THE JOURNAL OF FINANCE • VOL. LXIII, NO. 2 • APRIL 2008

The Term Structure of Real Rates and Expected Inflation

ANDREW ANG, GEERT BEKAERT, and MIN WEI*

ABSTRACT

Changes in nominal interest rates must be due to either movements in real interest rates, expected inflation, or the inflation risk premium. We develop a term structure model with regime switches, time-varying prices of risk, and inflation to identify these components of the nominal yield curve. We find that the unconditional real rate curve in the United States is fairly flat around 1.3%. In one real rate regime, the real term structure is steeply downward sloping. An inflation risk premium that increases with maturity fully accounts for the generally upward sloping nominal term structure.

THE REAL INTEREST RATE AND EXPECTED INFLATION are two key economic variables; yet, their dynamic behavior is essentially unobserved. A large empirical literature has yielded surprisingly few generally accepted stylized facts. For example, while theoretical research often assumes that the real interest rate is constant, empirical estimates for the real interest rate process vary between constancy as in Fama (1975), mean-reverting behavior (Hamilton (1985)), or a unit root process (Rose (1988)). There seems to be more consensus on the fact that real rate variation, if it exists at all, should only affect the short end of the term structure whereas the variation in long-term interest rates is primarily affected by shocks to expected inflation (see, among others, Fama (1990) and Mishkin

*Ang is with Columbia University and NBER. Bekaert is with Columbia University, CEPR and NBER. Wei is with the Federal Reserve Board of Governors. We thank Kobi Boudoukh, Qiang Dai, Rob Engle, Martin Evans, Rene Garcia, Bob Hodrick, Refet Gürkaynak, Monika Piazzesi, Bill Schwert, Ken Singleton, Peter Vlaar, Ken West, and Mungo Wilson for helpful discussions, and seminar participants at the American Finance Association, Asian Finance Association, Barclays Capital Annual Global Inflation-Linked Conference, CIREQ and CIRANO-MITACS conference on Macroeconomics and Finance, Empirical Finance Conference at the LSE, European Finance Association, FRBSF-Stanford University conference on Interest Rates and Monetary Policy, HKUST Finance Symposium, Washington University-St. Louis Federal Reserve conference on State-Space Models, Regime-Switching and Identification, Bank of England, Bank of Norway, Campbell and Company, University of Amsterdam, Columbia University, Cornell University, Erasmus University, European Central Bank, Federal Reserve Bank of Kansas, Federal Reserve Board of Governors, Financial Engines, HEC Lausanne, Indiana University, IMF, London Business School, National University of Singapore, NYU, Oakhill Platinum Partners, PIMCO, Singapore Management University, Tilburg University, UCL-CORE at Louvain-la-Neuve, University of Gent, University of Illinois, University of Michigan, University of Rochester, University of Washington, UCLA, UC Riverside, UC San Diego, USC, and the World Bank. Andrew Ang and Geert Bekaert both acknowledge funding from the National Science Foundation. Additional results and further technical details are available in the NBER working paper version of this article.

(1990)), although this is disputed by Pennacchi (1991). Another phenomenon that has received wide attention is the Mundell (1963) and Tobin (1965) effect: The correlation between real rates and (expected) inflation appears to be negative.

In this article, we seek to establish a comprehensive set of stylized facts regarding real rates, expected inflation, and inflation risk premiums, and to determine their relative importance for determining the U.S. nominal term structure. To infer the behavior of these variables, we use a model with three distinguishing features. First, we specify a no-arbitrage term structure model with both nominal bond yields and inflation data to efficiently identify the term structure of real rates and inflation risk premia. Second, our model accommodates regime-switching (RS) behavior, but still produces closed-form solutions for bond prices. We go beyond the extant RS literature by attempting to identify the real and nominal sources of the regime switches. Third, the model accommodates flexible time-varying risk premiums crucial for matching time-varying bond premia (see, for example, Dai and Singleton (2002)). These features allow our model to fit the dynamics of inflation and nominal interest rates.

This paper is organized as follows. Section I develops the model and discusses the effect of regime switches on real yields and inflation risk premia. In Section II, we detail the specification tests used to select the best model, analyze factor dynamics, and report parameter estimates. Section III contains the main economic results, which can be summarized as follows:

1. Unconditionally, the term structure of real rates assumes a fairly flat shape around 1.3%, with a slight hump, peaking at a 1-year maturity. However, there are some regimes in which the real rate curve is downward sloping.
2. Real rates are quite variable at short maturities but smooth and persistent at long maturities. There is no significant real term spread.
3. The real short rate is negatively correlated with both expected and unexpected inflation, but the statistical evidence for a Mundell–Tobin effect is weak.
4. The model matches an unconditional upward-sloping nominal yield curve by generating an inflation risk premium that is increasing in maturity.
5. Nominal interest rates do not behave procyclically across NBER business cycles but our model-implied real rates do.
6. The decompositions of nominal yields into real yields and inflation components at various horizons indicate that variation in inflation compensation (expected inflation and inflation risk premia) explains about 80% of the variation in nominal rates at both short and long maturities.
7. Inflation compensation is the main determinant of nominal interest rate spreads at long horizons.

Finally, Section IV concludes.

I. A Real and Nominal Term Structure Model with Regime Switches

A. Decomposing Nominal Yields

The nominal yield on a zero-coupon bond of maturity n, y_t^n, can be decomposed into a real yield, $\hat{y}_t^n$, and inflation compensation, $\pi_{t,n}^e$. The real yield represents the yield on a zero-coupon bond perfectly indexed against inflation. Inflation compensation reflects expected inflation, $\mathrm{E}_t(\pi_{t+n,n})$, and an inflation risk premium, $\varphi_{t,n}$ (ignoring Jensen's inequality terms):

$$\begin{aligned} y_t^n &= \hat{y}_t^n + \pi_{t,n}^e \\ &= \hat{y}_t^n + \mathrm{E}_t(\pi_{t+n,n}) + \varphi_{t,n}, \end{aligned} \tag{1}$$

where $\mathrm{E}_t(\pi_{t+n,n})$ is expected inflation from t to $t+n$, that is,

$$\mathrm{E}_t(\pi_{t+n,n}) = \frac{1}{n}\mathrm{E}_t(\pi_{t+1} + \cdots + \pi_{t+n}),$$

and π_{t+1} is one-period inflation from t to $t+1$.

The goal of this article is to achieve this decomposition of nominal yields, y_t^n, into real and inflation components ($\hat{y}_t^n$, $\mathrm{E}_t(\pi_{t+n,n})$, and $\varphi_{t,n}$) for U.S. data. Unfortunately, we do not observe real rates for most of the U.S. sample. Inflation-indexed bonds (the Treasury Income Protection Securities or TIPS) have traded only since 1997 and the market faced considerable liquidity problems in its early days (see Roll (2004)). Consequently, our endeavor faces an identification problem as we must estimate two unknown quantities—real rates and inflation risk premia—from only nominal yields. We obtain identification by using a no-arbitrage term structure model that imposes restrictions on the nominal yields. That is, the movements of long-term yields are linked to the dynamics of both short-term yields and inflation. These pricing restrictions uniquely identify the dynamics of real rates and inflation risk premiums using data on inflation and nominal yields. To pin down the average level of real rates, we further restrict the one-period inflation risk premium to be zero.

The remainder of this section sets up the model to identify the various components of nominal yields. Section I.B presents the term structure model and discusses the economic background of our factors and parametric assumptions. Importantly, both the empirical literature and economic logic suggest that the process generating inflation and real rates may undergo discrete shifts over time, which we model using an RS model following Hamilton (1989). We present solutions to bond prices in Section I.C and discuss how regime switches affect our decomposition in Section I.D. Section I.E briefly covers econometric and identification issues. Finally, Section I.F discusses how our work relates to the literature.

B. The Model

B.1. State Variable Dynamics

We employ a three-factor representation of yields, which is the number of factors often used to match term structure dynamics in the finance literature (see, for example, Dai and Singleton (2000)). We incorporate an observed inflation factor, denoted by π_t, which switches regimes. The other two factors are unobservable term structure factors. One factor, f_t, represents a latent RS term structure factor. The other latent factor is denoted by q_t and represents a time-varying but regime-invariant price of risk factor, which directly enters into the risk prices (see below). The factor q_t plays two roles. First, it helps time-varying expected excess returns on long-term bonds, as demonstrated by Dai and Singleton (2002).[1] Second, q_t also accounts for part of the time variation of inflation risk premia, as we show below.

We stack the state variables in the 3×1 vector $X_t = (q_t\ f_t\ \pi_t)'$, which follows

$$X_{t+1} = \mu(s_{t+1}) + \Phi X_t + \Sigma(s_{t+1})\varepsilon_{t+1}, \tag{2}$$

where s_{t+1} indicates the regime prevailing at time $t+1$ and

$$\mu(s_t) = \begin{bmatrix} \mu_q \\ \mu_f(s_t) \\ \mu_\pi(s_t) \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_{qq} & 0 & 0 \\ \Phi_{fq} & \Phi_{ff} & 0 \\ \Phi_{\pi q} & \Phi_{\pi f} & \Phi_{\pi\pi} \end{bmatrix}, \quad \Sigma(s_t) = \begin{bmatrix} \sigma_q & 0 & 0 \\ 0 & \sigma_f(s_t) & 0 \\ 0 & 0 & \sigma_\pi(s_t) \end{bmatrix}. \tag{3}$$

The regime variable represents K different regimes, $s_t = 1, \ldots, K$, and follows a Markov chain with a constant transition probability matrix $\Pi = \{p_{ij} = Pr(s_{t+1} = j \mid s_t = i)\}$. These regimes are independent of the shocks ε_{t+1} in equation (2).

In equation (3), the conditional mean and volatility of f_t and π_t switch regimes, but the conditional mean and volatility of q_t do not. The feedback parameters for all variables in the companion form Φ also do not switch across regimes. These restrictions are necessary to permit closed-form solutions for bond prices.

We order the factors so that the latent factors appear first. As a consequence, expected inflation depends on lagged inflation, other information captured by the latent variables, as well as a nonlinear drift term. The inflation forecasting literature strongly suggests that expected inflation depends on more than just lagged inflation (see, for example, Stockton and Glassman (1987)). In addition, by placing inflation last in the system, the reduced-form process for inflation involves moving average terms. The autocorrelogram of inflation in data is well approximated by a low order ARMA process.

B.2. Real Short Rate Dynamics

We specify the real short rate, $\hat{r}_t$, to be affine in the state variables:

$$\hat{r}_t = \delta_0 + \delta_1' X_t. \tag{4}$$

[1] Fama and Bliss (1987), Campbell and Shiller (1991), Bekaert, Hodrick, and Marshall (1997), and Cochrane and Piazzesi (2005), among many others, document time variation in expected excess holding period returns of long-term bonds.

For reference, we let $\delta_1 = (\delta_q\, \delta_f\, \delta_\pi)'$. The real rate process nests the special cases of a constant real rate ($\delta_1 = 0_{3\times 1}$), advocated by Fama (1975), and mean-reverting real rates within a single regime ($\delta_f = \delta_\pi = 0$), following Hamilton (1985). Allowing nonzero δ_f or δ_π causes the real rate to switch regimes. If $\delta_q \neq 0$, then the time-varying price of risk can directly influence the real rate, as it would in any equilibrium model with growth. In general, if $\delta_\pi \neq 0$, then money neutrality is rejected and real interest rates are functions of inflation.

The model allows for arbitrary correlation between the real rate and inflation. To gain some intuition, we compute the conditional covariance between real rates and actual or expected inflation for an affine model without regime switches. First, δ_π primarily drives the covariance between real rates and unexpected inflation. That is, $\text{cov}_t(\hat{r}_{t+1}, \pi_{t+1}) = \delta_\pi \sigma_\pi^2$. Second, without regimes, the covariance between expected inflation and real rates is given by

$$\text{cov}_t(\hat{r}_{t+1}, \text{E}_{t+1}(\pi_{t+2})) = \delta_q \Phi_{\pi q} \sigma_q^2 + \delta_f \Phi_{\pi f} \sigma_f^2 + \delta_\pi \Phi_{\pi\pi} \sigma_\pi^2.$$

The Mundell–Tobin effect predicts this covariance to be negative, whereas an activist Taylor (1993) rule would predict it to be positive, as the monetary authority raises real rates in response to high expected inflation (see, for example, Clarida, Galí, and Gertler (2000)). Clearly, the sign of the covariance is parameter dependent, and a negative δ_π does not suffice to obtain a Mundell–Tobin effect.

To compare the conditional covariance between real rates and expected inflation in our model with regimes, we derive $\text{cov}_t(\hat{r}_{t+1}, \text{E}_{t+1}(\pi_{t+2}) | s_t = i)$ for $K = 2$ regimes to be

$$\begin{aligned}
\text{cov}_t(\hat{r}_{t+1}, \text{E}_{t+1}(\pi_{t+2}) | s_t = i) = {} & \delta_q \Phi_{\pi q} \sigma_q^2 \\
& + \delta_f \Phi_{\pi f} \left[\textstyle\sum_{j=1}^2 p_{ij} \sigma_f^2(j) + p_{i1} p_{i2} (\mu_f(1) - \mu_f(2))^2 \right] \\
& + \delta_\pi \Phi_{\pi\pi} \left[\textstyle\sum_{j=1}^2 p_{ij} \sigma_\pi^2(j) + p_{i1} p_{i2} (\mu_\pi(1) - \mu_\pi(2))^2 \right] \\
& + \delta_f \delta_\pi \Phi_{\pi f} \Phi_{\pi\pi} p_{i1} p_{i2} [(\mu_\pi(1) - \mu_\pi(2))(\mu_f(1) - \mu_f(2))].
\end{aligned}$$

Relative to the one-regime model, the contribution of the factor variances for the RS factors now depends on the regime prevailing at time t and has two components namely, an average of the two regime-dependent factor variances and a term measuring the volatility impact of a change in the regime-dependent drifts. In addition, there is a new factor contributing to the covariance that comes from the covariance between these regime-dependent drifts for f_t and π_t.

B.3. Pricing Kernel and Prices of Risk

We specify the real pricing kernel to take the form

$$\widehat{m}_{t+1} = \log \widehat{M}_{t+1} = -\hat{r}_t - \frac{1}{2} \lambda_t(s_{t+1})' \lambda_t(s_{t+1}) - \lambda_t(s_{t+1})' \varepsilon_{t+1}, \tag{5}$$

where the vector of time-varying and RS prices of risk $\lambda_t(s_{t+1})$ is given by

$$\lambda_t(s_{t+1}) = (\gamma_t \quad \lambda(s_{t+1})')',$$

where $\lambda(s_{t+1})$ is a 2×1 vector of RS prices of risk $\lambda(s_{t+1}) = (\lambda_f(s_{t+1})\,\lambda_\pi(s_{t+1}))'$ and the scalar γ_t takes the form

$$\gamma_t = \gamma_0 + \gamma_1 q_t = \gamma_0 + \gamma_1 e_1' X_t, \tag{6}$$

where e_i represents a vector of zeros with a "1" in the ith position. In this formulation, the prices of risk of f_t and π_t change across regimes. The variable q_t controls the time variation of the price of risk associated with γ_t in equation (6) but does not switch regimes. Allowing γ_t to switch across regimes results in the loss of closed-form solutions for bond prices.

We formulate the nominal pricing kernel in the standard way as $M_{t+1} = \widehat{M}_{t+1} P_t / P_{t+1}$:

$$m_{t+1} = \log M_{t+1} = -\hat{r}_t - \frac{1}{2}\lambda_t(s_{t+1})'\lambda_t(s_{t+1}) - \lambda_t(s_{t+1})'\varepsilon_{t+1} - e_3' X_{t+1}. \tag{7}$$

B.4. Real Factor and Inflation Regimes

We introduce two different regime variables, $s_t^f \in \{1, 2\}$, affecting the drift and variance of the f_t process, and $s_t^\pi \in \{1, 2\}$, affecting the drift and variance of the inflation process. Since both the f_t and π_t factors enter the real short rate in equation (4), the real short rate contains both f_t and π_t regime components. This modeling choice accommodates the possibility that s_t^f captures changes of regimes in real factors. Since f_t enters the conditional mean of inflation in equation (2), the f_t regime also potentially affects expected inflation and can capture nonlinear expected inflation components not directly related to past inflation realizations.

The model with s_t^f and s_t^π can be rewritten using an aggregate regime variable $s_t \in \{1, 2, 3, 4\}$ to account for all possible combinations of $\{s_t^f, s_t^\pi\} = \{(1, 1), (1, 2), (2, 1), (2, 2)\}$. Hence, our model has $K = 4$ regimes. To reduce the number of parameters in the 4×4 transition probability matrix, we consider two restricted models of the correlation between s_t^f and s_t^π. Case A represents the simplest case of independent regimes.[2]

In an alternative case C, we specify a restricted form of the transition probability matrix so that the inflation regime at $t + 1$ depends on the stance of the f_{t+1} regime as well as the previous inflation environment, but we restrict future f_{t+1} regimes to depend only on current f_t regimes. Intuitively, this specification can capture periods in which aggressive real rates, for example, captured by a regime with high f_t, could successfully stave off a regime of high

[2] Ang, Bekaert, and Wei (2007) consider another restricted case of correlated s_t^f and s_t^π regimes. This fits the data less well than Case C presented here.

inflation. This leads to the following conditional transition probability:

$$\begin{aligned} &Pr\left(s^f_{t+1} = j, s^\pi_{t+1} = k \,|\, s^f_t = m, s^\pi_t = n\right) \\ &= Pr\left(s^\pi_{t+1} = k \,|\, s^f_{t+1} = j, s^f_t = m, s^\pi_t = n\right) \times Pr\left(s^f_{t+1} = j \,|\, s^f_t = m, s^\pi_t = n\right) \\ &= Pr\left(s^\pi_{t+1} = k \,|\, s^f_{t+1} = j, s^\pi_t = n\right) \times Pr\left(s^f_{t+1} = j \,|\, s^f_t = m\right), \end{aligned} \tag{8}$$

where we assume that $Pr(s^\pi_{t+1}|s^f{}_{t+1}, s^f_t, s^\pi_t) = Pr(s^\pi_{t+1}|s^f{}_{t+1}, s^\pi_t)$ and $Pr(s^f{}_{t+1}\,|s^f_t, s^\pi_t) = Pr(s^f{}_{t+1}\,|s^f_t)$. We denote $Pr(s^f{}_{t+1} = 1|s^f_t = 1) = p^f$ and $Pr(s^f{}_{t+1} = 2|s^f_t = 2) = q^f$ and parameterize $Pr(s^\pi_{t+1} = k\,|s^f_t = m, s^\pi_t = n)$ as $p^{``j",``m"}$, where

$$j = \begin{cases} A & \text{if } s^\pi_{t+1} = s^f_{t+1} = 1 \\ B & \text{if } s^\pi_{t+1} = s^f_{t+1} = 2. \end{cases}$$

The "j"-component captures (potentially positive) correlation between the f_t and π_t regimes. The "m"-component captures persistence in π_t regimes:

$$m = \begin{cases} A & \text{if } s^\pi_t = 1 \\ B & \text{if } s^\pi_t = 2. \end{cases}$$

This formulation can capture instances in which a high real rate regime, as captured by the high f_t regime, contemporaneously influences the inflation regime. Using the notation introduced above, the transition probability matrix Π for Case C takes the form:

$$\begin{array}{c} \\ [s_t=1] \\ [s_t=2] \\ [s_t=3] \\ [s_t=4] \end{array} \begin{array}{c} \begin{array}{cccc} [s_{t+1}=1] & [s_{t+1}=2] & [s_{t+1}=3] & [s_{t+1}=4] \end{array} \\ \begin{bmatrix} p^f p^{AA} & p^f(1-p^{AA}) & (1-p^f)(1-p^{BA}) & (1-p^f)p^{BA} \\ p^f p^{AB} & p^f(1-p^{AB}) & (1-p^f)(1-p^{BB}) & (1-p^f)p^{BB} \\ (1-q^f)p^{AA} & (1-q^f)(1-p^{AA}) & q^f(1-p^{BA}) & q^f p^{BA} \\ (1-q^f)p^{AB} & (1-q^f)(1-p^{AB}) & q^f(1-p^{BB}) & q^f p^{BB} \end{bmatrix} \end{array}$$

This model has four additional parameters relative to the model with independent real and inflation regimes. We can test Case C against the null of the independent regime Case A by testing the restrictions

$$H_0 : p^{BA} = 1 - p^{AA} \text{ and } p^{BB} = 1 - p^{AB}.$$

We find evidence to reject the case of independent regimes in favor of this case with a p-value of 0.033. Thus, our benchmark specification uses the probability transition matrix of Case C.

