Estimating Term Structure Equations Using Macroeconomic Variables

Ray C. Fair

Estimating Term Structure Equations Using Macroeconomic Variables

By

Ray C. Fair

January 2008

COWLES FOUNDATION DISCUSSION PAPER NO. 1634

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Estimating Term Structure Equations Using Macroeconomic Variables

Ray C. Fair*

December 2007

Abstract

This paper begins with the expectations theory of the term structure of interest rates with constant term premia and then postulates how expectations of future short term interest rates are formed. Expectations depend in part on predictions from a set of VAR equations and in part on the current and two lagged values of the short term interest rate. The results suggest that there is relevant independent information in both the VAR equations' predictions and the current and two lagged values of the short rate. The model fits the long term interest rate data well, including the 2004-2006 period, which some have found a puzzle. The properties of the model are consistent with the response of the long term U.S. Treasury bond rate to surprise price and employment announcements. The overall results suggest that long term rates can be fairly well explained by modeling expectation formation of future short term rates.

1 Introduction

The expectations theory of the term structure of interest rates says that long rates depend on expected future short rates. As Campbell and Shiller (1991) point out,

---

*Cowles Foundation and International Center for Finance, Yale University, New Haven, CT 06520-8281. Voice: 203-432-3715; Fax: 203-432-6167; email: ray.fair@yale.edu; website: http://fairmodel.econ.yale.edu. I am indebted to William Brainard, John Campbell, Andrew Lo, and Robert Shiller for helpful comments.

this theory is sometimes taken to include the hypothesis that expectations are rational and sometimes not. Tests generally reject the hypothesis that expectations are rational,[1] which is then a rejection of the expectations theory of the term structure if the theory is taken to include the hypothesis of rational expectations. This paper takes a somewhat different approach from the recent literature in estimating term structure equations. It assumes that the expectations theory holds with constant term premia and models how expectations are formed. Expectations are assumed to be based in part on predictions from a four-equation VAR model. Conditional on predictions from this model, four term structure equations are estimated by full information maximum likelihood (FIML). The overall model can be used to make predictions of interest rates of different maturities, and these predictions can be compared to predictions from other models.

The model is presented in Section 2; estimation is discussed in Section 3; and the estimates and prediction comparisons are presented in Section 4. Section 4 also examines how well the model predicts the period since 2000. Section 5 examines some of the properties of the model and compares these to the effects on long term interest rates from surprise price and employment announcements. The data are quarterly and four maturities are considered: one year, three years, five years, and ten years. The variables and notation used in this paper are presented in Table 1. The estimation period is 1963:2–2006:4, for a total of 175 quarterly observations. The interest rate data are for the last day of the quarter.

Recent work analyzing the term structure has begun to consider adding more

[1] In fact, King and Krumann (2002), fn. 17, p. 61, cite an unnamed monetary economist who argues that the expectations theory of the term structure has been rejected so many times that it should never be used!

macro variables to the analysis than simply short and long term interest rates,[2] and this study is in this spirit. Through the VAR equations agents use data on unemployment, inflation, and a cost shock variable to help form their expectations of future short rates, which then affect long rates. Contrary to much recent work, however, the term premia are assumed to be constant. None of the fluctuations in long rates are attributed in this study to fluctuations in term premia. The emphasis is instead on fluctuations in expectations of future short rates. Also, contrary to much recent work, no latent factors are postulated in this study.

2 The Model

The five interest rate variables in Table 1 in the model are: r_1, the three-month rate, r_4, the one-year rate, r_{12}, the three-year rate, r_{20}, the five-year rate, and r_{40}, the ten-year rate. The subscripts refer to quarters, rather than years or months, since the data are quarterly, and the interest rates are at quarterly rates. The interest rates other than r_1 will be called "long rates." The interest rate variables were chosen to maximize the length of the estimation period. The available data allow the estimation period to begin in 1963:2. Choosing more long rates would have

[2] See, for example, Kozicki and Tinsley (2001), Dewachter and Lyrio (2006), and Rudebusch, Sack, and Swanson (2007). Early work in this area, such as Sargent (1979), assumed only interest rates in the information sets of agents. Cochrane and Piazzesi (2006), who work with an affine-yield model, argue (p. 49) that a natural next step in the analysis is to incorporate other information, such as inflation, about long-term interest rate expectations. Rudebusch and Wu (2003), again working with an affine-yield model, interpret their latent term structure level factor as a medium-term central bank inflation target and their latent slope factor as cyclical variation in inflation and output gaps. In the present paper, as discussed next, the macro variables in the VAR equations are unemployment, inflation, and a cost shock variable. In a very early paper Modigliani and Shiller (1973) estimated term structure equations with a corporate bond yield on the left hand side and current and lagged values of the commercial paper rate and the inflation rate on the right hand side.

