Real Risk, Inflation Risk, and the Term Structure

Martin D. D. Evans*

Department of Economics,

Georgetown University,

Washington DC 20057

Email: evansm1@gunet.georgetown.edu

(202) 687-1570

and the

N.B.E.R.

First Draft July 1998

This Draft January 2002

Abstract

I develop and estimate a general equilibrium model for the term structures of nominal and real interest rates in the UK that incorporates Markov-switching. The model allows for non-neutralities, nonlinear dynamics, and flexibility in the dynamics of the risk premia - features that are all present in the data. I use the model to assess how accurately the term structure reflects changing expectations of future yields and inflation. This analysis shows that the presence of time-varying risk premia make it very hard to accurately track changes in the expected path of real or nominal yields over horizons of less than five years. By contrast, variations in inflation expected over the next two to three years are very accurately reflected by changes in spread between real and nominal yields, or by changes in nominal yields alone. Over longer horizons, the term structures closely track changing expectations regarding future nominal and real yields but not future inflation.

Keywords: Term Structure, Risk Premia, Inflation Risk, Markov-Switching

JEL Codes: G12,E43,E31,E42

1 Introduction

How accurately does the term structure of interest rates reflect expectations regarding future yields and inflation? This is an old and important question for researchers and policy-makers alike, but it has yet to be precisely answered. After more than a decade of regression-based tests rejecting forms of the expectations hypothesis, it appears that changing expectations and time-varying risk premia both contribute to the dynamics of the term structure.[2] To date, however, no consensus has emerged around a model that incorporates both facets. Without such a model, it is impossible to accurately assess the degree to which variations in the current term structure reflect changing expectations or risk premia.

This paper develops a new model with the aim of quantifying the influence of time-varying risk premia on the behavior of the UK term structure. Following Cox, Ingersoll and Ross (1985) (CIR), a large literature has developed using general equilibrium bond-pricing models to study the behavior of the US nominal term structure. In this paper, I focus on UK interest rates in order to exploit the information contained in the term structures of real and nominal yields. There has been a well-established market for both conventional and index-linked debt in the UK for the past seventeen years. In Evans (1998a) I showed how prices from this market could be used to construct nominal and real yield curves. These data provide information on the source of interest rate dynamics that cannot be found by studying the behavior of nominal rates alone. In particular, they allow us to separately identify the risk premia within the nominal and real term structures and the inflation risk premium linking nominal and real yields with expected inflation.[3] The behavior of these risk premia critically determines the accuracy with which expectations regarding future yields and inflation are reflected in the current term structure.

The model I present has its antecedents in the models of Vasicek (1977) and CIR. It is related to the Affine class of general equilibrium models that have been recently used by Backus, Foresi, Mozummdar and Wu (1997), Duffee (1998), Dai and Singleton (2000), Fisher and Gilles (1996), and Roberds and Whiteman (1999) to study the US term structure. All these models relate equilibrium bond prices to a stochastic discount factor, or pricing kernel, that in a representative agent model would be identified by the discounted intertemporal marginal rate of substitution. They generate time-varying risk premium by assuming that the pricing kernel process exhibits heteroskedasticity. The key feature that differentiates my model from the Affine class is that it incorporates Markov-switching into pricing kernel process. The first use of Markov-switching appears in Naik and Lee (1994), who extend Vasicek's model so that the mean and variance of the short rate switches.[4] Markov-switching plays a more extensive role in my model; it not only affects the mean and variance of real and nominal short rates, but also their degree of mean reversion, correlations with

---

[2] Recent surveys of this research include; Bekaert, Hodrick and Marshall (1997a), Campbell (1995) and Evans and Lewis (1994). A related literature considers statistical problems with the regression-based tests of the expectations hypothesis. For example, Evans and Lewis (1994) and Bekaert, Hodrick and Marshall (1997b) examine how changes in the time-series behavior of interest rates during the sample could affect the sample properties of standard tests. Although the evidence against the expectations hypothesis is weakened under these circumstances, it is not entirely eliminated.

[3] Earlier studies of UK real rates include Brown and Schaefer (1995), Arak and Kreichner (1985), Deacon and Derry (1994) and Barr and Campbell (1997). To account for the incomplete indexation of UK index-linked debt, these papers used a variety of assumptions about the behavior of risk premia to construct real yields. The analysis here uses real and nominal yield curves that are constructed from index-linked and nominal bond prices without assumptions concerning the behavior of the risk premia or inflation risk premia; see Evans (1998a) for details.

[4] Markov-switching models have also been used to study the term structure in conjunction with the expectations hypothesis by Hamilton (1988) and Sola and Driffill (1994). These model rule out time-varying risk premia and so ascribe all term structure movements to changing expectations regarding future yields.

inflation, and the link between the risk premia and volatility.[5] These facets allow the model greater flexibility to simultaneously account for the time series and cross-sectional behavior of yields. In particular, the model produces behavior in short-term real and nominal rates consistent with the evidence of nonlinear dynamics found by Ait-Sahalia (1996), Conley et al. (1997) and Stanton (1997) relating the degree of mean-reversion to the level of the short rate in US data. The model also allows the risk premia to vary independently of interest rate volatility. Duffee (1998) argues that the absence of this feature in Affine models contributes significantly to their poor empirical performance.

The incorporation of Markov-switching has another important benefit. Over the past two decades there have been a series of widely documented changes in UK monetary policy. For example, the UK's departure from the EMS in 1992 represented a significant change in policy regime. Such changes may well have resulted in a discrete shift in the behavior of inflation and its relation to real interest rates. Remolona, Wickens and Gong (1996) argue that the inflation risk premium fell by 30 percent after the UK left the EMS. The model I present allows for discrete shifts in the whole structure of the joint process for inflation and real rates and derives their implications for expectations regarding future interest rates and inflation.

The benefits from incorporating Markov switching do not come without some costs. In particular, the risk premia identified by the model are solely functions of the state variable governed by the switching process. As such, the risk premia can only take on a finite number of values. In principle this is not an important restriction on the behavior of the risk premia because the model can be solved for any finite number states. In practice, though, it is impossible to estimate a model with many states because it contains a very large number of parameters. Thus, the cost of my modeling approach arises from the fact that it may be impossible to estimate models with a sufficiently large number of states to adequately represent the dynamics of the true risk premia and their correlations with other variables. Fortunately, this does not appear to be a serious problem here. I find that a three-state model is able to closely replicate the statistical features of the data.

To take full advantage of the UK data, the model focuses on the behavior of both nominal and real yields and their interaction with inflation. In this respect it is most closely related to Remolona, Wickens and Gong (1996) who use UK data to estimate a generalized version of the CIR model. My model contains a real risk factor that identifies the short term real interest rate, and an inflation risk factor proportional to the expected rate of inflation. The joint switching process for the two risk factors allows for the presence of a time-varying correlation between inflation and real rates. The presence of this correlation contrasts with the strong neutrality assumption found in earlier models of inflation and nominal rates (see, for example, Pearson and Sun 1991, Gong and Remolona 1996), and is strongly supported by the model estimates.

I use the monthly yields on four real and four nominal bonds to estimate versions of the model with one, two and three states in the Markov-switching process. A formal comparison of the estimates reveals that the three state version of the model best characterizes the UK data. This model does a remarkably good job at matching the behavior of real and nominal yields over the sample period. It also identifies distinct differences in the behavior of the term structure across the three states. State one is characterized by upward sloping

---

[5] The model in this paper was developed independently and con-currently with a model by Bansal and Zhou (1999). They developed a Markov-switching extension of the CIR model for the U.S. nominal term structure. This study differs from their paper in its focus on the both the real and nominal term structures in the UK and the role of inflation risk. It also differs at technical a level. With the introduction of switching, the model falls outside the Affine class where analytically solutions for equilibrium bond prices are readily calculated. One advantage of the specification adopted here over Bansal and Zhou's model is that analytic solutions for equilibrium bond prices can still be found.

yield curves for both nominal and real rates. In state two, the real yield curve is inverted and the nominal curve is U-shaped. State three is also characterized by a U-shaped nominal curve but the real curve is now sharply positively sloped.

The next step in the analysis considers the question: What does the model imply about the ability of the real and nominal term structures to predict the future path of real and nominal yields? For this purpose, I use the model estimates to decompose the variance of the spread between long and short-term yields into a component due to changing yield expectations and a component due time-varying risk premia. The relative contribution of these components can also be estimated by the slope coefficients in familiar forecasting regressions. I estimate these regressions using real and nominal yields and compare the results against the predictions of the three state model. My principle findings are that:

- Time-varying risk premia make a significant contribution to the variance of nominal spreads. Changing expectations regarding future 12-month rates account for 20 to 98 percent of the spread's variance as the maturity of the long bond rises from 24 to 240 months. Expectations regarding future long-term yields only account for 3 to 73 percent of the variance. These estimates do not significantly differ from those implied by the forecasting regressions estimated in the data.

- Time-varying risk premia are somewhat less important in the real term structure. The model estimates imply that as the maturity of the long bond rises from 24 to 240 months, between 70 and 97 percent of the variance in the spread can be accounted for by changing expectations regarding 12-month yields, and 39 to 61 percent by expectations regarding long term yields. Although these estimates are somewhat higher than those obtained from the forecasting regressions, the difference is most probably due to measurement error bias in the regression estimates.

These findings indicate that predicting the future path of real or nominal yields with any accuracy is extremely difficult over horizons of less than 5 years. The link between the current term structure and expectations of future yields only approaches the simple relation implied by the expectations hypothesis at very long horizons.

The last step in my analysis asks: Can real and nominal yields provide a reliable indicator of inflation expectations? The answer to this question depends on the size and variability of the inflation risk premium linking nominal and real yields with expected inflation. I use the model estimates to compute the term structure of inflation risk and variance decompositions for nominal yields and the spread between nominal and real yields. I find that:

- The states identified by the model can be closely associated with three distinct inflation regimes: A regime of slowly rising inflation, quickly rising inflation, and slowly falling inflation.

- The spread between nominal and real yields provides an unreliable estimate of the *level* of inflation expectations because the size of the inflation risk premium differs significantly across states at all horizons. Depending on the state, the spread overstates the rate of expected inflation by between 1 and 0.6 percent at the one month horizon. At the ten year horizon, the spread understates the rate of expected inflation by between 1 and 3.5 percent.

- Variations in the inflation risk premium contribute little to variance of the spread over horizons ranging from 1 to 36 months. Beyond 5 years, variations in the inflation risk premium imply that changes in the spread understate the *change* in expected inflation by 11 to 32 percent.

- Variations in real rates and the inflation risk premium combine so that changes in nominal yields understate the variations in expected inflation at very short and long horizons. At the two to three year horizon, however, changes in real yields and the inflation risk premium offset one another so that nominal yields move almost one-to-one with expected inflation.

These results provide straightforward guidance on how best to draw accurate inferences about changing inflation expectations. Over horizons of one to twelve months, more accurate inferences can be derived from the spread between nominal and real yields than from nominal yields alone. For longer horizons, inferences based on the spread and nominal yields are quite similar. They are reasonably accurate over horizons of two to three years. Beyond this point, changes in the term structure increasingly understate changes in inflation expectations.

It is worth emphasizing that these results are derived from the maximum likelihood estimates of a general equilibrium bond-pricing model. This is a distinctly different approach from the many papers that use forecasting equations and time-series models to study the sources of term structure dynamics. For example, Fama (1990) and Mishkin (1990) used inflation forecasting equations to examine how much changing inflation expectations contributed to the variance of the US nominal yields, while Barr and Pesaran (1995) and Barr and Campbell (1997) calculated variance decompositions for the UK term structure based on Vector Autoregressions. The analysis presented here has two main advantages over these time-series based approaches. First, the model estimates incorporate information from both the time series and cross-sectional behavior of real and nominal yields. This enables investors' expectations to be estimated with much greater precision (given the presence of time-varying risk premia), than would be possible from a couple of yields, say, in a forecasting equation. This is an important consideration when studying the accuracy with which the term structure reflects long-horizon expectations of yields or inflation. The second advantage concerns possible instability in the time-series behavior of yields and inflation induced by policy changes.[6] Simple time series models will generally be unable to accurately estimate the expectations of investors who are anticipating the consequences of a policy change (see, for example, Evans 1998b). By contrast, estimates of investors' expectations identified by the Markov-switching model incorporate the effects of anticipated future shifts in the behavior of yields and inflation.

My analysis begins, in Section 2, with the presentation of the Markov-switching model. This section also discusses the distinctive features of the model. Econometric identification, estimation and testing issues are discussed in Section 3. Section 4 presents estimates of the one, two and three-state versions of the model, tests for the number of states, and compares the model estimates with the data. My analysis of the model estimates is presented in Section 5. Section 6 concludes.

---

[6] For evidence of instability in US data, see Evans and Lewis (1994, 1995) and Travalis and Wickens (1996); in UK data, see Remolona, Wickens and Gong (1996).

2 The Markov-Switching Model

The model I develop extends recent Affine models of the term structure to include Markov-switching as in Naik and Lee (1994). I take full advantage of the UK data by focusing on the behavior of both nominal and real yields and their interaction with inflation. I begin by describing the equilibrium pricing equations that lie at the heart of the model. Next, I present the dynamics of the model and solve for equilibrium bond prices. I then discuss the distinctive features of the model in terms of the behavior of spot rates and the risk premia.

2.1 Bond Pricing

Let M_{t+1} be a random variable that prices one-period state-contingent claims. If the economy admits no pure arbitrage opportunities, it can be shown that the one-period real return on all traded assets must satisfy

$$E_t[M_{t+1}\mathcal{R}^i_{t+1}] = 1, \tag{1}$$

where $\mathcal{R}^i_{t+1}$ is the gross real return on asset i between t and $t+1$. $E_t[.]$ denotes the expectation conditioned on investors' period t information set, $\mathcal{I}_t$. (Time periods are assumed to be discrete.) I shall refer to M_t as the real pricing kernel. In economies where there is a complete set of markets for state-contingent claims, there is a unique random variable $M_t > 0$ satisfying (1). Under other circumstances, this no-arbitrage condition still holds but for a range of M_ts (Duffie 1992). In economies with a representative agent, M_{t+1} is the discounted intertemporal marginal rate of substitution so that (1) also represents a first-order condition.

We can use (1) to find equations that price both real and nominal bonds. Let $Q^n_{k,t}$ denote the nominal price of a zero coupon bond at period t paying £1 at period $t+k$. The one period real return on this k-period bond is $(Q^n_{k-1,t+1}/Q^n_{k,t})(P_t/P_{t+1})$ where P_t is the (known) price level at t. Substituting this expression for $\mathcal{R}^i_{t+1}$ in (1) and rearranging gives (for $k > 0$),

$$Q^n_{k,t} = E_t\left[\frac{M_{t+1}P_t}{P_{t+1}}Q^n_{k-1,t+1}\right]. \tag{2}$$

We can derive a similar equation for real bonds. Let $Q^r_{k,t}$ denote the nominal price of a zero coupon bond at time t paying £(P_{t+k}/P_t) at period $t+k$. $Q^r_{k,t}$ also defines the real price of a claim to one unit of consumption at $t+k$. Now consider the real return from holding this k-period claim for one period. In $t+1$ the nominal price of a claim to £(P_{t+k}/P_{t+1}) is $Q^r_{k-1,t+1}$ so the price of a claim to £(P_{t+k}/P_t) must be $Q^r_{k-1,t+1}(P_{t+1}/P_t)$. The real return on holding the k-period claim is therefore $Q^r_{k-1,t+1}/Q^r_{k,t}$. Substituting this for $\mathcal{R}^i_{t+1}$ in (1) gives (for $k > 0$),

$$Q^r_{k,t} = E_t\left[M_{t+1}Q^r_{k-1,t+1}\right]. \tag{3}$$

Equations (2) and (3) determine the complete set of real and nominal bond prices in the economy in terms of the dynamics of the pricing kernel, M_t, and aggregate price level, P_t. Notice that $Q^r_{0,t}$ and $Q^n_{0,t}$ must equal unity. Hence, once the dynamics of the pricing kernel and the aggregate price level have been specified, we can use (2) and (3) to solve recursively for a complete set of nominal and real bond prices.

The analysis below examines the behavior of yields and risk premia. Let $q_{k,t}^j$ denote the log price of a k-period bond, $\ln Q_{k,t}^j$. Continuously compounded k-period real and nominal yields are defined by $y_{k,t}^r \equiv -\frac{1}{k}q_{k,t}^r$ and $y_{k,t}^n \equiv -\frac{1}{k}q_{k,t}^n$ respectively. In the case of one-period yields, I drop the k subscript and refer to y_t^j as the nominal ($j = n$) or real ($j = r$) spot rate. I focus on two sets of risk premia: the term premia, and the inflation risk premia. The former are defined as the expected excess log return on a k-period bond relative to the one-period yield, y_t^j, or spot rate:

$$\theta_{k,t}^j \equiv E_t\left[q_{k-1,t+1}^j - q_{k,t}^j\right] - y_t^j,$$

for $j = \{n, r\}$. Below I refer to $\theta_{k,t}^n$ and $\theta_{k,t}^r$ as the nominal and real term premia. The inflation risk premia is defined as

$$\psi_t \equiv y_t^n - E_t[\Delta p_{t+1}] - y_t^r,$$

where $\Delta p_{t+1} \equiv \ln(P_{t+1}/P_t)$ is the rate of inflation. This is the expected excess log real return on nominal bonds relative to the real rate over a one period horizon.

2.2   The Model

My model for the term structure uses (2) and (3) together with a specification for the dynamics of the pricing kernel and inflation. Specifically I assume that the log pricing kernel, $m_t \equiv \ln M_t$, follows

$$-m_{t+1} = \kappa_m(s_t) + z_{m,t} + \lambda_m(s_t)\omega_m(s_t)u_{m,t+1}, \tag{4}$$

$$z_{m,t+1} = \mu_m(s_{t+1}) + \alpha_m(s_t)(z_{m,t} - \mu_m(s_t)) + \omega_m(s_t)u_{mt+1}, \tag{5}$$

where $u_{m,t+1}$ is an $i.i.d.$ $N(0,1)$ shock. The terms $\kappa_m(.), \lambda_m(.), \mu_m(.), \alpha_m(.)$ and $\omega_m(.) \geq 0$ are functions of a discrete-valued variable s_t that follows an independent Markov process with constant transition probabilities. The process for inflation is also characterized by a switching structure:

$$\Delta p_{t+1} = \kappa_p(s_t) + z_{p,t} + \lambda_p(s_t)\left(\rho(s_t)\omega_m(s_t)u_{m,t+1} + \omega_p(s_t)u_{p,t+1}\right), \tag{6}$$

$$z_{p,t+1} = \mu_p(s_{t+1}) + \alpha_p(s_t)(z_{p,t} - \mu_p(s_t)) + \alpha_{pm}(s_t)(z_{m,t} - \mu_m(s_t)) \tag{7}$$
$$+ \rho(s_t)\omega_m(s_t)u_{m,t+1} + \omega_p(s_t)u_{p,t+1},$$

where $u_{p,t+1}$ is a $i.i.d.$ $N(0,1)$ shock. As above, $\kappa_p(.), \lambda_p(.), \rho(.), \mu_p(.), \alpha_p(.), \alpha_{pm}(.)$ and $\omega_p(.) \geq 0$ are all functions of s_t. Investors' information, $\mathcal{I}_t$, includes the parameters, the current values of the risk factors, $z_{m,t}$, and $z_{p,t}$, and the state variable, s_t.[7]

Equations (4)-(7) describe a recursive dynamic system. From (3) we see that real bond prices depend only on m_t so the behavior of the real term structure is determined by (4) and (5). I will refer to $z_{m,t}$ as the real risk factor. Nominal bond prices depend on both the real pricing kernel and inflation so both risk factors affect the behavior of the nominal term structure. I will refer to $z_{p,t}$ is the inflation risk factor.

