The Economic and Social Review, Vol. 21, No. 4, July, 1990, pp. 337-361

Modelling the Dynamics of the Term Structure of Interest Rates

JAMES M. STEELEY*
University of Keele

Abstract: In order to provide tractable bond pricing formulae, the arbitrage theories of the term structure make specific assumptions as to the number, identity and process generating the underlying forcing variables. This paper assesses the empirical plausibility of these common assumptions. It is found that there are three underlying factors, one more than is usually permitted. However, by careful examination of the dynamics of suitable instrumental variables to these factors, it is found that the further factor may be represented by the autoregressive conditional volatility of one of these factors. Thus, it can be readily integrated into existing two factor models.

I INTRODUCTION

The term structure of interest rates measures the relationship between the prices of a collection of default free pure discount bonds that differ only in their time to maturity. The determinants of this relationship have long been of considerable importance to economists. By providing a complete spectrum of interest rates across time, the term structure represents the whole market's expectations of future events. An explanation of the term structure offers us a method to extract and interpret this information, and to predict how certain underlying factors will affect the whole maturity range of interest rates.

*This paper is based upon Chapters 6-8 of my PhD thesis (1989). An earlier version appeared as paper FORC 89/2 in the Warwick Options Research Centre discussion paper series. I am very grateful to my supervisor Stewart Hodges and to Phil Dybvig, Robert Jarrow, Stephen Schaefer, participants at the University of Warwick Options Conference 1989, and the referees and editor of this journal for helpful comments on the earlier draft. The usual disclaimer applies. Financial support from the ESRC and the Department of Economics at the University of Warwick is acknowledged.

This paper is concerned with extracting and interpreting the information in the term structure when faced with a market comprised of coupon paying bonds, in particular, the market for long-term government fixed interest debt. The crucial difference between this type of market and the pure discount (zero coupon) market, is that the coupon payment feature means that the interest rates (or pure discount prices) are not directly observable. Furthermore, as we shall see, the irregularity of the dates on which coupons are paid compounds the problem further.

Nevertheless, the problem can be overcome, and Section II contains a summary of a new method for extracting the term structure in a market for coupon bonds. This is explained in detail in Steeley (1991). Section III considers the evolution of the term structure curve over time, and employs factor analytic techniques to identify the number of underlying variables influencing the path of the interest rates. Some indirect evidence as to the identity of these factors is discussed.

In order to produce tractable bond pricing formulae, the arbitrage theories of the term structure (Vasicek, 1977; Brennan and Schwartz, 1979; Cox, Ingersoll and Ross, 1985; Ho and Lee, 1986; and Heath, Jarrow and Morton, 1987) typically proceed by making explicit assumptions as to the underlying variables and the processes generating them. These variables tend to be interest rates of particular maturities, and in Section IV, the dynamics of interest rates that could act as instruments to our term structure factors are examined. Section V considers the implications of the findings for the future course of term structure modelling at both the theoretical and practical level.

II MEASURING THE TERM STRUCTURE OF INTEREST RATES

Studies to estimate the term structure of interest rates have used various methods of fitting the following discounting equation.

$$P_i = \frac{C_{i,1}}{(1+R_{i,1})} + \frac{C_{i,2}}{(1+R_{i,2})^2} + \ldots + \frac{C_{i,N(i)}}{(1+R_{i,N(i)})^{N(i)}}, \ \forall i. \qquad (1)$$

Here, bond i makes cash flow payments $C_{i,j}$ at times $j = 1, \ldots, N(i)$, where $N(i)$ is the maturity date of bond i. The set of corresponding spot rates, $R_{i,1}$, $R_{i,2}$, $R_{i,3}$, . . . , will be regarded as the term structure of interest rates in this market.

The particular measurement method chosen is largely determined by the intended use of the interest rate estimates, however, two techniques seem to prevail. Both estimate a linear approximation to the discount function, but differ in their choice of approximation function. McCulloch (1971) used polynomial spline functions, whereas Schaefer (1981) used a set of Bernstein polynomials.

Unless spline functions are carefully chosen, certain matrices formulated for use in the estimation are likely to be ill-conditioned. This section briefly summarises work, detailed elsewhere, that provides a form of approximation function not subject to this problem and which provides reliable estimates of the term structure of interest rates.[1]

2.1 *Methodological Review*

Conceptually, term structure estimation is reasonably straightforward. If we define the discount factor (pure discount price) appropriate to the time point at which bond i makes its jth cash flow payment, as

$$d_{i,j} = \frac{1}{(1 + R_{i,j})^j} \tag{2}$$

then it would seem natural to estimate the discount factors by applying the least squares algorithm to the standard discounting Equation (1). To apply this procedure, however, it is necessary that the number of bonds in the sample exceeds the number of payment dates in the sample. In the USA different bonds pay cash flows on about four key dates in a year, making this technique available, for at least a short-time horizon.[2] However, in the UK there is no such regularity of payment dates and, consequently, the least squares approach is not a realistic option.

Instead, the use of approximation functions means that rather than estimating each discount factor directly, we substitute the following linear approximation to the continuous discount function,

$$d(t) = \sum_{l=1}^{L} \alpha_l f_l(t) \tag{3}$$

and estimate the α_l coefficients that are applied to the L approximating functions chosen. On substitution of this function into our price Equation (1) we obtain,

$$P_i = \sum_{l=1}^{L} \alpha_l \sum_{j=1}^{N} C_{i,j} f_l(t). \tag{4}$$

We still have a linear regression equation but now we can choose how many coefficients we wish to estimate.

1. See Steeley (1991).
2. See Carleton and Cooper (1976).

In using spline functions for the linear approximation, extreme care is required when choosing the form of the component (basis) functions. Not all basis functions are capable of defining regressors useful for reliable estimation. Indeed, Powell (1981, pp. 227-228) shows that it is extremely bad practice to work with a function equivalent to that used by McCulloch (1971) as inaccuracies arise from the subtraction of large numbers, because it generates a regressors matrix that is nearly perfectly collinear.

2.2 *B-splines*

Instead, it is recommended that a basis of B-splines, which are identically zero over a large portion of the approximation space, be used.[3] The function

$$B_p^k(t) = \sum_{l=p}^{p+k+1} \left[\prod_{\substack{h=p \\ h \neq l}}^{p+k+1} \frac{1}{(t_h - t_l)} \right] \max[0,(t-t_l)] \quad -\infty < t < \infty \qquad (5)$$

is known as a k-order B-spline. The subscript p denotes that the function is only non-zero if t (here, time to a payment date) is in the interval (section of the approximation space) $[t_p, t_{p+k+1}]$. The borders between the n sections of the approximation space are known as knots, any approximation space will be spanned by n + k basis functions, and there will be non-zero portions of k + 1 functions in each section of the approximation space. Figure 1 provides example graphs for first, second and third order B-splines.

2.3 *Estimation and Results*

The details of the estimation procedure using B-splines are given in Steeley (1991) and are not reproduced here. However, certain choices made during this process will be discussed here as they directly influence the dimensionality properties of the interest rate data; a key factor in the ensuing analysis.