C. Bond Prices

Our model produces closed-form solutions for bond prices, enabling both efficient estimation and the ability to fully characterize real and nominal yields at all maturities without discretization error.

C.1. Real Bond Prices

In our model, the real zero-coupon bond price of maturity n conditional on regime $s_t = i$, $\widehat{P}_t^n(s_t = i)$, is given by

$$\widehat{P}_t^n(i) = \exp(\widehat{A}_n(i) + \widehat{B}_n X_t), \tag{9}$$

where $\widehat{A}_n(i)$ is dependent on regime $s_t = i$, $\widehat{B}_n$ is a $1 \times N$ vector, and N is the total number of factors in the model, including inflation. The expressions for $\widehat{A}_n(i)$ and $\widehat{B}_n$ are given in Appendix A. Since the real bond prices are given by (9), it follows that the real yields $\hat{y}_t^n(i)$ conditional on regime i are affine functions of X_t:

$$\hat{y}_t^n(i) = -\frac{\log(\widehat{P}_t^n)}{n} = -\frac{1}{n}(\widehat{A}_n(i) + \widehat{B}_n X_t). \tag{10}$$

While the expressions for $\widehat{A}_n(i)$ and $\widehat{B}_n$ are complex, some intuition can be gained on how the prices of risk affect each term. The prices of risk γ_0 and $\lambda(s_t)$ enter only the constant term in the yields $\widehat{A}_n(s_t)$, but affect this term in all regimes. More negative values of γ_0 or $\lambda(s_t)$ cause long maturity yields to be, on average, higher than short maturity yields. In addition, since the $\lambda(s_t)$ terms differ across regimes, $\lambda(s_t)$ also controls the regime-dependent level of the yield curve away from the unconditional shape of the yield curve. Thus, the model can accommodate the switching signs of term premiums documented by Boudoukh et al. (1999). The prices of risk affect the time variation in the yields through the parameter γ_1. This term only enters the $\widehat{B}_n$ terms. A more negative γ_1 means that long-term yields respond more to shocks in the price of risk factor q_t.

The pricing implications of (10), together with the assumed dynamics of X_t in (2), imply that the autoregressive dynamics of inflation and bond yields are constant over time, but the drifts vary through time, and shocks to inflation and real yields are heteroskedastic. Hence, our model is consistent with the macro models of Sims (1999, 2001) and Bernanke and Mihov (1998), who stress changing drifts, induced for example by changes in monetary policy, and heteroskedastic shocks. On the other hand, Cogley and Sargent (2001, 2005) advocate models with changes in the feedback parameters, induced for example by changes in systematic monetary policy, which we do not accommodate.

C.2. Nominal Bond Prices

Nominal bond prices take the form

$$P_t^n(i) = \exp(A_n(i) + B_n X_t) \tag{11}$$

for $P_t^n(i)$, the zero-coupon bond price of a nominal n-period bond conditional on regime i. The scalar $A_n(i)$ is dependent on regime $s_t = i$ and B_n is a $1 \times N$ vector. It follows that the nominal n-period yield conditional on regime i, $y_t^n(i)$, is an affine function of X_t:

$$y_t^n(i) = -\frac{\log(P_t^n)}{n} = -\frac{1}{n}(A_n(i) + B_n X_t). \tag{12}$$

Appendix B shows that the only difference between the $\widehat{A}_n(i)$ and $\widehat{B}_n$ terms for real bond prices and the $A_n(i)$ and B_n terms for nominal bond prices are due to terms that select inflation from X_t. Positive inflation shocks decrease nominal bond prices.

D. The Effect of Regime Switches

The key ingredient differentiating our model from the standard affine term structure paradigm is the presence of regimes. In this section, we develop intuition on how regimes affect the decomposition of nominal rates into real rate and inflation components.

D.1. Expected Inflation

In our model, one-period expected inflation, $\mathrm{E}_t(\pi_{t+1})$, takes the form

$$\begin{aligned} \mathrm{E}_t(\pi_{t+1}|s_t = i) &= e_3' \mathrm{E}[\mu(s_{t+1})|s_t = i] + e_3' \Phi X_t \\ &= \left(\sum_{j=1}^{K} p_{ij}\ \mu_\pi(j) \right) + e_3' \Phi X_t. \end{aligned} \tag{13}$$

This process is only different from a simple linear process because of the nonlinear drift, which can accommodate sudden discrete changes in expected inflation. Because expected inflation depends on f_t and π_t, the contemporaneous s_t^f and s_t^π regimes also both affect expected inflation.

D.2. Inflation Compensation

With only one regime, one-period inflation compensation, $\pi_{t,1}^e = y_t^1 - \hat{r}_t$, is given by

$$\pi_{t,1}^e = \left(\mu_\pi - \frac{1}{2}\sigma_\pi^2 - \sigma_\pi \lambda_\pi \right) + e_3' \Phi X_t.$$

With multiple regimes, inflation compensation is more complex:

$$\pi_{t,1}^e(i) = -\log \left[\sum_{j=1}^{K} p_{ij} \exp\left(-\mu_\pi(j) + \frac{1}{2}\sigma_\pi^2(j) + \sigma_\pi(j)\lambda_\pi(j) \right) \right] + e_3' \Phi X_t. \tag{14}$$

The last term in the exponential represents the one-period inflation risk premium, which is zero by assumption in our model. The $\frac{1}{2}\sigma_\pi^2(j)$ term is the standard Jensen's inequality term, which now becomes regime dependent. The

$-\mu_\pi(s_t)$ term represents the nonlinear, regime-dependent part of expected inflation. The last term, $e_3{}'\Phi\ X_t$, represents the time-varying part of expected inflation, which does not switch across regimes, and is the only term that is the same as in the affine model.

In comparing expected inflation in equation (13) with inflation compensation in equation (14), we see that the constant terms for $\pi^e_{t,1}$ and $\mathrm{E}_t(\pi_{t+1}\ |s_t)$ are different. The constants in the inflation compensation expression (14) reflect both a Jensen's inequality term $\frac{1}{2}\sigma_\pi^2(s_t)$ and a nonlinear term, driven by taking the log of a sum that is weighted by transition probabilities. Because exp (.) is a convex function, Veronesi and Yared (1999) call this effect a "convexity bias." Like the Jensen's term, this also makes $\pi^e_{t,1} < \mathrm{E}_t(\pi_{t+1})$. In our estimations, both the Jensen's term and the convexity bias amount to less than one basis point, even for longer maturities.

D.3. Real Term Spreads

The intuition for how regimes affect real term spreads can be readily gleaned from considering a two-period real bond. We first analyze the case of the real term spread, $\hat{y}^2_t - \hat{r}_t$, in an affine model without regime switches:

$$\hat{y}^2_t - \hat{r}_t = \frac{1}{2}\left(\mathrm{E}_t(\hat{r}_{t+1}) - \hat{r}_t\right) - \frac{1}{4}\mathrm{var}_t\left(\hat{r}_{t+1}\right) + \frac{1}{2}\mathrm{cov}_t\left(\widehat{m}_{t+1}, \hat{r}_{t+1}\right). \tag{15}$$

The first term, $(\mathrm{E}_t(\hat{r}_{t+1}) - \hat{r}_t)$, is an Expectations Hypothesis (EH) term, the second term, $\mathrm{var}_t(\hat{r}_{t+1})$, is a Jensen's inequality term, and the last term, $\mathrm{cov}_t(\widehat{m}_{t+1}, \hat{r}_{t+1})$, is the risk premium. In the single-regime affine setting, the last term is given by

$$\mathrm{cov}_t(\widehat{m}_{t+1}, \hat{r}_{t+1}) = -\gamma_0\sigma_q - \lambda_f\sigma_f - \gamma_1\sigma_q q_t. \tag{16}$$

Hence, the price of risk factor q_t determines the time variation in the term premium.

The RS model has a more complex expression for the two-period real term spread:

$$\begin{aligned}\hat{y}^2_t(i) - \hat{r}_t = {} & \frac{1}{2}(\mathrm{E}_t(\hat{r}_{t+1}|s_t = i) - \hat{r}_t) - \frac{1}{2}\left(\gamma_0\sigma_q + \gamma_1\sigma_q q_t\right) \\ & -\frac{1}{2}\log\left(\sum_{j=1}^{K} p_{ij}\exp\Big[-\delta_1'\Big(\mu(j) - \mathrm{E}\left[\mu(s_{t+1})|s_t = i\right]\Big)\right. \\ & \left. + \frac{1}{2}\delta_1'\Sigma(j)\Sigma(j)'\delta_1 + \lambda_f(j)\sigma_f(j)\Big]\right),\end{aligned} \tag{17}$$

for K regimes. First, the term spread now switches across regimes, explicitly shown by the dependence of $\hat{y}^2_t(i)$ on regime $s_t = i$. Not surprisingly, the EH term $(\mathrm{E}_t(\hat{r}_{t+1}|s_t = i) - \hat{r}_t)$ now switches across regimes. The time-varying price of risk term, $-\frac{1}{2}(\gamma_0\sigma_q + \gamma_1\sigma_q q_t)$, is the same as in (16) because the process for q_t does not switch regimes. The remaining terms in (17) are nonlinear, as they involve the log of the sum of an exponential function of regime-dependent terms that are weighted by transition probabilities. Within the nonlinear expression, the term $\frac{1}{2}\delta_1'\Sigma(j)\Sigma(j)'\delta_1$ represents a Jensen's inequality term, which is regime-dependent, and $\lambda_f(j)\sigma_f(j)$ represents a RS price of risk term. Thus, the average slope of the real yield curve can potentially change across regimes and produce a variety of regime-dependent shapes of the real yield curve, including flat, inverse-humped, upward-sloping, or downward-sloping yield curves. A new term in (17) that does not have a counterpart in (16) is $-\delta'_1(\mu(j) - \mathrm{E}\,[\mu\,(s_{t+1})\,|s_t = i])$, reflecting the "jump risk" of a change in the regime-dependent drift.

D.4. Inflation Risk Premia

The riskiness of nominal bonds is driven by the covariance between the real kernel and inflation: If inflation is high (purchasing power is low) when the pricing kernel realization (marginal utility in an equilibrium model) is high, nominal bonds are risky and the inflation risk premium is positive. It is tempting to conclude that the sign of the inflation risk premium determines the correlation between expected inflation and real rates. For example, a Mundell–Tobin effect implies that when a bad shock is experienced (an increase in real rates), the holders of nominal bonds experience a countervailing effect, namely, a decrease in expected inflation, which increases nominal bond prices. This intuition is not completely correct as we now discuss.

Consider the two-period pricing kernel, which depends on real rates both through its conditional mean and through real rate innovations. Interestingly, the effects of these two components are likely to act in opposite directions. High real rates decrease the conditional mean of the pricing kernel; but, if the price of risk is negative, positive shocks to the real rate should increase marginal utility. We first focus on the affine model. By splitting inflation into unexpected and expected inflation, we can decompose the two-period inflation risk premium, $\varphi_{t,2}$, into four components (ignoring the Jensen's inequality term):

$$\begin{aligned}\varphi_{t,2} = \frac{1}{2}\big[&-\mathrm{cov}_t(\hat{r}_{t+1}, \mathrm{E}_{t+1}(\pi_{t+2})) - \mathrm{cov}_t(\hat{r}_{t+1}, \pi_{t+1}) \\ &+ \mathrm{cov}_t(\widehat{m}_{t+1}, \mathrm{E}_{t+1}(\pi_{t+2})) + \mathrm{cov}_t(\widehat{m}_{t+1}, \pi_{t+1})\big]. \end{aligned} \tag{18}$$

The first two terms reveal that a negative correlation between real rates and both expected and unexpected inflation actually implies a positive risk premium. Nevertheless, a Mundell–Tobin effect does not necessarily imply a positive inflation risk premium because of the last two terms, which involve the innovations of the pricing kernel. In the affine model equivalent of our RS model, the last term is zero by assumption, but the third term is not and may swamp the others. In particular, for the affine specification:

$$\varphi_{t,2} = -\frac{1}{2}\left[\delta_\pi \sigma_\pi^2(1+\Phi_{\pi\pi}) + \Phi_{\pi q}(\sigma_q^2 + \gamma_1\sigma_q q_t) + \Phi_{\pi f}(\sigma_f^2 + \lambda_f \sigma_f)\right]. \tag{19}$$

Hence, the time variation in the inflation risk premium depends on q_t, and the mean premium depends on parameters that also determine the correlation between real rates and inflation. In particular, if the correlation between real rates and inflation is zero (requiring $\delta_\pi = \Phi_{\pi,q} = \Phi_{\pi,f} = 0$), then the inflation risk premium is also zero. Note that the price of risk λ_f plays a role in determining the inflation risk premium whereas it does not play a role in determining the correlation between real rates and expected inflation.

Naturally, the RS model has a richer expression for the inflation risk premium than equation (19).[3] Regime switches affect the inflation risk premium in two ways, through the RS price of risk, $\lambda_f(s_{t+1})$ and also through the RS means. This gives the inflation risk premium the ability to capture sudden shifts due to changing inflation environments.

E. Econometrics and Identification

We derive the likelihood function of the model in Appendix C. Our model implies a RS-VAR for inflation and yields with complex cross-equation restrictions imposed by the term structure model. Since the model has latent factors, identification restrictions must be imposed to estimate the model. We also discuss these issues in Appendix C. An important identification assumption is that we set the one-period inflation risk premium equal to zero, $\lambda_\pi(s_{t+1}) = 0$. This parameter identifies the average level of real rates and the inflation risk premium, and is very hard to identify without using real yields in the estimation. This restriction does not undermine the ability of the model to fit the dynamics of nominal interest rates and inflation well, as we show below. Models with nonzero λ_π give rise to lower and more implausible real rates than our estimates imply and have a poorer fit with the data.

Finally, we specify the dependence of the prices of risk for the f_t and π_t factors on s_t. Because we set $\lambda_\pi = 0$, we only need to model $\lambda_f(s_{t+1})$. In general, there are four possible λ_f parameters across the four s_{t+1} regimes. This potentially allows real and nominal yield curves to take on different unconditional shapes in different inflationary environments. When estimating a model where $\lambda(s_{t+1})$ varies over all regimes, a Wald test on the equality of $\lambda(s_{t+1})$ across s_{t+1}^π regimes is strongly rejected with a p-value less than 0.001, while a Wald test on the equality of $\lambda(s_{t+1})$ across $s^f{}_{t+1}$ regimes is not rejected at the 5% level. Hence, in our benchmark model, we consider prices of f risk to vary only across inflation regimes, s_{t+1}^π.

[3] The RS inflation risk premium is reported in Ang, Bekaert, and Wei (2007).

F. Related Models

To better appreciate the relative contribution of the model, we link it to three distinct literatures: (i) the extraction of real rates and expected inflation from nominal yields and realized inflation or inflation forecasts, (ii) the empirical RS literature on interest rates and inflation, and (iii) the theoretical term structure literature and equilibrium affine models in finance.

F.1. Time-Series Models

An earlier literature uses neither term structure data, nor a pricing kernel to obtain estimates of real rates and expected inflation. Mishkin (1981) and Huizinga and Mishkin (1986) simply project ex post real rates on instrumental variables. This approach is sensitive to measurement error and omitted variable bias. Other authors, such as Hamilton (1985), Fama and Gibbons (1982), and Burmeister, Wall and Hamilton (1986), use low-order ARIMA models and identify expected inflation and real rates using a Kalman filter under the assumption of rational expectations. The time-series processes driving real rates and expected inflation, with rational expectations, remain critical ingredients in our approach, but we use inflation data and the entire term structure to obtain more efficient identification. In addition, our approach identifies the inflation risk premium, which this literature cannot do.

F.2. Empirical Regime-Switching Models

Many articles document RS behavior in interest rates (see, among many others, Hamilton (1988), Gray (1996), Sola and Driffill (1994), Bekaert, Hodrick and Marshall (2001), and Ang and Bekaert (2002)) without analyzing the real and nominal sources of the regimes. Evans and Wachtel (1993) and Evans and Lewis (1995) document the existence of inflation regimes, whereas Garcia and Perron (1996) focus on real interest rate regimes. Our model simultaneously identifies inflation and real factor sources behind the regime switches and analyzes how they contribute to nominal interest rate variation.

F.3. Term Structure Models

Relative to the extensive term structure literature, our model appears to be the first to identify real interest rates and the components of inflation compensation in a model accommodating regime switches, while still admitting closed-form solutions. Most of the articles using a pricing model to obtain estimates of real rates and expected inflation have so far ignored RS behavior. This includes papers by Pennacchi (1991), Boudoukh (1993), and Buraschi and Jiltsov (2005) for U.S. data and Barr and Campbell (1997) and Evans (1998) for U.K. data. This is curious, because the early literature implicitly demonstrated the importance of accounting for potential structural or regime changes. For example, the Huizinga and Mishkin (1986) projections are unstable over the

1979–1982 period, and the slope coefficients of regressions of future inflation onto term spreads in Mishkin (1990) are substantially different pre- and post-1979, which is also recently confirmed by Goto and Torous (2003).

The articles that have formulated term structure models accommodating regime switches mostly focus only on the nominal term structure. Articles by Hamilton (1988), Bekaert, Hodrick, and Marshall (2001), Bansal and Zhou (2002), and Bansal, Tauchen, and Zhou (2004) allow for RS in mean reversion parameters that we do not, but their derived bond pricing solutions, using discretization or linearization, are only approximate. None of these models features a time-varying price of risk factor like q_t in our model. Naik and Lee (1994) and Landén (2000) present models with closed-form bond prices, but these models feature constant prices of risk and only shift the constant terms in the conditional mean.

