (33) imply that the own dynamics for f_t are an AR(1) (with, perhaps, stochastic volatility introduced by x_t), or equivalently the own dynamics for f_t^0 are an AR(p).

The joint dynamics of the macro factors (31) and latent factors (7) must satisfy (30). The fact that f_t does not appear in (7) does not guarantee that (30) holds. The Appendix describes parameter restrictions on K_{fx} and the latent-factor dynamics (7) that are sufficient to imply (30). Because K_{fx} and all of the components of (7) drop out of the estimation procedure, here it is sufficient to note that the set of models that satisfies these equations is non-empty.

An example may help clarify the role of latent factors in this setting. Assume that f_t contains current and lagged GDP growth. Imagine that the central bank decides to tighten monetary policy in response to events unrelated to the history of GDP growth. This tightening raises short-term rates today. Because GDP does not react immediately to the tightening, the increase in short-term rates shows up in the residual ω_t , and thus in the latent factors x_t . The tightened monetary policy affects future GDP, which is captured in the model by the matrix K_{fx} . There is no feedback from GDP to the latent factors because by construction, the latent factors contain news about future macro variables that is not embedded in current macro variables.

The model is completed with the specification of the pricing kernel. I use (9), which is the specification I used for the case of independent factors. I also use the same functional forms for risk compensation, given by (11) and (12).

Bond pricing formulas are calculated in the usual way. Guess the log-linear form (13) holds and apply the law of one price. The result is (14). Although the form of this equation is unchanged by the introduction of correlated factors, the interpretation is different. With correlated factors, the period- t expectation of f_{t+1} depends on both macro and latent factors. As in the case of independent factors, match coefficients from (14) in f_t . This step uses the special structure placed on the joint dynamics of f_t and x_t . Because $E_t(x_{t+1})$ does not

depend on f_t , this matching results in the recursion (15), as in the case of independent factors. Finally, by imposing assumption (17), the recursion for $B_{f,\tau}$ can be solved explicitly, producing (18), as in the case of independence.

Why are the macro factor loadings $B_{f,\tau}$ unchanged when the assumption of independence between macro and latent factors is dropped? The reason is the normalization that maximizes the explanatory power of macro factors for the short rate. Latent factors are not allowed to pick up any explanatory power for the term structure that is already picked up by the history of the macro factors. The only effect of introducing correlated factors is that the model's parameters can no longer be estimated with the technique described in Section 2.2. I describe a modified technique in the next subsection.

2.4 Partial term structure estimation with correlated factors

As in the case of independence, here the parameters δ_f , μ_0 , K_0 , and λ_{ff} can be estimated without imposing additional structure on the latent factors. There is one important difference. With independence, the expectation of the right side of (19) conditioned on Δf_t is zero. With correlated factors, this is no longer true because x_{t-1} may contain information about f_t . Therefore I take the expectation of (19) conditioned on f_{t-1} and apply (30):

$$E \left(\Delta y_{t,\tau} - \frac{-B'_{f,\tau}}{\eta\tau} \Delta f_t | f_{t-1} \right) = 0. \quad (34)$$

The corresponding moment vector for observation t is

$$h_t(\hat{\Phi}) = \begin{pmatrix} \Delta \tilde{y}_{t,\tau_1} - \frac{-\hat{B}'_{f,\tau_1}}{\eta\tau} \Delta \tilde{f}_t \\ \dots \\ \Delta \tilde{y}_{t,\tau_L} - \frac{-\hat{B}'_{f,\tau_L}}{\eta\tau} \Delta \tilde{f}_t \\ \Delta \tilde{f}_t^0 - \hat{\mu}_0 - \hat{K}_0 \tilde{f}_{t-1} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ \tilde{f}_{t-1} \end{pmatrix}. \quad (35)$$

Recall that with independence between macro and latent factors, the moment vector (23) was interpreted as moments of OLS regressions where cross-equation restrictions were imposed on the OLS parameter estimates. Almost the same interpretation can be applied to (35). The only difference is that the regressions of differenced yields on differenced macro factors are estimated with instrumental variables instead of OLS. The instruments are a constant and lagged macro factors. As with (23), no-arbitrage imposes cross-equation restrictions on the estimated parameters. Section 3 contains some additional discussion about the inappropriateness of OLS moment conditions when the latent factors contain information about future realizations of the macro factors.

The remainder of this section examines in detail some of the features of this model. I first discuss the intuition that underlies the instrumental variables approach.

2.5 Are all macro realizations alike?

Denote the instruments used in the moment condition (35) as $z'_{t-1} = \{1 f'_{t-1}\}$. The realization of f_t can be decomposed into the sum of three components: a component that is projected on z_{t-1} , a component that is in investors' information sets at t but is orthogonal to z_{t-1} , and a realized period- t shock. The model says that these three components have the same relation to the short rate. In other words, in (26) only f_t enters—the same vector δ_f multiplies each of these three components. Similarly, the time- t expectation of f_{t+1} and time- t compensation for risk do not depend separately on these three components, but only on their sum. Therefore bond prices are functions of f_t , not functions of the separate components of f_t .

The dependence on f_t instead of the individual components justifies the use of instrumental variables to estimate the loadings on the macro factors. Write the change in the τ -maturity yield from $t - 1$ to t as the sum of two pieces: a component that is projected on z_{t-1} and a residual. Substitute (31) into (19):

$$\Delta y_{t,\tau} = \left(\frac{-B'_{f,\tau}}{\eta\tau} \right) \left[E(\Delta f_t | z_{t-1}) \right] + \left\{ \left(\frac{-B'_{f,\tau}}{\eta\tau} \right) (-K_{fx}x_t + \Sigma_{f\epsilon_{f,t+1}}) + \left(\frac{-B'_{x,\tau}}{\eta\tau} \right) \Delta x_t \right\} \quad (36)$$

where

$$E(\Delta f_t | z_{t-1}) = \mu_f - K_{ff}f_{t-1}. \quad (37)$$

The residual term in curly brackets is orthogonal to f_{t-1} . Thus a regression of changes in yields on changes in the macro factors using instruments z_{t-1} produces consistent estimates of the (yield) loadings on the macro factors. The identifying assumption is that the loading on $E(\Delta f_t | z_{t-1})$ is the same as the loading on Δf_t .

Although this is the standard IV assumption, its validity in this setting is not guaranteed. As mentioned at the beginning of this section, the affine relation linking r_t and the factors is best thought of as a reduced form. In reality, there are a variety of structural shocks and policy responses that determine the realization of f_t . This paper offers no economic theory to justify why they all should have the same relation to bond yields. A more robust interpretation of the reduced form (26) is that it represents the average relation between the short rate and the individual components of f_t . Similarly, the IV regression represents the average relation between the short rate and the components of f_t spanned by z_{t-1} . The structure of the model (or, more precisely, the lack of structure) does not allow a test of the hypothesis that the relation between r_t and shocks to f_t is the same as the relation between

r_t and the predictable components of f_t .

2.6 Interpreting the restriction on the price of risk

What role is played by the assumption in (17)? Recall that under the physical measure the macro factors do not affect the dynamics of the latent factors. (This is true for both the case of independence and the case of correlated factors.) The combination of (17) and this feature of the physical dynamics implies that macro factors do not affect the dynamics of the latent factors under the equivalent martingale measure.

The purpose of this restriction is easy to see. The macro factor loadings $B_{f,\tau}$ are determined by the effect, under the equivalent martingale measure, of f_t on current and expected future short-term interest rates. The effect of f_t on the equivalent-martingale expectation of r_{t+j} depends on (a) the effect of f_t on equivalent-martingale expectations of f_{t+j} and x_{t+j} ; and (b) the parameters δ_f and δ_x , which determine the effect of f_t and x_t on the time- t short rate. Thus in general, we cannot determine $B_{f,\tau}$ without knowing δ_x . This leads us down the road of specifying the dynamics of latent factors. By setting $\lambda_{xf} = 0$, f_t has no effect on x_{t+j} under either the physical or the equivalent martingale measure. We can therefore determine the effect of f_t on bond prices without specifying features of the latent factors.

