

A COINTEGRATION ANALYSIS OF TREASURY BILL YIELDS

Anthony D. Hall, Heather M. Anderson, and Clive W. J. Granger*

Abstract—This paper shows that yields to maturity of U.S. Treasury bills are cointegrated, and that during periods when the Federal Reserve specifically targeted short-term interest rates, the spreads between yields of different maturity define the cointegrating vectors. This cointegrating relationship implies that a single non-stationary common factor underlies the time series behavior of each yield to maturity and that risk premia are stationary. An error correction model which uses spreads as the error correction terms is unstable over the Federal Reserve's policy regime changes, but a model using post 1982 data is stable and is shown to be useful for forecasting changes in yields.

I. Introduction

A topic which is discussed frequently in the term structure literature is that of the relationships between yields associated with bonds of different maturities. Arbitrage arguments, often augmented by considerations about risk are generally used to justify such relationships; the underlying problem is to explain the empirical observation that yields of different maturity appear to move together over time.

Formal empirical analysis of the relationships between yields of different maturities is not straightforward because nominal yields are not generally considered to be stochastically stationary. It has long been recognised that it is possible for sets of nonstationary variables to move together over time. Granger (1981) formalised this concept, defining such sets of variables as cointegrated variables, and since then various tests for cointegration and techniques for working with cointegrated variables have been developed.

The literature which relates cointegration to the theory of the term structure is currently small. A few authors have tested for (and found) cointegration between the yield on a long-term bond

and that on a short-term bond,¹ but the question of how one might further apply the theory of cointegration to study the term structure is largely unanswered. This study suggests that the term structure for U.S. Treasury bills is well modelled as a cointegrated system.

The organization of the paper is as follows. Section II relates the theory of cointegration, error correction models and common factors to well-known models of the term structure. Here, it is shown that if yields to maturity are integrated processes, then the term structure data are theoretically cointegrated. The cointegration expected here is of a special type, and the theoretical restrictions which should characterize this cointegration are derived and explored. Section III describes the data which have been used in this study. The empirical evidence that yields are cointegrated according to the predictions made in section II is presented in section IV. An estimated error correction model is presented to illustrate how this information can be utilised. The estimated model is statistically significant and is shown to be potentially useful for forecasting yields of Treasury bills. Section V concludes.

II. Theoretical Framework

A. Theory of the Term Structure

Let $R(k, t)$ be the continuously compounded yield to maturity of a k period pure discount bond, ($k = 1, 2, 3, \dots$) and let the forward rate $F(k, t)$ be the rate of return from contracting at time t to buy a one period pure discount bond which matures at time $t + k$. Then $F(1, t) = R(1, t)$, and forward rates can be recursively calculated from the Fisher-Hicks formulae,

$$R(k, t) = \frac{1}{k} \left[\sum_{j=1}^k F(j, t) \right]$$

$$\text{for } k = 1, 2, 3, \dots \quad (1)$$

¹ See, for instance, Campbell and Shiller (1987) or Engle and Granger (1987).

Received for publication July 23, 1990. Revision accepted for publication April 15, 1991.

* Australian National University, University of California, San Diego, and University of California, San Diego, respectively.

The authors wish to thank David Hendry and the referees for helpful comments on an earlier draft of this paper. Financial support for Hall from the Australian Research Council, Anderson from the P.E.O. International Peace Scholarship Fund, and for Granger from the National Science Foundation Grant SES 8902950 is gratefully acknowledged.

Forward rates $F(j, t)$ typically differ from the yield $R(1, t + j - 1)$ actually realised, so that investors may be assumed to rely on their expectations of $R(1, t + j - 1)$ when they choose between investing now or later. The relationship between forward and expected rates is assumed to be

$$F(j, t) = E_t[R(1, t + j - 1)] + \Lambda(j, t) \quad (2)$$

where E_t denotes expectations based on information available at time t and the $\Lambda(j, t)$ are premia, which may account for risk considerations or for investors' preferences about liquidity.

Substitution of equation (2) into equation (1) leads to a very general relationship between yields of different maturities, i.e.,

$$R(k, t) = \frac{1}{k} \left[\sum_{j=1}^k E_t[R(1, t + j - 1)] \right] + L(k, t),$$

$$\text{where } L(k, t) = \frac{1}{k} \left[\sum_{j=1}^k \Lambda(j, t) \right]. \quad (3)$$

This equation indicates that the yields of bonds with similar maturities will move together. Many of the traditional theories of the term structure focus on the properties of the premia $L(k, t)$. The pure expectations hypothesis asserts that the $L(k, t)$ are zero, while other versions of the expectations hypothesis assert that the premia are constant over time. Other assumptions about the premia would lead to different theories about the term structure, many of which are consistent with the framework described here.

Despite its simplicity, equation (3) does not provide an immediately useful basis for empirical studies of the term structure. None of the variables on the right hand side of this equation are directly measurable, and there is considerable empirical evidence that yields to maturity are integrated rather than stationary processes, so that conventional statistical analysis is not necessarily appropriate in this context.

B. Integration and Cointegration within the Term Structure

A series $X(t)$, which needs to be differenced d times before it has a stationary invertible ARMA representation, is said to be integrated of order

d , and this property is represented by the notation $X(t) \sim I(d)$. It is generally accepted that interest rates, and Treasury bill yields in particular, are well described as $I(1)$ processes.²

Given that the vector series $X(t)$ has only $I(1)$ components, it is sometimes possible to find vectors of constants $\alpha_1, \alpha_2, \dots, \alpha_r$ such that the linear combinations $\alpha'_i X(t)$ are all $I(0)$. In this case we say that $X(t)$ is cointegrated, and we define the vectors $\alpha_1, \alpha_2, \dots, \alpha_r$ to be cointegrating vectors. The space spanned by the cointegrating vectors is called the cointegration space.

