THOMAS J. SARGENT

Rational Expectations and the Term Structure of Interest Rates

I. INTRODUCTION*

THIS PAPER REPORTS SOME TESTS of two important hypotheses about the behavior of the term structure of interest rates. The first hypothesis is the "expectations hypothesis," which states that forward rates of interest are forced into equality with the short rates that investors expect to prevail in subsequent periods. The second hypothesis is that the expectations of investors are rational in the sense of John F. Muth [22].[1] By this we mean that investors' expectations are equivalent with the optimal forecasts of statistical theory for a certain specified class of statistical models. A convenient way to characterize a market that satisfies both of these hypotheses is as an "efficient market."[2] While the first hypothesis has been purportedly subjected to a bewildering variety of empirical "tests," it has only rarely[3] been tested within a framework that requires maintaining that expectations incorporate available information efficiently, the second hypothesis. The criterion of acceptable empirical results has generally been that of "plausibility," a criterion with unfortunately little discriminatory power.

*The research underlying this paper was supported by a grant from the Social Sciences Research Council. Helpful comments on an earlier draft were made by E. Philip Howrey and Richard Roll.

[1] Actually, what we are using in this paper is a poor man's version of Muth's argument, since the statistical models that are assumed to govern the interest rate have no analytical economic content, being naive autoregressive schemes. In this we are following the path taken in most previous work on the term structure. However, a more complete approach would take into account the likelihood that other variables such as the money supply, price level, and rate and composition of income contain information that is useful for predicting the interest rate.

[2] The phrase "efficient market" is due to Roll [26] and Fama [9].

[3] The most notable exception is Roll's excellent study of the behavior of U.S. Treasury Bill rates. See [26].

THOMAS J. SARGENT is associate professor of economics at the University of Minnesota and is economist with the National Bureau of Economic Research.

The result has been the proliferation of empirical studies, with few decisive results having been achieved.

The contention of this paper is that by also invoking the second hypothesis, namely that readily available information will not be wasted, much sharper tests of the expectations hypothesis are made possible. Needless to say, the requirement that available information be used efficiently is much weaker than the requirement that expectations be very accurate. It was against the imposition of that stronger requirement that David Meiselman was contending when he wrote that "anticipations may not be realized yet still determine the structure of rates in the manner asserted by the [expectations] theory."[4] Although this is correct, it nevertheless seems unwise to construct models that build in "irrational" expectations.

The implication of our two hypotheses is remarkably simple: they imply that a certain sequence of forward interest rates follows a martingale.[5] This important proposition, which is due to Paul Samuelson [27], and which was first implemented in the context of the term structure by Richard Roll [26], is discussed in Section II.[6] There we also discuss the relationship of David Meiselman's [19] important work to the Samuelson model considered here. As it turns out, Meiselman's equations are implied when things are restricted a bit more than they need to be to obtain Samuelson's martingale proposition. In Section III, we present an empirical test of whether Durand's basic yields satisfy the implications of our two hypotheses, together with some closely related tests of a certain class of "random-walk" models. In Section IV we return to Meiselman's model, arguing that there is an important (asymptotic) bias in estimates of the slopes in Meiselman's model. The presence of the bias helps explain some curious features of results that Meiselman and others have obtained. Finally, our conclusions are stated in Section V.

This paper is most closely related to the previous work of Stanley Diller [6] and Richard Roll [26]. It was Diller who first examined under what classes of stochastic processes governing the evolution of spot rates revision equations like Meiselman's would emerge as a consequence of "optimal" forecasting. Roll was the first economist to implement Samuelson's martingale theorem in the context of a study of the term structure. His treatment of that theorem is considerably broader than the one included here, being based on his extensive work on capital-market equilibrium theory. In addition, his empirical tests, conducted on the basis of weekly data on U.S. Treasury bill rates, do not assume that bill rates are covariance stationary. Instead, Roll argues that the evolution of bill rates is more adequately described by assuming that they are drawn from one of the stable distributions with infinite variance. While that specification is certainly an interesting one, abandon-

[4] Meiselman [19, p. 12].

