

COINTEGRATION AND TESTS OF PRESENT VALUE MODELS

John Y. Campbell
Princeton University

Robert J. Shiller
Yale University

Econometric Research Program
Research Memorandum No. 323

November 1985
Revised March 1986

We are grateful to Don Andrews, Gregory Chow, Rob Engle, Dick Meese and Ken West, and to participants in seminars at the University of California Berkeley, Columbia University, the Federal Reserve Bank of Philadelphia, the National Bureau of Economic Research, Princeton University and Rice University for helpful comments on an earlier version of this paper. We are responsible for any remaining errors.

Econometric Research Program
PRINCETON UNIVERSITY
207 Dickinson Hall
Princeton, NJ 08544

ABSTRACT

In a model where a variable Y_t is proportional to the present value, with constant discount rate, of expected future values of a variable y_t , the "spread" $S_t \equiv Y_t - \theta y_t$ will be stationary for some θ whether or not y_t must be differenced to induce stationarity. Thus, Y_t and y_t are cointegrated. The model implies that S_t is proportional to the optimal forecast of ΔY_{t+1} , and also to the optimal forecast of S_t^* , the present value of future Δy_t .

We use vector autoregressive methods, and recent literature on cointegrated processes, to test the model. When Y_t is the long-term interest rate and y_t the short-term interest rate, we find in postwar U.S. data that S_t behaves much like an optimal forecast of S_t^* even though as earlier research has shown it is negatively correlated with ΔY_{t+1} . When Y_t is a real stock price index and y_t the corresponding real dividend, using annual U.S. data for 1871-1986 we obtain less encouraging results for the model, although the results are sensitive to the assumed discount rate.

John Y. Campbell
Dept. of Economics
Princeton University
Dickinson Hall
Princeton, NJ 08544
(609)-452-4011

Robert J. Shiller
Cowles Foundation
Yale University
Box 2125 Yale Station
New Haven, CT 06520
(203)-436-0685

COINTEGRATION AND TESTS OF PRESENT VALUE MODELS

Present value models are among the simplest dynamic stochastic models of economics. A present value model for two variables, y_t and Y_t , states that Y_t is a linear function of the present discounted value of expected future y_t :

$$(1) Y_t = \theta(1-\delta) \sum_{i=0}^{\infty} \delta^i E_t y_{t+i} + c$$

c , the constant, θ , the coefficient of proportionality, and δ , the discount factor, are parameters which may be known a priori or may need to be estimated. E_t , here and in what follows, denotes mathematical expectation, conditional on the full public information set I_t which includes y_t and Y_t themselves and in general exceeds the information set H_t available to the econometrician. Models of this form include the expectations theory for interest rates (Y_t is the long-term yield and y_t the one-period rate), the present value model of stock prices (Y_t is the stock price and y_t the dividend), and with some modification, the permanent income theory of consumption.¹

Despite the simplicity of their structure, there is a surprising degree of controversy about the validity of present value models for bonds, stocks and other economic variables.² The controversy seems to be stimulat-

¹ The discounted sum in (1) extends to an infinite horizon. Most of the methods in this paper can be applied to the finite horizon case, at the cost of some additional complexity. Hansen and Sargent [1981a] refer to the class of linear models in which, conditional on I_t , there is no error term, as "exact linear rational expectations models".^t Present value models fall into this category. Throughout this paper we will treat conditional expectations as equivalent to linear projections on information.

² For bonds, see Sargent [1979], Shiller [1979, 1981a, 1985], Hansen and

ed by three problems which arise in testing (1). First, there are several test procedures in the literature; these include single-equation regression tests, tests of cross-equation restrictions on a vector autoregression (VAR), and variance bounds tests. It is not clear how these alternative approaches are related to one another.

Secondly, a statistical rejection of the model (1) may not have much economic significance. It is entirely possible that the model explains most of the variation in Y_t even if it is rejected at the 5% level. Most work on present value models concentrates on statistical testing rather than informal evaluation of the "fit" of the models.

Finally, the variables y_t and Y_t usually require some transformation before the theory of stationary stochastic processes can be applied. One approach is to remove a deterministic linear trend, but this can bias test procedures against the model (1) if in fact y_t and Y_t are stationary in first differences.³

In this paper we develop a test of the present value relation that is valid when the variables are stationary in first differences.⁴ The test exploits the recently developed theory of cointegrated processes (Granger and Engle [1985], Stock [1984]). It is fully efficient, testing all the implications of the model, and can be interpreted as a single-equation regres-

Sargent [1981a], Shiller, Campbell and Schoenholtz [1983], and Campbell and Shiller [1984]. For stocks, see Shiller [1981b, 1984], LeRoy and Porter [1981], West [1984], Marsh and Merton [1985], Scott [1985], Mankiw, Romer and Shapiro [1985] and Kleidon [1986].

³ This point is made for stocks by Marsh and Merton [1985] and Kleidon [1986]. Mankiw and Shapiro [1984] present a similar argument for the permanent income theory of consumption.

⁴ It might be attractive to model the variables y and Y as stationary in log first differences, or growth rates. However since the model (1) is linear in levels, a log specification is intractable.

sion or as a test of restrictions on a VAR. The VAR system can also be used to test variance bounds, and to assess the economic significance of deviations from (1). Thus the methods of the paper alleviate all three of the problems discussed above. We apply our methods to the present value models for bonds and stocks, while a companion piece by one of us (Campbell [1986]) studies the permanent income theory of consumption.

The organization of the paper is as follows. Section 1 discusses alternative tests of the present value relation when y_t and Y_t are stationary in first differences rather than levels. Section 2 presents relevant results from the literature on cointegration, which are used in our nonstationary test. Section 3 applies the method to data on bonds and stocks. Section 4 concludes.

1. Alternative Tests of the Present Value Relation

One straightforward way to test the model (1) is to use it to restrict the behavior of the variable $\xi_t \equiv Y_t - (1/\delta)(Y_{t-1} - \theta(1-\delta)y_{t-1})$. Substitution from (1) shows that apart from a constant (which will be suppressed below for exposition),

$$(2) \quad \xi_t = Y_t - E_{t-1}Y_t$$

From equation (2), ξ_t is the true innovation at time t in Y_t (that is, the innovation with respect to the full market information set I_t). ξ_t is observable knowing only Y_t, Y_{t-1}, y_{t-1} and the parameters of the model. This is a striking implication of the present value relation. We note that in general the model does not identify the true innovation in y_t .

Since the right hand side of (2) is orthogonal to all elements of the information set I_{t-1} , one can test the present value relation by regressing ξ_t on information in this set and testing that the coefficients are jointly zero. Unpredictability of ξ_t is equivalent to unpredictability of one-period excess returns on bonds and stocks.

An alternative approach is to use equation (1) to restrict the coefficients of a vector autoregressive system.⁵ This requires the choice of variables which are stationary, have a well-behaved joint time series representation,⁶ and allow all the implications of (1) to be tested.

⁵ See Sargent [1979], Baillie, Lippens and McMahon [1983] or Ito [1985]. These are all finite-horizon undiscounted applications.

⁶ In particular, they must have an invertible vector moving average (VMA) representation.

If y_t and Y_t are themselves stationary in levels, then a bivariate VAR for y_t and Y_t can be used. If these variables are stationary in first differences, however, the obvious choice of a VAR for Δy_t and ΔY_t turns out to be inappropriate for two reasons. First, one cannot impose all the restrictions of (1) on a first-differenced VAR. Secondly, as will be explained in the next section, under the model (1) no well-behaved VAR representation exists for Δy_t and ΔY_t .⁷

Instead of first-differencing, one can subtract θy_t from both sides of equation (1). Defining $S_t \equiv Y_t - \theta y_t$ and rearranging (1), one obtains

$$(3) \quad S_t = E_t S_t^*, \quad \text{where } S_t^* = \theta \sum_{i=1}^{\infty} \delta^i \Delta y_{t+i}$$

and

$$(4) \quad S_t = (\delta/(1-\delta)) E_t \Delta Y_{t+1}$$

If Δy_t and ΔY_t are stationary, it follows from (3) and (4) that S_t is stationary.⁸ We will refer to S_t as the "spread". In the case of the term structure, it is just the spread between short and long term interest rates; for stocks, it is the difference between the stock price, and a multiple of dividends. These two equations show two alternative interpretations of the spread S_t : by (3) it is the optimal forecast of S_t^* , a weight-

⁷ Melino [1983] and Shiller [1981b] criticized Sargent [1979] on essentially these grounds, although they did not have the benefit of the theory of cointegrated processes in constructing their argument. Hansen and Sargent [1981a] responded by using the long-short spread in a test of the expectations theory of the term structure, which is also the solution adopted in this paper. Hansen and Sargent used frequency domain methods and did not appeal to the theory of cointegration.