Assuming that yields to maturity are integrated $I(1)$ processes, the possibility that they might be cointegrated is seen by rearranging equation (3) to obtain

$$[R(k, t) - R(1, t)] = \frac{1}{k} \sum_{i=1}^{k-1} \sum_{j=1}^{j-i} E_t \Delta R(1, t + j) + L(k, t) \quad (4)$$

where $\Delta R(k, s) = R(k, s) - R(k, s - 1)$. The right hand side of equation (4) is stationary provided that $\Delta R(1, t)$ and the premia $L(k, t)$ are stationary. Given these conditions, it follows that the left hand side of equation (4) is stationary and that $(1, -1)$ is a cointegrating vector for $X(t) = [R(k, t), R(1, t)]$. This implies that each yield $R(k, t)$ is cointegrated with $R(1, t)$, and that the spreads between $R(k, t)$ and $R(1, t)$ are the stationary linear combinations of $X(t)$ which result from the cointegration of $X(t)$. We define the spread between the yields $R(i, t)$ and $R(j, t)$ as $S(i, j, t) = R(i, t) - R(j, t)$.

The cointegration implied by the above considerations is of a very special type. Specifically, the model predicts that any yield series is cointegrated with the one period yield, so that if we were to consider a set of n yield series (which included the one period yield), then each of the $(n - 1), n$ dimensional spread vectors contained in the set

$$\begin{aligned} & [(-1, 1, 0, \dots, 0)', (-1, 0, 1, 0, \dots, 0)', \\ & \dots, (-1, 0, \dots, 0, 1)'] \end{aligned}$$

is cointegrating for the (now augmented) vector

² See, for instance, Campbell and Shiller (1988), Stock and Watson (1988) or Engle and Granger (1987). For a formal analysis of Treasury Bill yields, see Anderson, Granger and Hall (1990).

$X(t) = [R(1, t), R(k_2, t), R(k_3, t), \dots, R(k_n, t)]$ (in which k_2, k_3, \dots, k_n are the maturities of the other $(n - 1)$ bills). As these spread vectors are linearly independent, the cointegration space has rank $(n - 1)$.

Given the above arguments, it is straightforward to show that the spread between any two yields will be cointegrating. The spread vector associated with any two yields is just a linear combination of two of the spread vectors defined using the one period yield, and since linear combinations of stationary variables are also stationary it follows that this more general spread vector is cointegrating. An implication of the finding that any spread is cointegrating is that any set of $(n - 1)$ linearly independent spread vectors defined in an n dimensional space will comprise a basis for the cointegrating space associated with $X(t) = [R(1, t), R(k_2, t), R(k_3, t), \dots, R(k_n, t)]$. Thus any set of n yields will have a cointegrating rank of $(n - 1)$.

This cointegration between yields of different maturity implies analogous cointegration between the one month holding returns associated with Treasury bills of different maturities. If $H(k, t + 1)$ is the continuously compounded rate of return from t to $t + 1$ (one month) on a Treasury bill with k months to maturity at t , then it is straightforward to demonstrate that

$$\begin{aligned} H(1, t + 1) &= R(1, t), \\ H(k, t + 1) &= k[R(k, t) - R(k - 1, t)] \\ &\quad - k[\Delta R(k - 1, t + 1)] \\ &\quad + R(k - 1, t + 1) \quad \text{for } k \geq 2, \end{aligned}$$

and that the return in excess of the one-month rate will be

$$\begin{aligned} H(k, t + 1) - H(1, t + 1) &= k[R(k, t) - R(k - 1, t)] \\ &\quad - (k - 1)[\Delta R(k - 1, t + 1)] \\ &\quad + [R(k - 1, t) - R(1, t)] \quad \text{for } k \geq 2. \end{aligned}$$

It follows from the properties of the yields that the holding returns are also $I(1)$ processes, that any set of n holding returns will have a cointegrating rank of $(n - 1)$. If this set includes the one-month holding return, the $(n - 1)$ "excess returns" will form a basis for the cointegrating space.

C. Modelling Cointegrated Data

It was shown in Engle and Granger (1987) that cointegration implies and is implied by an error correction representation, which in the case of the series $X(t) = [R(1, t), R(2, t), \dots, R(k, t)]$, can be expressed by the equation

$$\begin{aligned} \Delta X(t) &= -\delta[S(t - 1) - \mu] \\ &\quad + c(B)\Delta X(t - 1) + d(B)\epsilon(t) \end{aligned} \quad (5)$$

where δ is a non-zero $n \times (n - 1)$ matrix, $S(t)$ is an $(n - 1) \times 1$ vector of spreads, $c(B)$ and $d(B)$ are polynomials in the lag operator B , and $\epsilon(t)$ is a vector of white noise, which may be contemporaneously correlated. The vector $[S(t - 1) - \mu]$ is called the error correction term, while δ is a matrix of adjustment coefficients. Statistical significance of the δ will show that the error correction model is a valid representation of the data, and support the hypothesis that the spreads contained in $S(t)$ are cointegrating.