The RS term structure model by Dai, Singleton, and Yang (2006) incorporates regime-dependent mean reversions and state-dependent probabilities under the real measure, while still admitting closed-form bond prices. However, under the risk-neutral measure, both the mean reversion and the transition probabilities must be constants, exactly as in our formulation. Dai et al. allow for only two regimes, while we have a much richer RS specification. Another point of departure is that in their model, the evolution of the factors and the prices of risk depend on s_t rather than s_{t+1}. In contrast, our model specifies regime dependence using s_{t+1} as in Hamilton (1989), implying that the conditional variances of our factors embed a jump term reflecting the difference in conditional means across regimes. This conditional heteroskedasticity is absent in the Dai–Singleton–Yang parameterization. Our results show that the conditional means of inflation significantly differ across regimes, while the conditional variances do not, making the regime-dependent means an important source of inflation heteroskedasticity.

There are two related articles that use a term structure model with regime switches to investigate real and nominal yields. The first specification by Veronesi and Yared (1999) is quite restrictive as it only accommodates switches in the drifts. The second paper by Evans (2003) is most closely related to our article. He formulates a model with regime switches for U.K. real and nominal yields and inflation, but he does not accommodate time-varying prices of risk. Evans incorporates switches in mean-reversion parameters, but does not separate the sources of the regime switches into real factors and inflation.

II. Model Estimates

A. Data

We use 4-, 12- and 20-quarter maturity zero-coupon yield data from CRSP and the 1-quarter rate from the CRSP Fama risk-free rate file as our yield data. We compute inflation from the Consumer Price Index—All Urban Consumers (CPI-U, seasonally adjusted, 1982:Q4 = 100) from the Bureau of Labor Statistics. Our data span the sample from 1952:Q2 to 2004:Q4. Using monthly CPI

Table I
Nomenclature of Models

This table summarizes the models estimated. The affine models are single-regime models. In the two- and three-regime models, the real rate factor and inflation share the same regimes, so $s_t = s_t^f = s_t^\pi$, which take values from $\{1, 2\}$ or $\{1, 2, 3\}$, respectively. In the four- and six-regime models, the regimes s_t reflect switches in both s_t^f and s_t^π. In the four-regime model, $s_t^f \in \{1, 2\}$, $s_t^\pi \in \{1, 2\}$, and the probability transition matrix can be one of two cases, independent (Case A) and the correlated case (Case C) outlined in Section I.B.4. In the six-regime model, $s_t^f \in \{1, 2\}$, $s_t^\pi \in \{1, 2, 3\}$, and s_t^f and s_t^π are independent. The three-factor models contain the factors $X_t = (q_t \; f_t \; \pi_t)'$ with q_t a time-varying price of risk factor, f_t is a latent RS term structure factor, and π_t is inflation. The dynamics of X_t are outlined in Section B. The models denoted with w subscripts also contain an additional factor representing expected inflation. These models are described in Appendix D.

		Regime-Switching Models			
	Affine	Two Regimes	Three Regimes	Four Regimes	Six Regimes
3-Factor Models	I	II	III	IV^A, IV^C	VI
4-Factor Models	I_w	II_w	–	IV_w^A, IV_w^C	–

figures creates a timing problem because prices are collected over the course of the month and monthly inflation data are seasonal. Therefore, similar to Campbell and Viceira (2001), we sample all data at the quarterly frequency. For the benchmark model, we specify the 1-quarter and 20-quarter yields to be measured without error to extract the unobserved factors (see Chen and Scott (1993)). The other yields are specified to be measured with error and provide overidentifying restrictions for the term structure model.[4]

B. Model Nomenclature

In Table I, we describe the different term structure models we estimate. The top row represents models with the three factors $(q_t \; f_t \; \pi_t)'$. In the bottom row, we list alternative models that add an unobserved factor representing expected inflation, which we denote by w_t, that generalize classic ARMA models of expected inflation. We describe these models in Appendix D.

To gauge the contribution of regime switches, we estimate single-regime counterparts to the benchmark and unobserved expected inflation models. The single-regime models I and I_w are simply affine models. Model I is the single regime counterpart of the benchmark RS model IV, described in Section I. Model I_w is similar to the model estimated by Campbell and Viceira (2001), except that Campbell and Viceira assume that the inflation risk premium is constant, whereas in all our models the inflation risk premium is stochastic. We specifically contrast real rates and inflation risk premia from Model I_w with the real rates and inflation risk premia implied by our benchmark model below.

[4] We estimate several of our models using alternative schemes where other yields are assumed to be measured without error and find that the results are very similar.

The remaining models in Table I are RS models. Models *II* and II_w contain two regimes where $s_t^f = s_t^\pi$. Two regime models are the main specifications used in the empirical and term structure literature (see, for example, Bansal and Zhou (2002)). Model *III* considers a similar model but the regime variable can take on three values. Model *IV* represents the benchmark model, which has four regimes, with the different cases describing the correlation of the s_t^f and the s_t^π regimes (Cases A and C as described in Section II.B). Model *VI* contains two regimes for s_t^f that are independent of the three regimes for s_t^π.

C. Specification Tests

We report two specification tests of the models, an unconditional moment test and an in-sample serial correlation test for first and second moments in scaled residuals. The former is particularly important because we want to decompose the variation of nominal yields into real and expected inflation components. A well-specified model should imply unconditional means, variances, and autocorrelograms of inflation and yields close to the sample moments. We outline these tests in Appendix E.

Table II reports the results of these specification tests. Panel A focuses on matching inflation dynamics, while Panel B focuses on matching the dynamics of yields. Of all the models, only Model IV^C passes the inflation residual tests and fits the mean, variance, and autocorrelogram of inflation (using autocorrelations of lags 1, 5, and 10). About half of the models fail to match the autocorrelogram of inflation. Inflation features a relatively low first-order autocorrelation coefficient with very slowly decaying higher-order autocorrelations. Generally, the presence of regimes and the additional expected inflation factor help in matching this pattern. However, most of the models with the w-factor fail to match the mean and variance of inflation. While Model *VI* passes all moment tests, both residual tests reject strongly at the 1% level, eliminating this model. The match with inflation dynamics is extremely important as the estimated inflation process not only identifies expected inflation but also plays a critical role in identifying the inflation risk premium. This makes Model IV^C the prime candidate for the best model.

Panel B reports goodness-of-fit tests for two sets of yield moments, namely, the mean and variance of the spread and the long rate (all models fit the mean of the short rate by construction in the estimation procedure), and the autocorrelogram of the spread. Only four models fit the moments of yields and spreads: *I*, *III*, IV^A, and IV^C. Unfortunately, apart from model IV^C, these other models fail to match the inflation moments in Panel A.

We also report the residual test for the short rate and spread equations in Panel B. With the exception of model *VI*, most models produce reasonably well-behaved residuals. While model IV^C nails the dynamics of inflation in Panel A and closely matches term structure moments, the model's residual tests for short rates and spreads are significant at the 5% level, but not at the 1% level. Thus, there is some serial correlation and heteroskedasticity that

Table II

Specification Tests

This table reports moment and residual tests of inflation (Panel A) and of yields (Panel B), which are outlined in Appendix E. In the columns titled "Moment Tests," we report the p-values of goodness-of-fit χ^2 tests for various moments implied by the different models. In Panel A, the first moment test matches the mean and variance of inflation, whereas in Panel B, the first moment test matches the mean and variance of the long rate and the spread jointly. The long rate refers to the 20-quarter nominal rate y_t^{20} and the spread refers to $y_t^{20} - y_t^1$, for y_t^1 the 3-month short rate. The second autocorrelogram moment test matches autocorrelations at lags 1, 5, and 10. The columns titled "Residual Tests" report p-values of scaled residual tests for the different models. The first entry reports the p-value of a test of E $(\epsilon_t\ \epsilon_{t-1}) = 0$ and the second row reports the p-value of a GMM-based test of E $[(\epsilon_t^2 - 1)(\epsilon_{t-1}^2 - 1)] = 0$, where ϵ_t is a scaled residual. P-values less than 0.05 (0.01) are denoted by * (**). Table I contains the nomenclature of the various models.

Panel A: Matching Inflation Dynamics

	Moment Tests		
Model	Mean/Variance	Auto-correlogram	Residual Tests
I	0.00**	0.02*	0.00**
			0.08
I_w	0.08	0.00**	0.02*
			0.09
II	0.00**	0.01*	0.10
			0.17
II_w	0.00**	0.16	0.03*
			0.31
III	0.02*	0.02*	0.67
			0.22
IV^A	0.15	0.04*	0.16
			0.12
IV^C	0.60	0.08	0.21
			0.10
IV_w^A	0.00**	0.27	0.26
			0.26
IV_w^C	0.00**	0.18	0.22
			0.27
VI	0.50	0.13	0.00**
			0.00**

(*continued*)

remains present in the residuals. Consequently, the unconditional moments of unobserved real rates and inflation risk premia produced by model IV^C will imply nominal rates and inflation behavior close to that in the data, but the conditional dynamics of real short rates and inflation risk premia may be slightly more persistent or heteroskedastic than our estimates suggest.

D. Model Estimates

We focus on the benchmark model IV^C, which is the model that best fits the inflation and term structure data.[5] We discuss the parameter estimates,

[5] Estimates of other models are available upon request.

Table II—*Continued*
Specification Tests

Panel B: Matching Yield Dynamics

Model	Moment Tests: Mean/Var Long Rate/Spread	Moment Tests: Spread Autocorrelogram	Residual Tests: Short Rate	Residual Tests: Spread
I	0.78	0.14	0.19	0.14
			0.27	0.22
I_w	0.00**	0.26	0.47	0.34
			0.15	0.29
II	0.61	0.01**	0.05	0.65
			0.02*	0.15
II_w	0.00**	0.01*	0.52	0.48
			0.01**	0.34
III	0.12	0.09	0.05	0.05
			0.04*	0.05
IV^A	0.37	0.33	0.02*	0.96
			0.04*	0.08
IV^C	0.63	0.39	0.02*	0.34
			0.04*	0.03*
IV_w^A	0.00**	0.06	0.31	0.11
			0.08	0.35
IV_w^C	0.00**	0.24	0.33	0.07
			0.12	0.30
VI	0.04*	0.00**	0.01**	0.01*
			0.01**	0.00**

the implied factor dynamics, and the identification and interpretation of the regimes.

D.1. Parameter Estimates

Table III reports the parameter estimates. Inflation enters the real short rate equation (4) with a highly significant, negative coefficient of $\delta_\pi = -0.49$. In the companion form Φ of the VAR, the term structure latent factors q_t and f_t are both persistent, with correlations of 0.97 and 0.76, respectively. Their effects on the conditional mean of inflation and thus on expected inflation are positive with coefficients of 0.62 and 0.95, respectively. However, the coefficient on f_t is only borderline significant with a t-statistic of 1.85. Not surprisingly, lagged inflation also significantly enters the conditional mean of inflation, with a loading of 0.54. A test of money neutrality ($\delta_\pi = \Phi_{\pi,q} = \Phi_{\pi,f} = 0$) rejects with a p-value less than 0.001.

The conditional means and variances of the factors reveal that the first $s_t^f = 1$ regime is characterized by a low f_t mean and low standard deviation. Both the mean and standard deviations are significantly different across the two regimes at the 5% level. For the inflation process, the conditional mean of inflation is

Table III

Benchmark Model IV^C Parameter Estimates

The table reports estimates of the benchmark RS model IV^C with correlated s_t^f and s_t^π outlined in Section I. The stable probabilities of regime 1 to 4 are 0.725, 0.039, 0.197, and 0.038, with standard errors of 0.081, 0.029, 0.052, and 0.018, respectively. We reject the null of independent regimes (Case A) with a p-value of 0.033 using a likelihood ratio test.

Short Rate Equation $r_t = \delta_0 + \delta_1{}'X_t$				
		δ_1		
	δ_0	q	f	π
	0.008	1.000	1.000	-0.488
	(0.001)	–	–	(0.056)

Companion Form Φ				
		q	f	π
	q	0.975	0.000	0.000
		(0.014)	–	–
	f	0.000	0.762	0.000
		–	(0.012)	–
	π	0.618	0.954	0.538
		(0.164)	(0.516)	(0.064)

Conditional Means and Volatilities	Regime 1	Regime 2	P-value Test of Equality
$\mu_f(s_t^f) \times 100$	−0.010	0.034	0.037
	(0.005)	(0.016)	
$\mu_\pi(s_t^\pi) \times 100$	0.473	0.248	0.002
	(0.082)	(0.110)	
$\sigma_q \times 100$	0.094		–
	(0.011)		
$\sigma_f(s_t^f) \times 100$	0.078	0.175	0.000
	(0.019)	(0.047)	
$\sigma_\pi(s_t^\pi) \times 100$	0.498	0.573	0.249
	(0.028)	(0.063)	

(*continued*)

significantly different across the s_t^π regimes, with $s_t^\pi = 1$ being a relatively high inflation environment. However, there is no significant difference across regimes in the innovation variances. This does not mean that inflation is homoskedastic in this model. The regime-dependent means of f_t induce heteroskedastic inflation across the f_t factor regimes.

Table III also reports that the price of risk for the q_t factor is negative but imprecisely estimated. The prices of risk for the f_t factor are both significantly different from zero and significantly different across the two regimes. Moreover, they have a different sign in each regime, which may induce different term structure slopes across the regimes.

The transition probability matrix shows that the s_t^f regimes are persistent with probabilities $Pr(s^f{}_{t+1} = 1|s_t^f = 1) = 0.93$ and $Pr(s^f{}_{t+1} = 2|s_t^f = 2) = 0.77$.

Table III—*Continued*

Prices of Risk $\lambda(s_t^\pi) = (\gamma_1\, q_t\ \lambda_f(s_t^\pi)\ 0)'$			
	$\lambda_f(s_t^\pi)$		
γ_1	Regime 1	Regime 2	P-value Test of Equality
−17.1	−0.613	0.504	0.000
(15.7)	(0.097)	(0.151)	

Transition Probabilities Π				
	$s_{t+1}=1$	$s_{t+1}=2$	$s_{t+1}=3$	$s_{t+1}=4$
$s_t=1$	0.930	0.000	0.065	0.005
	(0.025)	(0.008)	(0.020)	(0.002)
$s_t=2$	0.125	0.804	0.019	0.052
	(0.030)	(0.029)	(0.007)	(0.016)
$s_t=3$	0.228	0.000	0.716	0.056
	(0.047)	(0.002)	(0.045)	(0.024)
$s_t=4$	0.031	0.197	0.205	0.567
	(0.010)	(0.041)	(0.039)	(0.064)
p^f	0.930	q^f	0.772	
	(0.021)		(0.047)	
p^{AA}	1.000	p^{AB}	0.135	
	(0.009)		(0.031)	
p^{AB}	0.865	p^{BB}	0.735	
	(0.031)		(0.055)	

Std Dev × 100 of Measurement Errors	
y_t^4	y_t^{12}
0.050	0.024
(0.003)	(0.001)

The probability $p^{AA} = Pr(s_{t+1}^\pi = 1 | s^f{}_{t+1} = 1,\ s_t^\pi = 1)$ is estimated to be one. Conditional on a period with a negative f_t and relatively high inflation (regime 1), we cannot transition into a period of lower expected inflation unless the f_t regime also shifts to the higher mean regime. Thus, the model assigns zero probability from transitioning from $s_t = 1 \equiv (s_t^f = 1, s_t^\pi = 1)$ to $s_{t+1} = 2 \equiv (s^f{}_{t+1} = 1,\ s_{t+1}^\pi = 2)$. Similarly, starting in regime 3, $s_t = 3 \equiv (s_t^f = 2,\ s_t^\pi = 1)$, we can transition into the low inflation regime ($s_{t+1}^\pi = 2$) only with a realization of $s^f{}_{t+1} = 2$, where f_t is high and volatile. We demonstrate below that this behavior has a plausible economic interpretation.

D.2. Factor Behavior

Table IV reports the relative contributions of the different factors driving the short rate, long yield, term spread, and inflation dynamics in the model. The price of risk factor q_t is relatively highly correlated with both inflation and the nominal short rate, but shows little correlation with the nominal spread.

Table IV

Factor Behavior

The table reports various unconditional moments of the three factors: the time-varying price of risk factor q_t, the RS factor f_t, and inflation π_t, from the benchmark model IV^C. The short rate refers to the 1-quarter nominal yield and the spread refers to the 20-quarter nominal term spread. The row labelled "Data π" refers to actual inflation data. The numbers between parentheses are standard errors reflecting parameter uncertainty from the estimation, computed using the delta method. The variance decomposition of the real rate is computed as $\mathrm{cov}(r_t, z_t)/\mathrm{var}(r_t)$, with z_t respectively q_t, f_t, and $\delta_\pi\ \pi_t$. The variance decomposition of expected inflation is computed as $\mathrm{cov}(\mathrm{E}_t[\pi_{t+1}], z_t)/\mathrm{var}(\mathrm{E}_t[\pi_{t+1}])$, with z_t respectively $\Phi_{\pi q}\, q_t$, $\Phi_{\pi f}\, f_t$, and $\Phi_{\pi\pi}\, \pi_t$. Panel B reports multivariate projection coefficients of inflation on the lagged short rate, spread, and inflation implied by the model and in the data. Standard errors in parentheses are computed using the delta method for the model-implied coefficients and are computed using GMM for the data coefficients.

Panel A: Moments of Factors

					Correlation with				
	St Dev	Auto	Contribution to Real Rate Variance	Contribution to Expected Inflation Variance	Inflation	Nominal Short Rate	Nominal Spread	Real Short Rate	Real Spread
q	1.70	0.98	0.51	0.28	0.61	0.90	−0.20	0.44	−0.09
	(0.55)	(0.01)	(0.35)	(0.08)	(0.11)	(0.05)	(0.07)	(0.21)	(0.02)
f	0.68	0.74	0.09	0.09	0.24	0.43	−0.99	0.19	−0.24
	(0.20)	(0.02)	(0.10)	(0.05)	(0.07)	(0.11)	(0.02)	(0.17)	(0.17)
π	3.50	0.76	0.40	0.62	1.00	0.69	−0.44	−0.34	0.59
	(0.42)	(0.05)	(0.36)	(0.08)	–	(0.08)	(0.06)	(0.29)	(0.12)
Data π	3.16	0.72				0.68	−0.37		

Panel B: Projection of Inflation on Lagged Instruments

	Inflation	Nominal Short Rate	Nominal Spread
Model	0.52	0.39	−0.08
	(0.06)	(0.07)	(0.17)
Data	0.49	0.29	−0.39
	(0.06)	(0.07)	(0.15)

In other words, q_t can be interpreted as a level factor. The RS term structure factor f_t is highly correlated with the nominal spread, in absolute value, so f_t is a slope factor. The factor f_t is also less variable and less persistent than q_t. Consequently, f_t does not play a large role in the dynamics of the real rate, only accounting for 9% of its variation. The most variable factor is inflation, which accounts for 51% of the variation of the real rate. Inflation is negatively correlated with the real short rate, at −34%, as a result of the negative $\delta_\pi = -0.49$ coefficient, while q_t is positively correlated with the real short rate (44%). The model produces a 69% (−44%) correlation between inflation and the nominal short rate (nominal 5-year spread), which matches the data correlation of 68% (−37%) very closely.