---

[7]This assumption rules out the possibility that investors have to learn about the current process for the risk factors. Allowing for learning in the model (i.e., by excluding s_t from $\mathcal{I}_t$) would greatly add to its complexity and make estimation intractable. For a discussion of the modelling problems induced by the introduction of learning, see Evans (1998b).

This model is a multivariate version of the Vasicek (1977) model extended to incorporate Markov switching.[8] As is well-known, the Vasicek model implies that all the risk premia are constant. In this model, both the term premia and the inflation risk premia vary with the state variable s_t. This feature differentiates the model from a large class of term structure models following CIR and provides a very flexible framework for modeling the dynamics of the term structures.

The relationship between the pricing kernel and inflation plays an important role in the analysis. If investors (correctly) perceive that the real pricing kernel and inflation evolve independently, the price of a nominal bond is equal to the price of a real bond multiplied by the expectation of the future real value of money (Campbell, Lo and MacKinlay, 1997). Although the model admits this possibility when $\alpha_{pm}(s) = \rho(s) = 0$, this restriction imposes a strong neutrality assumption on the data (in the absence of state variations). In particular, the restriction implies that; (i) real yields are uncorrelated with inflation, and (ii) there is no inflation risk premium. The first implication is easily demonstrated if we assume a single state. Equations (5) and (7) then imply that

$$Cov(z_{m,t}, z_{p,t}) = \left(\frac{\alpha_{pm}\alpha_m}{1 - \alpha_m^2} + \rho \right) \frac{\omega_m^2}{(1 - \alpha_p\alpha_m)},$$

(where the state-dependence of the parameters has been omitted for clarity). This covariance is proportional to the covariance between (expected) inflation and real yields. So when $\alpha_{pm} = \rho = 0$, real yields cannot be correlated with inflation. The second implication follows from the fact that $\rho(s)$ governs the covariance between innovation in the pricing kernel and inflation (see equations (4) and (6) above). As I discuss below, $\rho(s)$ affects the inflation hedging properties of nominal bonds, which in turn, determine the inflation risk premium. In particular, when $\rho(s) = 0$, nominal bonds have no hedging value and the inflation risk premium equals zero. Both implications of the neutrality assumption appear at odds with the UK data. The results in Evans (1998a) support the presence of a time-varying inflation risk premium and a negative correlation between inflation and real yields.

To solve for equilibrium bond prices, let $x_t' \equiv [-m_t, \Delta p_t]$, $z_t' \equiv [z_{m,t}, z_{p,t}]$, and $u_t' \equiv [u_{m,t}, u_{p,t}]$ so that (4) - (7) can be written in vector form as

$$
\begin{aligned}
x_{t+1} &= \kappa(s_t) + z_t + \Lambda(s_t)\Omega^{1/2}(s_t)u_{t+1}, \\
z_{t+1} &= \mu(s_{t+1}) + \alpha(s_t)(z_t - \mu(s_t)) + \Omega^{1/2}(s_t)u_{t+1},
\end{aligned}
\tag{8}
$$

where $\kappa(s)' = [\kappa_m(s), \kappa_p(s)], \mu(s)' = [\mu_m(s), \mu_p(s)],$

$$\alpha(s) = \begin{bmatrix} \alpha_m(s) & 0 \\ \alpha_{pm}(s) & \alpha_p(s) \end{bmatrix}, \ \Lambda(s) = \begin{bmatrix} \lambda_m(s) & 0 \\ 0 & \lambda_p(s) \end{bmatrix}, \ \text{and } \Omega^{1/2}(s) = \begin{bmatrix} \omega_m(s) & 0 \\ \rho(s)\omega_m(s) & \omega_p(s) \end{bmatrix}.$$

---

[8] The development of term structure models in discrete time is now standard; see, for example, Campbell, Lo, and MacKinlay (1996, Chapter 11) and Sun (1992). The Vasicek model has served as the basis for other models linking yields and inflation including; Pennacchi (1991), Foresi, Penati and Pennacchi (1996) and Campbell and Viceria (2001).

The equilibrium conditions in (2) and (3) can now be written as

$$Q_{k,t}^j = E_t \left[\exp\left(-d^j x_{t+1}\right) Q_{k-1,t+1}^j \right] \qquad j = \{r, n\}, \tag{9}$$

with $d^r = [1, 0]$ and $d^n = [1, 1]$. As in the Vasicek model, equilibrium bond prices in this model depend only on $d^j \left(\kappa(s) + \mu(s)\right)$ so the elements in $\kappa(s)$ and $\mu(s)$ cannot be identified separately from term structure data alone. To resolve this indeterminacy, I choose the elements of $\kappa(s)$ so that $d^j \kappa(s) = \frac{1}{2} d^j \Lambda(s) \Omega(s) \Lambda(s)' d^{j'}$ for $j = \{n, j\}$. This choice implies that $y_t^j = d^j z_t$ so the real risk factor identifies the real spot rate, and the sum of the real and inflation risk factors equals the nominal spot rate.

Solving (9) recursively with the aid of (8) gives the following expression for equilibrium log bonds prices:

$$-q_{k,t}^j = A_k^j(s_t) + B_k^j(s_t) z_t, \qquad j = \{n, r\}, \qquad k = 0, 1, \dots \tag{10}$$

where $A_k^j(.)$ is a scalar and $B_k^j(.)$ is a 1×2 vector of functions that depend on the state variable s, and the maturity of the bond, k. Because s is a discrete-valued variable, the $A_k^j(.)$ and $B_k^j(.)$ functions are completely described by the state-dependent parameters $A_k^j(s)$ and $B_k^j(s)$ for $s \in S$, where S is the set of possible states. These parameters follow the recursions

$$A_k^j(s) = E_s \left[A_{k-1}^j(\tilde{s}) + B_{k-1}^j(\tilde{s},) \left(\mu(\tilde{s}) - \alpha(s)\mu(s)\right) \right] + \theta_k^j(s),$$

$$\tag{11}$$

$$B_k^j(s) = E_s \left[B_{k-1}^j(\tilde{s})\alpha(s) \right] + d^j,$$

where $E_s\{f(\tilde{s})\} = \sum_{\tilde{s}=1,0} f(\tilde{s}) \Pr(s_{t+1} = \tilde{s} | s_t = s)$, with $A_1^j(s) = 0$ and $B_1^j = [1, 1]$. The appendix provides a detailed derivation of these recursions and contains a description of the state-dependent function $\theta_k^j(.)$ that determines the term premium on a k-period bond: $\theta_{k,t}^j = \theta_k^j(s_t)$. In particular, the appendix shows that the $\theta_k^j(.)$ function depends on the values of $B_{k-1}^j(s)$, $\Lambda(s)$, $\Omega(s)$ and $\kappa(s)$ for $s = 1, 2, \dots$ so that (11) defines a set of nonlinear recursions for the $A_k^j(s)$ and $B_k^j(s)$ parameters.[9]

2.3 Features

This model differs from CIR-type models in its implications for the behavior of spot rates and the risk premia. Consider first the behavior of the real spot rate. Combining the equilibrium condition $y_t^r = z_{m,t}$ with (5) gives

$$y_{t+1}^r = \mu_m(s_{t+1}) + \alpha_m(s_t)(y_t^r - \mu_m(s_t)) + \nu_{t+1}, \tag{12}$$

where $\nu_{t+1} \sim N(0, \omega_m^2(s_t))$. (12) shows the real spot rate following a switching AR(1) process with heteroskedastic innovations. This process introduces two features that are absent in CIR-type models. First it breaks the link between the level and volatility of the spot rate. Volatility may increase or decrease with the level of the spot rate depending on the form of the $\omega_m(.)$, $\mu_m(.)$ and $\alpha_m(.)$ functions. This also means that the level and conditional variance of yields need not display the same degree of persistence. Persistence in

---

[9] The appendix also contains derivations for many of the results presented below along with details of the methods used to identify, estimate and test the model.

the level depends on the form of the $\alpha_m(.)$ function and the persistence in s_t, whereas persistence in volatility only depends on the latter. In CIR-type models, by contrast, volatility is a linear function of the spot rate so volatility must display the same degree of persistence as the level.

Switching also introduces nonlinearity into the spot rate process. In particular, the drift function $E\left[y_{t+1}^r|y_t^r\right]$ can now be nonlinear in y_t^r. Intuitively, a rise in y_t^r increases the forecast of y_{t+1}^r given s_{t+1} and s_t by $\alpha(s_t)$, and changes the probability distribution of s_{t+1} and s_t. The combined effect determines how $E\left[y_{t+1}^r|y_t^r\right]$ changes and may differ according to the level of y_t^r. This means, for example, that real yields could display greater mean reversion the further y_t^r is from its unconditional mean. Ang and Bekaert (1998) study this effect using a switching specification like (12) to model nominal interest rates. Their estimates of the implied drift functions closely correspond to the estimates obtained by Ait-Sahalia (1996), Conley et al. (1997) and Stanton (1997) using non-parametric methods. The presence of switching allows the model to capture nonlinearity in the dynamics of spot rates and permits us to study their implications for the behavior of the term structure.

The model also differs from CIR-type models in the way it links the term premia to the behavior of spot rates. To illustrate this point, consider the following approximation to the term premium on a two-period real bond:

$$
\begin{aligned}
\theta_{2,t}^r + \frac{1}{2}Var_t(q_{t+1}^r) &\simeq -Cov_t(m_{t+1}, q_{t+1}^r), \\
&= -\lambda_m(s_t)\omega_m^2(s_t),
\end{aligned} \tag{13}
$$

where $Var_t(.)$ and $Cov_t(.)$ denote the variance and covariance conditioned on time t information, $\mathcal{I}_t$.[10] The variance term on the left is a Jensen Inequality adjustment that appears because the term premium was defined in terms of log returns. The right hand side of (13) identifies the hedging value of real bonds. Recall that m_{t+1} is the log of the real intertemporal marginal rate of substitution in representative agent models. So when the covariance on the right is positive, long-term real bonds provide a hedge against states where marginal utility is high, and the premium is smaller to compensate. According to the model, this covariance is proportional to the within-state variance of the spot rate, $\omega_m^2(s_t)$, and so varies over time. The term premium also varies through the price of risk parameter $\lambda_m(s_t)$.[11] This second source of variation is absent in CIR-type models and adds greater flexibility to the dynamics of the real term premium. In particular, because $\omega_m^2(s_t)$ must be non-negative, the sign of the term premium is determined by $\lambda_m(s_t)$, and can therefore change signs over the sample.

---

[10] (13) and (14) are derived by taking a lognormal approximation to (9) for the $k = 2$ case, (i.e., by assuming that m_{t+1} and q_{t+1}^j have a joint normal distribution conditioned on $\mathcal{I}_t$). The approximation error arises because $q_{t+1}^j = -d^j z_{t+1}$ and the conditional distribution of z_{t+1} is non-normal unless there is a single state. The model estimates and the empirical analysis below are based on the exact term premia, $\theta_k^j(s_t)$, derived in the Appendix. I present the approximations here because the $\theta_k^j(.)$ function is too complex to develop much intuition about the role of Markov-switching.

[11] The price of real risk is the ratio of the expected excess log return on a real bond, plus one half its own variance to adjust for Jensen's Inequality, to the standard deviation of the excess log return on the bond. In this model, the price of risk on a real two-period bond is $-\lambda_m(s_t)\omega_m(s_t)$, and so varies with the state via $\lambda_m(s_t)$ and $\omega_m(s_t)$. In CIR-type models, all variations in the price of risk come through the standard deviation of the spot rate that is proportional to the square root of the state variable.

The term premium on a two-period nominal bond may be approximated by

$$
\begin{aligned}
\theta_{t,2}^n + \frac{1}{2}Var_t(q_{t+1}^n) &\simeq -Cov_t(m_{t+1} - \Delta p_{t+1}, q_{t+1}^n), \\
&= -(\lambda_m(s_t) + \lambda_p(s_t)\rho(s_t))(1 + \rho(s_t))\omega_m^2(s_t) - \lambda_p(s_t)\omega_p^2(s_t).
\end{aligned}
\tag{14}
$$

In this case, the term premium depends on the within-state variance terms, $\omega_m^2(.)$ and $\omega_p^2(.)$ and so could change signs even if $\lambda_m(.)$ and $\lambda_p(.)$ remained constant. The switching model used here allows for greater flexibility in the dynamics of the term premia via variations in $\rho()$ and the risk price terms. In this respect, the model resembles the Semi-Affine class of models developed by Duffee (1998) that introduces a more general specification for the time-varying price of risk into a CIR-type structure.

A final feature of the model worth noting concerns the state-dependence between the real pricing kernel and inflation. This is governed by the functions $\alpha_{pm}(.)$ and $\rho(.)$. Variations in the current real rate affect expectations of future inflation according to the value of $\alpha_{pm}(s)$. Although the micro foundations of the inflation process are not specified in the model, it is not unreasonable to think that variations in $\alpha_{pm}(s)$, $\alpha_p(s)$ and $\mu_p(s)$ could reflect the effects of changing monetary policy regimes.

The value of $\rho(s_t)$ affects the covariance between inflation and the pricing kernel. This is the key determinant of the inflation risk premium, ψ_t. Combining (8) and (9) with the solution for real and nominal spot rates gives

$$
\psi_t + \frac{1}{2}Var_t(\Delta p_{t+1}) = Cov_t(\Delta p_{t+1}, m_{t+1}) = -\lambda_m(s_t)\lambda_p(s_t)\rho(s_t)\omega_m^2(s_t).
\tag{15}
$$

As above, the variance term on the left is a Jensen's inequality adjustment. The covariance term on the right identifies the real hedging value of nominal bonds. In a representative agent model, a positive covariance implies that the realized real return on nominal bonds will be unexpectedly low in states where marginal utility is high. This makes nominal bonds less attractive to investors so the equilibrium inflation risk premium has to rise to compensate. In this model the (adjusted) inflation risk premia has four sources of variation: the within-state variance of the real spot rate, $\omega_m^2(s_t)$, the risk price terms $\lambda_m(s_t)$ and $\lambda_p(s_t)$, and $\rho(s_t)$. Clearly, the inflation risk premium can vary independently of both the nominal and real term premia and the variance of spot rates.

To summarize, the switching model introduces a great deal of flexibility into modeling the term structure. It accommodates nonlinearities in the dynamics of spot rates and adds flexibility to the relationship between the risk premia and volatility. The cost of this added flexibility comes in two forms. First, there are no parameter restrictions to insure that nominal yields are bounded above zero even in the continuous time limit. In principle this problem could be mitigated by making the s_t process dependent on the level of nominal yields through the transition probabilities. By this means, the volatility of nominal yields could approach zero with the level of yields in the manner of CIR-type models. Unfortunately, a modification of this type would make the model much less tractable. I regard the possibility of negative nominal yields to be a small price to pay for tractability and flexibility of the model.[12] The second cost concerns the behavior of

---

[12] Backus, Foresi, Mozumdar and Wu (1997) and Dai and Singleton (2000) make the same argument in context of their models. My estimates of the three state model imply a 2.8 percent probability that 12-month nominal yields are negative. The probability falls quickly with maturity, reaching 0.2 percent at 84 months.

the state variable. The derivation of the parameter recursions in (11) characterizing equilibrium bond prices critically relies on the assumption that s_t follows a discrete-valued process. Although s_t can take on any finite number of states in principle, in practice estimating models with many states is impossible because they contain a very large number of parameters. I consider models with one, two and three states below and show that the three-state model closely replicates the statistical features of the UK data.The Empirical Model

2.4 Estimation

The model is estimated by maximum likelihood using the yields on real and nominal bonds of 1, 3, 5 and 7 year maturities. As in other studies (e.g., Duffee 1998, and Campbell and Viceira 2001), I introduce a pricing error into the equation for equilibrium yields when estimating the model. Specifically, I assume that the observed yields, $\hat{y}_{t,k}^j$, are related to the theoretically determined yields, $y_{k,t} \equiv -\frac{1}{k} q_{k,t}^j$, by $\hat{y}_{k,t}^j = y_{k,t}^j + \xi_{k,t}^j$ where $\xi_{k,t}^j \sim i.i.d. N(0, \Sigma_k^j)$ for $j = \{r, n\}$.[13] The vector of observed yields, $\hat{y}_t = [\hat{y}_{k,t}^j]$, is then related to the risk factors and the state variable by

$$\hat{y}_t = \mathbf{A}(s_t) + \mathbf{B}(s_t) z_t + \xi_t, \tag{16}$$

where the i'th rows of $\mathbf{A}(s_t)$ and $\mathbf{B}(s_t)$ are given by $A_k^j(s_t)/k$ and $B_k^j(s_t)/k$ and satisfy the recursions in (11). The other equations in the model comprise (8), governing the dynamics of z_t, and the Markov process for s_t.

When there is one state, (i.e., $s_t = 1$), (8) and (16) constitute a state space form in which the vector of yields is related to the unobserved risk factors. In this case maximum likelihood estimates of the parameters can be obtained using the Kalman Filter, as in Pennacchi (1991). When there is more than one state, the vector of observed yields is now a function of both z_t and s_t, which are unobserved and follow non-gaussian processes. Kim (1993) provides a method for approximating the likelihood under these circumstances based on an extension of the Kalman Filter and Hamilton's (1988) algorithm for Markov processes. When I initially used this technique to estimate the two and three-state versions of the model, I found that the estimated variances of the pricing errors for 3 year real and nominal yields were very small. To obtain greater precision, I then re-estimated the models without these pricing errors. In this case, z_t can be inferred directly from $\hat{y}_t$ for each state, so s_t becomes the only unobservable variable in the model. With this simplification, the exact likelihood can be calculated with the Hamilton algorithm. The second set of estimates obtained in this manner are almost identical to the first and are reported in the tables below.

2.5 Identification

A notable feature of the model is that it utilizes data on real and nominal yields but not inflation. This specification choice has one advantage and one disadvantage. The advantage is that we do not have to deal with complications caused by the reporting lag in the Retail Price Index. The lag means that the RPI for month t, P_t, is reported two weeks into month $t+1$. As a result, we cannot simply add the equation for Δp_t to

---

[13] Recall that $\hat{y}_{t,k}^j$ almost surely contains a sampling error because it is derived from an estimated yield curve.

(16) because this would have the counter-factual implication that P_t is an element of investor's information, $\mathcal{I}_t$. By excluding the inflation data, we avoid having to model the degree to which investors anticipate the value of P_t when pricing bonds at the end of the month.[14]

The disadvantage of omitting inflation is that we cannot identify the rate of expected inflation or the inflation risk premium from the parameter estimates without a further restriction. To see why, suppose we amend (8) to

$$
\begin{aligned}
x_{t+1} &= \phi + \kappa(s_t) + \bar{z}_t + \Lambda(s_t)\Omega^{1/2}(s_t)u_{t+1} + \Sigma^{1/2}e_{t+1}, \\
\bar{z}_{t+1} &= \varphi(s_{t+1}) + \alpha(s_t)(\bar{z}_t - \varphi(s_t)) + \Omega^{1/2}(s_t)u_{t+1},
\end{aligned}
$$

where e_{t+1} is a vector of $i.i.d.N(0,1)$ shocks, $\phi' = \left[\phi_m, \phi_p\right]$ is a vector of constants and $\varphi(s) = \mu(s) - \phi$. If we choose ϕ_m and ϕ_p such that $d^j\phi = \frac{1}{2}d^j\Sigma d^{j\prime}$ for $j = \{n, r\}$, it is easy to show that equilibrium bond prices satisfy (10) with $\bar{z}_t$ replacing z_t and $\varphi(s)$ replacing $\mu(s)$ in the parameter recursions (11). Adding homoskedastic shocks to the pricing kernel and inflation processes in this manner has no effect on the dynamics of real or nominal yields. All it does is shift the long run levels of the risk factors from $\mu(s)$ to $\varphi(s)$. This means that we cannot identify the parameters in Σ (or equivalently ϕ) from the behavior of yields alone. And, since the expected rate of inflation is given by $E_t\Delta p_{t+1} = (d^n - d^r)(\kappa(s_t) + \phi + \bar{z}_t)$, we cannot therefore identify the rate of expected inflation, or the inflation risk premia, $\psi_\tau \equiv y_t^n - y_t^r - E_t\Delta p_{t+1}$.