The dimensionality of the estimated term structure is determined by the number of approximation function coefficients as, deliberately, we are approximating a whole spectrum of rates by a limited number of basis functions. The reason for limiting the number of basis functions was to ensure sufficient degrees of freedom in estimation. The number of basis functions is given by the sum of the order of the B-splines used and the number of sections within the approximation space. Therefore, an increase in either of these two parameters, given a fixed supply of bonds, is made at the expense of reduced degrees of freedom during estimation.

3. Here, the approximation space is the interval across time from zero to the term of the longest maturity bond in the sample.

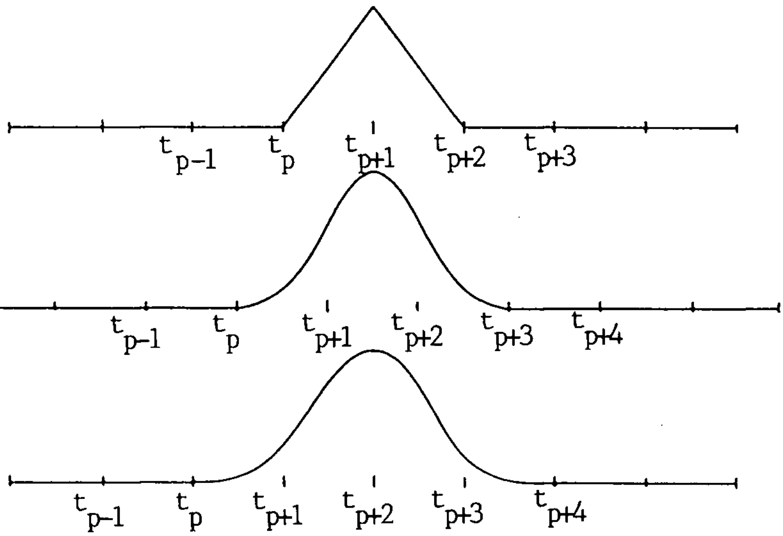

Figure 1: *B-splines of degrees one, two and three*

Cubic B-splines were used. This is the minimum to permit the estimation of forward rate curves with continuous first derivatives. The sectioning of the approximation space was chosen to minimise the standard errors of the estimated interest rates.[4] These choices provide six approximation function coefficients to estimate and set the dimensionality of the interest rate estimates to the same number.

4. The details of this procedure are provided in Steeley (1991).

Figure 2 shows the time path of the B-spline coefficients when the term structure is fitted week by week over the two years from 31 October 1985 to 15 October 1987.[5] We used a sample of 45 UK government gilt-edged stocks, chosen to have fixed coupon and repayment structures and to avoid the type of tax effects examined by Schaefer (1981). The data on prices (closing values for the day), coupons and redemption dates were obtained from Datastream.

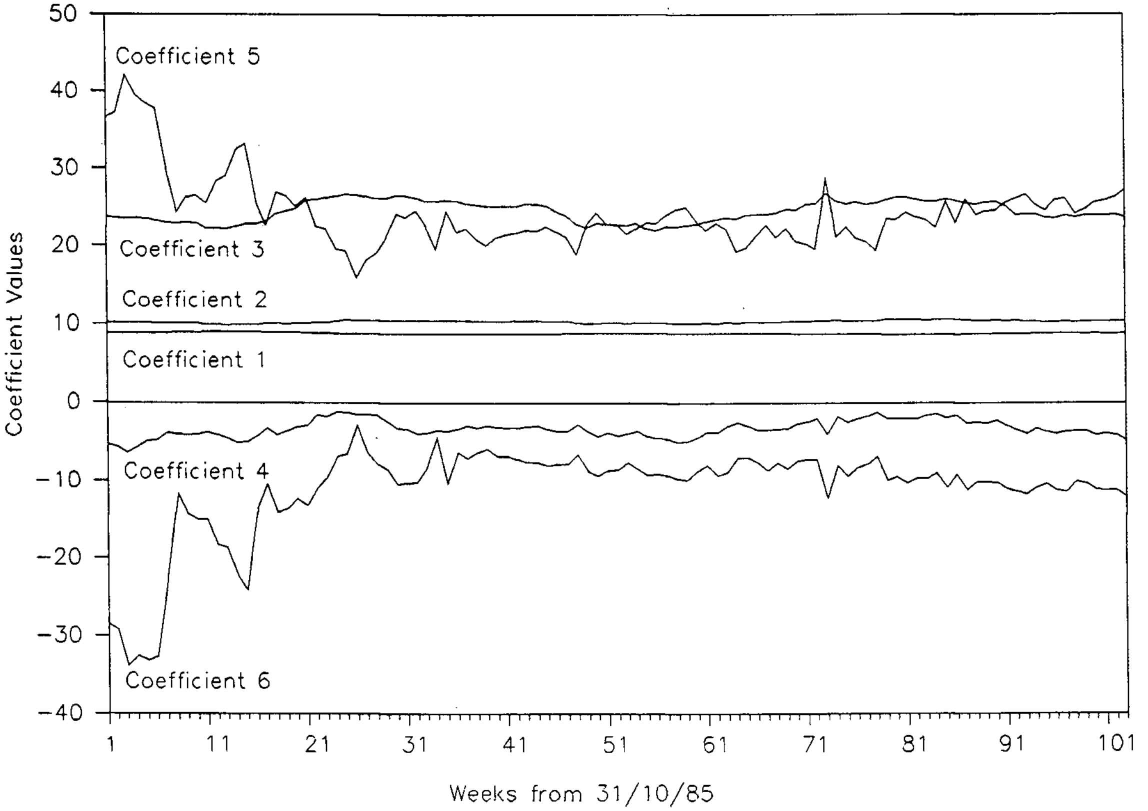

Figure 2: *Time Path of B-spline Coefficients 31/10/85 – 15/10/87*

5. This amounts to 103 observations.

The time path of the term structures of spot rates derived from these coefficients is shown in Figure 3. The graph is drawn for annually spaced maturities, from 1 to 18 years.[6] The term structure appears humped and to have moved extensively in a parallel fashion.

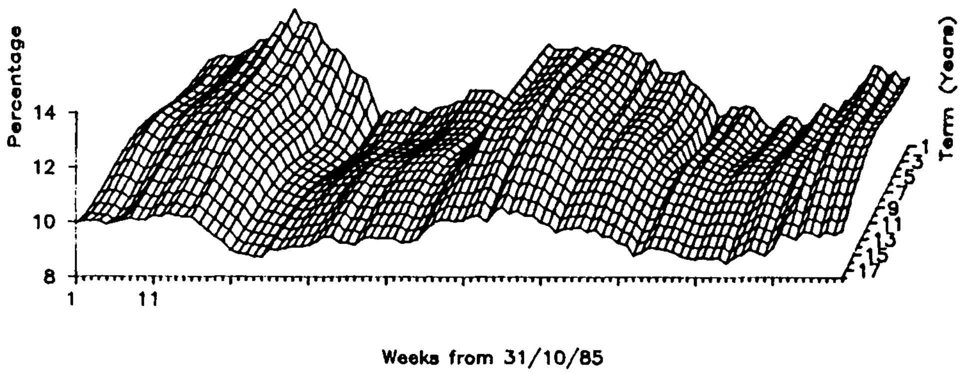

Figure 3: *Time Path of the Term Structure 31/10/85 – 15/10/87*

From Figure 2, we see that of the six basis functions driving the term structure, two seem to experience much more flunctuation than the others. This suggests that the co-movements in the set of spot interest rates may be adequately characterised by fewer than six factors. This question is addressed in the next section.