[5] A sequence $\{x_t, x_{t+1}, x_{t+2}, \ldots\}$ is said to follow a martingale if

$$E(x_{t+1} | x_t, x_{t-1}, x_{t-2}, \ldots) = x_t,$$

where E is the mathematical expectation operator. For discussions of the properties of martingales, see Doob [7] and Feller [10].

[6] Mandelbrot [17] should be regarded as the codiscoverer of the martingale model, along with Samuelson.

ing the assumption of covariance stationarity has its costs. In Section II we try to show that if the assumption of covariance stationarity is retained, a wide variety of empirical work on the term structure can be interpreted as testing Samuelson's model.

II. RESTRICTIONS IMPLIED BY THE TWO HYPOTHESES

A. The Martingale Theorem

It is necessary to assume that the spot one-period rate, R_t, can be characterized by the probability distribution function

$$\begin{aligned}\text{Probability } [R_{t+j} \leqslant R \mid R_t = r_0, R_{t-1} = r_1, \ldots] \\ = P(R, r_0, r_1, \ldots; j) \qquad j = 1, 2, 3, \ldots.\end{aligned} \tag{1}$$

The probability function is assumed to be independent of calendar time. That is all. The process need not be normal. It can even be stable Paretian, having infinite variance, provided that its mean exists (which rules out Cauchy processes). It need not be linear.

Notice that all the $P(R, \ldots; j)$'s for $j = 2, 3, \ldots, \infty$ can be calculated recursively from the one-span-forward probability distribution $P(R, \ldots; 1)$:

$$P(R, r_0, r_1, \ldots; j) = \int_{-\infty}^{\infty} P(R, z, r_0, r_1, \ldots; j-1)\, dP(z, r_0, r_1, \ldots; 1). \tag{2}$$

Let ${}_{t+j}F_t$ be the forward interest rate on one-period loans that prevails at time t for loans made at time $t+j$. We now impose *both* of our hypotheses by requiring that

$$\begin{aligned}{}_{t+j}F_t &= E\,[R_{t+j} | R_t, R_{t-1}, \ldots] \\ &= \int_{-\infty}^{\infty} R\,dP(R_t, R_{t-1}, \ldots; j)\end{aligned} \tag{3}$$

where E denotes mathematical expectation. Equation (3) states that the market equates the j-period forward rate to the expectation formed on the basis of the probability distribution describing the evolution of the spot rate. Not only are expectations supposed to determine the yield curve, but they are assumed to be based on all the information available, namely the P's. Samuelson's theorem states that under the assumed conditions the following sequence follows a martingale:

$$\{{}_{t+j}F_t,\ {}_{t+j}F_{t+1}, \ldots, {}_{t+j}F_{t+j-1}, R_{t+j}\} \tag{4}$$

The martingale is defined by the condition

$$E({}_{t+j}F_{t+1} \mid R_t, R_{t-1}, \ldots) = {}_{t+j}F_t \tag{5}$$

for all j. That is, the expected change in the forward rate applying to loans at a *given* date in the future is zero.[7] The theorem is important because it permits us to utilize the properties of martingale sequences[8] in constructing tests of the expectations theory of the term structure. Two of these properties are particularly useful. First, if the sequence (4) follows a martingale, then

$$E({}_{t+j}F_t - {}_{t+j}F_{t-1}) = 0, \qquad j = 0, 1, 2, \ldots \tag{7}$$

which, in the jargon of the term structure literature, rules out "liquidity premiums." This is a well-known implication of the expectations theory that has been exploited often in empirical work (see Wood [35] and Kessel [15]).[9]

A second property of a martingale is that its increments are uncorrelated, though not necessarily statistically independent.[10] That is, where cov (x,y) denotes the covariance between x and y, it is a characteristic of the martingale (4) that

$$\gamma(1) = \text{cov}({}_{t+j}F_t - {}_{t+j}F_{t-1}, {}_{t+j}F_{t-1} - {}_{t+j}F_{t-2}) = 0, j = 0, 1, 2, \ldots .$$

