The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure

Robert J. Shiller

University of Pennsylvania and National Bureau of Economic Research

Models which represent long-term interest rates as long averages of expected short-term interest rates imply, because of the smoothing implicit in the averaging, that long rates should not be too volatile. The volatility of actual long-term interest rates, as measured by the variance of short-term holding yields on long-term bonds, appears to exceed limits imposed by the models. Such excess volatility implies a kind of forecastability for long rates. Long rates show a slight tendency to fall when they are high relative to short rates rather than rise as predicted by expectations models.

I. Introduction

An argument that often seems to be implicit in popular criticisms of rational expectations models of the term structure of interest rates is that long-term interest rates are too "volatile" to accord with the averaging inherent in the models. With these expectations models, the long-term interest rate can be approximately represented as a long average of rationally expected future short-term rates plus a liquidity premium term. Long linear moving averages tend to smooth out the series averaged, and this tendency would seem to extend to

This paper was supported by the Federal Reserve Bank of Philadelphia, while I was visiting scholar there, and by the National Science Foundation under grant #SOC 77-26798. The views expressed here are solely mine and do not necessarily represent the views of the supporting agencies. I am particularly indebted to Franco Modigliani, who first pointed out to me an example of the kind of regression which appears here in table 3. The paper also benefited from discussions with Benjamin Friedman, Gary Gillum, Stephen LeRoy, Richard Roll, and many others. Leslie Appleton and Edward Hendricks provided research assistance.

[*Journal of Political Economy*, 1979, vol. 87, no. 6]

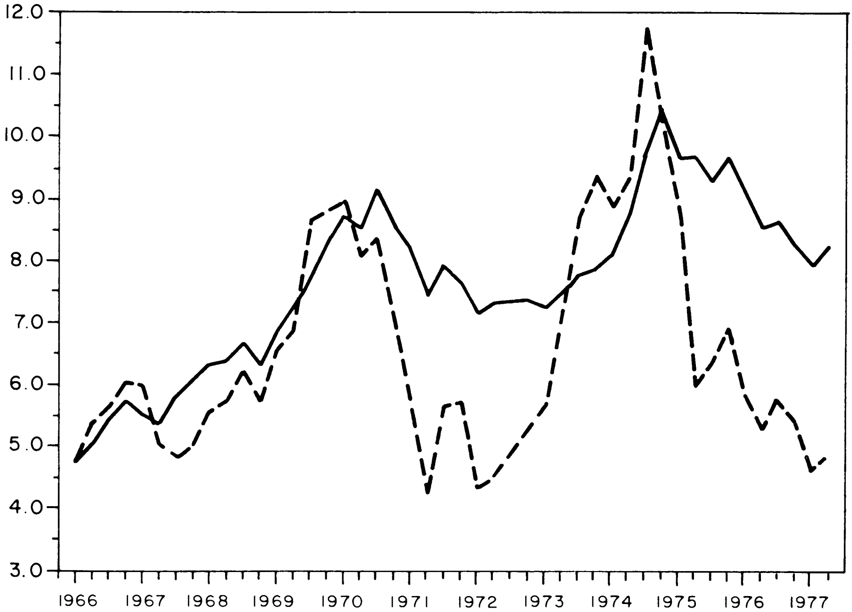

FIG. 1.—U.S. long-term interest (R) (solid line) and short-term interest rate (r) (dotted line) quarterly, 1966:I–1977:II. See Appendix B, data set 1.

nonlinear averaging schemes inherent in alternative versions of the model. In addition, if rational expectations represent a conditional mean or "conditional average," they should tend to change dramatically only when important new information arrives, which could not be too often. This tendency would seem to extend to alternative measures of conditional central tendency which might be used to represent public expectations. It would thus seem that observed volatility of interest rates would have to be ascribed to factors not usually represented in these models. The liquidity premium is usually described as reflecting public attitudes toward and perceptions of risk and is usually assumed constant or modeled as slow moving.

Observed long-term interest rates series are not much smoother than short-rate series, as can be seen, for example, in figure 1, which displays a long-rate series for high-grade bonds with over 20 years to maturity (solid line) and a 4–6-month short-rate series (dotted line).[1] As a result of the choppy behavior of long-term interest rates, the short-term holding yield on long-term bonds, which is related to the percentage change in the long-term interest rate, has a very high variance. Figure 2 shows the approximate annualized one-quarter

[1] This is a plot of data set 1, described in Appendix B.

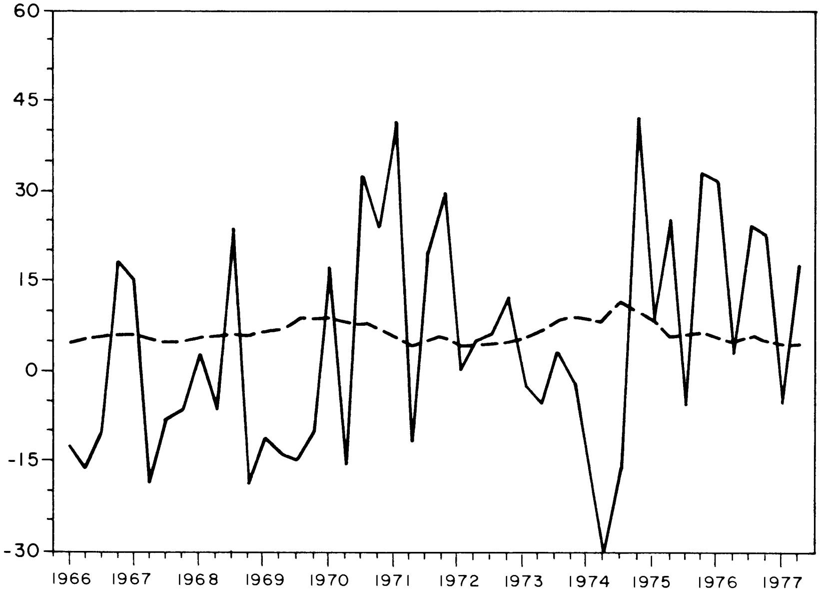

FIG. 2.—Approximate one-quarter holding yield (H) on the long-term bonds from fig. 1 (solid line) and the short-term interest rate (r) (dotted line) quarterly, 1966:I–1977:II. See Appendix B, data set 1, and n. 2.

holding yield on long-term bonds, as computed from the long-term interest rate series from figure 1 (solid line),[2] and the same short-term interest rate (dotted line). Note that the vertical axis in figure 2 has smaller units than that of figure 1, because the one-period holding yield is so volatile. The standard deviation of the holding yield is in fact 18.6 percentage points, and the holding yield ranges in this sample from −29 percent to +42 percent. Recent U.K. data also show great short-term holding-yield volatility. For British Consols, for which the time to maturity is infinite, the standard deviation of the annualized quarterly holding yield from 1956:I to 1977:II is 25.8 percentage points, and the holding yield ranges from −53 percent to +108 percent.[3] Culbertson (1957), in his well-known critique of expectations models of the term structure, remarked in connection with

[2] Computed using expression (5), below, where $R_t^{(n)}$ is the long-term interest rate in fig. 1 (divided by 400) and $R_{t+1}^{(n-1)}$ is the long-term interest rate for the following quarter (divided by 400), and C is taken as $R_t^{(n)}$. After computation, the holding yield is remultiplied by 400 to convert to annual percent. See n. 9 below concerning approximation error in this measure of holding yield.

[3] Computed from the expression in n. 7 below without the approximation error referred to in n. 2, using data set 5 described in Appendix B. The range divided by standard deviation is high, indicating heteroscedasticity as discussed below.

a graph of holding yields like our figure 2, "what sort of expectations, one might ask, could possibly have produced this result?"[4]

My purpose here is to develop the robust properties of a broad class of expectations models for the random behavior of long-term interest rates, and to examine whether observed rates are too volatile to accord with these models. LeRoy and Porter (1979) have independently shown some ways of evaluating such a model in the context of their analogous claim that stock prices are too volatile to accord with a model which makes stock prices equal the present discounted value of expected earnings. I have also found support (Shiller 1979) for a similar claim regarding common stock price volatility using methodology that draws on this paper. The volatility of bond prices is in a sense more basic an issue than that of stock prices, since (high-quality) bond-price movements depend only on variations in the rate of discount, whereas stock prices can change due either to a change in the rate of discount or to a change in expected earnings.

Any claim that long-term interest rates (or, in the case of the other papers, stock prices) are too volatile contradicts a large literature confirming "market efficiency" or "random-walk" behavior of certain sequences. The claim that bond markets are efficient or that long-term interest rates are approximately a random walk was made, for example, by Granger and Rees (1968), Bierwag and Grove (1971), Laffer and Zecher (1975), Phillips and Pippenger (1976), Sargent (1976, 1979), Mishkin (1978), and Pesando (1978). However, we shall see that even if long-term interest rates are much too volatile, conventional tests of "market efficiency" may be weak. Other studies have, in any event, provided evidence contrary to some forms of simple models of bond-market efficiency (Roll 1970; Sargent 1972; Cargill 1975; and Fama 1976), so the claim that market efficiency fails due to excess volatility should not be regarded as highly improbable based on past literature.

In Section II below I will discuss the basic linearized model and show how it can be derived as an approximation from a number of different versions of the expectations model. In Section III I will derive some inequality restrictions which show in what sense the expectations model implies long-term interest rate series must be stable and smooth, and I will contrast this with observed behavior of long rates.

In Section IV I test some market-efficiency restrictions by running simple regressions. If long rates are too volatile to accord with the

[4] See Culbertson (1957, p. 508). Culbertson, however, did not clearly state the argument we make here.

theory, then they must also be forecastable in a way inconsistent with the theory, and this is tested here. Implications of the model for the spectral density of interest rates are derived in Appendix A.

II. A Linearized Expectations Model

The linearized model we shall study relates the n-period interest rate (yield to maturity on n-period bonds) $R_t^{(n)}$ to a weighted average of expected future one-period (short-term) interest rates $r_t, r_{t+1}, \ldots$:

$$R_t^{(n)} = \frac{1-\gamma}{1-\gamma^n} \sum_{K=0}^{n-1} \gamma^K E_t(r_{t+K}) + \Phi_n, \tag{1}$$

where γ is a constant $0 < \gamma < 1$ and Φ_n is a constant "liquidity premium."[5] Here, E_t is the expectations operator conditional on information available at time t, which includes all current and lagged interest rates. Linear models relating long rates to expected future short rates should be familiar to most readers; however, the precise form I have chosen here requires some explanation, since the form is important for the analysis which follows. Most simple linear term structure models relate long-term interest rates to an unweighted simple average of expected short rates. Those models are indeed appropriate for pure (no coupon) discount bonds. In contrast, this equation involves weights which describe a truncated exponential (or "Koyck") distribution scaled so that the sum of the coefficients is one. Expected short-term interest rates in the near future carry more weight in determining the long yield than do expected short-term interest rates in the more distant future. We will set $\gamma = 1/(1 + \bar{R})$, and then (1) relates $R_t^{(n)}$ to the "present value" of future short-term interest rates discounted by $\bar{R}$. This model is intended for coupon-carrying bonds which are selling near par, or for consols with $n = \infty$. Since longer-term bonds which are available do carry coupons, and since yield series for bonds selling near par or for consols are available, this expression suits our purposes.

