IMF WORKING PAPER

© 1991 International Monetary Fund

WP/91/134   INTERNATIONAL MONETARY FUND

Research Department

Modelling the Yield Curve

Prepared by Mark P. Taylor 1/

Authorized for Distribution by Paul R. Masson

December 1991

Abstract

We test and estimate a variety of alternative models of the yield
curve, using weekly, high-quality U.K. data. We extend the Campbell-
Shiller technique to the overlapping data case and apply it to reject
the pure expectations hypothesis under rational expectations. We also
find that risk measures, in the form of conditional interest rate
volatility, are unable to explain the term premium. A simple, market
segmentation approach is, however, moderately successful in explaining
the term premium.

Table of Contents   Page

List of Tables

List of Figures

Summary

This paper employs a high-quality, weekly data base on U.K. interest rates on short- and long-term government instruments, to test and estimate a variety of alternative models of the yield curve. It applies a test of a hybrid model that is not contingent on the method used by agents to form expectations, as well as a more stringent test of the pure expectations model assuming rational expectations. It then estimates a GARCH-in-mean model of the term premium and also an empirical market segmentation model.

Among other innovations, the present study applies the vector auto-regressive methodology, introduced by Campbell and Shiller for testing present value models, to the case where the frequency of data observation is finer than the maturity of the short-term instrument. In addition, this study is the first to use the interest rate data generated by the "new" Bank of England method of measuring the yield curve, which is thought to be superior to previously published data on U.K. Treasury Bond yields.

The study finds that a hybrid model of the term structure, consistent with a range of alternative hypotheses concerning the term premium, cannot be rejected. The pure expectations model of the term structure under the assumption of rational expectations is, however, easily rejected for all of the maturities considered. The rejection of the rational expectations assumption, however, may be due not to a failure of rational expectations but to a failure of the assumption of a zero term premium.

Because term premia may reflect risk aversion, the paper attempts to model risk premia in the term structure. Although there is some sign of conditional heteroscedasticity in the holding period return series, a GARCH-in-mean model is unsuccessful in explaining the term premium. More encouraging results are achieved when a simple empirical formulation of the market segmentation approach is estimated. The relative amount of U.K. government fixed interest debt outstanding at the relevant maturity is a significant determinant of the excess holding return. The preferred model, therefore, is a hybrid expectations-market segmentation model in which both expectations of future short rates and supply and demand conditions at the relevant maturity determine the long-short rate spread.

I. Introduction

The behavior of the term structure of interest rates is crucial to a proper understanding of the working of an advanced macroeconomy. Financial economists and financial market practitioners are interested in the term structure because of its dual interpretation--the pricing of bonds at different maturities--which may have implications for arbitrage opportunities and concepts of market efficiency. Macroeconomists and policy makers, on the other hand, must examine the effects of alternative macroeconomic policies on the term structure: for example, whilst it is acknowledged that the authorities may find it easier to control short-term rates, it is longer-term rates which may have more of an impact on real investment flows (see e.g., Clarida and Friedman (1983)). Similarly, Taylor (1987) suggests that the stability of the demand for money may depend crucially on term structure effects whilst Boughton (1988) makes a similar claim with respect to exchange rate equations.

As we shall discuss presently, the most popular approach to the term structure--the expectations model--has the strong implication (in its purest form) that the authorities can in general affect the level but not the shape of the yield curve (unless expectations themselves are affected). If correct, this theory therefore has clear implications for the transmission mechanism of monetary policy. Moreover, under this hypothesis, the yield on a bond of any maturity becomes a sufficient statistic to describe the whole term structure, for any given pattern of expectations.

In this paper we test and estimate, using a high-quality, weekly data base on U.K. interest rates on short and long-term government instruments, a variety of alternative models of the yield curve. In particular, we apply a test of a hybrid model of the model of the term structure which is not contingent on the method used by agents to form expectations, as well as a more stringent test of the pure expectations model assuming rational expectations. We then estimate a GARCH-in-mean model of the term premium and also an empirical model which is held to be representative of the market segmentation or preferred habitat model of the yield curve.

Among the innovations in the present study, we develop the vector auto-regressive methodology introduced by Campbell and Shiller (1987) for testing present value models, to the case where the frequency of data observation is finer than the maturity of the short-term instrument. In addition, this study is the first to use the interest rate data generated by the 'new' Bank of England method of measuring the yield curve, which is thought to be superior to data on Treasury Bond yields previously published by the Bank (Bank of England (1990)).

The remainder of the paper is set out as follows. In the next section we develop a model of the term structure which is a hybrid in the sense that it is consistent with a number of hypotheses concerning term premia. A simple test of the hybrid model, which makes only very weak assumptions concerning expectations formation is then discussed in section three and,

after the data is discussed in section four, is applied in section five. In section six we develop and apply tests of the expectations model assuming rational expectations and allowing for at most constant term premia. In section seven we develop and estimate a risk premium model of the term premium, using the conditional volatility in interest rates as a risk measure. In section eight we estimate a very simple model of the term premium which is in the spirit of the market segmentation model. A final section concludes.

II. A Hybrid Model of the Term Structure

In this section we develop a model of the term structure which can be viewed as a hybrid of the expectations model of the term structure and either the risk premium or the market segmentation approach.

A number of writers are credited with originating the expectations theory (ET); the essentials of the view may be found in Lutz (1940), Hicks (1946), Fisher (1930), and Keynes (1930). 1/ Basically, the ET postulates that the slope of the yield curve may be explained by agents' expectations of future short-term interest rates: a positive yield gap (that is long-term rates above short-term rates) may be explained by the expectation that short-term interest rates are going to rise and, conversely, a negative yield gap may be explained by the expectations that short-term interest rates are going to fall.

One way of developing the ET formally is, following Shiller (1979) to take a Taylor series approximation of the holding period return and to set this equal to the one-period rate plus a term premium.

The redemption yield on an n-period bond with a current value of $p_t^{(n)}$, a redemption value of unity and paying a coupon of c per period is defined implicitly be the relation:

$$(1) \quad p_t^{(n)} = \sum_{i=1}^{n} \frac{c}{(1+R_t^{(n)})^i} + \frac{1}{(1+R_t^{(n)})^n}$$

The holding period return, $H_t^{(n)}$--i.e., the return to holding an n-period bond for one period--is just the capital gain plus the coupon payment, as a percentage of the market price of the n-period bond at time t:

1/ Shiller and McCulloch (1987) trace the ET as far back as Fisher (1896).

$$(2) \quad H_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)} - P_t^{(n)} + c}{P_t^{(n)}}$$

Using equation (1), this can be written as

$$(3) \quad H_{t+1}^{(n)} = \left[c + \frac{c}{R_{t+1}^{(n-1)}} + \frac{R_{t+1}^{(n-1)} - c}{R_{t+1}^{(n-1)}\left[1+R_{t+1}^{(n-1)}\right]^{n-1}} \right]\left(\frac{c}{R_t^{(n)}} + \frac{R_t^{(n)} - c}{R_t^{(n)}\left[1+R_t^{(n)}\right]^{n}} \right)^{-1} -1$$

Equation (3) can be approximated by taking a Taylor series expansion around the par yield, i.e., $R_t^{(n)} = R_{t+1}^{(n-1)} = c = \bar{R}$, say, and truncating after the linear term. This yields:

$$(4) \quad H_{t+1}^{(n)} = \frac{R_t^{(n)} - \gamma_n R_{t+1}^{(n-1)}}{1-\gamma_n}$$

where

$$\gamma_n = \gamma(1-\gamma^{n-1})(1-\gamma^n)^{-1}$$

$$\gamma = (1+\bar{R})^{-1}$$

Equation (4) follows from the definition of the redemption yield implicitly given in (1) and the definition of the holding period return given in (2). We now make a behavioral assumption, namely that the expected one-period holding return, $_tH_{t+1}^{(n)e}$, must be equal to the known one-period yield, r_t, plus a term premium $\phi^{(n)}$:

$$(5) \quad _tH_{t+1}^{(n)e} = r_t + \phi^{(n)}$$

The term premium can be thought of as a 'market completion premium' (Kane (1980)), since it expresses the value of whatever services serve to 'complete' markets for bonds of different maturities. According to the pure ET, expectations of future short rates completely explain long rates, with market completion services unnecessary; thus, the term premium in (5) will be identically equal to zero. Allowing the term premium to differ from zero implies the presence of valuable market completion services such as

risk-bearing or habitat-displacement. Since, in this and the next section we make no strong assumptions concerning term premia (except that they are stationary series), the model we are about to develop can be thought of as a 'hybrid'. The later sections of the paper are concerned with placing restrictions on the term premia or attempting to model them. 1/