More generally, the covariances of increments are zero for all lags $k \geqslant 1$:

$$\begin{aligned} \gamma(k) &= \text{cov}({}_{t+j}F_t - {}_{t+j}F_{t-1}, {}_{t+j}F_{t-k} - {}_{t+j}F_{t-k-1}) = 0 \\ j &= 0, 1, 2, \ldots ; \qquad k = 1, 2, 3, \ldots . \end{aligned} \tag{8}$$

[7] Here is Samuelson's proof. First, let the function $f()$ denote the expectation ${}_{t+j}F_{t+1}$,

$$\begin{aligned} {}_{t+j}F_{t+1} &= \int_{-\infty}^{\infty} R dP(R, R_{t+1}, R_t, \ldots ; j-1) \\ &= f(R_{t+1}, R_t, \ldots). \end{aligned} \tag{6}$$

The expectation of ${}_{t+j}F_{t+1}$, conditional on information available at time t, is given by

$$E[{}_{t+j}F_{t+1} | R_t, R_{t-1}, \ldots] = \int_{-\infty}^{\infty} f(z, R_t, R_{t-1}, \ldots) dP(z, R_t, R_{t-1}, \ldots ; 1).$$

Substituting (6) into the above expression yields $\int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} R dP(R, z, R_t, R_{t-1}, \ldots ; j-1)]$ $dP(z, R_t, R_{t-1}, \ldots ; 1)$. Changing the order of integration gives $\int_{-\infty}^{\infty} R d\, [\int_{-\infty}^{\infty} P(R, z, R_t, \ldots ;$ $j-1) dP(z, R_t, \ldots ; 1)]$. But by equation (2) this equals

$$\int_{-\infty}^{\infty} R dP(R, R_t, \ldots ; j) = {}_{t+j}F_t ,$$

which proves the theorem.

[8] See Feller [10] and Doob [7].

[9] However the theoretical foundations of (7) have rarely been explicitly exposed. Again, an exception is Roll [26].

[10] See Doob [7] and Samuelson [27].

The implications of equation (8) can be summarized compactly by noting that it implies that the spectral density of the increments of the sequence (4) is flat or "white." The spectral density $s(w)$ is defined as the following weighted sum of the $\gamma(k)$'s:

$$s(w) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma(k) e^{-iwk} = \frac{1}{2\pi} (\gamma(0) + 2 \sum_{k=1}^{\infty} \gamma(k) \cos kw).$$

But since $\gamma(k)$ is zero for all k not equal to zero,

$$s(w) = \frac{\gamma(0)}{2\pi}, \tag{9}$$

which establishes the spectral "whiteness" of increments in a martingale sequence.

It is this restriction on the increments of the forward rate sequence (4) that we will exploit to perform empirical tests of our two hypotheses. We do not use or even require the condition (7), since it is fairly well established that (7) seems inconsistent with the data, due to the presence of "liquidity premiums."[11] Yet (7) can be violated and (8) and (9) remain valid. For example, it is sufficient that the liquidity premiums be on the average constant through time, which implies the weaker condition

$$E({}_{t+j}F_t \mid R_{t-1}, R_{t-2}, \ldots) = \lambda_j + {}_{t+j}F_{t-1}, \text{ or } E({}_{t+j}F_t - {}_{t+j}F_{t-1}) = \lambda_j, \quad \lambda_j \leqslant 0, \tag{7'}$$

where the λ_j's are constant through time. A sequence obeying (7′) is said to be a "submartingale," a process that has uncorrelated increments and so obeys (8) and (9).