The error correction model has a very sensible economic interpretation in this context. Equation (5) shows that although yields on bonds of different maturity may diverge in the short run, the yields will adjust when the spreads between them deviate from the equilibrium value μ , so that in the long run yields of different maturity will move together. The error correction model does not necessarily imply that yields adjust because the spreads between them are out of equilibrium. As Campbell and Shiller (1987, 1988) point out in the context of their present value models of the term structure, the spreads might measure anticipated changes in yields. Using the short yields as an example, this merely implies that agents have more information in the spread for forecasting changes in short yields, than is available in the history of short yields alone. Thus the spreads are useful for forecasting changes in short yields, and the error correction model arises because of agents' forward looking behavior.

An alternative interpretation of the cointegration between yields of different maturities arises from the relationship between cointegration and common factors. Stock and Watson (1988) show that when there are $(n - p)$ linearly independently cointegrating vectors for a set of n $I(1)$ variables, then each of these n variables can be expressed as a linear combination of p $I(1)$ common factors and an $I(0)$ component. Applying

this result to the current context, we expect that there will be a single nonstationary common factor in yields of different maturity. Denoting the $I(1)$ common factor by $W(t)$, a simple representation of how it links the yields curve is given by

$$R(1, t) = A(1, t) + b_1W(t)$$

$$R(2, t) = A(2, t) + b_2W(t)$$

...

$$R(n, t) = A(n, t) + b_nW(t)$$

in which the $A(i, t)$ are $I(0)$ variables. Since $W(t)$ is $I(1)$ while the $A(i, t)$ are $I(0)$, the observed long-run movement in each yield series is primarily due to movement in the common factor. Thus $W(t)$ "drives" the time series behavior of each yield and determines how the entire yield curve changes over time. There may be a number of additional factors that explain the variation in the $I(0)$ variables $A(i, t)$, but these factors will be stationary and dominated by the nonstationary factor $W(t)$.

The assertion that the same common variable underlies the time series behavior of each yield to maturity is not new to the literature on the term structure. Cox, Ingersoll and Ross (1985) build a continuous time general equilibrium model of real yields to maturity in which the instantaneous interest rate is common to all yields. In the discrete time model developed in this paper it is emphasized that there is only one nonstationary $I(1)$ common variable. Here, one could interpret this nonstationary common factor as the one period yield, or for that matter, any of the other period yields. It is also appropriate to think of this common factor as something exogenous to the system of yields such as inflation, measures of monetary growth, or measures of investment.

III. The Data

The analysis has been conducted on the nominal yield to maturity data from the Fama Twelve Month Treasury Bill Term Structure File of the U.S. Government Securities File of the Center for Research in Securities Prices (CRSP) at the University of Chicago. The file contains twelve yield series on Treasury bills; one series for bills with one month to maturity, another for bills with two months to maturity, and so on to a series with twelve months to maturity. Full details of how the

file has been constructed are given in the CRSP documentation. These data are particularly appropriate for an investigation of the term structure. The observed yield on each bill has been derived from the price of that bill on a given day (the last trading day of the month), so that the data relate to bills which are identical in all respects other than term, and unlike many yield data sets, the raw data have been neither interpolated over time nor interpolated over maturities.

The nominal yield series studied here have been derived by taking the average of bid and asked quotes. The yields are standardized to a 30.4 day basis, and are expressed in percentages. The sample used consists of 228 observations for each series, dating from January 1970 until December 1988, but the series on yields to maturity for twelve month bonds has not been used because many of the observations were missing.³

The sample covers three monetary regimes which are distinguished by the degree of interest rate targeting undertaken by the Federal Reserve. The first regime, covering the period up to and including September 1979, corresponds to a period during which the Federal Reserve was targeting interest rates. The period from October 1979 to September 1982 covers the Federal Reserve's "new operating procedures," when it ceased targeting interest rates. The final regime, from October 1982 onwards, corresponds to the abandonment of the "new operating procedures" and the resumption of partial interest rate targeting. Plots of the yield data and differenced yield data for the four yields of shortest term are provided in charts 1 and 2. These are representative of all the yields to maturity, and they illustrate the similar behavior of the yields over the sample period. In particular, they illustrate that the yields were considerably more volatile during the "new operating procedures" regime than they have been at other times. Most of the analysis is based on full sample, but in view of the regime changes described above, and of empirical evidence that these caused structural changes in the term structure,⁴ three subsets corresponding to

³ Two observations for the eleven month bill and one for the ten month bill were also missing. These missing values were interpolated from the observed movements in the yield of the nine month bill.

⁴ See, for instance, Huizinga and Mishkin (1986) or Hardouvelis (1988).

CHART 1.—YIELDS TO MATURITY (% PER MONTH)

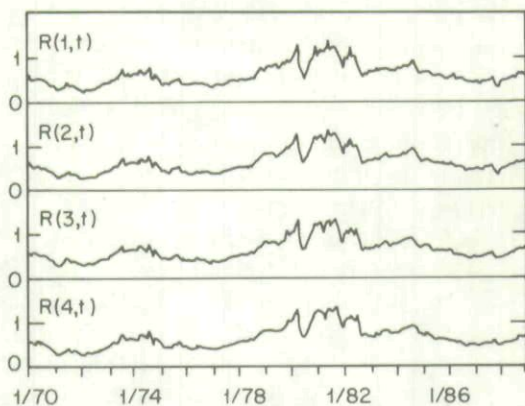
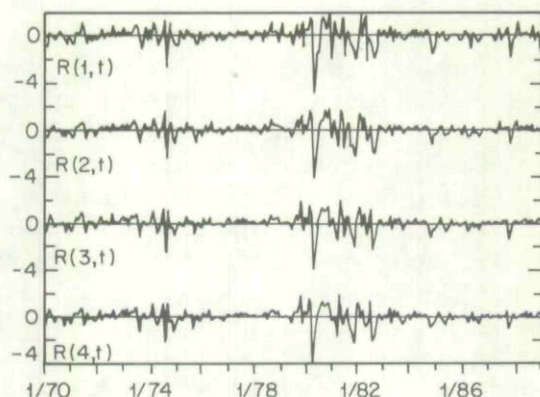


