Applied Financial Economics, 1995, **5**, 273–283

The expectations theory of the term structure: a cointegration/causality analysis of US interest rates

ROBERT J. MANDENO and DAVID E. A. GILES

Bankers Trust, N.Z. Ltd., Auckland, New Zealand and *Department of Economics, University of Victoria, PO Box 3050, Victoria, BC, Canada, V8W 3P5*

This study analyses monthly US Treasury short and long term interest rate data over the period 1950–82. We test for Granger causality between the short and long rate series in the context of an appropriately formulated vector autoregression which takes account of the non-stationarity of the data, and possible cointegration between the series. Unlike earlier studies, an allowance is also made for the seasonality of the monthly data, and special attention is paid to the presence of a major structural break in the data. Bootstrap simulation techniques are used to obtain critical values for unit root and cointegration tests which allow for this structural change. We find strong evidence that long rates caused short rates over the sample period, and some evidence of causality in the reverse direction. This supports the expectations theory of the term structure.

I. INTRODUCTION

There has been extensive research conducted in the area of developing and testing models and hypotheses to explain the relationship between short and long term interest rates. Generally, market participants have held the conventional belief that changes in long term interest rates follow changes in short term interest rates (e.g., Mankiw, 1986). However, the short-to-long interpretation is by no means a consensus (e.g., McFadyen *et al.*, 1991). A great deal of the recent literature has investigated the expectations hypothesis of the term structure where the long term rate is assumed to be a weighted average of expected future short rates. If such a relation holds then long term interest rates may contain useful information for predicting short term rates. The question of whether the expectations theory of the term structure is valid has obvious important implications for policy makers, financial institutions and any debt market participants with exposure to interest rate risk. Thus, there is good reason for investigating short and long term interest rates for the existence and direction of any causal relationships.

This study involves analysing US interest rate data for possible cointegration/causality between short and long term rates. To test for causality we estimate a vector autoregressive (VAR) model for the two series and test the restrictions implied by non-causality. Special attention is paid to data seasonality and structural breaks in the series. The layout of the rest of the paper is as follows. Section II provides some insight into the expectations theory of the term structure of interest rates, and summarizes some of the major previous empirical results. Section III considers the data used, and highlights some of the special problems they present. The econometric methodology applied in this study is described in detail in section IV, and the results are presented in section V. The final section provides some concluding remarks.

II. THE TERM STRUCTURE OF INTEREST RATES

The literature on the term structure of interest rates provides many different theories including preferred habitat, liquidity preference and expectations theories. Of these the preferred habitat is the most general and asserts that the investor prefers holding debt securities with maturities matching their investment planning horizon, but will hold

0960–3107

shorter or longer maturities if they expect to earn a risk premium (McFadyen *et al.*, 1991).

The expectations theory is a special case of the preferred habitat theory in which no risk premium is needed to entice investors to mismatch their planning horizon and the time until maturity. Thus, rational investors may either 'roll over' a series of one period bonds r_t or invest in an n-period bond R_t^n (Meiselman, 1962). From this traditional interpretation of the expectations theory we can write the long term rate as the present value of expected future short rates as follows:

$$R_t^n = \frac{1-a}{1-a^n} \sum_{k=0}^{n-1} a^k E_t(r_{t+k}) \tag{1}$$

where $a = 1/(1 + R^*)$ and R^* is the mean long bond rate.

This assertion that long rates are composed of predicted short rates, combined with the assumption that these predictions are reasonable and possibly unbiased, suggests that long rates contain information that would be useful in predicting future short rates. Thus, we would expect a long to short causal relationship under this hypothesis. Further, McFadyen *et al.* (1991) claim that the expectations hypothesis is consistent with both bidirectional causality and a long to short causal relationship.

Beginning with Macaulay (1938) much of the empirical research on this topic has focused on the expectations theory of the term structure. Macaulay found no evidence to support this hypothesis, and many more sophisticated studies conducted later on US postwar data also rejected this hypothesis (e.g., Campbell and Shiller, 1984; Fama, 1984; and Mankiw, 1986). Recently, however, Froot (1989) has used survey data on interest rate expectations in the US and has found some support for the theory, as have McFadyen *et al.* (1991) and Wallace and Warner (1993). Hamilton (1988) has also studied this problem with US data. MacDonald and Speight (1988, 1991) analysed UK, Belgian, Canadian, German and US data, and also found *some* evidence to support this theory.

In some of these studies the predictive power of the 'spread' between long and short term interest rates in forecasting future short term interest rate changes is used as the basis for testing the expectations theory of the term structure. Some authors, including Campbell and Shiller (1987) and Fama and Bliss (1987), have found the spread to have had some positive predictive power, particularly for short rate changes further into the future. In contrast, Taylor (1992) provides results from this approach (and based on UK data) which reject the expectations theory (and the risk premium model of the term structure), but support a market segmentation model instead. Karfakis and Moschos (1993) use modern cointegration techniques of the type that we adopt, to analyse the predictive power of the 'spread' with quarterly Australian data, and also find some support for the expectations theory. In contrast, using similar techniques, Shea (1992) finds negative results with US data.

The focal point of studies of this type involves testing whether the 'spread' 'causes' changes in the short rate, where causality is defined in the sense of Granger (1969): Y_t causes X_t if we are better able to predict Y_t by using past information about X, as well as about Y, than if we just use past information about Y itself. The motivation for this is clear from Equation 1: subtracting the short rate, r_t, from both sides of that equation we see that the spread ($S_t = R_t^n - r_t$) can be represented (linearly) in terms of expected future *changes* in the short rate of interest (Δr_t). As Taylor (1992, p. 530) notes, under the expectations hypothesis '... S_t is an optimal forecast of future Δr_t conditional on the full information set of the agents. So, 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 Δr_t.' So, following Campbell and Shiller (1987), the spread should linearly Granger-cause Δr_t.