Panel A also reports how the different factors contribute to the expected inflation dynamics. The latent factor components play an important role in the dynamics of expected inflation, with q_t and f_t accounting for 37% of the variance of expected inflation. Inflation itself accounts for 62% of the variance of inflation. Expected inflation also has a nonlinear RS component. We calculate the contribution of regimes to the variance of expected inflation by computing the variance of expected inflation assuming we never transition from regime 1, relative to the variance of expected inflation from the full model. Unconditionally, RS accounts for 12% of the variance of expected inflation. We also show later that regimes are critical for capturing sudden decreases in expected inflation occurring occasionally during the sample.

The implied processes for expected inflation and actual inflation are both very persistent. The first-order autocorrelation coefficient of one-quarter expected inflation is 0.89, which implies a monthly autocorrelation coefficient of 0.96 under the null of an AR(1). The autocorrelations decay slowly to 0.51 at 10 quarters. Fama and Schwert (1977) also note the strong persistence of expected inflation using time-series techniques to extract expected inflation estimates. For actual inflation, the first-order autocorrelation implied by the model is 0.76 and it is 0.35 at 10 quarters, matching the data almost perfectly at 0.72 and 0.35, respectively.[6] It is this very persistent nature of inflation that many of the other models cannot match. For example, in model I_w, similar to Campbell and Viceira (2001), the autocorrelations of actual inflation are 0.48 and 0.20 at 1 and 10 lags, respectively.

Because the factors are highly correlated with inflation, the nominal short rate, and the nominal spread, these three variables should capture a substantial proportion of the variance of expected inflation in our model. To verify this implication of our model with the data, we project inflation onto the short rate, spread, and past inflation both in the data and in the model. Panel B of Table IV reports these results. When the short rate increases by 1%, the model signals an increase in expected inflation of 39 basis points. A 1% increase in the spread predicts an eight basis point decrease in expected inflation. These patterns are consistent with what is observed in the data, but the response to an increase in the spread is somewhat stronger in the data. Past inflation has a coefficient of 0.52, matching the data coefficient of 0.49 almost exactly.

The model also matches other predictive regressions of future inflation. For example, Mishkin (1990) regresses the difference between the future n-period inflation rate and the one-period inflation rate onto the the n-quarter term spread. In the data, this coefficient takes on a value of 0.98 with a standard error of 0.36 for a horizon of 1 year. The model-implied coefficient is 0.97. Thus, we are confident that the model matches the dynamics of expected inflation well.

[6] The autocorrelations of inflation vary only modestly across regimes, with the first-order autocorrelation of inflation being highest in regime $s_t = 1$ at 0.77 and lowest in regime $s_t = 4$ at 0.74.

Table V

Real Rates, Inflation Compensation, and Nominal Rates across Regimes

We report means and standard deviations for real short rates, $\hat{r}_t$, the 20-quarter real term spread, $\hat{y}_t^{20} - \hat{r}_t$, 1-quarter ahead inflation compensation, $\pi^e_{t,1}$, and nominal short rates, r_t, implied by model IV^C across each of the four regimes. The regime $s_t = 1$ corresponds to $(s_t^f = 1, s_t^\pi = 1)$, $s_t = 2$ to $(s_t^f = 1, s_t^\pi = 2)$, $s_t = 3$ to $(s_t^f = 2, s_t^\pi = 1)$, and $s_t = 4$ to $(s_t^f = 2, s_t^\pi = 2)$. Standard errors reported in parentheses are computed using the delta method.

		Regime			
		$s_t = 1$	$s_t = 2$	$s_t = 3$	$s_t = 4$
Real Short Rate $\hat{r}_t$	Mean	1.14	1.98	1.34	1.97
		(0.39)	(0.53)	(0.35)	(0.45)
	Std Dev	1.40	1.55	1.55	1.68
		(0.22)	(0.29)	(0.25)	(0.29)
Real Term Spread $\hat{y}_t^{20} - \hat{r}_t$	Mean	0.15	−0.39	−0.03	−0.45
		(0.31)	(0.21)	(0.28)	(0.16)
	Std Dev	1.12	1.26	1.31	1.42
		(0.17)	(0.25)	(0.22)	(0.25)
Inflation Compensation $\pi^e_{t,1}$	Mean	3.92	2.46	4.43	3.20
		(0.38)	(0.79)	(0.39)	(0.67)
	Std Dev	2.75	2.95	3.01	3.13
		(0.50)	(0.51)	(0.48)	(0.49)
Nominal Short rate r_t	Mean	5.06	4.45	5.77	5.17
		(0.08)	(0.38)	(0.17)	(0.34)
	Std Dev	3.04	3.12	3.47	3.50
		(0.74)	(0.73)	(0.65)	(0.65)

D.3. Regime Interpretation

How do we interpret the behavior of the regime variable in economic terms? In Table V, we describe the behavior of real short rates, one-quarter ahead inflation compensation (which is virtually identical to one-period expected inflation except for Jensen's inequality terms), and nominal short rates across regimes. This information leads to the following regime characterization:

		Real Short Rates	Inflation	% Time
$s_t = 1$	$s_t^f = 1, s_t^\pi = 1$	Low and Stable	High and Stable	72%
$s_t = 2$	$s_t^f = 1, s_t^\pi = 2$	High and Stable	Low and Stable	4%
$s_t = 3$	$s_t^f = 2, s_t^\pi = 1$	Low and Volatile	High and Volatile	20%
$s_t = 4$	$s_t^f = 2, s_t^\pi = 2$	High and Volatile	Low and Volatile	4%

All the levels (low or high) and variability (stable or volatile) are relative statements, so caution must be taken in the interpretation. The last column lists the proportion of time spent in each regime in the sample based on the population probabilities.[7] The means of both real rates and inflation are driven mostly by the s_t^{π} regime, while their volatilities are driven by the s_t^f regime.

The first regime is a low real rate-high inflation regime, where both real rates and inflation are not very volatile. We spend most of our time in this regime. As we will see, it is better to characterize the relatively high inflation regime as a "normal regime" and the low inflation regime as a "disinflation regime." The volatilities of real short rates, inflation compensation, and nominal short rates are all lowest in regime 1. The regime with the second-largest stable probability is regime 3, which is also a low real rate regime. In this regime, the mean of inflation compensation is highest. Thus, in population we spend around 90% of the time in low real rate environments. Regimes 2 and 4 are characterized by relatively high and volatile real short rates. The inflation compensation in these regimes is relatively low. Table V shows that these regimes are also associated with downward-sloping term structures of real yields. Consequently, the transition probability estimates imply that passing through a downward-sloping real yield curve is necessary to reach the regime with relatively low inflation. Finally, regime 4 has the highest volatility of real rates, inflation compensation, and nominal rates.

D.4. Regimes over Time

In Figure 1, we plot the short rate, long rate, and inflation over the sample in the top panel and the smoothed regime probabilities in the bottom panel over the sample period. From 1952 to 1978, the estimation switches between $s_t = 1$ and $s_t = 3$. Recall that these regimes feature relatively low real rates and high inflation. In regime 3, inflation has its highest mean and is quite volatile, leading to high and volatile nominal rates. These regimes precede the recessions of 1960, 1970, and 1975.

Post-1978, the model switches between all four regimes. The period around 1979 to 1982 of monetary targeting is mostly associated with regime 4, characterized by the highest volatility of real rates and inflation and a downward sloping real yield curve. Before the economy transitions to regime $s_t = 2$ in 1982, with high real rates and low and more stable inflation, there are a few jumps into the higher inflation regime 3.

Post-1982, regimes 2 and 4, with lower expected inflation, occur regularly. These regimes are associated with rapid decreases in inflation and downward-sloping real yield curves. From a Taylor (1993) rule perspective, these regimes may reflect periods in which an activist monetary policy of raising real rates, especially through actions at the short-end of the yield curve, achieved disinflation. Several features of the occurrence of these regimes are consistent with

[7] If we identify the regimes through the sample by using the ex post smoothed regime probabilities, then we spend less time in regime $s_t = 1$ in sample than the population frequency. Unlike traditional two-regime estimations, like Gray (1996) and Bansal and Zhou (2002), this is not caused purely by switching out of $s_t = 1$ during the monetary targeting period of 1979 to 1982. In contrast, our model produces more recurring switches into regimes $s_t = 2$ and $s_t = 4$. Such switches also occur during the early 1990s and early 2000s, which we discuss below.

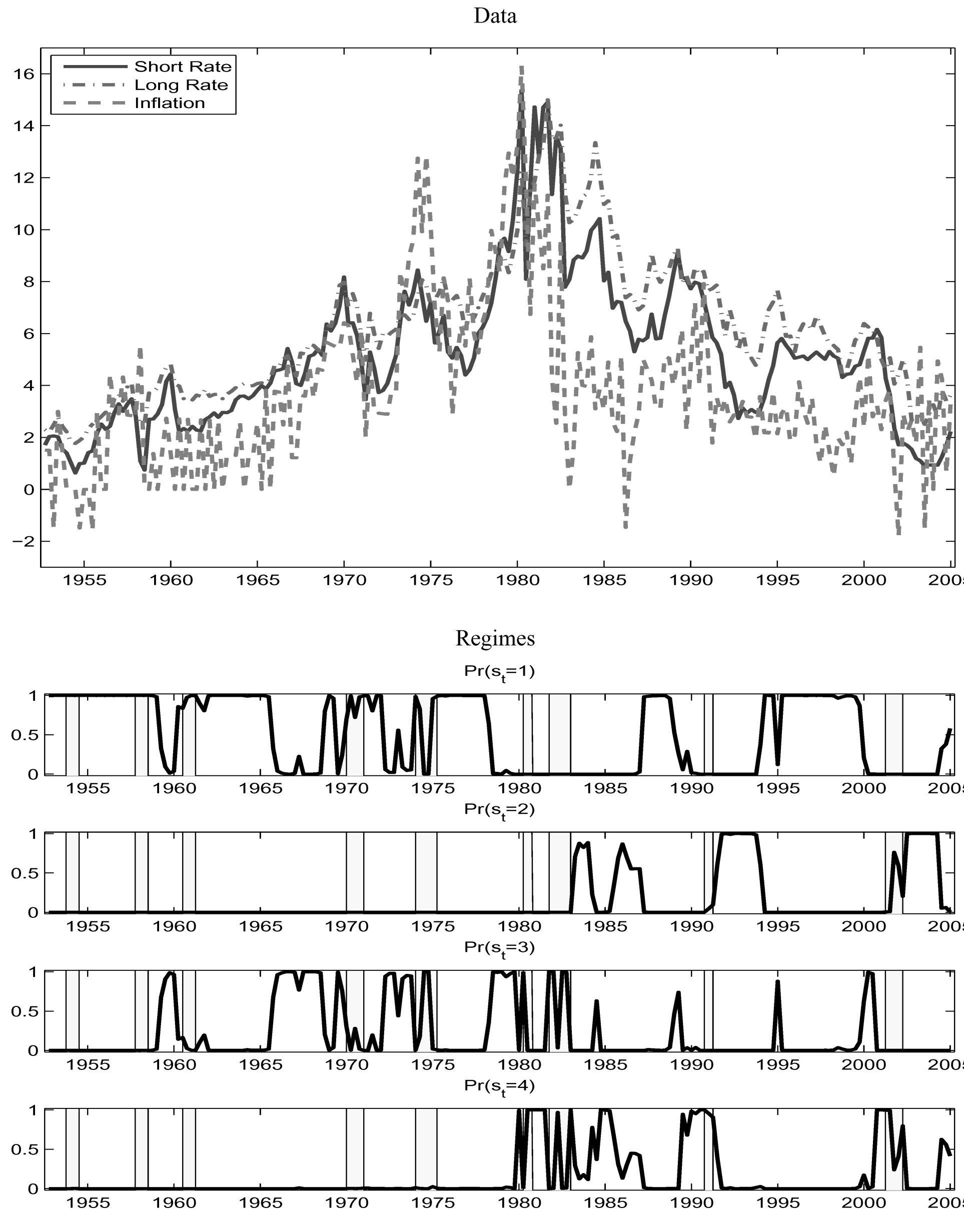

Figure 1. Smoothed regime probabilities, all regimes. The top graph plots the nominal short rate (1-quarter yield) and nominal long rate (20-quarter yield) together with quarter-on-quarter inflation. The top panel's y-axis units are annualized and are in percentages. In the bottom graph, we plot the smoothed probabilities of each of the four regimes, $Pr(s_t = i|I_T)$, conditioning on data over the entire sample, from the benchmark model IV^C. NBER recessions are indicated by shaded bars.

this interpretation. First, transitioning into regimes 2 and 4 requires high real rates. Second, these regimes only occur after the Volcker period, which is consistent with Nelson (2004) and Meltzer (2005), who argue that U.S. monetary authorities had sufficient credibility to change inflation behavior only after 1979. Third, it is also consistent with the econometric analysis of the Taylor rule in Bikbov (2005), Boivin (2006), and Cho and Moreno (2006), among others, who document a structural break from accommodating to activist monetary policies around 1980.

Towards the end of the 1980s we transition back to the normal regime 1, but just before the 1990 to 1991 recession the economy enters into regime 4, followed by regime 2, which lasts until 1994. During the late 1990s, the normal regime $s_t = 1$ prevails with normal, stable inflation and low real rates. During the early 2000s, quarter-on-quarter inflation was briefly negative, and the model transitions to the disinflation regimes $s_t = 2$ and $s_t = 4$ around the time of the 2001 recession. At the end of the sample, December 2004, the model seems to be transitioning back to the normal $s_t = 1$ regime.

In Figure 2, we sum the four s_t regimes into their s_t^f and s_t^π sources. In the top panel, we graph the real short and long 20-quarter real rates, together with one-period expected inflation and long-term inflation compensation for comparison. The real short rate exhibits considerable short-term variation, sometimes decreasing and increasing sharply. There are sharp decreases of real rates in the 1958 and 1975 recessions and after the 2001 recession. Real rates are highly volatile around the 1979–1982 period and increase sharply during the 1980 and 1983 recessions.[8] Consistent with the older literature like Mishkin (1981), real rates are generally low from the 1950s until 1980. The sharp increase in the early 1980s to above 7% was temporary, but it took until after 2001 before real rates reached the low levels common before 1980. Over 1961–1986, Garcia and Perron (1996) find three nonrecurring regimes for real rates: 1961–1973, 1973–1980, and 1980–1986. In Figure 2, these periods roughly correspond to low but stable real rates, very low to negative and volatile real rates, and high and volatile real rate periods. We generate this behavior with recurring s_t^f and s_t^π regimes. The Garcia–Perron model could not generate the gradual decrease in real rates observed since the 1980s. The long real rate shows less time variation, but the same secular effects that drive the variation of the short real rate are visible.

In the middle panel of Figure 2, we plot the smoothed regime probabilities for the regime $s_t^f = 1$, which is the low volatility f_t regime associated with relatively high nominal term spreads. The high variability $s_t^f = 2$ regime occurs just prior to the 1960 recession, during the OPEC oil shocks of the early 1970s, during the 1979–1982 period of monetary targeting, during the 1984 Volcker disinflation, in the 1991 recession, briefly in 1995, and in 2000.

In the bottom panel of Figure 2, the smoothed regime probabilities of s_t^π look very different from the regime probabilities of s_t^f, indicating the potential

[8] The 95% standard error bands computed using the delta method are very tight and well within 20 basis points, so we omit them for clarity.

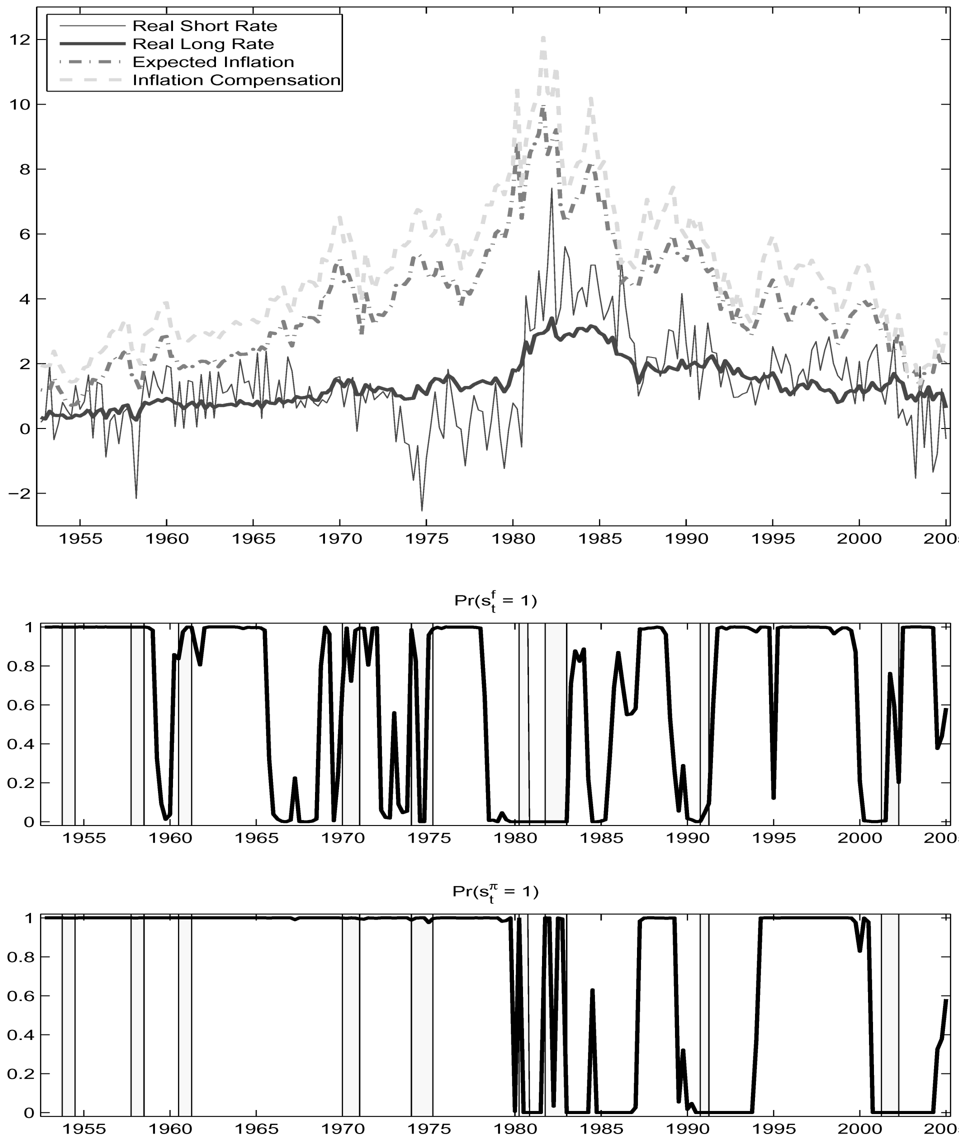

Figure 2. Smoothed regime probabilities. The top panel graphs the real short rate, $\hat{r}_t$, real long rate, $\hat{y}^{20}_t$, 1-quarter expected inflation, $\mathrm{E}_t(\pi_{t+1})$, and long-term inflation compensation, $\pi^e_{t,20}$, all implied from the benchmark model IV^C. The top panel's y-axis units are annualized and are in percentages. The middle and bottom panels plot smoothed regime probabilities using information from the whole sample. The middle panel shows the smoothed probabilities $Pr(s^f_t = 1|I_T)$ of the f factor regimes, s^f_t. The bottom panel graphs the smoothed probabilities $Pr(s^\pi_t = 1|I_T)$ of the inflation factor regime, s^π_t. NBER recessions are indicated by shaded bars.

importance of separating the real and inflation regime variables. We transition to $s^\pi_t = 2$, the disinflation regime, only after 1979, with the 1979–1982 period featuring some sudden and short-lived transitions to $s^\pi_t = 2$. The second inflation regime also occurs after 1985, during a sustained period in the early 1990s, and after 2000. In this last recession, there were significant risks of deflation. Clearly, the model accommodates rapid decreases in inflation by a transition to the second regime.[9]

Standard two-regime models of nominal interest rates (both empirical and term structure models) predominantly select the late 1970s and early 1980s as one regime change. These two-regime models identify the pre-1979 period and the period after the mid-1980s as a low mean, low volatility regime (see, for example, Gray (1996), Ang and Bekaert (2002), and Dai et al. (2006)). Our regimes for real factors and inflation have more frequent switches than two-regime models. In fact, the famous 1979–1982 episode is a period of both high real rates and high inflation in the late 1970s (regime 3), combined with high real rates and a transition to the second inflation regime caused by a dramatic decrease in inflation in the early 1980s (regime 4). Hence, our regime identification does not seem to be driven by a single period, but rather reflects a series of recurring regimes.