Equating (4) and (5) yields:

$$(6) \quad R_t^{(n)} = \gamma_n \, {}_t R_{t+1}^{(n-1)e} + (1-\gamma_n)(r_t + \phi^{(n)})$$

where ${}_t R_{t+1}^{(n-1)}$ denotes the current expectation of the redemption yield on an (n-1)-period bond next period. Assuming the redemption value of the bond at time t+n is 1 (so that $R_{t+n-1}^{(1)} = r_{t+n-1}$), equation (6) can be solved recursively to yield:

$$(7a) \quad R_t^{(n)} = \frac{1-\gamma}{1-\gamma^n} \sum_{i=0}^{n-1} \gamma^i \, {}_t r_{t+1}^e + \Phi^{(n)}$$

$$(7b) \quad \Phi^{(n)} = \frac{1-\gamma}{1-\gamma^n} \sum_{i=0}^{n-1} \gamma^i \phi^{(n-i)}$$

where ${}_t r_{t+i}^e$ denotes the current expectation of the one-period rate i periods ahead. Equations (7a) and (7b) say that the yield to maturity on an n-period bond will be a forward Koyck lag of expected future one-period rates, with expected rates given less and less weight further into the future, plus a term premium which will itself be a function of the one period holding return term premia on bonds of maturity of one to n periods.

Equations (7) are derived under the assumption that the period of observation is the same as the period to maturity of the short instrument. In the empirical work reported in this paper, we use high-quality data which is sampled weekly while the short-term rate of interest is the discount on three-month U.K. Treasury Bills. However, by regrouping terms in (7), they can easily be shown to hold, mutatis mutandis, for the overlapping data

---

1/ Although we shall not use time subscripts on $\phi^{(n)}$, nothing in this section requires the term premia to be constant--we require only the weaker assumption of stationarity.

case. Let y be the number of years to maturity of the long bond and r denote the discount on a 3-month (i.e., 13-week) T-Bill, then if the interval of observation is one week, equation (7a) can be written:

$$(8) \quad R_t^{(52y)} = \frac{1-\gamma}{1-\gamma^{52y}} \sum_{i=0}^{4y-1} \gamma^{13i} \, {}_tr^e_{t+13i} + \Phi^{(52y)}$$

where $\Phi^{(52y)}$ is a function of the thirteen-week term premia, analogous to (7b).

For large y, $\gamma^{52y} \approx 0$, so that (8) may be closely approximated by

$$(9) \quad R_t^{(52y)} - r_t = \sum_{i=1}^{4y-1} \gamma^{13i} \, {}_t\Delta_{13} \, {}_tr^e_{t+13i} + \Phi^{(52y)}$$

where Δ_{13} denotes the thirteenth-difference operator.

III. A Test of the Hybrid Model

Whatever assumptions are made concerning the term premium, it seems highly unlikely that it will be a nonstationary processes. Subject to the assumption that the term premia are at most stationary processes, and a similar assumption regarding agents' forecasting errors, we outline in this section a test of the hybrid model of the term structure which has the interesting feature of being nonspecific with regard to the method used by agents to form expectations: as noted above, the only assumption regarding expectations is that the forecasting error made by agents is a stationary, $I(0)$ process.

Equation (9) demonstrates that, if both long and short rates are $I(1)$ series, then the long-short spread must be $I(0)$--subject only to the qualification that forecasting errors are $I(0)$ and that the term premium is $I(0)$. In other words, the expectations model of the term structure implies that long and short rates must be cointegrated with a cointegrating parameter vector of unity (Campbell and Shiller (1987)).

Thus, a simple test of the hybrid model is simply to test for the stationarity of the spread. If this is found to be $I(1)$ when long and short rates are, then this would constitute a rejection of the model, regardless of the method used to form expectations, subject only to the caveat that

agents' forecasting errors are stationary. 1/ If the hybrid model is not rejected at this stage, then it will be worthwhile proceeding to an analysis of the various hypotheses concerning the term premia.

IV. Data

Data was obtained from the Bank of England on the three-month Treasury Bill discount rate and for redemption yields on U.K. Treasury Bonds of maturity 10, 15 and 20 years. The data is sampled weekly (Wednesday 3:00 p.m. rates) from the first week in January 1985 until the last week in November 1989--a total of 253 data points. The redemption yields were measured using the 'new' Bank of England method of measuring the yield curve (Bank of England (1990)). Data on the three-month T-Bill discount and the 15-year T-Bond redemption yield are graphed in Figure 1, while Figure 2 shows the 10-year and 20-year T-Bond redemption yields.

V. Results of the Hybrid Model Tests

Perron (1988) has stressed the importance of distinguishing between stochastic and nonstochastic trends when applying unit root tests. In particular, Perron demonstrates that if a series is stationary about a linear trend but no allowance for this is made in the construction of the unit root test, then the probability of a type II error (failure to reject the unit root hypothesis when it is false) will be high. 2/ Perron suggests the following strategy for testing for unit root behavior in a series y_t. The following regression is estimated by OLS:

$$(10) \quad y_t = \kappa + \lambda(t-T/2) + \delta y_{t-1} + u_t$$

where the sample size is T+1 and u_t may be serially correlated and heterogeneously distributed. The semi-nonparametric test statistics developed by Phillips (1987a,b) and Phillips and Perron (1986) can then be used to test the following hypotheses:

1/ Chow (1989) argues for the importance of allowing for alternative methods of expectations formation in the context of present value models. Taylor (1991) demonstrates the stationarity of forecast errors (when the variable being forecast is I(1)) for a wide variety of expectations formation mechanisms.

2/ Alternatively expressed, the test will lack power. The intuition behind Perron's formal proof can be seen as follows. Suppose the true data-generating process is $y_t = \alpha + \beta t + u_t$, where u_t is stationary white noise--i.e., y is stationary about a linear trend. If we estimate the AR(1) model $y_t = \gamma + \rho y_{t-1} + \epsilon_t$ then ρ will be forced to unity, so that the AR(1) model is equivalent to $y_t = y_0 + \gamma t + \epsilon_t$, where $\epsilon_t = \Sigma_0^t \epsilon_t$, which approximates a linear trend.