B. A More Restrictive Form of the Model

It is interesting to explore the implications of further restricting the probability distribution of (1) in the following manner: we assume that the one-period spot rate R_t follows a discrete (covariance) stationary stochastic process with finite variance. Then use of Wold's [33] famous theorem permits us to decompose R_t into two mutually uncorrelated processes

$$R_t = \xi_t + \epsilon_t$$

where ξ_t is a deterministic component, predictable with zero mean squared error given an adequate number of its past values, and ϵ_t is a one-sided moving sum of un-

[11] See Kessel [15], Wood [34], and Cagan [2].

correlated random variables,

$$\epsilon_t = \sum_{k=0}^{\infty} c_k u_{t-k}$$

where

$$E(u_t) = 0$$

$$E(u_t^2) = \sigma_u^2$$

$$E(u_t, u_{t-s}) = 0, \qquad s \neq 0.$$

We make the natural assumption that ξ_t is predicted perfectly by investors, an assumption that is the only one compatible with our maintaining Muth's hypothesis that information is not wasted.

Hence one-period spot rates are governed by

$$R_t = \xi_t + \sum_{k=0}^{\infty} c_k u_{t-k} \,. \tag{10}$$

At time t, the minimum-mean-squared-error forecast of the one-period spot rate to prevail at time $t+j$, which equals $E(R_{t+j} | R_t, R_{t-1}, \ldots)$ and which determines ${}_{t+j}F_t$, is given by [12]

[12] Essentially, (11) is derived by replacing $u_{t+j}, u_{t+j-1}, \ldots, u_{t+1}$ by their expected values of zero in the expression

$$R_{t+j} = \xi_{t+j} + \sum_{k=0}^{j-1} c_k u_{t+j-k} + \sum_{k=0}^{\infty} c_{k+j} u_{t-k} \,.$$

More formally, let the forecast be given by

$${}_{t+j}F_t = \sum_{k=0}^{\infty} d_k u_{t-k} + \xi_{t+j} \,.$$

The mean squared error is

$$\begin{aligned} E({}_{t+j}F_t - R_{t+j})^2 &= E\left(\sum_{k=0}^{\infty} d_k u_{t-k} - \sum_{k=0}^{\infty} c_k u_{t+j-k}\right)^2 \\ &= E\left(-\sum_{k=0}^{j-1} c_k u_{t-k} + \sum_{k=j}^{\infty} (d_{k-j} - c_k) u_t\right)^2 \\ &= \sigma_u^2 \sum_{k=0}^{j-1} c_k^2 + \sigma_u^2 \sum_{k=j}^{\infty} (d_{k-j} - c_k)^2 \end{aligned}$$

For fixed c_k's the above expression is minimized by setting $d_{k-j} = c_k$ for all k greater than or equal to j. The fundamental reference on the subject of forecasting processes like (10) is Whittle [32].

$$_{t+j}F_t = \xi_{t+j} + \sum_{k=0}^{\infty} c_{k+j}u_{t-k} \,. \tag{11}$$

This is the particular version of (3) corresponding to the process (10). Similarly,

$$_{t+j}F_{t-1} = \xi_{t+j} + \sum_{k=1}^{\infty} c_{k+j}u_{t-k}. \tag{12}$$

To establish the nature of the increments of the process in (4), we subtract (12) from (11), which yields

$$_{t+j}F_t - {}_{t+j}F_{t-1} = c_j u_t; \qquad j = 0, 1, 2, \ldots . \tag{13}$$

Equation (13) obviously satisfies the martingale condition (5). It is even more restrictive, however, implying that over time each increment $({}_{t+j}F_t - {}_{t+j}F_{t-1})$ is an independent, identically distributed random variable with variance $c_j^2 \sigma_u^2$. This implies that the spectrum of each increment is flat, an implication that can be tested empirically quite easily. Equation (12) states that the j-period-forward rate on one period loans follows an additive "random walk."[13]

It is the more restrictive (but still very general) version of the model that specifies that the P's of (1) can be described by (10) that seems to be implicit in some of the best empirical work on the term structure. To take an outstanding example, suppose that, following Meiselman [19], we estimate "revision" equations of the form

$$_{t+j}F_t - {}_{t+j}F_{t-1} = \beta_j(R_t - {}_tF_{t-1}), \qquad j = 1, 2, \ldots . \tag{14}$$