Table 1
The Data Used

Data from Federal Reserve Board, H.15		
(annual rates, last business day of the quarter)		
R_1^*	$=$	three-month Treasury bill rate, discount basis.
R_1	$=$	$[(365/360)R_1^*]/(100 - .25R_1^*)$.
R_4	$=$	one-year yield on U.S. Treasury securities.
R_{12}	$=$	three-year yield on U.S. Treasury securities.
R_{20}	$=$	five-year yield on U.S. Treasury securities.
R_{40}	$=$	ten-year yield on U.S. Treasury securities.
Macro Data		
Quarterly Averages		
UR	$=$	Unemployment rate. BLS data; variable UR in Fair (2004).
PF	$=$	Nonfarm price deflator. BEA data; variable PF in Fair (2004).
PIM	$=$	Import price deflator. BEA data; variable PIM in Fair (2004).
Variables in the Model		
r_1	$=$	$(1+R_1)^{.25}-1$
r_4	$=$	$(1+R_4)^{.25}-1$
r_{12}	$=$	$(1+R_{12})^{.25}-1$
r_{20}	$=$	$(1+R_{20})^{.25}-1$
r_{40}	$=$	$(1+R_{40})^{.25}-1$
u	$=$	UR
π	$=$	$\log(PF/PF_{-1})$
s	$=$	$\log(PIM/PIM_{-1})$

- Estimation period is 1963:2–2006:4, 175 observations.
- As noted in the table, an adjustment was made to the three-month rate, which is on a discount basis, to convert it to a yield.

shortened this period. Also, there are no gaps in the data for any of the five variables, which is not true for some of the other long rates. The other variables in the model are π, the domestic inflation rate, u, the unemployment rate, and s, the percentage change in the price of imports, a cost shock variable.

The Term Structure Equations

The interest rates r_4, r_{12}, r_{20}, and r_{40} are yields on coupon bonds, and because of this the linearized expectations model in Shiller (1979) is used for the term structure equations, which handles this problem. In the following equations γ is $1/(1+\bar{r})$, where $\bar{r}$ is taken to be .015, which is roughly the mean of r_1 in the sample period (at a quarterly rate). The four equations are:

$$r_{4t} = \frac{1-\gamma}{1-\gamma^4}(r_{1t} + \gamma r^e_{1t+1} + \gamma^2 r^e_{1t+2} + \gamma^3 r^e_{1t+3}) + \delta_1 \tag{1}$$

$$r_{12t} = \frac{1-\gamma}{1-\gamma^{12}}(r_{1t} + \gamma r^e_{1t+1} + \gamma^2 r^e_{1t+2} + \cdots + \gamma^{11} r^e_{1t+11}) + \delta_2 \tag{2}$$

$$r_{20t} = \frac{1-\gamma}{1-\gamma^{20}}(r_{1t} + \gamma r^e_{1t+1} + \gamma^2 r^e_{1t+2} + \cdots + \gamma^{19} r^e_{1t+19}) + \delta_3 \tag{3}$$

$$r_{40t} = \frac{1-\gamma}{1-\gamma^{40}}(r_{1t} + \gamma r^e_{1t+1} + \gamma^2 r^e_{1t+2} + \cdots + \gamma^{39} r^e_{1t+39}) + \delta_4 \tag{4}$$

t refers to the last day of quarter t. The expectations, denoted by a superscript e, are assumed to be made on this day. For example, r^e_{1t+1} is the expectation made on the last day of quarter t of the three-month rate that will exist on the last day of quarter $t+1$. The δ coefficients are the term premia. They are assumed to be constant across time, but possibly different across equations.

Equations (1)–(4) are standard term structure equations aside from the use of γ to adjust for the bonds being coupon bonds. If the bonds were zero-coupon bonds, then γ is 1 and the equations are the same as equation (10.2.10) in Campbell, Lo, and MacKinlay (1997), p. 417, except for the addition of the δ coefficients.

The VAR Equations

The four variables in the VAR equations are: the three-month interest rate, r_1, the inflation rate, π, the unemployment rate, u, and the cost shock variable, s. The right hand side variables in each equation include a constant term and four lagged values of each variable. The predictions are assumed to be made at the end of quarter t for quarters $t+1$ through $t+39$, where quarter $t+39$ is the last quarter for which expectations of r_1 are needed in the term structure equations. The variables r_{1t}, π_t, u_t, and s_t are assumed to be known at the end of quarter t when the predictions for quarters $t+1$ and beyond are made. The equations are:

$$\begin{aligned} r_{1t+1} &= f_6(cnst, r_{1t}, r_{1t-1}, r_{1t-2}, r_{1t-3}, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \\ &\quad u_t, u_{t-1}, u_{t-2}, u_{t-3}, s_t, s_{t-1}, s_{t-2}, s_{t-3}), +u_{6t+1} \end{aligned} \tag{5}$$

$$\begin{aligned} \pi_{t+1} &= f_7(cnst, r_{1t}, r_{1t-1}, r_{1t-2}, r_{1t-3}, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \\ &\quad u_t, u_{t-1}, u_{t-2}, u_{t-3}, s_t, s_{t-1}, s_{t-2}, s_{t-3}), +u_{7t+1} \end{aligned} \tag{6}$$

$$\begin{aligned} u_{t+1} &= f_8(cnst, r_{1t}, r_{1t-1}, r_{1t-2}, r_{1t-3}, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \\ &\quad u_t, u_{t-1}, u_{t-2}, u_{t-3}, s_t, s_{t-1}, s_{t-2}, s_{t-3}), +u_{8t+1} \end{aligned} \tag{7}$$

$$\begin{aligned} s_{t+1} &= f_9(cnst, r_{1t}, r_{1t-1}, r_{1t-2}, r_{1t-3}, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \\ &\quad u_t, u_{t-1}, u_{t-2}, u_{t-3}, s_t, s_{t-1}, s_{t-2}, s_{t-3}), +u_{9t+1} \end{aligned} \tag{8}$$

The f_i functions are linear, and $cnst$ denotes the constant term. The subscript $t+1$ has been used to emphasize the fact that the VAR equations are used to predict quarter $t+1$ (at the end of quarter t) given knowledge of the variables for quarter t.