III COMMON FACTORS IN THE DYNAMICS OF THE TERM STRUCTURE

To establish the number of common factors affecting the movements in the term structure, we apply a principal components decomposition to the covariance matrix of the B-spline coefficients.[7] The eigenvalues and vectors of

6. Given the paucity of data at the long end of the gilt-edged market, 18 years is the maximum maturity for reliable estimation.

7. The method of principal components formally resembles factor analysis. See Lawley and Maxwell (1971) for further details.

this matrix are given in Table 1. The eigenvectors measure the sensitivity of the B-spline coefficients to the underlying factors, and the eigenvalues divided by the number of variables measure the fraction of the variance of the coefficients explained by the factors. The factors themselves have the property that they are mean zero, have unit variance and are orthogonal.

Table 1: *Principal Components Analysis of B-spline Coefficients*

B-spline Coefficient	*Eigenvalues of Covariance Matrix of B-spline Coefficients*					
	4.14672	1.31376	0.33384	0.18485	0.02084	0.00000
	Corresponding Eigenvectors (columns)					
1	0.89126	0.40308	0.18420	0.09607	0.00596	0.00013
2	-0.88470	-0.39402	-0.21249	-0.12989	-0.00520	-0.00013
3	-0.81456	-0.44512	0.18772	0.32577	-0.01345	-0.00012
4	-0.87482	0.14210	0.42049	-0.19008	0.03953	-0.00012
5	0.75425	-0.64768	0.01700	-0.00592	0.10731	0.00011
6	-0.75629	0.60118	-0.20598	0.12826	0.08790	-0.00012

The results of the principal components decomposition are striking. There is one component that dominates the co-variation in the B-spline coefficients. This one component accounts for almost 70 per cent of the variance of the six coefficients. The second and third largest components explain a further 22 per cent and 5 per cent of the variance. Thus over 95 per cent of the variation in the B-spline coefficients can be explained by three component factors.

As with all principal component analyses, the components themselves are not directly identified. However, it is possible to provide some indirect evidence of their impact on the term structure. The eigenvectors represent the sensitivity of the B-spline coefficients to changes in the underlying components. As the components are normalised on calculation, we can interpret the addition (or subtraction) of the eigenvectors to the sample average of the coefficients, as the effect on the mean coefficients of a change in the underlying factors of magnitude two standard errors. By appropriate transformation through Equation (3), we can find the impact of changes in the components on the term structure.

A more direct way, which avoids a number of scaling difficulties with the above indirect route, is to conduct a principal components analysis on the spot rates themselves, thereby obtaining the sensitivities of the term structure to changes in the components (i.e., the eigenvectors) without further transformations. By providing, albeit indirectly, a further factor analysis of the B-spline coefficients, this procedure can also be used to verify the earlier findings.

The choice of interest rates to use in the principal components analysis is essentially arbitrary, but to be consistent with the earlier decomposition, should contain at least six interest rates that also reflect the full spread of maturities in the sample. For caution, we chose to analyse the set of 18 annually spaced spot rates, shown in Figure 3. Once again, only three components are significant. But, the degree of explanation is increased to almost 100 per cent. This is not surprising as, effectively, we have less variation to explain in the eighteen specified points on the curve than in the six coefficients that are intended to represent all points on the curve. The eigenvalues and eigenvectors for the first three components are provided in Table 2.

Table 2: *Principal Components Analysis of the Spot Rates*

Maturity (years)	*Eigenvalues 1-3 for Spot Rates*		
	17.43960	0.37931	0.12885
	Eigenvectors (columns)		
1	0.93789	0.30396	0.15790
2	0.96451	0.24627	0.09352
3	0.98130	0.18818	0.03210
4	0.98950	0.13468	-0.02155
5	0.99223	0.09146	-0.06219
6	0.99312	0.06183	-0.08675
7	0.99377	0.04060	-0.09939
8	0.99426	0.02256	-0.10320
9	0.99453	0.00409	-0.09913
10	0.99478	-0.01853	-0.08731
11	0.99507	-0.04803	-0.06735
12	0.99436	-0.08328	-0.04010
13	0.98141	-0.12069	-0.00694
14	0.97596	-0.15558	0.02945
15	0.97975	-0.18125	0.06541
16	0.97631	-0.18951	0.09536
17	0.97806	-0.17021	0.11227
18	0.97896	-0.11072	0.10617

These eigenvectors are directly the sensitivities of the spot rates to the factors. So, we impact the average spot rate curve by ±2 times the eigenvector for each of the three factors. The graphs representing these effects are given in Figures 4-6. It appears that the first component corresponds to roughly parallel shifts in the spot rate curve. This strongly accords with the manner in which the spot rate curve moved throughout this sample (see Figure 3). The second factor appears to represent changes in the slope of the spot rate curve, while the third factor corresponds to changes in the curvature of the spot rate curve, the well-documented term structure "twist". Two recent unpublished US studies, by Dybvig (1989) and Litterman and Scheinkman (1988) present similar conclusions for US Treasury issues.