To resolve this identification problem, I set the parameters in Σ so that the sample average of three year nominal and real yields equals the long run average implied by the model parameters. This is a minimal rational expectations assumption. Importantly, as the appendix shows, it has no impact on the differences in the behavior of term structure across regimes, the dynamics of yields, or the dynamics of the term and inflation risk premia. The model estimates reported below are based on this normalization.

2.6   Testing For Markov-Switching

In the next section I present estimates of one, two and three-state versions of the model. To assess their relative performance, we will need to test for Markov-switching. Standard hypothesis tests (i.e., Likelihood Ratio, Wald, and Lagrange Multiplier) cannot be used to test for the presence of switching between multiple states in the model specified by (8) and (16). The reason is that unidentified nuisance parameters present under the null hypothesis of fewer states invalidate the use of standard asymptotic theory (see, Hamilton 1988, and Hansen 1992). To overcome this problem, I follow Garcia and Perron (1996) by utilizing the test proposed by Gallant (1977) to compare models with different number of states. Under this procedure (described in the appendix), a large set of predicted values for the yields are calculated from estimates of the model with more states using randomly drawn values for the unidentified parameters. Several principle components are then extracted from this set of yields, added to the model with fewer states and their significance judged according to an F-test.[15]

---

[14] Evans (1998a) contains a further discussion of the potential problems caused by the reporting lag in the RPI. That paper also examines how realized inflation relates to the behavior of real and nominal yields. The focus of my analysis below is on the relationship between the term structures and inflation expectations.

[15] Hansen (1992) has also developed a test for switching but, as the appendix explains, it is too computationally intensive to apply here.

As in Garcia and Perron (1996), I also compare different versions of the model with the Davidson and MacKinnon J-test. To illustrate, let $\hat{y}_t^{s_1}$ and $\hat{y}_t^{s_2}$ denote the predicted values for the observed yields from estimates of a one and two-state version of the model respectively. The J-test is computed by first estimating the matrix regression

$$\hat{y}_t = (I - \boldsymbol{\beta})\hat{y}_t^{s_1} + \boldsymbol{\beta}\hat{y}_t^{s_2} + w_t,$$

where $\boldsymbol{\beta} = \text{diag}(\boldsymbol{\beta}_i)$, and then testing for the joint significance of $\boldsymbol{\beta}_i$ (the individual regression coefficients). The idea behind this test is that under the null of one state, predictions from the two-state version should not account for any of the discrepancy between observed yields and the predictions of the one-state model.

3 Empirical Results

3.1 Data

The analysis in this paper uses data on nominal and real yield curves derived from the secondary market prices of nominal and index-linked bonds that trade in the UK on the last business day of the month from January 1983 until November 1995. The nominal yields come from The Bank of England and are constructed using the method described in Deacon and Derry (1994) while the real yields come from Evans (1998a). The procedure for calculating real yields is summarized in the appendix. As there were relatively few nominal or index-linked bonds with short maturities trading during the sample period, it is not possible to precisely estimate the short end of the real and nominal yield curves. Estimated yields for one and two month bonds would surely contain significant sampling errors. To minimize the possible influence of these errors, I will focus on the behavior of yields for bonds with maturities of at least 12 months.[16]

Table 1 reports summary statistics on the log yields for nominal and real bonds on the last business day of the month from January 1983 until November 1995. The upper panel of the table shows that the nominal yield curve was on average mildly upward sloping while the real yield curve was downward sloping. Short-term yields are much more volatile than long-term yields in both term structures but volatility falls more quickly along the real term structure. From the skewness and kurtosis statistics, the unconditional distributions for both sets of yields appear non-normal. Variations in nominal yields of all maturities are very persistent as measured by the high values of the sample autocorrelations. In the case of real yields, persistence increases with maturity but remains below the level displayed by nominal yields.

3.2 Model Estimates

Table 2 reports the maximum likelihood estimates for one, two and three-state versions of the model. The upper rows show the state-dependent parameters of the process for the real and inflation risk factors expressed in annual percentage points. The parameters $\varphi_m(s)$ and $\varphi_p(s)$ respectively determine the long-run level of the real and inflation risk factors in each state. The estimates imply sizable cross-state differences in these

---

[16] The poor coverage of the UK market at the short-end of the maturity spectrum is widely recognized. Barr and Campbell (1995), for example, supplement the data on government bond prices with the one and three-month interbank rates to obtain their term structure estimates. They note, however, that these rates probably include a risk premium relative to the equivalent-maturity government bond, and so are not ideal.

Table 1: Summary Statistics

k months Nominal yields: $y_{t,k}^n$	mean	st.d.	skewness	kurtosis	Autocorrelations		
					lag 1	lag 2	lag 3
12	9.468	2.292	-0.265	2.479	0.958	0.912	0.874
24	9.464	1.923	-0.476	2.657	0.951	0.893	0.845
36	9.548	1.742	-0.541	2.712	0.945	0.877	0.821
60	9.626	1.636	-0.525	2.687	0.941	0.869	0.809
84	9.680	1.559	-0.481	2.637	0.939	0.867	0.805
120	9.665	1.252	-0.314	2.645	0.929	0.864	0.799
Real yields: $y_{t,k}^r$	mean	st.d.	skewness	kurtosis	lag 1	lag 2	lag 3
12	5.031	2.992	1.242	4.859	0.491	0.443	0.441
24	4.426	1.465	0.926	4.441	0.547	0.480	0.455
36	4.246	0.996	0.583	3.975	0.610	0.519	0.472
60	4.122	0.660	0.068	3.436	0.718	0.589	0.510
84	4.067	0.538	-0.175	3.287	0.791	0.646	0.552
120	4.009	0.465	-0.287	3.191	0.856	0.713	0.615

Notes: Sample statistics for nominal and real yields derived from the secondary market prices of nominal and index-linked bonds that trade in the UK on the last business day of the month from January 1983 until November 1995. The yields are calculated as $y_{k,t}^r \equiv -\frac{1200}{k} \ln Q_{k,t}^r$ and $y_{k,t}^n \equiv -\frac{1200}{k} \ln Q_{k,t}^n$. The asymptotic standard errors for the skewness and kurtosis statistics are 0.197 and 0.395.

long-run levels for both risk factors in the multiple state models. For example, in the three-state model, the largest cross-state difference between the long run levels of the real and inflation risk factors are approximately 8 and 10 percent. The parameters in the $\alpha(s)$ matrix determine the degree of within-state mean-reversion in the risk factors. There are much smaller differences in these estimates across states. The estimates of $\alpha_m(s)$ and $\alpha_p(s)$ are close to unity and the estimates of $\alpha_{pm}(s)$ are positive and significant. Thus, the estimated within-state rate of mean-reversion is very low for both risk factors.

The next three rows of the table report estimates of covariance matrix for the risk factor innovations. In the multi-state models, the estimates of $\omega_m(s)$, $\omega_p(s)$ and $\rho(s)$ differ from state to state and imply the presence of state-dependent heteroskedasticity in the innovations to the risk factors. The largest cross-state differences appear in the estimates of $\omega_m(s)$ and $\rho(s)$. The former parameter identifies the standard deviation of innovations to the real risk factor that varies from approximately 1.6 to 2.3 percent in the three state model. Since $y_t^r = z_{mt}$, $\omega_m(s)$ is also the standard deviation of innovations to the real spot rate. The estimates of $\rho(s)$ range from approximately -0.88 to -1.44 and are all statistically significant. Together, the estimates of $\omega_m(s)$, $\omega_p(s)$ and $\rho(s)$ imply that the correlation between the innovations to the real and inflation risk factors (given by $\rho(s)\omega_m(s)(\omega_p^2(s) + \rho(s)^2\omega_m^2(s))^{-1/2}$) are -0.95, -0.97 and -0.96 is states one, two and three respectively. Recall that real yields will only vary independently of inflation risk if $\alpha_{pm}(s) = 0$ and $\rho(s) = 0$. This neutrality restriction can be rejected with a high significance level based on the estimates of $\alpha_{pm}(s)$ and $\rho(s)$ in all three models. The negative values for $\rho(s)$ also imply that nominal spot rates are much less volatile than real rates within a state. Within-state innovations in nominal rates are equal to the sum of the innovations to both risk factors and so their standard deviation is given by $\left((1 + \rho(s))^2\omega_m^2(s) + \omega_p^2(s)\right)^{1/2}$.

Table 2: Model Estimates

	I: One State	II: Two States s=1	II: Two States s=2	III: Three States s=1	III: Three States s=2	III: Three States s=3
log likelihood	-1007.08	-778.357		-700.582		
$\varphi(.)$						
$\varphi_m(.)$	6.493 (0.371)	1.433 (0.449)	6.340 (0.270)	2.387 (0.269)	7.633 (0.348)	-0.573 (0.324)
$\varphi_p(.)$	8.666 (0.512)	10.300 (0.734)	3.304 (0.519)	6.278 (0.505)	1.003 (0.519)	10.938 (0.553)
$\alpha(.)$						
$\alpha_m(.)$	0.954 (0.001)	0.933 (0.001)	0.970 (0.002)	0.945 (0.001)	0.962 (0.003)	0.946 (0.003)
$\alpha_p(.)$	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)	0.990 (0.001)
$\alpha_{pm}(.)$	0.034 (0.001)	0.039 (0.001)	0.065 (0.002)	0.038 (0.001)	0.076 (0.005)	0.035 (0.007)
$\Omega^{1/2}(.)$						
$\omega_m(.)$	1.685 (0.016)	2.593 (0.030)	1.093 (0.010)	2.146 (0.020)	1.594 (0.029)	2.337 (0.047)
$\omega_p(.)$	0.639 (0.003)	0.679 (0.004)	0.517 (0.005)	0.633 (0.004)	0.577 (0.009)	0.612 (0.006)
$\rho(.)$	-0.960 (0.003)	-0.845 (0.003)	-1.335 (0.011)	-0.922 (0.004)	-1.441 (0.012)	-0.881 (0.005)
$\Lambda(.)$						
$\lambda_m(.)$	-1.825 (7.148)	-64.823 (3.278)	25.589 (24.875)	-61.095 (4.095)	77.361 (17.361)	-80.434 (3.787)
$\lambda_p(.)$	-0.309 (13.578)	-322.966 (50.767)	273.237 (32.943)	-213.878 (15.138)	488.398 (44.237)	-692.866 (266.017)
Π						
$s=1$	1.00	0.980 (0.001)	0.061 (0.002)	0.980 (0.001)	0.062 (0.004)	0.020 (0.000)
$s=2$		0.020 (0.001)	0.939 (0.002)	0.003 (0.000)	0.919 (0.003)	0.022 (0.005)
$s=3$				0.017 (0.000)	0.019 (0.004)	0.958 (0.002)
$E[y_t^r]$	6.493 (0.371)	1.433 (0.449)	6.340 (0.270)	2.387 (0.269)	7.633 (0.348)	-0.573 (0.324)
$E[y_t^n]$	15.159 (0.392)	11.733 (0.415)	9.644 (0.607)	8.665 (0.424)	8.636 (0.469)	10.365 (0.486)
$Var_t(y_{t+1}^r)^{1/2}$	1.685 (0.016)	2.593 (0.030)	1.093 (0.010)	2.146 (0.020)	1.594 (0.029)	2.337 (0.047)
$Var_t(y_{t+1}^n)^{1/2}$	0.643 (0.003)	0.789 (0.005)	0.633 (0.009)	0.655 (0.004)	0.909 (0.021)	0.673 (0.006)

Notes: The structure of each model is given by equations (8) and (16). The upper portion of the table reports estimates of the parameters in the state-dependent functions. The middle portion reports estimate of the transition probability matrix governing the Markov process, Π. The probability $\Pr(s_{t+1} = i | s_t = j)$ is reported in row i column j of the matrix. The lower panel shows the expected value for short rates in each state, $E[y_t^j]$, and the standard deviation, $Var_t(y_{t+1}^j)^{1/2}$, for $j = \{r, n\}$. Asymptotic standard errors are reported in parenthesis.

This is estimated to be equal to 65, 91 and 67 basis points in states one, two and three respectively.

The term and inflation risk premia are governed by the covariance parameters and the price of risk parameters, $\lambda_m(s)$ and $\lambda_p(s)$. In the one state model, both parameters are insignificantly different from zero. In the multi-state models the estimates are statistically significant and vary considerably across regimes. Recall that variations in the price of risk parameters add flexibility to the relationship between the term premium and volatility. In the case of the real term premium on two-month bonds, the estimates of the three-state model imply a positive premium of 24 and 37 basis points in states ones and three and a negative premium of 16 basis points in state two. Since real spot rates exhibit least volatility in state two, this implies that the term premium can change sign and be positively correlated with volatility - a combination of features that single factor CIR-type models cannot replicate. In the case of the three state model, the estimates imply a negative nominal term premium (on two-month bonds) of 72 and 7 basis points in states two and three, and a positive premium of 3 basis bonds in state one. Again, there is no simple relationship between the premium and the volatility of spot rates. Cross-state differences in the estimates of $\lambda_m(s)$, $\lambda_p(s)$ $\rho(s)$ and $\omega_m(s)$ all contribute to the inflation risk premium identified in equation (15) above. The estimates from the three-state model imply that the risk premium is equal to 1.09, 1.15 and 1.24 percent in states one, two and three respectively.[17]

The table reports estimates of the Markov transition probability matrix, Π, in the rows following the price of risk parameters. The diagonal elements of Π identify the probability of s_t remaining in the same state from one month until the next. In both multi-state models, estimates of these probabilities are close to one in every state. This means that variations in the state contribute more to the variance of the pricing kernel and inflation over longer horizons. Consequently, the possibility of a change in state impacts more on the behavior of long-term rather than short-term yields.[18]

The last four rows show how the behavior of real and nominal spot rates differs across states. In state s, the vector of risk factors converge to $\varphi(s)$ so the within-state long-run level for the spot rate is $E[y_t^j] = d^j \varphi(s)$. These averages differ significantly across states (and models). Concentrating on the case of the three-state model, the average real spot rate rises from 2.3 percent in state one to 7.6 percent in state two, and then falls to -0.5 percent in state three. Nominal rates, by contrast, are approximately equal to 8.6 percent in states one and two, and rise to 10.4 percent in state three. We can also characterize these differences in terms of the average level of the inflation risk factor, $\varphi_p(s) = E[y_t^n] - E[y_t^r]$. From this perspective, states one and three are characterized by much higher average levels of the inflation risk factor than state two. The standard deviation of innovations to spot rates within each state are reported in the last two rows. There is no simple relationship between the within-state averages and these volatility measures. In the case of real rates, the state with the lowest average has the highest volatility and vice-versa. Nominal rates, by contrast, display most volatility in state two where the average level is similar to state one.

Although many of the parameter estimates reported in Table 2 are broadly similar across the three model versions, there are some striking differences. In particular some of the parameter estimates in $\varphi(s)$, $\Omega(s)$ and $\Lambda(s)$ differ significantly from model to model. Since these parameters play an important role in

---

[17] Another way to interpret the parameter estimates is in terms of the price of risk. Estimates from the three-state model imply that the price of real (nominal) risk is 0.109 (0.047),-0.102 (-0.792), and 0.156 (-0.104) in states one, two and three.

[18] To conserve space, Table 2 omits the estimates of the pricing error variances. With the exception of the 12-month real yield, the estimated variances are very small. A table comparing the fitted values from each model against the data is provided in the appendix.

<div align="center">

Table 3: Tests

</div>

	Tests for the number of States			
	One vs. Two States	Two vs. Three States		
Gallant	61.534**	65.550**		
J-tests				
y_{12}^r	6.060**	3.484**		
y_{60}^r	7.267**	5.342**		
y_{84}^r	8.580**	1.392 *		
y_{12}^n	15.728**	1.688 *		
y_{60}^n	15.407**	0.856		
y_{84}^n	12.351**	1.481		
Joint	14.511**	15.774**		
	Coefficient Tests			
	Two State Model	Three State Model		
	1, 2	1, 2	1, 3	2, 3
$\varphi_m(.)$	74.627**	561.174**	176.772**	727.838**
$\varphi_p(.)$	497.943**	1022.666**	260.791**	1051.475**
$\alpha_m(.)$	165.349**	25.728**	0.001	10.270**
$\alpha_p(.)$	0.001	0.001	0.001	0.001
$\alpha_{pm}(.)$	102.421**	98.232**	0.507	50.533**
$\omega_m(.)$	2340.863**	271.187**	17.928**	170.093**
$\omega_p(.)$	653.186**	34.519**	8.078**	9.762
$\rho(.)$	1743.243**	1531.375**	45.378**	1823.851**
$\lambda_m(.)$	16.722**	80.401**	81.345**	101.071**
$\lambda_p(.)$	63.089**	228.562**	3.087	18.168**

Notes: The upper panel reports Gallant and $J-$test statistics for the null hypothesis of one state against two states in the left hand column, and two states against three states in the right hand column. Details of the Gallant test are provided in the appendix. The lower panels report Wald test statistics for the null hypothesis that $\varkappa(s_i) = \varkappa(s_j)$ for the state-dependent parameter $\varkappa(.)$ with s_i and s_j equal to the states listed at the head of each column. "*" and "**" respectively denote significance at the 5 percent and 1 percent levels.

determining the behavior of yields and the risk premia, it is important to exam which version of the model best characterizes the data. To this end, the upper panel of Table 3 presents the results of the Markov-switching tests. The left hand column reports tests of the one-state null hypothesis against the alternative of two states. Based on Gallant's test, this null is strongly rejected. The right hand column reports tests of the two-state null against the alternative of three states. This null hypothesis is also strongly rejected by Gallant's test. In both cases, the marginal significance level of the test statistic is less that 1 percent. The next rows of the table report J-test statistics for six yields individually and jointly for one verse two states in the left hand column, and two versus three states in the right hand column. These tests consider whether the difference between the actual yield on, say, a one year real bond and the prediction of the model with fewer states can be improved upon with the predictions from the model with more states. The results in the left hand column show that the two-state model improves on the ability of the one-state model with

respect to all the yields; the marginal significance level of the test statistics are less than 1 percent in every case. In the right hand column, the test statistics are significant for all but the cases of long-term nominal yields. Taken together, these results indicate that the three-state version of the model best characterizes the data. In particular, the three-state model appears to capture aspects of the real term structure that were not adequately represented by the two-state model.