Other authors have tested for causality between the short and long rates directly, using a similar methodology to that employed in this study, as a means of testing the expectations theory of the term structure. Provided that the non-stationarity characteristics of the data are taken into account, the rationale for this form of causality testing follows from the above discussion if one addresses Equation 1 directly, rather than manipulating that relationship in terms of the spread. For example, Krol (1987) tests for Granger causality between long and short term Euro-dollar interest rates, and fails to find any such relationship. More recently, Lütkepohl and Reimers (1992) have analysed US data for the period from 1957–90 and concluded that the long term rate was clearly causal for the short term rate and possibly there existed feedback between the two. However, two criticisms that can be made of their study are, (1) their analysis is carried out on monthly data with no consideration made for the possibility of seasonal integration and cointegration; and (2) there is no allowance made for readily observed structural breaks in the time series data. These are matters which are addressed explicitly in this paper.[1]

Many economic time series have been found to be (homogeneous) non-stationary–differencing such data an appropriate number of times (say, d) yields a stationary series. In this case the original series is said to be 'integrated of order d', or 'I(d)' (Engle and Granger, 1987, p. 252). A series that is I(1) has a unit root (at the zero frequency) in the level of the series. It is well known that regression analysis involving non-stationary time series generally leads to spurious regressions, which are typically characterized by a low Durbin–Watson statistic and a high R^2. The natural thing to do,

[1]Other related studies which have appeared recently in this journal include those of Moazzami (1991) Owen (1993) and Margaritis (1994).

then, for a series found to be I(d), is to difference the data d times before fitting the regression. Usually a linear combination of two series which are both I(1) will result in a series which is also I(1) (Engle and Granger, 1987). However, in certain cases there exists a linear combination of the two I(1) series that yields a stationary series (i.e., I(0)). In this case, the series are said to be 'cointegrated' (Granger, 1981), and there is an 'error correction' mechanism that relates the series. It is essential to establish whether there exists a cointegrating relationship between the data, as this affects the appropriate formulation of any functional model, such as the VAR models which form the basis for standard tests for (Granger) causality. We will be interested in the possibilities of both uni-directional and bi-directional causality (e.g., Guilkey and Salemi, 1982) between short term and long term interest rates.

III. DATA ISSUES

Monthly series for the period from February 1950 to December 1982 are analysed in this study. This sample period was chosen to allow a five year lag after the Second World War to avoid any adverse interference, while providing a sufficiently long span to facilitate the identification of possible unit roots and cointegrating relationships. The 1982 termination point coincides with a change in the operating regime of the Federal Reserve from framing short-run monetary policy in terms of targeting non-borrowed reserves to targeting borrowed reserves. The short term interest rate is measured by the one month US Treasury bill rate and long term rate is the monthly yield to maturity of 20 year US Treasury bonds (both expressed as a percentage per annum), from various issues of *Business Conditions Digest* (US Department of Commerce). In total, each monthly series comprises 395 observations, and these are shown in Fig. 1.

We see that both series trend upwards over this period with the long rate, in general, exceeding the short rate, and the latter data displaying significantly more variability than the former. As well as having more pronounced peaks and troughs, the short rate series shows greater noise in the sense of small local perturbations. This suggests the possibility of seasonal variations in the short rate series. On the other hand, the long rate series shows no noticeable seasonal fluctuations. To ascertain the extent of any seasonality we applied a seasonal filter to both series using the ratio to moving average method. The geometric mean of the seasonal factors for each month confirmed that there is more seasonality in the short rate series. However even for the short rate series the seasonal variation is relatively minor.

Another distinguishing feature of Fig. 1 is the presence of quite definite structural breaks throughout the sample period. In the period from 1950–79 there are five distinct structural breaks. Typically, these breaks are characterized by the short rate plummeting after a period of sustained growth. In each instance, the long rate also falls, but by a much smaller amount. In the period following 1979, both series show extreme volatility with another quite marked break occurring in 1980. The presence of these six pronounced structural breaks poses a problem for the ensuing analysis. In dealing with monthly data it is necessary to investigate the possibility of unit roots at seasonal frequencies prior to testing for cointegrating and causality, and some important contributions in this area include Ghysels (1990), Ghysels and Perron (1993), Engle *et al.* (1993), and various contributions in Hylleberg (1992). However, the theoretical econometrics literature has not yet dealt fully with the problem of integration/cointegration/causality testing for monthly data when structural breaks are present.

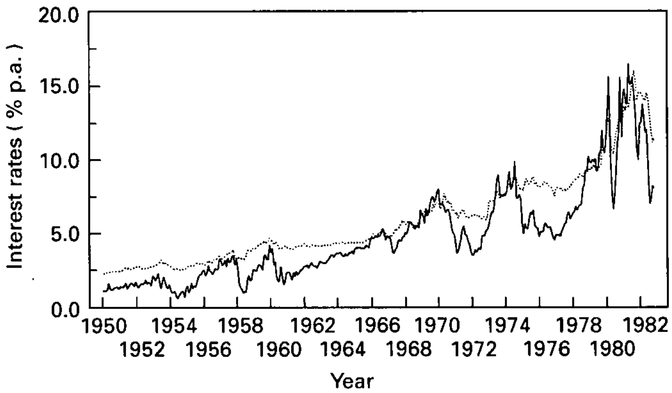

Fig. 1. *Short term (———) and long term (·········) interest rates (February 1950 to December 1982)*

One possible way of overcoming this problem would be to divide the data into seven sub-samples, defined by the breaks, and analyse these periods separately. However this would result in a substantial reduction in the power of the various tests (which may lead to unreliable results). Given that cointegration is a long run phenomenon, and that we want to investigate possible causal relationships between the interest rate series, a more desirable option is to annualize each series. This is the approach we take in this study.

For any month of the year an annual short rate and long rate series can be formed from the year-to-year observations in that particular month. This may appear to reduce the informational content of the data. However, in the context of testing for unit roots, this is not actually the case. Shiller and Perron (1985) show that in testing for a unit root the power of the Dickey–Fuller (1981) 't-test' is determined primarily by the span of the data, rather than by the number of observations. They even show that the power of the test can be destroyed by having too high a frequency of observations over a given span.[2]

[2]Also, see Choi (1992).

Effectively, we use all observations on both interest rate variables, but in the form of twelve separate series for each variable. We find that the short rate series is more variable than the long rate series over the sample period. Another distinctive feature which is exhibited by both series, and in all months, is the marked structural break occurring around 1977. Although it is difficult to establish the exact form of the break from these graphs, it appears as though it could possibly be a change in the level of the series combined with a change in the rate of growth. This is extremely important in terms of the appropriate methodology to adopt when testing for causality, as discussed in depth in Section IV. It must be noted that by annualizing the series we are sidestepping the whole problem of checking for unit roots at seasonal frequencies. Thus, although we still have a structural break to account for, there are techniques available to overcome this when testing for the usual (zero frequency) order of integration (Perron, 1989, 1990).