Figure 1.   Redemption Yield on Fifteen-Year U.K. T-Bonds (Solid Line)
and Three-Month U.K. T-Bill Discount (Broken Line)
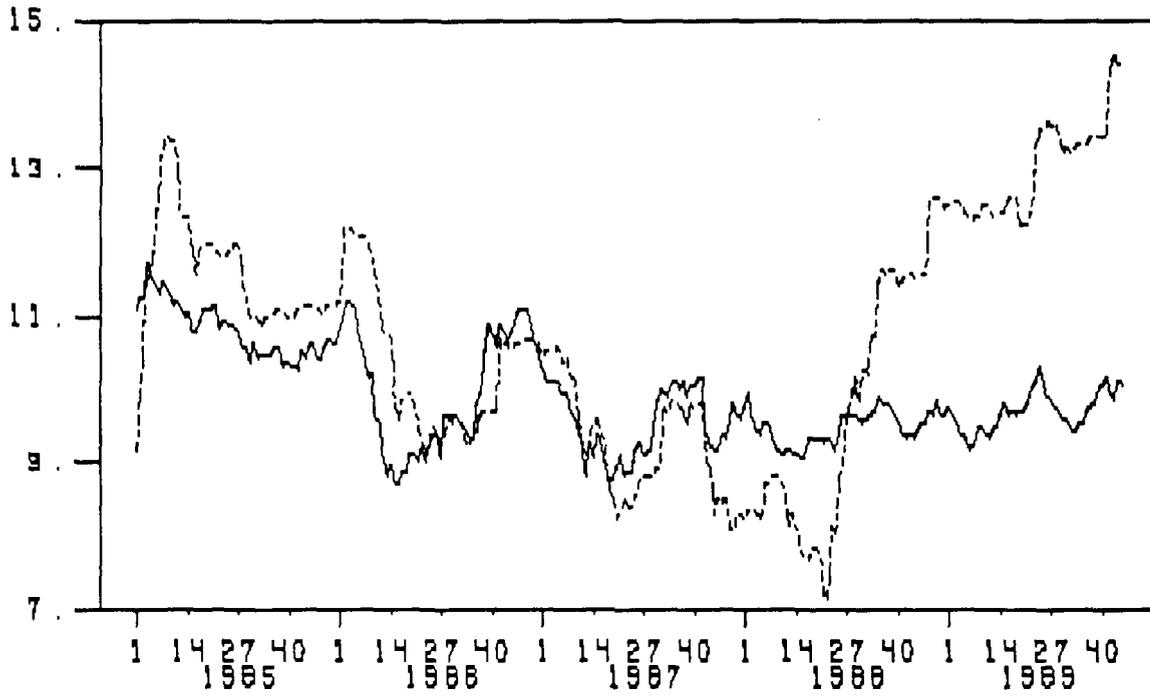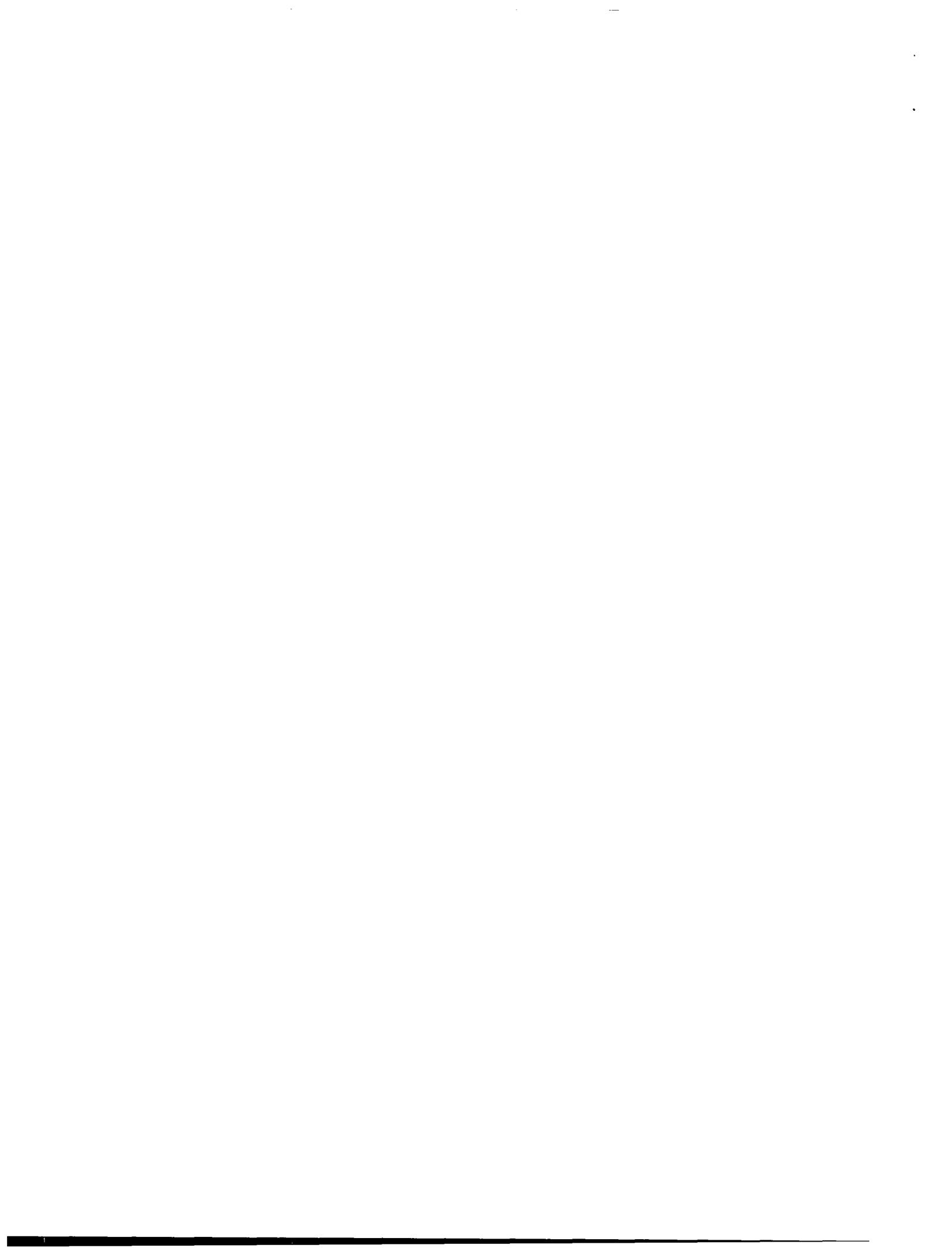
Figure 2.   Redemption on Ten-Year U.K. T-Bonds (Solid Line)
and Twenty-Year U.K. T-Bonds (Broken Line)
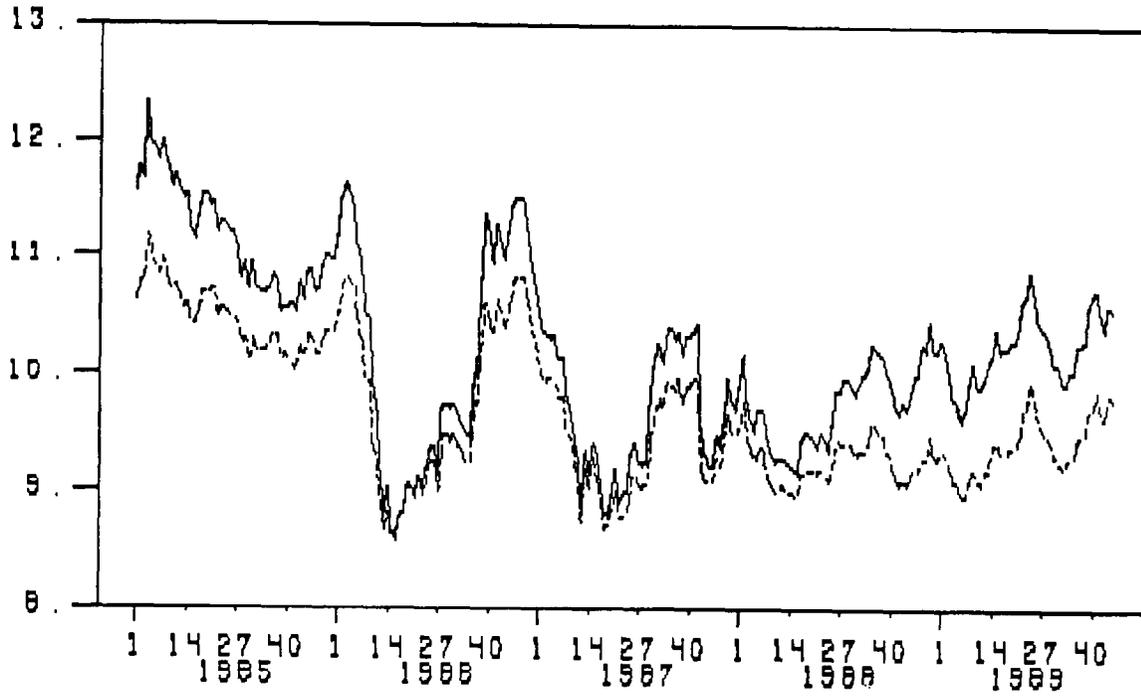
$$H_A:\lambda=1; \quad H_B:(\kappa,\lambda,\delta)=(0,0,1); \quad H_C:(\lambda,\delta)=(0,1)$$

The appropriate test statistics are in fact transforms of the standard t-statistic for H and of the standard F-statistics for H_B and H_C (and we denote them $Z(\tau_\tau)$, $Z(\Phi_2)$ and $Z(\Phi_3)$ respectively). If the unit root hypothesis can be rejected at this juncture, there is no need to proceed. If it cannot, however, then greater test power may be obtained by estimating the regression

$$(11) \quad y_t = \tilde{\kappa} + \tilde{\delta}y_{t-1} + \tilde{u}_t$$

and testing the hypotheses

$$(12) \quad H_D:\bar{\delta}=1 \text{ and } H_E:(\tilde{\kappa},\tilde{\delta}) = (0,1)$$

using the Phillips-Perron transforms of the relevant t-statistic and F-statistic ($Z(\tau_\mu)$ and $Z(\Phi_1)$). This is only valid, however, if the drift term in (10), κ, is zero since $Z(\tau_\mu)$ and $Z(\Phi_1)$ are not invariant with respect to κ. Thus, the statistics $Z(\tau_\mu)$ and $Z(\Phi_1)$ should only be used to provide additional evidence on the unit root hypothesis if the value of $Z(\Phi_2)$ suggests that H_B cannot be rejected (see Perron (1988)).