CHART 2.—DIFFERENCED YIELDS TO MATURITY (% PER MONTH)



the monetary regimes have also been analyzed. The SHAZAM (White (1978)) and PC-GIVE (Hendry (1989)) computer packages were used for the computations.

IV. The Empirical Evidence

A. Time Series Properties of Individual Yields

Augmented Dickey-Fuller unit root test statistics were computed for each of the eleven yield series, and the details of this analysis can be found in Anderson, Granger and Hall (1990). The full sample test statistics show no evidence against the null hypothesis that there is a unit root in yield levels, but the data clearly reject the null hypothesis that there is a unit root in the differences. When the three subsamples are examined, the same pattern emerges for each of the eleven yield series. A reasonable conclusion is that each

yield to maturity is an $I(1)$ process, over each of the Federal Reserve's monetary regimes.⁵

B. Cointegration Analysis

We now consider the hypotheses of interest, namely, that the yields are cointegrated with $(n - 1)$ cointegrating vectors corresponding to any set of n yields, and that the cointegrating vectors are the spread vectors.⁶ Johansen (1988) and Johansen and Juselius (1990) have developed likelihood based procedures which test for cointegration, estimate the cointegrating vectors and permit the testing of restrictions on the cointegrating vectors. These techniques have been applied to test the hypotheses of interest.

The results for the analysis which uses the eleven yields to maturity are presented in table 1. Johansen's λ -max and trace statistics accept the restriction that the rank of the cointegrating space is not more than ten, but strongly reject the hypotheses that the rank is not more than nine. This supports the proposition suggested by the theory that there are ten cointegrating vectors for the set of eleven yields. Conditional on there being ten cointegrating vectors, the null hypothesis that ten linearly independent spreads formed from the eleven yields comprise a basis for the cointegration space is rejected. This likelihood ratio test statistic is distributed as a chi-squared random variable with ten degrees of freedom under the null hypothesis. The value of the test statistic is 30.28, compared to its 5% critical value

⁵ The conclusion that yields to maturity are integrated processes cannot be true in a very strict sense because integrated series are unbounded, while nominal yields are bounded below by zero. Nevertheless, it is evident from the data that the statistical characteristics of yields are closer to those of $I(1)$ series than they are to $I(0)$ series, so that for the purposes of building models of the term structure it is appropriate to treat these yield series as if they were $I(1)$.

⁶ A unit root analysis of spreads provides indirect evidence on these hypotheses and such an analysis of each spread between all of the yields can be found in Anderson, Granger and Hall (1990). The spreads are found to be stationary over the full sample and in the first and third subsamples, consistent with the proposal that each spread is cointegrating. In the second subsample, many of the spreads are found to be nonstationary. The yields may still be cointegrating in this subsample, but the spreads do not define the cointegrating vectors. While relevant in this context, a unit root test on a spread tests the null hypothesis that the vector $[-1, 1]$ is not cointegrating, rather than the required null that the vector $[-1, 1]$ is cointegrating. As well, for sets of more than two yields, the unit root tests do not test the joint hypotheses that the spread vectors are cointegrating.

TABLE 1.—HYPOTHESIS TESTS TO DETERMINE THE COINTEGRATING RANK
FOR THE SET OF YIELDS $R(1, t), \dots, R(11, t)$
FULL SAMPLE (1970:3–1988:12)

Null Hypothesis about Rank r	λ -max Test Statistic	5% Critical Value	trace Test Statistic	5% Critical Value
$r \leq 10$	6.33	8.08	6.33	8.08
$r \leq 9$	29.38	14.10	35.71	17.84

Note: The critical values are from Johansen and Juselius (1990).

of 18.31. There are two plausible explanations for this rejection; either the spreads are not cointegrating, contradicting the theory, or the rejection has been caused by problems associated with the changes in the monetary regimes.

To investigate the first possibility, subsets of the spreads were tested to see whether singly or jointly they are contained in the cointegrating space. A selection of all the possible tests of hypothesis involving subsets of the various spreads between the eleven yields are summarized in table 2. In this table, the first column lists m yields, and the null hypothesis in each case is that the $(m - 1)$ linearly independent spreads formed from these yields are contained in the cointegration space. These likelihood ratio test statistics are all conditional on the rank of cointegration space being ten. The first row reports the test statistic that the ten linearly independent spreads span the cointegration space. The next block of test statistics considers the null hypotheses that an individual spread belongs to the cointegrating space. We report the tests for all spreads involving the one-month yield and all spreads involving adjacent maturities. For the tests involving the one-month yield, the null is accepted for seven out of the ten spreads. We find that the spreads $S(2, 1, t)$, $S(3, 1, t)$ and $S(4, 1, t)$ are not cointegrating. For the tests involving adjacent maturities, the null is accepted for six of the ten spreads, and in this instance the spreads $S(2, 1, t)$, $S(3, 2, t)$, $S(7, 6, t)$ and $S(11, 10, t)$ are not cointegrating. The next block of table 2 reports the test statistics obtained when we progressively increase the number of yields (k) in the subset, and test the null hypotheses that a set of $(k - 1)$ linearly independent spreads formed from these yields belongs to the cointegrating space. All of these joint hypotheses are rejected. The final block of statistics reports test statistics on all possible subsets of spreads involving the four shortest yields