IV. TESTING METHODOLOGY

The testing procedures in this study incorporate some novel features that have been developed to examine the effects of the structural break while investigating the possibility of causality between short and long term interest rates. In addition, the data have been analysed in three different ways:

(a) *Full sample period (1950–82) – naive approach*: Here, the tests are applied over the full sample period, with only partial account being taken for the effects of the structural break. This provides a reference point for gaining some insight into the importance of correctly accommodating the break in the data into the analysis.
(b) *Sub-sample (1950–76)*: This approach involves testing over the sample period leading up to, but not including, the structural break,[3] with the disadvantage that some relevant information is ignored. To conserve space we do not report the results of this part of the study, but full details are given in Mandeno and Giles (1994).
(c) *Full sample period (1950–82) – refined analysis*: This involves testing over the full sample period with a rigorous allowance being made for the structural break. The novel aspects of this analysis are detailed below, and the conclusions reached here should be of primary interest.

The following strategy is adopted twelve times – for the two annualized series recorded in each month – in each of cases (a) to (c) above: we determine the order of integration of the variables; test for cointegration between the variables (where applicable); estimate a VAR model; and apply tests for causality. This analysis must be conducted in this order to ensure that the VAR model is properly specified and the tests are valid. The particular unit root testing procedure we have adopted is best described first in the context of cases (a) and (b) above. Each series is tested for order of integration using the augmented Dickey–Fuller (ADF) 't-test', with a nominal significance level of 10% to ensure reasonable power. For the same reason, we follow Dickey and Pantula (1987) and test from high to low order of integration, allowing the highest feasible order of integration for any series to be I(3). That is, we test H_0: I(3) versus H_a: I(2), H_0: I(2) versus H_a: I(1), etc., as appropriate. At each stage we use the strategy suggested by Dolado *et al.* (1990) to determine the inclusion of drift and/or trend terms in the Dickey–Fuller (DF) regressions:

$$\Delta^k y_t = \alpha + \beta t + \gamma \Delta^{k-1} y_{t-1} + \sum_{j=1}^{p} \beta_j \Delta^k y_{t-j} + \varepsilon_t \qquad (2)$$

The level of augmentation (p) is chosen to ensure that the residuals from Equation 2 are white noise on the basis of the associated autocorrelogram–this procedure is used by Giles *et al.* (1992) and Giles (1994), for example, and has been shown by Dods and Giles (1995) to result in low size-distortion for the ADF test in samples of our size.

In case (c), where the presence of the structural break is taken into account formally, the approach we adopt essentially mimics that of Perron (1989, 1990). He employs the idea of 'intervention' analysis, as suggested by Box and Tiao (1975). Their methodology showed that 'aberrant' or 'outlying' events could be separated from the noise function and be modelled as 'interventions' in the deterministic part of the general time series model. Perron shows that failing to properly accommodate a structural break when testing for the order of integration will result in the integration testing being biased towards non-rejection of the null hypothesis of a unit root. Here, we treat the structural break in 1977 as an exogenous event or an 'intervention', but as the exact form of the break is not clear, for each series we allow for the most general possible case of an exogenous change in the level of the series combined with an exogenous change in the rate of growth. Rather than use Perron's critical values for the (modified) ADF test, which are valid only asymptotically and for very stylized forms of structural change in the data, we generate critical values which are personalized to our sample size and to the particular form of break in each series by using bootstrap simulation. To summarize this process, first we take each series y_t and estimate via OLS:

$$y_t = \mu_1 + y_{t-1} + \beta DTB_t + (\mu_2 - \mu_1) DU_t + e_t \qquad (3)$$

where, DTB is one if $t = 1977$, and zero otherwise; and DU_t is one if $t \geqslant 1977$ and zero otherwise.

In Equation 3, the term βDTB_t allows for a one time change in the level of the series in 1977, while the term

[3]There are insufficient data points to facilitate a corresponding analysis of the post-1976 sub-sample.

$(\mu_2 - \mu_1)DU_t$ accommodates a change in the rate of growth after 1976.

Then we use the estimated coefficients for μ_1, β and μ_2, to generate the series y_t^*:

$$y_t^* = \hat{\mu}_1 + y_{t-1}^* + \hat{\beta}DTB_t + (\hat{\mu}_2 - \hat{\mu}_1)DU_t + e_t^* \quad (4)$$

where e^* is the series formed by randomly sampling with replacement from the residuals generated in equation 3, and multiplying[4] by a scale factor of $(T/(T-k))$. The advantage of sampling with replacement from the residual series in Equation 3 is that we generate a series with the same distribution. This is more desirable than imposing an exogenous assumption on the underlying distribution of the residuals.

The y_t^* series generated in this way has the null hypothesis of a unit root imposed. The next step is to estimate the DF regression:

$$\Delta y_t^* = \gamma y_{t-1}^* + v_t \quad (5)$$

By repeating this process 10 000 times we generate a distribution for the 't-ratios' on γ, and hence the exact 10% critical value. This process is applied to the individual short and long rate series in each month of the year. Hence the critical values generated in each case are tailored to the exact form and magnitude of the structural breaks in each of the 24 separate series. Then, to test the hypothesis $H_0: y_t \sim I(1)$ versus $H_a: y_t \sim I(0)$ for each series, we simply estimate the DF regression:

$$\Delta y_t = \gamma y_{t-1} + v_t \quad (6)$$

Comparing the t-ratio on γ (t_γ) against the bootstrapped critical value completes the testing procedure.

If the short and long rate series in a particular month are integrated of the same order, we must investigate the possibility of a cointegrating relationship between them. This is straightforward in cases (a) and (b) above. Due to the reasonably small sample, we estimate both the forward *and* reverse cointegrating regressions (with and without linear trend terms), although asymptotically it will make no difference which way these regressions are normalized. The no-drift/trend ADF test (with p determined as above) is then applied to the residuals of the cointegrating regressions, with the 10% critical values again being taken from MacKinnon (1991).

Testing for cointegration in case (c), where allowance is made for the structural break in the data, is non-standard. We again begin by estimating the cointegrating regressions, then to assess the combined effect of the structural breaks in both the short and long rates on the associated residuals, we plot the latter and check for a significant break in 1977. If there is a break in these series, this must be allowed for by adjusting the critical value when testing if the residuals are I(0). The most general form of break would again be a change in the level of the series combined with a change in the rate of growth. Further, as each residuals series is just a linear combination of the two interest rates series, which both display breaks in 1977, any break in the residuals also must occur at this time. Hence, we can simply apply the same bootstrap simulation procedure as above to the cointegrating regression residuals (rather than the original data) to obtain modified critical values for the ADF cointegration test which take account of the precise forms of structural breaks in our data. If two time series are both I(1) and are cointegrated, then there must exist either uni-directional or bi-directional Granger causality between them (Engle and Granger, 1987). In principle, this provides a cross-check on our subsequent analysis, although the low power of the ADF test in small samples may produce an ambiguous result.