⁸ (3) and (4) also show that under the model, if Δy is stationary then ΔY must be stationary.

ed average of future changes in y ; by (4), it is proportional to the optimal forecast of the change in Y .

Equation (4) suggests a second regression test of the present value relation. If ΔY_{t+1} is regressed on S_t and other variables, the coefficient on S_t should be $(1-\delta)/\delta$ and the coefficients on the other variables should be zero. Shiller, Campbell and Schoenholtz [1983] tested the expectations theory of the term structure with a regression of this sort. In this paper we focus primarily on equation (3). The restrictions of (3) can be tested by estimating a VAR representation for Δy_t and S_t :

$$(5) \quad \begin{bmatrix} \Delta y_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

where the polynomials in the lag operator $a(L)$, $b(L)$, $c(L)$ and $d(L)$ are all of order p . Equation (5) can be stacked into a first-order system

$$(6) \quad \begin{bmatrix} \Delta y_t \\ \cdot \\ \cdot \\ \Delta y_{t-p+1} \\ S_t \\ \cdot \\ \cdot \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & a_p & b_1 & \dots & b_p \\ 1 & & & & & \\ \cdot & & & & & \\ \cdot & & & & & \\ & & & 1 & & \\ c_1 & \dots & c_p & d_1 & \dots & d_p \\ \cdot & & & & & \\ \cdot & & & & & \\ & & & & & 1 \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \cdot \\ \cdot \\ \Delta y_{t-p} \\ S_{t-1} \\ \cdot \\ \cdot \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ 0 \\ \cdot \\ 0 \\ u_{2t} \\ 0 \\ \cdot \\ 0 \end{bmatrix}$$

which can be written more succinctly as $z_t = Az_{t-1} + v_t$. The matrix A is called the companion matrix of the VAR. For all i ,

$$E[z_{t+i} | H_t] = A^i z_t,$$

where H_t is the VAR's limited information set $\{z_t, z_{t-1}, \dots\}$, a subset of I_t . As elsewhere in the paper, we are taking conditional expectations to be linear projections on information.

It is simple to prove that if the present value relation holds, S_t must linearly Granger cause Δy_t (and therefore Y_t must Granger cause y_t) unless S_t is itself an exact linear function of current and lagged Δy_t . For suppose that S_t does not Granger cause Δy_t . Then $E[\Delta y_{t+i} | H_t] = E[\Delta y_{t+i} | \Delta y_t, \Delta y_{t-1}, \dots]$ for all i , and from (3) $E[S_t | H_t] = E[S_t | \Delta y_t, \Delta y_{t-1}, \dots]$. But because S_t is itself in the information set H_t , $S_t = E[S_t | H_t]$ so S_t is an exact linear function of current and lagged Δy_t . This implies that the variance-covariance matrix of u_{1t} and u_{2t} , Ω , is singular.

The intuitive explanation for this result is that S_t is an optimal forecast of future values of Δy_t , conditional on agents' full information set. S_t will therefore have incremental explanatory power for future Δy_t if agents have information useful for forecasting Δy_t beyond the history of that variable. If agents do not have such information, they form S_t as an exact linear function of current and lagged Δy_t .

Equation (3) can be projected onto the information set H_t , and rewritten as

$$g' z_t = \theta \sum_{i=1}^{\infty} \delta^i h' A^i z_t$$

where g' and h' are row vectors with $2p$ elements, all of which are zero except for the $p+1$ 'st element of g' and the 1st element of h' , which are unity. If this expression is to hold for general z_t (that is, for nonsingular Ω), it must be the case that

$$(7) \quad g' = \theta \sum_{i=0}^{\infty} \delta^i h'A^i = \theta h'\delta A [I - \delta A]^{-1}$$

The restrictions of equation (7) appear to be highly nonlinear cross-equation restrictions of the type described by Hansen and Sargent [1981b] as the "hallmark" of rational expectations models. However it turns out that (7) can be simplified so that its restrictions are linear and easily interpreted. Postmultiplying both sides of (7) by $[I - \delta A]$, one obtains

$$(8) \quad g'[I - \delta A] = \theta h'\delta A$$

From the structure of the matrix A , the constraints imposed by (8) on individual coefficients are $c_i = -\theta a_i$, $i=1\dots p$, $d_1 = 1/\delta - \theta b_1$, and $d_i = -\theta b_i$, $i=2\dots p$. By adding $\theta \Delta y_t$ to S_t , one can interpret these restrictions. They state that $\xi_t = S_t - (1/\delta)S_{t-1} + \theta \Delta y_t$ is unpredictable given lagged Δy_t and S_t , which is implied by equation (2).⁹ In our empirical application, we obtain a Wald test statistic for equation (8) which is numerically identical to the Wald test statistic for a regression of ξ_t on lagged Δy_t and S_t .¹⁰

⁹ This interpretation of the VAR restrictions does not hold in the undiscounted finite-horizon VAR model of Sargent [1979]. The assumption that Δy and ΔY are stationary rules out rational speculative bubbles (West [1984], Hamilton and Whiteman [1985]), which would allow ξ to be unpredictable even if (1) does not hold.

¹⁰ However this statistic will not be numerically identical to the Wald statistic for a test of equation (7), even though (7) and (8) are algebraically equivalent restrictions. Nonlinear transformations of restrictions can change the numerical values of Wald statistics, and as Gregory and Veall [1985] point out, can dramatically alter their power. We report Wald statistics for (8) in the tables which summarize our em-

The major advantage of the VAR framework is that it can be used to generate alternative measures of the economic importance, not merely the statistical significance, of deviations from the present value relation. To see this more clearly, suppose that the present value model is false so that $E_t \xi_{t+i} \neq 0$ for $i \geq 1$. Then equations (3) and (4) no longer hold. We define the theoretical spread as the optimal forecast, given the information set H_t , of the present value of all future changes in y :

$$(9) \quad S'_t \equiv E[S^*_t | H_t] = h' \delta A [I - \delta A]^{-1} z_t.$$

We then have

$$(10) \quad S_t - S'_t = \sum_{i=1}^{\infty} \delta^i E[\xi_{t+i} | H_t]$$

and

$$(11) \quad S_t - (\delta/(1-\delta)) E[\Delta Y_{t+1} | H_t] = (1/(1-\delta)) E[\xi_{t+1} | H_t].$$

Equations (10) and (11) measure deviations from the model in two different ways. The metric of equation (11) is the difference between S_t and the optimal forecast, given the information set H_t , of the one-period change in Y . Equation (11) shows that this difference is large if excess returns are predictable one period in advance.

The metric of equation (10) is the difference between S_t and the theoretical spread, which is large if the present value of all future excess returns is predictable. By this measure, a large deviation from the model requires not only that movements in ξ be predictable one period in advance,

pirical results, and Wald statistics for (7) in footnotes.

but that they be predictable many periods in advance. Loosely speaking, predictable excess returns must be persistent as well as variable.¹¹

We use the VAR framework not only to conduct statistical tests of the present value relation, but to evaluate its failures using the metric of equation (10). We display time series plots of the spread S_t and the theoretical spread S'_t , the unrestricted VAR forecast of the present value of future changes in y . If the present value model is true, these variables should differ only because of sampling error. Large observed differences in the time series movements of the two variables imply (subject to sampling error) economically important deviations from the model.

The VAR framework can also be used to test the present value model against more specific alternatives. Volatility tests, for example, are designed to test against the alternative that Y_t or some transformation of it "moves too much". We present two different volatility tests. The first is just a test that the ratio $\text{Var}(S_t)/\text{Var}(S'_t)$ is unity. This ratio, together with its standard error, can be computed from the VAR system.¹² Under the present value model, the ratio should be one, but would be larger than one if the spread is too volatile relative to information about future y .

We obtain a second volatility test as follows. Let us define ξ'_t as θ times the innovation from $t-1$ to t in the expected present value of Δy , conditional on the VAR information set:

¹¹ The terminology of our earlier paper (Campbell and Shiller [1984]) may be helpful in understanding (10) and (11). The right hand side of (11) is proportional to what we called the one-period "holding premium", and the right hand side of (10) is what we called the "rolling premium".