The preceding analysis suggests that the $\lambda_{xf} = 0$ is critical to the applicability of the estimation technique. But in a broader sense, it is more like an identification assumption that allows us to disentangle variations in price of risk of macro factors from variations in the price of risk of latent factors. Recall that the effect of a shock to some individual factor i on the term structure is determined by the factor's loading $B_{i,\tau}$. If the price of factor- i risk depends in part on factor k , then variations in the level of the k^{th} factor correspond to variations in risk premia on the τ -maturity bond that are proportional to $B_{i,\tau}$. Now imagine there are two factors, i and j , with the same factor loadings: $B_{i,\tau} = B_{j,\tau} \forall \tau$. It is impossible to distinguish between the sensitivity of the price of factor i risk with respect to k from the sensitivity of the price of factor j risk with respect to k . Only their sum is identified. The reason is that variations in k produce the identical effects on risk premia across the term structure, regardless of whether the sensitivity is associated with factor i or factor j .

In models where all factors are latent, this indeterminacy is not a practical problem because no reasonable estimation procedure will produce two latent factors that have identical effects on the term structure. But when we exogeneously specify some of the factors, we open the door to this kind of indeterminacy. For example, the combination of macro and term structure data allow us to distinguish between a parallel shock to the term structure that is associated with a shock to GDP and a parallel shock that is unaccompanied by a

shock to GDP.

In a setting with multiple macro factors the problem is more severe than this example illustrates. If I can write some latent factor's loadings as a linear combination of the loadings on the macro factors, prices of risk cannot be uniquely determined. To show this I expand the recursion formula (15). Denote element i of the macro factor loading vector $B_{f,\tau}$ by $B_{f(i),\tau}$. Denote row i of λ_{ff} by $\lambda_{ff(i)}$. (This is a row vector.) Similar notation denotes row i of λ_{xf} . Then (15) can be written as

$$\begin{aligned} B'_{f,\tau} &= -\eta\delta'_f + B'_{f,\tau-1}(I - K_{ff}) \\ &\quad - B_{f(1),\tau-1}\lambda_{ff(1)} - B_{f(2),\tau-1}\lambda_{ff(2)} - \cdots - B_{f(n_0),\tau-1}\lambda_{ff(n_0)} \\ &\quad - B_{x(1),\tau-1}\lambda_{xf(1)} - B_{x(2),\tau-1}\lambda_{xf(2)} - \cdots - B_{x(n_x),\tau-1}\lambda_{xf(n_x)}. \end{aligned} \quad (38)$$

Now assume that the loading for, say, the first latent factor is a linear combination of the macro factor loadings:

$$B_{x(1),\tau} = \gamma_1 B_{f(1),\tau} + \cdots + \gamma_{n_0} B_{f(n_0),\tau} \quad \forall \tau. \quad (39)$$

We can then substitute the latent factor loading $B_{x(1),\tau-1}$ out of (38):

$$\begin{aligned} B'_{f,\tau} &= -\eta\delta'_f + B'_{f,\tau-1}(I - K_{ff}) \\ &\quad - B_{f(1),\tau-1}(\lambda_{ff(1)} + \gamma_1\lambda_{xf(1)}) - \cdots - B_{f(n_0),\tau-1}(\lambda_{ff(n_0)} + \gamma_{n_0}\lambda_{xf(1)}) \\ &\quad - B_{x(2),\tau-1}\lambda_{xf(2)} - \cdots - B_{x(n_x),\tau-1}\lambda_{xf(n_x)}. \end{aligned} \quad (40)$$

The same kind of substitution can be used for any latent factor that has loadings which can be written as a linear combination of the macro factor loadings.

The main message of this discussion is that the interpretation of λ_{ff} is not as rigid as the model's mathematics implies. In the model, element (i, j) of this matrix is the sensitivity in the price of risk of factor i associated with factor j . But it is more accurately interpreted as the sensitivity to factor j in the price of risk of any shock to the term structure that behaves like a shock to factor i .

2.7 Relaxing the affine structure

The affine dynamics of the latent factors x_t are not essential. The only reason the affine form is used is to guarantee joint log-normality of bond prices and the pricing kernel. Log-normality is what allows us to produce the recursion (14) from the law of one price. A more

general framework replaces (30) and (31) with

$$f_{t+1} - f_t = \mu_f + K_{ff}f_t + K_{fx}(x_t) + \Sigma_f \epsilon_{f,t+1}, \quad (41)$$

where $K_{fx}(x_t)$ is some function of the latent factors that satisfies

$$E(K_{fx}(x_t)|f_t) = 0. \quad (42)$$

The latent factor dynamics (7) can be replaced with

$$x_{t+1} - x_t = K_{xx}(x_t) + \Sigma_x S_x(x_t) \epsilon_{x,t+1}, \quad (43)$$

where $K_{xx}(x_t)$ and $S_x(x_t)$ are general functions of the latent factors. The innovations $\epsilon_{f,t+1}$ and $\epsilon_{x,t+1}$ are independent multivariate standard normal shocks. Replace the affine form for log-bond prices (13) with

$$\log P_{t,\tau} = A_\tau + B'_{f,\tau} f_t + w_\tau(x_t), \quad (44)$$

where $w_\tau(x_t)$ is a general function of x_t . The one-step-ahead forecast of this function is

$$w_\tau(x_{t+1}) = E_t(w_\tau(x_{t+1})) + \varepsilon_{\tau,t+1}, \quad (45)$$

where the residual $\varepsilon_{\tau,t+1}$ is a normally-distributed shock that is independent of $\epsilon_{f,t+1}$. The expectation of this function conditioned on contemporaneous and lagged macro factors is zero:

$$E(w_\tau(x_t)|f_t, \dots, f_{t-\infty}) = 0. \quad (46)$$

Equation (44) with $\tau = 1$ replaces the short rate equation (26). The form of the stochastic discount factor (9) is unchanged. The law of one price then implies that

$$\begin{aligned} A_\tau + B'_{f,\tau} f_t + w_\tau(x_t) &= A_1 + B'_{f,1} f_t + w_1(x_t) + A_{\tau-1} \\ &+ B'_{f,\tau-1} E_t(f_{t+1}) + E_t(w_{\tau-1}(x_{t+1})) \\ &+ \frac{1}{2} (B'_{f,\tau-1} \Sigma_f \Sigma'_f B'_{f,\tau-1} + \text{Var}_t(\varepsilon_{\tau-1,t+1})) \\ &- B'_{f,\tau-1} \Sigma_f \Lambda_{ft} - \text{Cov}_t(\varepsilon_{\tau-1,t+1}, \Lambda'_{xt} \epsilon_{x,t+1}). \end{aligned} \quad (47)$$

The final assumptions concern expectations conditioned on macro factors:

$$E(\text{Var}_t(\varepsilon_{\tau-1,t+1})|f_t) = V_{1,\tau-1}, \quad E(\text{Cov}_t(\varepsilon_{\tau-1,t+1}, \Lambda'_{xt} \epsilon_{x,t+1})|f_t) = V_{2,\tau-1}, \quad (48)$$

$$\Sigma_f E(\Lambda_{ft}|f_t) = \begin{pmatrix} \lambda_f + \lambda_{ff}f_t \\ 0 \end{pmatrix}. \quad (49)$$

With these assumptions, take the expectation of (47) conditioned on f_t . All terms involving latent factors drop out (or become constants), producing

$$B'_{f,\tau}f_t = \kappa_\tau + B'_{f,1}f_t + B'_{f,\tau-1}(I - K_{ff} - \lambda_{ff})f_t \quad (50)$$

where κ_τ is a maturity-dependent constant. Matching coefficients in f_t produces the bond-pricing formula (18) with $\eta\delta_f = -B_{f,1}$.

2.8 What can we do with partial dynamics of the term structure?

The cost of estimating only part of the dynamics of the term structure is that certain questions cannot be answered. It is helpful to illustrate the kinds of questions that can be addressed with this estimation procedure.

- How does the expected time-path of r_t depend on f_t ?

The response of the short rate to a macroeconomic shock is $\delta'_f \epsilon_{f,t}$. The expected change in the short rate from t to $t + j$, conditioned on f_t , is

$$E(r_{t+j} - r_t|f_t) = \delta'_f (I - (I - K_{ff})^j) (K_{ff}^{-1}\mu_f - f_t). \quad (51)$$

Note that this j -ahead forecast is not a minimum-variance forecast. There is additional information in the term structure (such as the current level of the short rate) that is ignored in forming this conditional expectation. Therefore the partial term structure dynamics should not be used to forecast, but rather to interpret the link between macroeconomic variables and the term structure.

- How do risk premia on bonds vary with f_t ?