We use the usual methodology in testing for Granger causality, in which a VAR model is estimated and a test for the joint significance of a particular subset of coefficients is applied. The first important consideration is the formulation of an *appropriate* VAR model between the short rate (x) and the long rate (y). The most general form of the VAR model used here for month h is given by[5]:

$$\Delta x_{t,h} = \alpha_1 + (\alpha_2 D1) + (\alpha_3 tr) + (\alpha_4 tD2) + \sum_{i=1}^{m} \beta_i \Delta x_{t-i,h} + \sum_{j=1}^{n} \gamma_j \Delta y_{t-j,h} + (\pi z_{t-1}) + u_t \quad (7)$$

$$\Delta y_{t,h} = a_1 + (a_2 D1) + (a_3 tr) + (a_4 tD2) + \sum_{i=1}^{q} b_i \Delta y_{t-i,h} + \sum_{j=1}^{r} c_j \Delta x_{t-j,h} + (p z_{t-1}) + v_t \quad (8)$$

where,

$$D1 = 0; \quad t = 1950, \ldots, 1976$$
$$\phantom{D1 = {}} 1; \quad t = 1977, \ldots, 1982$$
$$tD2 = 0; \quad t = 1950, \ldots, 1976$$
$$\phantom{tD2 = {}} tr; \quad t = 1977, \ldots, 1982$$
$$tr = \text{linear time trend},$$
$$z_{t-1} = \text{error correction term (see below).}$$

The dummy variables are included in the model to account for the structural break. The model is refined by individually removing the regressors $D1$, tr and $tD2$ if they are found to be insignificant on the basis of two-sided 10% standard Normal tests (to allow for the presence of lagged dependent variables in Equations 7 and 8).

Because at least one of the series is found to be I(1) in virtually all months (see Section V), all VAR models are estimated in first difference form. As a precaution, to ensure there is no evidence of over-differencing, or other general model mis-specification, all of the VAR residuals are

[4]This factor adjusts the variance of e^* to allow for the loss of k degrees of freedom in estimating Equation 3.

[5]Terms in parentheses are not included in the VAR model for every month (see below).

checked for autocorrelation at various orders using the Lagrange multiplier (LM) statistic. One important question is whether or not to include an 'error correction term' in the models. The Granger Representation Theorem (Engle and Granger, 1987) implies that if x and y are I(1) and cointegrated then we can either construct the VAR in terms of the levels of the data, or construct the VAR in terms of the first differences of the variables but also include an 'error correction term' (the lagged residuals from the appropriate cointegrating regression) in each equation. The *appropriate* cointegrating regression is the forward regression (i.e., x is the dependent variable) in the case of the forward VAR Equation 7 and vice versa for the reverse Equation 8. A trend is included in the cointegrating regression only where it is found to be significant. In the results below, the error correction residuals series are labelled z_1 (z_2) for the forward (backward) no-trend case, and z_3 (z_4) for the forward (backward) trend case respectively.

The selection of appropriate lag lengths in the VAR model is crucial. A trade-off exists, as under-parameterization will bias the results and over-parameterization will reduce the power of the tests (McFadyen *et al.*, 1991). Many methods (e.g., Akaike, 1979; Sims, 1980) have been suggested for determining these lag lengths. We follow the procedure suggested by Hsiao (1979, 1981) as it provides, as a by-product, a direct cross-check on the outcome of the ensuing causality tests. This procedure involves minimizing Akaike's 'final prediction error' (fpe) for each VAR equation. To illustrate, in Equation 7 we must determine two optimal lag lengths, m^* and n^*. These are found by a two-step procedure in which we first set $n = 0$ and find the value $m = m^*$ that minimizes the fpe; and then fix $m = m^*$ and vary n to find the value, n^*, which minimizes fpe (m^*, n). (Note that if fpe (m^*) > fpe(m^*, n^*), this suggests informally that y 'Granger-causes' x.)

Each equation of the VAR model can be estimated individually using OLS, or, if the model is treated as a system of 'seemingly unrelated regression equations' (SURE), there may be efficiency gains if full maximum likelihood (ML) estimation is used. To see which estimator is appropriate, we test the contemporaneous error covariance matrix for diagonality using the standard asymptotic likelihood ratio (LR) and Breusch-Pagan Lagrange multiplier (BPLM) tests. Once we have estimated the appropriate VAR model in each month the next step is to test for causality. To establish whether y Granger causes x we test the hypothesis H_0: '$\gamma_1 = \gamma_2 = \ldots = \gamma_n = (\pi) = 0$' versus H_a: 'not H_0'. Similarly to determine whether x Granger causes y we test H_0: '$c_1 = c_2 = \ldots = c_r = (p) = 0$' versus H_a: 'not H_0'.

Lütkepohl and Reimers (1992) use the results of Toda and Phillips (1993) to show that for *bivariate* non-stationary cointegrated systems the standard Wald test statistic for linear restrictions will still be asymptotically χ^2 under the null. Furthermore, with regard to finite sample properties, Geweke *et al.* (1983) compared three alternative tests for causality and concluded that Wald variants of the basic Granger test were superior to the other tests studied. For these reasons the Wald test is applied in this study. A variant of the Wald test suggested by Sims (1980) involving a degree of freedom correction is also employed. As a further refinement to the testing procedure outlined above, a bootstrap simulation experiment (with 5000 replications) was conducted to eliminate any size distortion in the Wald test for our particular models. As we are analysing a relatively small sample, there is the possibility of significant deviations of the true size of the test from its nominal value. This experiment yielded exact 10% critical values for the Wald causality tests for each of cases (a), (b) and (c) (with allowances for structural breaks and error correction terms, as appropriate), and each month's models.

V. RESULTS

All of the following results were obtained using the SHAZAM (1993) package. Table 1a and b presents the results from the ADF testing for the short and long term rates respectively when the full sample is used. In each case, 12 annual interest rate series are tested, all spanning the same period, but with each series measured in a different month.[6] We see that in eight months the short rate series is stationary, and I(1) in the remaining four. On the other hand the long rate is I(1) in each month of the year. In interpreting these tables, attention should be paid to the information provided there in relation to the inclusion of drift and/or trend terms in the DF regressions, as this affects the appropriate critical values.