¹² We use the formula for the variance-covariance matrix of the vector AR-1 process z : $\text{Vec}(Z) = \text{Inverse}(I - A \otimes A) \text{Vec}(V)$, where Z is the variance-covariance matrix of z and V is the variance-covariance matrix of v . This formula requires inverting a $4p^2$ by $4p^2$ matrix, so we apply it only to our stock market data for which we use relatively short lag lengths.

$$\begin{aligned}
(12) \quad \xi'_t &\equiv \theta \sum_{i=0}^{\infty} (E[\Delta y_{t+i} | H_t] - E[\Delta y_{t+i} | H_{t-1}]) \\
&= S'_t - (1/\delta)S'_{t-1} + \theta \Delta y_t
\end{aligned}$$

Under the present value model $\xi'_t = \xi_t$ since $S'_t = S_t$. We construct the ratio $\text{Var}(\xi_t)/\text{Var}(\xi'_t)$, again with standard error. The model implies that this ratio should be one, while the notion that stock prices are too volatile suggests it will be greater than one.¹³ We call the first of our variance ratios the "levels variance ratio" and the second the "innovations variance ratio".

The fact that a linear combination S_t of y_t and Y_t is stationary in its level, even though y_t and Y_t are individually stationary only in first differences, turns out to be important for understanding present value models. In the language of time series analysis, the vector $x_t = [y_t \ Y_t]'$ is cointegrated. Cointegrated vectors have a number of important properties which we now discuss.

¹³ This measure of "excess" volatility is analogous to that in West [1984].

2. Properties of Cointegrated Vectors

In this section we summarize the theory of cointegrated processes, and show how it applies to present value models.

Definition (Granger and Engle [1985]). A vector x_t is said to be cointegrated of order (d,b) , denoted $x_t \text{ CI}(d,b)$, if (i) all components of x_t are integrated of order d (stationary in d 'th differences), and (ii) there exists at least one vector $\alpha (\neq 0)$ such that $\alpha'x_t$ is integrated of order $d-b$, $b > 0$.

When y_t is stationary in first differences, the vector $x_t = [y_t \ Y_t]'$ is $\text{CI}(1,1)$ if the present value model holds. The $\text{CI}(1,1)$ case is the one that has been studied almost exclusively in the theoretical literature, and the results which follow apply to it.

Cointegrated systems of order $(1,1)$ have two unusual properties. These concern the existence of vector time series representations for the cointegrated variables, and the estimation of unknown elements of the vector α . Both properties turn out to be relevant for testing present value models.

The first important property of a cointegrated vector is that no invertible vector moving average (VMA) representation, and therefore no finite VAR representation, exists for Δx_t . The reason is that if there were an invertible VMA representation, no linear combination of x_t could be stationary.

More formally, write $\Delta x_t = K(L)\varepsilon_t = I\varepsilon_t + K_1\varepsilon_{t-1} + \dots$. Invertibility requires that all roots of the determinant of $K(z)$ lie outside the unit circle, so $K(1) = I + K_1 + \dots$ must be nonsingular. The variance-covari-

ance matrix of ε_t , V , must also be nonsingular. Now if the variance of $\alpha'x_t$ exists, it will be given by

$$\text{Var}(\alpha'x_t) = \sum_{i=0}^{\infty} \alpha' C_i V C_i' \alpha \quad \text{where } C_i = I + K_1 + \dots + K_i.$$

Since the limit of C_i as $i \rightarrow \infty$ is $K(1)$, the terms in the summation above approach a nonzero limit and the variance of $\alpha'x_t$ will not be finite.

It follows from this that if an economic theory imposes cointegration on a set of nonstationary variables, simple first differencing of all the variables does not lead to a well-behaved system for statistical modelling. Fortunately, there is a simple solution to the difficulty which is to include $\alpha'x_t$ in a VAR along with a subset of the elements of Δx_t . An equation which relates the change in an element of x_t to its own lags and lags of $\alpha'x_t$ is called an error-correction model for that element of x_t . The VAR proposed in the previous section to test present value models is an error-correction model for y_t , along with an equation describing the evolution of $\alpha'x_t$.

The second major result from the theory of cointegration concerns the "cointegrating vector" α . In a present value model, α is unique up to a scalar normalization, and is proportional to $[-\theta \ 1]'$. Stock [1984] proves that if there is a single unknown element of α , a variety of methods provide estimates which converge to the true parameter at a rate proportional to the sample size T (rather than \sqrt{T} as in ordinary cases). The reason for this is that asymptotically all linear combinations of the elements of x_t other than $\alpha'x_t$ have infinite variance.

The practical implication is that an unknown element of α may be estimated in a first-stage regression and then treated as known in second-stage

procedures, whose asymptotic standard errors will still be correct. This is extremely useful in carrying out the VAR tests of the previous section. In the case of stock prices, for example, the present value model constrains $\theta = \delta/(1-\delta)$, so one can estimate the discount factor from a preliminary regression and then treat it as known in testing the model.

Two types of preliminary regression have been proposed for estimating the unknown parameter θ . The first, called the cointegrating regression by Granger and Engle [1985], is just a regression of Y_t on y_t . The second is an error-correction regression of Δy_t or ΔY_t on lagged changes in and levels of y_t and Y_t . In the first case, one estimates θ as the coefficient on y_t , while in the second case one takes the ratio of the coefficient on lagged y_t to that on lagged Y_t .

One might argue that use of the error-correction regression is preferable because it accounts more fully for the short-run dynamics of Y_t and y_t . However it has an important disadvantage. For any cointegrated vector with two elements, there are two possible error-correction regressions, one for Δy_t and one for ΔY_t . Cointegration alone does not rule out that in one of these regressions, lagged Y_t and y_t have zero coefficients in the population, so that the coefficient ratio fails to identify the desired parameter.¹⁴ Of course, under the present value model the error-correction equation for Δy_t has nonzero coefficients (because $\alpha'x_t$ Granger causes Δy_t), but this is not implied by all plausible alternatives. Accordingly we rely primarily on the cointegrating regression to identify θ .

¹⁴ Cointegration does rule out that the coefficients are zero in both error-correction regressions.

One may want to conduct a formal statistical test of the hypothesis that x_t is not cointegrated. This turns out to pose some difficult statistical problems. If the cointegrating vector α is known, one can use a Dickey-Fuller test (Dickey and Fuller [1981]), regressing the change in $\alpha'x_t$ on a constant and a single lagged level. The t statistics and F statistic are biased upwards relative to the t and F distributions, but Dickey and Fuller provide significance levels based on a Monte Carlo study. If the statistics are sufficiently high, the hypothesis of no cointegration is rejected.

If the cointegrating vector is not known but must be estimated from a cointegrating regression, the Dickey-Fuller significance levels need to be further adjusted. Granger and Engle [1985] analyze a variety of tests which use the residual from the cointegrating regression, an estimate of $\alpha'x_t$. We report two of their test statistics, one based on the Dickey-Fuller regression and one which augments that regression with four lagged dependent variables. Granger and Engle provide significance levels for these tests, again based on a Monte Carlo study. The Monte Carlo results are based on 10,000 replications of 100 observations of independent random walks, with 4 lagged residual changes included in the test.

3. Testing the Model in Bond and Stock Markets

In this section we apply the methods developed above to test present value models for bonds and stocks. The model for bonds, usually referred to as the "expectations theory of the term structure", is a special case of equation (1) in which the parameters θ and δ are known a priori (θ equals one, and δ is a parameter of linearization), while the constant c is a liquidity premium unrestricted by the model.¹⁵

We test the present value model for bonds on a monthly U.S. Treasury 20-year yield series, available from 1959 to 1983 from Salomon Brothers' Analytical Record of Yields and Yield Spreads. The short rate used is a 1-month Treasury bill rate, obtained from the U.S. Treasury Bulletin. These data were previously studied in Campbell and Shiller [1984]; Shiller, Campbell and Schoenholtz [1983] worked with very similar data. We present empirical results both for the full sample 1959:1-1983:11, and for a short sample ending in 1978:9 which is more likely to correspond to a single interest rate regime.¹⁶

The present value model for stocks is a special case of equation (1) in which θ is known to equal $\delta/(1-\delta)$. The model restricts the constant c to be zero. The discount factor δ is not known a priori, but can be inferred by estimating the cointegrating vector for stock prices and dividends; a consistent estimate is also provided by the sample mean return on stocks.

¹⁵ The linearization required to write the expectations theory in this form is explained in Shiller [1979] and Shiller, Campbell and Schoenholtz [1983].

¹⁶ For both samples, the parameter of linearization δ is set equal to $1/(1+R)$, with R at 0.0587/12 (the mean 20-year bond rate in the short sample, expressed at a monthly rate).