Expectation Formation

The VAR equations may or may not approximate well how agents actually form their future expectations, and the following specification allows this to be tested. It distinguishes between the VAR equations' *predictions* of future short term interest rates and the agents' *expectations* of these rates. For a given set of coefficients and initial conditions and setting all error terms to zero, the four VAR equations can be solved at the end of quarter t for values of r_1 for quarters $t+1$ and beyond. Let r^{ee}_{1t+i} denote the prediction from these four equations for r_{1t+i} ($i = 1, \ldots, 39$). Agents' expectations that enter equations (1)–(4), which have a superscript e, are not necessarily assumed to be the same as these predictions, which have a superscript ee. Agents are instead assumed to form their expectations (at the end of quarter t) in the following way:

$$
\begin{aligned}
r^e_{1t+1} &= \alpha_1 r^{ee}_{1t+1} + \beta_{1,1} r_{1t} + \beta_{1,2} r_{1t-1} + \beta_{1,3} r_{1t-2} + \zeta_1 + v_{1t} \\
r^e_{1t+2} &= \alpha_2 r^{ee}_{1t+2} + \beta_{2,1} r_{1t} + \beta_{2,2} r_{1t-1} + \beta_{2,3} r_{1t-2} + \zeta_2 + v_{2t} \\
&\vdots \\
r^e_{1t+39} &= \alpha_{39} r^{ee}_{1t+39} + \beta_{39,1} r_{1t} + \beta_{39,2} r_{1t-1} + \beta_{39,3} r_{1t-2} + \zeta_{39} + v_{39t}
\end{aligned}
\tag{9}
$$

Each equation in (9) says that agents' expectation of a future value of r_1 is a function of the VAR equations' prediction of this value, of the actual (observed) values of r_1 for quarters t, $t-1$, and $t-2$, and of a constant term. The error term, v, reflects all the factors that affect expectations that are not captured in the right hand side variables.

If the VAR predictions contain no relevant information not in r_{1t}, r_{1t-1}, r_{1t-2}, and the constant, then the α coefficients are 0. If, on the other hand, r_{1t}, r_{1t-1}, and r_{1t-2} contain no relevant information not in the VAR predictions and the constant, then the β coefficients are 0. If the β coefficients are 0 and the α coefficients are 1, then (9) says agents' expectation for a particular quarter equals the VAR prediction aside from a possible constant and error. This specification thus allows some flexibility in modeling how expectations are formed. Agents' expectations are not forced to be exactly the VAR equations' predictions. The VAR equations may be just one input into the expectation process.

3 Estimation

In equations (1)–(4), let $\lambda_i = (1+\gamma)/(1+\gamma^i), i = 4, 12, 20, 40$. It will be useful for estimation purposes to write equations (1)–(4) as:

$$\frac{r_{4t}}{\lambda_4} - r_{1t} = \gamma r^e_{1t+1} + \gamma^2 r^e_{1t+2} + \gamma^3 r^e_{1t+3} + \frac{\delta_1}{\lambda_4} \qquad (1)'$$

$$\frac{r_{12t}}{\lambda_{12}} - \frac{r_{4t}}{\lambda_4} = \gamma^4 r^e_{1t+4} + \cdots + \gamma^{11} r^e_{1t+11} + \frac{\delta_2}{\lambda_{12}} - \frac{\delta_1}{\lambda_4} \qquad (2)'$$

$$\frac{r_{20t}}{\lambda_{20}} - \frac{r_{12t}}{\lambda_{12}} = \gamma^{12} r^e_{1t+12} + \cdots + \gamma^{19} r^e_{1t+19} + \frac{\delta_3}{\lambda_{20}} - \frac{\delta_2}{\lambda_{12}} \qquad (3)'$$

$$\frac{r_{40t}}{\lambda_{40}} - \frac{r_{20t}}{\lambda_{20}} = \gamma^{20} r^e_{1t+20} + \cdots + \gamma^{39} r^e_{1t+39} + \frac{\delta_4}{\lambda_{40}} - \frac{\delta_3}{\lambda_{12}} \qquad (4)'$$

The 39 expectation equations in (9) can then can be substituted into equations (1)′–(4)′ to yield:

$$\begin{aligned}\frac{r_{4t}}{\lambda_4} - r_{1t} &= \alpha_1 \gamma r^{ee}_{1t+1} + \alpha_2 \gamma^2 r^{ee}_{1t+2} + \alpha_3 \gamma^3 r^{ee}_{1t+3} + \theta_{1,1} r_{1t} + \theta_{1,2} r_{1,t-1} \\ &\quad + \theta_{1,3} r_{1,t-2} + \psi_1 + (\gamma v_{1t} + \gamma^2 v_{2t} + \gamma^3 v_{3t}) \qquad (1)''\end{aligned}$$

$$\begin{aligned}\frac{r_{12t}}{\lambda_{12}} - \frac{r_{4t}}{\lambda_4} &= \alpha_4 \gamma^4 r^{ee}_{1t+4} + \cdots + \alpha_{11} \gamma^{11} r^{ee}_{1t+11} + \theta_{2,1} r_{1t} + \theta_{2,2} r_{1,t-1} \\ &\quad + \theta_{2,3} r_{1,t-2} + \psi_2 + (\gamma^4 v_{4t} + \cdots + \gamma^{11} v_{11t}) \qquad (2)''\end{aligned}$$

$$\begin{aligned}\frac{r_{20t}}{\lambda_{20}} - \frac{r_{12t}}{\lambda_{12}} &= \alpha_{12} \gamma^{12} r^{ee}_{1t+12} + \cdots + \alpha_{19} \gamma^{19} r^{ee}_{1t+19} + \theta_{3,1} r_{1t} + \theta_{3,2} r_{1,t-1} \\ &\quad + \theta_{3,3} r_{1,t-2} + \psi_3 + (\gamma^{12} v_{12t} + \cdots + \gamma^{19} v_{19t}) \qquad (3)''\end{aligned}$$