Table 2a and b presents the results (in more abbreviated form) for case (c). Here we find that the short rate series are I(0) in all months except November, December and January, while the long rate series are I(1) in all months except September, October and November. These results exhibit *runs* of months with the same outcome, rather than a random scattering of different outcomes, suggesting a localized effect which is not apparent in the remainder of the year.

From Table 1a and b, the short and long rates are both I(1) in February through to May, so the existence of a cointegrating relationship is possible and is tested for only in these months. The results in Table 3 show that there is strong evidence of cointegration in April and May with weaker evidence in February and March. Wallace and Warner (1993) also provide strong evidence of such cointegration. We have conservatively assumed that there exists cointegration in all four months, and included

[6] The January 1950 observation was not available for either the short and long rates, hence the slightly smaller sample size in that month.

Table 1. *ADF tests for order of integration (full sample naive approach a)*

		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
a. *Short term interest rates*													
Notes	T	32	33	33	33	33	33	33	33	33	33	33	33
$H_0: I(3)$	$p^{\#}$	3	2	2	3	2	2	2	3	0	0	0	0
versus	t	− 5.30[a]	− 5.55[a]	− 5.06[a]	− 3.91[a]	− 5.90[a]	− 5.32[a]	− 5.47[a]	− 5.84[a]	− 7.76[a]	− 6.68[a]	− 7.43[a]	− 8.73[a]
$H_a: I(2)$		Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej
$H_0: I(2)$	$p^{\#}$	2	0	1	0	0	0	0	0	2	0	2	0
versus	t	− 5.87[a]	− 4.56[a]	− 5.21[a]	− 5.25[a]	− 5.99[a]	− 9.67[a]	− 9.29[a]	− 9.04[a]	− 5.47[a]	− 5.25[a]	− 4.85[a]	− 5.08[a]
$H_a: I(1)$		Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej
$H_0: I(1)$	$p^{\#}$	1	0	2	0	0	0	0	0	0	0	1	0
versus	t	− 4.42[a]	0.91[c]	1.43[c]	0.80[c]	0.68[c]	− 3.99[a]	− 4.27[a]	− 5.15[a]	− 5.10[a]	− 3.94[a]	− 3.50[a]	− 3.34[a]
$H_a: I(0)$		Rej	Acc	Acc	Acc	Acc	Rej	Rej	Rej	Rej	Rej	Rej	Rej
Conclude		I(0)	I(1)	I(1)	I(1)	I(1)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)
b. *Long term interest rates*													
Notes	T	32	33	33	33	33	33	33	33	33	33	33	33
$H_0: I(3)$	$p^{\#}$	0	0	0	2	0	2	1	1	0	0	0	0
versus	t	− 7.96[a]	− 8.20[a]	− 7.84[a]	− 5.70[a]	− 6.12[a]	− 5.11[a]	− 7.04[a]	− 6.33[a]	− 4.53[a]	− 3.72[a]	− 3.62[a]	− 4.32[a]
$H_a: I(2)$		Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej
$H_0: I(2)$	$p^{\#}$	0	0	0	0	0	0	0	0	0	0	0	0
versus	t	− 3.72[a]	− 3.95[a]	− 4.62[a]	− 5.09[a]	− 4.91[a]	− 5.02[a]	− 5.51[a]	− 5.58[a]	− 5.09[a]	− 4.92[a]	− 4.05[a]	− 3.50[a]
$H_a: I(1)$		Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej
$H_0: I(1)$	$p^{\#}$	0	0	0	3	0	0	0	0	0	0	0	0
versus	t	3.15[b]	3.43[b]	2.69[b]	3.16[b]	1.63[b]	1.85[b]	1.75[b]	2.70[c]	1.56[c]	0.94[c]	0.98[c]	1.54[c]
$H_a: I(0)$		Acc	Acc	Acc	Acc	Acc	Acc	Acc	Acc	Acc	Acc	Acc	Acc
Conclude		I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)

[a]DF '*t*-test' applied in drift/trend DF regression.
[b]DF '*t*-test' applied in drift/no-trend DF regression.
[c]DF '*t*-test' applied in no-drift/no-trend DF regression.
$^{\#}p$ is the augmentation level in the particular form of DF regression (with respect to drift/trend) from which the *t*-statistic is obtained.

Table 2. *Order of integration tests (full sample refined approach c)*

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
a. *Short term interest rates*												
t_y	− 4.17	− 4.50	− 4.51	− 4.71	− 5.22	− 8.22	− 7.72	− 7.12	− 6.01	− 4.93	− 4.09	− 3.80
10% crit[a]	− 4.24	− 4.13	− 4.06	− 4.10	− 4.08	− 4.11	− 4.09	− 4.19	− 4.15	− 4.17	− 4.11	− 4.13
Conclude	I(1)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	I(0)	I(1)	I(1)
b. *Long term interest rates*												
t_y	− 3.07	− 3.10	− 3.67	− 3.54	− 3.58	− 3.39	− 3.88	− 4.07	− 4.79	− 5.07	− 4.52	− 3.92
10% crit[a]	− 4.10	− 4.09	− 4.13	− 4.12	− 4.14	− 4.11	− 4.11	− 4.13	− 4.20	− 4.19	− 4.18	− 4.20
Conclude	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(1)	I(0)	I(0)	I(0)	I(1)

[a]These 10% critical values are personalized to each series and are generated using the bootstrap simulation techniques described in Section IV.

an error correction term in the VARs even if it is not necessary. This is better than risking erroneously excluding a necessary error correction term. Under-specifying the model will result in inconsistent parameter estimates, while over-specifying the model will result in parameter estimates which are consistent (though biased here due to the presence of lagged values of the dependent variable as regressors).

Table 3. *Cointegrating regression ADF tests (full sample naive approach a)*

T		Feb 33	Mar 33	Apr 33	May 33
z_1	p	1	1	1	0
	t	-3.88^a	-4.13^a	-4.46^a	-3.98^a
z_2	p	0	0	1	0
	t_y	-2.98	-3.11	-3.80^a	-3.52^a
z_3	p	0	0	1	0
	t_{trend}	-0.92	-1.52	-1.88^b	-1.10
	t_y	-3.50	-3.88^a	-4.54^a	-3.94^a
z_4	p	0	0	0	2
	t_{trend}	5.31^b	5.49^b	5.81^b	5.23^b
	t_y	-2.00	-2.27	-2.14	-2.03

Notes: [a]Reject $H_0: \gamma = 0$ using MacKinnon (1991) 10% critical value: z is I(0)
[b]Time trend significant using 2-sided standard normal 10% critical value.