TABLE 2.—TESTS THAT SPREAD
VECTORS ARE COINTEGRATING
FULL SAMPLE (1970:3–1988:12)

Spreads between	Test Statistic	DF	5% Critical Value
$R(1)$ through $R(11)$	30.28	10	18.31
$R(1), R(2)$	4.39	1	3.84
$R(1), R(3)$	6.56	1	3.84
$R(1), R(4)$	4.36	1	3.84
$R(1), R(5)$	2.48	1	3.84
$R(1), R(6)$	1.54	1	3.84
$R(1), R(7)$	0.65	1	3.84
$R(1), R(8)$	0.34	1	3.84
$R(1), R(9)$	0.21	1	3.84
$R(1), R(10)$	0.11	1	3.84
$R(1), R(11)$	0.01	1	3.84
$R(2), R(3)$	6.95	1	3.84
$R(3), R(4)$	0.00	1	3.84
$R(4), R(5)$	0.26	1	3.84
$R(5), R(6)$	0.73	1	3.84
$R(6), R(7)$	5.49	1	3.84
$R(7), R(8)$	2.51	1	3.84
$R(8), R(9)$	0.72	1	3.84
$R(9), R(10)$	1.15	1	3.84
$R(10), R(11)$	4.57	1	3.84
$R(1)$ through $R(3)$	7.27	2	5.99
$R(1)$ through $R(4)$	13.31	3	7.81
$R(1)$ through $R(5)$	14.34	4	9.49
$R(1)$ through $R(6)$	14.34	5	11.07
$R(1)$ through $R(7)$	21.96	6	12.59
$R(1)$ through $R(8)$	21.96	7	14.07
$R(1)$ through $R(9)$	21.97	8	15.51
$R(1)$ through $R(10)$	22.58	9	18.31
$R(2), R(4)$	2.30	1	3.84
$R(1), R(2), R(4)$	4.81	2	5.99
$R(1), R(3), R(4)$	7.84	2	5.99
$R(2), R(3), R(4)$	12.81	2	5.99

Note: $R(k)$ is the yield to maturity of a k period bill. Column one lists m yields. The null hypothesis in each case is that $(m - 1)$ linearly independent spreads formed from these yields belong in the cointegration space. The tests are conditional on the rank of the cointegration space being 10, and the test statistics have a Chi-squared distribution with DF degrees of freedom.

to maturity. Again we find a mixture of acceptances and rejection of the null hypotheses. In general, rejections seem to occur when the spread involves either the one-month, two-month, three-month or eleven-month yields.

A subsample analysis has not been performed in the eleven variable case due to degrees of

freedom considerations. In order to analyze the possible effects of the changes in the Federal Reserve's operating procedures, a detailed analysis of the four shortest yields has been performed. Table 3 reports the results of the tests to determine the cointegrating rank of these four yields. Over the full sample, the tests accept the null hypothesis that the rank of the cointegrating space is not more than three, but reject the null that the rank is not more than two. This confirms, as the theory predicts, that the four shortest yields are cointegrated and that the cointegrating rank is three. This result is repeated in the first and third subsample, but in the sample during which the new procedures were operating the tests suggest that the cointegrating rank is two. Hypothe-

sis tests that the spreads are contained in the cointegration space are reported in table 4. For the full sample, conditional on there being three cointegrating vectors, the hypothesis that three linearly independent spreads span the cointegrating space is rejected, and an analysis of subsets of these spreads also leads to some rejections. These results are consistent with the results of testing the same hypotheses in the eleven yield model. However, in the first and third subsamples, we can accept the hypothesis that the spreads form a basis for the cointegration space. With only one exception each of the hypotheses that subsets of these spreads are contained in the cointegration space is not rejected. These results are consistent with the predictions of the theory. On the other

TABLE 3.—HYPOTHESIS TESTS TO DETERMINE THE COINTEGRATING RANK FOR THE SET OF YIELDS $R(1, t)$, $R(2, t)$, $R(3, t)$ AND $R(4, t)$

Sample	Null Hypothesis about Rank r	λ -max Test Statistic	5% Critical Value	<i>trace</i> Test Statistic	5% Critical Value
Full Sample 70:3–88:12	$r \leq 3$	6.42	8.08	6.42	8.08
	$r \leq 2$	50.79	14.60	57.20	17.84
First Sample 70:3–79:9	$r \leq 3$	0.31	8.08	0.31	8.08
	$r \leq 2$	39.27	14.60	39.58	17.84
Second Sample 79:10–82:9	$r \leq 3$	3.22	8.08	3.22	8.08
	$r \leq 2$	10.50	14.60	13.72	17.84
	$r \leq 1$	20.14	21.28	33.87	31.26
	$r = 0$	28.56	27.34	62.43	48.42
Third Sample 82:10–88:12	$r \leq 3$	1.64	8.08	1.64	8.08
	$r \leq 2$	16.97	14.60	18.61	17.84

Note: The critical values are from Johansen and Juselius (1990).