III. The Term Structure of Real Rates and Expected Inflation

We describe the behavior of real yields in Section III.A. Section III.B discusses the behavior of expected inflation and inflation risk premia. Combining real yields with expected inflation and inflation risk premia produces the nominal yield curve, which we discuss in Section III.C, before turning to variance decompositions in Section III.D.

A. The Behavior of Real Yields

A.1. The Real Term Structure

We examine the real term structure in Figure 3 and Table VI. Figure 3 graphs the regime-dependent real term structure. Every point on the curve for regime i represents the expected real zero-coupon bond yield conditional on regime i, $(\mathrm{E}[\hat{y}^n_t|s_t = i])$.[10] The unconditional real yield curve is graphed in the circles, which show a slightly humped real curve peaking around a 1-year maturity before converging to 1.3%. Panel A of Table VI reports that in the normal regime $(s_t = 1)$, the long-term rate curve assumes the same shape but is shifted slightly downwards, ranging from 1.14% at a 3-month horizon to 1.29% at a 5-year horizon.

[9] The inflation regime identifications of Evans and Wachtel (1993) and Evans and Lewis (1995) are not directly comparable as their models feature a random walk component in one regime (with no drift) and an AR(1) model in the other.

[10] These computations are detailed in Ang et al. (2007). It is also possible to compute the more extreme case $\mathrm{E}[\hat{y}^n_t|s_t = i, \forall t]$, that is, assuming that the process never leaves regime i. These curves have similar shapes to the ones shown in the figures but lie at different levels.

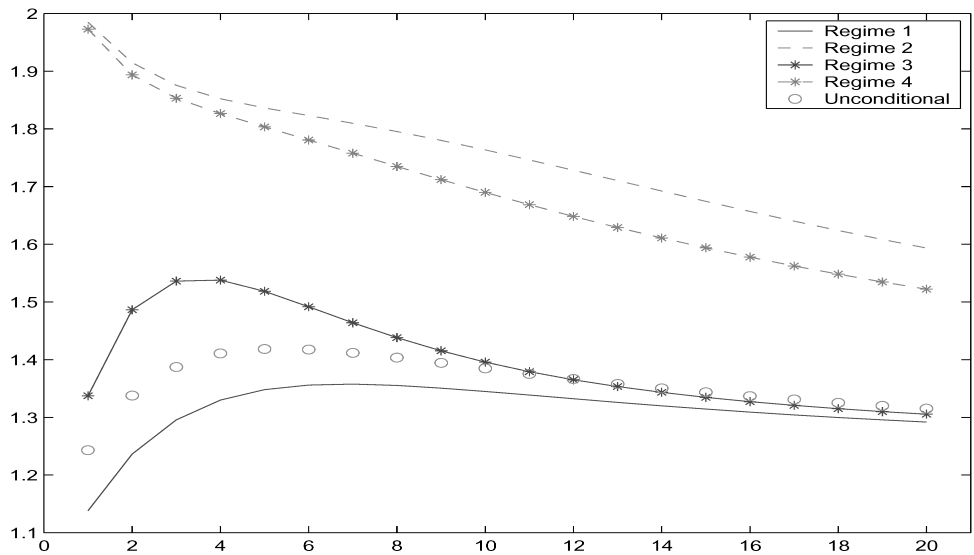

Figure 3. Real-term structure. We graph the real yield curve, conditional on each regime and the unconditional real yield curve implied from the benchmark model IV^C. The x-axis displays maturities in quarters of a year. The y-axis units are annualized and are in percentages.

In regimes 2 and 4, real rates start just below 2% at a 1-quarter maturity and decline to 1.59% for regime 2 and 1.52% for regime 4 at a 20-quarter maturity. Finally, regime 3, a low real rate-high inflation and volatile regime, has a humped, nonlinear, real term structure. This real yield curve peaks at 1.54% at the 1-year maturity before declining to the same level as the unconditional yield curve at 20 quarters. Thus, we uncover our first claim:

CLAIM 1: *Unconditionally, the term structure of real rates assumes a fairly flat shape around 1.3%, with a slight hump, peaking at a 1-year maturity. However, there are some regimes in which the real rate curve is downward sloping.*

Panel A of Table VI also reports that while the standard deviation of real short rates is lowest in regime 1 at 1.40%, the standard deviations of real long rates are approximately the same across regimes, at 0.55%. We compute unconditional moments of real yields in Panel B, which shows that the unconditional standard deviation of the real short rate (20-quarter real yield) is 1.46% (0.55%). These moments solidly reject the hypothesis that the real short rate is constant, but at long horizons real yields are much more stable and persistent. This is reflected in the autocorrelations of the real short rate and 20-quarter real rate, which are 60% and 94%, respectively. The mean of the 20-quarter real term spread is only 7 basis points. The standard error is only 28 basis points, so that the real term structure cannot account for the 1.00% nominal term spread in the data. Hence:

CLAIM 2: *Real rates are quite variable at short maturities but smooth and persistent at long maturities. There is no significant real term spread.*

Table VI

Characteristics of Real Rates

The table reports various moments of the real rate, implied from model IV^C. Panel A reports the conditional mean and standard deviation of real rates of various maturities in quarters across regimes. Panel B reports the unconditional mean, standard deviation, and autocorrelation of real yields. Panel C reports the correlation of real yields with actual and unexpected inflation implied from the model. We report the conditional correlation of real yields with actual inflation, $\mathrm{corr}(\hat{y}^n_{t+1}, \pi_{t+1}|s_t)$, and the conditional correlation of real yields with expected inflation, $\mathrm{corr}(\hat{y}^n_{t+1}, \mathrm{E}_{t+1}(\pi_{t+1+n,n})|s_t)$. Standard errors reported in parentheses are computed using the delta method.

Panel A: Conditional Moments

	Regime $s_t = 1$		Regime $s_t = 2$		Regime $s_t = 3$		Regime $s_t = 4$	
Maturity Qtrs	Mean	St Dev	Mean	St Dev	Mean	St Dev	Mean	St Dev
1	1.14	1.40	1.98	1.55	1.34	1.55	1.97	1.68
	(0.39)	(0.22)	(0.53)	(0.29)	(0.35)	(0.25)	(0.45)	(0.29)
4	1.33	0.86	1.85	0.93	1.54	0.89	1.83	0.94
	(0.38)	(0.25)	(0.54)	(0.25)	(0.40)	(0.25)	(0.48)	(0.25)
20	1.29	0.55	1.59	0.56	1.31	0.55	1.52	0.56
	(0.39)	(0.32)	(0.42)	(0.32)	(0.44)	(0.32)	(0.41)	(0.49)

Panel B: Unconditional Moments

Maturity Qtrs	Mean	St Dev	Auto
1	1.24	1.46	0.60
	(0.38)	(0.23)	(0.08)
4	1.41	0.88	0.73
	(0.38)	(0.25)	(0.13)
20	1.32	0.55	0.94
	(0.40)	(0.32)	(0.05)
Spread 20-1	0.07	1.19	0.52
	(0.28)	(0.18)	(0.06)

(continued)

Table VI—*Continued*

Panel C: Correlations with Actual and Expected Inflation

Maturity	Regime $s_t = 1$		Regime $s_t = 2$		Regime $s_t = 3$		Regime $s_t = 4$		Unconditional	
Qtrs	Actual	Expected	Actual	Expected	Actual	Expected	Actual	Expected	Actual	Expected
1	−0.34	−0.02	−0.47	−0.12	−0.40	0.03	−0.49	−0.06	−0.34	−0.03
	(0.31)	(0.24)	(0.31)	(0.25)	(0.37)	(0.28)	(0.35)	(0.29)	(0.29)	(0.31)
4	−0.11	0.14	−0.26	0.02	−0.17	0.16	−0.29	0.06	−0.13	0.16
	(0.43)	(0.32)	(0.45)	(0.36)	(0.56)	(0.44)	(0.55)	(0.46)	(0.43)	(0.44)
20	0.41	0.46	0.30	0.38	0.46	0.54	0.34	0.45	0.37	0.57
	(0.26)	(0.20)	(0.33)	(0.26)	(0.36)	(0.29)	(0.41)	(0.35)	(0.29)	(0.28)

A.2. The Correlation of Real Rates and Inflation.

Panel C of Table VI reports conditional and unconditional correlations of real rates and inflation. At the 1-quarter horizon, the conditional correlation of real rates with actual inflation is negative in all regimes and hence also unconditionally. The negative estimate for δ_π mostly drives this result. The correlations with expected inflation are smaller in absolute value, but still mostly negative. However, the differences across regimes are not large in economic terms and the correlations are overall not significantly different from zero. Consequently, we do not find strong statistical evidence for a Mundell–Tobin effect:

CLAIM 3: *The real short rate is negatively correlated with both expected and unexpected inflation, but the statistical evidence for a Mundell–Tobin effect is weak.*

This negative correlation between real rates and inflation is consistent with earlier studies such as Huizinga and Mishkin (1986) and Fama and Gibbons (1982), but their analysis implicitly assumes a zero inflation risk premium so their instrumented real rates may partially embed inflation risk premiums. The small Mundell–Tobin effect we estimate is consistent with Pennachi (1991), who uses a two-factor affine model of real rates and expected inflation, but opposite in sign to Barr and Campbell (1997), who use U.K. interest rates and find that the unconditional correlation between real rates and inflation is small but positive. As each regime records a negative correlation between real rates and inflation, we do not find any evidence that the sign of the correlation has changed over time, unlike what Goto and Torous (2006) find using an empirical model that neither employs term structure information nor precludes arbitrage.

The correlations between real yields and actual or expected inflation turn robustly positive at long horizons. Some of these correlations are statistically significant, although again most are not precisely estimated. The positive signs at long horizons result from the positive feedback effect of the Φ coefficients dominating the negative effect of the δ_π coefficient in the short rate equation. This indicates that the Mundell–Tobin effect is only a short-horizon phenomenon. Over long horizons, real yields and inflation are positively correlated.

A.3. The Effect of Regimes on Real Rates

Introducing regimes allows a further nonlinear mapping between latent factors and nominal yields not available in a traditional affine model, so that the dynamics of real long yields are not just linear transformations of nominal yield factors. To compare the effect of incorporating regimes, we contrast our model-implied real yields with those implied by model I_w. Figure 4 plots real yields from models I_w and IV^C, and we characterize the differences between the real yields from each model in Table VII.

Panel A of Table VII reports the population moments of real yields from models I_w and IV^C. The mean real short rate in model I_w is 1.42%, very close to the 1.39% mean of the 1-quarter real yield for a similar model estimated by Campbell and Viceira (2001). This is slightly higher, but very similar to the

Table VII

Effect of Regimes on Real Rates

The table reports various characteristics of real yields from model I_w, an affine model similar to Campbell and Viceira (2001), and our model IV^C. In Panel A we report population means, standard deviations, and autocorrelations of real 1-quarter short rates and real 20-quarter long yields, together with their correlation. Standard errors reported in parentheses are computed using the delta method. In Panel B, we report statistics on the differences between the real yields implied by model I_w and model IV^C over the sample.

Panel A: Real Yield Characteristics

		Model I_w	Model IV^C
Real Short Rate $\hat{r}_t$	Mean	1.42	1.24
		(0.31)	(0.38)
	St Dev	1.59	1.46
		(0.29)	(0.23)
	Auto	0.72	0.60
		(0.09)	(0.08)
Real Long Rate $\hat{y}_t^{20}$	Mean	1.69	1.32
		(0.30)	(0.40)
	St Dev	1.04	0.55
		(0.34)	(0.32)
	Auto	0.96	0.94
		(0.02)	(0.05)
Correlation $\hat{r}_t$, $\hat{y}_t^{20}$		0.79	0.64
		(0.08)	(0.06)

Panel B: Comparisons of I_w and IV^C over the Sample

Real Short Rate $\hat{r}_t$ Differences	Std Dev	1.40
	Min	−2.61
	Max	6.01
Real Long Rate $\hat{y}_t^{20}$ Differences	Std Dev	0.54
	Min	−1.06
	Max	1.85

mean level of short rates from our model IV^C, at 1.24%. The standard deviations of real short rates are also similar across the two models, at 1.59% and 1.46%, for models I_w and IV^C, respectively. However, Model I_w's real short rates are somewhat more persistent, at 0.72, than the autocorrelation of short rates from model IV^C, at 0.60. There are bigger differences for population moments for real long yields between the models. The long end of the real yield curve for model I_w is, on average, 40 basis points higher than for model IV^C and twice as variable, with standard deviations of 1.04% and 0.55%, respectively. The correlation between short and long real rates is higher for model I_w, at 0.79, than for model IV^C, at 0.64. Thus, the addition of regimes has important consequences for inferring long real rates.

Figure 4 plots the real short and long yields over the sample from the two models. The top panel shows that the real short rates from models I_w and

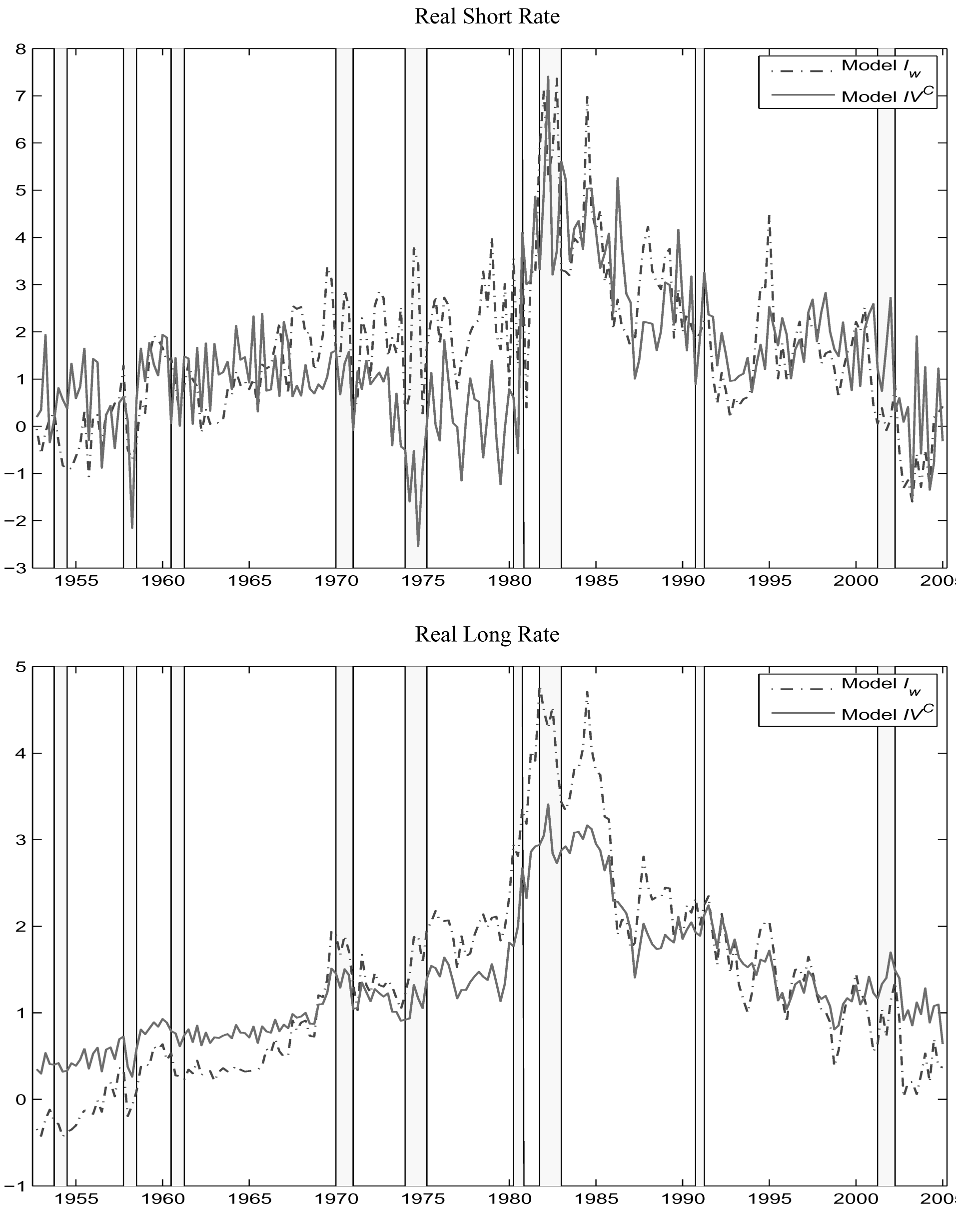

Figure 4. Comparing Model IV^C real yields with Model I_w. The figure compares the 1-quarter real short rate (5-year real long yield) of the benchmark model IV^C and model I_w in the top (bottom) panel over the sample period.

IV^C follow the same secular trends, but the correlation between the two model implied real rates is only 0.57. The main difference occurs during the late 1970s. Model IV^C documents that real short rates were fairly low during this period, consistent with the early estimates of Mishkin (1981) and Garcia and Perron (1996). In contrast, model I_w's real rates are much higher during this period. To quantify these differences, Panel B of Table VII reports summary statistics on the difference between $\hat{r}_t$ from model I_w and $\hat{r}_t$ from model IV^C. The largest difference of 6.01% occurs during the 1974 recession. In the bottom panel of Figure 4, we graph the real long yield from the two models. While the higher variability of the I_w-implied real long yield is apparent, the two models clearly share the same trends. In fact, the real long rates from the two models have a 0.95 correlation.

In a traditional affine model, there is a direct linear mapping between the latent factors and nominal yields, which may imply that real rates, which are linear combinations of the latent factors, are highly correlated with nominal yields. This is the case for model I_w. The bottom panel of Figure 4 shows that real long yields from model I_w start from below zero in 1952 and reach close to 5% in 1981, before declining to 30 basis points in 2005. These long real rates are highly correlated with long nominal rates, with a correlation coefficient of 0.98. Incorporating regimes in model IV^C reduces the correlation between real and nominal long rates to 90%. In contrast to model I_w, real long yields implied by model IV^C are more stable and have never been negative. This appears to be a more economically reasonable characterization of real long yields.