Table 4. *Cointegrating regression ADF tests (full sample refined approach c)*

T		Jan 32	Dec 33
z_1	t_y	-3.77	-4.39^a
	crit[b]	-4.31	-4.10
z_2	t_y	-3.67	-4.46^a
	crit[b]	-4.33	-4.13
z_3	t_{trend}	0.81	-0.14
	t_y	-3.85	-4.40^a
	crit[b]	-4.32	-4.15
z_4	t_{trend}	4.14^c	5.23^c
	t_y	-3.17	-4.65^a
	crit[b]	-4.29	-4.16

Notes: [a]Reject H_0: $\gamma = 0$ using bootstrap 10% critical value: z is I(0).
[b]10% personalized critical value generated through bootstrap experiment.
[c]Time trend significant using 2-sided standard normal 10% critical value.

From Table 2a and b, January and December are the only months in which both series are I(1) in case (c), where the full sample is used with account being taken of the structural break. In this case we begin by plotting the residual series z_1 to z_4 in these months and checking if a significant break occurs in 1977. These graphs revealed no obvious breaks, but to be conservative we have allowed for a possible break in 1977 when testing for cointegration. The associated results appear in Table 4, and they show that the short and long rate series *are not* cointegrated in January (as z_1 to z_4 are all I(1)) but *are* cointegrated in December (as z_1 to z_4 are all I(0)).

These various stationarity properties of the data determine the appropriate formulation of the VAR models. Table 5a and b summarizes the relevant results relating to the forward (Equation 7) and reverse (Equation 8) VAR models over the full sample period. Error correction terms are included only in the months February to May. Table 5c presents the results relating to the choice of estimation method. In all months except April, the contemporaneous error covariance matrix displayed significant non-diagonality, so the equations were estimated as a system by maximum likelihood in all months (even though this was not strictly necessary in April).

Table 5d and e summarizes the outcomes from the fpe, Wald and Sims' tests. The results from these three tests form the basis for the conclusions regarding causality at the bottom of each section of the table. In all months except January, the outcomes from all three tests coincide. In January, however, there is some conflict in the test outcomes. To signify this, parentheses are placed around the conclusion reported. (This convention is used in all of the following tables.) From Table 5d we see that there is very strong evidence of causality from long-to-short interest rates, with the null hypothesis of no causality rejected in ten out of the twelve months. Similarly from Table 5e there is also very strong evidence supporting a short-to-long causal relationship, this being supported in ten out of the twelve months. Therefore, we conclude that this naive analysis over the full sample period strongly supports the existence of bi-directional causality between long and short term interest rates.

Table 6a–e presents the results obtained using the preferred 'refined' approach, in which the full sample is used but with a proper allowance for the observed structural break in 1977. As the methodology employed here is identical to full-sample approach (a), the only VAR models that required re-estimating were those in which error correction terms were either removed (February–May) or inserted (December) as a result of the refined integration/cointegration analysis. In the forward VAR regressions the conclusions are the same as when using the naive approach. However, in the reverse regression the refined approach results in different conclusions for the months February through to May, with the null hypothesis of no causality no longer being rejected. The net outcome is ten out of twelve rejections of the null in the forward direction and only six out of twelve rejections in the reverse direction. In comparison with the sub-sample analysis, the refined analysis shows much stronger evidence of causality in both directions. This result can probably be attributed to the increased power of the tests to reject the null when applied over the larger sample period.

Table 5. *Full sample naive approach a*

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
a: *VAR construction – forward regression*												
z_{t-1}	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
$D1$	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
tr	No	No	Yes	No	Yes	No	No	Yes	No	Yes	No	No
$tD2$	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No
m^*	3	1	5	5	4	4	4	3	6	1	1	5
fpe(m^*)	1.72	1.99	1.89	1.26	2.50	2.85	2.72	3.21	4.30	3.30	2.26	3.04
n^*	2	3	1	4	5	2	3	3	2	5	4	4
fpe(m^*, n^*)	1.66	0.97	1.13	0.77	1.49	1.86	1.59	1.47	2.76	1.33	1.47	2.99
b: *VAR construction – reverse regression*												
z_{t-1}	No	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No
$D1$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
tr	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	Yes
$tD2$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
q^*	1	2	2	3	2	3	2	3	6	5	5	5
fpe(q^*)	0.31	0.24	0.28	0.33	0.47	0.42	0.45	0.55	0.96	1.08	1.08	0.52
r^*	1	5	5	7	6	1	2	1	2	1	8	1
fpe(q^*, r^*)	0.29	0.19	0.21	0.22	0.38	0.35	0.41	0.41	0.62	1.13	0.65	0.39
c: *Estimation method*[a]												
BPLM	7.50	3.01	8.42	1.25	7.67	8.37	8.76	12.05	10.58	16.64	8.98	12.79
LR	13.57	4.75	11.26	1.44	19.11	11.72	12.78	16.16	25.58	25.86	24.74	32.22
Est.	ML	ML	ML	OLS[b]	ML	ML	ML	ML	ML	ML	ML	ML
d: *Causality tests – forward regression*												
fpe(m^*, n^*) > fpe(m^*)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wald	3.03	44.15[c]	12.58[c]	19.37[c]	23.91[c]	15.26[c]	21.14[c]	40.00[c]	18.68[c]	54.27[c]	24.47[c]	1.47
χ^2_h 10%[d]	4.61	7.78	4.61	9.24	10.64	4.61	6.25	6.25	4.61	9.24	7.78	7.78
True size	14.1	14.9	15.3	21.1	21.2	13.5	14.0	14.9	19.7	15.1	14.4	19.4
Prob	26.1	0.0	1.3	4.0	2.8	0.2	0.3	0.0	0.8	0.0	0.3	84.2
Sims[e]	25.22[f]	21.59[f]	9.05[f]	10.78[f]	12.24[f]	11.24[f]	13.77[f]	19.98[f]	9.35[f]	25.64[f]	16.67[f]	1.26
Concl.	(Acc)	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Acc
e: *Causality tests – reverse regression*												
fpe(m^*, n^*) > fpe(m^*)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Wald	3.02[c]	21.42[c]	27.15[c]	32.22[c]	27.69[c]	5.88[c]	6.56[c]	9.19[c]	17.29[c]	0.29	12.17	11.94[c]
χ^2_h 10%[d]	2.71	10.64	10.64	13.36	12.02	2.71	4.61	2.71	4.61	2.71	14.68	4.61
True size	12.3	17.9	17.8	21.3	19.0	11.1	12.8	12.4	17.0	11.9	17.8	5.0
Prob	9.5	2.2	1.1	2.1	1.5	2.9	5.4	0.8	1.1	58.7	26.0	0.3
Sims[e]	2.38	23.26[f]	15.30[f]	14.51[f]	13.51[f]	25.73[f]	5.40[f]	7.56[f]	7.88[f]	0.29	8.41	8.32[f]
Concl.	(Rej)	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Rej	Acc	Acc	Rej