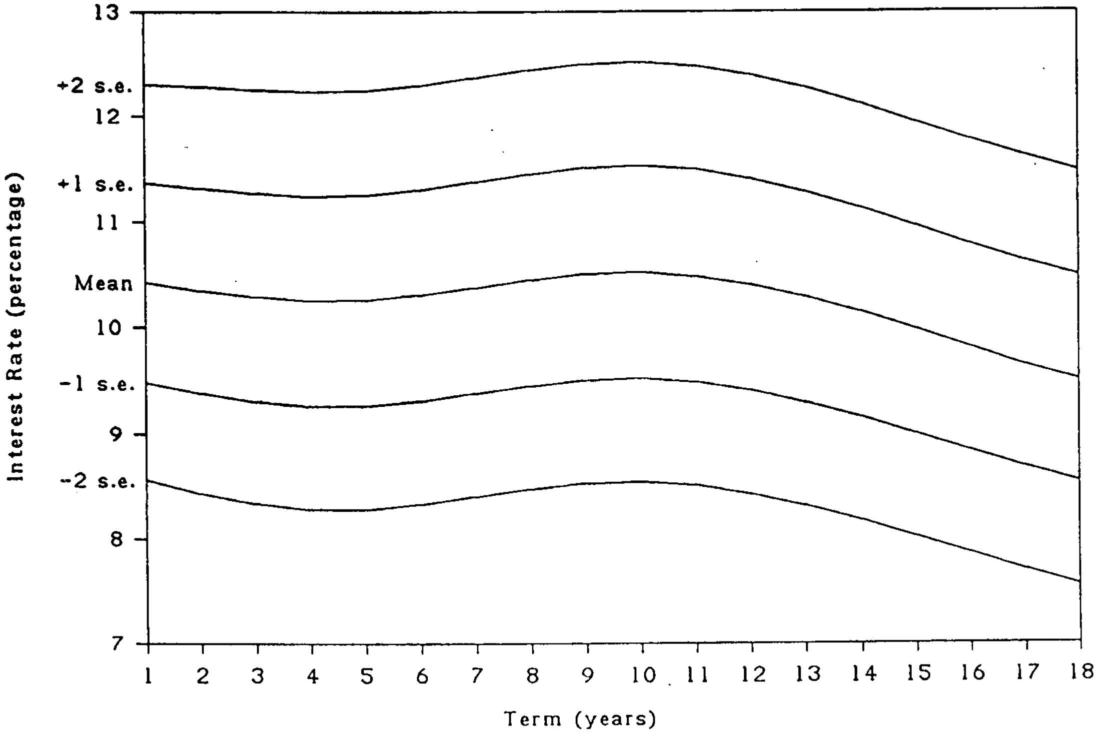

Figure 4: *Impact of Component 1 on Mean Term Structure*

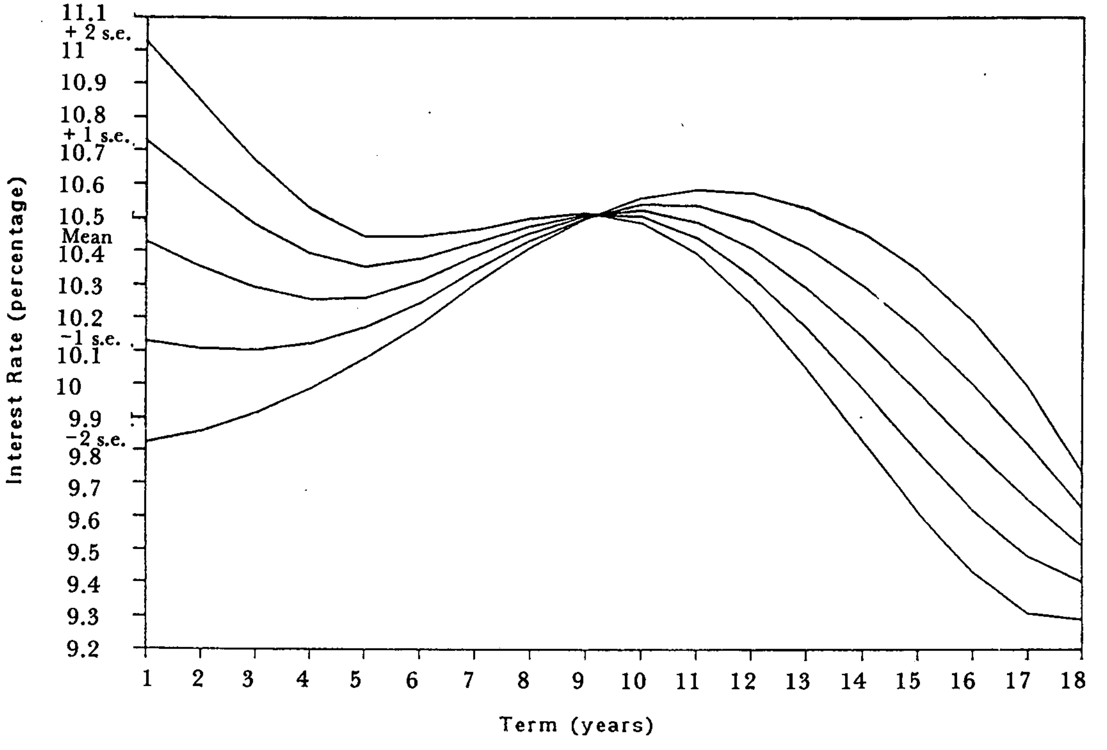

Figure 5: *Impact of Component 2 on Mean Term Structure*

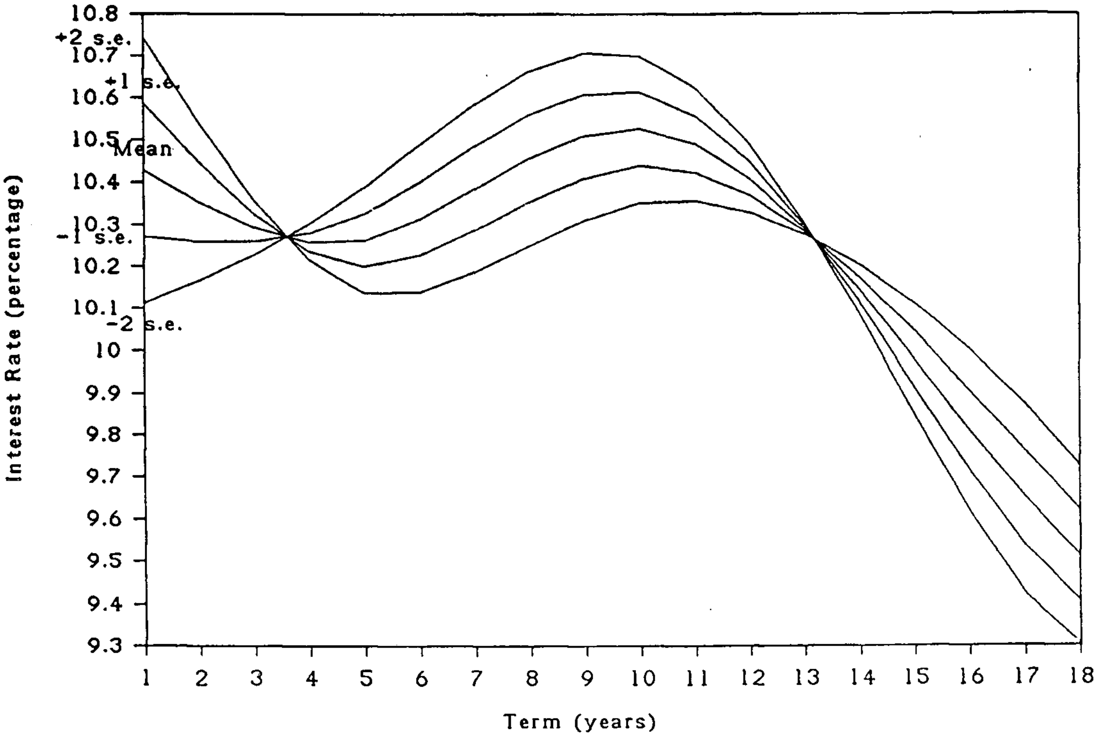

Figure 6: *Impact of Component 3 on Mean Term Structure*

IV MODELLING THE TERM STRUCTURE

In order to generate tractable bond pricing formulae, the arbitrage theories of the term structure make specific assumptions as to the underlying factors and the processes driving them. These variables tend to be interest rates of particular maturities. In this section we examine the dynamic properties of certain interest rate processes that could act as instruments to our underlying factors. Although, working in the spirit of these term structure theories, our modelling process permits a much more general class of stochastic process for the interest rates than is usually assumed. In particular, we consider in detail the variance of the innovations in the interest rate process.

In the one factor models of, for example, Vasicek (1977) and Cox, Ingersoll and Ross (1985), the driving force behind the term structure is assumed to be the instantaneous short interest rate. In the two factor models of, for example, Brennan and Schwartz (1979) and Schaefer and Schwartz (1984), the two factors are represented by the long rate and the short rate, and the long rate and the spread (the difference) between the long rate and the short rate, respectively. The use of the spread rate by Schaefer and Schwartz, which is essentially no more than a redefinition of variables compared to the Brennan and Schwartz model, follows from their proof that a key simplification to obtaining an approximate analytical solution can be obtained if the two state variables are orthogonal. This was after Ayres and Barry (1979, 1980), who had noticed the regularity with which these two rates were uncorrelated. Schaefer (1980) and Nelson and Schaefer (1983) provide further support of this phenomenon.[8]

In the spirit of these one and two factor "interest rate" based models, we will analyse the dynamics of the shortest maturity rate in our sample (the one year rate), the longest maturity rate (the eighteen year rate) and the spread between the two rates. We shall denote their value at time t as r_t, l_t, s_t, respectively. Summary statistics for the series, of 103 observations, are given in Table 3. Time series graphs of the three series appear in Figure 7.