The partial nature of the estimated model does not allow us to say anything about mean excess returns to bonds. However, it does allow us to determine how variations in f_t correspond to variations in expected excess returns. The expected excess log return to a τ -maturity bond held from t to $t + 1$, conditioned on f_t , is

$$E(\log P_{t+1,\tau-1} - \log P_{t,\tau} - \eta r_t|f_t) = \kappa_{\tau-1} + B'_{f,\tau-1} \begin{pmatrix} \lambda_{ff} \\ 0 \end{pmatrix} f_t. \quad (52)$$

The constant term $\kappa_{\tau-1}$ is unrestricted.

- How does the term structure react to an innovation in a macroeconomic variable?

Consider the expectation of the τ -maturity annualized bond yield $y_{t,\tau}$ conditional on the contemporaneous f_t . (In other words, we observe f_t but we have no information about the latent factors.) The expectation is

$$y_{t,\tau} = E(y_{t,\tau}|f_t) + \nu_{t,\tau}, \quad (53)$$

$$E(y_{t,\tau}|f_t) = a_\tau + \frac{1}{\tau} \delta'_f (I - (I - K_{ff}^q)^\tau) (K_{ff}^q)^{-1} f_t, \quad (54)$$

$$a_\tau = -(1/\tau)A_\tau, \quad \nu_{t,\tau} = -(1/\tau)B'_{x,\tau}x_t. \quad (55)$$

The constant term a_τ is unrestricted.

- What does f_t tell us about the future evolution of the term structure?

The j -period-ahead forecast of the change in the yield on a bond with constant maturity τ is

$$\begin{aligned} E(y_{t+j,\tau} - y_{t,\tau}|f_t) = \\ \frac{1}{\tau} \delta'_f (I - (I - K_{ff}^q)^\tau) (K_{ff}^q)^{-1} (I - (I - K_{ff})^j) (K_{ff}^{-1} \mu_f - f_t). \end{aligned} \quad (56)$$

- Is the empirical failure of the expectations hypothesis associated with f_t ?

Campbell and Shiller (1991) estimated regressions of the form

$$y_{t+s,l-s} - y_{t,l} = b_0 + b_1 \frac{s}{l-s} (y_{t,l} - y_{t,s}) + e_{t+s,l,s} \quad (57)$$

for maturities $l > s$. Under the weak form of the expectations hypothesis the coefficient b_1 should equal one, but in the data it is often negative. A common interpretation of this result is that bond risk premia and the slope of the term structure are positively correlated. Is this also true of the variability in the term structure that is associated with variations in f_t ? Consider estimating (57) using f_t as instruments. The conditional expectation of yield spread on the right of (57) can be expressed as

$$E(y_{t,l} - y_{t,s}|f_t) = \theta_{l,s} + \left(-\frac{1}{l} B_{f,l} + \frac{1}{s} B_{f,s} \right)' f_t \quad (58)$$

where $\theta_{l,s}$ is an unrestricted constant. The conditional expectation of the left side of (57)

can be expressed as

$$E(y_{t+s,l-s} - y_{t,l}|f_t) = \phi_{l,s} + \frac{s}{l-s} E(y_{t,l} - y_{t,s}|f_t) - \frac{1}{l-s} B'_{f,l-s} ((I - K_{ff})^s - (I - K_{ff}^q)^s) f_t \quad (59)$$

where $\phi_{l,s}$ is an unrestricted constant. If $\lambda_{ff} = 0$, then $K_{ff} = K_{ff}^q$ and the final term in (59) is identically zero. In this case, the population estimate of b_1 from IV estimation of (57) is one. More generally, the population regression coefficient is

$$\hat{b}_1 = 1 - \frac{1}{s} \left[\left(-\frac{1}{l} B_{f,l} + \frac{1}{s} B_{f,s} \right)' \text{Var}(f_t) \left(-\frac{1}{l} B_{f,l} + \frac{1}{s} B_{f,s} \right) \right]^{-1} \times \left(-\frac{1}{l} B_{f,l} + \frac{1}{s} B_{f,s} \right)' \text{Var}(f_t) ((I - K_{ff})^s - (I - K_{ff}^q)^s)' B_{f,l-s} \quad (60)$$

where $\text{Var}(f_t)$ is the unconditional variance-covariance matrix of f_t . Given this variance and the parameters of the term structure model, the regression coefficient can be computed.

In the next section I illustrate some of these applications by using the model to study the joint dynamics of inflation and the term structure.

3 An application: Inflation and the term structure

Researchers have long studied the relation between inflation and bond yields. I reexamine this relation using the model of correlated factors developed in Section 2.3. The vector of macro factors consists of current and lagged inflation:

$$f_t = \begin{pmatrix} \pi_t & \dots & \pi_{t-(p-1)} \end{pmatrix}. \quad (61)$$

Thus the short rate equation (26) looks something like a Taylor (1993) rule regression. The Taylor rule adds a measure of the period- t output gap to this equation and, depending on the implementation, may include only contemporaneous inflation or impose constraints on the parameters.³ The main difference between the empirical analysis here and the usual empirical implementations of the Taylor rule is that I use information from the term structure to both refine the estimate of the short rate's loading on inflation δ_f and to simultaneously estimate the sensitivity of the price of interest rate risk to the level of inflation. Ang and Piazzesi (2003) also investigate the latter issue, although their methodology is quite different.

³For example, the short rate in quarter t is often expressed as an affine function of inflation during the past year, implying that f_t contains lags zero through three of quarterly inflation and that $\delta_{f(i)} = \delta_{f(j)}$, $i \neq j$.

I first describe the data used to estimate these relations. I then discuss how to choose the number of lags of inflation included in f_t . This leads into a comparison of the estimation techniques used when the residual of (26) is either independent or correlated with future inflation. I then present the results of the estimation.

3.1 The data

I use quarterly data from 1960 through 2003. The beginning date is chosen to match the sample used by Clarida, Galí, and Gertler (2000) in their empirical study of the Taylor rule. Inflation in quarter t is measured by the change in the log of the personal consumption expenditure (PCE) chained price index from $t - 1$ to t . I define quarter- t bond yields as yields as of the end of last month in the quarter. This choice is a compromise between two reasonable alternatives: using average yields within a quarter, as inflation is measured, or using yields observed some time after the end of the quarter, so that we are sure the yields incorporate the information in the announced inflation rate for the previous quarter. The short-term rate is the three-month yield from the Center for Research in Security Prices (CRSP) riskfree rate file. Yields on zero-coupon bonds with maturities of one and five years are taken from the CRSP Fama-Bliss file. Inflation and bond yields are continuously compounded and expressed as annual rates.

Table 1 reports summary statistics for various subperiods. I divide the full sample at two points: after 1979Q2 and after 1983Q4. The first break point corresponds to the beginning of the Volcker tenure at the Fed and the accompanying disinflation. There is substantial evidence that a regime change in the joint dynamics of inflation and interest rates occurred at that time.⁴ This break point is also used by Clarida et al. (2000). The second break point corresponds to the end of the disinflation. Its precise placement is somewhat arbitrary because it is harder to determine when the disinflation ended than when it began. I use 1983Q4 so that there are enough observations to identify the model's parameters during the disinflationary period.

Many characteristics of these data are common to all three periods, including the high persistence of both inflation and yields. I assume that both interest rates and inflation are stationary processes. Although this assumption is typical in both the term structure and Taylor rule literatures, it is motivated more by economic intuition (and econometric convenience) than by statistical evidence. Unit root tests typically fail to reject the hypothesis of nonstationarity for either interest rates or inflation. Contemporaneous correlations between changes in inflation and changes in interest rates are fairly low, ranging from about 0.25

⁴See, e.g., Gray (1996) and the earlier research he cites.

in the early sample to about 0.10 in the late sample. In Section 3.3 I discuss why these correlations understate the true relation between inflation and interest rates.

The focus on the three-month, one-year, and five-year yields is motivated by the following considerations. The three-month maturity is the shortest consistent with the quarter-length periods used in the model and the five-year maturity is the longest zero-coupon bond available from CRSP. The one-year yield is at about the midpoint between these two years—not in terms of maturity but in terms of comovement. We see from Table 1 that in both the disinflationary and post-disinflation periods, the correlation between quarterly changes in one-year yields and three-month yields is within a percentage point of the corresponding correlation between one-year yields and five-year yields. During the pre-Volcker period, variations in the one-year yield are a little closer to variations in the long end of the term structure than the short end.