[a]The *BPLM* and *LR* statistics are compared against a χ^2_1, critical value.
[b]*ML* also applied here, even though it is not strictly necessary.
[c]Wald Test Statistic exceeds χ^2_h 10% critical value.
[d]h = number of restrictions; e.g. in forward VAR, $h = n^*$.
[e]Sims' test statistic compared against χ^2_h critical value.
[f]Sims' test statistic exceeds χ^2_h 10% critical value.

Table 6. *Full sample refined approach c*

	Feb	Mar	Apr	May	Dec
z_{t-1}	No	No	No	No	Yes
$D1$	Yes	Yes	Yes	Yes	No
tr	No	No	Yes	Yes	Yes
$tD2$	Yes	Yes	Yes	Yes	Yes
m^*	5	5	5	4	6
fpe(m^*)	2.04	1.73	1.33	2.41	2.87
n^*	3	4	4	5	1
fpe(m^*, n^*)	1.34	1.35	0.92	1.68	2.85
b: *VAR construction – reverse regression*					
z_{t-1}	No	No	No	No	Yes
$D1$	Yes	Yes	Yes	Yes	Yes
tr	Yes	No	Yes	Yes	Yes
$tD2$	Yes	Yes	Yes	Yes	Yes
q^*	3	3	3	3	2
fpe(q^*)	0.23	0.29	0.31	0.48	0.33
r^*	1	5	5	2	6
fpe(q^*, r^*)	0.25	0.30	0.34	0.51	0.35
c: *Estimation method*[a]					
BPLM	10.75	12.42	6.19	10.93	7.06
LR	16.86	17.90	9.53	20.78	11.02
Est.	ML	ML	ML	ML	ML
d: *Causality tests – forward regression*					
fpe(m^*, n^*) > fpe(m^*)	Yes	Yes	Yes	Yes	No
Wald	23.00[d]	12.35[d]	13.55[d]	21.04[d]	3.08
χ_h^2 10%[b]	6.25	7.78	7.78	9.24	4.61
True size	17.0	19.7	20.5	19.6	18.8
Prob	0.5	7.5	6.6	2.2	29.1
Sims[c]	11.97[e]	9.15[e]	9.72[e]	12.41[e]	3.72
Concl.	Rej	Rej	Rej	Rej	Acc
e: *Causality tests – reverse regression*					
fpe(m^*, n^*) > fpe(m^*)	No	No	No	No	No
Wald	0.13	11.32[d]	9.72	1.76	55.31[d]
χ_h^2 10%[b]	2.71	9.24	9.24	4.61	12.02
True size	11.3	17.3	19.2	13.6	18.1
Prob	72.6	9.3	22.0	43.9	0.1
Sims[c]	0.13	8.50	6.63	1.67	22.19
Concl.	Acc	(Acc)	Acc	Acc	Rej

[a]The BPLM and LR statistics are compared against a χ_1^2, critical value.
[b]h = number of restrictions: e.g. in forward VAR, $h = n^*$.
[c]Sims' test statistic compared against χ_h^2 critical value.
[d]Wald test statistic exceeds χ_h^2 10% critical value.
[e]Sims' test statistic exceeds χ_h^2 10% critical value.

VI. CONCLUSIONS

The three approaches that we have taken in this study yield somewhat differing results regarding cointegrating and causal relationships between US Treasury short and long term interest rates. Of these three approaches, our 'refined' analysis is the most comprehensive, and so the results attained through this method form the basis of our conclusions regarding causality. That is, there is *strong evidence* of a causal relationship from long term to short term interest rates; and there is *some evidence* of causality in the reverse direction, so it is possible that a feedback relationship exists. These conclusions accord with those reached by Lütkepohl and Reimers (1992) although we have employed a more refined approach here. Consistent with other work (e.g., Froot, 1989), but in contrast to many earlier findings (e.g., Campbell and Shiller, 1984), these results support the expectations theory of the term structure of interest rates. Another notable feature of our results relates to the different outcomes when using different months as base-points for annual measurement. Interestingly, we find *runs* of months with the same outcome, suggesting a localized effect associated with consecutive months. From a policy viewpoint, this deserves further attention.

In comparing the outcomes of our three approaches, the differences between the sub-sample and refined full-sample results are especially interesting. Clearly, the post-break observations in these series contain vital information which has totally reshaped the conclusions reached. This highlights the importance of using more sophisticated analysis to explicitly incorporate a structural break rather than simply using a pre/post-break sub-sample.

ACKNOWLEDGEMENTS

We are grateful to Judith Giles, Doug Pearce and Mike Veall for helpful discussions in the course of undertaking this work, and to the referee for constructive comments and suggestions on an earlier draft.

REFERENCES

Akaike, H. (1979) A Bayesian extension of the minimum AIC procedure of autoregressive model fitting, *Biometrika*, **66**, 237–42.

Box, G. E. P. and Tiao, G. C. (1975) Intervention analysis with applications to economic and environmental problems, *Journal of the American Statistical Association*, **70**, 70–79.

Campbell, J. Y. and Shiller, R. J. (1984) A simple account of the behaviour of long term interest rates, *American Economic Review* (Papers and Proceedings), **74**, 44–48.

Campbell, J. Y. and Shiller, R. J. (1987) Cointegration and tests of present value models, *Quarterly Journal of Economics*, **95**, 1062–88.

Choi, I. (1992) Effects of data aggregation on the power of tests for a unit root, *Economics Letters*, **40**, 397–401.

Dickey, D. A. and Fuller, W. A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, **49**, 1057–72.

Dickey, D. A. and Pantula, S. G. (1987) Determining the order of differencing in autoregressive processes, *Journal of Business and Economic Statistics*, **5**, 455–61.

Dods, J. and Giles, D. E. A. (1995) Alternative strategies for 'augmenting' the Dickey–Fuller test: size-robustness in the face of pre-testing, *Journal of Statistical Computation and Simulation*, forthcoming.