One difficulty with this formulation for stocks is that Y_t and y_t are not measured contemporaneously. Y_t is a beginning-of-period stock price and y_t is paid some time within period t . Literal application of the methods outlined in section 1 would require us to assume that y_t is known to the market at the start of period t ; but this might lead us to a spurious rejection of the model if in fact y_t is only known at the start of period $t+1$. Intuitively, it is not hard to "predict" excess returns using ex post information. In order to avoid this problem, we modify the procedures of section 1 by constructing a variable $SL_t \equiv Y_t - \theta y_{t-1}$. We use this variable in our tests and alter the cross-equation restrictions appropriately. The dependent variables in the VAR are now SL_t and Δy_{t-1} , both of which are in the information set at the start of time t but not at the start of time $t-1$ under our conservative assumption about the market's information.¹⁷ Since $SL_t = S_t + \theta \Delta y_t$, it is of course stationary if S_t and Δy_t are.

We test the model for stocks using time series data for real annual prices and dividends on a broad stock index from 1871 to 1986. Y_t is the Standard and Poor Composite Stock Price Index for January, divided by the January producer price index scaled so that the 1967 producer price index equals 100. (Before 1900 an annual average producer price index is used). The nominal dividend series is, starting in 1926, dividends per share adjusted to index, four quarter total, for the Standard and Poor Composite Index. The nominal dividend before 1926 is taken from Cowles [1939], who extended the Standard and Poor series back in time.¹⁸ Finally, y_t is the

¹⁷ Engle and Watson [1985] do some similar regressions to ours, using a similar data set on stock prices and dividends. They use the variable S rather than SL ; their results differ from ours in that they find no evidence of Granger causality from S to Δy , but they do not reject the present value model more strongly than we do.

nominal dividend series, divided by the annual average producer price index scaled so that the 1967 producer price index equals 100.

We begin in Tables 1A and 1B by running cointegrating and error-correction regressions to try to identify the parameter θ from the data.¹⁹ In the case of the term structure, the expectations theory imposes $\theta=1$, but it is of interest to see whether the data deliver an estimate close to this number.

The estimates of θ in Table 1A are not particularly close to those implied by the expectations theory. The cointegrating regression estimates θ to be 0.735 in the short sample, and 0.854 in the full sample.²⁰ The error-correction regression gives imprecise estimates of θ which are larger than unity.

Granger and Engle's [1985] tests for cointegration, based on the residual from the cointegrating regression, yield ambiguous results. Their ξ_2 statistic is the t statistic on the lagged residual in a regression with

¹⁸ The dividend data differ slightly from those used in Shiller [1981], West [1984], Mankiw, Romer and Shapiro [1985] and others. It has recently come to our attention that the second, 1939 edition of Cowles' book contains some corrections to the dividend series presented in the original 1938 edition, and these corrections are incorporated here.

¹⁹ Standard errors in these and all subsequent tables are White's [1984] heteroskedasticity-consistent standard errors. However the statistics used to test for cointegration do not correct for heteroskedasticity, as the Monte Carlo studies which generated their significance levels did not do so. Variance ratios are also unconditional variances which ignore heteroskedasticity. There is generally little difference between White's standard errors and the ordinary ones in our regressions.

²⁰ The standard errors reported for these coefficients are very small; the short sample estimate is 10 standard errors from one, while the full sample estimate is 6 standard errors from one. However these standard errors are not corrected for serial correlation in the cointegrating regression residual, and thus are likely to be far too small. Also the coefficients and standard errors in the cointegrating regression suffer from finite-sample bias as Stock [1984] points out.

the change in the residual as the dependent variable and a constant and the lagged residual as independent variables. Their ξ_3 statistic is the t statistic on the lagged residual in a regression which adds four lagged changes in the residual to the previous regression as independent variables. Both statistics fail to reject the null of no cointegration at the 10% level in the short sample, and only the ξ_2 statistic rejects at this level in the full sample.

These weak results may be due to loss of power when the cointegrating vector is estimated rather than imposed. Dickey and Fuller's [1981] $\hat{\phi}_1$ statistic is the F statistic for joint significance of the coefficients in a regression of the change in a series on a constant and the lagged level of the series. When we set $\theta = 1$ and use the $\hat{\phi}_1$ statistic to test for stationarity of the spread, we find that we can reject the null that the spread follows a random walk at the 5% level in both sample periods. This evidence encourages us to proceed, imposing θ equal to one.

In Table 1B we estimate the cointegrating vector for the data on stock prices and dividends. The cointegrating regression estimates θ at 31.092; the corresponding real discount rate (the reciprocal of θ) is 3.2%, which is lower than the average dividend-price ratio and considerably lower than the sample mean return of 8.2%.²¹ The error-correction regression delivers a fairly similar estimate of θ , 37.021 with an implied real discount rate of 2.7%. Granger and Engle's tests for no cointegration give mixed re-

²¹ The θ estimate which corresponds to the sample mean return is 12.195. The higher estimate in the cointegrating regression is associated with a negative constant term; under the present value model, the constant should be proportional to the unconditional mean change in dividends, so it should be positive rather than negative. An estimated discount rate lower than the mean dividend-price ratio is consistent with the model only if dividends are expected to decline through time, the historical rise being due to sampling error.

sults; the ξ_2 statistic rejects at the 5% level, while the ξ_3 statistic narrowly fails to reject at the 10% level. We proceed to construct SL_t using discount rates of 8.2% and 3.2% as a check on the robustness of our methods.

In Table 2A we report summary statistics for a VAR test of the expectations theory of the term structure. The VAR includes Δy_t and S_t as variables, and the number of lags is chosen by the Akaike Information Criterion (AIC).²² VARs are estimated for the short sample 1959-78, and the full sample 1959-83; they have 11 and 6 lags respectively.

In both sample periods the lagged variables have a fair degree of explanatory power for the change in short rates. The R^2 for the Δy_t equation is 21.6% in the short sample, and 17.1% in the full sample. This argues against the view of Mankiw and Miron [1985] that short rate changes are essentially unpredictable in the postwar period in the U.S. Furthermore, there is strong evidence that spreads Granger cause short rate changes as they should do if the expectations theory is true. The hypothesis of no Granger causality can be rejected at the 0.01% level for the short sample and the 0.3% level for the full sample. Short rate changes also Granger cause spreads, although the significance levels are somewhat weaker at 0.3% and 1.3% respectively.

A formal test of the expectations theory restrictions in equation (8) (which are just the restrictions of equation (2)) rejects very strongly. The null that excess returns on long bonds are unpredictable can be reject-

²² That is, we pick the number of lags to minimize [- ln likelihood + number of parameters] in the vector autoregression. Sawa [1978] has argued that the AIC tends to choose models of higher order than the true model, but states that the bias is negligible when $p < T/10$ as it is here. The test statistics in Tables 2A and 2B are not highly sensitive to small changes in the lag length of the VAR system.

ed at less than the 0.005% level in the short sample, and at the 0.03% level in the full sample. The corresponding R^2 values for excess returns are 26.3% and 16.7% respectively.²³

However this negative result for the model is softened by the results of two further exercises. First, a test of the innovations variance equality (12) does not reject. Table 2A reports the innovations variance ratio, along with an asymptotic standard error for this number. If the ratio is greater than one, long-term bond returns "move too much". We find that the ratio is 1.160 in the short sample but only 0.502 in the long sample, and in both samples the asymptotic standard error is very large.²⁴

Secondly, when we use the VAR to informally evaluate the model, we find that the model has considerable ability to explain the time series movements of the spread. In the short sample the standard deviations of the spread S_t , and the unrestricted VAR forecast of short rate changes S'_t , are very close at 1.060 and 1.067 respectively. The two variables have a high correlation of 0.978. In the long sample the spread appears to move too much; its standard deviation is 1.320, while the standard deviation of the unrestricted forecast is only 0.717. However the two variables still have a high correlation of 0.956. A plot following the tables illustrates the comovement of S_t and S'_t in the short sample.²⁵

²³ Nonlinear Wald tests of equation (7) reject at significance levels of less than 0.005% in the short sample, and 8.4% in the full sample.

²⁴ Runkle [1984] also finds large asymptotic standard errors on variance decompositions from VAR models.