Table 3: *Summary Statistics for Interest Rate Series*

Variable	*Mean*	*Std. Err.*	*Skewness*	*Kurtosis*	*Normality*+	*Corr*#
r	10.4082	1.0867	0.2656	5.0033	5.3147	0.624
l	9.5005	0.5192	0.0095	4.8798	4.2299	1.000
s	-0.9078	0.6617	-0.5162	5.4020	5.9312	-0.077

Notes: + Bera-Jarque Normality Test statistic distributed as $\chi^2(2)$. The 5 per cent critical value is 5.99.

The correlation between first differences in the interest rate and the long rate: the orthogonality test.

8. Models incorporating non-interest rate variables have been considered, by Richard (1978) and Cox, Ingersoll and Ross (1985). See Brennan and Schwartz (1979) for a comparison of the solution properties of those models with their "interest rate" based model.

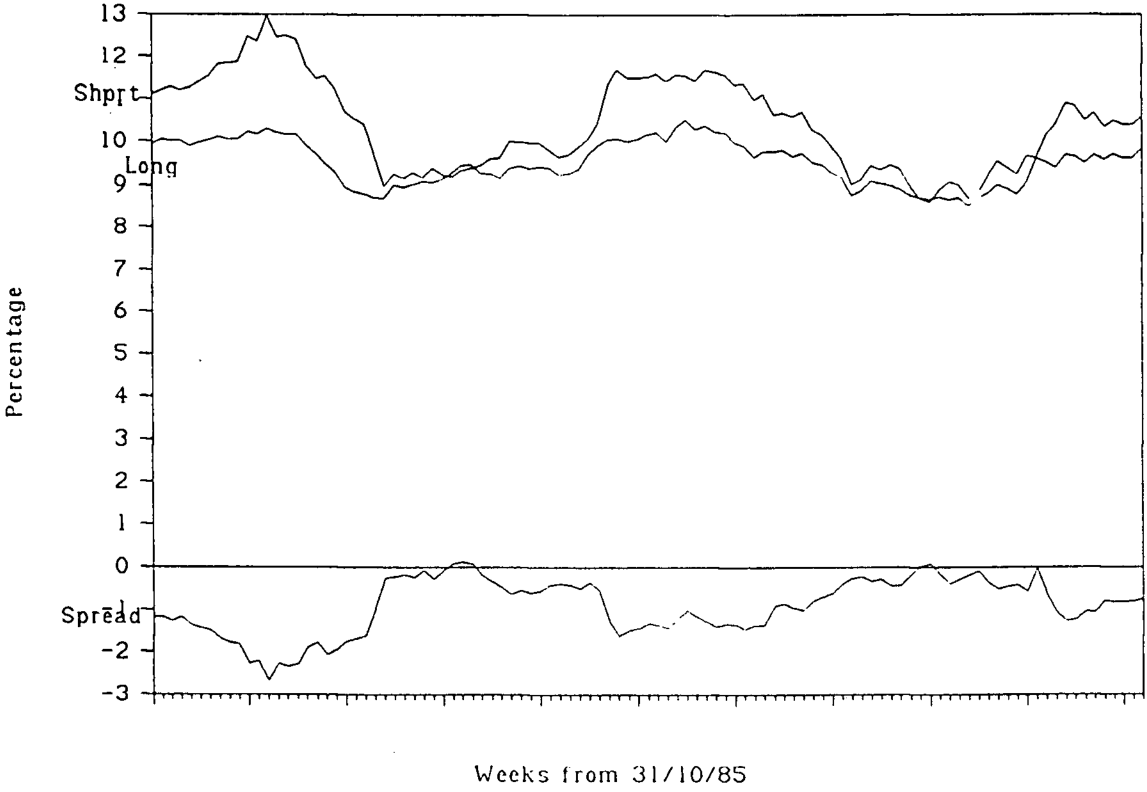

Figure 7: *Path of Short, Long and Spread Rates*
31/10/85 – 15/10/87

From Table 3, it can be seen that, as illustrated by Figure 3, the mean long rate is below the mean short rate. This is characteristic of a humped term structure, and gives rise to a negative mean spread rate.[9] Although the kurtosis figures are all in excess of the value 3 that would be expected if the distributions were normal, by the Bera-Jarque (1982) Lagrange Multiplier test, we cannot reject the hypothesis that the interest rates are normally distributed. It is the hypothesis of normal distributions that is essential for the construction of hypothesis tests and maximum likelihood estimators of the parameters of the interest rate stochastic processes, that are to follow.

The correlations given in Table 3 indicate that we cannot dispute the widely observed result that the spread interest rate and long rate are orthogonal, while the same is not true of the short and long rates.

9. Such humped structures with negative spreads are, of course, not the most usual shape of term structure found in fixed interest security markets.

Tests of the arbitrage term structure theories, for example Brennan and Schwartz (1982) and Brown and Dybvig (1986), use those particular interest rate stochastic processes specified in the original theoretical development, in order to test the implied pricing formulae. Here, I wish to deviate from this traditional starting point and consider more general dynamic processes from the outset. This form of modelling, the general-to-specific approach, is becoming more widely adopted in much econometric analysis.[10] It begins by specifying a very general dynamic model as the maintained hypothesis, and continues by sequentially imposing economically meaningful restrictions on this hypothesis. By modelling from general-to-specific rather than the reverse, we avoid the well-established biases and inefficiencies from accepting an under-parameterised model, yet maintain the property of consistency in our estimated parameters.

Hence, the general process that will be estimated for each of the three series is,

$$\Delta x_t = \alpha(x^* - x_{t-1}) + \sum_{j=1}^{j=n} \beta_j \Delta x_{t-j} + u_t \qquad (6)$$

where x_t is the value of the state variable at time t, u_t is the random innovation to the process at time t, α is the coefficient of mean reversion, and x^* is the long-run mean value of the process. The restriction $\beta_j = 0 \forall j$ gives the discrete time equivalent of the simple diffusion process used in many of the theoretical models, that is,

$$dx = \alpha(x^* - x)dt + \sigma(x)dz \qquad (7)$$

which has instantaneous drift x^*, instantaneous variance $\sigma^2(x)$, and where z(t) is a Wiener process with zero mean and variance dt. The Vasicek (1977) model has $\sigma(x) = \sigma$, a constant; the Cox, Ingersoll and Ross (1985) model has $\sigma(x) = \sigma\sqrt{x}$; and the Dothan (1978) model has $\sigma(x) = \sigma x$.

We may distinguish empirically between the various possible forms of Equation (7), by determining the form of the variance term. The Breusch-Pagan (1979) test for heteroscedasticity may be used for this purpose. Under the null hypothesis of homoscedasticity in the u_t process, $T.R^2$, where T is the number of observations and R^2 is the coefficient of determination, from the regression of u_t^2 on the k variables thought likely to influence the variance, is distributed as $\chi^2(k-1)$. Appropriate choice of variables allows us to distinguish between different forms of heteroscedasticity.