$$\begin{aligned}\frac{r_{40t}}{\lambda_{40}} - \frac{r_{20t}}{\lambda_{20}} &= \alpha_{20} \gamma^{20} r^{ee}_{1t+20} + \cdots + \alpha_{39} \gamma^{39} r^{ee}_{1t+39} + \theta_{4,1} r_{1t} + \theta_{4,2} r_{1,t-1} \\ &\quad + \theta_{4,3} r_{1,t-2} + \psi_4 + (\gamma^{20} v_{20t} + \cdots + \gamma^{39} v_{39t}) \qquad (4)''\end{aligned}$$

where $\theta_{1,i} = \sum_{j=1}^{3} \beta_{j,i}\gamma^j$, $\theta_{2,i} = \sum_{j=4}^{11} \beta_{j,i}\gamma^j$, $\theta_{3,i} = \sum_{j=12}^{19} \beta_{j,i}\gamma^j$, and $\theta_{4,i} = \sum_{j=20}^{39} \beta_{j,i}\gamma^j$, $i = 1, 2, 3$. Also, $\psi_1 = \delta_1/\lambda_4 + \sum_{j=1}^{3} \zeta_j\gamma^j$, $\psi_2 = \delta_2/\lambda_{12} - \delta_1/\lambda_4 + \sum_{j=4}^{11} \zeta_j\gamma^j$, $\psi_3 = \delta_3/\lambda_{20} - \delta_2/\lambda_{12} + \sum_{j=12}^{19} \zeta_j\gamma^j$, and $\psi_4 = \delta_4/\lambda_{40} - \delta_3/\lambda_{20} + \sum_{j=20}^{39} \zeta_j\gamma^j$.

Timing

Before considering estimation further, it is important to be clear on the timing that is assumed in the model. At the end of quarter t agents solve the VAR equations for quarters $t+1$ through $t+39$, given a set of coefficients and assuming zero errors. They are assumed to know the values of r_1, π, u, and s for quarters t and back. This solution yields predictions of 39 future values of r_1—the values with superscript ee. Given these values and given the actual values of r_{1t}, r_{1t-1}, and r_{1t-2}, equations (1)$''$–(4)$''$ are solved for r_{4t}, r_{12t}, r_{20t}, and r_{40t}. This assumes that r_{1t} is known before the other four rates are determined (say, a few minutes before). Note that r_{1t} is not predicted from the VAR equations, where the first quarter predicted is $t+1$. It is simply assumed to be known at the end of quarter t for purposes of determining r_{4t}, r_{12t}, r_{20t}, and r_{40t} in equations (1)$''$–(4)$''$. So the timing is: agents predict quarters $t+1$ and beyond knowing r_{1t}, π_t, u_t, and s_t, and then given these predictions and the actual value of r_{1t}, the long rates are determined

Restrictions

Two sets of restrictions were imposed on equations (1)″–(4)″ before estimation. These restrictions would not be needed if there were 39 interest rates (two-quarter, three-quarter, ..., 40-quarter) instead of only four. The first set concerns the v error terms. These errors pick up the effects on expectations that are not captured by the variables in equations (9). It may be that these errors are serially correlated, and to test for this the error terms are assumed to be first-order serially correlated with the restrictions that 1) v_1, v_2 and v_3 have the same serial correlation coefficient, 2) $v_4, \ldots, v_{11}$ have the same serial correlation coefficient, 3) $v_{12}, \ldots, v_{19}$ have the same serial correlation coefficient, and 4) $v_{20}, \ldots, v_{39}$ have the same serial correlation coefficient, denoted ρ_1, ρ_2, ρ_3, and ρ_4 respectively. . Let $\mu_1 = \gamma v_1 + \gamma^2 v_2 + \gamma^3 v_3$, $\mu_2 = \gamma^4 v_4 + \cdots + \gamma^{11} v_{11}$, $\mu_3 = \gamma^{12} v_{12} + \cdots + \gamma^{19} v_{19}$, and $\mu_4 = \gamma^{20} v_{20} + \cdots + \gamma^{39} v_{39}$. The serial correlation assumptions are then:

$$\mu_{1t} = \rho_1 \mu_{1t-1} + \epsilon_{1t} \tag{10}$$

$$\mu_{2t} = \rho_2 \mu_{1t-1} + \epsilon_{2t} \tag{11}$$

$$\mu_{3t} = \rho_3 \mu_{1t-1} + \epsilon_{3t} \tag{12}$$

$$\mu_{4t} = \rho_4 \mu_{1t-1} + \epsilon_{4t} \tag{13}$$

where the ϵ error terms are assumed to be iid.

The second set of restrictions concerns the α coefficients. There are 39 of them, which is too many to estimate individually given that there are only four equations. Instead, four coefficients were estimated unrestricted: α_1, α_4, α_{12}, and α_{20}. The restrictions imposed are that 1) α_2 and α_3 equal α_1, 2) α_5 through α_{11} equal α_4, 3) α_{13} through α_{19} equal α_{12}, and 4) α_{21} through α_{39} equal α_{20}. This means that two restrictions are imposed on equation $(1)''$, seven each are imposed on equations $(2)''$ and $(3)''$, and 19 are imposed on equation $(4)''$.