²⁵ The high correlation of these variables in postwar U.S. data might also have been inferred from results in Modigliani and Shiller [1973] (see particularly figure 6). Despite the evidence reported in Modigliani and Shiller and in the present paper, one of us (Shiller [1979]) presented evidence suggesting that long-term interest rates are too volatile to accord with the expectations theory. By contrast with Modigliani and

These results suggest that tests of predictability of returns are highly sensitive to deviations from the expectations theory; they are so sensitive, in fact, that they may obscure some of the merits of the theory. An example illustrates the point. Suppose long and short rates differ from the expectations theory in the following manner: $S_t = S'_t + w_t$, where w_t is serially uncorrelated noise. As Campbell and Shiller [1984] point out, a regression of ΔY_{t+1} on S_t may find that the coefficient has the opposite sign to that predicted by (4), even if the variance of w_t is quite small. In our data, as reported in Table 2A, we find that this regression delivers the wrong sign in both sample periods. However a regression of S^*_t on S_t will find that the coefficient has the same sign as predicted by (3), and downward bias caused by w_t will be small if the variance of w_t is small. Moreover the innovations variance ratio $\text{Var}(\xi_t)/\text{Var}(\xi'_t)$ may not be much greater than one. In this example the metric of equation (10) reveals the strengths of the expectations theory which are obscured by the metric of equation (11).²⁶

In Table 2B we repeat the above exercises for stock prices and dividends. We work with one sample period, but two discount rates. The Akaike Criterion selects a 4-lag representation for the data when the sample mean discount rate 8.2% is used, and a 2-lag representation when the cointegrating regression discount rate 3.2% is used.

Shiller and the present paper, Shiller [1979] assumed that levels of short rates are stationary, an assumption more clearly appropriate for prewar data sets.

²⁶ We do not claim that this example is literally correct for our data. The model $S=S'+w$ can be tested, for any MA(q) process for w , by regressing ξ on information known $q+2$ periods earlier. We found that this test rejected the model for q up to 8 using the bond data for 1959-78.

The VAR estimates suggest that dividend changes are rather highly predictable; the R^2 values for the equations which explain them are around 40%. There is very strong evidence that price-dividend differences Granger cause dividend changes, which is what one would expect if there is any truth to the present value model for stock prices.

We conduct two formal tests of the model. The first restricts the mean of the price-dividend difference, while the second leaves the mean unconstrained and restricts only the dynamics of the variable. (In the case of the term structure, the mean spread is always unconstrained because we allow a constant risk premium).

The results of these tests include some statistical rejections at conventional significance levels, but they are not nearly as strong as the rejections in the term structure. The pattern of results is sensitive to the choice of discount rate. When the sample mean return is used, the mean restriction on SL_t is satisfied almost exactly. Therefore the test of only the dynamic restrictions in equation (8) rejects more strongly, at the 4.7% level as compared with the 7.2% level for the full set of restrictions. When the discount rate from the cointegrating regression is used, the complete set of restrictions is rejected at the 1.1% level while the significance level for the dynamic restrictions is only 21.8%.²⁷

These tests are "portmanteau" tests of the present value model against an unspecified alternative. We also present variance ratios in order to test against the specific alternative that stock prices "move too much" in levels or innovations. The point estimate of the levels variance ratio

²⁷ Nonlinear Wald tests of the dynamic restrictions in the form (7), rather than (8), reject at less than the 0.005% level for the 8.2% discount rate, and at the 7.3% level for the 3.2% discount rate.

$\text{Var}(SL_t)/\text{Var}(SL'_t)$ is dramatically different from unity, at 68.91, when the sample mean discount rate is used. Unsurprisingly the variance ratio is smaller when future dividend changes are discounted at the lower rate estimated by the cointegrating regression, but it is still considerable at 4.76. However the asymptotic standard errors on these ratios are huge, and one cannot reject the hypothesis that both of them equal unity.

The innovations variance ratios $\text{Var}(\xi_t)/\text{Var}(\xi'_t)$ are also estimated larger than unity, and here the standard errors are less extreme. In the sample mean discount rate case, one can reject at the 5% level the hypothesis that the innovation variance ratio is unity; it is estimated to be 11.27, with a standard error of 4.49. With the lower discount rate, the ratio is estimated at 1.41, with a standard error of 0.44.

Plots of the price-dividend difference and the unrestricted VAR forecast of dividend changes give a visual image of these variance results. At an 8.2% discount rate, SL_t and SL'_t are negatively correlated and the excess volatility of stock prices is very dramatic. At a 3.2% discount rate SL_t and SL'_t have a correlation of 0.911 and the excess volatility is much less dramatic. For both discount rates however, a regression of ΔY_{t+1} on SL_t gives a coefficient estimate with a negative sign rather than the positive sign implied by the present value model.

To compare our results on volatility with results using earlier methods, we also computed sample values of S^*_t using two alternative terminal conditions: $S^*_T=0$ and $S^*_T=S_T$, where T is the last observation in our sample. We computed SL^*_t analogously. Equation (3) implies $\sigma(S^*_t) > \sigma(S_t)$ and $\sigma(SL^*_t) > \sigma(SL_t)$. For the bond data in the period 1959-78, $\sigma(S^*_t) = 1.269$ for $S^*_T=0$ or 1.217 for $S^*_T=S_T$, while $\sigma(S_t) = 1.060$ so the inequality

is satisfied. For the stock data at an 8.2% discount rate, $\sigma(SL_t^*) = 3.837$ for $SL_T^*=0$ or 7.928 for $SL_T^*=SL_T$, while $\sigma(SL_t) = 15.506$ so the inequality is sharply violated. The inequality is again satisfied by the stock data at a 3.2% discount rate, where $\sigma(SL_t^*) = 11.207$ or 12.888 and $\sigma(SL_t) = 9.937$.

Following Scott [1985], we also regressed S_t^* on S_t and a constant. If the present value model is true, the coefficient on S_t should be one. The same holds for the corresponding regression with SL_t^* and SL_t . For bonds in 1959-78, we estimated the coefficient at 0.77 for $S_T^*=0$ or 0.81 for $S_T^*=S_T$; for stocks at an 8.2% discount rate we estimated it at -0.06 for $SL_T^*=0$ or 0.16 for $SL_T^*=SL_T$, while for stocks at a 3.2% discount rate we estimated it at -0.01 or 0.02. Thus the results using S_t^* and SL_t^* generally support the conclusion that the present value model for bonds fits the data comparatively well, whereas the model for stocks has a poor fit even though it cannot be rejected statistically at high levels of confidence.

We close with two caveats about the plots and summary statistics generated by the VAR system. First, we do not have standard error bands for the plotted VAR forecast or its correlation with the spread. The very high asymptotic standard errors for the levels variance ratios, however, indicate that the VAR forecast is not precisely estimated.²⁸ Thus the favorable visual impression created by a high correlation between the spread and the VAR forecast may be due to chance rather than to real merits of the present

²⁸ We found further evidence for this when we examined other data sets. In one annual data set on postwar U.S. interest rates, we found that the correlation between S and S' was negative. In this case short-run movements in S corresponded to short-run movements in S' , but S' had a distinct downtrend over the sample.

value model.

Secondly, the VAR simulation method may be misleading if the wrong value of θ is chosen so that the spread variable is nonstationary. For example, if θ is chosen too large, the movements of S_t are dominated by the movements of $-\theta y_t$. The VAR results are then approximately those which one would get if one regressed Δy_t and $-\theta y_t$ on lagged values of these variables. It is well known that in finite samples estimates of autoregressive parameters for nonstationary variables are biased downwards, and this problem will afflict the VAR if θ is too large.

In a simple case where y_t follows an AR-1 process with a unit root, and the VAR includes one lag only, one can show that the estimated VAR companion matrix will have first column zero and second column $[(1-\rho)/\theta \quad \rho]'$, where ρ is a downward biased estimate of the unit root. This companion matrix satisfies the restrictions of equation (9) almost exactly, whatever the behavior of the variable Y_t . A symptom of this misspecification would be that mean returns would not obey the model, even though the dynamics of returns would appear to satisfy the restrictions.

It is possible that a problem of this sort affects our results for the stock market when we use a low 3.2% discount rate corresponding to a high θ of 31.092. The cointegrating regression which generates this θ estimate - a regression of the level of Y on the level of y - is dominated by the enormous postwar hump in stock prices. Since this hump coincided with a much milder hump in real dividends, the regression estimates a coefficient for y which is much larger than the historical average price-dividend ratio. The negative intercept prevents the fitted value from overpredicting Y over the sample period as a whole. As a result, over the bulk of the

sample period, the spread SL_t is distinctly negatively correlated with the lagged dividend.²⁹ The VAR estimates place considerable weight on this earlier part of the sample period, because the dividend equation is specified in terms of dividend changes which are more variable before 1946. Thus the high correlation of SL_t and SL'_t may be to some extent spurious. The regression of SL^*_t on SL_t , by contrast, is a levels regression which is dominated by the postwar hump in stock prices; in this regression we find the coefficient to be essentially zero rather than one as required by the model.

²⁹ Over the period 1871 to 1946, the spread has a correlation of -0.7 with the lagged dividend when θ is set equal to 31.092.