The lower panel of Table 3 reports tests for cross-state differences in the parameters. These tests do not speak to question of how many states should be in the model. Rather they provide statistical evidence on how switching between states affects the dynamics of yields and the various risk premia. The table reports Wald tests for the null hypothesis that $\varkappa(s_i) = \varkappa(s_j)$ for the state-dependent parameter $\varkappa(.)$ with s_i and s_j equal to the states listed at the head of each column.[19] In both the three and two-state models, the only parameter not to display any significant cross-state difference is $\alpha_p(s)$, the coefficient on the lagged inflation risk factor in the inflation risk factor process. This coefficient is uniformly estimated at 0.99 across all states and versions of the model. All the other parameters differ significantly across at least two states.

3.3 The Three-State Model

In light of the results reported in Table 3, I focus below on the estimates of the three-state model. Figure 1 plots estimates of the state variable, s_t. As the plot shows, there are numerous changes in the estimated value of the state variable throughout the sample. The estimated matrix of transition probabilities implies that the unconditional probability of being in state one, two and three is 0.604, 0.098 and 0.298. Over the sample, the corresponding empirical frequencies are 0.559, 0.21 and 0.231. Thus, the model estimates imply that over the sample state one occurred a little less and state two a little more frequently than was consistent with the expectations embedded in the term structure.

Differences in the behavior of the term structure across states are displayed in Figure ??. The plots show the average yield curve in each state calculated as $y_k^j = (A_k^j(s) + B_k^j(s)\varphi(s))/k$ for $j = \{r, n\}$. In the case of real yields (shown on the left), the slope of yield curve is positive in state one, negative (inverted) in state two, and positive in state three with greater curvature than in state one. All three curves converge to a yield of approximately 5.3 percent at long maturities. The average nominal curves (shown on the right) are somewhat different. The state one curve is concave and positively sloped, while the state two and three curves are "U-shaped". The state two curve falls to a minimum of approximately 7.9 percent at 12 months before rising steadily above the state three curve and towards the state one curve. In state three, yields are initially higher and fall gradually towards a minimum of approximately 9.6 percent at 36 months. There is much less convergence in the three curves than is exhibited by the real curves at long maturities. Overall, these plots show quite different cross-sectional relationship between yields across states.

---

[19] Because all the parameters of the model are identified under the null if multiple states are present, the restrictions on $\varkappa(s)$ can be examined with a conventional Wald test, as in Engel and Hamilton (1990).
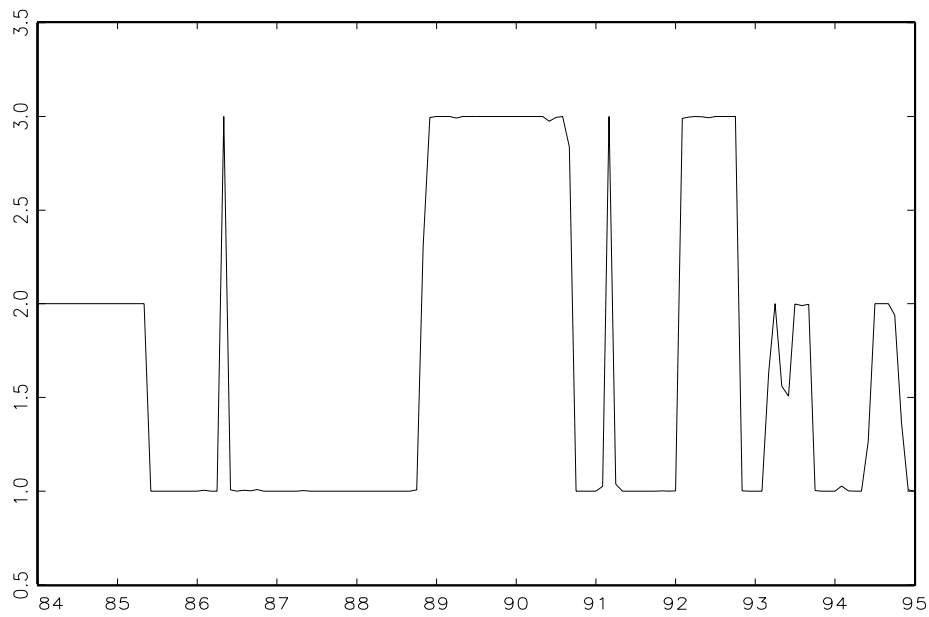
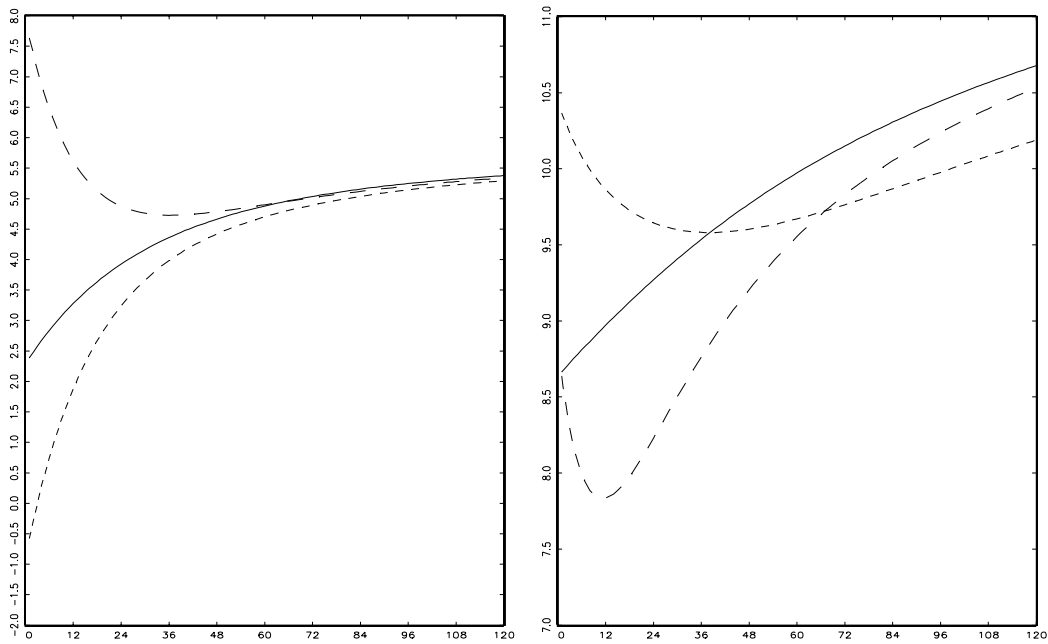Figure 1: Estimates of the state, s_t.

Figure 2: Average yield curves for $s = 1$ (solid line), $s = 2$ (dashed line), and $s = 3$ (short dashed line). Real yields are shown in the left hand plot and nominal yields are shown in the right hand plot.

19

The model estimates also imply different dynamic behavior of the real and inflation risk factors across states. To illustrate this Figure 3 plots the autocorrelation function for the real and inflation risk factors within each state.[20] The autocorrelations in the upper two plots show that both risk factors are more slowly mean reverting in state two than in states one and three. Recall from Table 2 that state two is also characterized by the highest average for the real risk factor, and the lowest average for the inflation risk factor. This means that in state two realizations of the real spot rate (i.e., $y_t^r = z_{m,t}$) are higher on average and display less mean-reversion than in other states. The autocorrelation function for the sum of the real and inflation risk factors is shown in the lower plot. All three autocorrelation functions fall more slowly than the real risk functions with the state two function again being slowest. These plots imply that realizations of the nominal spot rate (i.e., $y_t^n = z_{m,t} + z_{p,t}$) will display the slowest rate of mean reversion in state two.
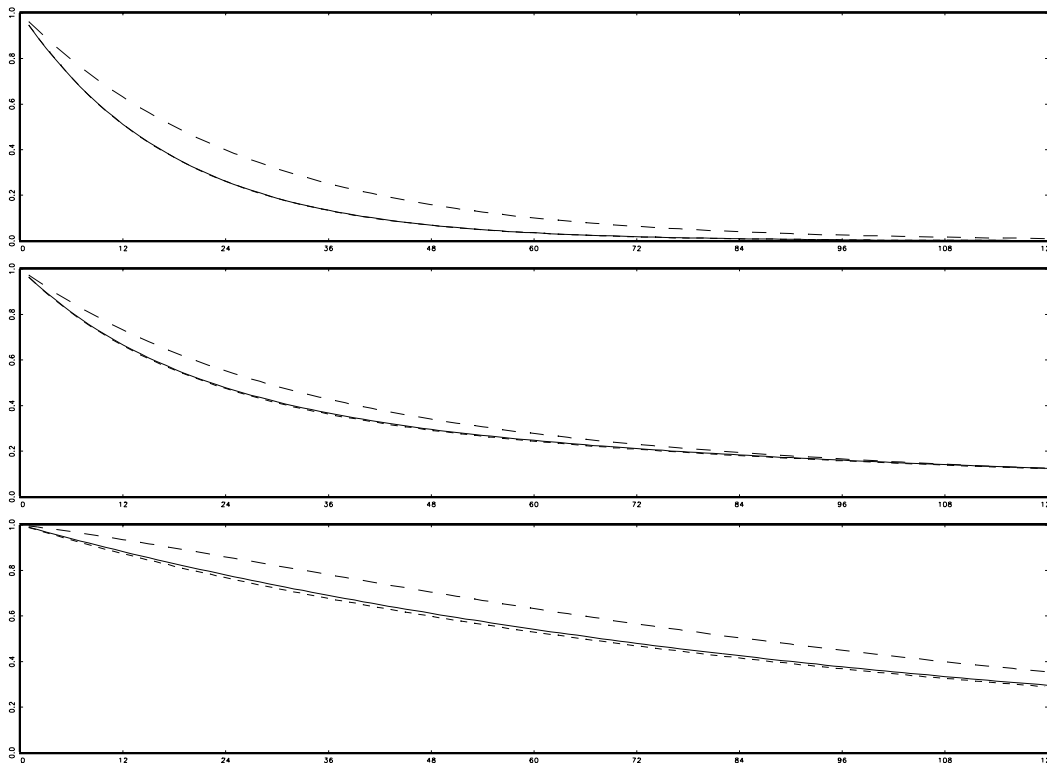

Figure 3: Autocorrelation functions for $z_{m,t}$ (top plot), $z_{p,t}$ (middle plot), and $z_{m,t} + z_{p,t}$ (bottom plot) within state $s = 1$ (solid line), $s = 2$ (dashed line) and $s = 3$ (short dashed line).

A visual comparison of the yields and the predicted values from the three-state model is provided in Figure 4. The overall impression from these plots is that the model quite accurately characterizes the behavior of both the real and nominal term structures. The largest differences appear in the plot for 12-month real yields. Actual yields are more volatile than the model predictions; the sample standard deviation of actual

___

[20] These autocorrelations are computed as $\ell \Phi_k(s)\ell' / [\ell \Phi_0(s)\ell']$ where the vector ℓ picks out the risk factor(s) from z_t, $\Phi_k(s) = \alpha(s)^k \Phi_0(s)$ and $\Phi_0(s) = \alpha(s)\Phi_0(s)\alpha(s)' + \Omega(s)$ for $s = \{1, 2, 3\}$.

yields is 3.0 percent compared to 1.86 percent for the model predictions. The model does a better job of replicating the volatility of the other yields. Differences between the standard deviations for actual and predicted yields are all less than 0.03 percent. The plots also reveal periods where the model over or under-predicts actual yields for several months. For example, the model over-predicts 12-month nominal yields for much of 1994 to 1996. Over the whole sample, however, the averages of actual and predicted yields are very similar (see appendix table for details).
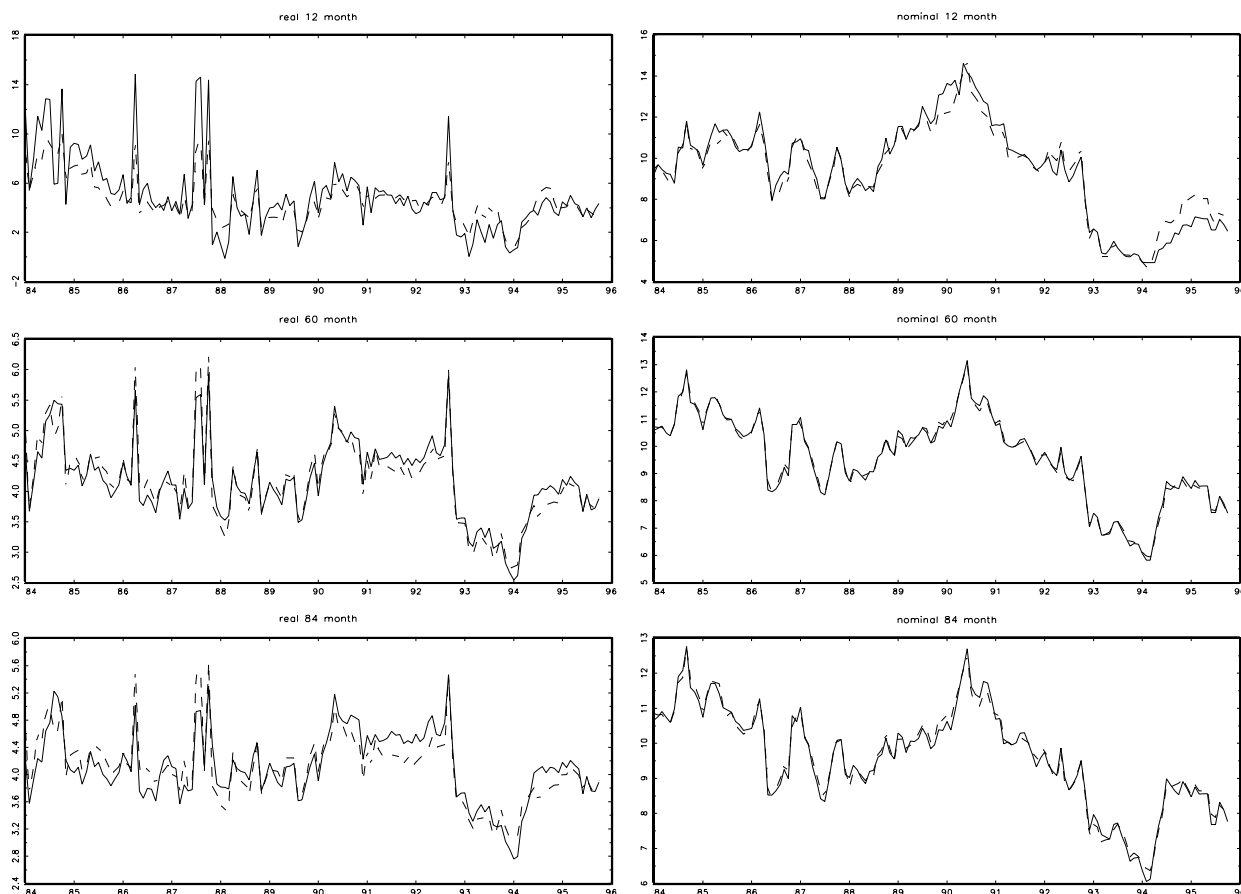
Figure 4: Actual yeilds (solid plot) and predicted yeilds (dashed plot) from estimates of the three state model.

4 Analysis

In this section, I use the estimates of the three-state model to address two questions: What does the model imply about the ability of the real and nominal term structures to predict the future path of real and nominal yields? Can movements in real and nominal yields provide a reliable indicator of changing inflation expectations? The answer to the first question depends on the behavior of the real and nominal term premia. Since these risk premia are time-varying in the model, changes in the current term structure will be a less

than completely reliable guide to changing expectations regarding yields. The issue is the *degree* to which changes in the current term structure reflect changing yield expectations. The answer to the second question depends on the behavior of both the term and inflation risk premium. If all these risk premia were constant, then changes in inflationary expectations would be reflected one-to-one in the spread between nominal and real yields. Below I examine how variations in the risk premia implied by the model estimates affect term structure based inferences about inflation expectations at different horizons.

4.1 How Variable are the Term Premia?

We can use some well-known regressions to see how variations in the term premia affect inferences concerning future yields. For this purpose it is useful to consider the term premia defined relative to the h-month rather than the 1-month yield. This is the expected excess return on holding a k-month bond for h $(< k)$ months relative to the h-month yield:

$$\Theta_{k,t}^j \equiv \frac{1}{h} E_t[q_{k-h,t+h}^j - q_{k,t}^j] - y_{h,t}^j \qquad j = \{n, r\}.$$

According to the Markov-switching model, $\Theta_{k,t}^j$ can be expressed as just a function of k and the current state s_t; $\Theta_k^j(s_t)$. By focusing on $\Theta_k^j(s_t)$ we can examine the variability of the term premium without having to work with the data on one-month real and nominal yields.

To derive the regressions, assume that the maturity of the long-term bond is equal to $k = \tau h$ months. Substituting the identity $q_{k,t}^j \equiv -k y_{k,t}^j$ into the definition above and iterating forward gives

$$\nabla y_{k,t}^j \quad = \quad y_{k,t}^j - y_{h,t}^j = \sum_{i=1}^{\tau-1} (\tfrac{\tau-i}{\tau}) E_t \Delta^h y_{h,t+hi}^j + \frac{1}{\tau} \sum_{i=0}^{\tau-2} E_t \Theta_{k-hi}^j(s_{t+hi}), \qquad (17)$$

$$\nabla y_{k,t}^j \quad = \quad y_{k,t}^j - y_{h,t}^j = (\tau - 1)(E_t y_{k-h,t+h}^j - y_{k,t}^j) + \Theta_k^j(s_t). \qquad (18)$$

The first equation shows the spread between the k and h-month yields equal to the weighted average of expected changes in the future h-month yield, $E_t \Delta^h y_{h,t+hi}^j$, plus the average of the current and expected future term premia, $E_t \Theta_{k-hi}^j(s_{t+hi})$. In the second, the spread is equal to the expected change in the long-term yield plus the term premium. Multiplying both sides of these equations by the spread and taking expectations of the resulting expressions gives

$$Var(\nabla y_{k,t}^j) \quad = \quad \sum_{i=1}^{\tau-1} (\tfrac{\tau-i}{\tau}) Cov(E_t \Delta^h y_{h,t+hi}^j, \nabla y_{k,t}^j) + \frac{1}{\tau} \sum_{i=0}^{\tau-2} Cov(E_t \Theta_{k-hi}^j(s_{t+hi}), \nabla y_{k,t}^j), \qquad (19)$$

$$Var(\nabla y_{k,t}^j) \quad = \quad (\tau - 1) Cov(E_t y_{k-h,t+h}^j - y_{k,t}^j, \nabla y_{k,t}^j) + Cov(\Theta_k^j(s_t), \nabla y_{k,t}^j). \qquad (20)$$

These equations provide us with two variance decompositions for the spread. In (19) the variance comprises the weighted sum of the covariances between the spread and expected future changes in h-month yields, and the average of the covariance between the spread and the (expected) term premia. In (20) the variance comprises the covariance between the spread and the expected change in the long-term yield,

and the covariance between the spread and the term premium. Under the expectations hypothesis, term premia are constant, so the second terms in (19) and (20) equal zero. Under these circumstances, changing expectations regarding future yields account for all the variance in the spread. In the Markov-switching model both terms in each decomposition contribute to the variance of the spread because the term premia are time-varying. Thus, the question to be addressed is: How large is the contribution of changing yield expectations to the variance of the spread given the presence of the time-varying term premia identified by the model estimates?