Note that our model is not specific to a particular time interval chosen, for example, whether quarterly or annual. If the model holds for a given time interval, then it holds for a higher time interval as well. By regrouping terms in (1) where m is replaced by mn and then using (1) again where n is replaced by m, it is easily verified that:

$$R_t^{(mn)} = \frac{1-\hat{\gamma}}{1-\hat{\gamma}^n} \sum_{K=0}^{n-1} \hat{\gamma}^K E_t[R_{t+Km}^{(m)}] + \hat{\Phi}_n,$$

[5] Throughout this paper, superscripts are distinguished from exponents by superfluous parentheses. Here $r_t = R_t^{(1)}$. I use lowercase r_t to denote the one-period rate, for notational convenience. Later, I shall use unsuperscripted uppercase R to denote the perpetuity rate, i.e., $R_t = R_t^{(\infty)}$.

where $\hat{\gamma} = \gamma^m$ and $\hat{\Phi}_n = \Phi_{mn} - \Phi_m$. Hence, mn-period rates are related to a truncated Koyck average of expected m-period rates over the next n time intervals, where time intervals are m periods long.

Expression (1) is a linearization of any of a number of versions of the rational expectations model as applied to bonds which carry coupons and mature in n periods. To see this, we begin with a few definitions. For coupon-carrying bonds, for which we normalize the principal at maturity at 1.00, and for which coupon rate per period is denoted by C, the present value $V_t^{(n)}$ of future coupons and principal is defined by:

$$V_t^{(n)} \equiv C \sum_{K=0}^{n-1} \prod_{j=0}^{K} (1 + r_{t+j})^{-1} + \prod_{j=0}^{n-1} (1 + r_{t+j})^{-1} \equiv V^{(n)}(\vec{r}_t), \tag{2}$$

where $\vec{r}_t$ is defined as the vector $(r_t, r_{t+1}, \ldots, r_{t+n-1})$, and $V^{(n)}(\cdot)$ will refer to the present value function here defined. I have assumed here that coupons are paid once per period starting after one period. The first term in the expression is the present value of the stream of coupon payments; the second term is the present value of the payment of principal ($1.00) at maturity.

The yield to maturity or long-term interest rate $R_t^{(n)}$ on an n-period bond is determined by the requirement that the price $P_t^{(n)}$ of the bond is the present value of coupons and principal discounted by $R_t^{(n)}$; that is,[6]

$$P_t^{(n)} = V^{(n)}[R_t^{(n)}] = \frac{C}{R_t^{(n)}} + \frac{R_t^{(n)} - C}{R_t^{(n)}[1 + R_t^{(n)}]^n}. \tag{3}$$

If $R_t^{(n)} = C$, $P_t^{(n)} = 1$, and conversely. Such bonds, whose price today equals the principal paid at maturity, are selling "at par." Since new bonds are issued at par, our newly issued and recently offered bond yield averages refer to such bonds.

The one-period holding yield $H_t^{(n)}$ is equal to the capital gain $P_{t+1}^{(n-1)} - P_t^{(n)}$ (note than an n-period bond at time t becomes an $[n - 1]$-period bond at time $t + 1$) plus the coupon payment C at the end of the period divided by the price $P_t^{(n)}$ at time t (to convert to a rate of return):[7]

$$H_t^{(n)} = \frac{P_{t+1}^{(n-1)} - P_t^{(n)} + C}{P_t^{(n)}}. \tag{4}$$

[6] The $V^{(n)}[R_t^{(n)}]$ refers to the function $V^{(n)}$ in which $R_t^{(n)}$ is substituted for each of $r_t, r_{t+1}, \ldots, r_{t+n-1}$; $R_t^{(n)}$ is the single real positive root to eq. (3).

[7] It has been pointed out that if numerator and denominator in the ratio (4) are jointly normally distributed, the ratio will not have a finite variance. One should not be misled by this fact. If the mean of the denominator is large relative to its standard deviation, the distribution function of the ratio approximates the normal. Anyway, the denominator cannot be normally distributed since price cannot be negative.

This may be rewritten, using (3), in terms of the yields to maturity $R_t^{(n)}$ and $R_{t+1}^{(n-1)}$:[8]

$$H_t^{(n)} = \left\{C + \frac{C}{R_{t+1}^{(n-1)}} + \frac{R_{t+1}^{(n-1)} - C}{R_{t+1}^{(n-1)}[1 + R_{t+1}^{(n-1)}]^{n-1}}\right\} \Big/ \left\{\frac{C}{R_t^{(n)}} + \frac{R_t^{(n)} - C}{R_t^{(n)}[1 + R_t^{(n)}]^n}\right\} - 1. \tag{5}$$

The simplest way to motivate expression (1) is to consider a model which relates the expected one-period holding yield to the short-term interest rate:

$$E_t[H_t^{(n)}] = r_t + \phi^{(n)}, \tag{6}$$

where $\phi^{(n)}$ is a constant. In the one-period Sharpe-Lintner mean-variance capital asset pricing model, as applied to the bond market by Roll (1971), McCallum (1975), and Friend, Westerfield, and Granito (1978), $\phi^{(n)}$ will equal $\beta^{(n)}[E(R_m) - r]$, where R_m is the return on the market portfolio and $\beta^{(n)}$ is equal to the covariance between $H_t^{(n)}$ and the return on the market portfolio divided by the variance of the return on the market portfolio. The capital asset pricing model itself does not necessarily imply that $\phi^{(n)}$ is constant, although the usual tests of the capital asset pricing model as in Friend and Blume (1970) or Black, Jensen, and Scholes (1972), as well as the above cited applications to the bond market, assume this is true, at least over certain time intervals. If we make this assumption, then substituting (5) into (6) gives us a first-order nonlinear rational expectations model relating $R^{(n)}$ and r. The nonlinearities, however, create fundamental problems. Our approach will be to linearize expression (5) around $R_t^{(n)} = R_{t+1}^{(n-1)} = \bar{R} = C$ (i.e., take a Taylor expansion truncated after the linear term) to give us a linearized holding yield $\tilde{H}_t^{(n)}$ which will approximate $H_t^{(n)}$. This procedure gives us:[9]

[8] If the bond is a perpetuity, $n = \infty$ and expression (5) reduces to $H_t = R_t - \Delta R_{t+1}/R_{t+1}$.

[9] The approximation error introduced by the linearization (7) is not large. The correlation coefficients between H and $\tilde{H}$ for data sets 1–6 as computed in table 1 are .993, .994, .990, .997, .947, and .978, respectively. Models which describe pure discount bonds and which equalize expected log holding yields seem to avoid the necessity for such approximation. That apparent advantage is illusory, however, for if these models are to be similarly robust to variations in assumptions, as discussed below, the same sort of linearization arguments must be made. Another approximation error, which affects our estimated holding-period yields based on newly issued or recently offered yield series, even without the linearization, is introduced by my practice of substituting the yield average at time t for $R_t^{(n)}$ and the yield average at time $t + 1$ for $R_{t+1}^{(n-1)}$ in expression (5) or (7). The problem with this practice is that the maturity date and coupon are not kept constant from period to period in the yield averages, and, in fact, coupons roughly equal current yields. The error introduced by failing to keep the maturity date constant is certainly negligible. There is no measurable difference in yield between, e.g., 25-year

$$\tilde{H}_t^{(n)} = \frac{R_t^{(n)} - \gamma_n R_{t+1}^{(n-1)}}{1 - \gamma_n}, \tag{7}$$

where $\gamma_n = \{1 + \bar{R}[1 - 1/(1 + \bar{R})^{n-1}]^{-1}\}^{-1} = \gamma(1 - \gamma^{n-1})/(1 - \gamma^n)$. Substituting this expression for $\tilde{H}_t^{(n)}$ in place of $H_t^{(n)}$ in (6) and rearranging gives

$$R_t^{(n)} = \gamma_n E_t[R_{t+1}^{(n-1)}] + (1 - \gamma_n)[r_t + \phi^{(n)}],$$

which is a first-order linear rational expectations difference equation in $R_t^{(n)}$ with variable coefficients. Such a model may be solved using familiar methods in the rational expectations literature—as surveyed, for example, in Shiller (1978)—by a method of recursive substitution and with a terminal value condition for the maturity date. To do this, one merely substitutes in place of $E_t R_{t+1}^{(n-1)}$ in the above expression the expected value of the expression obtained by replacing t by $t + 1$ and n by $n - 1$. After doing this, one then replaces $E_t[R_{t+2}^{(n-2)}]$ in the resulting expression, and so on. The resulting solution (involving a terminal value condition that $R_{t+n-1}^{[1]} = r_{t+n-1}$, or that, in effect, the price of the bond is 1.00 at $t + n$) is expression (1) with

$$\Phi_n = \frac{1 - \gamma}{1 - \gamma^n} \sum_{K=0}^{n-1} \gamma^K \phi^{(n-K)}. \tag{8}$$

The model (1) is thus a consequence of the capital asset pricing model under our linearization assumption coupled with the additional assumption introduced by the terminal condition. The capital asset pricing model itself is not (in contrast to model [1]) invariant to changes in the time interval chosen, that is, to the investment horizon of the representative investor. Roll (1971), in his capital asset pricing model of the bond market, tried to find from the data what is the "representative" investment horizon of investors. My linearization sidesteps this problem.