Following Dickey and Pantula (1987) we tested sequentially for two, one and zero unit roots. The results of applying the Perron strategy to the data are given in Table 1: all of the interest rate series examined appear to be unit root series.

Table 2 contains results of unit root tests applied to the short-long spread series for the four maturities of long bonds considered. In every case the unit root hypothesis is easily rejected. Thus, the hybrid model cannot be rejected. We now develop further the various hypotheses concerning the term premia.

VI. Testing the Expectations Model Assuming Rational Expectations

More stringent tests of the expectations model require that the investigator assume a particular mechanism by which agents form their expectations. Any tests that are carried out are thus contingent on this assumption. In this section we test the ET under the assumption that bond market participants are endowed with rational expectations. According to the rational expectations hypothesis, agents' subjective expectations are identical to the mathematical expectation, conditional on the information set available at the time the forecast is made, Ω_t:

$$_t\Delta_{13} \, _tr^e_{t+i} = E(\Delta_{13}r_{t+i}|\Omega_t)$$

Table 1. Unit Root Tests

Interest Rate	Statistic	$\Delta^2 R_t$	ΔR_t	R_t
Three-month T-Bill	$Z(\tau_\mu)$:	-34.531	-13.790	-0.913
	$Z(\Phi_1)$:	594.752	95.197	0.722
	$Z(\tau_\tau)$:	-34.371	-13.884	-1.064
	$Z(\Phi_2)$:	393.433	64.380	0.879
	$Z(\Phi_3)$:	590.113	96.505	1.009
Ten-year T-Bond	$Z(\tau_\mu)$:	-39.501	-16.107	-0.521
	$Z(\Phi_1)$:	778.637	129.693	3.040
	$Z(\tau_\tau)$:	-39.262	-16.074	-3.857
	$Z(\Phi_2)$:	512.830	86.103	6.694
	$Z(\Phi_3)$:	727.022	129.154	7.448
Fifteen-year T-Bond	$Z(\tau_\mu)$:	-38.264	-14.590	-2.659
	$Z(\Phi_1)$:	730.498	106.491	3.576
	$Z(\tau_\tau)$:	-38.061	-14.610	-2.552
	$Z(\Phi_2)$:	482.163	71.196	2.412
	$Z(\Phi_3)$:	723.243	106.792	3.583
Twenty-year T-Bond	$Z(\tau_\mu)$:	-39.542	-15.100	-2.640
	$Z(\Phi_1)$:	779.956	114.065	3.542
	$Z(\tau_\tau)$:	-39.327	-15.115	-2.700
	$Z(\Phi_2)$:	514.766	76.202	2.649
	$Z(\Phi_3)$:	772.148	114.301	3.920

The null hypotheses and test statistics are discussed in Section 4.1 and
defined in Perron (1988). Allowance for up to fourth-order serial
correlation was made and a Bartlett lag window used in order to ensure
positive definiteness (Newey and West (1987)). The critical values are
as follows (Fuller (1976), Dickey and Fuller (1981)):

Significance levels (in percent):	Critical Values			
	10	5	2.5	1
$Z(\tau_\mu)$	-2.57	-2.88	-3.14	-3.46
$Z(\Phi_1)$	3.78	4.59	5.38	6.43
$Z(\tau_\tau)$	-3.13	-3.43	-3.69	-3.99
$Z(\Phi_2)$	4.03	4.68	5.31	6.09
$Z(\Phi_3)$	5.34	6.25	7.16	8.27

Table 2. Testing for Stationarity of the Long-Short Spread

Long Rate Maturity	τ_μ	$Z(\tau_\mu)$
Ten years	-4.109	-5.223
Fifteen years	-4.333	-4.998
Twenty years	-4.239	-4.902

Note: τ_μ and $Z(\tau_\mu)$ denote the Dickey-Fuller and modified (Phillips-Perron) Dickey-Fuller test statistics for a unit root in the spread series respectively. In constructing the $Z(\tau_\mu)$ statistic we have allowed for up to fourth-order serial correlation and used a Bartlett lag window to ensure positive definiteness (Newey and West (1987)). The null hypothesis in each case is that the spread is I(1). See note to Table 1 for critical values.

Thus the pure (rational) expectations model of the term structure (RETS), in which the term premia are set identically to zero, implies

$$(13) \quad S_t = \sum_{i=1}^{4y-1} \gamma^{13i} E(\Delta_{13} r_{t+13i} | \Omega_t)$$

where $S_t = (R^{(52y)} - r)_t$ denotes the long-short spread.

Now since S_t and $\Delta_{13} r_t$ are stationary, I(0), series, there should exist a bivariate Wold representation (Hannan (1970)), which can be approximated by a vector auto-regression (VAR) of appropriate lag depth m:

$$\Delta_{13} r_t = \sum_{i=1}^{m} \alpha_i \Delta_{13} r_{t-i} + \sum_{i=1}^{m} \beta_i S_{t-i} + u_t$$

(14)

$$S_t = \sum_{i=1}^{m} \gamma_i \Delta_{13} r_{t-i} + \sum_{i=1}^{m} \delta_i S_{t-i} + v_t$$

As noted by Campbell and Shiller (1987), a weak implication of the RETS model is that the spread should linearly Granger-cause changes in the short rate. In terms of the estimated version of (14), this means that the ß coefficients in (14) should be significantly different from zero. The intuition for this is that since, from (13), S_t is an optimal forecast of future $\Delta_{13} r_t$ conditional on the full information set of agents, if agents have information useful in forecasting future short rate changes beyond the history of that variable, it will be reflected in S_t. If they do not, then S_t must be an exact linear function of current and lagged $\Delta_{13} r_t$. 1/

Table 3 contains results of tests of Granger causality from S_t to $\Delta_{13} r_t$ for all of the long bond maturities considered. In order to ensure that the tests were not biased or lacked power because of an inappropriate choice of lag length for the VAR, we considered all lag lengths from four to thirteen. We then considered heteroscedastic-robust linear Wald statistics of the hypothesis

---

1/ Which, in the words of Campbell and Shiller (1987), "is a stochastic singularity which we do not observe in the data."

Table 3. Granger Causality Tests From S_t to $\Delta_{13}r_t$

VAR Lag Length	Linear Wald Statistics $H_0:S_t$ does not Granger-Cause $\Delta_{13}r_t$		
	Ten years	Fifteen years	Twenty years
4	6.87 (0.14)	5.13 (0.27)	4.43 (0.35)
5	6.84 (0.23)	4.86 (0.43)	3.87 (0.57)
6	6.91 (0.33)	4.93 (0.55)	4.05 (0.67)
7	10.33 (0.17)	6.88 (0.44)	6.11 (0.53)
8	8.03 (0.43)	5.94 (0.65)	5.31 (0.72)
9	9.15 (0.42)	6.70 (0.67)	6.24 (0.72)
10	12.82 (0.23)	10.74 (0.38)	9.89 (0.45)
11	13.79 (0.24)	12.51 (0.33)	12.38 (0.34)
12	16.12 (0.19)	15.98 (0.19)	15.83 (0.20)
13	18.16 (0.15)	17.41 (0.18)	16.85 (0.21)

Notes: Linear Wald statistics are central χ^2 variates with degrees of freedom equal to the VAR lag length under the null hypothesis. Figures in parentheses denote marginal significance levels.

(15) H_0: $\beta_i = 0$, $i = 1, .., m$.

As can be seen from Table 3, in no case could the null hypothesis of no Granger causality from spreads to short rate changes be rejected at any reasonable nominal test size, for any of the three long bond maturities, for any choice of lag length. Moreover, this finding was reinforced by an inspection of the individual regression coefficients on the lagged spread terms in the short rate change equations: in no case was any individual spread coefficient significantly different from zero at the 5 percent level. 1/ This evidence is extremely unfavorable to the expectations hypothesis.