B. The Behavior of Inflation and Inflation Risk

B.1. The Term Structure of Expected Inflation

Table VIII reports some characteristics of inflation compensation, $\pi^e_{t,n}$, expected inflation, $\mathrm{E}_t(\pi_{t+n,n})$, and the inflation risk premium, $\varphi_{t,n}$. We focus first on the inflation compensation estimates. The most striking feature in Table VIII is that the term structure of inflation compensation slopes upwards in all regimes. Regime $s_t = 1$ is the normal regime, and in this regime the inflation compensation spread is $\pi^e_{t,20} - \pi^e_{t,1} = 1.17\%$, very close to the unconditional inflation compensation spread of 1.14%. In regimes $s_t = 2$ and $s_t = 4$, inflation compensation starts at a lower level because these are the regimes with downward-sloping real yield curves and a disinflationary environment. However, the inflation compensation spreads are roughly comparable to the unconditional compensation spread, at 1.34% and 1.16% for regimes $s_t = 2$ and $s_t = 4$, respectively. We report the term structure of expected inflation in the second panel of Table VIII. Expected inflation always approaches the unconditional mean of inflation as the horizon increases in all regimes, because inflation is a stationary process.

B.2. The Inflation Risk Premium

Since the term structure of inflation compensation is upward sloping but expected inflation converges to long-run unconditional expected inflation, the

Table VIII
Inflation Compensation, Expected Inflation, and Inflation Risk Premiums

The table reports means of inflation compensation, the difference between nominal and real yields, expected inflation, and the inflation risk premium all implied from the benchmark model IV^C. Standard errors reported in parentheses are computed using the delta method.

Qtrs	$s_t = 1$	$s_t = 2$	$s_t = 3$	$s_t = 4$	Unconditional
Inflation Compensation $\pi^e_{t,n}$					
1	3.92	2.46	4.43	3.20	3.94
	(0.38)	(0.78)	(0.39)	(0.67)	(0.38)
4	4.20	2.49	4.95	3.34	4.25
	(0.34)	(0.70)	(0.39)	(0.59)	(0.35)
20	5.09	3.80	5.45	4.36	5.08
	(0.41)	(0.45)	(0.43)	(0.42)	(0.38)
Expected Inflation $E_t(\pi_{t+n,n})$					
1	3.93	2.47	4.44	3.21	3.94
	(0.38)	(0.79)	(0.39)	(0.67)	(0.38)
4	3.89	2.63	4.48	3.47	3.94
	(0.38)	(0.73)	(0.41)	(0.65)	(0.38)
20	3.91	3.39	4.20	3.82	3.94
	(0.38)	(0.49)	(0.39)	(0.46)	(0.38)
Inflation Risk Premium $\varphi_{t,n}$					
4	0.31	−0.14	0.47	−0.13	0.31
	(0.09)	(0.06)	(0.15)	(0.09)	(0.10)
20	1.18	0.42	1.25	0.55	1.14
	(0.36)	(0.23)	(0.42)	(0.31)	(0.36)

increasing term structure of inflation compensation is due to an inflation risk premium:

CLAIM 4: *The model matches an unconditional upward-sloping nominal yield curve by generating an inflation risk premium that is increasing in maturity.*

The third panel of Table VIII reports statistics on the inflation risk premium $\varphi_{t,n}$. In the normal regime $s_t = 1$ and unconditionally, the 5-year inflation risk premium is around 1.15%, which is almost the same magnitude as the 5-year term spread generated by the model of 1.21%. The inflation risk premium is higher in regime $s_t = 3$ with higher and more variable inflation than in regime $s_t = 1$. In the high real rate regimes $s_t = 2$ and $s_t = 4$, the inflation risk premium is less than 55 basis points. In regime $s_t = 4$, the inflation risk premium is not statistically different from zero. In Campbell and Viceira's (2001) one-regime setting, $\varphi_{t,40}$ is approximately 0.42%, accounting for about half of their model-implied 40-quarter nominal term spread of 0.88%.[11] We obtain inflation risk premiums of this low magnitude only in high real rate regimes, and in normal

[11] Campbell and Viceira (2001) report that the difference in expected holding-period returns on 10-year nominal bonds over nominal 3-month T-bills in excess of the expected holding-period returns on 10-year real bonds over the real 3-month short rate is approximately 1.1% and define

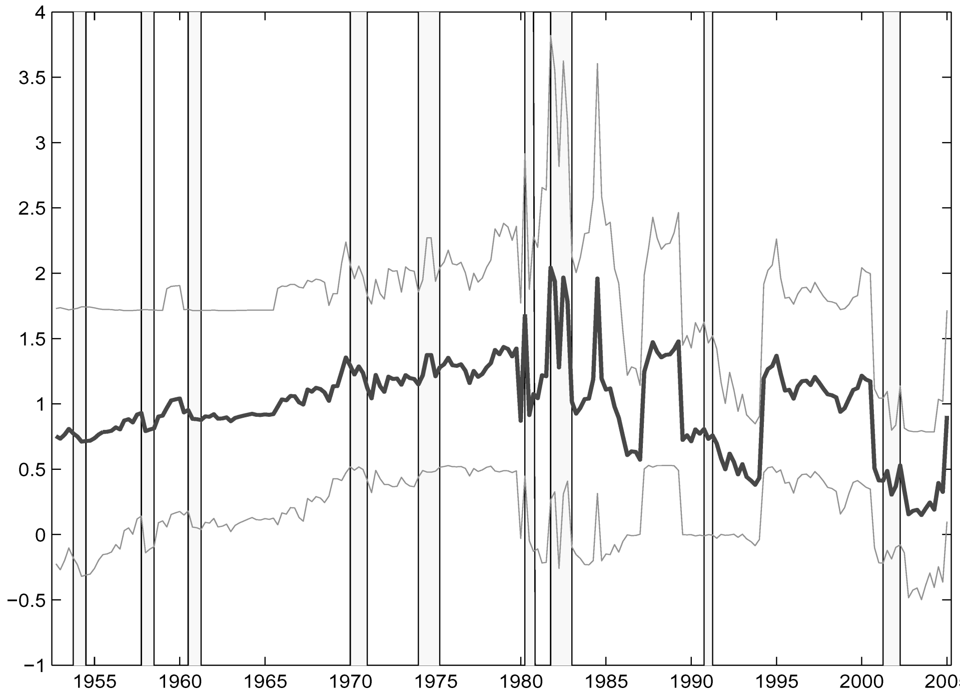

Figure 5. Inflation risk premiums. The figure graphs the time-series of the 20-quarter inflation risk premium, $\varphi_{t,20}$, with two standard error bounds, implied from the benchmark model IV^C. NBER recessions are indicated by shaded bars.

times assign almost all of the positive nominal yield spread to inflation risk premiums.

The time variation of the inflation risk premium is correlated with the time variation of the price of risk factor, q_t, but the correlation of the inflation risk premium with q_t is small, at 9.5% for a 20-quarter maturity. To calculate the proportion of the variance of $\varphi_{t,20}$ due to regime changes, we compare the unconditional variance of $\varphi_{t,20}$ varying across all four regimes with the variance of $\varphi_{t,20}$ if the model never switched from $s_t = 1$. We find that a significant fraction, namely 40%, of the variation of $\varphi_{t,20}$ is due to regime changes.

Figure 5 graphs the 20-quarter inflation risk premium over time and shows that the inflation risk premium decreased in every recession. During the 1981 to 1983 recession, the inflation premium is very volatile, increasing and decreasing by over 75 basis points. The general trend is that the premium rose very gradually from the 1950s until the late 1970s before entering a very volatile period during the monetary targeting period from 1979 to the early 1980s. It is then that the premium reached a peak of 2.04%. While the trend since then has been downward, there have been large swings in the premium. From a

this to be the inflation risk premium. In our model, the corresponding number for this quantity at a 20-quarter maturity is $\mathrm{E}[\ln(P^{19}_{t+1}/P^{20}_t) - y^1_t] - \mathrm{E}[\ln(\hat{P}^{19}_{t+1}/\hat{P}^{20}_t) - \hat{r}_t] = 1.46\%$.

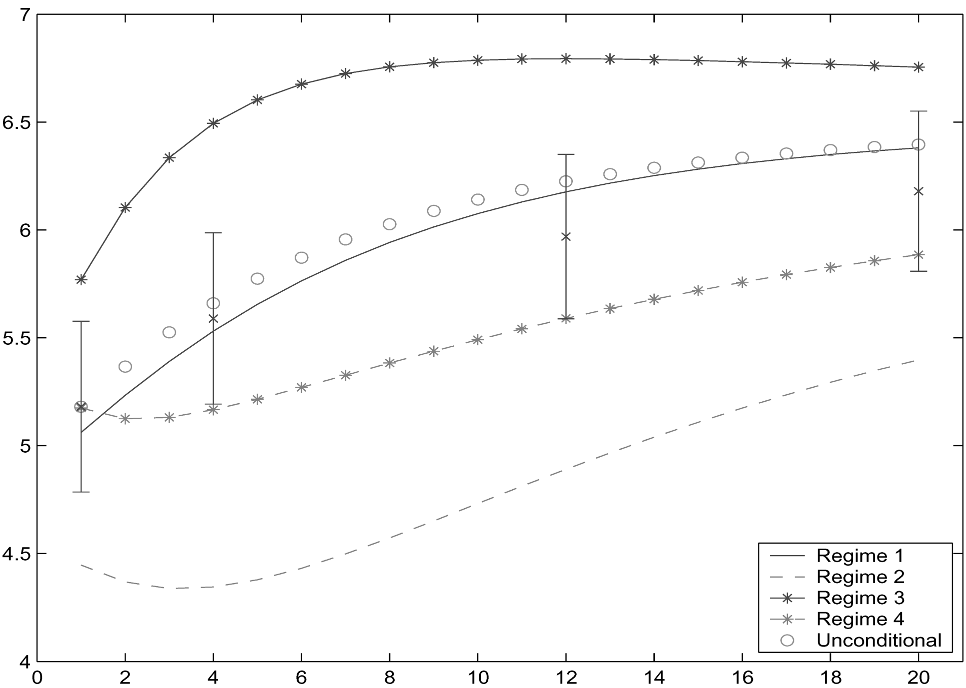

Figure 6. Nominal term structure The figure graphs the nominal yield curve, conditional on each regime and the unconditional nominal yield curve from the benchmark model IV^C. The x-axis displays maturities in quarters of a year. The y-axis units are annualized and are in percentages. Average yields from data are represented by "x," with 95% confidence intervals represented by vertical bars.

temporary low of 50 basis points in the mid-1980s it shot above 1%, coinciding with the halting of the large dollar appreciation of the early 1980s. The inflation premium dropped back to around 50 basis points in the late eighties and reached a low of 0.38% in 1993. The sharp drops in the inflation risk premium coincide with transitions to regimes with high real short rates. During 1994, the premium shot back up to 1.37% at the same time the Federal Reserve started to raise interest rates. During the late 1990s bull market inflation risk premiums were fairly stable and declined to 0.15% after the 2001 recession when there were fears of deflation. At the end of the sample in December 2004, the inflation risk premium started to increase again edging close to 1%.

C. Nominal Term Structure

Figure 6 graphs the average nominal yield curve. The unconditional yield curve is upward sloping, with the slope flattening out for longer maturities. The benchmark model produces a nominal term spread of $y_t^{20} - y_t^1 = 1.21\%$, well inside a one-standard error bound of the 1.00% term spread in data. Strikingly, in no regime does the benchmark model generate a conditional downward-sloping

Table IX

Conditional Moments across NBER Business Cycles

The table reports various sample moments of real rates, nominal rates, and inflation compensation from the benchmark model IV^C, conditional on expansions and recessions as defined by the NBER. Standard errors reported in parentheses are computed using the delta method on sample moments.

		Mean		Std Dev	
	Maturity Qtrs	Expansion	Recession	Expansion	Recession
Real Rates $\hat{y}_t^n$	1	1.45	1.23	1.30	2.06
		(0.20)	(0.20)	(0.04)	(0.08)
	20	1.33	1.43	0.65	0.87
		(0.38)	(0.38)	(0.18)	(0.25)
Nominal Rates y_t^n	1	5.03	5.95	2.59	4.07
		(0.09)	(0.14)	(0.27)	(0.41)
	20	6.05	6.85	2.46	3.71
		(0.20)	(0.22)	(0.26)	(0.38)
Inflation Compensation $\pi_{t,n}^e$	1	3.57	4.73	2.23	3.62
		(0.19)	(0.17)	(0.18)	(0.28)
	20	4.72	5.42	1.89	2.93
		(0.37)	(0.39)	(0.38)	(0.57)

nominal yield curve. In regimes $s_t = 2$ and $s_t = 4$, the real rate term structure is downward-sloping, but the upward-sloping term structure of inflation risk premiums completely counteracts this effect. Thus, regimes are important for the shape of the real, not nominal yield curve.

The first regime (low real rate-normal inflation regime) displays a nominal yield curve that almost matches the unconditional term structure. In the second regime, the yield curve is shifted downwards but is more steep because rates are lower than in the first regime due to lower expected inflation and inflation risk. In the third regime, the term structure is steeply upward sloping at the short end but then becomes flat and slightly downward sloping for maturities extending beyond 10 quarters. Nominal interest rates are the highest in this regime because in this regime, expected inflation is high and the level of real rates is about the same as in regime 1. In regime 4, the real interest rate curve is downward sloping, starting at a high level. Inflation compensation, however, is low in this regime (resulting in nominal yields of an average level), and is upward sloping, making the nominal yield curve upward sloping on average. Yet, in both regimes 2 and 4, a slight J-curve effect is visible at short maturities with nominal rates decreasing slightly before starting to increase.

Interest rates are often associated with the business cycle. The business cycle dates reported by the NBER are regarded as benchmark dates by both academics and practitioners. According to the conventional wisdom, interest rates are procyclical and spreads countercyclical (see, for example, Fama (1990)). Table IX shows that this is incorrect when measuring business cycles using NBER recessions and expansions. Interest rates are overall larger during

Table X
Unconditional Variance Decomposition of Nominal Yields

The table reports unconditional variance decompositions of nominal yields, y_t^n, into real rate, expected inflation, and inflation risk premium components, denoted by $\hat{y}_t^n$, $\mathrm{E}_t(\pi_{t,n})$, and $\varphi_{t,n}$, respectively, implied from model IV^C. This is done using the equation:

$$1 = \frac{\mathrm{var}(y_t^n, y_t^n)}{\mathrm{var}(y_t^n)} = \frac{\mathrm{cov}(\hat{y}_t^n, y_t^n) + \mathrm{cov}(\mathrm{E}_t(\pi_{t+n,n}), y_t^n) + \mathrm{cov}(\varphi_{t,n}, y_t^n)}{\mathrm{var}(y_t^n)}.$$

Standard errors reported in parentheses are computed using the delta method on population moments.

Maturity Qtrs	Real Rates	Expected Inflation	Inflation Risk
1	0.20	0.80	0.00
	(0.09)	(0.09)	(0.00)
20	0.20	0.71	0.10
	(0.09)	(0.09)	(0.08)

NBER recessions. However, when we focus on real rates, the conventional story obtains:

CLAIM 5: *Nominal interest rates do not behave procyclically across NBER business cycles but our model-implied real rates do.*

This can only be the case if expected inflation is countercyclical. Table IX shows that this is indeed the case, with inflation compensation averaging 4.73% in recessions but only 3.57% in expansions. Veronesi and Yared (1999) also find that real rates are procyclical in an RS model. In contrast, the real rates implied by model I_w are actually countercyclical, averaging 1.58% (1.80%) across NBER expansions (recessions). Thus, the presence of the regimes helps to induce the procyclical behavior of real rates. Finally, Table IX also illustrates that recessions are characterized by more volatility in real rates, nominal rates, and inflation.

D. Variance Decompositions

Table X reports the population variance decomposition of the nominal yield into real rates and inflation compensation. The variance decompositions, conditioning on the regime, are very similar across regimes and so are not reported. The results show that:

CLAIM 6: *The decompositions of nominal yields into real yields and inflation components at various horizons indicate that variation in inflation compensation (expected inflation and inflation risk premia) explains about 80% of the variation in nominal rates at both short and long maturities.*

This is at odds with the conventional wisdom that expected inflation primarily affects long-term bonds (see, among others, Fama (1975) and Mishkin

Table XI

Unconditional Variance Decomposition of Nominal Yield Spreads

The table reports unconditional variance decompositions of nominal yield spreads, $y_t^n - y_t^1$, into real rate, expected inflation, and inflation risk premium components, denoted by $\hat{y}_t^n - \hat{r}_t$, $\mathrm{E}_t(\pi_{t+n,n}) - \mathrm{E}_t(\pi_{t+1})$, and $\varphi_{t,n}$, respectively, implied from model IV^C. This is done using the equation:

$$1 = \frac{\mathrm{var}(y_t^n - y_t^1, y_t^n - y_t^1)}{\mathrm{var}(y_t^n - y_t^1)} = \frac{\mathrm{cov}(\hat{y}_t^n - \hat{r}_t, y_t^n - y_t^1) + \mathrm{cov}(\mathrm{E}_t(\pi_{t+n,n}) - \mathrm{E}_t(\pi_{t+1}), y_t^n - y_t^1) + \mathrm{cov}(\varphi_{t,n}, y_t^n - y_t^1)}{\mathrm{var}(y_t^n - y_t^1)}.$$

Standard errors reported in parentheses are computed using the delta method on population moments.

Panel A: Unconditional

Maturity Qtrs	Real Rates	Expected Inflation	Inflation Risk
4	0.44	0.56	−0.01
	(0.15)	(0.15)	(0.00)
20	0.19	0.85	−0.05
	(0.18)	(0.18)	(0.02)

Panel B: Conditional on Regime

Maturity Qtrs	Real Rates	Expected Inflation	Inflation Risk	Real Rates	Expected Inflation	Inflation Risk
	Regime $s_t = 1$			Regime $s_t = 2$		
4	0.14	0.87	−0.01	0.08	0.93	−0.01
	(0.19)	(0.19)	(0.00)	(0.22)	(0.22)	(0.00)
20	0.04	1.03	−0.08	−0.02	1.07	−0.05
	(0.20)	(0.20)	(0.03)	(0.22)	(0.22)	(0.03)
	Regime $s_t = 3$			Regime $s_t = 4$		
4	0.69	0.32	−0.00	0.64	0.36	−0.00
	(0.12)	(0.12)	(0.00)	(0.13)	(0.13)	(0.00)
20	0.31	0.71	−0.02	0.29	0.73	−0.02
	(0.16)	(0.16)	(0.01)	(0.17)	(0.17)	(0.01)

(1981)). The single-regime model I_w attributes even less of the variance of long-term yields to inflation components: At a 20-quarter maturity, variation in real yields accounts for 37% of movements in nominal rates compared to 28% at a 1-quarter maturity. This may be caused by the poor match of inflation dynamics using an affine model calibrated to inflation data. Pennachi's (1991) affine model identifies expected inflation from survey data and finds that expected inflation and inflation risk show little variation across horizons. Table X also reports that the inflation risk premium accounts for 10% of the variance of a 20-quarter maturity nominal yield.