Yields on bonds of intermediate maturities are available, but including them has two consequences. First, adding additional moment conditions expands the wedge between finite-sample and asymptotic properties of GMM estimation. Second, using yields on bonds of similar maturities increases the likelihood that the model’s parameter estimates will be determined by economically unimportant properties of these yields. Efficient GMM estimation emphasizes the linear combinations of yields that are statistically most informative about the model. Moments involving yield spreads on similar-maturity bonds are likely to be highly informative because such spreads exhibit little volatility. If the model is right and the yields are observed without noise, including bonds of similar maturities is a good way to pin down the parameters. But we know the model is only an approximation to reality, and the zero-coupon bond yields are interpolated. I therefore use a small number of points on the yield curve that capture its general shape.⁵

Monthly observations of inflation and yields are also available. Monthly data contain more information but their use requires both more parameters and more GMM moment conditions. The number of inflation lags p included in the vector f_t must capture both the autoregressive properties of inflation and the relation between lagged inflation and current bond yields. These properties are driven more by calendar time than by frequency of observation. Thus shifting to monthly data will triple both the amount of available data and the lag length p . With $n_0 = 1$ (a single contemporaneous macro variable) and L bond yields, the

⁵A comparison with maximum likelihood term structure estimation may be helpful. One method used to estimate an n -factor term structure model is to assume that n points on the term structure are observed without error and other points are contaminated by measurement error. In principle, any n maturities will work, yet in practice the n maturities are widely spaced in order to force the model to fit the overall shape of the term structure. The estimation procedure used in this paper does not rely on ad hoc noise, but as a consequence it is more difficult to use information from many points on the term structure.

number of moment conditions in (23) is $p(L + 1) + 1$ and the number of moment conditions in (35) is $(p + 1)(L + 1)$. The number of parameters is $1 + 3p$. (The AR(p) description of inflation uses $1 + p$ parameters and there are p parameters in both δ_f and λ_{ff} .) Thus the number of moment conditions and parameters increases almost proportionally with p . Put differently, the number of data points per moment condition (and per parameter) increases only slightly if monthly data are used. For the sake of parsimony, I use quarterly data.

3.2 The choice of lag length

The lag length p must be at least as large as the number of lags necessary to capture the autoregressive properties of inflation. I estimated autoregressions using up to six lags and calculated the Akaike and Bayesian Information Criteria (AIC and BIC) for each. For the full sample, both criteria are minimized with three lags. For the early sample, both criteria are minimized with a single lag. For the late sample, the AIC is minimized with three lags and the BIC is minimized with a single lag. (None of these results are reported in any table.)

The lag length must also be sufficiently large so that the residual of (26) is uncorrelated with lags of inflation. In other words, adding lags to (61) should not increase the explanatory power of current and lagged inflation. There is no consensus in the Taylor rule literature as to the proper lag length. (That literature typically interprets lags in terms of slow reaction of the Fed to inflation and output.) Using different econometric frameworks, Clarida et al. (2000), Rudebusch (2002), and English, Nelson, and Sack (2003) arrive at different conclusions about the persistence of the Fed's reaction function. A recent review of the evidence is in Sack and Wieland (2000).

We might be tempted to treat (26) as a regression equation and use the information criteria to choose the appropriate lag length. But estimation of (26) is problematic for the same reason that estimation of the Taylor rule is problematic: the residual exhibits very high serial correlation. To illustrate the problem, consider estimation of (26) with $p = 3$ over the period 1984 through 2003. The estimated equation is

$$r_t = 1.82 + 0.53\pi_t + 0.33\pi_{t-1} + 0.41\pi_{t-2} + \omega_t. \quad (62)$$

The first-order autocorrelation of ω_t is 0.9. This high autocorrelation makes it difficult to test hypotheses and construct reliable standard errors. Accordingly, I defer further discussion of the choice of p in order to discuss in more detail the methods we can use to estimate the parameters of (26). The choice of method critically depends on the relation between the residual ω_t and future inflation.

3.3 The relation between inflation and the short rate

A natural method to correct for the high autocorrelation of ω_t is to first-difference (26):

$$r_t - r_{t-1} = \delta'_f(f_t - f_{t-1}) + (\omega_t - \omega_{t-1}). \quad (63)$$

When f_t consists of lags of inflation, the residual of (63) is much closer to white noise than is the residual of (26). If we adopt the assumption that ω_{t-1} is orthogonal to f_t , (63) can be estimated with OLS. However, this assumption is inconsistent with both intuition and evidence.⁶ Investors at time $t - 1$ have more information about inflation during t than is contained in the history of inflation. Since investors care about real returns, presumably the short rate at $t - 1$ (which is a nominal return earned during period t) depends on this information. If so, ω_{t-1} will be positively correlated with f_t . Therefore $f_t - f_{t-1}$ is negatively correlated with $\omega_t - \omega_{t-1}$ and the OLS estimate of δ_f is biased. Similarly, contemporaneous correlations between changes in inflation and changes in bond yields are relatively small because news about next period's inflation rate dampens these correlations.

As discussed in Section 2.5, estimation with a particular set of instruments avoids this bias, subject to the maintained hypothesis that (63) is correctly specified. The instruments are a constant and f_{t-1} . Table 2 reports results of estimating (63) with these instruments when f_t contains lags zero to two of quarterly inflation. Standard errors are adjusted for generalized heteroskedasticity and four lags of moving average residuals using the technique of Newey and West (1987a).⁷ The results for the full sample are puzzling. The sign of the estimated relation (negative) is wrong and the standard errors are huge. Moreover, the fitted residuals are positively correlated with contemporaneous changes in inflation. The intuition behind the bias in OLS coefficients implies that this correlation should be negative.

By contrast, the subsample results are in line with our intuition, and contradict the results from the full sample. In both the early and late periods, the coefficient on the contemporaneous change in inflation is about 0.5. (This is also true in the disinflationary period, but the middle-period results are shown only for completeness—there are too few observations to draw any conclusions.) The coefficients on lagged changes in inflation are also positive in both of these subperiods, while the correlations between fitted residuals and contemporaneous changes in inflation are strongly negative. The negative correlation implies that short-term rates lead inflation. Further evidence of this predictability is the positive correlation between fitted residuals and the next quarter's change in inflation. All of these

⁶A large empirical literature beginning with Fama (1975) considers the forecast power of interest rates for inflation.

⁷The sample autocorrelations of the residuals (not reported in any table) are fairly close to zero at all lags.

results are consistent with our intuition about the relation between inflation and interest rates.

What explains the anomalous full-sample results? The assumptions underlying the IV regression are not satisfied over the full sample because the relation between the instruments and the explanatory variables has changed over time. In other words, inflation dynamics over this period are not stable, as we can see from Table 3. In the early subperiod, inflation basically follows an AR(1). In the late subperiod, inflation dynamics are more complicated. The idea of the IV regression is that changes in short-term rates are projected on expected changes in inflation, where expectations are conditioned on lagged inflation. For the purposes of the regression, this expectation is proxied by an in-sample projection of changes in inflation on lagged inflation. Because inflation dynamics have varied over the period, true conditional expectations do not correspond to the full-sample projection. This problem is avoided by splitting the sample into subperiods that exhibit stable dynamics.

I now return to the question of lag length. I reestimated the IV regressions in Table 2, adding a fourth lag of differenced inflation as an explanatory variable and a fourth lag of inflation as an instrument. For all of the regressions, I could not reject the hypothesis that the coefficient on the additional lag of differenced inflation is zero. I therefore set $p = 3$, which is sufficient to capture the dynamics of inflation in both the early and late subperiods.