Then (13) implies that the right and left hand sides of (14) are perfectly correlated. It is in this sense that high R^2's in estimates of Meiselman's equation can be said to confirm the expectations hypothesis. Moreover, note that, maintaining (13), the least squares estimate of β_j is

$$\hat{\beta}_j = \frac{\sum_t c_j u_t \cdot c_0 u_t}{\sum_t (c_0 u_t)^2}$$

$$= \frac{c_j c_0 \Sigma u_t^2}{c_0^2 \Sigma u_t^2} \tag{15}$$

$$\hat{\beta}_j = \frac{c_j}{c_0}.$$

[13]Cootner [4] is a useful reference on random walks applied to asset prices.

Thus, as Stanley Diller [6] has pointed out, if (13) is maintained the family of β's supplies estimates of the c_k's that seem to be generating the data. Whether or not estimating (14) is as a practical matter a reliable means of recovering the c_k's is another question, one that we return to in Section IV.

Other empirical studies of the term structure have frequently analyzed yields to maturity directly, rather than forward rates (e.g. Modigliani and Sutch [21], Malkiel [16], Granger and Rees [13], Wood [34]). Consequently, it is of interest to spell out the implications of this section's somewhat restricted form of the model for the behavior of yields to maturity. Yields to maturity are related to forward rates by the Fisher-Hicks formula,

$$R_{nt} = \frac{1}{n}(R_{1t} + {}_{t+1}F_{1t} + \ldots + {}_{t+n-1}F_{1t})$$

where R_{nt} is the yield to maturity on an n-period bond.[14] Substituting (10) and (11) into the above formula yields

$$R_{nt} = \frac{1}{n}(\xi_t + \xi_{t+1} + \ldots + \xi_{t+n-1}) + \frac{1}{n}\sum_{k=0}^{\infty}(c_k + c_{k+1} + \ldots + c_{k+n-1})\,u_{t-k}$$

or (16)

$$R_{nt} = \theta_{nt} + \sum_{k=0}^{\infty} d_{nk}u_{t-k}$$

where $\theta_{nt} = \frac{1}{n}(\xi_t + \xi_{t+1} + \ldots + \xi_{t+n-1})$, $d_{nk} = (c_k + c_{k+1} + \ldots + c_{k+n-1})/n$.

Relation (16) states that the non-deterministic parts of the yields to maturity for all maturities can be expressed as one-sided moving sums of the *same* serially independent, identically distributed random variable. The particular weights in the moving sum, of course, will in general depend on maturity.

Suppose now that all yields are purely non-deterministic, so that θ_{nt} is zero for all n and t, or that the deterministic parts have been removed by a "de-trending" operation. Then (16) implies that the coherence between yields to maturity for any two maturities is equal to unity over all frequency bands.[15] Thus, taking Fourier transforms of each side of (16) yields

$$R_n(w) = D_n(w)U(w) \tag{17}$$

[14] The formula given is actually an arithmetic approximation to Hicks's formula.

[15] The coherence measures the proportion of variance in one series occurring over some frequency band that can be explained by the variation in another series over the same frequency band. It is analogous to the R^2 statistic of correlation analysis, and like the R^2, it is bounded by zero and unity.

where $R_n(w) = \sum_t R_{nt}\, e^{-iwt}$

$$D_n(w) = \sum_k d_{nk}\, e^{-iwk}$$

$$U(w) = \sum_t u_t\, e^{-iwt}\,.$$

The spectral density of the yield to maturity on n-period bonds is given by

$$\begin{aligned} f_{nn}(w) &= E\,|R_n(w)|^2 \\ &= E\,|D_n(w)\,U(w)|^2 \\ &= f_{uu}(w)\,|D_n(w)|^2 \end{aligned}$$

where E is the mathematical expectation operator, the vertical bars denote the amplitude of the included quantity and $f_{uu}(w)$ is the spectral density of the disturbance process, u, which is known to be "white noise," so that