In Table XI, we decompose the variation of nominal term spreads into real rate, expected inflation, and inflation risk premium components. Unconditionally, inflation components account for 55% of the 4-quarter term spread and 80% of the 20-quarter term spread variance. For term spreads, inflation shocks only dominate at the long end of the yield curve. In the regimes with relatively stable real rates (regimes 1 and 2), inflation components account for over 100%

of the variance of long-term spreads. In regimes 3 and 4, real rates are more volatile, and expected inflation accounts for approximately 35% of the variation in the 4-quarter term spread, increasing to over 70% for the 20-quarter term spread. Hence, the conventional wisdom that inflation is more important for the long end of the yield curve holds, not for the level of yields, but for term spreads:

CLAIM 7: *Inflation compensation is the main determinant of nominal interest rate spreads at long horizons*.

The intuition behind this result is that the long and short ends of nominal yields have large exposure to common factors, including the factors driving inflation and inflation risk. It is only after controlling for an average effect, or by computing a term spread, that we can observe relative differences at different parts of the yield curve. Thus, only after computing the term spread do we isolate the factors that differentially affect long yields controlling for the short rate exposure. The finding that inflation components are the main driver of term spreads is not dependent on having regimes in the term structure model. Mishkin (1990, 1992) finds consistent evidence with simple regressions using inflation changes and term spreads, as do Ang, Dong, and Piazzesi (2006) in a single-regime affine model. In model I_w, the attribution of the unconditional variance of the 20-quarter term spread to inflation is also close to 100%.

IV. Conclusion

In this article, we develop a term structure model that embeds regime switches in both real and nominal factors, and incorporates time-varying prices of risk. The model that provides the best fit with the data has correlated regimes coming from separate real factor and inflation sources.

We find that the real rate curve is fairly flat but slightly humped, with an average real rate of around 1.3%. The real short rate has an unconditional variability of 1.46% and has an autocorrelation of 60%. In some regimes, the real rate curve is downward sloping. In these regimes, expected inflation is low. The term structure of inflation compensation, the difference between nominal and real yields, is upward sloping. This is due to an upward-sloping inflation risk premium, which is unconditionally 1.14% on average. We find that expected inflation and inflation risk account for 80% of the variation in nominal yields at both short and long maturities. However, nominal term spreads are primarily driven by changes in expected inflation, particularly during normal times.

It is interesting to note that our results are qualitatively consistent with Roll's (2004) analysis on TIPS data, over the very short sample period since TIPS began trading. Consistent with our results, Roll also finds that the nominal yield curve is more steeply sloped than the real curve, which is also mostly flat over our overlapping sample periods. Roll also shows direct evidence of an inflation premium that increases with maturity.

Our work here is only the beginning of a research agenda. In future work, we could use our model to link the often-discussed deviations from the Expectations Hypothesis (see, for example, Campbell and Shiller (1991)) to deviations from the Fisher Hypothesis (Mishkin (1992)). Although we have made one step in the direction of identifying the economic sources of regime switches in interest rates, more could be done. In particular, a more explicit examination of the role of business cycle variation and changes in monetary policy as sources of the regime switches is an interesting topic for further research.

Appendix A. Real Bond Prices

Let N_1 be the number of unobserved state variables in the model ($N_1 = 3$ for the stochastic inflation model, $N_1 = 2$ otherwise) and $N = N_1 + 1$ be the total number of factors, including inflation. The following proposition describes how our model implies closed-form real bond prices.

PROPOSITION A: *Let $X_t = (q_t\ f_t\ \pi_t)'$ or $X_t = (q_t\ f_t\ w_t\ \pi_t)'$ follow (2), with the real short rate (4) and real pricing kernel (5) with prices of risk (6). The regimes s_t follow a Markov chain with transition probability matrix $\Pi = \{p_{ij}\}$. Then the real zero-coupon bond price for period n conditional on regime i, $\widehat{P}_t^n(s_t = i)$, is given by*

$$\widehat{P}_t^n(i) = \exp(\widehat{A}_n(i) + \widehat{B}_n X_t). \tag{A1}$$

The scalar $\widehat{A}_n(i)$ is dependent on regime $s_t = i$ and $\widehat{B}_n$ is a $1 \times N$ vector that is partitioned as $\widehat{B}_n = [\widehat{B}_{nq}\ \widehat{B}_{nx}]$, where $\widehat{B}_{nq}$ corresponds to the q variable and $\widehat{B}_{nx}$ corresponds to the other variables in X_t. The coefficients $\widehat{A}_n(i)$ and $\widehat{B}_n$ are given by

$$\begin{aligned}\widehat{A}_{n+1}(i) &= -\left(\delta_0 + \widehat{B}_{nq}'\sigma_q\gamma_0\right) + \log\sum_j p_{ij}\exp\Big(\widehat{A}_n(j) + \widehat{B}_n\mu(j) \\ &\qquad -\widehat{B}_{nx}\Sigma_x(j)\lambda(j) + \frac{1}{2}\widehat{B}_n\Sigma(j)\Sigma(j)'\widehat{B}_n'\Big) \\ \widehat{B}_{n+1} &= -\delta_1' + \widehat{B}_n\Phi - \widehat{B}_{nq}\sigma_q\gamma_1 e_1',\end{aligned} \tag{A2}$$

where e_i denotes a vector of zeros with a "1" in the ith place and $\Sigma_x(i)$ refers to the lower $N_1 \times N_1$ matrix of $\Sigma(i)$ corresponding to the non-q_t variables in X_t. The starting values for $\widehat{A}_n(i)$ and $\widehat{B}_n$ are

$$\begin{aligned}\widehat{A}_1(i) &= -\delta_0 \\ \widehat{B}_1 &= -\delta_1'.\end{aligned} \tag{A3}$$

Proof: We first derive the initial values in (A3):

$$\begin{aligned} P_t^1(i) &= \textstyle\sum_j p_{ij} \mathrm{E}_t\left[\widehat{M}_{t+1} | S_{t+1} = j\right] \\ &= \sum_j p_{ij} \exp\left(-r_t - \frac{1}{2}\lambda_t(j)' \lambda_t(j) - \lambda_t(j)' \varepsilon_{t+1}\right) \\ &= \exp\left(-\delta_0 - \delta_1' X_t\right). \end{aligned} \tag{A4}$$

Hence,

$$\widehat{P}_t^1(i) = \exp(\widehat{A}_1(i) + \widehat{B}_1 X_t),$$

where $A_1(i)$ and B_1 take the form in (A3).

We prove the recursion (A2) by induction. We assume that (A1) holds for maturity n and examine $\widehat{P}_t^{n+1}(i)$:

$$\begin{aligned} \widehat{P}_t^{n+1}(i) &= \sum_j p_{ij} \mathrm{E}_t \exp\left[-r_t - \frac{1}{2}\lambda_t(j)' \lambda_t(j) - \lambda_t(j)' \varepsilon_{t+1} + \widehat{A}_n(j) + \widehat{B}_n X_{t+1}\right], \\ &= \sum_j p_{ij} \mathrm{E}_t \exp\left[-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t(j)' \lambda_t(j) - \lambda_t(j)' \varepsilon_{t+1} + \widehat{A}_n(j)\right. \\ &\quad \left. + \widehat{B}_n\left(\mu(j) + \Phi X_t + \Sigma(j) \varepsilon_{t+1}\right)\right]. \end{aligned} \tag{A5}$$

Evaluating the expectation, we have

$$\begin{aligned} \widehat{P}_t^{n+1}(i) &= \sum_j p_{ij} \exp\left[-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t(j)' \lambda_t(j) + \widehat{A}_n(j) + \widehat{B}_n \mu(j)\right. \\ &\quad \left. + \widehat{B}_n \Phi X_t + \frac{1}{2}\left(\widehat{B}_n \Sigma(j) - \lambda_t(j)'\right)\left(\widehat{B}_n \Sigma(j) - \lambda_t(j)'\right)'\right] \\ &= \exp\left[-\delta_0 + \left(\widehat{B}_n \Phi - \delta_1'\right) X_t\right] \\ &\quad \times \sum_j p_{ij} \exp\left[\widehat{A}_n(j) + \widehat{B}_n \mu(j) - \widehat{B}_n \Sigma(j) \lambda_t(j) + \frac{1}{2}\widehat{B}_n \Sigma(j) \Sigma(j) \widehat{B}_n'\right]. \end{aligned} \tag{A6}$$

But, we can write

$$\begin{aligned} \widehat{B}_n \Sigma(j) \lambda_t(j) &= [\widehat{B}_{nq}\ \widehat{B}_{nx}] \begin{bmatrix} \sigma_q\left(\gamma_0 + \gamma_1 e_1' X_t\right) \\ \Sigma_x(j) \lambda(j) \end{bmatrix} \\ &= \widehat{B}_{nq} \sigma_q\left(\gamma_0 + \gamma_1 e_1' X_t\right) + \widehat{B}_{nx} \Sigma_x(j) \lambda(j). \end{aligned} \tag{A7}$$

Expanding and collecting terms, we can then write

$$\widehat{P}_t^n(i) = \exp(\widehat{A}_n(i) + \widehat{B}_n X_t),$$

where $\widehat{A}_n(i)$ and $\widehat{B}_n$ take the form of (A2). Q.E.D.

Appendix B. Nominal Bond Prices

Following the notation of Appendix A, let N_1 be the number of unobserved state variables in the model ($N_1 = 3$ for the stochastic inflation model, $N_1 = 2$ otherwise) and $N = N_1 + 1$ be the total number of factors including inflation. The following proposition describes how our model implies closed-form nominal bond prices.

PROPOSITION B: *Let $X_t = (q_t\ f_t\ \pi_t)'$ or $X_t = (q_t\ f_t\ w_t\ \pi_t)'$ follow (2), with the real short rate (4) and real pricing kernel (5) with prices of risk (6). The regimes s_t follow a Markov chain with transition probability matrix $\Pi = \{p_{ij}\}$. Then the nominal zero-coupon bond price for period n conditional on regime i, $P_t^n\ (s_t = i)$, is given by:*

$$P_t^n(i) = \exp(A_n(i) + B_n X_t), \tag{B1}$$

where the scalar $A_n(i)$ is dependent on regime $s_t = i$ and B_n is an $N \times 1$ vector:

$$\begin{aligned} A_{n+1}(i) = &-\left(\delta_0 + B'_{nq}\sigma_q\gamma_0\right) + \log\sum_j p_{ij}\exp\Big(A_n(j) + \left(B_n - e'_N\right)\mu(j) \\ &-\left(B_{nx} - e'_{N_1}\right)\Sigma_x(j)\lambda(j) + \frac{1}{2}\left(B_n - e'_N\right)\Sigma(j)\Sigma(j)\left(B_n - e'_N\right)'\Big) \\ B_{n+1} = &-\delta'_1 + \left(B_n - e'_N\right)\Phi - B_{nq}\sigma_q\gamma_1 e'_1, \end{aligned} \tag{B2}$$

where e_i denotes a vector of zeros with a "1" in the ith place, $A(i)$ is a scalar dependent on regime $s_t = i$, B_n is a row vector, which is partitioned as $B_n = [B_{nq}\ B_{nx}]$, where B_{nq} corresponds to the q variable, and $\Sigma_x(i)$ refers to the lower $N_1 \times N_1$ matrix of $\Sigma(i)$ corresponding to the non-q_t variables in X_t. The starting values for $A_n(i)$ and B_n are

$$\begin{aligned} A_1(i) &= -\delta_0 + \log\sum_j p_{ij}\exp\left(-e'_N\mu(j) + \frac{1}{2}e'_N\Sigma(j)\Sigma(j)'e_N + e'_{N_1}\Sigma_x(j)\lambda(j)\right) \\ B_1 &= -\left(\delta'_1 + e'_N\Phi\right). \end{aligned} \tag{B3}$$

Proof: We first derive the initial values (B3) by directly evaluating

$$\begin{aligned} P_t^1(i) &= \sum_j p_{ij}E_t\left[\widehat{M}_{t+1}|S_{t+1} = j\right] \\ &= \sum_j p_{ij}\exp\Big(-r_t - \frac{1}{2}\lambda_t(j)'\lambda_t(j) - \lambda_t(j)'\varepsilon_{t+1} - e'_N\left(\mu(j)\right. \\ &\quad \left.+\Phi X_t + \Sigma(j)\varepsilon_{t+1}\right)\Big) \\ &= \exp\left(-\delta_0 - \delta'_1 X_t - e'_N\Phi X_t\right) \\ &\quad \times\sum_j p_{ij}\exp\left(-e'_N\mu(j) - e'_N\Sigma(j)\varepsilon_{t+1} - \frac{1}{2}\lambda_t(j)'\lambda_t(j) - \lambda_t(j)'\varepsilon_{t+1}\right) \\ &= \exp\left(-\delta_0 - \delta'_1 X_t - e'_N\Phi X_t\right) \\ &\quad \times\sum_j p_{ij}\exp\left(-e'_N\mu(j) + \frac{1}{2}e'_N\Sigma(j)\Sigma(j)'e_N + e'_N\Sigma(j)\lambda_t(j)\right). \end{aligned} \tag{B4}$$

Note that $e'_N \Sigma(j)\lambda_t(j) = e'_{N_1}\Sigma_x(j)\lambda(j)$. Hence

$$P_t^1(i) = \exp\left(A_1(i) + B_1 X_t\right),$$

where $A_1(i)$ and B_1 are given by (B3).

To prove the general recursion we use proof by induction:

$$\begin{aligned}
P_t^{n+1}(i) &= \sum_j p_{ij} E_t\Bigg[\exp\left(-r_t - \frac{1}{2}\lambda_t(j)'\lambda_t(j) - \lambda_t(j)'\varepsilon_{t+1} - e'_N X_{t+1}\right) \\
&\qquad \exp\left(A_n(j) + B_n X_{t+1}\right)\Bigg] \\
&= \sum_j p_{ij} E_t\Bigg[\exp\Bigg(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t(j)'\lambda_t(j) - \lambda_t(j)'\varepsilon_{t+1} + A_n(j) \\
&\qquad + \left(B_n - e'_N\right)\left(\mu(j) + \Phi X_t + \Sigma(j)\varepsilon_{t+1}\right)\Bigg)\Bigg] \\
&= \sum_j p_{ij} \exp\Bigg(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t(j)'\lambda_t(j) + A_n(j) + \left(B_n - e'_N\right)\mu(j) \\
&\qquad + \left(B_n - e'_N\right)\Phi X_t + \frac{1}{2}\left(\left(B_n - e'_N\right)\Sigma(j) - \lambda_t(j)'\right) \\
&\qquad \times \left(\left(B_n - e'_N\right)\Sigma(j) - \lambda_t(j)'\right)'\Bigg) \\
&= \exp\left(-\delta_0 + \left(\left(B_n - e'_N\right)\Phi - \delta_1'\right)X_t\right)\sum_j p_{ij}\exp\Bigg(A_n(j) + \left(B_n - e'_N\right)\mu(j) \\
&\qquad - \left(B_n - e'_N\right)\Sigma(j)\lambda_t(j) + \frac{1}{2}\left(B_n - e'_N\right)\Sigma(j)\Sigma(j)\left(B_n - e'_N\right)'\Bigg).
\end{aligned} \tag{B5}$$

Now note that

$$\begin{aligned}
\left(B_n - e'_N\right)\Sigma(j)\lambda_t(j) &= \left(B_n - e'_N\right)\begin{bmatrix}\sigma_q\left(\gamma_0 + \gamma_1 e_1' X_t\right) \\ \Sigma_x(j)\lambda(j)\end{bmatrix} \\
&= \begin{bmatrix} B_{nq} \\ B_{nx} - e'_{N_1}\end{bmatrix}\begin{bmatrix}\sigma_q\left(\gamma_0 + \gamma_1 e_1' X_t\right) \\ \Sigma_x(j)\lambda(j)\end{bmatrix} \\
&= B_{nq}\sigma_q\left(\gamma_0 + \gamma_1 e_1' X_t\right) + \left(B_{nx} - e'_{N_1}\right)\Sigma_x(j)\lambda(j),
\end{aligned} \tag{B6}$$

where $B_n = [B_{nq} \quad B_{nx}]$.

Hence, collecting terms and substituting (B6) into (B5), we have:

$$P_t^{n+1}(i) = \exp\left[A_{n+1}(i) + B_{n+1} X_t\right],$$

where $A_n(i)$ and B_n are given by (B2). Q.E.D.

Appendix C. Likelihood Function and Identification

A. Likelihood Function

We specify the set of nominal yields without measurement error as $Y_{1t}(N_1 \times 1)$ and the remaining yields as $Y_{2t}(N_2 \times 1)$. There are as many yields measured without error as there are latent factors in X_t. The complete set of yields are denoted as $Y_t = (Y'_{1t}\, Y'_{2t})'$ with dimension $M \times 1$, where $M = N_1 + N_2$. Note that the total number of factors in X_t is $N = N_1 + 1$, since the last factor, inflation, is observable.

Given the expression for nominal yields in (11), the yields observed without error and inflation, $Z_t = (Y'_{1t}\, \pi_t)'$, take the form

$$Z_t = \mathcal{A}_1(s_t) + \mathcal{B}_1 X_t, \tag{C1}$$

$$\mathcal{A}_1(s_t) = \begin{bmatrix} \mathcal{A}_n(s_t) \\ 0 \end{bmatrix}, \qquad \mathcal{B}_1 = \begin{bmatrix} \mathcal{B}_n \\ e'_N \end{bmatrix}, \tag{C2}$$

where $\mathcal{A}_n(s_t)$ is the $N_1 \times 1$ vector stacking the $-\ A_n(s_t)/n$ terms for the N_1 yields observed without error, and $\mathcal{B}_n$ is a $N_1 \times N$ matrix that stacks the $-B_n/n$ vectors for the two yields observed without error. Then we can invert for the unobservable factors:

$$X_t = \mathcal{B}^{-1}(Z_t - \mathcal{A}_1(s_t)). \tag{C3}$$

Substituting this into (C1) and using the dynamics of X_t in (2), we can write

$$Z_t = c(s_t, s_{t-1}) + \Psi Z_{t-1} + \Omega(s_t)\epsilon_t, \tag{C4}$$

where

$$\begin{aligned} c(s_t, s_{t-1}) &= \mathcal{A}_1(s_t) + \mathcal{B}_1\mu(s_t) - \mathcal{B}_1\Phi\mathcal{B}_1^{-1}\mathcal{A}_1(s_{t-1}) \\ \Psi &= \mathcal{B}_1\Phi\mathcal{B}_1^{-1} \\ \Omega(s_t) &= \mathcal{B}_1\Sigma(s_t). \end{aligned}$$

Note that our model implies an RS-VAR for the observable variables with complex cross-equation restrictions.

The yields Y_{2t} observed with error have the form

$$Y_{2t} = \mathcal{A}_2(s_t) + \mathcal{B}_2 X_t + u_t, \tag{C5}$$

where $\mathcal{A}_2$ and $\mathcal{B}_2(s_t)$ follow from Proposition B and u is the measurement error, $u_t \sim N(0, V)$, where V is a diagonal matrix. We can solve for u_t in equation (C5) using the inverted factor process (C3). We assume that u_t is uncorrelated with the errors ε_t in (2).