Note that the δ, ζ, and β coefficients are not identified. The lack of identification of the β coefficients is again a consequence of having data for only four long rates rather than 39. A key question for purposes of this paper is whether the estimates of the α coefficients are significant. In other words, is there relevant information in the VAR predictions that is not in r_{1t}, r_{1t-1}, and r_{1t-2}?

Using these two sets of restrictions, equations $(1)''$–$(4)''$ can be written:

$$\begin{aligned}
\frac{r_{4t}}{\lambda_4} - r_{1t} &= \alpha_1(\gamma r^{ee}_{1t+1} + \gamma^2 r^{ee}_{1t+2} + \gamma^3 r^{ee}_{1t+3}) + \theta_{1,1} r_{1t} + \theta_{1,2} r_{1,t-1} \\
&\quad + \theta_{1,3} r_{1,t-2} + \psi_1 + \mu_{1t} && (1)''' \\
\frac{r_{12t}}{\lambda_{12}} - \frac{r_{4t}}{\lambda_4} &= \alpha_4(\gamma^4 r^{ee}_{1t+4} + \cdots + \gamma^{11} r^{ee}_{1t+11}) + \theta_{2,1} r_{1t} + \theta_{2,2} r_{1,t-1} \\
&\quad + \theta_{2,3} r_{1,t-2} + \psi_2 + \mu_{2t} && (2)''' \\
\frac{r_{20t}}{\lambda_{20}} - \frac{r_{12t}}{\lambda_{12}} &= \alpha_{12}(\gamma^{12} r^{ee}_{1t+12} + \cdots + \gamma^{19} r^{ee}_{1t+19}) + \theta_{3,1} r_{1t} + \theta_{3,2} r_{1,t-1} \\
&\quad + \theta_{3,3} r_{1,t-2} + \psi_3 + \mu_{3t} && (3)''' \\
\frac{r_{40t}}{\lambda_{40}} - \frac{r_{20t}}{\lambda_{20}} &= \alpha_{20}(\gamma^{20} r^{ee}_{1t+20} + \cdots + \gamma^{39} r^{ee}_{1t+39}) + \theta_{4,1} r_{1t} + \theta_{4,2} r_{1,t-1} \\
&\quad + \theta_{4,3} r_{1,t-2} + \psi_4 + \mu_{4t} && (4)'''
\end{aligned}$$

The term structure model consists of equations $(1)'''$– $(4)'''$, where the error terms are assumed to be first order serially correlated as in (10)–(13). There are 24 unrestricted coefficients to estimate, counting the four serial correlation coefficients in (10)–(13). Again, γ is constant; it is 1/1.015.

Computational Issues

The four VAR equations were estimated by OLS for the 1963:2–2006:4 period, 175 quarters. Then for each quarter between 1963:3 and 2007:1 a dynamic simulation was run for 39 quarters ahead. For example, for the period beginning in 1963:3 a simulation was run for 1963:3 thorough 1973:1 using only information available from 1963:2 back. This yields the 39 predictions relevant for quarter 1963:2. This process is then repeated 174 more times. The last simulation, which begins in 2007:1, uses data from 2006:4 back and predicts through 2016:3.

Under the assumption that the errors terms $\epsilon_{1t}, \epsilon_{2t}, \epsilon_{3t}$, and ϵ_{4t} are jointly normally distributed with zero means and some covariance matrix Σ, the 24 coefficients can be estimated by full information maximum likelihood (FIML). The FIML estimation of nonlinear models with rational expectations is discussed in Fair and Taylor (FT) (1990), and the present estimation problem is a special case of the general problem considered in FT. It is special in that once the predictions are computed from the VAR equations, they can be used in the term structure equations with no feedback to the VAR equations. So given the predictions from the VAR equations, this is a standard FIML estimation problem.

Once a procedure is available for computing the value of the likelihood function for a given set of coefficients, the estimation problem can be turned over to a nonlinear maximization algorithm. These algorithms search over sets of coefficients to find the set that maximizes the objective function. For the FIML estimation of large models the algorithm that I have found best to use is the Parke (1982) algorithm, and it has been used for the work below.

4 The Results

Coefficient Estimates

The 24 coefficient estimates for the term structure equations are presented in Table 2.[3] The variance-covariance matrix of the coefficient estimates is the inverse of the matrix of the second derivatives of the log of the likelihood function. The 24×24 second derivative matrix was computed numerically after the maximum of the likelihood function had been reached.

The α estimates are individually significant, and the hypothesis that the α's are all zero is strongly rejected, with a p-value of .00004. The estimates are not, however, close to 1, ranging from .149 to .340. r_{1t} is highly significant, and r_{1t-1} and r_{1t-2} are significant or close to significant except in equation (1)$'''$.[4] There is thus relevant independent information in both the VAR predictions and the current and two lagged values of r_1 regarding the expected future values of r_1.

The $\theta_{1,i}$ coefficients are the weighted sum of three β's; the $\theta_{2,i}$ and $\theta_{3,i}$ coefficients are the weighted sum of eight β's; and the $\theta_{4,i}$ coefficients are the weighted sum of 20 β's, $i = 1, 2, 3$. Although the weights are declining, one would probably expect the $\theta_{1,i}$ estimates to be the smallest for a given i, which is the case in Table 2. One would also expect the $\theta_{3,i}$ estimates to be smaller than the $\theta_{2,i}$ estimates for a given i because of the declining weights, which is also the case in Table 2 except for $i = 2$. Finally, one would expect the $\theta_{4,i}$ estimates to be the largest because of

[3]To save space, the 68 coefficient estimates for the VAR equations are not presented. Remember that the VAR equations are estimated first by OLS (for the same 1963:2–2006:4 period), and then the predictions from these equations are used for the estimates in Table 2.