A further test of the expectations model can be developed as a set of cross-equation restrictions on the VAR (14), following Campbell and Shiller (1987). The VAR system (14) can be re-parameterized into companion form:

$$
(16) \quad
\begin{bmatrix}
\Delta_{13}r_t \\
\hline
\Delta_{13}r_{t-1} \\
\cdot \\
\cdot \\
\Delta_{13}r_{t-m+1} \\
\hline
s_t \\
\hline
s_{t-1} \\
\cdot \\
\cdot \\
s_{t-m+1}
\end{bmatrix}
=
\begin{bmatrix}
\alpha_1\ \alpha_2 \cdots\ |\ \alpha_m & |\ \beta_1\ \beta_2 \cdots\ |\ \beta_m \\
\hline
 & \\
I\quad |\quad 0 & |\quad 0\quad |\quad 0 \\
\quad | & |\quad\quad | \\
\quad | & |\quad\quad | \\
\hline
\gamma_1\ \gamma_2 \cdots\ |\ \gamma_m & |\ \delta_1\delta_2 \cdots\ |\ \delta_m \\
\hline
 & \\
0\quad |\quad 0 & |\quad I\quad |\quad 0 \\
\quad | & |\quad\quad |
\end{bmatrix}
\begin{bmatrix}
\Delta_{13}r_{t-1} \\
\hline
\Delta_{13}r_{t-2} \\
\\
\Delta_{13}r_{t-m} \\
\hline
s_{t-1} \\
\hline
s_{t-2} \\
\cdot \\
\cdot \\
s_{t-m}
\end{bmatrix}
+
\begin{bmatrix}
u_t \\
\hline
0 \\
\\
\\
\hline
v_t \\
\hline
0 \\
\\
0
\end{bmatrix}
$$

where I is the (m-1)-dimensional identity matrix. The system (16) can be written compactly in an obvious notation as

(17) $Z_t = A Z_{t-1} + V_t$

Now define (2mx1) selection vectors g and h which have unity in the first and (m+1)th elements respectively, and zeros elsewhere, so that:

1/ Asymptotic t-ratios were constructed using a White (1980) heteroscedasticity correction to the covariance matrix.

$$g'Z_t = \Delta_{13}r_t$$

(18)

$$h'Z_t = S_t$$

Let H_t be a restricted information set consisting only of current and lagged values of S_t and $\Delta_{13}r_t$:

$$H_t = \{S_t, S_{t-1}, \ldots, \Delta_{13}r_t, \Delta_{13}r_{t-1}, \ldots\} \subseteq \Omega_t$$

We then have, by the chain rule of forecasting and projecting onto H_t:

$$E(\Delta_{13}r_{t+13i}|H_t) = g'\Lambda^{13i}Z_t$$

and

$$E(S_{t+i}|H_t) = h'\Lambda^i Z_t$$

Thus, projecting both sides of (13) onto H_t and applying the law of iterated mathematical expectations, we have 1/

$$(19) \quad h'Z_t = g'\gamma^{13}\Lambda^{13}(I-\gamma^{13}\Lambda^{13})^{-1}Z_t$$

where I is now the 2m-dimensional identity matrix. If (19) is to hold nontrivially, the following parameter restrictions are imposed on the VAR:

$$(20) \quad h' - g'\gamma^{13}\Lambda^{13}(I-\gamma^{13}\Lambda^{13})^{-1} = 0$$

1/ In deriving (19) we have summed the matrix geometric progression, using $\gamma^n \approx 0$ for large n:

$$\sum_{i=0}^{4y-2} \gamma^{13i}\Lambda^{13i} \approx (I-\gamma^{13}\Lambda^{13})^{-1}$$

Expression (20) is thus a set of 2m cross-equation parameter restrictions which the VAR representation for $(\Delta_{13}r_t, S_t)'$ must satisfy if the RETS is correct. [1]/

A thirteenth-order VAR appeared to be an adequate time series representation of the S_t and $\Delta_{13}r_t$ series, for all three long bond maturities considered, in terms of generating adequate diagnostic statistics. [2]/

In addition to tests of the cross-equation RETS restrictions on the VAR, we also computed a variance bounds test suggested by Campbell and Shiller (1987). Define the theoretical spread, S_t' as the optimal forecast of the right hand side of (13) given the information set H_t, which is just the right hand side of (19):

$$(21) \quad S_t' = g'\gamma^{13}\Lambda^{13}(I-\gamma^{13}\Lambda^{13})^{-1}Z_t$$

The restrictions (20) test the hypothesis

$$(22) \quad H_0: S_t' = S_t, \text{ for all } t$$

Formal tests of (22) (equivalently, (20)) may lead to rejection of the RETS because of tiny deviations from the null hypothesis, such as minor data imperfections or the use of linearizations. Such deviations from the null hypothesis may be statistically important but economically uninteresting. A more informal way of assessing the adequacy of the model is simply to compare the time series movements of S_t and S_t' graphically: economically important deviations from the model should show up as manifest differences in the time series behavior of the two series.

We also computed the variance ratio $Var(S_t)/Var(S_t')$ together with its standard error. This ratio would be expected to be unity if the RETS model were correct, whilst a value in excess of unity would indicate that there is 'excess volatility' in the sense that the spread is too volatile relative to information about future short rates.

1/ Because a constant intercept term was included in our estimated VARs, our tests in fact allow for at most *constant* term premia.

2/ Although it is quite likely that this lag depth was an over-parameterization in some cases, this would have the effect of reducing the test power, so that any rejections that do occur will hold *a fortiori*. Moreover, we obtained qualitatively identical results for all lag lengths between four and thirteen.

Table 4 reports tests of the RETS which were carried out by calculating the (heteroscedastic-robust) Wald statistic for the restrictions on the VAR representation of $(\Delta_{13}r_t, S_t)'$ as given by equation (20), 1/ as well as estimates of the variance ratio. The RETS is massively rejected in every case on the basis of formal tests of the restrictions (20). Moreover, although the variance ratios are in no case massively different from one, they are in each case more than two standard errors away from unity, thus indicating the presence of excess volatility in the long-short spread.

Time series plots of theoretical against actual spreads provide less than compelling evidence in favor of the RETS model. Figure 3 graphs the mean deviation series for the theoretical and actual spread, using the yield on ten year bonds and the three month T-Bill rate: 2/ the two series, whilst clearly related to one another, show a tendency to diverge from one another in a very marked fashion during certain periods. This visual evidence strongly suggests that the formal rejection of the RETS model is due to economically important departures from the model.

Overall, the results reported in this section suggest formal rejection of the RETS model on the basis of simple Granger causality tests from spreads to short rate changes, 3/ on the basis of tests of cross-equation restrictions on the vector autoregressive representation of spreads and short rate changes and on the basis of variance ratio tests. Our excess volatility tests also support the 'tail wags dog' conjecture of Shiller et al (1983) that long-term interest rates tend to overreact to information relevant only to short-term rates. Moreover, a visual comparison of actual and theoretical spreads suggests that the rejection of the RETS is due to economically important deviations from the model.

These results are in stark contrast to those obtained by MacDonald and Speight (1988), who have also tested the RETS on U.K. data using the vector autoregressive technique, and conclude that their evidence 'is sufficient to justify postulating that long and short interest rates in the U.K. are determined by the expectations model of the term structure'. It seems quite likely that the additional test power which is obtained by using overlapping data (MacDonald and Speight use nonoverlapping data) may account for this difference. Additionally, the data used in this paper are likely to be superior to those used by MacDonald and Speight (Bank of England (1990)).

1/ $\bar{R}$ was estimated as the mean redemption yield for each maturity over the sample period as a whole. Qualitatively identical results were obtained using standard Wald tests (i.e., without allowance for heteroscedasticity).

2/ Graphs of the theoretical and actual spread for the other two T-Bond maturities qualitatively similar to Figure 3 and so are not shown.

3/ Strictly speaking, the failure of spreads to Granger cause short rate changes does not constitute a *rejection* of the RETS model, but a failure to confirm it.

Table 4. Tests of the Rational Expectations Model
of the Term Structure

Long Rate Maturity	Wald Statistic $\chi^2(26)$	$Var(S_t)/Var(S_t')$
Ten years	3,471.46 (0.00)	1.503 (0.141)
Fifteen years	3,000.73 (0.00)	1.684 (0.158)
Twenty years	3,445.38 (0.00)	1.755 (0.165)

Notes: Wald statistics are heteroscedastic-robust (Taylor (1989)) and
are central chi-square variates with 26 degrees of freedom under the null
hypothesis. Figures in parentheses below Wald statistics denote marginal
significance levels, those below variance ratios denote standard errors.