$$f_{uu}(w) = \frac{\sigma_u^2}{2\pi}\,.$$

The cross spectrum between yields to maturity on n-period and j-period bonds is given by

$$f_{nj}(w) = E\,[R_n^*(w)\,R_j(w)]$$

where the asterisk denotes complex conjugation. Thus,

$$\begin{aligned} f_{nj}(w) &= E(D_n^*(w)\,U^*(w)\,D_j(w)\,U(w)) \\ &= f_{uu}(w)\,D_n^*(w)\,D_j(w)\,. \end{aligned}$$

The coherence coefficient at frequency band w, $\mathrm{coh}(w)$, is defined as the squared amplitude of the cross spectrum divided by the product of the values of the spectral density at that frequency band:

$$\begin{aligned} \mathrm{coh}(w) &= \frac{f_{uu}{}^2\,|D_n^*(w)\,D_j(w)|^2}{f_{uu}(w)\,|D_n(w)|^2\, f_{uu}(w)\,|D_j(w)|^2} \\ &= 1\,. \end{aligned}$$

Thus, on our hypotheses, the coherence between (the non-deterministic parts of) yields to maturity for any two maturities equals unity. Several empirical studies have estimated cross spectra for pairs of yields to maturity of various maturities (e.g. Granger and Rees [13] and Sargent [28]). The calculations above establish that the coherence coefficients estimated in such studies constitute evidence capable of disconfirming the version of the expectation hypothesis being discussed in this section.[16]

The foregoing implications of our hypotheses can be expressed in an equivalent way in terms of the distributed lag relationships that must exist between any two yields to maturity. In particular, (16) implies that a yield of any maturity can be completely explained by a distributed lag function of any other yield. Let us assume again that θ_{nt} equals zero for all n and t, or that any deterministic components of the series have been removed.[17] Then (17) implies that

$$R_n(w) = \frac{D_n(w)}{D_j(w)} R_j(w) .$$

Taking the inverse Fourier transform of the above expression establishes that[18]

$$R_{nt} = H(L)R_{jt} = \sum_{i=0}^{\infty} h_i R_{jt-i} \tag{18}$$

[16]The presence of errors of observation obviously would reduce the implied coherence below unity. Thus, in place of (16), suppose that R_{2t} is governed by

$$R_{2t} = \sum_{k=0}^{\infty} d_{2k} u_{t-k} + \epsilon_t,$$

where ϵ_t is an error of observation assumed to be white noise uncorrelated with u_t. Then the coherence between R_{1t} and R_{2t} is given by

$$\text{coh}(w) = 1 - \frac{f_{\epsilon\epsilon}(w)}{f_{22}(w)}$$

where $f_{\epsilon\epsilon}(w)$ is the spectrum of ϵ (which equals $\sigma_\epsilon^2/2\pi$ where σ_ϵ^2 is the variance of ϵ).

Some economists have calculated the phase of the cross spectrum between yields to maturity, sometimes arguing that on the expectations hypothesis the longer rate ought to lead the shorter rate. However, unless further restrictions on the R-process are added to those that have been introduced in the text, it is impossible to predict the sign of the phase of the cross spectrum.

[17]We also assume that the roots of $\Sigma_{k=0}^{\infty} d_{jk}L^k = 0$ lie outside the unit circle for all j.

[18]Alternatively, note that assuming $\theta_{nt} = 0$, (16) becomes

$$R_{nt} = \left(\sum_{i=0}^{\infty} d_{ni}L^i\right)u_t; \qquad n = 1, 2, \ldots .$$

Assuming that $\Sigma_{i=0}^{\infty} d_{ji}L^i$ is invertible, u_t can be written

$$u_t = \frac{1}{\sum_{i=0}^{\infty} d_{ji}L^i} R_{jt},$$

where $H(L) = \dfrac{\sum_{k=0}^{\infty} d_{nk} L^k}{\sum_{k=0}^{\infty} d_{jk} L^k} = \sum_{i=0}^{\infty} h_i L^i$,

and where L is the lag operator defined by $L^n x_t = x_{t-n}$: The one-sided character of the impulse-response function $H(L)$ follows from the one-sided nature of the d's. The hypotheses set forth in this section thus imply that it is possible entirely to explain the behavior of yields to maturity via the class of one-sided distributed lag functions (18). Fitting distributed lags directly to data on yields to maturity has been the approach followed by de Leeuw [5], Malkiel [16] and Modigliani and Sutch [21].