3.4 Details of model estimation

To summarize, the relevant components of the term structure model are:

$$f_t = \begin{pmatrix} \pi_t & \pi_{t-1} & \pi_{t-2} \end{pmatrix}, \quad (64)$$

$$r_t - r_{t-1} = \delta_f'(f_t - f_{t-1}) + (\omega_t - \omega_{t-1}), \quad (65)$$

$$E(\pi_t | f_{t-1}) - \pi_{t-1} = \mu_0 + K_0 f_{t-1} + \xi_t, \quad (66)$$

$$E^q(\pi_t | f_t) - \pi_{t-1} = \mu_0^q + (K_0 + \lambda_{ff})f_{t-1} + \xi_t^q. \quad (67)$$

The identified parameters are scalar μ_0 and the vectors δ_f , K_0 , and λ_{ff} , each of which has three elements. Instead of reporting estimates of K_0 , I report the implied coefficients of the AR(3) for inflation,

$$\rho \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} - K_0. \quad (68)$$

The model allows the price of interest rate risk to depend on both contemporaneous and lagged inflation. I report results only for the special case where the price of risk depends on contemporaneous inflation, or $\lambda_{ff(2)} = \lambda_{ff(3)} = 0$. There are two reasons. First, for both the

pre-Volcker and post-disinflation periods, the more general functional form does not provide any statistically significant improvement in fit. Second, estimation of the general model over alternative subperiods sometimes produces an estimate of the equivalent-martingale feedback matrix K_{ff}^q that fits the observed bond yields well, but implies wildly implausible behavior for yields that are not included in the estimation.⁸

I estimate the model separately over the pre-Volcker period 1960Q1 through 1979Q2 and the post-disinflation period 1984Q1 through 2003Q4. For completeness, I also estimate the model over the disinflationary period, although the sample period is too short to draw any meaningful conclusions. In fact, for this eighteen-quarter period I set $p = 2$ because there are too few observations to estimate the model using the moments for $p = 3$. The GMM methodology is described in Section 2.4. The moment vector is (35). I use two iterations of GMM. For the first iteration, the weighting matrix is the inverse of the sample covariance matrix of the moments evaluated at “regression/constant risk premia” parameters. These parameters are determined by an AR(3) regression of inflation, IV estimation of (63), and $\lambda_{ff} = 0$. The parameter estimates produced by this first iteration are used to construct an asymptotically efficient weighting matrix and the parameters are estimated again. The covariance matrix of the moment vector is estimated using the robust method of Newey and West (1987a) with four moving average lags. The solution to the GMM optimization problem requires nonlinear optimization. To find the global minimum, I randomly generate 20 starting values. For each starting value, I use simplex to get in a well-behaved neighborhood of the parameter estimates. I then use a derivative-based algorithm to improve the accuracy of the estimates.

3.5 Results

The results are displayed in Table 4. Panel A reports parameter estimates and Panel B reports specification tests. The first specification test is the Hansen (1982) J -test of over-identifying restrictions. The second is a likelihood ratio test of the hypothesis $\delta_f = \rho$. This condition implies that short rates can be written as

$$r_t = \delta_0 + E(\pi_{t+1}|f_t) + \omega_t. \quad (69)$$

In other words, ex ante real short rates are uncorrelated with expected inflation. It has an asymptotic $\chi^2(p)$ distribution under the null. The third is a Lagrange multiplier test of the hypothesis that the price of risk depends on the first two elements of f_t instead of just the

⁸When this occurs, some eigenvalues of $I - K_{ff}^q$ are typically imaginary with absolute values outside of the unit circle.

first element. It has an asymptotic $\chi^2(1)$ distribution under the null. The latter two test statistics are derived in Newey and West (1987b).

There are three main conclusions to draw from these results. First, in both the early and late periods there is a strong positive relation between the short rate and inflation. Of course, we do not need a no-arbitrage model to tell us this; standard methods such as the IV regressions in Table 2 also document this relation. The value of imposing no-arbitrage is that the precision of the estimated relation is improved. The standard errors on δ_f in Table 4 are all smaller than the corresponding standard errors produced by the IV regressions. Also note that in both periods, the magnitude of the estimated relation is stronger when no-arbitrage is imposed than when it is not imposed. Below I discuss what features of the data contribute to this pattern.

The second conclusion is that short rates and expected inflation move almost one-for-one. A comparison of the estimated δ_f vectors with the estimated AR(3) parameters reveals they are almost identical in the early period. The correspondence is not quite as close in the late period, but the hypothesis that $\delta_f = \rho$ cannot be rejected in either period. This conclusion is surprising, since earlier research such as Clarida et al. (2000), Rudebusch (2002), and Goto and Torous (2003) documented that short rates have been much more sensitive to inflation rates in the post-deflationary period than prior to Volcker's tenure. Below I argue that the difference between my results and earlier evidence is driven by the longer sample period used here.

The third conclusion is that there is a modest relation between inflation and bond risk premia. The estimates of $\lambda_{ff(1)}$ are positive, implying that higher inflation corresponds to lower bond risk premia.⁹ The estimate is statistically different from zero in the early sample but not in the late sample. To get a sense of the magnitude of the reported coefficients, consider the standard deviation of expected excess quarterly log returns to a five-year bond. The standard deviation implied by the model can be computed with a combination of the formula for expected excess returns (52) and the sample variance of the inflation state vector f_t . For the early sample, the implied standard deviation is 13.3 basis points, or 53 basis points on an annual basis. For the late sample, the implied standard deviation is only 5 basis points on an annual basis.

What features of the data drive the high estimated sensitivity of the short rate and the low sensitivity of risk premia? To explore this question, I take a closer look at the behavior of bond yields during 1984Q1 through 2003Q4. Table 5 reports estimates of the relation

⁹The log price loadings on inflation, $B_{f,\tau}$, are negative (higher inflation implies lower prices). From (52), positive $\lambda_{ff(1)}$ implies a negative relation between expected excess log bond returns and inflation.

between one-year and five-year bond yields and f_t :

$$y_{t,\tau} = b_{0,\tau} + b'_{\tau}f_t + e_t. \tag{70}$$

The vector b_{τ} is calculated with three techniques. First, I difference (70) and estimate it with instrumental variables, paralleling the estimation of (63). Second, I use the IV estimate of (63) from Table 2, the estimated AR(3) of inflation from Table 3, and the assumption that risk premia are invariant to f_t . The vector b_{τ} is then given by no-arbitrage, but I do not impose the requirement that the computed vector for the one-year yield is consistent with the vector for the five-year yield. Third, I use the parameter estimates of the no-arbitrage model reported in Table 4 to compute b_{τ} . No standard errors are reported in Table 5 because I am interested in the magnitudes of the estimates rather than their precision.

Intuitively, estimation of the no-arbitrage model with GMM produces loadings of yields on f_t that are as close as possible to the IV estimates of these factor loadings, subject to the requirement of no-arbitrage. A comparison of the first row of Table 5 with the second reveals that the one-year yield is more sensitive to f_t than is implied by the IV estimates of short-rate dynamics and constant risk premia. In fact, these IV estimates are larger than the corresponding IV estimates for the short rate reported in Table 2. To fit the IV estimates for the one-year yield, either the short rate needs to be more responsive to inflation or risk premia need to be high when inflation is high.

If we attempt to reconcile the IV estimates for the short rate and the one-year yield simply by adjusting the risk premia, no-arbitrage requires that the loadings for the five-year yield exceed the loadings for the one-year yield. (In other words, inflation must be nonstationary under the equivalent-martingale measure.) A comparison of the first and fourth rows of Table 5 reveals that this is counterfactual. Therefore GMM estimation picks short-rate loadings δ_f that exceed the corresponding IV estimates, trading off fitting the short rate with fitting the longer-maturity yields. The model-implied loadings for the one-year and five-year yields (the table's third and sixth rows) are fairly close to the IV-estimated loadings, although the coefficients on contemporaneous inflation are too high and the coefficients on lagged inflation are too low. These loadings are produced with a value of λ_{ff} close to zero. If risk premia increased when inflation increased (negative λ_{ff}), the loadings on inflation would be larger. This would produce a better fit for the loadings on lagged inflation but a worse fit for the loadings on contemporaneous inflation.

As mentioned above, much research documents the high sensitivity of interest rates to inflation in the Volcker and Greenspan tenures. The results here do not support this result. The reason is that I have more recent data at my disposal. In Table 6, I report estimation

results for the post-deflationary period, with different ending points. The ending point of 1996Q4 is the same used in Clarida et al. (2000). Consistent with their evidence, the no-arbitrage results for this sample implies a very high sensitivity of the short rate to inflation. The sum of the coefficients on lags zero through two of the short rate exceeds two. Adding five years of data (an ending point of 2001Q4) does not substantially affect these results. However, including data for 2002 dramatically changes the results. With this sample, the estimated loadings on inflation are economically small and statistically indistinguishable from zero.