The linear expression (1) may also serve as an approximation to a number of other versions of the expectations model of the term structure, so long as interest rates do not vary too much. The model discussed above, in which expected one-period holding yields $H_t^{(n)}$ equal the short rate plus a constant, may be written as $P_t^{(n)} = E_t V^{(n)}(\vec{r}_t + \vec{\phi})$, where $\vec{\phi} = [\phi^{(n)}, \phi^{(n-1)}, \ldots, \phi^{(1)}]$. An alternative model is one in which forward rates equal expected future spot rates plus a liquidity

---

bonds and 25¼-year bonds. A bigger error is introduced by the fact that the coupon rate is not kept constant between t and $t + 1$ in the yield averages. There is a relationship between coupon and yield for individual bonds, as a study by Shiller and Modigliani (1979) concluded. The relationship, which appears to be due to the differential taxation of capital gains versus income, works in the direction of causing our measures H and $\tilde{H}$ to slightly overstate actual holding yields by, in effect, purifying our series from tax effects.

premium, which can be written as $P_t^{(n)} = V^{(n)}(E_t\vec{r}_t + \vec{\phi})$. Other models include a model in which the expected total return from holding a bond n periods (reinvesting coupons at the short rate) equals the expected total return from investing in a sequence of shorts for n periods,

$$P_t^{(n)} = E_t\left[V^{(n)}(\vec{r}_t)\prod_{i=0}^{n-1}(1 + r_{t+i})\right] \Big/ E_t\left[\prod_{i=0}^{n-1}(1 + r_{t+i})\right],$$

or a model in which the yield $R_t^{(n)}$ equals the expected yield from holding a sequence of shorts which are liquidated in a manner which matches the coupon principal payout structure of par long-term bonds, $P_t^{(n)} = V^{(n)}[R_t^{(n)}]$, where

$$R_t^{(n)} = E_t\left\{\left[1 - \prod_{i=0}^{n-1}(1 + r_{t+i})^{-1}\right] \Big/ \sum_{K=0}^{n-1}\prod_{j=0}^{K}(1 + r_{t+j})^{-1}\right\}.^{10}$$

All of these models have the common property that if the expectations operator is replaced by the number 1, the models reduce to $P_t^{(n)} = V^{(n)}(\vec{r}_t + \vec{\phi})$ for some vector $\vec{\phi}$; that is, under perfect certainty they reduce (except for the liquidity premium) to the present-value formula itself. One can readily verify that if the present-value expression (2) is linearized around $\vec{r}_t = \bar{R} = C$, we get:

$$V_t^{(n)} \cong 1 - \sum_{K=0}^{n-1}\left(\frac{1}{1 + \bar{R}}\right)^{K+1}(r_{t+K} - \bar{R}).$$

If the variation in interest rates is not too large, all of the above models can thus be written as[11]

$$P_t^{(n)} \cong 1 - \sum_{K=0}^{n-1}\left(\frac{1}{1 + \bar{R}}\right)^{K+1}[E_t r_{t+K} + \phi^{(n-K)} - \bar{R}].$$

[10] The four special cases considered here are the four versions of the rational expectations model of the term structure suggested by Cox, Ingersoll, and Ross (1977), although their analysis considered only pure discount bonds. With coupon bonds, the last model mentioned is a little more difficult to understand, since one cannot ex ante plan such a sequence of shorts. Ex post, one can see how one could have invested \$1.00 at time t in shorts, withdrawn C dollars at each subsequent period, and be left with \$1.00 at maturity. This model sets $R_t^{(n)}$ to the ex ante expectation of the C which will achieve this.

[11] All of these models can be written in the form $P_t^{(n)} = g^{(n)}(E_t, \vec{r}_t)$ such that $g^{(n)}(1, \vec{r}_t) = V^{(n)}(\vec{r}_t + \vec{\phi})$. They can then be linearized using a two-step procedure. First, one linearizes all subexpressions in $g^{(n)}(E_t, \vec{r}_t)$ that do not contain (but are premultiplied by) E_t around $r_t + \phi^{(n)} = r_{t+1} + \phi^{(n-1)} = \ldots = r_{t+n-1} + \phi^{(1)} = \bar{R} = C$. The expectations operator which premultiplies the expressions can then be brought inside (using the distributive law), yielding an approximate expression for P_t^n in $E_t[r_t + \phi^{(n)}]$, $E_t[r_{t+1} + \phi^{(n-1)}] \ldots E_t[r_{t+n-1} + \phi^{(1)}]$. The second step is to linearize this approximate expression for $P_t^{(n)}$ around $E_t[r_t + \phi^{(n)}] = E_t[r_{t+1} + \phi^{(n-1)}] = \ldots = E_t[r_{t+n-1} + \phi^{(1)}] = \bar{R} = C$. This yields an expression which is linear in $E_t(\vec{r}_t + \vec{\phi})$. Since we know that $g^{(n)}(1, \vec{r}_t) = V^{(n)}(\vec{r}_t + \vec{\phi})$, we know that this linearization evaluated at $E_t = 1$ is just the linearization of $V^{(n)}(\vec{r}_t + \vec{\phi})$ around $r_t + \phi^{(n)} = r_{t+1} + \phi^{(n-1)} = \ldots = r_{t+n} + \phi^{(1)} = \bar{R} = C$. Hence, our linearized $g^{(n)}(E_t, \vec{r}_t)$ must be E_t times the linearized $V^{(n)}(\vec{r}_t + \vec{\phi})$.

Similarly, the expression (3) for yield can be linearized around $R_t^{(n)} = \bar{R} = C$ as

$$P_t^{(n)} \cong 1 - \frac{1}{\bar{R}}\left[1 - \frac{1}{(1+\bar{R})^n}\right][R_t^{(n)} - \bar{R}].$$

Equating the above two expressions and solving for $R_t^{(n)}$ yields expression (1) with $\gamma = 1/(1 + \bar{R})$ and Φ as given by expression (8). Expression (1) is thus an accurate characterization of all these rational expectations models of the term structure whenever the variation in short-term interest rates is not too large. Our results below concerning the relative volatility of long rates when compared to that of short rates will be robust characterizations of all these models whenever the level of interest rate volatility is not too high.

III. The Volatility of Interest Rates in the Linearized Expectations Model

We will assume now for ease of exposition that the long-term bond under consideration is a perpetuity, that is, $n = \infty$, so that $\gamma_n = \gamma$, and we will drop superscripts for R and H. Results here can be routinely extended to the case of finite maturity bonds.

The simplest (albeit unrealistic) assumption one can make regarding the formation of expectations is that there is perfect knowledge of future one-period rates of interest and hence perfect knowledge of future long-term rates as well. Then, $R_t = R_t^*$, where R_t^* is an "ex post rational rate" analogous to that defined in Shiller and Siegel (1977) given by:

$$R_t^* = (1 - \gamma) \sum_{K=0}^{\infty} \gamma^K r_{t+K}, \tag{9}$$

which is just expression (1) for $n = \infty$ and $\Phi = 0$ where the expectations operator has been dropped. Here R_t^* is a weighted moving average of r_t.

In figure 3 I have plotted the ex post rational rate R^* based on the short-rate series shown in figure 1 and the assumption that $\gamma = .98$ and R^* at the end of the sample equals the average short rate over this sample. That is, I used $R_t^* = \gamma R_{t+1}^* + (1 - \gamma)r_t$, working backward from the terminal value of R^*. One notes the dramatically reduced amplitude for this long-rate series R^* compared with that actually observed for the long-rate series R, and that the short-run movements in the long rate R that we observed in figure 1 seem totally absent from R^*. This is entirely as we would expect, since we know (see Appendix A) that the moving average in (9) reduces cycles of wave

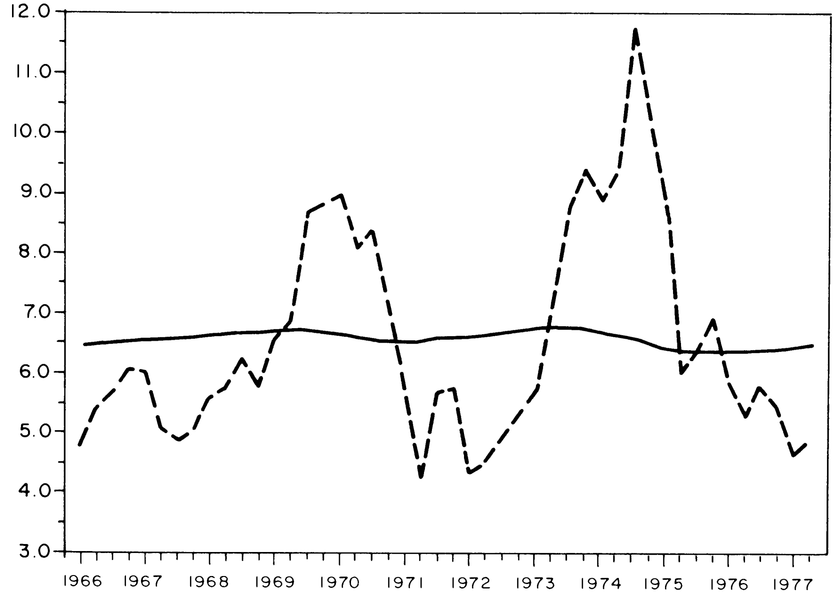

FIG. 3.—Ex post rational long-term interest rate (R^*) (solid line), or what long rates would have been if bonds were priced at present value with actual subsequent short rates (r) (dotted line). See expression (9), text; short rate from fig. 1.

length 5–6 years by a factor of about .08 and of very short wave lengths by a factor of about .01.[12]

One might also note a curious aspect of the behavior of the long-term interest rate in figure 3. Whenever the long rate is above the short rate it is rising, and whenever it is below it, it is falling. The greater the spread between the long rate and the short rate the greater the rate of increase of the long rate. The reason for this

[12] Figure 3 is based on a perpetuity assumption, but the basic result on the smoothing of long cycles carries over to finite maturity bonds. Based on the gain of the filter in expression (1) for $n = 100$ quarters and $\gamma = .98$ (appropriate for this data set), we find that for frequencies in the vicinity of 5 years amplitude is reduced by a factor of .065–.085, roughly as illustrated in fig. 3. The reason for the similarity is clear: 100 quarters is close enough to infinity ($.98^{100} = .13$) that the bond is approximately a consol. Since R_t^* varies little, the linearization (7) which underlies (9) is quite accurate. The exact yield of a consol whose price is given by the present value formula $V(\vec{r} + \vec{\phi})$ (rather than our linearized approximation) under the perfect-certainty assumption and an assumption about Φ consistent with the $\bar{R}$ chosen in fig. 3 looks virtually indistinguishable, differing (except for Φ) by no more than one basis point throughout from the series plotted here. If the linearization (7) is not accurate enough over the range that actual R varies, then that fact itself is a disconfirmation of the model, not of our use of the linearization in describing the model. If we had chosen a higher terminal value for R^*, then we would in effect add an exponential trend to the R^* plotted in fig. 3, and fig. 3 would then represent deviations from the trend. Thus, R^* is smooth regardless of our assumptions about interest rates beyond the sample.

behavior is not hard to find. Whenever the long rate is above the short rate, the long bond has a higher current yield (coupon divided by price) which must be offset by an expected capital loss if expected holding-period returns are to be equalized. A capital loss of course requires an increase in long rates. Conversely, when long rates are low relative to short rates, there must be an expected capital gain, that is, a decline in long rates. If one compares this behavior in figure 3 with the actual behavior of the long rate in figure 1, one again sees a striking difference.