The one-sided character of the distributed lag functions in (18) suggests a test of the version of the expectations hypothesis presently under consideration, a test that it may sometimes be convenient to employ. That test would involve estimating the "two-sided" distributed lag functions

$$R_{nt} = \sum_{i=-m_1}^{m_2} h_i R_{jt-i} + w_t, \qquad \text{for} \qquad n \neq j, \tag{19}$$

where m_1 and m_2 are positive parameters and w_t is a statistical residual. Significant estimated coefficients on future values of the right-hand variable in (19) would indicate that it is not possible adequately to represent both R_{nt} and R_{jt} as one-sided moving averages of the *same* single white noise. From a generalization of the theorem of Wold cited above, it is known that any pair of indeterministic, covariance stationary stochastic processes, say R_{jt} and R_{nt}, can be represented as pairs of different one-sided moving sums of the same two independent white noises:

$$\begin{aligned} R_{jt} &= \sum_{i=0}^{\infty} a_i u_{t-i} + \sum_{i=0}^{\infty} b_i \epsilon_{t-i} \\ R_{nt} &= \sum_{i=0}^{\infty} d_i u_{t-i} + \sum_{i=0}^{\infty} f_i \epsilon_{t-i} \end{aligned} \tag{20}$$

---

which implies that R_{nt} can be expressed as

$$R_{nt} = \left[\frac{\sum_{i=0}^{\infty} d_{ni} L^i}{\sum_{i} d_{jt} L^i} \right] R_{jt}.$$

where the a's, b's, d's and f's are parameters and ϵ and u are two mutually independent white noises. Sims [29] has proved a theorem that implies that the distributed lag (19) and the reverse distributed lag with R_{nt} on the right and R_{jt} on the left will both be one-sided if and only if all of the b's and all of the f's, or alternatively all of the a's and all of the d's in (20) equal zero.[19] By fitting such two-sided distributed lags it is thus possible to determine whether all yields to maturity can be expressed as one-sided moving sums of the same white noise, as is implied by the version of the expectations theory being considered here.

In conclusion, many empirical models of the term structure can be rationalized within the framework presented in this section. In particular, the contributions of Wood [34], de Leeuw [5], Malkiel [16], Bierwag and Grove [1], and Modigliani and Sutch [21] all fall within this framework. The models of Modigliani and Sutch [21] and Bierwag and Grove [1] seem to be the most general, being less dependent on particular a priori assumptions about parameters. Yet while all of those studies can be rationalized in a general way by appealing to this framework, the tests that they have reported make no use at all of the second of our hypotheses—that forecasts incorporate information efficiently.

It should be noted that our two hypotheses, even in the special form of this section, do not imply things that are commonly thought to be their implications. Thus, they do not imply that the spot one-period rate, R_t, follows a random walk, which would mean that $R_t - R_{t-1}$ is serially uncorrelated. They do not imply that the j-period spot rate follows a random walk for any finite j.[20] Moreover, the "fair game" property built into the model clearly does not mean that spot rates cannot be described by a stable stochastic difference equation.[21]

III. TESTS OF SAMUELSON'S MODEL

A. Spectral Densities of Some Forward Rate Sequences

In this section we report tests of the expectations hypothesis by calculating the spectral densities of forward rate sequences that should be "white" on our main-