Figure 1 helps to explain these results. Panel A is constructed using the parameter estimates from the no-arbitrage model estimated over the 1984Q1 through 2001Q4 period. It plots one-quarter-ahead forecasts of changes in the five-year yield.¹⁰ The last two years are out-of-sample forecasts in the sense that the model is estimated without these data, although the forecast formed at quarter $t - 1$ uses inflation data through quarter $t - 1$. Panel B shows the corresponding realization of the change in the five-year yield. (Note that the scales of the two figures do not correspond; realizations are much more volatile than forecasts.)

Inflation was very low during early 2002. Therefore the AR model of inflation forecasted rising inflation in late 2002, and correspondingly rising bond yields. In Panel A, two of the largest predicted changes in the five year bond yield are the predictions formed in 2002Q1 and 2002Q2 for changes in 2002Q2 and 2002Q3, respectively. But bond yields fell substantially during 2002. In, fact, the largest decline in the five-year yield during the entire sample period occurs between 2002Q2 and 2002Q3. Thus the forecasts are spectacularly wrong in 2002. The forecast accuracy improves in 2003, which is why estimation over the entire period finds a statistically strong relation between inflation and bond yields.

3.6 Does the evidence support the model?

In both the pre-Volcker and post-disinflation periods, the formal tests of the overidentifying restrictions do not come close to rejecting the model. Yet in a broader sense, these results reinforce existing evidence that a single regime is an unsatisfactory description of the joint dynamics of inflation and the term structure. Estimation of the model over the entire period 1960Q1 through 2003Q4 produces inflation factor loadings δ_f that are negative, much like the IV estimates reported in Table 2 for the entire sample. (The full-sample results of the no-arbitrage model are not reported in any table.) As previously discussed in the context of these IV estimates, the problem with the full sample is that inflation dynamics have varied substantially over time.

¹⁰The predicted changes are consistently negative because the model is fitting the general fall in interest rates during the sample period.

A more general no-arbitrage model needs to incorporate regime changes in inflation dynamics. Unfortunately, tractable bond pricing in a regime-switching framework requires a number of restrictions on the nature of the regime switching; not all of the components of the dynamics are allowed to switch regimes. The requirement of tractability leads to a variety of nonnested regime-switching models. For example, the model of Ang and Bekaert (2003) cannot accommodate changing factor dynamics, while the model of Dai, Singleton, and Yang (2003) allows for changing dynamics only by imposing tight restrictions on the dynamics of the price of interest rate risk.¹¹ The results here suggest that a relatively simple regime-switching framework can accurately fit these data. There is no need to allow for regime changes in the compensation investors require to face inflation risk. In addition, the short rate's sensitivity to one-step-ahead forecasted inflation can be constant across regimes. The only component that must shift regimes is the AR process followed by inflation. Whether these simple dynamics are consistent with a tractable bond-pricing framework is an open question.

4 Concluding comments

This paper makes two contributions to the term structure literature. The first contribution is a methodological framework to investigate the relation between the term structure and other observable variables. The framework imposes no-arbitrage without requiring the estimation of the complete description of the term structure's dynamics. Therefore it can be used to describe the dynamics of expected returns to bonds conditional on the observable variables. The framework is simple to implement with GMM because it is essentially a set of regressions that are estimated with either OLS or instrumental variables. Cross-equation restrictions implied by no-arbitrage allow us to infer the parameters of the model from these regressions.

The second contribution is to use this framework to examine the relation between inflation and the term structure. The results suggest a simple description of this relation: short-term interest rates move in tandem with expected inflation, and risk premia are largely unaffected by inflation. Nonetheless, the relation between inflation and the term structure is unstable over time because the dynamics of inflation (used to determine expected inflation) are unstable. Hence these results add to the already large body of evidence pointing to the importance of modeling regime shifts in interest rate dynamics.

¹¹Ang and Bekaert (2003) discuss the modeling advantages and disadvantages of regime-switching factor dynamics.

Appendix

This appendix contains a formal dynamic model of correlated macro and latent factors. The latent factors affect the dynamics of the macro factors, while the expectation of the latent factors conditioned on macro factors satisfies (30) in the text. The framework presented here is not the only way in which to introduce correlations between f_t and x_t while satisfying (30), but it is nonetheless fairly general.

There are two types of latent factors in this model. The first type, which I label $x_{0,t}$, creates variation in short-term interest rates that is independent of the macro factors, as in Section 2.1. The second type of latent factor affects both short-term interest rates and future realizations of the macro factors. The dynamics of the first type, $x_{0,t}$ are simple to express because they do not depend on other factors. Formally,

$$x_{0,t+1} - x_{0,t} = \mu_{x0} - K_{x0}x_{0,t} + \Sigma_{x0}S_{x0t}\epsilon_{x0,t+1}, \quad (71)$$

where S_{x0t} is a diagonal matrix that depends on $x_{0,t}$.

The joint dynamics of the macro factors and the second type of latent factors are somewhat more complicated than those of $x_{0,t}$. At time t , investors observe some signals that contain information about future realizations of the macro variables. Some signals will show up quickly in the macro variables; others will show up only after a considerable lag. Formally, I assume that investors observe a vector of shocks $\epsilon_{x,i,t}$, $i = 1, \dots, d$ at time t . (For ease of discussion I treat the individual shock $\epsilon_{x,i,t}$ as a scalar, but treating it as a vector introduces no complications other than those of notation.) These shocks are independent standard normal variables conditioned on investors' information at $t-1$. Shock $\epsilon_{x,i,t}$ is news about the realization of f_{t+i}^0 .

Stack lags zero through $i-1$ of the shock $\epsilon_{x,i,t}$ into the vector $x_{i,t}$:

$$x_{i,t} = \left(\begin{array}{cccc} \epsilon_{x,i,t} & \epsilon_{x,i,t-1} & \dots & \epsilon_{x,i,t-(i-1)} \end{array} \right)', \quad i = 1, \dots, d. \quad (72)$$

The dynamics of $x_{i,t}$ are, in first-order companion form,

$$x_{i,t+1} - x_{i,t} = - \left(\begin{array}{cccc} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & \dots & 0 & 0 \\ & & \dots & & \\ 0 & 0 & \dots & -1 & 1 \end{array} \right) x_{i,t} + \left(\begin{array}{c} \epsilon_{x,i,t+1} \\ 0 \end{array} \right). \quad (73)$$

Equation (73) simply reflects the definition of $x_{i,t}$ and the fact that $\epsilon_{x,i,t+1}$ is a shock.

The entire set of latent factors is

$$x_t = \begin{pmatrix} x'_{0,t} & x'_{1,t} & \dots & x'_{d,t} \end{pmatrix}'. \quad (74)$$

Recall that r_t is affine in f_t and x_t . Therefore all of the shocks $\epsilon_{x,i,t-j}, j < i$ are allowed to affect r_t directly. Put differently, the short rate can react to the information observed by investors before it is incorporated into the macro variables.

At time $t+i$, the macro variables f_{t+i}^0 react to $\epsilon_{x,i,t}$. The key restriction built into their relation is that after $t+1$, the shock has no direct effect on the macro variables. It only affects these macro variables indirectly, through the persistence of the macro variables themselves. The macro dynamics satisfy (31) in the text, where the matrix K_{fx} is

$$K_{fx} = \begin{pmatrix} K_{fx0} & K_{fx1} & \dots & K_{fxd} \\ 0 & 0 & \dots & 0 \end{pmatrix}, \quad (75)$$

$$K_{fx0} = 0, \quad K_{fxi} = \begin{pmatrix} 0 & k_i \end{pmatrix}. \quad (76)$$

The submatrices of zeros in the second row of K_{fx} are a consequence of the first-order companion form of (31). The matrix K_{fx0} is zero because the latent factors $x_{0,t}$ are independent of the macro factors. The submatrix of zeros in K_{fxi} is $n_0 \times (i-1)$ and k_i is a vector of length n_0 . This structure implies that the shock $\epsilon_{x,i,t}$ does not affect the macro factors until $t+i$, at which point its effect is determined by the elements of k_i .

It is easy to verify that these dynamics satisfy (30). The key intuition is that the vector $x_{i,t}$ contains shocks that show up in the macro factors at $t+1, \dots, t+i$. Thus it is independent of $f_{t-j}, j \geq 0$.