We can estimate how much changing yield expectations contribute to the variance of the spread in the data with two simple regressions:

$$\sum_{i=1}^{\tau-1} (\tfrac{\tau-i}{\tau})\Delta^h y_{h,t+hi}^j \;=\; a_0 + a_1 \nabla y_{k,t}^j + u_{t+k},$$

(21)

$$y_{k-h,t+h}^j - y_{k,t}^j \;=\; b_0 + b_1 \tfrac{1}{\tau-1}\nabla y_{k,t}^j + e_{t+h}, \qquad j = \{r,n\}.$$

Regressions similar to these have often been used to test the expectations hypothesis, (see Campbell and Shiller 1991 for a survey). Here they provide us with estimates of the variance decompositions. In particular, under the rational expectations, the OLS estimates of a_1 and b_1 respectively equal the first terms in (19) and (20) expressed as a faction of $Var(\nabla y_{k,t}^j)$.[21] These estimates can be compared against the predictions of the model in two ways. First, we can use the model estimates to compute each of the terms in (19) and (20) (see appendix for details). This provides us with asymptotic variance decompositions conditioned on the parameter estimates. In other words, these decompositions match those derived from the regressions in (21) estimated with an infinitely large data sample generated by the model. Second, we can use the model estimates to simulate artificial data series equal in length to the data sample. The regressions in (21) can then be run with these artificial data to derive a further set of variance decomposition estimates. Repeating this procedure a large number of times will provide us with a Monte Carlo distribution of finite-sample estimates for the variance decompositions based on the model.

The left hand columns of Table 4 report the estimated regression coefficients a_1 and b_1 using $h = 12$ with asymptotic standard errors computed under the null of the expectations hypothesis (i.e., $a_1 = 1$ and $b_1 = 1$). In the case of nominal yields, the estimates of a_1 range between 0.64 and 1.07 while the estimates of b_1 are somewhat lower, ranging between 0.28 and 0.74. Although these estimates suggest that variations in expected future 12-month yields contribute more to the variance of the spread than do varying expectations concerning future long-term yields, none of the coefficients are estimated with a great deal of precision. In fact, using the standard errors in column (ii), we cannot reject the null of $a_1 = 1$ or $b_1 = 1$ in any case at the 5 percent significance level. While these findings are consistent with the results reported in Barr and Pesaran (1995) and Barr and Campbell (1997) and Bekaert, Hodrick and Marshall (1997a), they are rather different from the results reported for the US term structure. In US data the estimates of a_1 for nominal

---

[21] To see this explicitly, note that $Cov(E_t\Delta^h y_{t+hi}^j, \nabla y_{k,t}^j) = Cov(\Delta^h y_{t+hi}^j, \nabla y_{k,t}^j)$ and $Cov(E_t y_{k-h,t+h}^j - y_{k,t}^j, \nabla y_{k,t}^j) = Cov(y_{k-h,t+h}^j - y_{k,t}^j, \nabla y_{k,t}^j)$ under the rational expectations assumption that forecast errors are uncorrelated with variables in the information set $\mathcal{I}_t$ (e.g., $\nabla y_{t,k}^j$). Making these substitutions in (19) and (20) and dividing both sides by the variance of the spread gives $1 = \hat{a}_1 + (1 - \hat{a}_1)$ and $1 = \hat{b}_1 + (1 - \hat{b}_1)$ where $\hat{a}_1$ and $\hat{b}_1$ are the OLS estimates of a_1 and b_1.

yields approach unity as the forecasting horizon k rises while the estimates of b_1 become increasingly negative (see, for example, Campbell and Shiller 1991, and Evans and Lewis 1994). The estimates of a_1 and b_1 in Table 4 point less strongly to the presence of time-varying term premia than do the US data.

The asymptotic contribution of changing yield expectations to the variance of spread implied by the model estimates is reported in column (v). In the case of nominal yields, changing expectations of future 12-month yields account for more of the variance of the spread as the maturity of the long-term bond rises; the implied estimates of a_1 rise from 20 percent at 24 months to 70 percent at 84 months. Changing expectations regarding long-term yields contribute far less to the spread variance; the estimates of b_1 range from 3 to 18 percent.[22] It is clear from these results that the model estimates ascribe an important role to time-varying term premia in the behavior of nominal yields.[23] In fact, a comparison of the estimates in columns (ii) and (v), seems to indicate that the model ascribes too little of the spread's variance to changing yield expectations. However, this is not the case. Column (ii) shows finite-sample estimates of the variance contribution, whereas column (v) reports the asymptotic variance contribution based on the model estimates. To directly compare the implications of the model against the estimates in (ii) we need to account for the effects of estimating the regression in (21) in a finite sample. For this purpose, column (vi) reports the probability of estimating a slope coefficient less than or equal to the value shown in (ii) based on the Monte Carlo distribution of regression parameters generated by the model estimates. These probabilities range between 18 and 27 percent. With this perspective, the important contribution of time-varying term premium to the variance of the spread implied be the model estimates is not at odds with the behavior of nominal yields over the sample period.[24]

The lower panel of Table 4 reports results for real yields. Here the regression estimates of a_1 range from 0.61 to 0.66 while the estimates of b_1 range from 0.22 to 0.33. These estimates are smaller than their nominal counterparts and are estimated with greater precision. And, as a result, we can reject the null of $a_1 = 1$ or $b_1 = 1$ in all cases at the 5 percent level.[25] The asymptotic contribution of changing yield expectations to the variance of the spread implied by the model range from 69 to 90 percent in the case of a_1 and from 39 to 50 percent in the case of b_1. These estimates imply that time-varying term premium contribute less to the variance of the spread asymptotically than is indicated by the OLS coefficients in column (ii). The p-values in column (vi) show that based on the model estimates, the probability of observing coefficients as low as those in (ii) in a finite sample is rather small; between 2.6 and 9.9 percent. These probabilities are much smaller than those reported for nominal yields in the upper panel of the table. They indicate that the estimated model is less able to replicate the finite sample behavior of real yields than nominal yields. This

---

[22] To see the implications of the model estimates over long horizons, the table also shows the values of a_1 and b_1 for the case where the maturity k of the long bond is 240 months. I did not run regressions for this case because the data sample only covers 15 years.

[23] This finding is consistent with the results in Steeley (1997, 2000) that provide direct evidence for the presence of time-varying risk premia in the UK nominal term structure. The presence of time-varying risk premia is also implied by the model estimates reported by Remolona, Wickens and Gong (1996).

[24] Accounting for the yield-spread regression results obtained in US data with a term structure model has proved difficult using CIR-type models. In particular, Roberds and Whiteman (1999) found that these models could not simultaneously replicate the regression results and account for other key features of the data. Dai and Singleton (2000) report greater success using Semi-Affine models that, like the model here, allow for greater flexibility in modeling of the market price of risk (see above).

[25] Similar findings are reported in Evans (1998) using an alternative forecasting equation for real yields. Earlier evidence on the presence of time-varying real term premia includes the CIR-model estimates in Brown and Schaefer (1994) and the VAR decompositions in Barr and Pesaran (1995).

Table 4: Yield Spread Decomposition Results						

Nominal Yields						
		Data Sample			Three-State Model	
Parameter	k	Estimate	Std.	R^2	Implied Parameter	P-value (%)
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
a_1	24	0.640	(0.233)	0.11	0.197	18.080
	36	0.820	(0.244)	0.17	0.361	20.420
	60	0.971	(0.231)	0.34	0.580	24.160
	84	1.071	(0.226)	0.24	0.705	27.240
	240				0.981	
b_1	24	0.280	(0.572)	0.01	0.030	18.080
	36	0.490	(0.654)	0.01	0.056	21.720
	60	0.743	(0.920)	0.02	0.125	24.260
	84	0.700	(1.175)	0.01	0.176	25.180
	240				0.729	
Real Yields						
		Data Sample			Three-State Model	
Parameter	k	Estimate	Std.	R^2	Implied Parameter	P-value (%)
a_1	24	0.613	(0.084)	0.31	0.694	4.340
	36	0.661	(0.038)	0.29	0.761	4.120
	60	0.662	(0.050)	0.28	0.851	7.280
	84	0.652	(0.084)	0.26	0.900	9.920
	240				0.972	
b_1	24	0.225	(0.191)	0.01	0.388	4.340
	36	0.336	(0.154)	0.05	0.402	4.080
	60	0.337	(0.147)	0.05	0.454	3.300
	84	0.291	(0.158)	0.03	0.502	2.620
	240				0.607	

Notes: Columns (ii) - (iv) report the estimated slope coefficients from the regressions in (21) with $h = 12$. The maturity of the long-term yield is indicated by k in column (i). Standard errors, reported in parentheses, are adjusted for conditional heteroskedasticity and serial correlation using the methods of Hansen (1982) and Newey-West (1987); an MA($k - 1$) in the case of a_1 and an MA($h - 1$) in the case of b_1. Column (v) reports the asymptotic/population regression coefficient implied by the ML estimates of the three state model. The probability of estimating a coefficient smaller than the estimate in column (ii) under the null hypothesis of the estimated three state model is reported in column (vi).

is not entirely surprising. Recall that the model was least able to replicate the high volatility of 12-month real yields. If some of this "excess volatility" is attributable to a lack of precision at the short end of the real yield curve, then the real yield spreads used in the regression will include a measurement error and the OLS coefficients in column (ii) will be biased towards zero. Because the model allows for these measurement errors, the parameter estimates used to compute the asymptotic values of a_1 and b_1 in (v) or the p-values in (vi) are free from this bias. Thus, the difficulty in precisely calculating short term real yields from the available data on index-linked bonds may explain why the p-values are smaller in the lower panel than in the upper panel of the table.

Overall, these results show that in large samples time-varying term premia make a significant contribution to the variance of both nominal and real spreads. In the nominal term structure, changing expectations regarding future 12-month rates account for 20 to 98 percent of the spread's variance as the maturity of the long bond rises from 24 to 240 months. Expectations' regarding future long-term yields only account for 3 to 73 percent of the variance. Time-varying term premia are somewhat less important in the real term structure. Our model estimates imply that between 70 and 97 percent of the variance in the spread can be accounted for by changing expectations regarding 12-month yields, and 39 to 61 percent by expectations regarding long term yields. These findings suggest that predicting the future path of real or nominal yields with any accuracy is extremely difficult over horizons of 5 years or less. The link between the current term structure and expectations of future yields only approaches the simple relation implied by the expectations hypothesis at very long horizons.

4.2 How Important is Inflation Risk?

I now turn to the second question: Can movements in real and nominal yields provide a reliable indicator of changing inflation expectations? To address this question we need to consider the links between nominal yields, real yields and expected inflation over different horizons. At the one month horizon, the spread between nominal and real spot rates is

$$y_t^n - y_t^r = E_t \Delta p_{t+1} + \psi_t,$$

and so varies with the expected monthly rate of inflation and the inflation risk premium. Over longer horizons, the nominal-real spread depends on the rate of expected inflation, the term premia in both the nominal and real term structures, and the inflation risk premium. This can be seen by combining the equation above with the nominal and real versions of (17) with $h = 1$:

$$
\begin{align}
y_{k,t}^n - y_{k,t}^r &= E_t \pi_{k,t+k} + \Psi_{k,t}, \tag{22}\\
\Psi_{k,t} &= \frac{1}{k} \sum_{i=0}^{k-1} E_t \left(\theta_{k-i,t+i}^n - \theta_{k-i,t+i}^r + \psi_{t+i} \right).
\end{align}
$$

The first term on the right is the expected rate of inflation over the next k months, $\frac{1}{k} E_t \Delta^k p_{t+k}$, the second is the multi-period inflation risk premium, $\Psi_{k,t}$. This is the average of the expected difference between the term premia on nominal and real bonds, $\theta_{k,t}^n - \theta_{k,t}^r$, plus the (one period) inflation risk premium, ψ_t. The Markov-switching model implies that $\Psi_{k,t}$ can be expressed as just a function of k and the current state

s_t; $\Psi_k^j(s_t)$. When $\Psi_k^j(s_t) = 0$, (22) describes the k-month version of the Fisher Equation.

Equation (22) has two important implications for the relation between yields and expected inflation. First, the presence of time-varying term premia, $\theta_{k,t}^j$, and inflation risk premia, ψ_t, doesn't necessarily imply that the spread between nominal and real yields will be an unreliable indication of inflationary expectations. If forecasts of $\theta_{k-i,t+i}^n - \theta_{k-i,t+i}^r + \psi_{t+i}$ are approximately constant for horizons i greater than say i^*, variations in $\Psi_{k,t}$ will become very small as k becomes much larger than i^* because $\Psi_{k,t}$ depends on the average of the forecasts at all horizons $i < k$. Notice that this differs from the requirement needed for the expectations hypothesis to approximately hold at long horizons. For this we need that the variations in $\frac{1}{k}\sum_{i=0}^{k-1} E_t\theta_{k-i,t+i}^j$ die out as k rises (see equation (19) above with $h = 1$). Forecasts of the term premia may be quite variable at particular horizon i, while forecasts of the difference between the nominal and real premia may be constant. Although the real and nominal term structures provide unreliable guidance concerning expected future yields at all but the longest horizon, it is still possible for these same term structures to provide (more) reliable inferences about inflationary expectations over short horizons.

The second implication of equation (22) concerns the relation between the size of the multi-period inflation risk premium $\Psi_{k,t}$ and k, which might be termed "the term structure of inflation risk". An important determinant of the size of $\Psi_{k,t}$ is the difference between the nominal and real term premia. Recall that $\theta_{k,t}^j$ identifies the expected excess log return on holding a $k-$period bond relative to the spot rate. The difference between $\theta_{k,t}^n$ and $\theta_{k,t}^r$ therefore indicates the degree to which investors are compensated for inflation risk inherent in holding long-term nominal bonds. Some intuition into the factors governing the difference can be gained by considering the case where k equals 2. Using (13) and (14) above, we can write

$$
\begin{aligned}
\theta_{2,t}^n - \theta_{2,t}^r \quad \simeq \quad & \frac{1}{2}\left(Var_t(q_{t+1}^r) - Var_t(q_{t+1}^n)\right) \\
& + Cov_t\left(\Delta p_{t+1}, q_{t+1}^n\right) + Cov_t\left(m_{t+1}, q_{t+1}^r - q_{t+1}^n\right).
\end{aligned}
\tag{23}
$$

The variance terms in the first line arise from Jensen's inequality because the term premium are defined in terms of log returns. The second line shows that inflation risk contributes to the nominal term premium through two components; the covariance between inflation and the future price of the nominal bond, and the covariance between the real kernel and the difference between the future prices of real and nominal bonds. The first term arises because an unexpected rise in prices between t and $t+1$ reduces the real value of the nominal long-term bond at $t+1$. This makes holding long term nominal bonds less attractive so the nominal term premium must rise to compensate. The second term adds to the premium because a positive covariance implies that the unexpected return on holding long-term real rather than nominal bonds is high when the future pricing kernel is unexpectedly high. In a representative agent model, this occurs when marginal utility is high, so once again the nominal term premium must rise to compensate.

The upper panel of Figure 5 shows estimates of $\Psi_k(s)$ plotted against the horizon k in months for $s = \{1, 2, 3\}$. Here we see that the term structure of inflation risk has a negative slope in all three states. The inflation risk premium is highest in state one, starting at approximately 1 percent when $k = 1$ and falling to minus 1.8 percent at $k = 120$. The premia in states two and three are quite similar; falling from approximately 0.6 percent at $k = 1$ to minus 3.5 percent at $k = 120$. These estimates imply that the spread between nominal and real yields overstates the rate of expected inflation by between 1 and 0.6 percent at the one month horizon. At long horizons, the opposite is true. At the ten year horizon the spread understates the

rate of expected inflation by between 1 and 3.5 percent. Overall, these plots show that the Fisher equation is a rather poor approximation to the relation between the level of nominal yields, real yields and expected inflation in the UK.
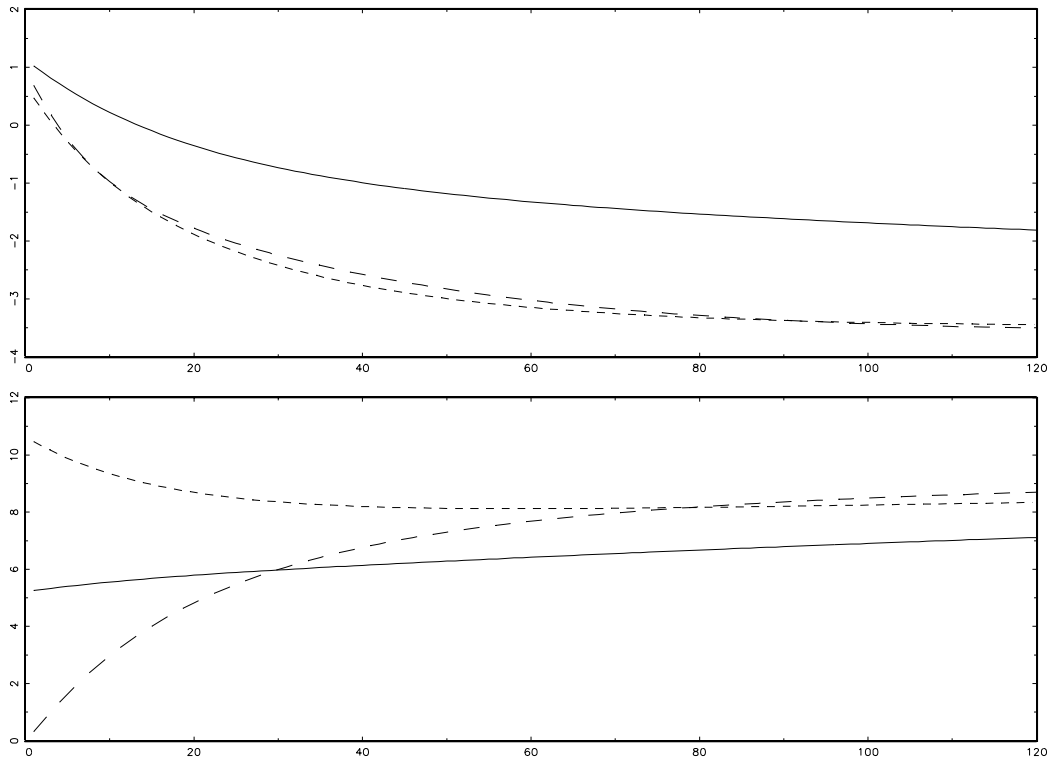

Figure 5: The multi-period inflation risk premium $\Psi_{k,t}$ (top plot), and average rate of expected inflation $E_t\pi_{k,t+k}$ (bottom plot), in states; $s = 1$ (solid line), $s = 2$ (dashed line) and $s = 3$ (short dashed line).

The lower panel of Figure 5 shows the average rate of expected inflation in each state plotted against the horizon k in months. Expected inflation rises with the horizon k in states one and two, and falls in state three. There are considerable cross-state differences in expectations at both short and long horizons. At one month, the expected rate of inflation ranges from 0.3 percent in state two to 10.3 percent in state three while the range at ten years is 7.1 to 8.7 percent. These pronounced differences suggest that the states identified by the model can be closely associated with three distinct inflation regimes: A regime of slowly rising inflation in state one, quickly rising inflation in state two, and slowly falling inflation in state three. Of course nothing hangs on this taxonomy. The point to emphasize from the plot is simply that very different sets of inflationary expectations are embedded in the dynamics of term structure.