These striking differences between the behavior of the long rate implied by (9) and that actually observed suggest that the model (1) is incorrect too. However, we shall see that the inclusion of the expectations operator in the model (1) causes nontrivial complications, and when these are taken into account the case for excess volatility is not as simple to prove as one might have expected.

The model (1) makes $R_t = E_t R_t^* + \Phi$. This means the forecast error $R_t^* + \Phi - R_t$ must be uncorrelated with information known at time t, which includes all current and lagged interest rates:

$$E[(R_t^* + \Phi - R_t) \cdot R_{t-\tau}] = 0, \qquad \tau \geqslant 0;$$
$$E[(R_t^* + \Phi - R_t) \cdot r_{t-\tau}] = 0, \qquad \tau \geqslant 0. \tag{10}$$

Here, the t subscript on the expectations operator has been dropped since these are unconditional expectations. Implications of these restrictions for the spectrum of the bivariate process (R_t, r_t) are discussed in Appendix A. One might think of testing these restrictions by regressing $R_t^* - R_t$ onto a constant and R_t, R_{t-1}, R_{t-2}, and r_t, r_{t-1}, r_{t-2}, . . . (or any subset of these). The theoretical regression coefficients (except for the intercept) must be zero. It should be obvious that, with the data plotted in the figures, not all coefficients would be zero. Since R_t^* is very stable, and R_t very volatile, movements in $R_t^* - R_t$ correspond closely to movements in R_t, and hence $R_t^* - R_t$ and R_t would show correlation approaching -1. The residuals in such regressions are, however, serially correlated so ordinary significance tests are not valid. Along lines suggested by generalized least squares, the data may be transformed to eliminate the serial correlation by subtracting γ times the led value from the current value of all variables. Such a "generalized least-squares" regression would then amount to regressing $\tilde{H}_t - r_t$ onto transformed right-hand variables.[13] Except for a

[13] We can write $R_t^* = G(F)r_t$, where F is the forward operator defined by $F^K r_t = r_{t+K}$, and $G(F)$ is a polynomial in the forward operator. Then, $G(F) = (1 - \gamma)/(1 - \gamma F)$. One can invert the polynomial and one finds $r_t = G(F)^{-1}R_t^* = [(1 - \gamma F)/(1 - \gamma)]R_t^*$. The linearized holding-period yield $\tilde{H}_t$ can be written $\tilde{H}_t = [(1 - \gamma F)/(1 - \gamma)] R_t$, and hence $R_t = [(1 - \gamma)/(1 - \gamma F)]\tilde{H}_t$. Therefore, $(R_t - R_t^*) = [(1 - \gamma)/(1 - \gamma F)](\tilde{H}_t - r_t)$. Since by our model $(\tilde{H}_t - r_t)$ is a forecast error whose lagged values are known at time t, $\tilde{H}_t - r_t$

constant due to Φ, $\tilde{H}_t - r_t$ is, like $R_t^* - R_t$, a forecast error which should be uncorrelated with all information at time t but which, unlike $R_t^* - R_t$, is serially uncorrelated since its lagged value is known at time t. There is, however, still one important difference between this model and the usual generalized least-squares model, namely, the residual $\tilde{H}_t - r_t$ is uncorrelated with all current and past, but not future, interest rates, and hence the variables $R_t - \gamma R_{t+1}$ and $r_t - \gamma r_{t+1}$ cannot be included in the regression. The correct procedure to test the model is then to regress $\tilde{H}_t - r_t$ onto a constant and $R_{t-1} - \gamma R_t, R_{t-2} - \gamma R_{t-1}, \ldots$, and $r_{t-1} - \gamma r_t, r_{t-2} - \gamma r_{t-1}, \ldots$, or, just onto a constant and $R_t, R_{t-1}, \ldots, r_t, r_{t-1}, \ldots$, and do a significance test such as a conventional F-test on the coefficients. I will perform tests along these lines in Section IV below. For the remainder of this section, we will concern ourselves instead with the implications of (10) for the volatility of interest rates.

The first implication of (10) for interest rate volatility has already been suggested. If a regression of $R_t^* - R_t$ onto a constant and R_t is to yield a zero coefficient for R_t, then var(R) must be less than var(R^*), as Shiller (1972) and LeRoy and Porter (1979) noted. Since $(R_t^* - R_t)$ must be uncorrelated with R, $\text{var}(R_t^*) = \text{var}(R_t) + \text{var}(R_t^* - R_t)$. Moreover, since R_t^* is a moving average of r_t, it must have a smaller variance than r_t, and hence $\text{var}(R_t) \leqslant \text{var}(R_t^*) \leqslant \text{var}(r)$.

This means that the model (1) must imply an even smaller amplitude for the long-term interest rate than that predicted by equation (9). It seems clear, then, that the large amplitude of movements of the long rate as observed in figure 1 could not be reconciled with the behavior of short rates if short rates are expected to swing up and down in the future with repeated episodes roughly like those observed in the cycles from 1967 to 1971 and 1972 to 1976. We cannot rule out, however, that other behavior of the short rate is expected, unless we model the stochastic behavior of the short rate. If r is expected to have some very slow (long-cycle) movements in the future, then these movements will not be effectively reduced by the moving average (9), so R^* and R may yet have a fairly high variance.

The inequality $\text{var}(R) < \text{var}(R^*)$ puts limits on the total amplitude of the long-rate series. It does not tell us whether the long-rate series need be a "smooth" series. Intuition would suggest that the smoothness observed in figure 3 should extend, at least in some sense, to the model (1). We can show that this is the case by finding an upper bound to the variance of the linearized holding-period yield $\tilde{H}$ or of $\tilde{H} - r$. In doing this, we make no assumptions, it should be emphasized,

---

is serially uncorrelated, and hence the generalized least-squares transformation yields serially uncorrelated residuals. This also implies that $\text{var}(R_t - R_t^*) = [(1-\gamma)^2/(1-\gamma^2)]$ var $(\tilde{H}_t - r_t)$, as was noted by LeRoy and Porter (1979) and formed the basis of the test by which they rejected their model of stock prices.

about the nature of the random processes or the information used in forecasting, such as the bivariate ARIMA forecasts assumed by LeRoy and Porter (1979). We assume only (10) and that processes are stationary.

From (10) and our discussion above, we know that $\mathrm{cov}(\tilde{H}_t - r_t, R_t) = 0$. Using the definition (7) of $\tilde{H}_t$ for $n = \infty$ (so that $\gamma_n = \gamma$), we then see that:

$$\mathrm{cov}(R_{t+1}, R_t) = \frac{1}{\gamma}\mathrm{var}(R_t) - \frac{(1-\gamma)}{\gamma}\rho_{rR}\sqrt{\mathrm{var}(R_t)}\sqrt{\mathrm{var}(r_t)}, \tag{11}$$

where ρ_{rR} is the correlation coefficient between r_t and R_t. We then take the expression for the variance of the holding yield $\tilde{H}_t$:[14] $\mathrm{var}(\tilde{H}_t) = \mathrm{var}[(R_t - \gamma R_{t+1})/(1-\gamma)] = [(1+\gamma^2)\,\mathrm{var}(R_t) - 2\gamma\mathrm{cov}(R_t, R_{t+1})]/(1-\gamma)^2$ and substitute into this expression the expression (11) for $\mathrm{cov}(R_t, R_{t+1})$ and maximize (by differentiating and setting to zero) the resulting expression with respect to $\mathrm{var}(R_t)$. The second derivative must be negative, since our model implies $\rho_{rR} > 0$. We find that the maximum is $V_{\tilde{H}} = \mathrm{var}(r)\rho_{rR}^2/(1-\gamma^2)$. Since positive semidefiniteness requires that $\rho_{rR}^2 \leqslant 1$, our model then implies that:

$$\sigma(\tilde{H}_t) \leqslant a\sigma(r_t), \tag{I.1}$$

where $a = (1-\gamma^2)^{-1/2}$ and σ denotes standard deviation. Since $\rho_{rR}^2\mathrm{var}(r) = \mathrm{var}(\hat{r})$, where $\hat{r}$ is the fitted value of a regression of r_t on R_t, we also have the stronger inequality:

$$\sigma(\tilde{H}_t) \leqslant a\sigma(\hat{r}_t). \tag{I.1$'$}$$

The coefficient a may be rather larger than the one expected (for $\gamma = .98$ with quarterly data, $a \cong 5$), so fairly high holding-yield volatility is consistent with the model. Still a is finite and (I.1) and (I.1′) can be tested. The coefficient a depends on our choice of $\gamma = 1/(1+\bar{R})$ and depends ultimately on the interest rate $\bar{R}$ we linearize around. However, for small $\bar{R}$, $a \cong 1/\sqrt{2\bar{R}}$, so that a is not highly sensitive to the choice of $\bar{R}$. If we varied the log of $\bar{R}$ by as much as $\pm$ .30, the log of a would vary only by about $\pm$ .15. This will generally not affect our results.

These inequalities quantify the smoothing behavior of the model, since $\sigma(\tilde{H}_t)$ is high when R is a choppy unsmooth series. Clearly, the permissible standard deviation of $\tilde{H}$ increases with the standard deviation of r, and only that component of the variance of r, namely, the variance of $\hat{r}$, that correlates with R is relevant. The correlation coefficient ρ_{rR} measures how much information about future r is

[14] Stationarity requires $\mathrm{var}(R_t) = \mathrm{var}(R_{t+1})$. Stationarity means, in the theory of stochastic processes, that the unconditional distribution of R_t does not change, and hence that R does not explode. The conditional variance $E_t(R_t - \bar{R})^2$ may yet change.

implicit in current r_t. If ρ_{rR} is one, then the r process is first-order autoregressive $\tilde{r}_t = \lambda\tilde{r}_{t-1} + \epsilon_t$ and $E_t(\tilde{r}_{t+j}) = \lambda^j\tilde{r}_t, j \geqslant 0$, and $\tilde{R}_t = \tilde{r}_t(1 - \gamma)/(1 - \gamma\lambda)$, where ~ denotes demeaned series. The maximum possible $\sigma(\tilde{H})$ for given $\sigma(r)$ then occurs with $\lambda = \gamma$, and here $\sigma(R) = \sigma(r)/(1 + \gamma) \cong \frac{1}{2}\sigma(r)$.[15]

By the same sort of reasoning, we can also show that restrictions on the smoothness of the long-rate series can be found which do not require knowledge of var(r_t) and would hold, technically, even if var(r_t) were infinite (r_t is unstationary) so long as var(Δr_t) is known and finite.[16] As long as Δr_t is stationary, $\tilde{H}_t - r_t$ and $R_t - r_t$ will also be stationary, and we can put limits on var($\tilde{H}_t - r_t$) given var(Δr_t).