Dolado, J. J., Jenkinson, T. and Sosvilla-Rivero, S. (1990) Cointegration and unit roots, *Journal of Economic Surveys*, **4**, 249–73.

Engle, R. F. and Granger, C. W. J. (1987) Cointegration and error correction: representation, estimation and testing, *Econometrica*, **55**, 251–76.

Engle, R. F., Granger, C. W. J., Hylleberg, S. and Lee, H. S. (1993) Seasonal cointegration – The Japanese consumption function, *Journal of Econometrics*, **55**, 275–98.

Fama, E. F. (1984) Premiums in bond returns, *Journal of Financial Economics*, **13**, 529–46.

Fama, E. F. and Bliss, R. R. (1987) The information in long-maturity forward rates, *American Economic Review*, **77**, 680–92.

Froot, K. (1989) New hope for the expectations hypothesis of the term structure of interest rates, *Journal of Finance*, **44**, 283–305.

Geweke, J., Meese, R. and Dent, W. (1983) Comparing alternative tests of causality in temporal systems–analytic results and experimental evidence, *Journal of Econometrics*, **21**, 161–94.

Ghysels, E. (1990) Unit root tests and the statistical pitfalls of seasonal adjustment: the case of US postwar real gross national product, *Journal of Business and Economic Statistics*, **8**, 145–52.

Ghysels, E. and Perron, P. (1993) The effect of seasonal adjustment filters on tests for a unit root, *Journal of Econometrics*, **55**, 57–98.

Giles, J. A. (1994) Another look at the evidence on foreign aid led economic growth, *Applied Economics Letters*, **1**, 194–99.

Giles, D. E. A., Giles, J. A. and McCann, E. (1992) Causality, unit roots and export-led growth: the New Zealand experience, *Journal of International Trade and Economic Development*, **1**, 195–218.

Granger, C. W. J. (1969) Investigating causal relations by econometric models and cross spectral methods, *Econometrica*, **37**, 424–38.

Granger, C. W. J. (1981) Some properties of time series data and their use in econometric model specification, *Journal of Econometrics*, **28**, 121–30.

Guilkey, D. K. and Salemi, M. K. (1982) Small sample properties of three tests for Granger-causal ordering in a bivariate stochastic system, *Review of Economics and Statistics*, **64**, 668–80.

Hamilton, J. D. (1988) Rational expectations econometric analysis of changes in regime – an investigation of the term structure of interest rates, *Journal of Economic Dynamics and Control*, **12**, 385–423.

Hsiao, C. (1979) Autoregressive modelling of Canadian money and income data, *Journal of the American Statistical Association*, **74**, 553–60.

Hsiao, C. (1981) Autoregressive modelling and money-income causality detection, *Journal of Monetary Economics*, **7**, 85–106.

Hylleberg, S. (ed.) (1992) *Modelling Seasonality*, Oxford University Press, Oxford.

Karfakis, C. I. and Moschos, D. M. (1993) The information content of the yield curve in Australia, *Working Paper* No. 185, Department of Economics, University of Sydney.

Krol, R. (1987) The term structure of eurodollar interest rates and its relationship to the US treasury-bill market, *Journal of International Money and Finance*, **6**, 339–54.

Lütkepohl, H. and Reimers, H. (1992) Granger-causality in cointegrated VAR processes-the case of the term structure, *Economic Letters*, **40**, 263–68.

Macaulay, F. R. (1938) *Some Theoretical Problems Suggested by Movements of Interest Rates, Bond Yields and Stock Prices in the United States*, National Bureau of Economic Research, New York.

MacDonald, R. and Speight, A. (1988) The term structure of interest rates in the UK, *Bulletin of Economic Research*, **40**, 287–99.

MacDonald, R. and Speight, A. (1991) The term structure of interest rates under rational expectations: some international evidence, *Applied Financial Economics*, **1**, 211–21.

McFadyen, J., Pickerill, K. and Devaney, M. (1991) The expectations hypothesis of the term structure: more evidence, *Journal of Economics and Business*, **43**, 79–85.

MacKinnon, J. G. (1991) Critical values for cointegration tests, in R. F. Engle and C. W. J. Granger (eds.), *Long-Run Economic Relationships*, Oxford University Press, Oxford, pp. 267–76.

Mandeno, R. J. and Giles, D. E. A. (1994) The expectations theory of the term structure: a cointegration/causality analysis of US interest rates, *Discussion Paper* 93–10, Department of Economics, University of Victoria.

Mankiw, N. (1986) The term structure of interest rates revisited, *Brookings Papers on Economic Activity*, **1**, 223–47.

Margaritis, D. (1994) Time-varying risk premia in the term structure of interest rates in New Zealand, *Applied Financial Economics*, **4**, 111–20.

Meiselman, D. (1962) *The Term Structure of Interest Rates*, Prentice-Hall, Englewood Cliffs.

Moazzami, B. (1991) The Fisher equation controversy re-examined, *Applied Financial Economics*, **1**, 129–33.

Owen, P. D. (1993) Cointegration analysis of the Fisher hypothesis: the role of the real rate and the Fisher identity, *Applied Financial Economics*, **3**, 21–26.

Perron, P. (1989) The Great Crash, the oil price shock, and the unit root hypothesis, *Econometrica*, **57**, 1361–401.

Perron, P. (1990) Testing for a unit root in a time series with a changing mean, *Journal of Business and Economic Statistics*, **8**, 153–62.

SHAZAM (1993) *SHAZAM Econometrics Computer Program: User's Reference Manual, Version 7.0*, McGraw-Hill, New York.

Shea, G. S. (1992) Benchmarking the expectations hypothesis of the interest-rate term structure: an analysis of cointegration vectors, *Journal of Business and Economic Statistics*, **10**, 347–66.

Shiller, R. J. and Perron, P. (1985) Testing the random walk hypothesis - power vs. frequency of observation, *Economics Letters*, **18**, 381–86.

Sims, C. A. (1980) Macroeconomics and reality, *Econometrica*, **48**, 1–48.

Taylor, M. P. (1992) Modelling the yield curve, *Economic Journal*, **90**, 524–37.

Toda, H. Y. and Phillips, P. C. B. (1993) Vector autoregressions and causality, *Econometrica*, **61**, 1367–93.

Wallace, M. S. and Warner, J. T. (1993) The Fisher effect and the term structure of interest rates: tests of cointegration, *Review of Economics and Statistics*, **75**, 320–24.