Figure 3. Actual Spread (Solid Line) and Theoretical Spread
(Broken Line) Between Ten-Year T-Bond Redemption
Yield and Three-Month T-Bill Discount
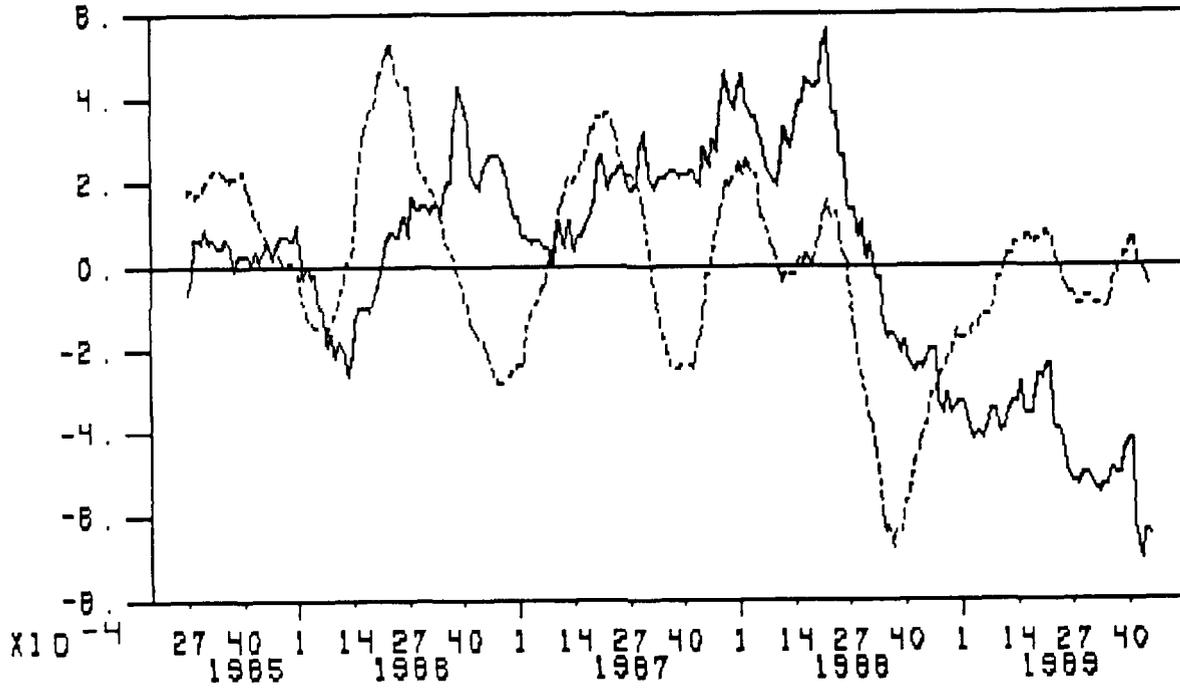
The results of this section are, however, in keeping with a large number of U.S. studies (see Shiller and McCulloch (1987) for a survey). 1/

Since all of the test procedures discussed and applied in this section allow for at most constant term premia, one way of rationalizing the rejection of the RETS model is to argue that there are time-varying risk premia present which are revealed in the data as significant excess volatility in the spread. This argument is pursued in the next section.

VII. Modelling Risk in the Term Structure

One possible cause of the rejection of the RETS is that the term premium is time-varying. Although, as Shiller and McCulloch (1987) note, the theory of the term structure of interest rates is insufficiently developed to suggest precisely which variables ought to explain variation in the term premium, one possibility is that it reflects risk-averse behavior on the part of bond market participants. 2/ Under this interpretation, bond market participants require a reward for bearing risk in the holding of bonds so that a risk premium may separate the holding period return to the long bond and the short rate.

One way of assessing the risk to holding long bonds is to examine the variability of returns. Thus, in a number of US studies, measures of the variability of interest rates have been used, with some degree of success, in statistically explaining variability in term premia (Modigliani and Shiller (1973); Fama (1976); Mishkin (1982); Shiller, Campbell and Schoenholtz (1983); Jones and Roley (1983); Bodie, Kane and McDonald (1984); Engle, Lilien and Robins (1987)).

In this section we follow Engle, et al. (1987) and attempt to model the term premia as dependent on risk arising from unanticipated variability in interest rates. A measure of the extent of unanticipated movements is obtained as the conditional standard deviation of the holding return. We also follow Engle, et al. in modelling the conditional variance of the holding return as an autoregressive, conditionally heteroscedastic (ARCH) process. Our analysis differs from that of Engle, et al. in two important

---

1/ Shiller et al. (1983) point out, however, that despite these rejections the expectations theory continually resurfaces in policy discussions: "We are reminded of the Tom and Jerry cartoons that precede feature films at movie theaters. The villain, Tom the cat, may be buried under a ton of boulders, blasted through a brick wall (leaving a cat-shaped hole), or flattened by a steamroller. Yet seconds later he is up again plotting his evil deeds" (op cit., page 175). The robustness of the theory is presumably due to lack of a sufficiently robust and well-developed alternative.

2/ Shiller and McCulloch (1987) demonstrate that the pure ET model is based on the assumption of risk neutrality.

respects, however. First, since we use a sampling frequency finer than the term to maturity of our short-term interest rate, moving average forecast errors are generated. Secondly, we apply the generalized ARCH (GARCH) formulation due to Bollerslev (1986). The GARCH model allows a much broader class of lag structure than the simple ARCH formulation (and should also give a more parsimonious representation of the data).

The relationship between the term premium and the holding period return is described by equation (5). Because the short rate we use in our applied work is the 13-week Treasury Bill rate, the holding-period return we consider is also of 13 weeks duration. 1/ Thus, the 13-week holding period return to a bond with y years remaining to maturity, $H_{t+13}^{(52y)}$ must differ from the term premium by the 13-week forecasting error, ξ_{t+13} say:

$$(23) \quad H_{t+13}^{(52y)} - r_t = \phi_t^{(52y)} + \xi_{t+13}$$

Because of overlapping news items, any twelve successive values of the error process will in general be correlated. This is consistent with assuming that ξ_{t+13} follows a moving average process of order twelve.

The generalized ARCH-in-mean, or GARCH-m, model of the term premium is given by equation (23) and equations (24)-(27)

$$(24) \quad \phi_t^{(52y)} = \kappa + \lambda \{E(\xi_{t+13}^2|\Omega_t)\}^{1/2}$$

$$(25) \quad \xi_{t+13} = \epsilon_{t+13} + \sum_{i=1}^{12} \vartheta_i \epsilon_{t+13-i}$$

$$(26) \quad \epsilon_{t+i}|\Omega_{t+i-1} \sim IN(0, h_{t+i}^2)$$

$$(27) \quad h_{t+i}^2 = \pi_0^2 + \pi_1^2 \epsilon_{t+i-1}^2 + \pi_2^2 h_{t+i-1}^2$$

1/ Approximate thirteen-week holding returns were generated using a formula analogous to (4)--see Shiller, Campbell, and Schoenholtz (1983), p. 179. Shiller et al. (1983) compare exact holding period returns with this linearization and show the approximation to be extremely close.

Equation (24) defines the term premium as a linear function of the conditional standard deviation of the thirteen-week ahead forecast error. Equation (25) reflects the fact that the pattern of serial correlation of a thirteen-step ahead forecast error is consistent with that of a twelfth-order moving average process. Equations (26) and (27) define the conditional variance of the (weekly) innovations as a GARCH(1,1) model, with parameters expressed as squares in (27) in order to impose non-negativity. For values of π_2^2 less than unity, (27) implies that the conditional variance is an infinite function of past, squared innovations, with geometrically declining weights. It seems intuitively reasonable that agents should discount information in the distant past more heavily--the effective "memory length" is determined by the magnitude of π_2^2. 1/

Stacking all the parameters of the system (23)-(27) into a single vector:

$$\Psi = (\kappa,\ \lambda,\ \vartheta_1,\ \vartheta_2, \ldots, \vartheta_{12},\ \pi_0,\ \pi_1,\ \pi_2)'$$

and applying Schweppe's (1965) prediction error decomposition form of the likelihood function, the log-likelihood for a sample of T observations (conditional on initial observations) is, ignoring the constant term:

$$(28) \quad L(\Psi) = -(1/2) \sum_{t=1}^{T} [\ln(h_t^2(\Psi)) + \epsilon_t^2/h_t^2]$$

1/ The additional nonlinearity induced in the model by using overlapping data is considerable. For example, the conditional standard deviation in (24) is given by the square root of:

$$E(\xi_{t+13}^2 | \Omega_t) = E(h_{t+13}^2 | \Omega_t) + \sum_{i=1}^{12} \vartheta_i^2 E(h_{t+13-i}^2 | \Omega_t)$$

The conditional variance of ϵ_{t+1}, h_{t+1}^2 $[=E(h_{t+1}^2 | \Omega_t)]$ is given by the GARCH equation (27), but the other terms in this equation must be calculated recursively from the formula

$$E(h_{t+i}^2 | \Omega_t) = \pi_o^2 + (\pi_1^2 + \pi_2^2) E(h_{t+i-1}^2 | \Omega_t)$$

The gain in efficiency resulting from increasing the sample size thirteen-fold was thought to be important enough to outweigh these disadvantages, however.