[4]The hypothesis that the four coefficients for r_{1t-2} are all zero was rejected, with a p-value of .0068.

Table 2
Coefficient Estimates for Equations (1)$'''$– (4)$'''$

	(1)$'''$	(2)$'''$	(3)$'''$	(4)$'''$
r^{ee}_{1t+j}'s—$\alpha_1, \alpha_4, \alpha_{12}, \alpha_{20}$:	.340	.149	.231	.237
	(4.59)	(2.60)	(3.42)	(2.09)
r_{1t}—$\theta_{1,1}, \theta_{2,1}, \theta_{3,1}, \theta_{4,1}$:	1.949	3.670	2.512	3.561
	(11.26)	(10.78)	(9.45)	(7.11)
r_{1t-1}—$\theta_{1,2}, \theta_{2,2}, \theta_{3,2}, \theta_{4,2}$:	.021	.564	.688	1.577
	(0.19)	(1.93)	(2.66)	(3.19)
r_{1t-2}—$\theta_{1,3}, \theta_{2,3}, \theta_{3,3}, \theta_{4,3}$:	.030	.775	.465	.836
	(0.29)	(2.76)	(1.85)	(1.72)
cnst—$\psi_1, \psi_2, \psi_3, \psi_4$:	.0031	.0338	.0385	.1031
	(2.24)	(4.82)	(4.80)	(4.00)
$\rho_1, \rho_2, \rho_3, \rho_4$:	.507	.736	.781	.850
	(9.15)	(21.07)	(25.52)	(31.47)

- χ^2 test of hypothesis that all α's are zero: value = 25.41, 4 degrees of freedom, p-value = .00004.
- FIML estimation.
- Estimation period is 1963:2–2006:4.
- t-statistics are in parentheses.

the larger sum, and this is the case in Table 2 except for $\theta_{4,1}$ versus $\theta_{2,1}$. Similarly, one would expect the same pattern for the estimates of the constant term. This is the case except that the estimate of ψ_3 is slightly larger than the estimate of ψ_2.

The estimates of the serial correlation coefficients are large and highly significant. This means that the error terms in (9) are serially correlated. The high degree of serial correlation in term structure equations is a persistent problem. It has been noticed in papers ranging in time from Modigliani and Shiller (1973) to Dewachter and Lyrio (2006). If one interprets the serially correlated errors as reflecting serially correlated omitted variables, the results suggest that there are important omitted variables in the expectation equations (9). This is discussed further in the Conclusion.

Root Mean Squared Errors

Once the model is estimated, it can be used to make predictions of the four long rates, and these can be compared to predictions from other models. For present purposes, two other models have been used. The first is the random walk (RW) model, where each rate equals last quarter's rate:

$$r_{it} = r_{it-1}, \quad i = 4, 12, 20, 40 \tag{14}$$

This model does not use information on r_{1t}, which the model in this paper does, and an alternative model that incorporates this information is one in which the change in each rate equals the change in r_{1t}:

$$r_{it} = r_{it-1} + r_{1t} - r_{1t-1}, \quad i = 4, 12, 20, 40 \tag{15}$$

This model will be called "random walk plus" (RW+).

Root mean squared error (RMSEs) are presented in Table 3. The prediction period is the same as the estimation period, 1963:2–2006:4, 175 observations. The errors are in percentage points at annual rates. The RMSEs for the present model range from .367 percentage points for r_4 to .490 percentage points for r_{20}. They are all smaller than the corresponding errors for RW and RW+, and so the model does noticeably better than RW and RW+. Note that the RW and RW+ models use information on the lagged long rate, which the present model does not use directly.

Table 3
Root Mean Squared Errors

	Present Model	RW	RW+
r_{4t}	.367	1.019	.461
r_{12t}	.473	.842	.629
r_{20t}	.490	.768	.692
r_{40t}	.470	.639	.765

- Prediction period is 1963:2–2006:4.
- Errors are in percentage points at annual rates.

However, the present model uses information on the lagged errors through the serial correlation coefficients, which in effect incorporates information on the lagged long rate.

Backus and Wright (2007), among others, have puzzled over the fact that as the Fed raised the short term interest rate from 2004 to 2006, long term rates did not rise. They argued that the most likely explanation is a fall in the term premia. In the present model the term premia are assumed to be constant, and so it is interesting to see how the model predicts the 2004-2006 period. Is this period really a puzzle?

To examine this question both the VAR equations and the term structure equations were reestimated for the estimation period ending in 1999:4. Then predictions of the long rates were made for the 2000:1–2006:4 period, which are outside sample predictions. Results for the ten-year rate, r_{40}, are presented in Table 4. Also presented in the table is the actual value of r_1. This was a period in which r_1 fell rapidly through 2004:1 and then rose rapidly after that. How well did the model predict r_{40} in this period? Remember that the predictions are bases in part on the 39-quarter-ahead predictions from the VAR equations and in part on the current