4. Conclusions

In this paper we have shown how a present value model may be tested when the variables of the model follow linear stochastic processes which are stationary in first differences rather than in levels. If the present value model is true, a linear combination of the variables is stationary. The paper draws on the theory of cointegrated processes to estimate this linear combination or spread and test the model.

We also propose an informal method for evaluating the "fit" of a present value model. A VAR is used to construct an optimal unrestricted forecast, and this is compared with a restricted forecast from the model. We compute the standard deviations and correlation of the two forecasts, and plot their historical movements.

We apply our methods to the controversial present value models for stocks and bonds. We find that both models can be rejected statistically at conventional significance levels, with much stronger evidence for bonds. However in our data set the spread between long and short term interest rates seems to move quite closely with the present value of expected short rate changes. Deviations from the present value model for bonds do not appear to be highly persistent. In contrast, our evaluation of the present value model for stocks indicates that the spread between stock prices and dividends moves too much, and that deviations from the present value model are quite persistent, although the strength of the evidence for this depends sensitively on the discount rate assumed in the test.

TABLE 1A

ESTIMATION OF THE COINTEGRATING VECTOR
AND TEST FOR COINTEGRATION
IN THE TERM STRUCTURE

Sample 1959-78

$$1) \quad Y_t = 2.364 + 0.735 y_t \quad R^2 = 0.640$$

(0.120) (0.028) Estimate of $\theta = 0.735$

$$2) \quad \Delta y_t = 0.064 - 0.215 \Delta y_{t-1} + 0.391 \Delta Y_{t-1}$$

(0.111) (0.117) (0.154)

$$- 0.061 y_{t-1} + 0.041 Y_{t-1} \quad R^2 = 0.057$$

(0.034) (0.026) Estimate of $\theta = 1.488$

Tests of no cointegration:

Granger and Engle [1985] ξ_2 statistic for equation (1) residual: 2.69.
10% significance level 3.03, 5% level 3.37, 1% level 4.07.

Granger and Engle [1985] ξ_3 statistic for equation (1) residual: 1.96.
10% significance level 2.84, 5% level 3.17, 1% level 3.77.

Dickey and Fuller [1981] Φ_1 statistic for spread: 5.21.

10% significance level 3.81, 5% level 4.63, 1% level 6.52.

TABLE 1A (CONTINUED)

Sample 1959-83

$$3) \quad Y_t = 1.995 + 0.854 y_t \quad R^2 = 0.815$$

(0.133) (0.026)

Estimate of $\theta = 0.854$

$$4) \quad \Delta y_t = 0.061 - 0.198 \Delta y_{t-1} + 0.884 \Delta Y_{t-1}$$

(0.140) (0.119) (0.243)

$$- 0.132 y_{t-1} + 0.102 Y_{t-1} \quad R^2 = 0.112$$

(0.058) (0.053)

Estimate of $\theta = 1.294$

Tests of no cointegration:

Granger and Engle [1985] ξ_2 statistic for equation (3) residual: 4.02.

10% significance level 3.03, 5% level 3.37, 1% level 4.07.

Granger and Engle [1985] ξ_3 statistic for equation (3) residual: 2.74.

10% significance level 2.84, 5% level 3.17, 1% level 3.77.

Dickey and Fuller [1981] ϕ_1 statistic for spread: 10.63.

10% significance level 3.81, 5% level 4.63, 1% level 6.52.

TABLE 1B

ESTIMATION OF THE COINTEGRATING VECTOR
AND TEST FOR COINTEGRATION
IN THE STOCK MARKET

Sample 1871-1986

$$1) \quad Y_t = -12.979 + 31.092 y_t$$

(2.080) (1.268)

$R^2 = 0.842$
Estimate of $\theta = 31.092$
Implied discount rate = 3.2%

$$2) \quad \Delta y_t = 0.101 + 0.165 \Delta y_{t-1} + 0.010 \Delta Y_t$$

(0.033) (0.089) (0.002)

$$- 0.157 y_{t-1} + 0.004 Y_t$$

(0.057) (0.002)

$R^2 = 0.373$
Estimate of $\theta = 37.021$
Implied discount rate = 2.7%

3) Sample mean return = 8.2%
Corresponding estimate of $\theta = 12.195$

Tests of no cointegration:

Granger and Engle [1985] ξ_2 statistic for equation (1) residual: 3.58.
10% significance level 3.03, 5% level 3.37, 1% level 4.07.

Granger and Engle [1985] ξ_3 statistic for equation (1) residual: 2.64.
10% significance level 2.84, 5% level 3.17, 1% level 3.77.

TABLE 2A

TESTS OF PRESENT VALUE MODEL
IN THE TERM STRUCTURESample 1959-78

Akaike Criterion selects 11-lag VAR.

Δy equation $R^2 = 0.216$.

S Granger causes Δy at 0.01% level.

S equation $R^2 = 0.877$.

Δy Granger causes S at 0.3% level.

Test of present value model Chi-Squared(22) = 83.02, P-Value < 0.005%.
Innovations variance ratio $\text{Var}(\xi)/\text{Var}(\xi')$ = 1.160, standard error 1.146

Summary statistics: $E(\Delta y) = 0.016$
 $E(S) = 1.144$
 $E(S') = 0.016$
 $\sigma(\Delta y) = 0.442$
 $\sigma(S) = 1.060$
 $\sigma(S') = 1.067$
 $\rho(S, S') = 0.978$

Coefficient on S in a regression of ΔY on a constant and S: -0.020

Sample 1959-83

Akaike Criterion selects 6-lag VAR.

Δy equation $R^2 = 0.171$.

S Granger causes Δy at 0.3% level.

S equation $R^2 = 0.772$.

Δy Granger causes S at 1.3% level.

Test of present value model Chi-Squared(12) = 35.63, P-Value = 0.03%.
Innovations variance ratio $\text{Var}(\xi)/\text{Var}(\xi')$ = 0.502, standard error 0.506

Summary statistics: $E(\Delta y) = 0.021$
 $E(S) = 1.138$
 $E(S') = 0.021$
 $\sigma(\Delta y) = 0.793$
 $\sigma(S) = 1.320$
 $\sigma(S') = 0.717$
 $\rho(S, S') = 0.956$

Coefficient on S in a regression of ΔY on a constant and S: -0.039

TABLE 2B

TESTS OF PRESENT VALUE MODEL
IN THE STOCK MARKETSample 1871-1986 $\theta = 12.195$ (8.2% discount rate): Akaike Criterion selects 4-lag VAR. Δy equation $R^2 = 0.400$.SL Granger causes Δy at $< 0.001\%$ level.SL equation $R^2 = 0.837$ Δy Granger causes SL at 63.3% level.

Test of present value model with mean restriction:

Chi-Squared(9) = 15.74, P-Value = 7.2%.

Test of present value model without mean restriction:

Chi-Squared(8) = 15.72, P-Value = 4.7%.

Levels variance ratio $\text{Var}(S)/\text{Var}(S') = 68.91$, standard error 77.93Innovations variance ratio $\text{Var}(\xi)/\text{Var}(\xi') = 11.27$, standard error 4.49

Summary statistics:

$E(\Delta y)$	=	0.017
$E(SL)$	=	16.068
$E(SL')$	=	2.563
$\sigma(\Delta y)$	=	0.168
$\sigma(SL)$	=	15.506
$\sigma(SL')$	=	1.891
$\rho(SL, SL')$	=	-0.459

Coefficient on SL in a regression of ΔY on a constant and SL: -0.064 $\theta = 31.092$ (3.2% discount rate): Akaike Criterion selects 2-lag VAR. Δy equation $R^2 = 0.378$.SL Granger causes Δy at $< 0.001\%$ level.SL equation $R^2 = 0.516$. Δy Granger causes SL at 1.8% level.

Test of present value model with mean restriction:

Chi-Squared(5) = 14.90, P-Value = 1.1%.