Although the analytic derivatives of (28) can be computed (see Engle, Lilien and Robins (1987)), variable-metric algorithms which employ numerical derivatives were available and performed sufficiently well to obviate the need to employ more complex methods. 1/ Under standard regularity conditions (Crowder (1976)), maximization of (28) will yield estimates with the usual maximum likelihood properties.

The maximum likelihood estimates of model (23)-(27) for each of the maturities considered are tabulated in Table 5. The results are surprisingly uniform across all of the maturities. In particular, although there are signs of significant conditional heteroscedasticity in the data, the slope coefficient for the GARCH-M term, λ, is in each case insignificant, indicating that the conditional heteroscedasticity does not help explain the term premium.

Although we find no support for the risk premium explanation of variation in term premia, using a recently developed and popular method of modelling risk premia, this does not preclude the possibility that another specification of the risk premium may meet with greater success. For example, a standard result in financial economics is that agents should only be rewarded for bearing systematic risk--i.e., risk which cannot be effectively eliminated by holding a sufficiently diversified portfolio. A logical extension of the above approach would thus be to apply a time-varying beta capital asset pricing model to the term structure (see Bollerslev, Engle and Wooldridge (1987); Hall, Miles and Taylor (1988)). We leave this on the agenda for future research and, in the next section, turn to an examination of the 'market segmentation' explanation of term premia.

VIII. Modelling the Market Segmentation Model of the Term Structure

According to the market segmentation or preferred habitat approach to the term structure (see e.g., Modigliani and Sutch (1966)), a bond market participant will have preferences for lending or borrowing at a certain term which might be termed the 'preferred habitat' of that agent. Thus, there exists a separate supply and demand for loanable funds at each habitat and although traders may be 'tempted out of their natural habitat by the lure of higher expected returns' (ibid.) preferences will be sufficiently strong to allow the term premia to deviate from zero. 2/

1/ The likelihood function was maximized using the Broyden-Fletcher-Goldfarb-Shanno positive secant update algorithm (Dennis and Schnabel (1983)). Presumably because of the high degree of nonlinearity involved this method occasionally failed (i.e., failed to converge), and the downhill simplex method of Nelder and Mead (1965) was used to restart the optimization process (see Press, et al. (1986)).

2/ Modigliani and Sutch (1966) argue that, on theoretical grounds, the term premia may change sign; see Shiller and McCulloch (1987) for further discussion and references.

Table 5. Maximum Likelihood Estimates of the GARCH-m Premium Model

$$H_{t+13}^{(52y)} - r_t = \phi_t^{(52y)} + \xi_t$$

$$\phi_t^{(52y)} = \kappa + \lambda \{E(\xi_{t+13}^2 | \Omega_t)\}^{1/2}$$

$$\xi_{t+13} = \epsilon_{t+13} + \sum_{i=1}^{12} \vartheta_i \, \epsilon_{t+13-i}$$

$$h_{t+i}^2 = \pi_0^2 + \pi_1^2 \epsilon_{t+i-1}^2 + \pi_2^2 h_{t+i-1}^2$$

Long Rate Maturity	κ	λ	π_0	π_1	π_2	$\sum_{i=1}^{12} \vartheta_i$	Q(18)
Ten years	0.004	0.008	0.063	0.441	0.322	0.417	14.65
	(0.002)	(0.007)	(0.022)	(0.053)	(0.142)		(0.69)
Fifteen years	0.007	0.006	0.023	0.411	0.404	0.885	17.39
	(0.003)	(0.018)	(0.012)	(0.032)	(0.012)		(0.50)
Twenty years	0.008	0.009	0.095	0.339	0.781	0.919	16.82
	(0.004)	(0.023)	(0.043)	(0.031)	(0.022)		(0.54)

Notes: Q(18) denotes the Ljung-Box test statistic applied to the estimated weekly innovations ϵ_t at 18 autocorrelations. Figures in parentheses below coefficient estimates are estimated asymptotic standard errors, those below the Ljung-Box statistics are marginal significance levels.

The market segmentation approach thus suggests that the term premium should be a function of the supply of and demand for fixed interest debt at the relevant maturity. Over most of the sample period, the U.K. government was deliberately applying the budget surplus to a redemption of fixed interest debt. In relative terms, it seems reasonable to assume that the demand for T-Bonds at each term has remained fairly stable over the period, so that reductions in the amount of debt outstanding might be expected to have raised bond prices at the relevant maturities and thus, ceteris paribus, to have depressed the ensuing holding return.

If $K_t^{(52y)}$ denote the amount of U.K. government fixed interest debt of y years maturity outstanding at time t and T_t denote the total amount of debt outstanding at time t, then a simple formulation of the market segmentation model would be

$$(29) \quad H_{t+13}^{(52y)} - r_t = \alpha + \beta f_1(K_t^{(52y)}) + \gamma f_2(T_t) + \xi_{t+13}$$

where f_1 and f_2 denote continuous, monotonically increasing functions. For simplicity, suppose that these functions are in fact logarithmic:

$$(30) \quad H_{t+13}^{(52y)} - r_t = \alpha + \beta \ln(K_t^{(52y)}) + \gamma \ln(T_t) + \xi_{t+13}$$

An increase in the amount of debt outstanding at a particular maturity would be expected to depress the current price of the relevant bond and so raise the holding period return. Thus, β is expected to be positive. An increase in the total amount of debt outstanding has two effects--a relative effect and an absolute effect. If, as the market segmentation approach suggests, the total amount of debt outstanding is only significant for determining the term premium in so far as the relative amount of debt outstanding at that maturity is affected, then β and γ would be expected to be equal in magnitude and opposite in sign.

Equation (30) was estimated for each of the maturities considered, using weekly data on total government debt outstanding at each maturity. 1/ An ordinary least squares estimator was used, with a method of moments correction to the estimated covariance matrix to allow for moving average errors (of at most twelfth order) and time-varying auto-covariances.

The estimation results (Table 6) are encouraging: the estimated coefficients are statistically significantly different from zero and are of a plausible sign and magnitude. Moreover, the hypothesis that the slope coefficients are equal and opposite can not be rejected at standard significance levels. This therefore provides evidence which is encouraging

---

1/ Data were also obtained from the Bank on the total nominal value of U.K. government fixed interest debt outstanding at each of the above data points, for three maturity bands: over eight and less than ten years, over ten and less than fifteen years and between fifteen and twenty years.

Table 6. Estimates of the Market Segmentation Model

$$H_{t+13}^{(52y)} - r_t = \alpha + \beta \ln(K_t^{(52y)}) + \gamma \ln(T_t) + \xi_{t+13}$$

Long rate Maturity	α	β	γ	$W(H_0:\beta+\gamma=0)$	R^2
Ten years	0.0004	0.0080	-0.0069	1.311	0.24
	(0.0002)	(0.0037)	(0.0031)	(0.25)	
Fifteen years	0.0007	0.0099	-0.0088	1.203	0.42
	(0.0003)	(0.0038)	(0.0031)	(0.27)	
Twenty years	0.0008	0.0138	-0.0099	1.025	0.25
	(0.0003)	(0.0062)	(0.0043)	(0.31)	

Notes: Estimates obtained using ordinary least squares with a method of moments correction to allow for moving average errors and possibly time-varying auto-covariances (Hansen (1982)). R^2 is the coefficient of determination, DW the Durbin-Watson statistic. W(.) is a linear Wald statistic, distributed as χ_2 with one degree of freedom under the null hypothesis. Figures in parentheses below coefficient estimates are estimated heteroscedastic-consistent standard errors, those below test statistics are marginal significance levels.

for the market segmentation approach: it would appear that the policy of repurchasing government debt has been largely responsible for the inversion of the U.K. yield curve over the sample period, over and above any effects government policy may have had upon expected future interest rates.