10. This type of modelling, due to Sargan (1964) and Hendry (1979), does not enjoy universal support. Alternative approaches have been developed by, for example, Sims (1980), Leamer (1978) and Zellner (1971).

The above procedure ignores the possibility that the variance term may be influenced by other variables and in more complex ways. A recent class of models, that are enjoying much current support as models for asset prices, are the autoregressive conditional heteroscedasticity (ARCH) models, first proposed by Engle (1982). The attractive property of these models is that they seem to capture the leptokurtosis observed in many unconditional distributions of financial market price variables. Furthermore, a recent study by Nelson (1989) has presented the general conditions for a sequence of finite dimensional discrete time Markov processes to converge to an Ito process, and has derived the diffusion limit of certain models of the ARCH class. This suggests that this type of discrete time process may be a better approximation, than simple Markov discrete time processes, to the simple diffusion processes assumed in many arbitrage theories.

An ARCH(q) model in the variance term (applied, for example, to Equation (7)) is given by

$$\Delta x_t = \alpha(x^* - x_{t-1}) + u_t, \quad u_t \mid \Omega_{t-1} \sim NI(0,h_t) \tag{8}$$

$$h_t = \theta_0 + \sum_{j=1}^{q} \theta_j u_{t-j}^2$$

where Ω_{t-1} is the information set at time t - 1 and h_t is the conditional variance which is a linear function of the last q squared innovations. A Lagrange Multiplier test for the presence of an ARCH(q) model is obtained by calculating $T.R^2$ from a regression of u_t^2 on $u_{t-1}^2, \ldots, u_{t-q}^2$. Under the null hypothesis of no ARCH effects, this statistic will have an asymptotic chi-squared distribution with q degrees of freedom.

An alternative more parsimonious model than the ARCH(q) model is the Generalized ARCH or GARCH model of Bollerslev (1986). The GARCH(p,q) model (also, for demonstration, applied to Equation (7)) is

$$\Delta x_t = \alpha(x^* - x_{t-1}) + u_t, \quad u_t \mid \Omega_{t-1} \sim NI(0,h_t) \tag{9}$$

$$h_t = \theta_0 + \sum_{j=1}^{q} \theta_j u_{t-j}^2 + \sum_{j=1}^{p} \phi_j h_{t-j}^2$$

Pantulla (1986) has shown that the variance equation can be expressed as

$$u_t^2 = \omega + \sum_{j=1}^{m} (\theta_j + \phi_j) u_{t-j}^2 - \sum_{j=1}^{p} \phi_j \nu_{t-j}^2 + \nu_t \tag{10}$$

where m = max(p,q) and ν_t is serially uncorrelated. Thus, u_t^2 will have the usual properties of an ARMA(m,p) process, so that identification tests for the orders of p and m can be carried out on the $\hat{u}_t^2$ series. The GARCH model can be shown to be an infinite order ARCH model with exponentially declining weights and, therefore, allows the estimation of high order ARCH models in a parsimonious manner.

There are three structural changes during our sample period and simple dummy variables are introduced to capture these changes. At the beginning of March 1986, the accrued interest on gilts became an item of income rather than capital for tax purposes. This paved the way for the abolition of capital gains tax on gilts on 2nd July 1986. The third change was the Big Bang deregulation of the market on Monday, October 27th 1986. The value of these variables at time t will be denoted $dAcc_t$, $dTax_t$ and dBB_t.

Table 4 provides the estimates of the parameters of Equation (6), estimated by ordinary least squares, for each of our three interest rate series.

Clearly, from Table 4, the general dynamic equation over-fits all the three interest rate processes, that is, there are a number of statistically insignificant variables in each equation.

A feature common to all the processes is that the dummy variable capturing any effect on the market following the Big Bang deregulation of October 1986 is statistically insignificant. As this dummy variable splits the sample period in approximately equal halves, it also acts as a useful check (maximising degrees of freedom) on the stability of the parameter estimates across the whole sample period.[11] Further, and more rigorous, diagnostic tests will be applied to the final versions of each of the interest rate processes.

The general-to-specific modelling process consists of conducting likelihood ratio tests comparing successively simplified specifications. If the sum of squared residuals from the regression of the next simplification is RRSS, and the sum of squared residuals from the current simplification (initially the general model) is URSS, then the likelihood ratio test comparing the two is given by

$$LR(k) = T \times \ln[RRSS/URSS] \sim \chi^2(k) \qquad (11)$$

where T is the number of observations, and k is the number of restrictions imposed on the current simplification to achieve the next meaningful simplification. If the value of the statistic is less than the appropriate critical value, the next simplification is not rejected. Intuitively, this test determines whether the next simplification is significantly different from the current specification.

11. The usual Chow test is not appropriate where a hypothesis of constant variance is not being maintained.

Table 4: *Least Squares Estimates of the Parameters of the General Dynamic Model for the Three Interest Rate Series*

RHS Variable	Δr_t	Δl_t	Δs_t
Constant+	1.0638	0.9620	-0.2627
	(3.196)	(2.453)	(-2.741)
α	0.0891	0.0971	0.1181
	(3.198)	(2.506)	(2.708)
Δx_{t-1}	0.1285	0.0997	0.1341
	(1.337)	(0.034)	(0.988)
Δx_{t-2}	0.1291	0.5216	-0.0174
	(1.214)	(0.546)	(-0.169)
Δx_{t-3}	0.0157	0.9392	-0.0766
	(0.155)	(0.952)	(-0.899)
Δx_{t-4}	0.0982	0.1664	0.2214
	(0.944)	(1.657)	(0.294)
$dAcc_t$	-0.2387	-0.0820	0.2614
	(-1.989)	(-1.445)	(2.547)
$dTax_t$	0.1627	0.0831	-0.1437
	(1.839)	(1.673)	(-1.756)
dBB_t	-0.0895	-0.0501	0.0667
	(-1.202)	(-1.169)	(1.115)
$\hat{\sigma}$	0.275	0.158	0.217
$\bar{R}^2$	0.09	0.04	0.07

Notes: Figures in parentheses are [zero-centred] "t-statistics", corrected for heteroscedasticity by White's (1980) heteroscedasticity-consistent covariance matrix estimator.
+This is the long run mean x* scaled (multiplied) by α.
The variable $\hat{\sigma}$ is the sample estimate of the volatility [standard deviation] parameter of the process, as though this were a constant value for all time.

If it is not, we choose the next simplification for the obvious reasons of parsimony. This test is the asymptotic equivalent of the familiar "F-test" for the addition of variables, though we are operating in reverse in this situation.

Following the above process, final specifications were settled upon for the three interest rate series, see Table 5.