Test of present value model without mean restriction:

Chi-Squared(4) = 5.75, P-Value = 21.8%

Levels variance ratio $\text{Var}(S)/\text{Var}(S') = 4.76$, standard error 3.85Innovations variance ratio $\text{Var}(\xi)/\text{Var}(\xi') = 1.41$, standard error 0.44

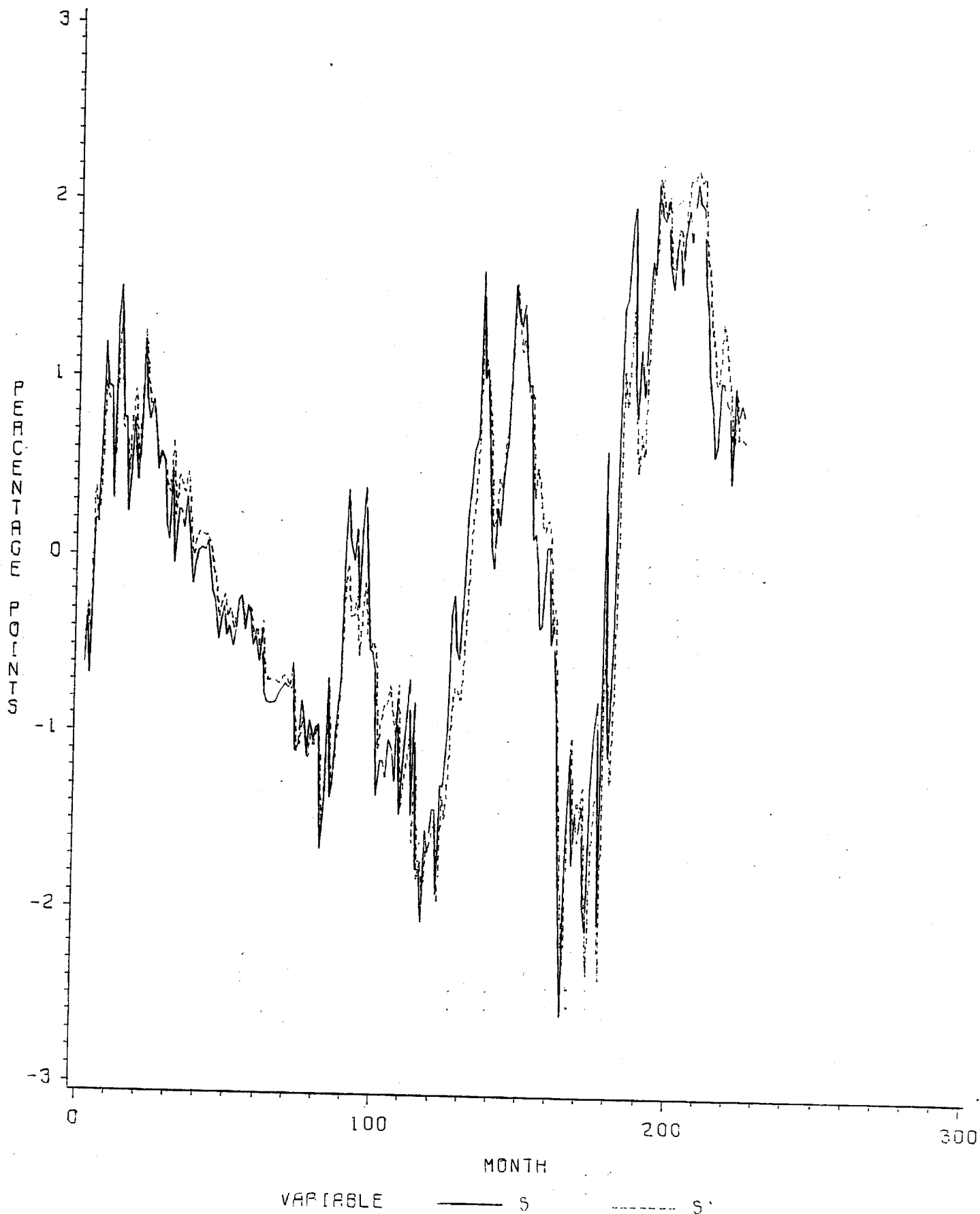
Summary statistics:

$E(\Delta y)$	=	0.017
$E(SL)$	=	-12.524
$E(SL')$	=	16.661
$\sigma(\Delta y)$	=	0.167
$\sigma(SL)$	=	9.937
$\sigma(SL')$	=	4.543
$\rho(SL, SL')$	=	0.911