Table 4
Outside Sample Predictions for 2000:1–2006:4

Quarter	r_1	$\hat{r_{40}}$	r_{40}	$\hat{r_{40}} - r_{40}$
2000.1	5.76	6.67	5.90	0.77
2000.2	5.75	6.17	5.90	0.27
2000.3	6.09	6.31	5.68	0.63
2000.4	5.77	5.90	5.02	0.87
2001.1	4.24	4.75	4.84	-0.09
2001.2	3.60	4.69	5.31	-0.62
2001.3	2.38	4.64	4.52	0.12
2001.4	1.73	4.10	4.98	-0.88
2002.1	1.78	4.80	5.31	-0.51
2002.2	1.69	5.09	4.77	0.31
2002.3	1.56	4.65	3.58	1.06
2002.4	1.21	3.59	3.78	-0.18
2003.1	1.13	3.87	3.78	0.09
2003.2	0.90	3.64	3.49	0.15
2003.3	0.94	3.66	3.90	-0.24
2003.4	0.94	3.94	4.20	-0.26
2004.1	0.94	4.22	3.81	0.41
2004.2	1.33	4.06	4.54	-0.49
2004.3	1.70	4.68	4.08	0.60
2004.4	2.20	4.46	4.17	0.28
2005.1	2.76	4.62	4.43	0.19
2005.2	3.09	4.82	3.88	0.94
2005.3	3.50	4.46	4.27	0.19
2005.4	4.03	4.83	4.32	0.51
2006.1	4.56	4.97	4.77	0.20
2006.2	4.91	5.36	5.05	0.31
2006.3	4.81	5.34	4.56	0.77
2006.4	4.93	4.99	4.63	0.36

- Estimation period was 1963:2–1999:4.
- Values are in percentage points at annual rates.
- RMSE = .541 for 2000:1–2003:4.
- RMSE = .494 for 2004:1–2006:4.
- Mean error = .094 for 2000:1–2003:4.
- Mean error = .357 for 2004:1–2006:4.

and two lagged values of r_1. The RMSE is .541 for the sub period 2000:1–2003:4 and .494 for the sub period 2004:1–2006:4. These compare to the RMSE in Table 3 for r_{40} of .470, and so the RMSEs are only slightly larger. The mean error is .094 for the first sub period and .357 for the second. The model thus overpredicts the ten-year rate by an average of .357 percentage points in the 2004–2006 period. This is consistent with Backus and Wright's puzzle that the long term rates were lower-than-expected in this period, but the size of the mean error is not large. There is at most only a modest puzzle here. In other words, a model that generates expectations of future short term rates as in (9) can account for much of the behavior of long term rates in the 2004–2006 period.

5 Properties of the Model

Two Experiments

To examine the properties of the model, two experiments were performed using the coefficient estimates in Table 2 (and the corresponding VAR estimates). First, the model was solved for a particular quarter t = 1991:4 with all relevant error terms set to zero.[5] Call this the "base" solution. Then for the first experiment the error term in equation (6)—the VAR equation for π—was taken to be .01 in $t + 1$ = 1992:1 and zero otherwise. All other error terms were still set to zero. The model was then solved. Call this the "π shock" solution. This shock is a one percentage point inflation shock. For the second experiment the error term in equation (7)—the VAR equation for the unemployment rate—was taken to be $-$.01 in $t + 1$ = 1992:1 and zero otherwise. All other error terms were still set to zero. The model was then solved. Call this the "u shock" solution. This shock is a one percentage point unemployment rate shock.

[5] Because the model is linear in variables, the results do not depend on the particular quarter, 1991:4, used. Any quarter will give the same results.

The results for the four long rates are presented in Table 5. Each value in the table is the predicted value from the shocked solution minus the predicted value from the base solution times 400 (to put the values in percentage points at an annual rate).

The π shock led to an increase in r_{4t} of 4.8 basis points. The other changes are 2.8, 2.5, and 1.8 basis points, respectively. The changes for the u shock are 14.0, 7.3, 4.5, and 2.8 basis points respectively. Both shocks thus led to an increase in the long rates for the current period. This is because the shocks for quarter $t + 1$ changed the VAR predictions of the values of r_1 for quarters $t + 2$ and beyond, which then affected the long rates through the term structure equations. The shocks led agents to expect higher short term rates in the future, which led to an increase in the current long term rates.

Comparison to Surprise Announcement Effects

The properties just described are consistent with the responses of long term interest rates to surprise announcements about prices and employment. In Fair (2003) I searched, using tick data on stock and bond prices and exchange rates, for announcements and events that led to large changes in prices within five minutes. The period examined was 1982–2000, and news wires were used for the searches. 221 announcements and events were found that led to large five minute changes in at least one of the five variables examined. The five variables were the S&P 500 stock price index, the 30-year U.S. Treasury bond price, and three exchange rates. The three exchange rates were the U.S. dollar relative to the Deutsche mark or

Table 5
Effects of a Price Shock and an Unemployment Shock

	p **shock**	u **shock**
r_4	.048	.140
r_{12}	.028	.073
r_{20}	.025	.045
r_{40}	.018	.028

- Values in percentage points at annual rates. .048 is 4.8 basis points.
- Price shock was 1.0 percentage points.
- Unemployment shock was -1.0 percentage points.

euro, the Japanese yen, and the British pound.

Table 3 in Fair (2003) lists all 221 announcements and events and their five minute effects. There are 11 CPI or PPI announcements in which inflation was higher-than-expected and 15 announcements in which inflation was lower than expected.[6] For all 11 higher-than-expected announcements the bond price fell (the 30 year interest rate rose), and for all 15 lower-than-expected announcements the bond price rose. A positive (negative) inflation shock thus leads to an increase (a decrease) in long term rates, which is consistent with the properties of the model. When there is, say, a positive inflation shock, agents expect short term rates to be larger in the future, which immediately increases long term rates.