TABLE 4.—TESTS THAT SPREAD VECTORS ARE COINTEGRATING

Spreads Between	Sample Period				DF	5% Critical Value
	70:3–88:12	70:3–79:9	79:10–82:9	82:10–88:12		
$R(1)$ through $R(4)$	14.66	5.56	...	1.87	3	7.81
$R(1), R(2)$	3.77	5.39	8.08	0.24	1	3.84
$R(1), R(3)$	5.49	2.83	8.04	0.02	1	3.84
$R(1), R(4)$	3.44	2.13	8.57	0.00	1	3.84
$R(2), R(3)$	6.20	0.00	7.29	0.15	1	3.84
$R(2), R(4)$	1.69	0.02	10.02	0.23	1	3.84
$R(3), R(4)$	0.45	0.05	12.92	0.29	1	3.84
$R(1), R(2), R(3)$	6.32	5.56	13.92	1.11	2	5.99
$R(1), R(2), R(4)$	3.86	5.48	21.50	1.66	2	5.99
$R(1), R(3), R(4)$	5.98	4.20	24.04	1.86	2	5.99
$R(2), R(3), R(4)$	14.09	0.05	23.21	0.30	2	5.99

Note: $R(k)$ is the yield to maturity of a k period bill. Column one list m yields. The null hypothesis in each case is that $(m - 1)$ linearly independent spreads formed from these yields belong in the cointegration space. For the full sample and subsamples (70:3–79:9) and (82:10–88:12), the test statistics are conditional on there being 3 cointegrating vectors. For the subsample (79:10–82:9), the test statistics are conditional on 2 cointegrating vectors. All test statistics have a chi-squared distribution with DF degrees of freedom.

hand, the results from an analysis of the second subsample are not consistent with the theory. Conditional on there being two cointegrating vectors, the tests indicate that none of the possible spread vectors are cointegrating.

On the basis of this evidence, we conclude that during periods in which the Federal Reserve has targeted interest rates as an instrument of monetary policy, the tests broadly support the predictions of the theory. We find $(n - 1)$ cointegrating vectors among each n yields to maturity, and it is reasonable to conclude that the spreads form a basis for the cointegrating space. This cointegrating relationship has the important implication that the risk or liquidity premia of Treasury bills are stationary $I(0)$ variables. This conclusion follows directly from consideration of equation (4) and the empirical evidence that yields are $I(1)$ and cointegrated processes, and the findings that the spreads between the yields define the cointegrating relationships.

These relationships appear to have broken down during the period of the new operating procedures. During this time, the Federal Reserve placed primary emphasis on controlling the growth of reserves available to depository institutions while greatly expanding the allowable range of fluctuations in the federal funds rate. This period experienced wide gyrations in quarterly monetary growth rates despite the announced policy of controlling the growth in monetary aggregates, unusually high real interest rates, changing inflation and deteriorating economic conditions. Short-term interest rates were influenced almost exclusively by the private sector. There was a marked increase in the short-run volatility of interest rates, a conventional measure of the risk of holding long-term debt, presumably with a substantial impact on risk or liquidity premia. Over this period we observe a change in the cointegrating relationships between yields on Treasury bills. Yields are still cointegrated, but the spreads no longer define the cointegrating relationships, and there appears to be at least one extra nonstationary common factor over this period. A reasonable explanation is that because of the uncertainty caused by the enhanced volatility in monetary growth, interest rates and economic activity resulting from the introduction of the new procedures, the risk or liquidity premia

became nonstationary over this period, causing a breakdown of the cointegrating relationships.

C. Error Correction Models

In this section we present an estimated error correction model using the four shortest yields to illustrate how these cointegration results might be utilised. The spreads are used to define the cointegrating vectors, but because this is not consistent with the data over the whole sample, estimation of the model is restricted to the period after the new operating procedures were abandoned.⁷

The error correction model presented here was derived by sequentially reducing a general unrestricted model that contained four lags of each of the differenced series. Details of the advantages of modeling by reduction may be found in Hendry (1989). Sets of $(n - 1)$ linearly independent spread vectors are not unique for $n \geq 2$, so that it was necessary to choose which spreads were to be used in the estimation of the error correction model. In theory it should not matter which spreads are used; in practice, we used the spreads $S(2, 1, t)$, $S(3, 2, t)$ and $S(4, 3, t)$, since these were the least correlated. In estimating the model, it was necessary to include two dummy variables, $D84$ and $D87$, to account for outliers which occurred in October 1984 and October 1987 (the second outlier is presumably due to the effects of the stock market crash).

Ordinary least squares (OLS) and full information maximum likelihood (FIML) model reductions lead to the same model specification and these estimates are presented in tables 5 and 6. Diagnostic statistics reveal little insample evidence of misspecification. The diagnostic test statistics are those produced by PC-GIVE and details of each test statistic can be found in Hendry (1989). Forecast Chow tests of the null hypothesis that there is no change in any parameter between the sample period (January 1983 until December 1987) and the forecast period (January 1988 until December 1988) show weak

⁷ As expected, error correction models estimated over the full sample showed evidence of instability. It may have been possible to obtain stable models of yields over longer samples by including (exogenous) volatility variables in the error correction model, or by introducing ARCH errors into the models, but these approaches were not tried here.