References

- Ang, Andrew, and Geert Bekaert, 2003, The term structure of real rates and expected inflation, Working paper, Columbia Business School.
- Ang, Andrew, and Monika Piazzesi, 2003, A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, *Journal of Monetary Economics* 50, 745-787.
- Ang, Andrew, Monika Piazzesi, and Min Wei, 2003, What does the yield curve tell us about GDP Growth?, Working paper, Columbia Business School.
- Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press, Princeton, 1997.
- Campbell, John Y., and Robert J. Shiller, 1991, Yield spreads and interest rate movements: A bird's eye view, *Review of Economic Studies* 58, 495-514.
- Clarida, Richard, Jordi Galí, and Mark Gertler, 2000, Monetary policy rules and macroeconomic stability: evidence and some theory, *Quarterly Journal of Economics* 115, 147-180.
- Cox, John C., Jonathan E. Ingersoll, Jr., and Stephen A. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385-407.
- Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2003, Regime shifts in a dynamic term structure model of U.S. Treasury bond yields, Working paper, Stern School of Business.
- Dewachter, Hans, and Marco Lyrio, 2002, Macro factors and the term structure of interest rates, Working paper, Catholic University of Leuven.
- Dewachter, Hans, Marco Lyrio, and Konstantijn Maes, 2002, Estimation of a joint model for the term structure of interest rates and the macroeconomy, Working paper, Catholic University of Leuven.
- Diebold, Francis X., Glenn D. Rudebusch, and S. Boragan Aruoba, 2003, The macroeconomy and the yield curve: A nonstructural analysis, Working paper, The Wharton School.
- Duffee, Gregory R., 2002, Term premia and interest rate forecasts in affine models, *Journal of Finance* 57, 405-443.
- Duffie, Darrell, and Rui Kan, 1996, A yield-factor model of interest rates, *Mathematical Finance* 6, 379-406.

- English, William B., William R. Nelson, and Brian P. Sack, 2003, Interpreting the significance of the lagged interest rate in estimated monetary policy rules, *Contributions to Macroeconomics* 3, Article 5.
- Fama, Eugene F., 1975, Short-term interest rates as predictors of inflation, *American Economic Review* 65, 269-282.
- Goto, Shingo, and Walter Torous, 2003, The conquest of U.S. inflation: Its implications for the Fisher hypothesis and the term structure of nominal interest rates, Working paper, University of South Carolina.
- Gray, Stephen F., 1996, Modeling the conditional distribution of interest rates as a regime-switching process, *Journal of Financial Economics* 42, 27-62.
- Hansen, Lars Peter, 1982, Large sample properties of generalized method of moment estimators, *Econometrica* 50, 1029-1053.
- Hansen, Lars Peter, and Kenneth J. Singleton, 1982, Generalized instrumental variables estimators of nonlinear rational expectations models, *Econometrica* 50, 1269-1286.
- Hördahl, Peter, Oreste Tristani, and David Vestin, 2002, A joint econometric model of macroeconomic and term structure dynamics, Working paper, European Central Bank.
- Kimmel, Robert L., Patrick Cheridito, and Damir Filipovic, 2004, Market price of risk specification for affine models: Theory and evidence, Working paper, Princeton University.
- Newey, Whitney K., and Kenneth D. West, 1987a, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703-708.
- Newey, Whitney K., and Kenneth D. West, 1987b, Hypothesis testing with efficient method of moments estimation, *International Economic Review* 28, 777-787.
- Piazzesi, Monika, 2003, Bond yields and the Federal Reserve, *Journal of Political Economy*, forthcoming.
- Rudebusch, Glenn D., 2002, Term structure evidence on interest rate smoothing and monetary policy inertia, *Journal of Monetary Economics* 49, 1161-1187.
- Rudebusch, Glenn D., and Tao Wu, 2003, A macro-finance model of the term structure, monetary policy, and the economy, Working paper, Federal Reserve Bank of San Francisco.

- Sack, Brian, and Volker Wieland, 2000, Interest-rate smoothing and optimal monetary policy: A review of recent empirical evidence, *Journal of Economics and Business* 52, 205-228.
- Taylor, John B., 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- Vasicek, Oldrich A., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177-188.

Variable	_ Levels _		Differences		Contemporaneous corr of differences		
	SD	ρ	SD	ρ	Inflation	3-mon yield	1-yr yield
1960Q1–1979Q2							
Inflation	2.80	0.90	1.23	-0.16	1.00		
3-mon yield	1.83	0.93	0.69	-0.02	0.26	1.00	
1-yr yield	1.83	0.92	0.71	-0.04	0.24	0.80	1.00
5-yr yield	1.63	0.95	0.50	-0.12	0.16	0.65	0.88
1979Q3–1983Q4							
Inflation	2.79	0.82	1.63	-0.39	1.00		
3-mon yield	2.66	0.39	2.99	-0.36	0.00	1.00	
1-yr yield	2.26	0.41	2.51	-0.42	0.16	0.93	1.00
5-yr yield	1.66	0.61	1.45	-0.28	0.19	0.82	0.94
1984Q1–2003Q4							
Inflation	1.26	0.62	1.10	-0.37	1.00		
3-mon yield	2.20	0.96	0.59	0.13	0.13	1.00	
1-yr yield	2.34	0.96	0.67	0.05	0.07	0.86	1.00
5-yr yield	2.17	0.95	0.66	-0.01	0.06	0.59	0.85

Table 1: Summary statistics for quarterly observations of inflation and Treasury bond yields

Inflation is the log change in the PCE chain-weighted price index. Zero-coupon Treasury yields are from CRSP. All data are continuously compounded and expressed in percent per year. Standard deviations are denoted SD and first-order autocorrelation coefficients are denoted ρ . “Differences” refers to quarterly changes.

Sample Period	Lag of change in inflation			Corr between residual and lead of change in inflation		
	0	1	2	0	1	2
1960Q1–2003Q4	−0.125 (0.543)	−0.108 (0.224)	0.117 (0.142)	0.24	0.07	0.00
1960Q1–1979Q2	0.592 (0.230)	0.015 (0.119)	0.109 (0.090)	−0.62	0.29	0.08
1979Q3–1983Q4	0.340 (0.611)	−0.259 (0.500)	0.677 (0.315)	−0.29	0.31	−0.15
1984Q1–2003Q4	0.436 (0.353)	0.287 (0.183)	0.229 (0.141)	−0.37	0.32	0.07

Table 2: Instrumental variable regressions of changes in three-month yields on changes in inflation

Quarterly changes in the three-month Treasury bill yield are regressed on lags zero through two of changes in inflation. No constant term is included. Yields are from CRSP and inflation is the log change in the PCE chain-weighted price index. Both are expressed as annual rates. The regressions are estimated with instrumental variables, where the instruments are a constant and lags one through three of quarterly inflation. Standard errors are in parentheses. They are adjusted for generalized heteroskedasticity and four lags of moving average residuals using the technique of Newey and West. The final three columns report sample correlations between fitted residuals and leads zero through two of changes in inflation. After accounting for lags, there are 173 observations for 1960Q1–2003Q4, 75 for 1960Q1–1979Q2, 15 for 1979Q3–1983Q4, and 77 for 1984Q1–2003Q4.

Sample period	Lag			SEE
	1	2	3	
1960Q1–2003Q4	0.638 (0.076)	0.115 (0.079)	0.178 (0.078)	1.147
1960Q1–1979Q2	0.804 (0.123)	0.066 (0.143)	0.078 (0.128)	1.233
1979Q3–1983Q4	0.388 (0.166)	0.058 (0.158)	0.436 (0.169)	1.342
1984Q1–2003Q4	0.425 (0.097)	0.100 (0.129)	0.265 (0.148)	0.959

Table 3: An AR(3) description of quarterly inflation

Inflation is measured by the quarterly change in the PCE chain-weighted price index. Estimation of the AR(3) is with OLS. Standard errors, adjusted for generalized heteroskedasticity, are in parentheses. The column labeled “SEE” reports the standard error of the estimate. After accounting for lags, there are 173 observations for 1960Q1–2003Q4, 75 for 1960Q1–1979Q2, 15 for 1979Q3–1985, and 77 for 1984Q1–2003Q4.