Table 5: *Least Squares Estimates of the Parameters of the Specific Dynamic Model for the Three Interest Rate Series*

RHS Variable	Δr_t	Δl_t	Δs_t
Constant[+]	0.9112	0.5932	-0.2360
	(3.010)	(2.026)	(-3.165)
α	0.0749	0.0624	0.1120
	(2.853)	(2.034)	(2.909)
Δx_{t-4}		0.2026	
		(1.924)	
$dAcc_t$	-0.2902		0.1614
	(-2.802)		(2.664)
$dTax_t$	0.1609		
	(2.204)		
x^*	12.157%	9.493%	-0.209%
x^*_{Acc}	8.285%		-0.064%
x^*_{Tax}	10.431%		
$\hat{\sigma}$	0.274	0.158	0.215
$\bar{R}^2$	0.06	0.04	0.06
F-Stat(dfs)	1.729 (5,92)	0.907 (6,92)	1,385 (6.96)
LR(df)	9.067 (5)	5.818 (6)	8.732 (6)
DW	1.693	1.751	1.695
LM(4)	5.529	4.004	3.365
BP(df)	9.253 (6)	2.794 (5)	2.273 (4)
ARCH(1)	2.198	0.017	10.358*
ARCH(4)	3.357	0.579	12.631*
ARCH(8)	5.488	2.215	13.478

Notes: Figures in parentheses underneath the parameter estimates are White (1980) heterpscedasticity-consistent [zero-centred] "t-statistics".

+ This is the long-run mean x* scaled (multiplied) by α.

x^*, x^*_{Acc} and x^*_{Tax} are the values of the long-run mean: before the change in tax treatment of accrued interest and removal of liability to capital gains tax; after the former but before the latter; and after both, respectively.

The F-stat and LR figures are tests of whether these simple specifications are significantly different from the general specifications (Table 4).

LM (Lagrange Multiplier test for autocorrelation), BP (Breusch-Pagan test for heteroscedasticity) and ARCH (autoregressive-conditional heteroscedasticiy test) are distributed as χ^2. The regressors in the BP test were the regressors in the model, plus their squares (except for the dummy variables, which would cause perfect multicollinearity) and cross products.

*Indicates statistical significance at the 5 per cent level.

The three specific equations pass all the diagnostic tests, using a 5 per cent significance level, with two exceptions and one near miss. The exceptions are the tests for first and up to fourth order ARCH effects on the spread model. Here the statistics exceed the critical values by a substantial margin. As the latter test is inclusive of the former, these two statistics together indicate clear evidence of a first order ARCH process in the innovations of the spread model. An examination of the autocorrelation function and partial autocorrelation function, in the manner suggested by Box and Jenkins (1976), confirms this view. The "near-miss" statistic is the Breusch-Pagan test for heteroscedasticity in the innovations in the short rate process. Indeed, although not picked up by the ARCH test, the autocorrelation and partial autocorrelation functions of the squared residuals indicate some serial dependence present, and the coefficient on the squared r_{t-1} variable in the Preusch-Pagan equation was statistically significantly different from zero.

Clearly, the spread process should be jointly modelled as the specific process in Table 5 together with a first order ARCH process. As the ARCH(1) process is a parsimonious model able to capture dependence in the variance of the innovations, and because the spread is a redefinition of the short rate, it is also used to model the short rate process, in the hope of flattening the squared residuals in that process also.

Joint generalised maximum likelihood estimates of the conditional mean and variance parameters for each of the short rate and the spread rate are given in Table 6. The standard errors are calculated from analytical first and second derivatives by the method of Berndt, Hall, Hall and Hausman (1974).

The ARCH models, Table 6, for both the spread and the short rate process pass all the diagnostic tests for non-constant variance. Thus these models have achieved the aim of flattening the residuals of the models in Table 5, the only specification problem with those models. Although the tax dummy variable appears insignificant in the short rate ARCH model, which would make both the short rate and spread rate follow similar processes, the same model without this dummy variable is significantly different from one with it present.

Having arrived at preferred specifications for the three possible instrumental variables to our three factors generating the term structure, it is time to interpret the results within this context. In Section III, it was found that the three underlying components could be interpreted as being variables having the following impacts on the term structure: a change in level, a change in slope and a change in curvature. Although we have analysed only two independent processes (because the spread is a combination of the short and long rates), by virtue of the ARCH variance processes in the short and spread rates, we can build a three factor model by combining the long rate with either of these two rates. But this does not imply that such processes necessarily make sense

Table 6: *Generalised Maximum Likelihood Estimates of the Parameters of the ARCH Models for the Short Rate and Spread Rate Series*

RHS Variable	Δr_t	Δs_t
Constant$^+$	1.0766	-0.2680
	(2.238)	(-3.229)
α	0.0856	-0.0937
	(2.195)	(-2.737)
$dAcc_t$	-0.3323	0.2374
	(-2.676)	(3.588)
$dTax_t$	0.1413	
	(1.711)	
$\hat{\theta}_0$	0.2454	0.1642
	(10.463)	(7.688)
$\hat{u}^2_{t-1}$	0.4244	0.6731
	(2.505)	(3.672)
x^*	12.574%	-2.860%
x^*_{Acc}	8.693%	-0.326%
x^*_{Tax}	10.343%	
BP(df)	4.522 (6)	0.803 (4)
LM(4)	1.696	1.118
ARCH(1)	0.033	0.914
ARCH(4)	0.382	1.149
ARCH(8)	0.717	1.893

Notes: Figures in parentheses are "t-statistics", calculated from analytical derivatives following Berndt, Hall, Hall and Hausman (1974).

+This is the long-run mean x* scaled (multiplied) by α.

x^*, x^*_{Acc} and x^*_{Tax} are the values of the long-run mean: before the change in tax treatment of accrued interest and removal of liability to capital gains tax; after the former but before the latter; and after both, respectively.

LM (Lagrange Multiplier test for autocorrelation), BP (Breusch-Pagan test for heteroscedasticity) and ARCH (autoregressive-conditional heteroscedasticity test) are distributed as χ^2. The regressors in the BP test were the regressors in the model, plus their squares (except for the dummy variables, which would cause perfect multicollinearity) and cross products.

as the three required instrumental variables. However, it is not unrealistic to suppose that a model comprising the long rate (for level), the spread rate (for slope), and the volatility of the spread rate (for curvature) might provide an adequate starting point. The development of such a model is beyond the bounds of this paper and, therefore, is left open to further research.

It is worth noting, of course, that we may have detected ARCH effects because we are approximating a continuous time diffusion process by a discrete time difference equation, as discussed earlier (see, also, Nelson (1989)). However, this does not detract from the view, of at least this author, that the conditional volatility of one of the state variables may well prove to be a useful third factor in established models of the term structure.

V SUMMARY AND IMPLICATIONS

This paper seeks to model the dynamics of the term structure of interest rates, that is, the set of prices of pure discount (zero-coupon) bonds that differ only in their time to maturity. However, the term structure that is to be modelled is that in a market for coupon-bearing bonds. In such a market, the pure discount prices are not directly observable, and this paper begins by outlining a new method for extracting the pure discount prices. The market chosen to illustrate this method, and analysed in subsequent sections was the market for British government fixed interest securities, or "Gilts".

By using factor analytic techniques, it is possible to determine that there are three key factors underlying, and driving, the term structure in the gilt-edged market. This result is consistent with recent unpublished US studies of their Treasury market. These factors can be interpreted as causing the term structure to change its level, its slope, and its curvature.