The term structures of inflation risk shown in Figure 5 imply that the spread between nominal and real yields provide an unreliable estimate of the *level* of inflation expectations because the size of the inflation risk premium differs significantly across states at all horizons. These cross-state differences also affect how accurately *variations* in inflation expectations are reflected in the nominal-real spread. To see this formally,

k	$\dfrac{Cov\left(y_{k,t}^{n}-y_{k,t}^{r},E_{t}\pi_{k,t+k}\right)}{Var(y_{k,t}^{n}-y_{k,t}^{r})}$	$\dfrac{Cov\left(y_{k,t}^{n},E_{t}\pi_{k,t+k}\right)}{Var(y_{k,t}^{n})}$	$\dfrac{Cov\left(y_{k,t}^{n},y_{k,t}^{r}\right)}{Var(y_{k,t}^{n})}$	$\dfrac{Cov\left(y_{k,t}^{n},\Psi_{k,t}\right)}{Var(y_{k,t}^{n})}$
	(i)	(ii)	(iii)	(iv)
1	0.966	1.452	-0.561	0.109
12	0.948	1.195	-0.291	0.096
24	0.940	0.989	-0.010	0.022
36	0.975	0.968	0.103	-0.071
60	1.111	1.111	0.128	-0.239
84	1.222	1.228	0.108	-0.336
120	1.319	1.320	0.088	-0.407

Table 5: Variance Decompositions

Variance decompositions calculated from the estimates of the three-state model (see appendix for details).

I compute two variance decompositions based on (22):

$$Var(y_{k,t}^{n}-y_{k,t}^{r}) = Cov\left(y_{k,t}^{n}-y_{k,t}^{r},E_{t}\pi_{k,t+k}\right)+Cov\left(y_{k,t}^{n}-y_{k,t}^{r},\Psi_{k,t}\right),$$

$$(24)$$

$$Var(y_{k,t}^{n}) = Cov\left(y_{k,t}^{n},E_{t}\pi_{k,t+k}\right)+Cov\left(y_{k,t}^{n},y_{k,t}^{r}\right)+Cov\left(y_{k,t}^{n},\Psi_{k,t}\right).$$

The first equation writes the variance of the spread as the sum of covariance with expected inflation and covariance with the multi-period inflation risk premium. Clearly, the closer the second covariance is to zero, the more accurately will variations in the spread reflect changing inflation expectations. The second equation provides a variance decomposition for nominal yields in terms of their covariance with expected inflation, real yields and the inflation risk premium. This equation allows us to examine the degree to which changes in inflation expectations can be accurately inferred from nominal yields alone.

Column (i) of Table 5 reports the contribution of changing inflation expectations to the variance of the spread based on the three-state model estimates. For horizons ranging from 1 to 36 months, variations in the inflation risk premium contribute little to variance of the spread. On average, a one percent increase in the spread is associated with a 0.94 to 0.97 percent increase in expected inflation, a 3 to 6 percent overstatement of the change in expected inflation. Variations in the inflation risk premium become more important at horizons beyond 5 years. As a result, changes in the spread understate the change in expected inflation by 11 to 32 percent. For example, at 84 months, a one percent increase in the spread is associated with a 1.22 percent increase in expected inflation.

Columns (ii) - (iv) of the table show how variations in expected inflation, real rates and the inflation risk premium contribute to the variance of nominal yields. Column (ii) shows that changes in nominal yields understate the variations in expected inflation at very short and long horizons. Columns (iii) and (iv) show the reason for this pattern. At short horizons real yields are very variable and negatively correlated with expected inflation. Thus, insofar as real shocks contribute to changing inflation expectations, an opposite movement in real yields offsets their effect on nominal yields. This is illustrated in the table by the large

29

negative covariance between short-term real and nominal yields for k equals 1 and 12. Real yields are much less variable at long horizons. Now variations in nominal yields primarily reflect changes in expected inflation and the inflation risk premium. Because these changes are negatively correlated (see Figure 5) the effect on nominal yields of an increase in expected inflation is partially offset by the associated fall in the inflation risk premium. At the two to three year horizons, the correlation between expected inflation and real yields is offset by the correlation between expected inflation and the inflation risk premium. At this point, nominal yields move almost one-to-one with expected inflation, a finding consistent with the Fisher hypothesis (Fisher 1930).

The results in Table 5 provide straightforward guidance on how best to draw accurate inferences about changing inflation expectations. Over horizons of one to twelve months, more accurate inferences can be derived from the spread between nominal and real yields than from nominal yields alone because short-term real yields are strongly (negatively) correlated with expected inflation. For longer horizons, inferences based on the spread and nominal yields are quite similar. They are reasonably accurate over horizons of two to three years. Beyond this point, changes in the term structure increasingly understate changes in inflation expectations.

5 Conclusion

This paper has examined how changing expectations regarding future yields and inflation are reflected in the term structures of nominal and real interest rates. For this purpose, I developed a bond-pricing model with Markov-switching that allows for changes in the degree of mean-reversion displayed by spot rates, time-varying correlations between real and nominal risk factors, and flexibility in the link between the dynamics of the risk premia and interest rate volatility. A formal comparison of a one, two and three-state version of the model revealed that the three state model best characterized the UK data. This model does a remarkably good job at matching the behavior of real and nominal yields over the sample period. It also identifies distinct differences in the behavior of the real and nominal term structures across the three states.

I then used the model estimates to quantify the importance of time-varying term premia in the real and nominal term structures, and time-varying inflation risk premia linking nominal and real yields with expected inflation. I found that the presence of time-varying term premia make changes in nominal and real yield spreads a very unreliable indicator of changing yield expectations over horizons less than 5 years. The link between the current term structure and expectations of future yields only approaches the relation implied by the expectations hypothesis at very long horizons. Inferences regarding expected inflation based on the Fisher Equation are similarly flawed. Estimates of the inflation risk premium imply that the spread between nominal and real yields provide unreliable estimates of the *level* of inflation expectations at all horizons. Accurate inferences about *changing* inflation expectations can be obtained using either the change in the spread between nominal and real yields, or the change in the nominal yield alone, over a 2 to 3 year horizon. Outside this window, changes in the term structure either understate or overstate the change in expected inflation by a considerable degree.

Overall, these results clearly indicate that time-varying term premia and inflation risk premia significantly contribute to the dynamics of real and nominal yields in the UK. The model links variations in the risk premia to changes in the dynamics of spot rates and expected inflation via switches in the state variable. However,

it does not relate switches in the state variable to economic fundamentals, such as the policy regime. In this sense, Markov-switching is something of a "black-box" that enables us to develop a theoretically consistent and empirically viable model. Future research will have to look inside this box if we are to deepen our understanding of the UK term structures.

References

Ait-Sahalia, Y., 1996, Testing Continuous-Time Model of the Spot Interest Rate, *Review of Financial Studies.*

Ang, A. and G. Bekaert, 1998, Regime Switches in Interest Rates, working paper, Stanford University.

Arak, M., and L. Kreichner, 1985, The real rate of interest: Inferences from the new UK indexed gilts, *International Economic Review.*

Backus, D., S Foresi, A. Mozummdar and L. Wu, 1997, Predictable Changes in Yields and Forward Rates, manuscript, Stern School of Business, NYU.

Bansal, R. and H. Zhou, 1999, Term Structure of Interest Rates with Regime Shifts, manuscript, Fuqua School of Business, Duke University.

Barr, D. and J. Campbell, 1997, Inflation, real interest rates, and the bond market: A study of UK nominal and index-linked government bond prices, *J. Monetary Economics.*

Barr, D. and B. Pesaran, 1995, An Assessment of the Relative Importance of Real Interest Rates, Inflation and Term Premia in Determining the Prices of Real and Nominal UK Bonds, Working Paper No. 32, The Bank of England.

Bekaert, G., R. Hodrick and D. Marshall, 1997a, On biases in tests of the expectations hypothesis of the term structure of interest rates, *Journal of Monetary Economics.*

Bekaert, G., R. Hodrick and D. Marshall, 1997b, "Peso Problem" Explanations for Term Structure Anomalies, N.B.E.R. working paper.

Brown, R. and S. Schaefer, 1995, The term structure of real interest rates and the Cox, Ingersol and Ross model, *Journal of Financial Economics.*

Campbell, J., 1995, Some lessons from the yeild curve, *Journal of Economic Perspectives.*

Campbell, J., A. Lo and A. MacKinlay, 1997, *The Econometrics of Financial Markets*, (Princeton University Press, Princeton NJ).

Campbell, J. and R. Shiller, 1991, Yield Spreads and Interest Rate Movements: A Bird's Eye View, *Review of Economic Studies.*

Campbell, J., and L. Viceira, 2001, Who should buy long-term bonds? *American Economic Review.*

Conley, T., Hansen, L., Luttmer, E., and Scheinkman, J., 1997, Short-term interest rates as subordinated diffusions, *Review of Financial Studies.*

Cox, J., J. Ingersoll and S. Ross, 1985, A theory of the Term Structure of Interest Rates, *Econometrica.*

Dai, Q., and K. Singleton, 2000, Expectation puzzles, time-varying risk premia, and dynamic models of the term structure, N.B.E.R. working paper no. 8167.

Deacon, M. and A. Derry, 1994, Estimating the Term Structure of Interest Rates, Working paper No. 24, Bank of England.

Duffee, G., 1998, Term premia and interest rate forecasts in Affine models, manuscript, Haas School of Business, UC Berkeley.

Duffie, D., 1992, *Dynamic Asset Pricing Theory* (Princeton University Press, Princeton NJ).

Engel, C. and J. Hamilton, 1990, Long Swings in the Dollar: Are they in the data and do markets know it? *American Economic Review.*

Evans, M., 1998a, The Term Structure of Real Interest Rates: New Estimates and Implications, *Journal of Finance.*

Evans, M., 1998b, Dividend Variability and Stock Market Swings, *Review of Economic Studies.*

Evans, M. and K. Lewis, 1994, Do Risk Premia Explain it all? Evidence from the Term Structure, *Journal of Monetary Economics.*

Evans, M., and K. Lewis, 1995, Do inflation expectations affect the real rate? *Journal of Finance.*

Fama, E., 1990, Term structure forecasts of interest rates, inflation, and real returns, *Journal of Monetary Economics.*

Fisher, M. and C. Gilles, 1996, Term Premia in Exponential-Affine Models of the Term Structure, manuscript, Board of Governors of the Federal Reserve System.

Foresi, S., A. Penati and G. Pennacchi, 1996, Estimating the Cost of US Indexed Bonds, manuscript, Stern School of Business, NYU.

Gallant, A., 1977, Testing a nonlinear regression specification: a nonregular case, *Journal of the American Statistical Association.*

Garcia, R. and P. Perron, 1996, An analysis of the real interest rate under regime shifts, *The Review of Economics and Statistics.*

Gong, F., and E. Remolona, 1996, Inflation Risk in the U.S. Yield Curve: The Usefulness of Indexed Bonds, manuscript, Federal Reserve Bank of New York.

Hamilton, J.D., 1988, Rational Expectations Analysis of Changes in Regimes: An Investigation of the Term Structure of Interest Rates, *Journal of Economic Dynamics and Control.*

Hamilton, J.D., 1994,*Time Series Analysis*, Princeton University Press, New Jersey.

Hansen, B.E., 1992, The Likelihood Ratio Test Under Nonstandard Conditions: Testing the Markov Switching Model of GNP, *Journal of Applied Econometrics.*

Hansen, L., 1982, Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*

Kim, C.J., 1993, Dynamic Linear Models with Markov-switching, *Journal of Econometrics.*

Mishkin, F., 1990, What does the term structure tell us about future inflation? *Journal of Monetary Economics.*

Naik, V. and M.H. Lee, 1994, The Yield Curve and Bond Option Prices with Discrete Shifts in Economic Regimes, manuscript, University of British Columbia.

Newey, W. and K. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica.*

Pearson, N. D. and T.-S. Sun 1991, An Empirical Examination of the Cox, Ingersoll, Ross Model of the Term Structure of Interest Rates, manuscript, University of Rochester and Columbia University.

Pennacchi, G. 1991, Identifying the Dynamics of Real Interest Rates and Inflation: Evidence using Survey Data, *Review of Financial Studies.*

Roberds, W. and C. H. Whiteman, 1999, Endogenous Term Premia and Anomalies in the Term Structure of Interest Rates: Explaining the Predictability Smile, *Journal Of Monetary Economics.*

Remolona, E., M. Wickens and F. Gong, 1996, What was the Market's View of UK Monetary Policy? Estimating Inflation Risk and Expected Inflation with Indexed Bonds, manuscript, University of York..

Sola, M and J. Driffill, 1994, Testing the Term Structure of Interest Rates Using a Vector Autoregression with Regime Switching, *Journal of Economic Dynamics and Control.*

Stanton, R., 1997, A nonparametric model of term structure dynamics and the market price of interest rate risk, *Journal of Finance.*

Steeley, J., 1997, A two-factor model of the UK yield curve, *The Manchester School, Supplement.*

Steeley, J., 2000, Endogenous time-varying risk premia in the gilt-edged market, manuscript, University of Stirling, UK.

Sun, T.-S., 1992, Real and Nominal Interest Rates: A Discrete-time Model and its Continuous-time Limit, *Review of Financial Studies.*

Tzavalis, E., and M. Wickens, 1996, Forecasting inflation from the term structure, *Journal of Empirical Finance.*

Vasicek, O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics.*

Appendix

Data

The analysis in this paper uses data on nominal and real yield curves derived from the secondary market prices of nominal and index-linked bonds that trade in the UK. The nominal yields come from The Bank of England and are constructed using the method described in Deacon and Derry (1994). Briefly, a no-arbitrage condition is used to link the prices of discount bonds to the prices of coupon-paying bonds seen in the market. Then, at each date, the parameters of a discount function are chosen to match observed prices against their theoretical values implied by the no-arbitrage condition. The yield curve for each period is then constructed from the estimated discount function.

The construction of the real yield curve is complicated by two factors. First, index-linked bonds issued by the UK government only provide incomplete indexation for the principle and coupon payments because there is an eight month indexation lag built into the payoff structure of the bonds. Second, there is a two week reporting lag in the price index. As a result, uncertainty about the current and future price index has some effect on the prices of index-linked bonds. Both these facts make it impossible to derive the real term structure directly from the observed prices of index-linked bonds. However, in Evans (1998a) I show how real yields can be constructed using a two-step procedure. First, the index-linked yield curve is calculated from market prices using a no-arbitrage technique like the one used to find the nominal term structure. Second, the effects of inflation uncertainty (arising from the indexation and reporting lags) are purged from the index-linked yields to derive estimates of the real yield curve. In step one, I utilize the nominal discount function estimated by The Bank of England together with the prices of index-linked debt. As an alternative, one could simultaneously estimate the discount functions for nominal and index-linked bonds. This would be a somewhat more complex undertaking but might have the advantage of further reducing estimation errors in the yield curves.

Bond Price Solution

To show that equations (10) and (11) describe equilibrium bond prices, I proceed in two steps. First I show that under the proposed solution, expected excess holding returns are a function of the maturity of the long-term bond and the current state. In the second step I derive the necessary restrictions on (10) for this property to hold.

Step 1: First write the equilibrium condition in (9) as

$$1 = E\left[\exp(-d^j x_{t+1} + q^j_{k-1,t+1} - q^j_{k,t})|\mathcal{I}_t\right], \tag{A1}$$

for $j = \{n, r\}$, where $\mathcal{I}_t$ is the investor's period t information set. For the case where $k = 1$, we can calculate the expectation directly by noting that the conditional distribution of x_{t+1} is normal, and $q^j_{0,t+1} = 0$. After some rearrangement, this gives us

$$y^j_t \equiv -q^j_{1,t} = d^j z_t + d^j \kappa(s_t) - \frac{1}{2}d^j \Lambda(s_t)\Omega(s_t)\Lambda(s_t)'d^{j'}.$$

With the normalization, $d^j \kappa(s) = \frac{1}{2}d^j \Lambda(s)\Omega(s)\Lambda(s)'d^{j'}$, this equation simplifies to $y^j_t = d^j z_t$ and bond prices satisfy (10) with $A^j_1(s_t) = 0$, and $B^j_1(s_t) = d^j$.

For the case where $k > 1$, I use the law of iterated expectations to rewrite (A1) as

$$1 = E\left[E[\exp(-d^j x_{t+1} + q^j_{k-1,t+1} - q^j_{k,t})|\mathcal{I}_t, s_{t+1}]\Big|\mathcal{I}_t\right]. \tag{A2}$$

Under the proposed solution, the joint distribution of x_{t+1} and $q^j_{k-1,t+1}$ conditioned on $\{\mathcal{I}_t, s_{t+1}\}$ is normal,

so the inner conditional expectation can be written as

$$\exp\left(E[(-d^j x_{t+1} + q^j_{k-1,t+1} - q^j_{k,t})|\mathcal{I}_t, s_{t+1}] + \tfrac{1}{2}Var(-d^j x_{t+1} + q^j_{k-1,t+1}|\mathcal{I}_t, s_{t+1})\right),$$

Substituting for x_{t+1} and $q^j_{k-1,t+1}$ using (8), (10) and the fact that $y^j_t = d^j z_t$, and combining the result with (A2) gives

$$\exp\left(d^j \kappa(s_t)\right) = E\left[\exp\left(E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}] + \Gamma^j_{k-1}(s_{t+1}, s_t)\right)\Big| \mathcal{I}_t\right], \tag{A3}$$

where $\delta^j_{k,t+1} \equiv q^j_{k-1,t+1} - q^j_{k,t} - y^j_t$ and

$$\Gamma^j_{k-1}(s_{t+1}, s_t) \equiv \tfrac{1}{2}\left(d^j \Lambda(s_t) + B^j_{k-1}(s_{t+1})\right)\Omega(s_t)\left(d^j \Lambda(s_t) + B^j_{k-1}(s_{t+1})\right)'.$$

Notice that $E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}]$ only depends on period $t+1$ information via s_{t+1}. Thus, given $\mathcal{I}_t$, there are S possible values for $E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}]$, corresponding to the S-states of the Markov process. Let $\Phi^j_{k,t}$ be an $S \times 1$ vector with i'th. element equal to $E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1} = i]$. (A3) can now be written as

$$\exp\left(d^j \kappa(s)\right) = E\left[\exp\left(\ell_{s_{t+1}}\Phi^j_{k,t} + \Gamma^j_{k-1}(s_{t+1}, s)\right)\Big| \mathcal{I}_t\right] = \sum_{\tilde{s}=1}^{S}\Pi_{\tilde{s},s}\exp\left(\ell_{\tilde{s}}\Phi^j_{k,t} + \Gamma^j_{k-1}(\tilde{s}, s)\right),$$

where ℓ_s denotes an $1 \times S$ vector with a one at element s and zeros elsewhere, and $\Pi_{\hat{s},s} \equiv \Pr(s_{t+1} = \tilde{s}|s_t = s)$. Stacking the S versions of this equation gives

$$\begin{bmatrix} e^{d^j \kappa(1)} \\ \vdots \\ e^{d^j \kappa(s)} \\ \vdots \\ e^{d^j \kappa(S)} \end{bmatrix} = \begin{bmatrix} \Pi_{1,1}e^{\Gamma^j_{k-1}(1,1)} & & \cdots & & \Pi_{S,1}e^{\Gamma^j_{k-1}(S,1)} \\ \vdots & \ddots & & & \vdots \\ \Pi_{1,s}e^{\Gamma^j_{k-1}(1,s)} & & \Pi_{s,s}e^{\Gamma^j_{k-1}(s,s)} & & \Pi_{S,s}e^{\Gamma^j_{k-1}(S,s)} \\ \vdots & & & \ddots & \vdots \\ \Pi_{1,s}e^{\Gamma^j_{k-1}(1,S)} & & \cdots & & \Pi_{S,S}e^{\Gamma^j_{k-1}(S,S)} \end{bmatrix} \begin{bmatrix} e^{\ell_1 \Phi^j_{k,t}} \\ \vdots \\ e^{\ell_s \Phi^j_{k,t}} \\ \vdots \\ e^{\ell_S \Phi^j_{k,t}} \end{bmatrix},$$

or, more compactly,

$$\exp_v\left(\mathcal{K}^j\right) = \mathcal{P}^j_{k-1}\exp_v\left(\Phi^j_{k,t}\right),$$

where $\exp_v(\chi)$ denotes the exponential operator applied to each element in a vector χ. Provided $\mathcal{P}^j_{k-1}$ is nonsingular, we can rewrite this matrix equation as

$$\Phi^j_{k,t} = \ln_v\left(\left(\mathcal{P}^j_{k-1}\right)^{-1}\exp_v\left(\mathcal{K}^j\right)\right), \tag{A4}$$

where $\ln_v(\chi)$ denotes the log operator applied to each element in a vector χ. Notice that the right hand side of this expression only depends on $\kappa(.)$, $\Gamma^j_{k-1}(.,.)$ and the transition probabilities. Thus, expected excess returns, $E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}]$, do not depend on z_t even though $z_t \in I_t$. The equilibrium term premium can now be easily found by applying the law of iterated expectations to the definition $\theta^j_{k,t} \equiv E[\delta^j_{k,t+1}|\mathcal{I}_t]$:

$$\begin{aligned} \theta^j_{k,t} &= E\left[E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}]\Big| \mathcal{I}_t\right] \\ &= \ell_{s_t}\Pi\Phi^j_{k,t} \end{aligned} \tag{A5}$$

where Π is the matrix of Markov transition probabilities with $\Pi_{i,j} \equiv \Pr(s_{t+1} = i|s_t = j)$. Combining (A4)

with (A5) gives

$$\theta_k^j(s) = \ell_s \Pi' \ln_v \left(\left(\mathcal{P}_{k-1}^j \right)^{-1} \exp_v \left(\mathcal{K}^j \right) \right), \tag{A6}$$

which is the value of the term premium in state s consistent with the equilibrium condition in (A1), the dynamics of x_{t+1}, and the bond price solution in (10).