To show this, we use the restriction $\mathrm{cov}(\tilde{H}_t - r_t, R_t - r_t) = 0$, which follows, as before, from (10). This restriction implies, instead of (11), that:

$$\mathrm{cov}(R_t - r_t, R_{t+1} - r_{t+1}) = \frac{1}{\gamma}\,\mathrm{var}(R_t - r_t) - \mathrm{cov}(R_t - r_t, \Delta r_{t+1}). \quad (11')$$

One may then substitute (11′) into $\mathrm{var}(\tilde{H}_t - r_t) = (1 - \gamma)^{-2}\,\mathrm{var}[(R_t - r_t) - \gamma(R_{t+1} - r_{t+1}) - \gamma\Delta r_{t+1}]$, which yields

$$\mathrm{var}(\tilde{H}_t - r_t) = (1 - \gamma)^{-2}[(\gamma^2 - 1)\,\mathrm{var}(R_t - r_t) + \gamma^2\,\mathrm{var}(\Delta r_t) + 2\gamma^2\rho_{R_t - r_t, \Delta r_t}\sqrt{\mathrm{var}(R_t - r_t)}\sqrt{\mathrm{var}(\Delta r_t)}].$$

If we maximize this expression with respect to both var($R_t - r_t$) and $\rho_{R_t - r_t, \Delta r_t}$, we get an upper bound to var($\tilde{H}_t - r_t$). Our model then implies that:

$$\sigma(\tilde{H}_t - r_t) \leqslant b\sigma(\Delta r_t), \qquad b = \gamma(1 - \gamma^2)^{-1/2}/(1 - \gamma) = a\gamma/(1 - \gamma) = a/\bar{R}. \quad (\mathrm{I}.2)$$

The upper bound is obtained if $\rho_{R_t - r_t, \Delta r_t} = 1$ and, in a case analogous to that which gave rise to the upper bound in (I.1), Δr_t is first-order autoregressive, $\Delta r_t = \gamma\Delta r_{t-1} + \epsilon_t$, and $E_t(\Delta r_{t+j}) = \gamma^j\Delta r_t, j \geqslant 0$. This upper bound, which is quite high, is obtained only when r_t is very strongly unstationary. In fact, for all of our data sets Δr_t is negatively correlated with $R_t - r_t$. If we require that $\rho_{R_t - r_t, \Delta r_t}$ be less than or equal to zero, the maximum var($\tilde{H}_t - r_t$) comes when var($R_t - r_t$) = 0, and so our model implies:

$$\sigma(\tilde{H}_t - r_t) \leqslant c\sigma(\Delta r_t), \qquad c = \gamma/(1 - \gamma) = 1/\bar{R}. \quad (\mathrm{I}.3)$$

[15] If r_t is regressed on a constant and r_{t-1}, the coefficient of r_{t-1} is .843, .966, .588, .845, .885, and .502 for data sets 1–6, respectively, always substantially below the corresponding γ_n in table 1.

[16] If r is expected to drift too far over the relevant horizon, our linearization argument for (1) may break down. This appears not to be a problem in our samples, since the value of R_t remained fairly near $\bar{R}$.

The upper bound occurs when Δr_t is white noise, r_t is a random walk, and $R_t = r_t + \Phi$.

The elements in the inequalities (I.1), (I.1′), (I.2), and (I.3) are examined for six data sets in table 1. All of the six data sets involve bonds that have a sufficiently large time to maturity that they may be approximated for our purposes as consols, and $\gamma_n \cong \gamma$.[17] However, γ_n (using for n the maturity of a representative bond in the sample and for $\bar{R}$ the average value of R over the sample) was substituted for γ in the above formulas so as not to overstate the holding-yield variance.[18]

We see from the sample standard deviations that for five of the six data sets $\sigma(\tilde{H}) > a\sigma(r)$, violating the inequality (I.1), and for all six data sets $\sigma(\tilde{H}) > a\sigma(\hat{r})$, violating the inequality (I.1′).[19] One also notes that for two data sets $\sigma(R) > \sigma(r)$. The inequalities (I.2) and (I.3) are, however, satisfied by the data.

That the inequalities (I.1) and (I.1′) are violated by the data constitutes some evidence against the model (1). That (I.2) and (I.3) are satisfied means that one cannot find evidence contrary to the model based only on the knowledge of $\sigma(\Delta r)$ and the simple arguments alluded to before about the smoothing imposed by the averaging in (1). In simple terms, it is conceivable that even if $\sigma(\Delta r)$ is very small, $\sigma(r)$ may be large, if r is expected to drift far above its historical range in the future (and Φ is sufficiently negative that R is not large). The

[17] A couple of other features of the data deserve mention. The bonds are callable, though for the Federal Reserve Series there is 5-year call protection, and for consols in data set 5 the call price is prohibitively high. Consideration of call provisions only strengthens our case. Call provisions ought to reduce the volatility of bond prices by shortening the effective maturity or putting an upper barrier on price. Differential taxation of interest and capital gains might, if our time period is long enough so that capital gains rates apply, suggest that $P_{t+1}^{(n-1)} - P_t^{(n)}$ in expression (4) be multiplied by $(1 - g)/(1 - i)$, where g is the effective rate of taxation on capital gains and i is the income tax rate of the "representative" bond investor (see Shiller and Modigliani 1979). This consideration again strengthens our case. Offering a tax advantage for capital gains means bond prices do not have to move as much to achieve equalization of (after-tax) returns.

[18] The approximation error in our use of finite maturity bonds then comes only in our use of $R_{t+1}^{(n)}$ rather than $R_{t+1}^{(n-1)}$ in computing $\tilde{H}_t^{(n)}$. This error should be negligible for long-term bonds. The inequalities where γ_n is substituted for γ can be derived in the same way for finite maturity bonds after n is substituted for $n - 1$ in the expression for $\tilde{H}_t^{(n)}$.

[19] The sample period for which the results are weakest is that from 1919 to 1959 in the United States, data set 4, apparently largely because of the period of the depression and World War II, when short rates were low and then officially pegged near zero, but long rates failed to fall so far. In this abnormal situation the market apparently correctly anticipated that the peg would end, and so here the expectations theory does the best. If the years of low short rates 1933–51 are omitted from the sample, $a\sigma(\hat{r})$ falls to 3.82, $a\sigma(r)$ to 4.32, and $\sigma(\tilde{H})$ rises to 6.81. That the inequality (I.1) is violated by this shorter sample is not due to the 1920–21 episode. Although short-term holding yields made large movements then, the short rate also moved dramatically. If the years 1920–21 are also omitted (as well as the years 1933–51) from data set 4, $\sigma(\tilde{H})$ falls from 6.81 to 6.09, but $a\sigma(r)$ falls even further, from 4.32 to 3.33, causing (I.1) to be violated even more strongly.

TABLE 1

Standard Deviations of Interest Rates and Holding-Period Returns

Data Set, Country, and Period	γ_n / n	$\sigma(R)$ / $\sigma(\Delta R)$	$\sigma(r)$ / $\sigma(\Delta r)$	$\sigma(H)$ / $\sigma_m(H)$	$\sigma(\tilde{H})$ / $\sigma_m(\tilde{H})$	$a\sigma(\hat{r})$ / $a\sigma(r)$	$\sigma(\tilde{H} - r)$ / $\sigma_m(\tilde{H} - r)$	$b\sigma(\Delta r)$ / $c\sigma(\Delta r)$
1. U.S., beginning of quarter, 1966:I–1977:II	.978	1.39	1.78	18.6	19.5	4.75	19.7	213.
	100	.429	.999	15.8	16.5	8.55	16.8	44.4
2. U.S., beginning of month, 1969:I–1974:I	.992	.433	1.77	27.7	27.4	6.25	27.5	503.
	288	.206	.512	24.0	23.6	14.03	23.8	63.5
3. U.S., beginning of year, 1960–76	.925	1.85	1.39	9.09	9.82	2.93	9.69	41.9
	25	.760	1.29	7.09	7.65	3.66	7.55	15.9
4. U.S., beginning of year, 1919–58	.940	1.04	1.86	5.21	5.48	4.73	5.31	43.5
	25	.327	.950	4.36	4.58	5.44	4.44	14.8
5. U.K., end of quarter, 1956:I–1977:II	.980	3.31	2.84	25.8	34.4	12.3	34.3	335.
	∞	.689	1.36	22.9	30.4	14.3	30.4	66.7
6. U.K., annual average, 1824–1929	.968	.596	1.17	4.27	4.95	2.28	4.97	144.
	∞	.159	1.20	3.82	4.43	4.66	4.44	36.2

NOTE.—The top row of terms in the column heads refers to the top line of each data set. The second row of terms refers to the second line of each data set. See Appendix B for description of data sets 1–6. The σ denotes sample standard deviation; σ_m denotes the lower bounds of a one-sided 95 percent confidence interval for the standard deviation, assuming observations are normal and uncorrelated. Except for data set 2 (for which see Appendix B), H is computed from expression (5) using n indicated and $C = R_t$. For all data sets, $\tilde{H}$ is computed from expression (7) using γ_n indicated; γ_n is computed as in (7) with n indicated using for R the mean of R over the sample period. All rates are expressed in percent per annum in this and subsequent tables and must be converted to rate per period to accord with notation in the text. The parameter a is defined as in inequality (I.1), b in (I.2) and c in (I.3), with γ as given in the first column.

short-run movements in R might reflect genuine new information about the large values of r that will come in the future.

Since $\tilde{H}_t - r_t$ and $\tilde{H}_t$ are approximately serially uncorrelated, we can put a lower bound on $\sigma(\tilde{H} - r)$ and $\sigma(\tilde{H})$ by computing a one-sided 95 percent confidence interval for them based on the χ^2 sampling distribution.[20] The lower bounds in table 1 are denoted $\sigma_m(\tilde{H} - r)$ and $\sigma_m(\tilde{H})$. Then if we accept that the true standard deviations lie above these values, $\sigma_m(\tilde{H})/a$ is the lowest possible value for $\sigma(r)$ and $\sigma(\hat{r})$ if the model (1) is to hold. For data sets 1, 2, 3, and 5 the standard deviation of the short rate would have to be roughly twice its historically observed value in order to justify this standard deviation of $\tilde{H}$, given the inequality (I.1).