Following Hamilton (1994), we redefine the states s_t^* to count all combinations of s_t and s_{t-1}, with the corresponding redefined transition probabilities $p_{ij}^* =$

$p(s^*_{t+1} = i \mid s^*_t = j)$. We rewrite (C4) and (C5) as:

$$\begin{aligned} Z_t &= c(s^*_t) + \Psi Z_{t-1} + \Omega(s^*_t)\epsilon_t, \\ Y_{2t} &= \mathcal{A}_2(s^*_t) + \mathcal{B}_2 X_t + u_t. \end{aligned} \tag{C6}$$

Now the standard Hamilton (1989, 1994) and Gray (1996) algorithms can be used to estimate the likelihood function. Since (C6) gives us the conditional distribution $f(\pi_t, Y^1_t \mid s^*_t = i, I_{t-1})$, we can write the likelihood as:

$$\begin{aligned} \mathcal{L} &= \prod_t \sum_{s^*_t} f(\pi_t, Y_{1t}, Y_{2t} \mid s^*_t, I_{t-1}) Pr(s^*_t \mid I_{t-1}) \\ &= \textstyle\prod_t \sum_{s^*_t} f(Z_t \mid s^*_t, I_{t-1}) f(Y_{2t} \mid \pi_t, Y_{1t}, s^*_t, I_{t-1}) Pr(s^*_t \mid I_{t-1}), \end{aligned} \tag{C7}$$

where

$$\begin{aligned} f(Z_t \mid s^*_t, I_{t-1}) &= (2\pi)^{-(N_1+1)/2} |\Omega(s^*_t)\Omega(s^*_t)'|^{-1/2} \\ &\quad \exp\left(-\frac{1}{2}(Z_t - c(s^*_t) - \Psi Z_{t-1})'[\Omega(s^*_t)\Omega(s^*_t)\prime]^{-1}(Y_{2t} - c(s^*_t) - \Psi Z_{t-1})\right) \end{aligned}$$

is the probability density function of Z_t conditional on s^*_t and

$$\begin{aligned} & f(Y_{2t} \mid \pi_t, Y_{1t}, s^*_t, I_{t-1}) \\ &\quad = (2\pi)^{-N_2/2} |V|^{-1/2} \exp\left(-\frac{1}{2}(Y_{2t} - \mathcal{A}_2(s^*_t) - \mathcal{B}_2 X_t)' V^{-1} (Y_{2t} - \mathcal{A}_2(s^*_t) - \mathcal{B}_2 X_t)\right) \end{aligned}$$

is the probability density function of the measurement errors conditional on s^*_t.

The ex ante probability $Pr(s^*_t = i \mid I_{t-1})$ is given by

$$Pr(s^*_t = i \mid I_{t-1}) = \sum_j p^*_{ji} Pr(s^*_{t-1} = j \mid I_{t-1}), \tag{C8}$$

which is updated using

$$\begin{aligned} Pr(s^*_t = j \mid I_t) &= \frac{f(Z_t, s^*_t = j \mid I_{t-1})}{f(Z_t \mid I_{t-1})} \\ &= \frac{f(Z_t \mid s^*_t = j, I_{t-1}) Pr(s^*_t = j \mid I_{t-1})}{\sum_k f(Z_t \mid s^*_t = k, I_{t-1}) Pr(s^*_t = k \mid I_{t-1})}. \end{aligned}$$

An alternative way to derive the likelihood function is to substitute (C3) into (C5). We then obtain an RS-VAR with complex cross-equation restrictions for all variables in the system $(Z_t{}' \, Y_{2t}{}')'$. Note that unlike a standard affine model, the likelihood is not simply the likelihood of the yields measured without error multiplied by the likelihood of the measurement errors. Instead, the regime variables must be integrated out of the likelihood function.

B. Identification

There are two identification problems. First, there are the usual identification conditions that must be imposed to estimate a model with latent variables, which have been derived for affine models by Dai and Singleton (2000). In a single-regime Gaussian model, Dai and Singleton show that identification can be accomplished by setting the conditional covariance to be a diagonal matrix and letting the correlations enter through the feedback matrix (Φ), which is parameterized to be lower triangular, which we do here.

The RS model complicates identification relative to an affine model. The parameterization in equations (2) to (7) already imposes some of the Dai and Singleton (2000) conditions, but some further restrictions are necessary. Since q_t and f_t are latent variables, they can be arbitrarily scaled. We set $\delta_1 = (\delta_q \, \delta_f \, \delta_\pi)' = (1\,1\,\delta_\pi)'$ in (4). Setting δ_q and δ_f to be constants allows σ_q and $\sigma_f(s_{t+1})$ to be estimated. Because q_t is an unobserved variable, estimating μ_q in (3) is equivalent to allowing γ_0 in (6) or δ_0 in (4) to be nonzero. Hence, q_t must have zero mean for identification. Therefore, we set $\mu_q = 0$, since q_t does not switch regimes. Similarly, because we estimate $\lambda_f(s_{t+1})$, we constrain f_t to have zero mean.

The resulting model is theoretically identified from the data, but it is well known that some parameters that are identified in theory can be very hard to estimate in small samples. This is especially true for price of risk parameters. Because we are using four nominal yields, we should be able to identify all three prices of risk. However, Dai and Singleton (2000) note that it is typically difficult to identify more than one constant price of risk. Hence, we set $\gamma_0 = 0$ in (6) and instead estimate the RS price of risk $\lambda_f(s_{t+1})$.

We also set $\Phi_{fq} = 0$ in equation (3). With this restriction, there are, in addition to inflation factors, two separate and easily identifiable sources of variation in interest rates: An RS factor and a time-varying price of risk factor. Identifying their relative contribution to interest rate dynamics becomes easy with this restriction and it is not immediately clear how a nonzero coefficient would help enrich the model.

As q_t and f_t are zero mean, the mean level of the real short rate in (4) is determined by the mean level of inflation multiplied by δ_π and the constant term δ_0. We set δ_0 to match the mean of the nominal short rate in the data, similar to Ang et al. (2006) and Dai et al. (2006).

Finally, we set the one-period price of inflation risk equal to zero, $\lambda_\pi(s_{t+1}) = 0$. Theoretically, this parameter is uniquely identified, but in practice the average level of real rates and the premium is largely indeterminate without further restrictions. It turns out that the first-order effect of λ_π on real rates and the inflation risk premium is similar and of opposite sign. Because of this, the parameter is not only hard to pin down, but also essentially prevents the identification of the average level of real rates and the average level of the inflation risk premium. Models with a positive one-period inflation risk premium will imply lower real rates and higher inflation premiums than the results we report.

Appendix D. A Regime-Switching Model with Stochastic Expected Inflation

In a final extension, motivated by the ARMA-model literature (see Fama and Gibbons (1982), Hamilton (1985)), we allow inflation to be composed of a stochastic expected inflation term plus a random shock:

$$\pi_{t+1} = w_t + \sigma_\pi \varepsilon^{\pi}_{t+1}, \qquad \text{Appendix D.}$$

where $w_t = \mathrm{E}_t\,[\pi_{t+1}]$ is the one-period-ahead expectation of future inflation. This can be accomplished in our framework by expanding the state variables to $X_t = (q_t\; f_t\; w_t\; \pi_t)'$, which follow the dynamics of equation (2), except now:

$$\mu(s_t) = \begin{bmatrix} \mu_q \\ \mu_f(s_t) \\ \mu_w(s_t) \\ 0 \end{bmatrix}, \quad \Phi = \begin{bmatrix} \Phi_{qq} & 0 & 0 & 0 \\ \Phi_{fq} & \Phi_{ff} & 0 & 0 \\ \Phi_{wq} & \Phi_{wf} & \Phi_{ww} & \Phi_{w\pi} \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{D1}$$

and $\Sigma(s_t)$ is a diagonal matrix with $(\sigma_q\, \sigma_f(s_t)\, \sigma_w(s_t)\, \sigma_\pi(s_t))'$ on the diagonal. Note that both the variance of inflation and the process of expected inflation are regime-dependent. Moreover, past inflation affects current expected inflation through $\Phi_{w\pi}$.

The real short rate and the regime transition probabilities are the same as in the benchmark model. The real pricing kernel also takes the same form as (5) with one difference: The regime-dependent part of the prices of risk in equation (6) is now given by

$$\lambda(i) = (\lambda_f(i)\; \lambda_w(i)\; \lambda_\pi(i))',$$

but we set $\lambda_w(i) = 0$ for identification.

Appendix E. Specification Tests

A. Moment Tests

To enable comparison across several nonnested models of how the moments implied from various models compare to the data, we introduce the point statistic:

$$H = (h - \bar{h})' \Sigma_h^{-1} (h - \bar{h}), \tag{E1}$$

where $\bar{h}$ are sample estimates of unconditional moments, h are the unconditional moments from the estimated model, and Σ_h is the covariance matrix of the sample estimates of the unconditional moments, estimated by GMM with four Newey and West (1987) lags. In this comparison, the moments implied by various models are compared to the data, with the data sampling error Σ_h held constant across the models. The moments we consider are the first and second moments of term spreads and long yields, the first and second moments of inflation, the autocorrelogram of term spreads, and the autocorrelogram of inflation.

Equation (E1) ignores the sampling error of the moments of the model, implied by the uncertainty in the parameter estimates, making our moment test informal. However, this allows the same weighting matrix, computed from the data, to be used across different models. If parameter uncertainty is also taken into account, we might fail to reject, not because the model accurately pins down the moments, but because of the large uncertainty in estimating the model parameters.

B. Residual Tests

We report two tests on in-sample scaled residuals ϵ_t of yields and inflation. The scaled residuals ϵ_t are not the same as the shocks ε_t in (2). For a variable x_t, the scaled residual is given by $\epsilon_t = (x_t - \mathrm{E}_{t-1}(x_t))/\sqrt{\mathrm{var}_{t-1}(x_t)}$, where x_t are yields or inflation. The conditional moments are computed using our RS model and involve ex ante probabilities $p(s_t = i|I_{t-1})$. Following Bekaert and Harvey (1997), we use a GMM test for serial correlation in scaled residuals ϵ_t:

$$E[\epsilon_t \, \epsilon_{t-1}] = 0. \tag{E2}$$

We also test for serial correlation in the second moments of the scaled residuals:

$$E\left[(\epsilon_t^2 - 1)(\epsilon_{t-1}^2 - 1)\right] = 0. \tag{E3}$$

REFERENCES

Ang, Andrew, and Geert Bekaert, 2002, Regime switches in interest rates, *Journal of Business and Economic Statistics* 20, 163–182.

Ang, Andrew, Geert Bekaert, and Min Wei, 2007, The term structure of real rates and expected inflation, NBER Working Paper 12930.

Ang, Andrew, Sen Dong, and Monika Piazzesi, 2006, No-arbitrage Taylor rules, Working paper, Columbia Business School.

Bansal, Ravi, George Tauchen, and Hao Zhou, 2004, Regime-shifts, risk premiums in the term structure, and the business cycle, *Journal of Business and Economic Statistics* 22, 396–409.

Bansal, Ravi, and Hao Zhou, 2002, Term structure of interest rates with regime shifts, *Journal of Finance* 57, 1997–2043.

Barr, David G., and John Y. Campbell, 1997, Inflation, real interest rates, and the bond market: A study of U.K. nominal and index-linked government bond prices, *Journal of Monetary Economics* 39, 361–383.

Bekaert, Geert, and Campbell R. Harvey, 1997, Emerging equity market volatility, *Journal of Financial Economics* 43, 29–77.

Bekaert, Geert, Robert J. Hodrick, and David Marshall, 1997, On biases in tests of the Expectations Hypothesis of the term structure of interest rates, *Journal of Financial Economics* 44, 309–348.

Bekaert, G., Robert J. Hodrick, and David Marshall, 2001, Peso problem explanations for term structure anomalies, *Journal of Monetary Economics* 48, 241–270.

Bernanke, Ben S., and Ilian Mihov, 1998, Measuring monetary policy, *Quarterly Journal of Economics* 113, 869–902.

Bikbov, Ruslan, 2005, Monetary policy regimes and the term structure of interest rates, Working paper, Columbia University.

Boivin, Jean, 2006, Has U.S. monetary policy changed? Evidence from drifting coefficients and real-time data, *Journal of Money, Credit and Banking* 38, 1149–1173.

Boudoukh, Jacob, 1993, An equilibrium-model of nominal bond prices with inflation-output correlation and stochastic volatility, *Journal of Money Credit and Banking* 25, 636–665.

Boudoukh, Jacob, Matthew Richardson, Tom Smith, and Robert F. Whitelaw, 1999, Regime shifts and bond returns, Working paper, NYU.

Buraschi, Andrea, and Alexei Jiltsov, 2005, Time-varying inflation risk premia and the Expectations Hypothesis: A monetary model of the Treasury yield curve, *Journal of Financial Economics* 75, 429–490.

Burmeister, Edwin, Kent D. Wall, and James D. Hamilton, 1986, Estimation of unobserved expected monthly inflation using Kalman filtering, *Journal of Business and Economic Statistics* 4, 147–160.

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.

Campbell, John Y., and Luis M. Viceira, 2001, Who should buy long-term bonds? *American Economic Review* 91, 99–127.

Chen, Ren-Raw, and Louis Scott, 1993, Maximum likelihood estimation for a multi-factor equilibrium model of the term structure of interest rates, *Journal of Fixed Income* 3, 14–31.

Cho, Seonghoon, and Antonio Moreno, 2006, A small-sample study of the new-Keynesian macro model, *Journal of Money, Credit and Banking* 38, 1461–1481.

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.

Cochrane, John, and Monika Piazzesi, 2005, Bond risk premia, *American Economic Review* 95, 138–160.

Cogley, Timothy, and Thomas J. Sargent, 2001, Evolving post-World War II U.S. inflation dynamics, in Bernanke, Ben S., and Kenneth S. Rogoff, eds.: *NBER Macroeconomics Annual 2001* (Cambridge, MA: MIT Press).

Cogley, Timothy, and Thomas J. Sargent, 2005, Drifts and volatilities: Monetary policies and outcomes in the post-WWII U.S., *Review of Economic Dynamics* 8, 263–302.

Dai, Qiang, and Kenneth J. Singleton, 2000, Specification analysis of affine term structure models, *Journal of Finance* 55, 1943–1978.

Dai, Qiang, and Kenneth J. Singleton, 2002, Expectation puzzles, time-varying risk premia, and affine models of the term structure, *Journal of Financial Economics* 63, 415–441.

Dai, Qiang, Kenneth J. Singleton, and Wei Yang, 2007, Are regime shifts priced in U.S. treasury markets? *Review of Financial Studies* 20, 1669–1706.

Evans, Martin D. D., 1998, Real rates, expected inflation, and inflation risk premia, *Journal of Finance* 53, 187–218.

Evans, Martin D. D., 2003, Real risk, inflation risk, and the term structure, *Economic Journal* 113, 345–389.

Evans, Martin D. D., and Karen Lewis, 1995, Do expected shifts in inflation affect estimates of the long-run Fisher relation? *Journal of Finance* 50, 225–253.

Evans, Martin D. D., and Paul Wachtel, 1993, Inflation regimes and the sources of inflation uncertainty, *Journal of Money, Credit and Banking* 25, 475–511.

Fama, Eugene F., 1975, Short-term interest rates as predictors of inflation, *American Economic Review* 65, 269–282.

Fama. Eugene F., 1990, Term-structure forecasts of interest rates, inflation and real returns, *Journal of Monetary Economics* 25, 59–76.

Fama, Eugene F., and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Fama, Eugene F., and Michael R. Gibbons, 1982, Inflation, real returns and capital investment, *Journal of Monetary Economics* 9, 297–323.

Fama, Eugene F., and G. William Schwert, 1977, Asset returns and inflation, *Journal of Financial Economics* 5, 115–146.

Garcia, Rene, and Pierre Perron, 1996, An analysis of the real interest rate under regime shifts, *Review of Economics and Statistics* 78, 111–125.

Goto, Shingo, and Walter N. Torous, 2006, 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.

Hamilton, James D., 1985, Uncovering financial market expectations of inflation, *Journal of Political Economy* 93, 1224–1241.

Hamilton, James D., 1988, Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates, *Journal of Economic Dynamics and Control* 12, 385–423.

Hamilton, James D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357–384.

Hamilton, James D., 1994, *Time Series Analysis* (Princeton, NJ: Princeton University Press).

Huizinga, John, and Frederic S. Mishkin, 1986, Monetary policy regime shifts and the unusual behavior of real interest rates, *Carnegie-Rochester Conference Series on Public Policy* 24, 231–274.

Landén, C., 2000, Bond pricing in a hidden Markov model of the short rate, *Finance and Stochastics* 4, 371–389.

Meltzer, Allan H., 2005, From inflation to more inflation, disinflation and low inflation, Keynote Address, Conference on Price Stability, Federal Reserve Bank of Chicago.

Mishkin, Frederic S., 1981, The real rate of interest: An empirical investigation, *Carnegie-Rochester Conference Series on Public Policy* 15, 151–200.

Mishkin, Frederic S., 1990, The information in the longer maturity term structure about future inflation, *Quarterly Journal of Economics* 105, 815–828.

Mishkin, Frederic S., 1992, Is the Fisher effect for real? *Journal of Monetary Economics* 30, 195–215.

Mundell, Robert, 1963, Inflation and real interest, *Journal of Political Economy* 71, 280–283.

Naik, Vasant, and Moon-Ho Lee, 1994, The yield curve and bond option prices with discrete shifts in economic regimes, Working paper, University of British Columbia.

Nelson, Edward, 2004, The great inflation of the seventies: What really happened? Federal Reserve Bank of St. Louis Working Paper 2004–001.

Newey, Whitney K., and Kenneth D. West, 1987, A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Pennacchi, George, 1991, Identifying the dynamics of real interest rates and inflation: Evidence using survey data, *Review of Financial Studies* 1, 53–86.

Roll, Richard, 2004, Empirical TIPS, *Financial Analysts Journal* 60, 31–53.

Rose, Andrew, 1988, Is the real interest rate stable? *Journal of Finance* 43, 1095–1112.

Sims, Christopher A., 1999, Drifts and breaks in monetary policy, Working paper, Yale University.

Sims, Christopher A., 2001, Comment on Sargent and Cogley's "Evolving post World War II U.S. inflation dynamics," in Bernanke, Ben S., and Kenneth S. Rogoff, eds.: *NBER Macroeconomics Annual 2001* (Cambridge, MA: MIT Press).

Sola, Martin, and John Driffill, 1994, Testing the term structure of interest rates using a vector autoregression with regime switching, *Journal of Economic Dynamics and Control* 18, 601–628.

Stockton, David J., and James E. Glassman, 1987, An evaluation of the forecast performance of alternative models of inflation, *Review of Economics and Statistics* 69, 108–117.

Taylor, John B., 1993, Discretion versus policy rules in practice, *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.

Tobin, James, 1965, Money and economic growth, *Econometrica* 33, 671–684.

Veronesi, Pietro, and Francis B. Yared, 1999, Short and long horizon term and inflation risk premia in the U.S. term structure: Evidence from an integrated model for nominal and real bond prices under regime shifts, CRSP Working Paper 508.