TABLE 5.—OLS ERROR CORRECTION MODEL FOR THE FOUR VARIABLE SYSTEM (1983:1–1987:12)

Explanatory Variable	Model for $\Delta R(1, t)$		Model for $\Delta R(2, t)$		Model for $\Delta R(3, t)$		Model for $\Delta R(4, t)$	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
$S(2, 1, t - 1)$.644	.205	—	—	—	—	—	—
$S(3, 2, t - 1)$.665	.296	.994	.272	—	—	—	—
$S(4, 3, t - 1)$	-.682	.438	-.268	.406	.664	.382	—	—
$\Delta R(2, t - 1)$.293	.108	.224	.100	—	—	—	—
Constant	-.014	.006	-.012	.005	-.004	.004	.0002	.004
D84	-.127	.030	-.113	.028	-.105	.029	-.098	.030
D87	-.176	.030	-.125	.028	-.126	.029	-.100	.030

Diagnostic Statistics								
Type	Model for $\Delta R(1, t)$		Model for $\Delta R(2, t)$		Model for $\Delta R(3, t)$		Model for $\Delta R(4, t)$	
	Distribution	Test	Distribution	Test	Distribution	Test	Distribution	Test
Dep. Var. S.D.		0.0444		0.0375		0.0350		0.0345
R^2		0.6050		0.5185		0.3812		0.2723
Standard Error		0.0295		0.0272		0.0283		0.0299
Serial Correlation	$F(12, 41)$	1.35	$F(12, 42)$	0.75	$F(12, 44)$	0.31	$F(12, 45)$	0.49
ARCH	$F(4, 45)$	2.40	$F(4, 46)$	1.45	$F(4, 48)$	0.41	$F(4, 49)$	0.19
Normality	$\chi^2(2)$	1.39	$\chi^2(2)$	0.41	$\chi^2(2)$	1.00	$\chi^2(2)$	1.51
Heteroskedasticity	$F(11, 41)$	1.00	$F(9, 44)$	0.51	$F(5, 50)$	0.52	$F(3, 53)$	0.45
Reset	$F(1, 52)$	0.60	$F(1, 53)$	0.21	$F(1, 55)$	0.85	$F(1, 56)$	0.11
Functional Form	n.a.		$F(11, 42)$	0.45	$F(4, 51)$	0.52	$F(2, 54)$	0.48
Chow	$F(12, 53)$	2.75 ^a	$F(12, 54)$	0.89	$F(12, 56)$	0.64	$F(12, 57)$	0.61

Note: — implies that in the reduction process the estimated coefficient was found to be insignificant. n.a. means that the statistic was not computed.
^a Significant at the 5% critical level.

evidence of change in the equation for $\Delta R(1, t)$. This is apparently due to an outlying observation for $\Delta R(1, t)$ in December 1988, which was more than five standard deviations away from the sample mean. Disregarding the effects of this outlying observation, further Chow tests provide no evidence that the estimated models are unstable.

Error correction terms have statistically significant coefficients, thereby confirming the cointegration found earlier and the validity of the error correction representation. It is interesting to note the manner in which the cointegrating vectors enter into each equation; the spreads are not relevant in the model for changes in the yield of

TABLE 6.—FIML ERROR CORRECTION MODEL FOR THE FOUR VARIABLE SYSTEM

Explanatory Variable	Model for $\Delta R(1, t)$		Model for $\Delta R(2, t)$		Model for $\Delta R(3, t)$		Model for $\Delta R(4, t)$	
	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
$S(2, 1, t - 1)$.864	.114	—	—	—	—	—	—
$S(3, 2, t - 1)$.602	.183	.983	.094	—	—	—	—
$S(4, 3, t - 1)$	-.873	.283	-.462	.188	.558	.106	—	—
$\Delta R(2, t - 1)$.186	.067	.092	.035	—	—	—	—
Constant	-.016	.005	-.011	.004	.003	.004	.0002	.004
D84	-.128	.029	-.116	.027	-.105	.028	-.098	.029
D87	-.171	.029	-.119	.027	-.126	.028	-.100	.029

Diagnostic Statistics				
Type	Model for $\Delta R(1, t)$	Model for $\Delta R(2, t)$	Model for $\Delta R(3, t)$	Model for $\Delta R(4, t)$
$P_{\text{actual/predicted}}$.7675	.7088	.6168	.5218
Standard Error	.0283	.0262	.0273	.0292

Note: — implies that in the reduction process the estimated coefficient was found to be insignificant.

longest maturity, but successively more spreads are needed to "explain" changes in yields as the term to maturity becomes shorter. This pattern is also found in other error correction models (not reported in this paper), estimated with different sets of yields. This type of model suggests that yields of longer maturities "drive" the term structure, with short-term yields adjusting to movements in the longer term yields. One interpretation of this observation is based on an expectations argument. The spreads between yields at the longer end of the term structure contain information about future shorter-term rates, and current short-term rates adjust according to this information.

D. Forecasts

The existence of an error correction model implies some Granger-causality in the system, which in turn suggests that the error correction model may be a useful forecasting tool. The error correction model estimated by FIML has been used to obtain 12 one-step ahead forecasts over the period 1988:1 to 1988:12, to illustrate its use for this purpose. These forecasts are compared with a set of naive no-change forecasts and the forecasts from an unrestricted second order vector autoregression (VAR). The dummy variables discussed above are included in the VAR model.