Step 2: To find the restrictions on $A_k^j(s)$ and $B_k^j(s)$ consistent with this expression for the term premium, I use (10) to substitute for $q_{k-1,t+1}^j$, $q_{k,t}^j$ and y_t^j in the identity $\theta_{k,t}^j \equiv E[\delta_{k,t+1}^j | \mathcal{I}_t]$. This gives

$$
\begin{aligned}
\theta_{k,t}^j &= A_k^j(s_t) + B_k^j(s_t) z_t - E\left[A_{k-1}^j(s_{t+1}) + B_{k-1}^j(s_{t+1}) z_{t+1} | \mathcal{I}_t \right] - d^j z_t \\
&= A_k^j(s_t) - E_{s_t}\left[A_{k-1}^j(\tilde{s}) + B_{k-1}^j(\tilde{s},) \left(\mu(\tilde{s}) - \alpha(s_t)\mu(s_t) \right) \right] + \left(B_k^j(s_t) - E_{s_t}\left[B_{k-1}^j(\tilde{s})\alpha(s_t) \right] - d^j \right) z_t
\end{aligned}
$$

where the second line follows from (8). Since the k-period term premium is a function of k and s_t, the term in parenthesis must equal a vector of zeros. Hence

$$
\begin{aligned}
A_k^j(s_t) &= E_{s_t}\left[A_{k-1}^j(\tilde{s}) + B_{k-1}^j(\tilde{s},) \left(\mu(\tilde{s}) - \alpha(s_t)\mu(s_t) \right) \right] + \theta_{k,t}^j, \\
B_k^j(s_t) &= E_{s_t}\left[B_{k-1}^j(\tilde{s})\alpha(s_t) \right] + d^j.
\end{aligned}
$$

The recursions for $A_k^j(s_t)$ and $B_k^j(s_t)$ in (11) follow from these equations with $\theta_{k,t}^j = \theta_k^j(s_t)$.

Comment: The form for the term premium function in (A6) assumes that $\mathcal{P}_{k-1}^j$ is nonsingular. Alternative derivations for the term premium function exist for cases were $\mathcal{P}_{k-1}^j$ is singular. To illustrate, consider the case where $S = 2$, $\alpha(1) = \alpha(2)$, and $\Pi_{11} + \Pi_{22} = 1$. Although $\mathcal{P}_{k-1}^j$ is singular for all k under these circumstances, the term premium function can still be readily found. In particular, it is easy to show that bond prices satisfy

$$
\begin{aligned}
q_{k,t}^j &= A_k^j(s_t) + B_k^j z_t, \\
A_k^j(s) &= \tfrac{1}{2} d^j \Lambda(s) \Omega(s) \left(d^j \Lambda(s) \right)' - \tfrac{1}{2}\left(d^j \Lambda(s) + B_{k-1}^j \right) \Omega(s) \left(d^j \Lambda(s) + B_{k-1}^j \right)' \\
&\quad - \ln\left\{ E_s \exp(-A_{k-1}^j(\tilde{s}) - B_{k-1}^j(\mu(\tilde{s}) - \alpha\mu(s))) \right\}, \\
B_k^j &= B_{k-1}^j \alpha + d^j,
\end{aligned}
$$

with $A_0^j(s) = 0$ and $B_0^j = [0,0]$. (To verify this, substitute the proposed solution into (A1) and equate coefficients.) The term premium function is therefore

$$
\begin{aligned}
\theta_k(s) &= \frac{1}{2}\left[d^j \Lambda(s)\Omega(s)(d^j\Lambda(s))' - (d^j\Lambda(s) + B_{k-1}^j)\Omega(s)(d^j\Lambda(s) + B_{k-1}^j)' \right] \\
&\quad - \left\{ \ln\left\{ E_s \exp(-A_{k-1}^j(\tilde{s}) - B_{k-1}^j(\mu(\tilde{s}) - \alpha\mu(s))) \right\} - E_s\left[A_{k-1}^j(\tilde{s}) + B_{k-1}^j \left(\mu(\tilde{s}) - \alpha\mu(s) \right) \right] \right\}.
\end{aligned}
$$

In principle, the bond price solution can be found in a similar manner for other cases where $\mathcal{P}_{k-1}^j$ is singular. In practice, singularity in $\mathcal{P}_{k-1}^j$ did not turn out to be a problem for the two and three-state models estimated in the paper so equilibrium bond prices are identified using (10) and (11) with (A6).

Risk Premia

If the joint distribution of x_{t+1} and q_{t+1}^j conditioned on $\mathcal{I}_t$ is normal, the equilibrium condition in (A1) with $k = 2$, becomes

$$E[(-d^j x_{t+1} + q_{t+1}^j - q_{2,t}^j)|\mathcal{I}_t] = -\tfrac{1}{2}Var(-d^j x_{t+1} + q_{t+1}^j|\mathcal{I}_t).$$

Substituting for x_{t+1} on the right hand side and simplifying, using the fact that $d^j \kappa(s) = \tfrac{1}{2}Var(-d^j x_{t+1})$, gives

$$\theta_{2,t}^j + \tfrac{1}{2}Var(q_{t+1}^j|\mathcal{I}_t) = Cov(d^j x_{t+1}, q_{t+1}^j|\mathcal{I}_t).$$

This expression only holds exactly in the case where; (i) there is a single state, or (ii) where $\mu(s) = \mu$. Under other circumstances, the conditional distribution for q_{t+1}^j is non-normal because q_{t+1}^j depends on s_{t+1} and u_{t+1}. The expression above is the basis for the approximations shown in the first lines of (13) and (14). The second lines are derived by substituting for x_{t+1} and $q_{t+1}^j = -d^j z_{t+1}$ in the covariance term:

$$Cov(d^j x_{t+1}, q_{t+1}^j|\mathcal{I}_t) = -d^j \Lambda(s_t) E[e_{t+1}, \mu(s_{t+1})'|\mathcal{I}_t]d^{j\prime} - d^j \Lambda(s_t)\Omega(s_t)d^{j\prime}.$$

The first in this expression equals zero because e_{t+1} is independent of s_{t+1}. The second lines in (13) and (14) are therefore equal to $-d^j \Lambda(s_t)\Omega(s_t)d^{j\prime}$ for $j = r$ and $j = n$ respectively.

The equation for the inflation risk premium in (15) is derived by combining the nominal and real versions of (A1) with $k = 1$:

$$E\left[\exp(-d^r x_{t+1} + y_t^r)|\mathcal{I}_t\right] = E\left[\exp(-d^n x_{t+1} + y_t^n)|\mathcal{I}_t\right].$$

Using the conditional normality of x_{t+1}, this expression simplifies to

$$
\begin{aligned}
\psi_t \equiv y_t^n - y_t^r - E[\Delta p_{t+1}|\mathcal{I}_t] &= \tfrac{1}{2}Var(d^r x_{t+1}|\mathcal{I}_t) - \tfrac{1}{2}Var(d^n x_{t+1}|\mathcal{I}_t), \\
&= -Cov(d^r x_{t+1}, (d^n - d^r)x_{t+1}|\mathcal{I}_t) - \tfrac{1}{2}Var((d^n - d^r)x_{t+1}|\mathcal{I}_t), \\
&= Cov(m_{t+1}, \Delta p_{t+1}|\mathcal{I}_t) - \tfrac{1}{2}Var(\Delta p_{t+1}|\mathcal{I}_t),
\end{aligned}
\tag{A7}
$$

which is the form of (15). Substituting for the variance terms in the first line gives

$$\psi_t = \tfrac{1}{2}d^r \Lambda(s_t)\Omega(s_t)\left(d^r \Lambda(s_t)\right)' - \tfrac{1}{2}d^n \Lambda(s_t)\Omega(s_t)\left(d^n \Lambda(s_t)\right)' = \psi(s_t) \tag{A8}$$

so the inflation risk premium can be written as a function of s_t.

Section V considers the multi-period versions of the term and inflation risk premia. Combining the definition $\theta_{k,t}^j \equiv E[q_{k-1,t+1}^j - q_{k,t}^j - y_t^j|\mathcal{I}_t]$ with the equilibrium relation $\theta_{k,t}^j = \theta_k(s_t)$, we have

$$-q_{k,t}^j = E\left[-q_{k-1,t+1}^j|\mathcal{I}_t\right] + y_t^j + \theta_k^j(s_t).$$

Iterating this equation forward (with $q_{0,t}^j = 0$) gives

$$q_{k,t}^j = -\sum_{i=0}^{k-1} E\left[y_{t+i}^j + \theta_{k-i}^j(s_{t+i})|\mathcal{I}_t\right]. \tag{A9}$$

We can use this equation to show that the multi-period term premium $\Theta_{k,t}^j \equiv \tfrac{1}{h}E[q_{k-h,t+h}^j - q_{k,t}^j|\mathcal{I}_t] - y_{h,t}^j$ is a just a function of k and s_t. Substituting (A9) into the definition for $\Theta_{k,t}^j$ and simplifying gives

$$\Theta_{k,t}^j = \frac{1}{h}\sum_{i=0}^{k-1} E\left[\theta_{k-i}^j(s_{t+i})|\mathcal{I}_t\right] - \frac{1}{h}\sum_{i=0}^{k-h-1} E\left[\theta_{k-h-i}^j(s_{t+h+i})|\mathcal{I}_t\right] - \frac{1}{h}\sum_{i=0}^{h-1} E\left[\theta_{h-i}^j(s_{t+i})|\mathcal{I}_t\right].$$

Since s_t follows a first order Markov process, all the conditional expectations in this expression are functions

of the current state, so $\Theta^j_{k,t}$ can be written as a function of the current state; $\Theta^j_k(s_t)$.

To derive equations (17) and (18), I write the multi-period premium in terms of yields and combine the result with $\Theta^j_{k,t} = \Theta^j_k(s_t)$ to give

$$\Theta^j_k(s_t) \equiv E[\tau y^j_{k,t} - (\tau - 1)y^j_{k-h,t+h}|\mathcal{I}_t] - y^j_{h,t}, \tag{A10}$$

where $\tau = k/h$. A simple rearrangement of this expression gives (18). To derive (17), I rewrite (A10) as

$$\tau\left(y^j_{k,t} - y^j_{h,t}\right) = \Theta^j_k(s_t) + (\tau - 1)E[y^j_{k-h,t+h} - y^j_{h,t+h}|\mathcal{I}_t] + (\tau - 1)E[\Delta^h y^j_{h,t+1}|\mathcal{I}_t]$$

and iterate forward.

The equation for the multi-period inflation risk premium in (22) comes from taking the difference between the nominal and real and nominal versions of (A9):

$$y^n_{k,t} - y^r_{k,t} = \frac{1}{k}\sum_{i=0}^{k-1} E\left[y^n_{t+i} - y^r_{t+i} + \theta^n_{k-i}(s_{t+i}) - \theta^r_{k-i}(s_{t+i})|\mathcal{I}_t\right].$$

Substituting for $y^n_{t+i} - y^r_{t+i}$ with (A8), gives

$$y^n_{k,t} - y^r_{k,t} = E[\pi_{k,t+k}|\mathcal{I}_t] + \Psi_{k,t},$$

$$\Psi_{k,t} \equiv \frac{1}{k}\sum_{i=0}^{k-1} E[\theta^n_{k-i}(s_{t+i}) - \theta^r_{k-i}(s_{t+i}) + \varphi(s_{t+i})|\mathcal{I}_t], \tag{A11}$$

which is the form of (22). As above, all the conditional expectations in the second line can be written as a function of s_t so $\Psi_{k,t}$ can also be expressed as a function of the current state; $\Psi^j_k(s_t)$.

Transforming the Model

The analysis in Section V uses long-horizon expectations of future yields and inflation. Computing these expectations requires forecasts of z_t which follows

$$z_{t+1} = \mu(s_{t+1}) + \alpha(s_t)(z_t - \mu(s_t)) + w_{t+1}, \tag{A12}$$

where $w_{t+1} = \Omega^{1/2}(s_t)u_{t+1}$. Finding long-horizon forecasts from this process is complicated by the presence of switching in $\alpha(.)$, because it effectively makes the process for z_t nonlinear. To alleviate this problem, I transform the model.

First, I represent the Markov process for s_t as a Vector Autoregression. Let $\zeta_t = \ell'_{s_t}$ so $\zeta_t = [1, 0.....]'$ if $s_t = 1$, and $\zeta_t = [0, 1, 0.....]'$ if $s_t = 2$, and so on (as in Hamilton 1994, p 679). We can now represent the S-state Markov process by

$$\zeta_{t+1} = \Pi\zeta_t + v_{t+1}, \tag{A13}$$

where $v_{t+1} = \zeta_{t+1} - E[\zeta_{t+1}|\zeta_t, \zeta_{t-1,..}]$.

Next, let $\eta_t = \zeta_t \otimes \gamma_t$ where $\gamma_t \equiv z_t - \mu(s_t)$. We can now rewrite (A12) as

$$\gamma_{t+1} = \alpha\eta_t + w_{t+1}, \tag{A14}$$

where $\alpha = [\alpha(1), ..\alpha(s), ...\alpha(S)]$. This equation gives us a linear forecast for γ_{t+1} in terms of η_t. To determine the multi-step forecasts we need to derive the dynamics for η_t. For this I use (A13) and (A14) to substitute

39

for ζ_{t+1} and γ_{t+1} in the definition of η_{t+1}. After some matrix algebra, this gives us

$$
\begin{aligned}
\eta_{t+1} &= (\Pi \otimes \alpha)(\zeta_t \otimes \eta_t) + \epsilon_{t+1}, \\
\epsilon_{t+1} &= v_{t+1} \otimes (\alpha \eta_t) + v_{t+1} \otimes w_{t+1} + (\Pi \zeta_t) \otimes w_{t+1}.
\end{aligned}
$$

It is straightforward to check that $\zeta_t \otimes \eta_t$ can be written as $\Upsilon \eta_t$ for a matrix Υ comprising of ones and zeros.[1] Combining these results, we can write (A13) - (A15) as a vector system:

$$
\begin{bmatrix} \gamma_{t+1} \\ \eta_{t+1} \\ \zeta_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \alpha & \mathbf{0} \\ \mathbf{0} & (\Pi \otimes \alpha)\Upsilon & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Pi \end{bmatrix} \begin{bmatrix} \gamma_t \\ \eta_t \\ \zeta_t \end{bmatrix} + \begin{bmatrix} w_{t+1} \\ \epsilon_{t+1} \\ v_{t+1} \end{bmatrix},
$$

or, more compactly,

$$
\mathcal{Z}_{t+1} = \mathcal{A}\mathcal{Z}_t + \mathcal{E}_{t+1}. \tag{A15}
$$

Note that under the assumptions of the model, $E[\mathcal{E}_{t+1}|\mathcal{I}_t]$ equals a vector of zeros. Hence (A15) implies that $E[\mathcal{Z}_{t+i}|\mathcal{I}_t] = \mathcal{A}^i \mathcal{Z}_t$.

The final step is to show that expected inflation and equilibrium yields can be written in terms of the elements of $\mathcal{Z}_t$ rather than s_t and z_t. Substituting $\gamma_t = z_t - \mu(s_t)$ into the bond pricing solution in (10), gives

$$
y_{k,t}^j = \tfrac{1}{k} C_k^j(s_t) + \tfrac{1}{k} B_k^j(s_t)\gamma_t,
$$

where $C_k^j(s) = [A_k^j(s) - B_k^j(s)\mu(s)]$ is a state-dependent scalar. Now note that $C_k^j(s_t)$ can be written as $[C_k^j(1), ... C_k^j(s), ... C_k^j(S)]\zeta_t$ and $B_k^j(s_t)\gamma_t$ can be written as $[B_k^j(1), ... B_k^j(s), ... B_k^j(S)]\eta_t$. Making these substitutions in the equation above gives

$$
\begin{aligned}
y_{t,k}^j &= \tfrac{1}{k}\left[0, 0, B_k^j(1), ..., B_k^j(s), ... B_k^j(S), C_k^j(1), ... C_k^j(s), ... C_k^j(S)\right]\mathcal{Z}_t, \\
&= \mathcal{H}_k^j, \mathcal{Z}_t. \tag{A16}
\end{aligned}
$$

The model identifies the expected rate of inflation as $E[\Delta p_{t+1}|\mathcal{I}_t] = \kappa_p(s_t) + \mu_p(s_t) + [0,1]\gamma_t$. Proceeding as above, we can write expected inflation in terms of $\mathcal{Z}_t$ as

$$
\begin{aligned}
E[\Delta p_{t+1}|\mathcal{I}_t] &= [0, 1, 0, ...0, ...0, (\kappa_p(1) + \mu_p(1)), .., (\kappa_p(s) + \mu_p(s)), ... (\kappa_p(S) + \mu_p(S))]\mathcal{Z}_t, \\
&= \mathcal{G}\mathcal{Z}_t. \tag{A17}
\end{aligned}
$$

Variance Decompositions

Table 4 reports estimates of the variance decompositions in (19) and (20) for the spread between long- and short-term yields $\nabla y_{k,t}^j = y_{k,t}^j - y_{h,t}^j$. Using (A15) and (A16), we have

$$
\begin{aligned}
Var(\nabla y_{k,t}^j) &= \left(\mathcal{H}_k^j - \mathcal{H}_h^j\right)\mathcal{S}\left(\mathcal{H}_k^j - \mathcal{H}_h^j\right)', \\
Cov(E_t y_{k-h,t+h}^j - y_{k,t}^j, \nabla y_{k,t}^j) &= \left(\mathcal{H}_{k-h}^j \mathcal{A}^h - \mathcal{H}_k^j\right)\mathcal{S}\left(\mathcal{H}_k^j - \mathcal{H}_h^j\right)',
\end{aligned}
$$

---

[1] For example, in the $S = 2$ case the possible values of ζ_t are [1,0]' and [0,1]', corresponding to $s_t = 1$, and $s_t = 2$. When $s_t = 1$, $\eta_t = [\gamma_t', 0, 0]'$ and $\zeta_t \otimes \eta_t = [\gamma_t', 0, 0, 0, 0, 0]'$. When $s_t = 2$, $\eta_t = [0, 0, \gamma_t']'$ and $\zeta_t \otimes \eta_t = [0, 0, 0, 0, 0, \gamma_t']'$. So in this case

$$
\Upsilon' = \begin{bmatrix} \mathbf{I}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{0}_2 & \mathbf{I}_2 \end{bmatrix}
$$

where $\mathbf{I}_2$ denotes the 2×2 identity matrix and $\mathbf{0}_2$ a 2×2 matrix of zeros.

$$\sum_{i=1}^{\tau-1}(\tfrac{\tau-i}{\tau})Cov(E_t\Delta^h y_{h,t+hi}^j,\nabla y_{k,t}^j)=\sum_{i=1}^{\tau-1}(\tfrac{\tau-i}{\tau})\mathcal{H}_h^j\left(\mathcal{A}^{hi}-\mathcal{A}^{h(i-1)}\right)\mathcal{S}\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)'$$

where $\mathcal{S}$ denotes the covariance of $\mathcal{Z}_t$. Column (v) of table 4 reports

$$a_1=\frac{\sum_{i=1}^{\tau-1}(\tfrac{\tau-i}{\tau})Cov(E_t\Delta^h y_{h,t+hi}^j,\nabla y_{k,t}^j)}{Var(\nabla y_{k,t}^j)}=\frac{\sum_{i=1}^{\tau-1}(\tfrac{\tau-i}{\tau})\mathcal{H}_h^j\left(\mathcal{A}^{hi}-\mathcal{A}^{h(i-1)}\right)\mathcal{S}\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)'}{\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)\mathcal{S}\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)'}$$

and

$$b_1=\frac{(\tau-1)Cov(E_t y_{k-h,t+h}^j-y_{k,t}^j,\nabla y_{k,t}^j)}{Var(\nabla y_{k,t}^j)}=\frac{(\tau-1)\left(\mathcal{H}_{k-h}^j\mathcal{A}^h-\mathcal{H}_k^j\right)\mathcal{S}\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)'}{\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)\mathcal{S}\left(\mathcal{H}_k^j-\mathcal{H}_h^j\right)'}$$

for $h=12$, with $\tau=k/12$. The matrices $\mathcal{A}$ and $\mathcal{H}_k^j$ are calculated directly from the estimates of the three-state model. The covariance matrix $\mathcal{S}$ is computed by Monte Carlo simulation of the $\mathcal{Z}_t$ process (based on the three state estimates) over 120,000 observations (i.e. 10,000 years of monthly data).