IV. Testing Equality Restrictions on the Cross-Covariance Functions

We have seen that in all data sets the sample standard deviations of $\tilde{H}_t$ and r_t do not satisfy the inequality (I.1′). This implies that the sample covariances do not satisfy the equality restriction $\text{cov}(\tilde{H}_t - r_t, R_t) = 0$, which was used to prove (I.1′) (at least as long as the sample variance of R_t approximately equals the sample variance of R_{t+1}). Yet we have noted that many authors have studied some of the covariance restrictions implied by the model (by running appropriate regressions) for various samples and concluded in favor of the rational expectations model. Why then was it not discovered by these authors that this covariance restriction was violated in the sample? I will offer two explanations. First (and foremost), even if the holding-period yield variance is much too high, the R^2 in the regressions which would reveal this may be very low. Second, many authors have not run the right regressions; that is, they tested some restrictions which are not relevant to the holding-yield variance inequalities.

To see the first point, consider, for example, the monthly data for which the point is most dramatic. With monthly data, the parameter γ_n will be very nearly one. In data set 2, it is .992. The maximum standard deviation of the one-period holding yield consistent with the rational expectations model occurs when R_t is given by the simple autoregression $R_t = \gamma_n R_{t-1} + \epsilon_t$, and then the standard deviation of $\tilde{H}_t$ is about 12 times the standard deviation of R_t. If, on the other hand, the correlation of R_t with its own lagged value is only slightly smaller, then the standard deviation of the one-period holding yield may be much higher. For example, if the correlation of R_t with R_{t+1} is .96,

[20] The χ^2 confidence interval depends on the normality assumption, which is open to question. Normality appears satisfied for most data sets (see n. 23 below).

then the standard deviation of $\tilde{H}$ is about 35 times that of R_t. In this situation, if ΔR_{t+1} is regressed on R_t, the R^2 (i.e., the proportion of the variance of ΔR_{t+1} explained by R_t) is only .02. This R^2 turns out to be very small (and may remain small if other independent variables are added), so that it is unlikely to be significant in small samples. Hence many would be led to conclude that R_t is approximately a random walk.

The second point is that many of the most elaborate studies of the covariance restrictions implied by the model, such as those by Sutch (1968), Shiller (1972), and Modigliani and Shiller (1973), simply did not test these particular restrictions. These studies tested, in essence, whether a regression of R onto current and lagged r and current and lagged inflation rates produces an equation (the "term-structure equation") which is the same as an optimal forecasting equation for R^* based on the same variables. Neither R nor $R - r$ appeared on the right-hand side of the equations. In fact, the serially uncorrelated residual in the term-structure equation (which was ascribed in part to exogenous shocks to supply and demand in various "habitats") meant that high R or high $R - r$ would indeed imply high $\tilde{H} - r$.

Sargent (1979), on the other hand, did test (in essence) these restrictions, along with others implied by the model, using a likelihood ratio test. His tests did not reject the hypothesis that all restrictions hold, perhaps due to low power of this test for the particular restrictions that we are concerned with here and perhaps to his use of a relatively short-term (5-year) bond to represent a long bond.

We now proceed to test restrictions in (10) by running regressions as described above. When one contemplates running such regressions, one must confront the fact that there are potentially an infinite number of coefficients in the model, yet only a finite amount of data. One must eliminate variables before running a regression. Since all coefficients are zero, it does not matter from the standpoint of the model which variables are eliminated. My approach was to eliminate all but R_t (table 2) or $(R_t - r_t)$ (table 3). It is good statistical methodology to concentrate the power of one's test onto an interesting alternative hypothesis. I have attempted to do this in consideration of the volatility arguments noted above (these restrictions produced the inequalities) and the understandability and simplicity of the alternative hypothesis. Sargent's (1979) regressions included, in effect, eight right-hand-side variables: the current and three lagged values of both R and r.[21]

[21] This interpretation does not appear in Sargent's paper, which emphasizes that his model tests some complicated nonlinear restrictions on the coefficients of an autoregression for (r_t, R_t). The nonlinearity of the restrictions is introduced by his use of the assumption that the R_t series refers to finite maturity pure discount bonds. Since time to

TABLE 2

REGRESSION OF SPREAD BETWEEN SHORT-TERM HOLDING YIELD OF LONG-TERM BONDS AND SHORT RATE ON THE LONG RATE ($H_t - r_t = A + B \cdot R_t + e_t$)

Data Set, Country, and Period	$\hat{A}$ (t_A)	$\hat{B}$ (t_B)	R^2	SE	D-W
1. U.S., beginning of quarter, 1966:I–1977:II	−40.2	4.97			
	(−2.72)*	(2.62)*	.135	17.7	1.63
2. U.S., beginning of month, 1969:I–1974:I	−139.	18.2			
	(−2.28)*	(2.27)*	.080	26.9	2.01
3. U.S., beginning of year, 1960–76	−8.81	1.32			
	(−1.14)	(1.11)	.076	8.81	1.48
4. U.S., beginning of year, 1919–58	−1.14	.748			
	(−.373)	(.954)	.023	5.08	1.59
5. U.K., end of quarter, 1956:I–1977:II	−14.7	1.54			
	(−2.03)*	(1.86)*	.040	25.2	1.88
6. U.K., annual average, 1824–1929	−2.50	.646			
	(−1.51)	(.913)	.008	4.32	1.25

NOTE.—Numbers in parentheses are t-statistics.
*Significant at the 5 percent level with a one-sided test. See Appendix B for source of data.

TABLE 3

REGRESSION OF CHANGE IN LONG RATE ON LONG-RATE–SHORT-RATE SPREAD ($R_{t+1} - R_t = A + B \cdot [R_t - r_t] + e_t$)

Data Set, Country, Period, and Dates	$\hat{A}$ (t_A)	$\hat{B}$ (t_B) [t'_B]	R^2	SE	D-W
1. U.S., beginning of quarter, 1966:I–1977:II	.211 (2.97)*	−.125 (−3.33)* [−3.93]*	.201	.388	2.21
2. U.S., beginning of month, 1969:I–1974:I	.039 (1.27)	−.020 (−1.22) [−1.68]*	.024	.206	2.17
3. U.S., beginning of year, 1960–76	.451 (1.49)	−.187 (−1.10) [−1.58]	.075	.755	1.88
4. U.S., beginning of year, 1919–58	.005 (.068)	−.028 (−.582) [−1.90]*	.009	.323	2.12
5. U.K., end of quarter, 1956:I–1977:II	.164 (2.06)*	−.120 (−2.81)* [−3.29]*	.086	.662	2.27
6. U.K., annual average, 1824–1929	.011 (.715)	.002 (.183) [−2.32]*	.0003	.162	1.46

NOTE.—Numbers in parentheses are t-statistics; [t'_B] is the t-statistic for the hypothesis $B = (1-\gamma_n)/\gamma_n$, where γ_n is as given in table 1.
*Significant at the 5 percent level with a one-sided test.

The regressions were run for each of the data sets using both $\tilde{H}_t$ and H_t. Results were very similar with the two different measures of one-period holding yields, and I chose to present in table 2 those using $H_t - r_t$ and R_t and in table 3 those using $\tilde{H}_t - r_t$ and $R_t - r_t$. In table 3, the results are reported as a regression of $R_{t+1} - R_t$ on $R_t - r_t$, and since $R_{t+1} - R_t$ is a linear combination of $\tilde{H}_t - r_t$ and the independent variable, it amounts to the same regression. The coefficient of $R_t - r_t$ should then equal $(1 - \gamma_n)/\gamma_n$, which is greater than zero.

One notes, since R_t enters positively into the determination of H_t and $\tilde{H}_t$, that R_t in effect appears on both sides of all of these equations. This might seem to suggest an upward "bias" for the slope coefficient in the regressions. In fact, however, under the rational expectations hypothesis the error term in the regressions is uncorrelated with the independent variable, and so under this hypothesis there is no bias. If, on the other hand, the rational expectations hypothesis is only partly true and R_t is influenced by other factors not in the rational expectations model, then we should not be surprised if such a "bias" emerges and the slope coefficient turns out to be significant.

Unfortunately, an upward bias in the coefficient of R_t or $(R_t - r_t)$ might also emerge in some of the data sets even if the rational expectations model is true. It is possible that measurement error in the R_t series could induce an upward bias in the coefficient of R_t or $(R_t - r_t)$. However, there should be little measurement error in these series. It is true, of course, that not all bonds of the same maturity and coupon sell for exactly the same price, so that different bond yield series may differ slightly, but this is not so much measurement error as a reflection of the very deviation from the rational expectations model that we are interested in describing.[22]

---

maturity is constant, $H_t^{(n)}$ cannot be expressed as a linear function of his $R_t^{(n)}$ series. We have eliminated the nonlinearity by discussing consols or long-term coupon-carrying bonds which may be approximated as consols. Our model then places simple linear restrictions on the autoregressive coefficients for (r_t, R_t), which are in effect summarized in (10).

[22] Suppose the true unobserved long rate (which we shall denote by $\tilde{R}_t$) behaves in accordance with the expectations theory, and that (for table 3) $\tilde{R}_{t+1} - \tilde{R}_t = a + b(\tilde{R}_t - r_t) + U_t$, where $b = (1 - \gamma_n)/\gamma_n$. If the error in measurement $e_t = R_t - \tilde{R}_t$ is uncorrelated with all other variables (including its own lagged values), and if U_t is uncorrelated with $R_t - r_t$, then the coefficients whose estimates appear in table 3 will be $\hat{b} = [b \operatorname{var}(\tilde{R} - r) - \operatorname{var}(e)]/[\operatorname{var}(\tilde{R} - r) + \operatorname{var}(e)]$. Using $\operatorname{var}(R - r) = \operatorname{var}(\tilde{R} - r) + \operatorname{var}(e)$ and solving for var (e) we find: $\operatorname{var}(e) = [(b - \hat{b}) \operatorname{var}(R - r)]/(1 + b)$. We can, using this expression, deduce how big the measurement error would have to be if it were to account for our results in table 3. Setting $b = (1 - \gamma_n)/\gamma_n$ from table 1, we then find that for data sets 1–6 the standard deviation of the measurement error would have to be 58, 27, 55, 32, 62, and 17 basis points, respectively. It is inconceivable that true measurement error could be this high. A 4 standard deviation range would be over 2 percentage points for several series. Discrepancies between different bond-yield averages purporting to measure

In table 2 we observe a significant positive coefficient (based on a one-tailed test at the 5 percent level) for R_t in data sets 1, 2, and 5.[23] Although other data sets were not significant at the 5 percent level, the pattern appears to be the same for all six data sets: a negative intercept and a positive slope coefficient.[24]

The R^2 is small in all of the table 2 regressions, as we would expect. The coefficients, however, are not small. For example, with data set 1 the coefficient of 4.97 means that if R_t rises by 1 percentage point, then the expected annualized one-quarter holding yield on long bonds relative to short returns rises by 4.97 percentage points. The R^2 is still small since there is so much unpredictable variation in the short-term holding yield.