Table 7 provides the summary statistics for these forecasts. The biases are all of the same order of magnitude and all forecast standard deviations are high, but the error correction model has smaller forecast standard deviations, leading to consistently smaller root mean square errors. The error correction model gives between a 4% and 16% reduction in root mean square error over the naive model, and smaller gains over the VAR. The improvement in forecasts using the error correction model are small (and not statistically significant), but they illustrate the potential of this type of model.

V. Conclusion

This paper shows that it is appropriate to model the term structure of U.S. Treasury bills as a cointegrated system. During monetary regimes characterized by stabilizing the short-run fluctuations in the federal funds rate, the spreads between yields of different maturity define the

TABLE 7.—SUMMARY STATISTICS
FOR 1 STEP AHEAD FORECAST ERRORS
(1988:1–1988:12)

Variable	Method	Statistic		
		Mean	St. Dev.	RMSE
$\Delta R(1)$	Naive	.0144	.0568	.0586
	VAR	.0070	.0543	.0548
	ECM	.0033	.0491	.0492
$\Delta R(2)$	Naive	.0188	.0239	.0304
	VAR	.0111	.0268	.0290
	ECM	.0151	.0228	.0274
$\Delta R(3)$	Naive	.0184	.0164	.0247
	VAR	.0130	.0195	.0235
	ECM	.0152	.0169	.0228
$\Delta R(4)$	Naive	.0164	.0182	.0245
	VAR	.0091	.0227	.0245
	ECM	.0151	.0181	.0236
RMSE Ratios				
Ratio	$\Delta R(1)$	$\Delta R(2)$	$\Delta R(3)$	$\Delta R(4)$
ECM/Naive	.8406	.9006	.9220	.9626
ECM/VAR	.8987	.9421	.9696	.9631

cointegrating vectors in this system. An error correction model implied by this cointegration is estimated, found to be statistically significant, and it seems to provide more accurate forecasts of yields than naive no-change forecasts, or forecasts based on a VAR. During the period of the new operating procedures, yields are still cointegrated, but the spreads no longer define the cointegrating vectors.

The type of cointegration found for monetary regimes that emphasize controlling short-term interest rates, has the important implications that the term or liquidity premia of Treasury bills are stationary processes and that a single nonstationary common factor underlies the time series behavior of each yield to maturity. The common factor cannot be uniquely identified, and it could be a linear combination of several $I(1)$ variables. It is worth emphasizing that this is a nonstationary factor and it may be possible to find a number of common stationary factors that are useful in explaining the behavior of Treasury bill yields.⁸ Further research may suggest a useful way of

⁸ For this reason, our analysis is consistent with estimated factor models that use the stationary excess holding returns of Treasury bills. With these data, Stambaugh (1989) finds two common factors while Knez, Litterman and Scheinkman (1988) report estimated models with three and four factors.

identifying the common nonstationary factor so that it can then be estimated and studied. Much might be learned about the term structure if this common factor can be related to economic variables such as monetary growth and/or inflation, and further research on the common factor interpretation of cointegration in the term structure will undoubtedly improve our understanding of how the term structure changes over time.

REFERENCES

- Anderson, Heather M., Clive W. J. Granger, and Anthony D. Hall, "Treasury Bill Yield Curves and Cointegration," University of California, San Diego Discussion Paper Number 90-24, July 1990.
- Campbell, John Y., and Robert J. Shiller, "Cointegration and Tests of Present Value Models," *Journal of Political Economy* 95 (Oct. 1987), 1062-1088.
- , "Interpreting Cointegrated Models," *Journal of Economic Dynamics and Control* 12 (July 1988), 505-522.
- Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica* 53 (Mar. 1985), 385-407.
- Engle, Robert F., and Clive W. J., Granger, "Cointegration and Error Correction Representation, Estimation, and Testing," *Econometrica* 55 (Mar. 1987), 251-276.
- Granger, Clive W. J., "Some Properties of Time Series Data and Their Use in Econometric Model Specification," *Journal of Econometrics* 16 (May 1981), 121-130.
- Hardouvelis, Gikas A., "The Predictive Power of the Term Structure during Recent Monetary Regimes," *The Journal of Finance* 43 (June 1988), 339-356.
- Hendry, David F., "PC-GIVE: An Interactive Econometric Modelling System," Institute of Economics and Statistics and Nuffield College, University of Oxford, Jan. 1989.
- Huizinga, John, and Frederic S. Mishkin, "Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates," *Carnegie-Rochester Conference on Public Policy* 24 (Spring 1986), 231-274.
- Johansen, Soren, "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamics and Control* 12 (July 1988), 231-254.
- Johansen, Soren, and Katarina Juselius, "Maximum Likelihood Estimation and Inference on Cointegration-with Applications to the Demand for Money," *Oxford Bulletin of Economics and Statistics* 52 (May 1990), 169-210.
- Knez, Peter, Robert Litterman, and José Scheinkman, "Explorations into Factors Explaining Money Market Returns," Goldman Sachs & Co., Discussion Paper No. 6, Sept. 1989.
- Stambaugh, Robert F., "The Information in Forward Rates: Implications for Models of the Term Structure," *Journal of Financial Economics* 21 (May 1988), 41-70.
- Stock, James H., and Mark W. Watson, "Testing for Common Trends," *Journal of the American Statistical Association* 83 (Dec. 1988), 1097-1107.
- White, Kenneth J., "A General Computer Program for Econometric Methods-SHAZAM," *Econometrica* 46 (Jan. 1978), 239-240.

Copyright of *Review of Economics & Statistics* is the property of MIT Press and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.