Coefficient on SL in a regression of ΔY on a constant and SL: -0.079

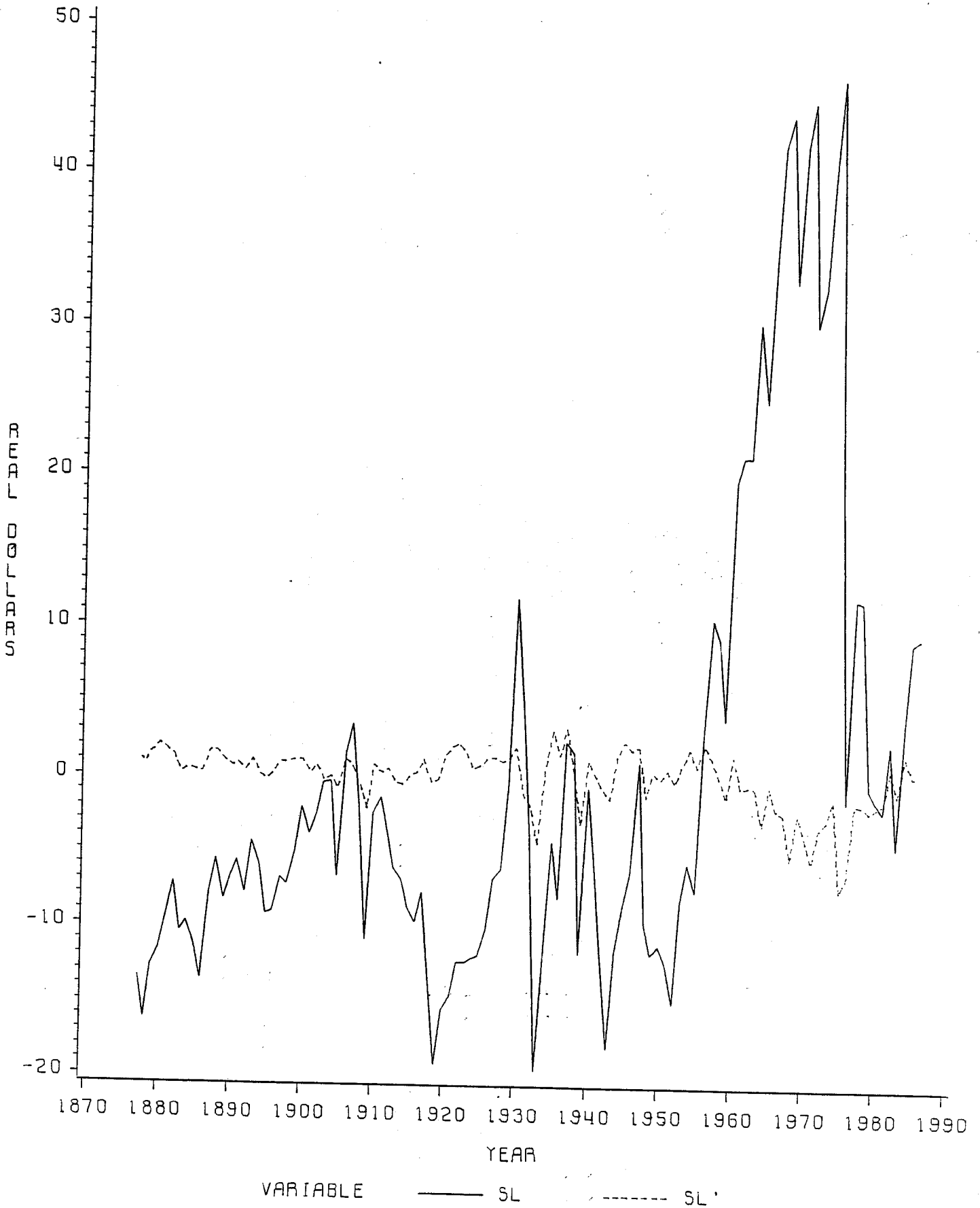
term structure

SHORT SAMPLE



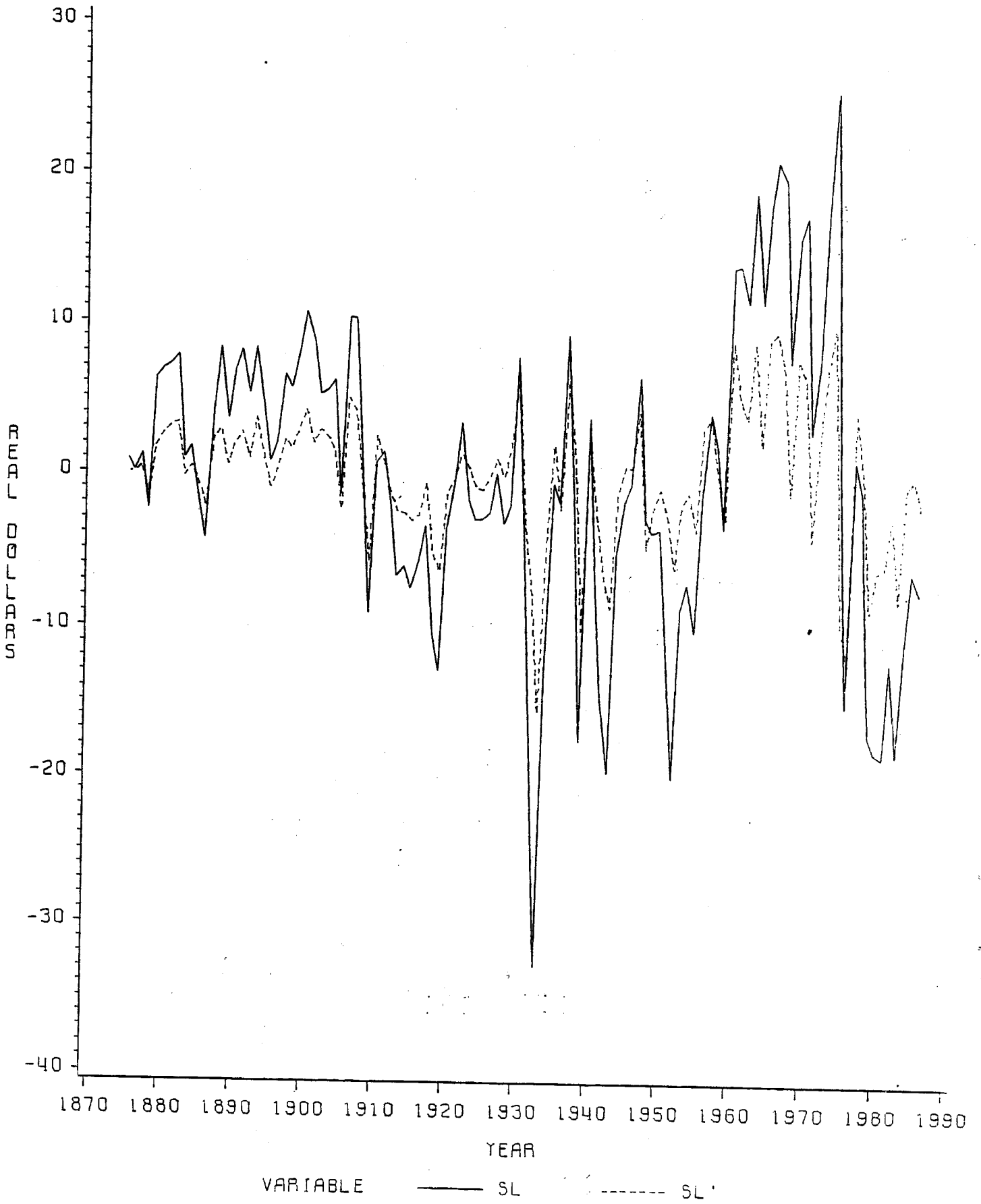
STOCK MARKET

DISCOUNT RATE 8.2%



STOCK MARKET

DISCOUNT RATE 3.2%



BIBLIOGRAPHY

Baillie, Richard T., Robert E. Lippens and Patrick C. McMahon, "Testing Rational Expectations and Efficiency in the Foreign Exchange Market", Econometrica Vol. 51 No. 3 pp. 553-564, May 1983.

Campbell, John Y., "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis", National Bureau of Economic Research Working Paper No. 1805, January 1986.

Campbell, John Y. and Robert J. Shiller, "A Simple Account of the Behavior of Long-Term Interest Rates", American Economic Review Vol. 74 No. 2 pp. 44-48, May 1984.

Cowles, Alfred, Common Stock Indexes, second edition, 1939.

Dickey, David A. and Wayne A. Fuller, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", Econometrica Vol. 49 No. 4 pp. 1057-1072, July 1981.

Engle, Robert F. and Mark Watson, "Applications of Kalman Filtering in Econometrics", paper presented at World Congress of the Econometric Society, August 1985.

Granger, Clive W.J. and Robert F. Engle, "Dynamic Model Specification with Equilibrium Constraints: Cointegration and Error-Correction", University of California at San Diego Discussion Paper 85-18, May 1985.

Gregory, Allan W. and Michael R. Veall, "Formulating Wald Tests of Nonlinear Restrictions", Econometrica Vol. 53 No. 6 pp. 1465-1468, November 1985.

Hamilton, James D. and Charles H. Whiteman, "The Observable Implications of Self-Fulfilling Expectations", Journal of Monetary Economics Vol. 16 No. 3 pp. 353-374, November 1985.

Hansen, Lars Peter and Thomas J. Sargent [1981a], "Exact Linear Rational Expectations Models: Specification and Estimation", Staff Report 71, Federal Reserve Bank of Minneapolis, November 1981.

Hansen, Lars Peter and Thomas J. Sargent [1981b], "Linear Rational Expectations Models for Dynamically Interrelated Variables", in Robert E. Lucas Jr. and Thomas J. Sargent eds., Rational Expectations and Econometric Practice, Minnesota, 1981.

Ito, Takatoshi, "Use of Time-Domain Vector Autoregressions to Test Uncovered Interest Parity", National Bureau of Economic Research Working Paper No. 1493, November 1984.

Kleidon, Allan W., "Variance Bounds Tests and Stock Price Valuation Models", forthcoming Journal of Political Economy 1986.

LeRoy, Stephen F. and Richard D. Porter, "The Present-Value Relation: Tests Based on Implied Variance Bounds", Econometrica Vol. 49 No. 3 pp. 555-574, May 1981.

Mankiw, N. Gregory and Jeffrey A. Miron, "The Changing Behavior of the Term Structure of Interest Rates", National Bureau of Economic Research Working Paper No. 1669, July 1985.

Mankiw, N. Gregory, David Romer and Matthew D. Shapiro, "An Unbiased Reexamination of Stock Market Volatility", Journal of Finance Vol. 40 No. 3 pp. 677-687, July 1985.

Mankiw, N. Gregory and Matthew D. Shapiro, "Trends, Random Walks, and Tests of the Permanent Income Hypothesis", Cowles Foundation Discussion Paper, Yale University, 1984.

Marsh, Terry A. and Robert C. Merton, "Dividend Variability and Variance Bounds Tests for the Rationality of Stock Market Prices", American Economic Review, forthcoming 1985.

Melino, Angelo, Essays on Estimation and Inference in Linear Rational Expectations Models, unpublished PhD dissertation, Harvard University, 1983.

Modigliani, Franco and Robert J. Shiller, "Inflation, Rational Expectations and the Term Structure of Interest Rates", Economica Vol. 40 No. 157 pp. 12-43, February 1973.

Runkle, David, "Vector Autoregression Reality", unpublished paper, Brown University, 1984.

Sargent, Thomas J., "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure", Journal of Monetary Economics Vol. 5 pp. 133-143, January 1979.

Sawa, Takamitsu, "Information Criteria for Discriminating Among Alternative Regression Models", Econometrica Vol. 46 No. 6 pp. 1273-1291, November 1978.

Scott, Louis O., "The Present Value Model of Stock Prices: Regression Tests and Monte Carlo Results", Review of Economics and Statistics Vol. 57 No. 4 pp. 599-605, November 1985.

Shiller, Robert J., "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure", Journal of Political Economy Vol. 87 No. 6 pp. 1190-1219, December 1979.

Shiller, Robert J. [1981a], "Alternative Tests of Rational Expectations Models: The Case of the Term Structure", Journal of Econometrics Vol. 16 pp. 71-87, May 1981.

Shiller, Robert J. [1981b], "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?", American Economic Review Vol. 71 No. 3 pp. 421-436, June 1981.

Shiller, Robert J., "Stock Prices and Social Dynamics", Brookings Papers on Economic Activity 1984:2 pp. 457-498, 1984.

Shiller, Robert J., "Conventional Valuation and the Term Structure of Interest Rates", National Bureau of Economic Research Working Paper No. 1610, April 1985.

Shiller, Robert J., John Y. Campbell and Kermit L. Schoenholtz, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates", Brookings Papers on Economic Activity 1983:1 pp. 173-217, 1983.

Stock, James H., "Asymptotic Properties of Least Squares Estimates of Cointegrating Vectors", unpublished paper, Harvard University, November 1984.

West, Kenneth D., "Speculative Bubbles and Stock Price Volatility", Financial Research Memorandum No. 54, Princeton University, December 1984.

White, Halbert, Asymptotic Theory for Econometricians, Academic Press, 1984.