In table 3, the coefficients of $R_t - r_t$ have a negative sign (contrary to the implication of the model [1]) in data sets 1–5 and are significantly below the theoretical value of $(1 - \gamma_n)/\gamma_n$ at the 5 percent level based on a one-tailed test for all data sets except data set 3.

The results in table 3 contradict what may be thought of as the essential characteristic of rational expectations models: that long-term interest rates tend on average to move in such a way as to equalize short-term holding yields. This characteristic, which we noted in connection with the perfect-certainty model illustrated in figure 3, carries over to our model (1) in the sense that long rates should rise on average when long rates are high relative to short rates and decline on average when long rates are low relative to short rates. Instead, long rates if anything move in the opposite direction, which means that when long rates are high relative to short rates they tend to move down in the subsequent period. The capital gain thus produced augments (rather than offsets) the advantage to holding long-term bonds when these bonds have higher current yield. This behavior is not consistent with our rational expectations models but is instead what we would expect to find if long rates are influenced by noise

---

roughly the same thing are nowhere near this high, and, as noted above, even these discrepancies are not due primarily to measurement error but instead to variations in actual price among similar high-quality bonds.

[23] Another potential problem in the evaluation of the t-statistics here is that the error term may be nonnormal. The studentized range test recommended by Fama and Roll (1971) as a test of normality gives statistics for $H_t - r_t$, for data sets 1–6, of 3.97, 4.61, 3.18, 4.40, 6.81, and 6.94, respectively. The ratio shows no evidence for nonnormality for data sets 1–4, but is significant at the 0.5 percent level for data sets 5 and 6. Examining the data suggests that for data set 5 the problem is one of increasing variance through time, rather than leptokurtosis. The equations for data set 5 were thus reestimated by generalized least squares by scaling the observations by $e^{-0.1t}$, $t = 1, \ldots, 86$. The t-statistic in the regression then drops to 1.33 in table 1, which is no longer significant at the 5 percent level, but t_b and t'_b in table 2 remain significant ($t_b = -1.88$ and $t'_b = -2.40$).

[24] Basu (1977) found an analogous result that price-earnings ratios are negatively correlated with corporate stock returns.

which causes long rates temporarily to rise relative to short rates and then fall to a more "normal" level.[25]

V. Summary and Conclusions

The goal that was set for this paper was in some ways ambitious. I sought to find simple ways of understanding whether the data are well described by any of a number of expectations models of the term structure. I was guided, however, by a plain fact that seemed to stand in glaring contradiction to these models: the fact that actual long-rate series (as illustrated in fig. 1) look completely different from ex post rational long-rate series (as illustrated in fig. 3).

Since the ex post rational long-rate series was very smooth, it seemed likely that a robust implication of the expectations models of the term structure would be that actual long-rate series should be similarly smooth. It is not easy, however, to give formal content to this implication for all of these models. The linearization which produced expression (1) enabled me to derive some inequalities which do this, subject, however, to the approximation error in the linearization. The upper bounds on the volatility of the long-rate series that these inequalities impose occur only in worst possible cases where the short rate has a specific autocovariance function. In using these inequalities to examine the model, we passed up the possibility of using the actual autocovariance function in conjunction with the expectations model to put a tighter limit on the volatility of the long-rate series.

Based on sample standard deviations (table 1), the inequalities (I.1) and (I.1′) implied by the linearized expectations model appear violated by the data. If we wish to test whether the population standard deviations violate the inequalities, then we must inquire whether the terms in the inequalities can be reliably measured under general assumptions in small samples. It was felt that it is reasonable to suppose that we can put a lower bound on the left side of the inequalities by a usual χ^2 one-sided confidence interval, since the variable whose standard deviation is measured is approximately serially uncorrelated. It is another matter to put an upper bound on the right-hand side of the inequalities, since we have no real information in small samples about possible trends or long cycles in interest rates. Indeed, some would claim that short-term interest rates may be un-

[25] This observation (which was first pointed out to me by Franco Modigliani) is analogous to one reported early by Shiller and Siegel (1977) that long-term bond yields move on average in a direction which exacerbates rather than mitigates the effect of inflation on real returns. Mishkin (1978) has further confirmed table 2 (as well as table 1) with high-quality U.S. short-term holding-yield data on intermediate term bonds, and with a heteroscedasticity correction, although his table 3 results are somewhat less significant than those reported here for roughly the same sample.

stationary and hence have infinite variance. The fact that the lower bound on the left-hand side exceeds the sample value of the right-hand side may be interpreted as safely telling us, then, that we must rely on such unobserved variance or expected explosive behavior of short rates if we wish to retain expectations models. This conclusion appears to hold for earlier time periods as well as the more recent.

The inequalities which characterize the smoothing behavior are not the only avenues for constructing tests of the model which are powerful against an alternative hypothesis that long rates are too volatile, and in fact small sample tests are available. Such regression tests (tables 2 and 3) do generally reject the expectations model in favor of an alternative hypothesis that long rates are disturbed by transient effects unrelated to expectations. The regression results might still be construed as offering some support for the expectations models in the sense that the R^2 are small. However, based on the results reported here, there is nothing more to be said for the expectations model. My table 3 regressions show that movements in long rates tend to be in a direction opposite to that predicted by the expectations models.

These negative results on the expectations model may be contrasted with earlier positive results by Sutch (1968), Shiller (1972), Modigliani and Shiller (1973), and Sargent (1979). Some interpretation of the contrast, in terms of the residual of the term-structure equation, was offered above. An attempt at further reconciliation of these apparently conflicting results will be the subject of another paper.

Appendix A

Model Restrictions on the Spectral Densities of Interest Rates

The ex post rational long rate R_t^* is defined in expression (9) as a moving average of r_t. The squared gain of this moving average, or "linear filter," is $g^2(\omega) = (1 - \gamma)^2/[1 - 2\gamma \cos(\omega) + \gamma^2]$, where ω is frequency $-\pi \leq \omega \leq \pi$. The squared gain is 1.00 at $\omega = 0$, declines monotonically as frequency increases, and reaches $[(1 - \gamma)/(1 + \gamma)]^2$ at $\omega = \pi$.[26] If γ is close to 1.00, the decline will be very dramatic, as my example illustrated. The spectrum of R^* thus lies below the spectrum of r everywhere except at $\omega = 0$, and is relatively much more concentrated in the lower frequencies. Spectral analysis of actual interest rate time series, however (e.g., Granger and Rees 1968), does not reveal any such attenuation of the high-frequency components in long-rate series.

One might have thought that our results with the perfect-certainty model (9)—that the spectrum of R^* must be more concentrated in the lower frequencies and must lie everywhere below the spectrum of r—should also carry

[26] If the interest rates r_{t+K} do not have finite second moments (as claimed, e.g., by Roll [1970] for forward rates), then we cannot use this analysis. It remains true if, as these authors estimated, the characteristic exponent is greater than one, that the dispersion of the rates will be reduced by averaging. Clark (1973), in any event, showed that a finite variance model explains speculative price data well.

over to the R in the general model (1). However, this is not the case, as a simple example will illustrate. Suppose r_t is a first-order moving average of a white-noise process ϵ_t: $\tilde{r}_t = \epsilon_t + \theta\epsilon_{t-1}$ $0 < \theta < 1$, and suppose the ϵ_t is also serially independent and no other information is available for forecasting, so that the optimal forecast is linear in current and lagged r. Then, in this special case, $E_t(\tilde{r}_{t+1}) = \theta\epsilon_t$ and $E_t(\tilde{r}_{t+K}) = 0$ $K > 1$ so that $\tilde{R}_t = (1 - \gamma)$ $[(1 + \gamma\theta)\epsilon_t + \theta\epsilon_{t-1}]$ will appear less smooth than r_t since the moving average weights for R_t are relatively more concentrated on ϵ_t. More precisely, the squared gain from r to R (which is also the ratio of their spectra) is $g^2(\omega) = (1 - \gamma)^2 [(1 + \gamma\theta)^2 + \theta^2 + 2\theta(1 + \gamma\theta)\cos(\omega)]/[1 + \theta^2 + 2\theta\cos(\omega)]$, which is a function that increases monotonically with ω for $0 \leqslant \omega \leqslant \pi$. One might also note that as θ approaches one, $g^2(\omega)$ approaches infinity at $\omega = \Pi$. This illustrates that there is no limit to the ratio of the spectra at a particular frequency, and that the spectrum of R need not lie everywhere below the spectrum of r as does the spectrum of R^*.

In spite of this counterexample, it remains true that R_t cannot have too much power at the high frequencies. In the above example, the variance of R_t is very small, so that even though the spectrum of R is relatively more concentrated in the higher frequencies and may have absolutely more power at the highest frequencies, the total power in the spectrum of R at the higher frequencies is still small. Inequality restrictions on the total power at the high frequencies for R_t can be derived from the restrictions that the model (1) places on the autocovariance function of the bivariate process (r_t, R_t).

The covariance restrictions (10) can be rewritten in terms of the cross-covariance functions:

$$C_{RR}(\tau) = (1 - \gamma)\sum_{K=0}^{\infty}\gamma^K C_{rR}(K + \tau), \qquad \tau \geqslant 0, \tag{A1}$$

$$C_{rR}(-\tau) = (1 - \gamma)\sum_{K=0}^{\infty}\gamma^K C_{rr}(K + \tau), \qquad \tau \geqslant 0, \tag{A2}$$

where for any pair of variables x and y, $C_{xy}(\tau)$ will refer to $E[(x_t - E(x)][y_{t-\tau} - E(y)]$. These expressions give all restrictions imposed by our model on the autocovariance function of the bivariate process (r_t, R_t). Since $C_{rR}(\tau)$ is not generally an even function of τ (as $C_{RR}[\tau]$ and $C_{rr}[\tau]$ are), these expressions do not suffice to define $C_{RR}(\tau)$ given $C_{rr}(\tau)$, and hence the spectrum of R_t is not determined by the spectrum of r_t as it was in the perfect-certainty case or first-order moving-average case discussed above. This is as we would expect, since we have not specified all of the information used in forecasting. Both of the special cases considered above are in fact consistent with the above equalities, though in each case additional restrictions are also involved. When future short-term interest rates are known with certainty, the relations (A1) and (A2) hold for all τ. When the expected future short rates are optimal linear forecasts based on current and lagged short rates only (as in the moving-average case considered above), then other restrictions can be shown to characterize the cross-covariance function; that is, R does not cause r in the Granger (1969) or Sims (1972) sense. Neither of these restrictions is assured in the general case, however, and I do not assume them here.