In Table 3 in Fair (2003) there are 28 employment announcements in which employment was stronger than expected and 25 announcements in which employment

[6]The 11 higher-than-expected announcements are 56, 57, 61, 69, 83, 92, 124, 164, 191, 210, and 215. The 15 lower-than-expected announcements are 59, 64, 72, 93, 108, 112, 115, 120, 125, 142, 148, 155, 161, 196, and 205. (Fair (2003), Table 3, pp. 324–325.)

was weaker than expected (announcements when the bond market was open).[7] The 30-year interest rate rose after the stronger than expected employment announcement in 22 of the 28 cases, and it fell after the weaker than expected announcement in 19 of the 25 cases. It is thus generally the case that stronger (weaker) than expected employment announcements lead to an increase (a decrease) in the 30-year rate. Again, these results are consistent with the properties of the model, where the negative shock to the VAR unemployment equation led to an increase in the long term rates. When the announcement is stronger than expected (negative shock to the unemployment equation) agents' expectations of future short term rates increase, which leads to an immediate change in long term rates, and vice versa when the announcement is weaker than expected (positive shock to the unemployment equation).

6 Conclusion

This paper begins with the expectations theory of the term structure of interest rates with constant term premia, equations (1)–(4), and then postulates how expectations of future short term interest rates are formed. The results in Table 2 suggest that

[7] The 28 stronger than expected announcements are 53, 55, 94, 105, 107, 118, 119, 121, 123, 129, 134, 136, 138, 139, 143, 144, 146, 147, 150, 154, 158, 163, 165, 187, 189, 199, 212, and 214. The 25 weaker than expected announcements are 58, 63, 68, 71, 73, 81, 82, 86, 88, 91, 97, 101, 103, 109, 111, 114, 116, 117, 128, 133, 141, 149, 194, 202, and 208. (Fair (2003), Table 3, pp. 325–328.) In a few cases it is not obvious whether an announcement is a positive or negative surprise, and so a few of these classifications may be incorrect. Also, the average hourly wage rate is announced at the same time as employment, and it may be in a few cases that the surprise was regarding the wage rate and not employment. The possible misclassifications and the wage rate announcements may explain some of the 12 cases discussed next where the 30-year rate did not change as expected.

there is relevant independent information in both the VAR equations' predictions of the future short term rates and the current and two lagged values of the short term rate. The results in Table 3 show that the model fits the data better than the random walk model and the model in which the change in the long rate is equal to the change in the short rate. The results in Table 4 show that the model fits the 2000-2006 period fairly well—there is not much of a puzzle regarding the long term rates in the 2004–2006 period. The properties of the model reported in Table 5 are consistent with the response of the 30-year U.S. Treasury bond rate to surprise price and employment announcements.

It future work it may be interesting to experiment with models other than the VAR model used here. No searching was done in this study over alternative models. The VAR equations were specified at the beginning of this study and never changed. The model that one is after is the model that best approximates what agents actually use to generate their expectations, not necessarily the model that best approximates the actual economy. The model need not be a VAR model. It will be interesting to see if models can be found that lead to smaller estimated serial correlation coefficients in Table 2. One could also experiment with adding other variables directly to the expectation equations (9). Finally, one could expand the number of long rates considered, possibly at a cost of shortening the estimation period, which would allow more α coefficients to be directly estimated. The overall results in this paper suggest that long term rates can be fairly well explained by modeling expectation formation of future short term rates.

References

[1] Backus, David K., and Jonathan H. Wright, 2007, "Cracking the Conundrum," Working Paper 2007–46, Finance and Economics Discussion Series, Federal Reserve Board, Washington D.C.

[2] Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

[3] Campbell, John Y., and Robert J. Shiller, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

[4] Cochrane, John, and Monika Piazzesi, 2006, "Decomposing the Yield Curve," unpublished manuscript, University of Chicago.

[5] Dewachter, Hans, and Marco Lyrio, 2006, "Learning, Macroeconomic Dynamics and the Term Structure of Interest Rates," June 16.

[6] Fair, Ray C., 2003, "Shock Effects on Stocks, Bonds, and Exchange Rates," *Journal of International Money and Finance*, 22, 307–341.

[7] Fair, Ray C., 2004, *Estimating How the Macroeconomy Works*. Cambridge, MA: Harvard University Press.

[8] Fair, Ray C., and John B. Taylor, 1990, "Full Information Estimation and Stochastic Simulation of Models with Rational Expectations," *Journal of Applied Econometrics*, 5, 381–392.

[9] King, Robert G., and Andre Kurmann, 2002, "Expectations and the Term Structure of Interest Rates: Evidence and Implications," *Federal Reserve Bank of Richmond Economic Quarterly*, 88, 49–95.

[10] Kozicki, Sharon, and P.A. Tinsley, 2001, "Shifting Endpoints in the Term Structure of Interest Rates," *Journal of Monetary Economics*, 47, 613–652.

[11] Modigliani, Franco, and Robert J. Shiller, 1973, "Inflation, Rational Expectations and the Term Structure of Interest Rates," *Economica*, 40, 12–43.

[12] Parke, William R., 1982, "An Algorithm for FIML and 3SLS Estimation of Large Nonlinear Models," *Econometrica*, 50, 81–96.

[13] Rudebusch, Glenn D., and Tao Wu, 2003, "A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy," Working Paper 2003-17, Federal Reserve Bank of San Francisco.

[14] Rudebusch, Glenn D., Brian P. Sack, and Eric T. Swanson, 2007, "Macroeconomic Implications of Changes in the Term Premium," *Federal Reserve Bank of St. Louis Review*, 89, 241–269.

[15] Sargent, Thomas J., 1979, "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure," *Journal of Monetary Economics*, 5, 133–143.