The k-period rate of expected inflation can be computed from (A15) and (A17) as

$$E[\pi_{k,t+k}|\mathcal{I}_t]\equiv\frac{1}{k}\sum_{i=1}^k E[\Delta p_{t+i}|\mathcal{I}_t]=\frac{1}{k}\sum_{i=1}^k\mathcal{G}\mathcal{A}^i\mathcal{Z}_t=\mathcal{F}_k\mathcal{Z}_t. \tag{A18}$$

Combining this equation with (A16), gives the following equation for the multi-period inflation risk premium:

$$\Psi_{k,t}=[\mathcal{H}_k^n-\mathcal{H}_k^r-\mathcal{F}_k]\,\mathcal{Z}_t. \tag{A19}$$

Since $\Psi_{k,t}$ is a function of s_t (see equation (A11) above), only the last S columns of $[\mathcal{H}_k^n-\mathcal{H}_k^r-\mathcal{F}_k]$ are nonzero and comprise $[\Psi_k(1),...,\Psi_k(s),...,\Psi_k(S)]$. The upper panel of Figure 5 plots $\Psi_k(s)$ against k for $s=\{1,2,3\}$.

To compute the variance decompositions reported in Table 5, I use (A16), (A18) and (A19) to write

$$\frac{Cov\left(y_{k,t}^n-y_{k,t}^r,E_t\pi_{k,t+k}\right)}{Var(y_{k,t}^n-y_{k,t}^r)}=\frac{(\mathcal{H}_k^n-\mathcal{H}_k^r)\mathcal{S}(\mathcal{F}_k)'}{(\mathcal{H}_k^n-\mathcal{H}_k^r)\mathcal{S}(\mathcal{H}_k^n-\mathcal{H}_k^r)'},$$

$$\frac{Cov\left(y_{k,t}^n,E_t\pi_{k,t+k}\right)}{Var(y_{k,t}^n)}=\frac{\mathcal{H}_k^n\mathcal{S}(\mathcal{F}_k)'}{\mathcal{H}_k^n\mathcal{S}(\mathcal{H}_k^n)'},$$

$$\frac{Cov\left(y_{k,t}^n,y_{k,t}^r\right)}{Var(y_{k,t}^n)}=\frac{\mathcal{H}_k^n\mathcal{S}(\mathcal{H}_k^r)'}{\mathcal{H}_k^n\mathcal{S}(\mathcal{H}_k^n)'},$$

$$\frac{Cov\left(y_{k,t}^n,\Psi_{k,t}\right)}{Var(y_{k,t}^n)}=\frac{\mathcal{H}_k^n\mathcal{S}(\mathcal{H}_k^n-\mathcal{H}_k^r-\mathcal{F}_k)'}{\mathcal{H}_k^n\mathcal{S}(\mathcal{H}_k^n)'},$$

The estimates in the table are based on the $\mathcal{H}_k^j$ and $\mathcal{F}_k$ calculated directly from the estimates of the three state model, and $\mathcal{S}$ is generated by the Monte Carlo simulation.

Estimation Details

The state space form of the model to be estimated is given by equations (8) and (16)

$$\hat{y}_t = \mathbf{A}(s_t) + \mathbf{B}(s_t)z_t + \xi_t, \tag{16}$$

$$z_{t+1} = \mu(s_{t+1}) + \alpha(s_t)(z_t - \mu(s_t)) + \Omega^{1/2}(s_t)u_{t+1}, \tag{8}$$

where $\xi_t \sim i.i.d.N(0, \mathbf{R})$ and $u_{t+1} \sim i.i.d.N(0, I)$. For the case with no regime-switching and pricing errors in all the yields (i.e., $\mathbf{R}$ has full rank), the parameter estimates can be found by maximizing the log sample likelihood that is formed recursively from the Kalman Filter as in Pennarchi (1991). This is the method I initially used to estimate the one-state version of the model. When switching is present and $\mathbf{R}$ has full rank, the Kalman Filter needs to be modified to accommodate the unobserved states, s_t. Kim (1993) describes how the sample likelihood can be approximated in this case by combining Kalman Filter equations with the Hamilton (1988) algorithm. This is the method I initially used to estimate the two and three-state versions of the model.

The exact sample likelihood function can be found when there are no pricing errors in the equations for three-year real and nominal yields. Let $\hat{y}_t = [\hat{y}'_{1,t}, \hat{y}'_{2,t}]'$ where $\hat{y}'_{1,t}$ is the vector of 3-year yields, with $\mathbf{A}(s) = [\mathbf{A}'_1(s), \mathbf{A}'_2(s)]', \mathbf{B}(s) = [\mathbf{B}'_1(s), \mathbf{B}'_2(s)]'$ and $\xi_t = [0, \xi'_{2,t}]'$ partitioned conformably. Given the recursive factor structure of the model, $\mathbf{B}_1(s)$ has rank 2, so the risk factors can be found directly as $z_t = \mathbf{B}_1^{-1}(s_t)(\hat{y}_{1,t} - \mathbf{A}_1(s_t))$. We can now rewrite the model as

$$\hat{y}_{1,t+1} = \mathbf{C}_1(s_{t+1}, s_t) + \mathbf{D}_1(s_{t+1}, s_t)\hat{y}_{1,t} + \varepsilon_{t+1}, \tag{A20}$$

$$\hat{y}_{2,t+1} = \mathbf{C}_2(s_{t+1}) + \mathbf{D}_2(s_{t+1})\hat{y}_{1,t+1} + \xi_{2,t+1}, \tag{A21}$$

where $\varepsilon_{t+1} \sim N(0, \mathbf{B}_1(s_{t+1})\Omega(s_t)\mathbf{B}'_1(s_{t+1}))$ with

$$\mathbf{C}_1(s_{t+1}, s_t) = \mathbf{B}_1(s_{t+1})(\mu(s_{t+1}) - \alpha(s_t)\mu(s_t)) + \mathbf{A}_1(s_{t+1}) - \mathbf{D}_1(s_{t+1}, s_t)\mathbf{A}_1(s_t),$$

$$\mathbf{D}_1(s_{t+1}, s_t) = \mathbf{B}_1(s_{t+1})\alpha(s_t)\mathbf{B}_1^{-1}(s_t),$$

$$\mathbf{C}_2(s_{t+1t}) = \mathbf{A}_2(s_{t+1}) - \mathbf{D}_2(s_{t+1})\mathbf{A}_1(s_{t+1}),$$

$$\mathbf{D}_2(s_{t+1}) = \mathbf{B}_2(s_{t+1})\mathbf{B}_1^{-1}(s_{t+1}).$$

(A20) and (A21) constitute a standard Markov switching model that can be estimated by maximum likelihood using the Hamilton (1988) algorithm. The parameter estimates reported in Table 2 were obtained in this manner.

Identification

Consider the model

$$x_{t+1} = \phi + \kappa(s_t) + \bar{z}_t + \Lambda(s_t)\Omega^{1/2}(s_t)u_{t+1} + \Sigma^{1/2}e_{t+1}, \tag{A22}$$

$$\bar{z}_{t+1} = \varphi(s_{t+1}) + \alpha(s_t)(\bar{z}_t - \varphi(s_t)) + \Omega^{1/2}(s_t)u_{t+1}, \tag{A23}$$

where e_{t+1} is a vector of i.i.d. $N(0,1)$ shocks, $\phi' = [\phi_m, \phi_p]$ is a vector of constants and $\varphi(s) = \mu(s) - \phi$. If we choose ϕ_m and ϕ_p such that $d^j\phi = \frac{1}{2}d^j\Sigma d^{j'}$ for $j = \{n, r\}$, equilibrium bond prices satisfy (10) with $\bar{z}_t$ replacing z_t and $\varphi(s)$ replacing $\mu(s)$ in the parameter recursions (11).

To substantiate this claim, I first use (A22) to substitute for x_{t+1} in (A1) for the case where $k = 1$. Simplifying the resulting expression gives,

$$y_t^j = d^j\bar{z}_t + \left(d^j(\kappa(s_t) + \phi) - \tfrac{1}{2}d^j\Lambda(s_t)\Omega(s_t)(\Lambda(s_t)d^j)' - \tfrac{1}{2}d^j\Sigma d^{j'}\right) = d^j\bar{z}_t.$$

Hence, short rates follow the same process except that the long run mean within each state is given by $d^j\varphi(s)$ rather than $d^j\mu(s)$.

Next, consider the prices of long term bonds. Using (10) with ($\bar{z}_t$ replacing z_t) to substituting for $q^j_{k-1,t+1}$, $q^j_{k,t}$ and y^j_t in the identity $\theta^j_{k,t} \equiv E[\delta^j_{k,t+1}|\mathcal{I}_t]$, gives

$$
\begin{aligned}
\theta^j_{k,t} &= A^j_k(s_t) + B^j_k(s_t)\bar{z}_t - E\left[A^j_{k-1}(s_{t+1}) + B^j_{k-1}(s_{t+1})\bar{z}_{t+1}|\mathcal{I}_t\right] - d^j\bar{z}_t, \\
&= A^j_k(s) - E_s\left[A^j_{k-1}(\tilde{s}) + B^j_{k-1}(\tilde{s},)\,(\varphi(\tilde{s}) - \alpha(s)\varphi(s))\right] + \left(B^j_k(s) - E_s\left[B^j_{k-1}(\tilde{s})\alpha(s)\right] - d^j\right)\bar{z}_t, \\
&= \theta_k(s_t),
\end{aligned}
$$

where the third line follows from the recursions for $A^j_k(s)$ and $B^j_k(s)$ in (11) with $\varphi(s)$ replacing $\mu(s)$. Thus, the term premium continues to be a functions of k and s_t. All that now remains is to verify that the term premium function remains the same. To show this, notice from (A22) and (A23) that

$$
\begin{aligned}
E[\exp(-d^j x_{t+1} + q^j_{k-1,t+1} - q^j_{k,t})|\mathcal{I}_t, s_{t+1}] = \exp\left(-d^j\kappa(s_t) + E[\delta^j_{k,t+1}|\mathcal{I}_t, s_{t+1}]\right) \\
\times \exp\left(\tfrac{1}{2}(d^j\Lambda(s_t) + B^j_{k-1}(s_{t+1}))\Omega(s_t)(d^j\Lambda(s_t) + B^j_{k-1}(s_{t+1}))' + \tfrac{1}{2}d^j\Sigma d^{j\prime} - d^j\phi\right).
\end{aligned}
$$

Since the last two terms cancel, the implications of the equilibrium condition in (A2) can still be written as (A4) and the term premium continues to be determined by (A6).

While the value of Σ (or equivalently ϕ) has no impact on the term premia, it does affect the level of the inflation risk premium. From (A7) we have

$$
\begin{aligned}
\psi_t &= \tfrac{1}{2}Var(d^r x_{t+1}|\mathcal{I}_t) - \tfrac{1}{2}Var(d^n x_{t+1}|\mathcal{I}_t), \\
&= \tfrac{1}{2}d^r\Lambda(s_t)\Omega(s_t)\,(d^r\Lambda(s_t))' - \tfrac{1}{2}d^n\Lambda(s_t)\Omega(s_t)\,(d^n\Lambda(s_t))' + \tfrac{1}{2}d^r\Sigma d^{r\prime} - \tfrac{1}{2}d^n\Sigma d^{n\prime}.
\end{aligned}
$$

The last two terms add a constant to the expression for the inflation risk premium derived in (A8) above.

Switching Tests

To test the null hypothesis for the null of an S_1-state model versus the alternative of a S_2-state model ($S_2 > S_1$), we need to deal with the presence of parameters that are only identified under the alternative. The idea behind the Gallant test is to first calculate estimates of the dependent variables from the S_2-state model using a range of values for the unidentified parameters (under the null). These estimates are then added to the S_1-state model and their significance judged according to an F-test. This procedure is implemented in four steps:

1. Let β_i represent a given set of values for the parameters of the S_2-state model. These values are chosen independently on a uniform grid for each parameter that ranges from 0.8 to 1.2 times the value of the maximum likelihood estimate. For $i = 1, 2, ...N$, use β_i to calculate the fitted values for real and nominal yields, $Y^i \equiv \{y^i_t\}^T_{t=1}$.

2. Extract d principal components from the matrix $Y \equiv [Y^1 \ldots Y^i \ldots Y^N]$ to the matrix $W \equiv \{w_t\}^T_{t=1}$. In practice, I set $N = 1000$ and $d = 5$.

3. Estimate the S_1-state model with the addition of these components, i.e.,

$$
\hat{y}_t = \mathbf{A}(s_t) + \mathbf{B}(s_t)z_t + \mathbf{C}w_{\mathbf{t}} + \xi_t,
$$

where $\mathbf{C}$ is a diagonal matrix. Call this the augmented S_1 model.

4. Let $\hat{y}^{S_1}_{t,}$ and $\hat{y}^{S_1^*}_{t,}$ denote the fitted values for the yields based on the S_1 model and the augmented

S_1model respectively. Compute the following residual sum of squares

$$RSS_{S_1} = \sum_{t=1}^{T}(\hat{y}_t - \hat{y}_{t,}^{S_1})'(\hat{y}_t - \hat{y}_{t,}^{S_1})$$

$$RSS_{S_1^*} = \sum_{t=1}^{T}(\hat{y}_t - \hat{y}_{t,}^{S_1^*})'(\hat{y}_t - \hat{y}_{t,}^{S_1^*}).$$

Under the null of no switching, the statistic

$$\frac{(RSS_{S_1} - RSS_{S_1^*})(T - d - k)}{RSS_{S_1^*}d}$$

is distributed $F(d, T - d - k)$ where k is the number of parameters in the S_1-state model.

To understand the practical problems with undertaking the Hansen (1992) test for the null of an S_1-state model versus the alternative of a S_2-state model ($S_2 > S_1$), consider the following partition of the parameter vector; $\beta = (\beta_a', \beta_d')'$, $\beta_a = (\beta_b', \beta_c')'$. We are interested in testing null that $\beta_b = 0$, where β_c is the vector of nuisance parameters unidentified under the null. β_d is the vector of parameters identified under both the null and the alternative hypothesis. Hansen's test statistic is

$$LR_T^* = \sup_{\beta_a}(LR_T(\beta_a)/\sqrt{V_T(\beta_a)}),$$

where

$$LR_T(\beta_a) = L_T(\beta_a, \hat{\beta}_d(\beta_a)) - L_T(0, \beta_b, \hat{\beta}_d(0, \beta_b)),$$

$$V_T(\beta_a) = \sum_{t=1}^{T} q_t(\beta_a, \hat{\beta}_d(\beta_a))^2,$$

$$q_t(\beta_a, \hat{\beta}_d(\beta_a)) = l_t(\beta_a, \hat{\beta}_d(\beta_a)) - l_t(0, \beta_b, \hat{\beta}_d(0, \beta_b)) - \frac{1}{T}LR_T(\beta_a),$$

with $\hat{\beta}_d(\beta_a) = \arg\max_{\beta_d} L_T(\beta_a, \beta_d)$ and $L_T(\beta_a, \beta_d) = \sum_{t=1}^{T} l_t$ is the sample likelihood.

Consider the case where $S_1 = 1$, and $S_2 = 2$. For the model studied here, the vector β_d includes the elements in $\{(\kappa(1), \mu(1), \alpha(1), \Lambda(1), \Omega(1), \Sigma\}$, $\beta_b = \Pi_{11} - 1$ and the vector β_c includes the elements of $\{\kappa(2), \mu(2), \alpha(2), \Lambda(2), \Omega(2), \Pi_{22}\}$. In practice, the supremum in LR_T^* must be taken over a finite grid for β_a, a 13-dimensional vector, and for each point in this grid one would have to calculate $\hat{\beta}_d(\beta_a)$ by concentrating the likelihood over the 18-dimensioned vector β_d. This is not a practical proposition given the complexity of the model. The computational demands of the test in the case where $S_1 = 2$, and $S_2 = 3$ are even larger.

	k	Data		One State		Two States		Three States	
		Mean	Std	Mean	Std.	Mean	Std.	Mean	Std.
real									
	12	4.987	3.002	4.470	1.557	4.674	1.693	4.500	1.761
	36	4.284	0.989	4.284	0.986	4.284	0.986	4.284	0.986
	60	4.173	0.632	4.152	0.683	4.152	0.682	4.154	0.663
	84	4.122	0.499	4.069	0.509	4.095	0.507	4.079	0.488
nominal									
	12	9.429	2.372	9.301	2.009	9.412	2.168	9.423	2.218
	36	9.474	1.781	9.474	1.781	9.474	1.781	9.474	1.781
	60	9.581	1.572	9.587	1.590	9.586	1.568	9.589	1.562
	84	9.605	1.420	9.660	1.427	9.636	1.403	9.634	1.401

Model Comparisons

Note: The table reports sample statistics for real and nominal yields calculated from the data, and the three models using the parameter estimates in Table2