In the arbitrage theories of the term structure, such factors are usually represented by instrumental variables in the form of interest rates of particular maturities. This permits tractable bond pricing solutions to be generated. The most elaborate models, for which practical solutions have been obtained, consist of two factors. The principal components analysis in this paper would suggest that these models are missing one key factor. By analyzing the dynamics of interest rates of particular terms, in the spirit of these models, and using the general-to-specific methodology, it is found that the dynamics of certain rates are well represented by ARCH (autoregressive conditional heteroscedasticity) models. It is believed that such stochastic volatility processes may hold the key to generating three factor models of the term structure that can outperform the existing two factor models in bond pricing, bond portfolio management (duration and immunisation) and bond option pricing.[12]

12. Readers with a particular interest in this area of research may like to consult a recent paper by Engle, Ng and Rothschild (1989) that also considers ARCH processes and factor analytic techniques. They factor analyse the covariance matrix of the innovations of an assumed conditional mean process in order to determine how many factors are influencing the conditional variance alone. This assists the accurate construction of the conditional variance portion of the model. Their empirical results for pricing US Treasury issues appear very promising indeed.

REFERENCES

AYRES, H.R., and J.Y. BARRY, 1979. "The Equilibrium Yield Curve for Government Securities", *Financial Analysts Journal*, May/June, pp. 31-39.

AYRES, H.R., and J.Y. BARRY, 1980. "A Theory of the US Treasury Market Equilibrium", *Management Science*, Vol. 26, pp. 539-569.

BERA, A.K., and C.M. JARQUE, 1982. "Model Specification Tests: A Simultaneous Approach", *Journal of Econometrics*, Vol. 20, pp. 59-82.

BERNDT, E.K., B.H. HALL, R.E. HALL, and J.A. HAUSMAN, 1974. "Estimation and Inference in Non-Linear Structural Models", *Annals of Economic and Social Measurement*, October, pp. 653-665.

BOLLERSLEV, T., 1986. "Generalized Autoregressive Conditional Heteroscedasticity", *Journal of Econometrics*, Vol. 31, pp. 307-327.

BOX, G.E.P., and G.M. JENKINS, 1976. *Time Series Analysis, Forecasting and Control*, San Francisco: Holden Day.

BRENNAN, M.J., and E.S. SCHWARTZ, 1979. "A Continuous-Time Approach to the Pricing of Bonds", *Journal of Banking and Finance*, Vol. 3, pp. 133-155.

BRENNAN, M.J., and E.S. SCHWARTZ, 1982. "An Equilibrium Model of Bond Pricing and a Test of Market Efficiency", *Journal of Financial and Quantitative Analysis*, Vol. 17, pp. 301-329.

BREUSCH, T.S., and A.R. PAGAN, 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation", *Econometrica*, Vol. 47, pp. 1287-1294.

BROWN, S.J., and P.H. DYBVIG, 1986. "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates", *Journal of Finance*, Vol. 51, pp. 617-630.

CARLETON, W.T., and I.A. COOPER, 1976. "Estimation and Uses of the Term Structure of Interest Rates", *Journal of Finance*, Vol. 31, pp. 1067-1083.

COX, J.C., J.E. INGERSOLL Jr., and S.A. ROSS, 1985. "A Theory of the Term Structure of Interest Rates", *Econometrica*, Vol. 53, pp. 385-407.

DOTHAN, L.U., 1978. "On the Term Structure of Interest Rates", *Journal of Financial Economics*, Vol. 6, pp. 59-69.

DYBVIG, P.H., 1989. "Bond and Bond Option Pricing based on the Current Term Structure", *unpublished discussion paper*, Washington University in St. Louis.

ENGLE, R.F., 1982. "Autogressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation", *Econometrica*, Vol. 50, pp. 987-1067.

ENGLE, R.F., V.K. NG, and M. ROTHSCHILD, 1989. "Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills", *unpublished discussion paper*, University of California at San Diego.

HEATH, D., R. JARROW, and A. MORTON, 1987. "Bond Pricing and the Term Structure of Interest Rates: A New Methodology", *unpublished discussion paper* (Cornell University).

HENDRY, D.F., 1979. "Predictive Failure and Econometric Modelling in Macroeconomics: The Transactions Demand for Money", in P. Ormerod (ed.), *Economic Modelling*, Heinemann, pp. 217-242.

HO, T.S., and S. LEE, 1986. "Term Structure Movements and Pricing Interest Rate Contingent Claims", *Journal of Finance*, Vol. 41, pp. 1011-1029.

LAWLEY, D.N., and A.E. MAXWELL, 1971. *Factor Analysis as a Statistical Method*, 2nd Edition, London: Butterworths.

LEAMER, E.E., 1978. *Specification Search: ad hoc Inference with Non-Experimental Data*, New York: Wiley.

LITTERMAN, R., and J. SCHEINKMAN, 1988. "Common Factors affecting Bond Returns", *Financial Strategies Group discussion paper*, Goldman Sachs.

McCULLOCH, J.H., 1971. "Measuring the Term Structure of Interest Rates", *Journal of Business*, Vol. 44, pp. 19-31.

NELSON, D.B., 1989. "ARCH Models as Diffusion Approximations", *Journal of Econometrics*, forthcoming.

NELSON, J., and S.M. SCHAEFER, 1983. "The Dynamics of the Term Structure and Alternative Portfolio Immunization Strategies", in G. Kaufman, G. Bierwag, and A. Toevs, *Innovations in Bond Portfolio Management: Duration Analysis and Immunization*, Greenwich, Conn.: JAI Press, pp. 239-265.

PANTULLA, S.G., 1986. "Comment", *Econometric Reviews*, Vol. 5, pp. 71-74.

POWELL, M.J.D., 1981. *Approximation Theory and Methods*, Cambridge University Press.

RICHARD, S.F., 1978. "An Arbitrage Model of the Term Structure of Interest Rates", *Journal of Financial Economics*, Vol. 6, pp. 33-57.

SARGAN, J.D., 1964. "Wages and Prices in the United Kingdom: A Study in Econometric Methodology", in P.E. Hart, G. Mills and J.K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*, London: Butterworths.

SCHAEFER, S.M., 1980. "Discussion", *Journal of Finance*, Vol. 35, pp. 417-419.

SCHAEFER, S.M., 1981. "Measuring a Tax-Specific Term Structure of Interest Rates in the Market for British Government Securities", *Economic Journal*, Vol. 91, pp. 415-438.

SCHAEFER, S.M., and E.S. SCHWARTZ, 1984. "A Two-Factor Model of the Term Structure: An Approximate Analytical Solution", *Journal of Financial and Quantitative Analysis*, Vol. 19, pp. 413-424.

SIMS, C.A., 1980. "Macroeconomics and Reality", *Econometrica*, Vol. 48, pp. 1-48.

STEELEY, J.M., 1989. "The Effects of Big Bang on the Gilt-Edged Market: Term Structure Movements and Market Efficiency", *unpublished PhD thesis*, University of Warwick.

STEELEY, J.M., 1991. "Estimating the Gilt-Edged Term Structure: Basis Splines and Confidence Intervals", *Journal of Business Finance and Accounting*, forthcoming.

VASICEK, O.A., 1977. "An Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, Vol. 5, pp. 177-188.

WHITE, H., 1980. "A Heteroscedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroscedasticity", *Econometrica*, Vol. 48, pp. 817-838.

ZELLNER, A., 1971. *An Introduction to Bayesian Inference in Econometrics*, New York: John Wiley and SAS.