IX. Conclusion

In this paper we have tested and estimated a variety of alternative models of the yield curve, using weekly, high-quality data on three-month U.K. T-Bill discounts and redemption yields on ten, fifteen and twenty-year U.K. Treasury Bonds.

At the most general level, it was found that a hybrid model of the term structure, consistent with a range of alternative hypotheses concerning the term premium, could not be rejected, when no particular assumptions were made concerning agents' expectations formation mechanisms, except that forecasting errors are stationary. In the remainder of the paper, we pursued various hypotheses concerning the term premium in order to distinguish among the models nested within this hybrid model.

The pure expectations model of the term structure under the assumption of rational expectations (RETS) was easily rejected for all of the maturities considered. The rejection of the RETS, however, may be due not to a failure of rational expectations but to a failure of the assumption of a zero term premium. We then went on to attempt to model the term premium.

One possibility is that term premia reflect risk aversion. We therefore attempted to model risk premia in the term structure by using the conditional volatility of interest rates as a risk measure. Although there was some sign of conditional heteroscedasticity in the holding period return series, a GARCH-in-mean model was, however, unsuccessful in explaining the term premium.

We did, however, achieve more encouraging results when a simple empirical formulation of the market segmentation approach was estimated. In particular, it was found that the relative amount of U.K. government fixed interest debt outstanding at the relevant maturity was a significant determinant of the excess holding return. Our preferred model, therefore, was a hybrid expectations-market segmentation model in which both expectations of future short rates and supply and demand conditions at the relevant maturity determine the long-short rate spread. 1/

The model we have estimated as representative of the market segmentation approach, although encouraging, is clearly ad hoc. Further work might concentrate on developing a more tightly specified model, drawing on theoretical work on the flow of funds. 2/

1/ Masson (1978), using a structural model, reports evidence that supports the market segmentation approach, using data on Canadian government bonds.

2/ See Shiller and McCulloch (1987) for a discussion of this literature in relation to the term structure.

References

Bank of England, "A New Yield Curve Model," <u>Bank of England Quarterly Bulletin</u> (February 1990), pp. 84-85.

Bodie, Z., A. Kane, and R. McDonald, "Why haven't Nominal Rates Declined?" <u>Financial Analysts' Journal</u>, Vol. 40 (1984), pp. 16-27.

Bollerslev, T., "Generalized Autoregressive Conditional Heteroscedasticity," <u>Journal of Econometrics</u>, Vol. 31 (1986), pp. 307-328.

Boughton, J.M., "Exchange Rates and the Term Structure of Interest Rates," (Washington: International Monetary Fund), <u>Staff Papers</u>, Vol. 35 (1988), pp. 36-62.

Campbell, J.Y., and R.J. Shiller, "Cointegration and Tests of Present Value Models," <u>Journal of Political Economy</u>, Vol. 95 (1987), pp. 1062-1088.

Chow, G.C., "Rational Versus Adaptive Expectations in Present Value Models," <u>Review of Economics and Statistics</u> (1989), pp. 376-383.

Clarida, R.H. and B.M. Friedman, "Why Have Short-Term Interest Rates Been So High?," <u>Brookings Papers on Economic Activity</u>, Vol. 2 (1983), pp. 553-578.

Crowder, M.J., "Maximum Likelihood Estimation for Dependent Observations," <u>Journal of the Royal Statistical Society</u>, Series B, Vol. 38 (1876), pp. 45-53.

Dennis, R., and A. Schnabel, <u>Numerical Methods for Unconstrained Optimization and Nonlinear Equations</u> (Englewood Cliffs, NJ: Prentice-Hall, 1983).

Dickey, D.A., and W.A. Fuller, "The Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," <u>Econometrica</u>, Vol. 49 (1981), pp. 1057-1072.

Engle, R.F., and C.W.J. Granger, "Co-Integration and Error Correction: Representation, Estimation and Testing," <u>Econometrica</u>, Vol. 55 (1987), pp. 251-277.

_____, D.M. Lilien, and R.P. Robins, "Estimating Time Varying Premia in the Term Structure: The ARCH-M Model," <u>Econometrica</u>, Vol. 55 (1987), pp. 391-407.

Dickey, D.A., and S.G. Pantula, "Determining the Order of Differencing in Autoregressive Processes," <u>Journal of Business and Economic Statistics</u>, Vol. 5 (1987), pp. 455-461.

Fama, E.F., "Forward Rates as Predictors of Future Spot Rates," <u>Journal of Financial Economics</u>, Vol. 3 (1976), pp. 361-367.

Fisher, I., "Appreciation and Interest," Publications of the American Economic Association (1896), pp. 23-29 and 88-92.

_____, Theory of Interest (New York: MacMillan, 1930).

Fuller, W.A., Introduction to Statistical Time Series (New York: Wiley, 1976).

Hannan, E.J., Multiple Time Series (New York: Wiley, 1970).

Hansen, L.P., "Large Sample Properties of Generalized Method of Moments Estimators," Econometrica, Vol. 50 (1982), pp. 1029-1054.

Hicks, J.R., Value and Capital, 2nd edition (Oxford: Oxford University Press, 1946).

Jones, D.S., and V. Roley, "Rational Expectations and the Expectations Model of the Term Structure: A Test Using Weekly Data," Journal of Monetary Economics, Vol. 12 (1983), pp. 453-365.

Kane, E.J., "Market Incompleteness and Divergences Between Forward and Futures Interest Rates," Journal of Finance, Vol. 35 (1983), pp. 221-234.

Keynes, J.M., Treatise on Money (New York: MacMillan, 1930).

Lutz, F.A., "The Structure of Interest Rates," Quarterly Journal of Economics, Vol. 55 (1940), pp. 33-63.

MacDonald, R., and A. Speight, "The Term Structure of Interest Rates in the UK," Bulletin of Economic Research, Vol. 40 (1988), pp. 287-299.

Masson, P., "Structural Models of the Demand for Bonds and the Term Structure of Interest Rates," Economica, Vol. 45 (1978), pp. 363-77.

Mishkin, F.S., "Monetary Policy and Short-Term Interest Rates: An Efficient Markets-Rational Expectations Approach," Journal of Monetary Economics, Vol. 37, 63-72.

Modigliani, F., and R. Sutch, "Innovations in Interest Rate Policy," American Economic Review, Vol. 56 (1966), pp. 178-197.

_____, and R.J. Shiller, "Inflation, Rational Expectations and the Term Structure of Interest Rates," Economica, Vol. 40 (1973), pp. 12-43.

Nelder, J.A., and R. Mead, "The Downhill Simplex Algorithm," Computer Journal, Vol. 7 (1965), pp. 308-320.

Newey, W.K., and K.D. West, "A Simple, Positive Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," Econometrica, Vol. 55 (1987), pp. 703-708.

Perron, P., "Trends and Random Walks in Macroeconomic Time Series: Further Evidence From A New Approach," Journal of Economic Dynamics and Control, Vol. 12 (1988), pp. 297-332.

Phillips, P.C.B. (1987a), "Time Series Regression with a Unit Root," Econometrica, Vol. 55 (1987), pp. 277-302.

_____ (1987b) "Asymptotic Expansions in Nonstationary Vector Autoregressions," Econometric Theory, Vol. 3 (1987), pp. 45-68.

_____, and P. Perron, "Testing for a Unit Root in Time Series Regression," Biometrika (1986).

Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, Numerical Recipes (Cambridge: Cambridge University Press, 1986).

Schweppe, F., "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Transactions on Information Theory, Vol. 11 (1965), pp. 61-70.

Shiller, R.J., "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," Journal of Political Economy, Vol. 87 (1979), pp. 1190-1219.

_____, J.Y. Campbell, and K.L. Schoenholtz, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," Brookings Papers on Economic Activity (1983), pp. 173-217.

_____, and J.H. McCulloch, "The Term Structure of Interest Rates," National Bureau of Economic Research, Working Paper No. 2341 (1987).

Taylor, Mark P., "Financial Innovation, Inflation and the Stability of the Demand for Broad Money in the United Kingdom," Bulletin of Economic Research, Vol. 39 (1987), pp. 225-233.

_____, "Vector Autoregressive Tests of Uncovered Interest Rate Parity with Allowance for Conditional Heteroscedasticity," Scottish Journal of Political Economy, Vol. 36 (1989), pp. 238-252.

_____, "The Hyperinflation Model of Money Demand Revisited," Journal of Money, Credit and Banking, Vol. 23 (1991), pp. 327-351.

White, H., "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," Econometrica, Vol. 48 (1980), pp. 817-838.