Panel A. Parameter estimates

Sample period	Loading of short rate — on inflation lag i : —			Coef i of — AR(p) for inflation —			Price of risk loading on inflation
	0	1	2	1	2	3	
1960Q1–1979Q2	0.762 (0.202)	0.186 (0.106)	0.076 (0.065)	0.760 (0.083)	0.170 (0.135)	0.086 (0.084)	0.116 (0.038)
1979Q3–1983Q4	0.499 (0.400)	-0.159 (0.291)	-	0.485 (0.069)	0.408 (0.054)	-	0.092 (0.698)
1984Q1–2003Q4	0.590 (0.275)	0.319 (0.149)	0.291 (0.113)	0.458 (0.075)	0.102 (0.093)	0.364 (0.073)	0.026 (0.095)

Panel B. Hypothesis tests

Sample period	Overidentifying moments	Equality of coefficients	Additional lag in price of risk
1960Q1–1979Q2	4.666 [0.793]	0.079 [0.994]	0.479 [0.489]
1979Q3–1983Q4	3.117 [0.794]	14.254 [0.001]	0.652 [0.420]
1984Q1–2003Q4	6.273 [0.617]	3.112 [0.375]	0.778 [0.378]

Table 4: Estimates from a no-arbitrage term structure model

The short rate is given by $r_t = \delta_0 + \delta'_f f_t + \omega_t$, where the vector f_t contains lags zero through two of quarterly inflation. Quarterly inflation follows an AR(3) process. Under the equivalent martingale measure, the first coefficient of this AR(3) equals the physical measure coefficient less the price of risk adjustment. Estimation is with GMM, using quarterly observations of inflation and yields on zero-coupon bonds with maturities of three months, one year, and five years. Standard errors are in parentheses. They are adjusted for generalized heteroskedasticity and four lags of moving-average residuals. The test of overidentifying moments is Hansen's J test. The test of equality of coefficients is an LR test of the hypothesis that δ_f equals the AR(3) coefficients. The test of an additional lag in the price of risk is an LM test that bond risk premia depend on the first lag of inflation in addition to current inflation. P -values of the test statistics are in square brackets.

Maturity	Method	Loading of the yield on inflation lag i :		
		0	1	2
One year	IV	0.460	0.321	0.258
One year	short rate/constant premia	0.437	0.224	0.149
One year	model	0.574	0.303	0.233
Five years	IV	0.379	0.216	0.257
Five years	short rate/constant premia	0.208	0.100	0.066
Five years	model	0.392	0.204	0.153

Table 5: Loadings of longer-term bond yields on current and lagged inflation, 1984Q1 through 2003Q4

The yield on a τ -maturity bond is expressed as $y_{t,\tau} = b_0 + b_1\pi_t + b_2\pi_{t-1} + b_3\pi_{t-2} + e_{t,\tau}$, where π_t is inflation during quarter t . Estimated coefficients are produced using three methods. With “IV,” the equation is first-differenced and estimated over 1984Q1 through 2003Q4 with instrumental variables. With “short rate/constant premia,” the coefficients are calculated using (a) the estimate of the corresponding expression for the short rate, (b) the estimate of the AR(3) dynamics of inflation, and (c) the assumption that risk premia are invariant to inflation. With “model,” the coefficients are calculated using a no-arbitrage model estimated over 1984Q1 through 2003Q4.

Sample period	Loading of short rate — on inflation lag i : —			Coef i of — AR(p) for inflation —			Price of risk loading on inflation
	0	1	2	1	2	3	
1984Q1–1996Q4	0.923 (0.235)	0.584 (0.131)	0.561 (0.130)	0.382 (0.057)	0.026 (0.103)	0.473 (0.083)	0.012 (0.079)
1984Q1–2001Q4	0.861 (0.256)	0.462 (0.125)	0.484 (0.122)	0.519 (0.074)	-0.018 (0.092)	0.477 (0.074)	0.053 (0.073)
1984Q1–2002Q4	0.157 (0.308)	0.090 (0.132)	0.139 (0.121)	0.544 (0.070)	-0.022 (0.084)	0.363 (0.104)	0.224 (0.356)

Table 6: Estimates from a no-arbitrage term structure model: Sample sensitivity

The short rate is given by $r_t = \delta_0 + \delta'_f f_t + \omega_t$, where the vector f_t contains lags zero through two of quarterly inflation. Quarterly inflation follows an AR(3) process. Under the equivalent martingale measure, the first coefficient of this AR(3) equals the physical measure coefficient less the price of risk adjustment. Estimation is with GMM, using quarterly observations of inflation and yields on zero-coupon bonds with maturities of three months, one year, and five years. Standard errors are in parentheses. They are adjusted for generalized heteroskedasticity and four lags of moving-average residuals.

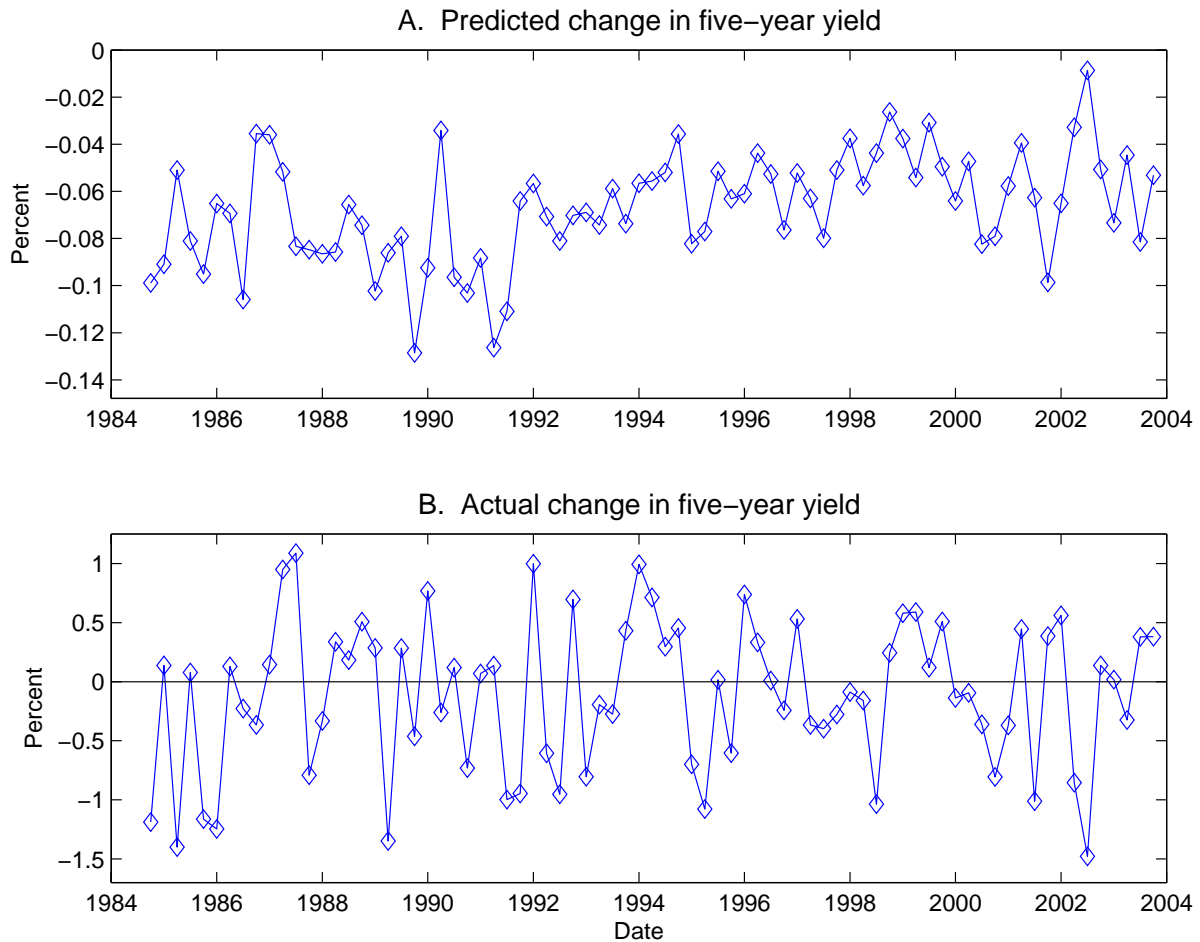


Figure 1: A comparison of forecasted and actual quarterly changes in bond yields

At the end of quarter $t - 1$, the change in the five-year bond yield from $t - 1$ to t is predicted using a no-arbitrage term structure model. The model is estimated using data through 2001Q4, while the one-quarter-ahead forecasts (plotted in Panel A) are constructed through 2003Q3. Panel B plots realized changes in yields. The plots are aligned so that the forecast made at $t - 1$ of the quarter t change corresponds to the quarter t realization of this change.