We can now see in what sense the restrictions (A1) and (A2) may be described as putting a limit on the high-frequency variance of R_t. Since the holding-period yield $\tilde{H}_t$ is derived by passing R_t through the linear filter $(1 - \gamma F)/(1 - \gamma)$, then the spectrum of $\tilde{H}_t$, $S_{\tilde{H}}(\omega)$, equals the spectrum of R_t, $S_R(\omega)$, times the squared gain of this filter which is $g^2(\omega) = [1 + \gamma^2 - 2\gamma\cos(\omega)]/(1 -$

$\gamma)^2$. This gain function rises monotonically with the absolute value of frequency. Squared gain is 1.0 at $\omega = 0$ and rises, for $\gamma = .98$ (corresponding to $\bar{R} = 8$ percent per annum with quarterly data), to about 10,000. The variance of $\tilde{H}_t$ is the integral from $-\pi$ to π of its spectrum, which is then the integral from $-\pi$ to π of $g^2(\omega)$ times the spectrum of R, so that, by (I.1):

$$\int_{-\pi}^{\pi} \frac{1 + \gamma^2 - 2\gamma\cos(\omega)}{(1 - \gamma)^2} S_R(\omega)\, d\omega \leq \frac{\text{var}(r_t)}{1 - \gamma^2}.$$

The left-hand side of this expression is a weighted integral of the spectrum of R_t with very high weights for high frequencies. Thus, an observation that the holding-period yield variance does not satisfy the inequality may generally be described as an observation that the high-frequency components of R_t are too strong to be consistent with the model (1).

Restrictions (I.2) and (I.3) may similarly be interpreted as inequality restrictions on a weighted integral of the spectrum of the bivariate process (Δr_t, $R_t - r_t$). Our observation that the variance of R must be less than the variance of R^* is also of this form. Other (presumably less easily interpreted) inequality restrictions of this form can also be derived from (A.1) and (A.2).

Appendix B

Sources of Data

Data are from the macro data library of the Federal Reserve System or the Federal Reserve *Bulletin* unless otherwise noted.

Data set 1 (quarterly, 1966:I–1977:III): The long-term interest rate R is the Federal Reserve recently offered AAA utility bond yield series (constructed by Kichline, Laub, and Stevens [1973]), and the short-term interest rate r is the 4–6-month prime commercial paper rate. Both series are for the first week of the quarter.

Data set 2 (monthly, 1969:I–1974:II): The long-term interest rate R is a series produced by Salomon Brothers for yields on the first of the month of a composite portfolio of Aa utilities and industrials (Leibowitz and Johannesen 1975, table XII). The annualized 1-month holding yield H (Leibowitz and Johannesen 1975, table XV) is not computed from the R series but from an average price series for the bonds using the actual average coupon and average maturity. The short-term interest rate r is the 90–119-day prime commercial paper rate starting in June 1972, and the 4–6-month prime commercial paper rate before that date, both for the first week of the month.

Data set 3 (annual, 1960–77): The long-term interest rate R is the Federal Reserve new issue Aaa utility yield series for the first month of the year, and the short-term interest rate r is the 12-month U.S. Treasury bill rate averaged over the first month of the year.

Data set 4 (annual, 1919–59): The long-term interest rate R is the Moody Aaa corporate bond yield average, and the short-term interest rate r is the 4–6-month prime commercial paper rate, both for the first month of the year. The sample was ended in 1959 to provide estimates over a sample period which does not overlap with those of data sets 1–3.

Data set 5 (quarterly, 1956:I–1977:III): The long-term interest rate R is the flat yield on 2½ percent British Consols as reported in *Financial Statistics* from the Central Statistical Office starting in 1962 and in the *London Times* for the earlier years. Observations are taken at the last Friday of the quarter. The coupon is paid on the fifth day of the following quarter. The short-term interest rate series r is the 3-month local authorities temporary loan rate for the last Friday of the quarter starting 1960:III and for the last Saturday before that as reported in the *Bank of England Statistical Abstract, Number I* (1970), table 29, and subsequent issues of the *Bank of England Quarterly Bulletin*.

Data set 6 (annual 1824–1930): The long-term interest rate R is the annual average rate of 3 percent British Consols through 1888 and on 2½ percent government annuities starting in 1889 (Homer 1963, table 19, col. 2, and table 57, col. 2). The short-rate r is, for 1824–44, Overend and Gurney's annual average first-class 3-month bill rates and, after 1844, the annual average rates (averaging maximum and minimum) for 3-month bank bills, both from Mitchell and Deane (1962, p. 460). The data series were terminated here in 1930 to provide an estimate for the period before the great depression.

References

Basu, S. "The Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Markets Hypothesis." *J. Finance* 32 (June 1977): 663–82.

Bierwag, G. O., and Grove, M. A. "A Model of the Structure of Prices of Marketable U.S. Treasury Securities." *J. Money, Credit and Banking* 3, no. 3 (August 1971): 605–29.

Black, Fischer; Jensen, Michael; and Scholes, Myron. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theory of Capital Markets*, edited by Michael Jensen. New York: Praeger, 1972.

Cargill, Thomas. "The Term Structure of Interest Rates: A Test of the Expectations Hypothesis." *J. Finance* 30 (June 1975): 761–71.

Clark, Peter K. "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices." *Econometrica* 41, no. 1 (January 1973): 135–55.

Cox, John C.; Ingersoll, Jonathan E., Jr.; and Ross, Stephen A. "A Theory of the Term Structure of Interest Rates and the Valuation of Interest-dependent Claims." Mimeographed. Stanford Univ., 1977.

Culbertson, John M. "The Term Structure of Interest Rates." *Q.J.E.* 71 (November 1957): 485–517.

Fama, Eugene F. "Forward Rates as Predictors of Future Spot Rates." *J. Financial Econ.* 3 (October 1976): 361–77.

Fama, Eugene F., and Roll, Richard. "Parameter Estimates for Symmetric Stable Distributions." *J. American Statis. Assoc.* 66, no. 334 (June 1971): 331–38.

Friend, Irwin, and Blume, Marshall E. "Measurement of Portfolio Performance under Uncertainty." *A.E.R.* 60, no. 4 (September 1970): 561–75.

Friend, Irwin; Westerfield, Randolf; and Granito, Michael. "New Evidence on the Capital Asset Pricing Model." *J. Finance* 33, no. 3 (June 1978): 903–17.

Granger, C. W. J. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods." *Econometrica* 37, no. 3 (July 1969): 424–38.

Granger, C. W. J., and Rees, H. J. B. "Spectral Analysis of the Term Structure of Interest Rates." *Rev. Econ. Studies* 35 (January 1968): 67–76.

Homer, Sidney. *A History of Interest Rates.* New Brunswick, N.J.: Rutgers Univ. Press, 1963.

Kichline, James L.; Laub, P. Michael; and Stevens, Guy V. G. "Obtaining the Yield on a Standard Bond from a Sample of Bonds with Heterogeneous Characteristics." Staff Econ. Study no. 77, Board Governors Federal Reserve System, 1973.

Laffer, Arthur B., and Zecher, Richard. "Some Evidence on the Formation, Efficiency and Accuracy of Anticipations of Nominal Yields." *J. Monetary Econ.* 1 (July 1975): 327–42.

Leibowitz, Martin L., and Johannesen, Richard I., Jr. "Introducing the Salomon Brothers Total Performance Index for the High Grade Long-Term Corporate Bond Market." Mimeographed. New York: Salomon Brothers, 1975.

LeRoy, Stephen, and Porter, Richard. "The Present Value Relation: Tests Based on Implied Variance Bounds." Mimeographed. Board Governors Federal Reserve System, 1979.

McCallum, John S. "The Expected Holding Period Return, Uncertainty and the Term Structure of Interest Rates." *J. Finance* 30, no. 2 (May 1975): 307–23.

Mishkin, Frederic S. "Efficient Markets Theory: Its Implications for Monetary Policy." *Brookings Papers Econ. Activity* 3 (1978): 707–52.

Mitchell, Brian R., and Deane, Phyllis. *Abstract of British Historical Statistics.* Cambridge: Cambridge Univ. Press, 1962.

Modigliani, Franco, and Shiller, Robert. "Inflation, Rational Expectations, and the Term Structure of Interest Rates." *Economica* 40 (February 1973): 12–43.

Pesando, James E. "On the Efficiency of the Bond Market: Some Canadian Evidence." *J.P.E.* 86, no. 6 (December 1978): 1057–76.

Phillips, Llad, and Pippenger, John. "Preferred Habitat vs. Efficient Market: A Test of Alternative Hypotheses." *Federal Reserve Bank St. Louis Rev.* 58, no. 5 (May 1976): 11–19.

Roll, Richard. *The Behavior of Interest Rates: An Application of the Efficient Market Model to U.S. Treasury Bills.* New York: Basic, 1970.

———. "Investment Diversification and Bond Maturity." *J. Finance* 26, no. 1 (March 1971): 51–66.

Sargent, Thomas J. "Rational Expectations and the Term Structure of Interest Rates." *J. Money, Credit and Banking* 4, no. 1, pt. 1 (February 1972): 74–97.

———. "A Classical Macroeconometric Model for the United States." *J.P.E.* 84, no. 2 (April 1976): 207–37.

———. "A Note on Maximum Likelihood Estimation of the Rational Expectations Model of the Term Structure." *J. Monetary Econ.* 5, no. 1 (January 1979): 133–43.

Shiller, Robert J. "Rational Expectations and the Term Structure of Interest Rates." Ph.D. dissertation, Massachusetts Inst. Tech., 1972.

———. "Rational Expectations and the Dynamic Structure of Macroeconomic Models." *J. Monetary Econ.* 4 (January 1978): 1–44.

———. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" Mimeographed. Univ. Pennsylvania, 1979.

Shiller, Robert J., and Modigliani, Franco. "Coupon and Tax Effects on New and Seasoned Bond Yields and the Measurement of the Cost of Debt Capital." *J. Financial Econ.*, vol. 7, no. 3 (December 1979).

Shiller, Robert J., and Siegel, Jeremy J. "The Gibson Paradox and Historical Movements in Real Interest Rates." *J.P.E.* 85, no. 5 (October 1977): 891–908.

Sims, Christopher A. "Money, Income and Causality." *A.E.R.* 62, no. 4 (September 1972): 540–52.

Sutch, Richard C. "Expectations, Risk and the Term Structure of Interest Rates." Ph.D. dissertation, Massachusetts Inst. Tech., 1968.