[19] See Sims [29]. Notice that the expectations hypothesis *implies* that in this context equation (10) is an adequate representation of the one-period rate. Thus, let u_t be the fundamental white noise process in the univariate (Wold) representation of ${}_{t+1}F_t$,

$$ {}_{t+1}F_t = A(L)u_t = \sum_{j=0}^{\infty} a_j L^j \quad u_t. $$

Then since ${}_{t+1}F_t$ is an efficient, unbiased forecast of R_{t+1}, we have

$$ R_{t+1} = A(L)u_t + \epsilon_{t+1} , $$

where ϵ_{t+1} has mean zero and is uncorrelated with u_s for all $s \neq t+1$. But in general the above expression can be an adequate representation of R_{t+1} only if $\epsilon_{t+1} = u_{t+1}$. Then the above expression becomes equivalent with (10) with $c_0 = 1$ and $c_k = a_{k-1}$ for $k \geqslant 1$. Thus, the expectations hypothesis implies that the forward rates ${}_{t+j}F_t$ contain no information that is useful in predicting R_{1t} over and above that contained in current and past values of R_{1t}.

[20] See Granger and Rees [13] for a statement of another view.

tained hypotheses. The data are those studied by Meiselman, Durand's [5] annual series of basic corporate yields for terms to maturity of one, two, . . . , ten years. We begin by testing the somewhat more restrictive version of the model that assumes that spot short rates follow a stationary process with finite variance and so can be described by equation (10). That specification implies that the variates $({}_{t+j}F_t - {}_{t+j}F_{t-1})$ are serially uncorrelated, and so have a flat or "white" spectrum. For the period 1901-1954, we have estimated the spectral densities of these variates for $j = 1, \ldots, 9$. The estimates were derived using eight lagged covariances and the standard covariance-cosine transformations in conjunction with a Parzen window. The results are recorded in Table 1. The F-statistic pertinent for testing against the null hypothesis of spectral whiteness is also given for each spectrum.[22] High values of F lead to rejection of the null hypothesis. At the 1 percent level of significance, serial independence must be rejected for all j's less than or equal to five, while at the 5 percent level of significance all but the spectra for j's of eight and nine are inconsistent with serial independence. Generally, serial dependence is seen to diminish as

TABLE 1

Spectra of Increments in Forward Rates, 1901–1954 $({}_{t+j}F_t - {}_{t+j}F_{t-1})$*

Period $j+1$ in Years	1	2	3	4	5	6	7	8	9
∞	.209	.123	.080	.051	.045	.026	.017	.017	.019
16.	.170	.097	.063	.042	.035	.021	.014	.014	.010
8.	.103	.055	.036	.026	.020	.014	.009	.010	.008
5.33	.064	.034	.024	.020	.012	.012	.008	.010	.009
4.00	.053	.029	.022	.018	.013	.014	.011	.013	.011
3.20	.061	.031	.023	.019	.017	.017	.015	.015	.013
2.67	.067	.033	.022	.018	.018	.017	.015	.014	.013
2.29	.055	.029	.019	.016	.015	.014	.012	.011	.012
2.00	.045	.025	.017	.015	.014	.014	.011	.009	.011
F-ratio:†	4.65	4.93	4.63	3.33	3.65	2.25	2.09	1.84	1.57

$F_{24,24}(.05) = 1.98, F_{24,24}(.01) = 2.66$

$^*{}_{t+0}F_t \equiv R_t$

†F-ratio equals highest value of spectrum of j divided by lowest value. The values recorded here are based on values of the spectrum calculated to more places than reported in the body of the table.

---

[21] See Gordon and Hynes [12] for a seemingly contrary view about the behavior of commodity prices.

[22] Where $\hat{s}(w)$ is an estimate of $s(w)$, $n\,\hat{s}(w)/s(w)$ is approximately χ^2 with n degrees of freedom. The number n is given by 3.7 times the number of data points divided by the number of lagged covariance terms used in calculating the spectrum. On the assumption that $\hat{s}(w_1)$ and $\hat{s}(w_2)$ are independent, the null hypothesis that $s(w_1) = s(w_2)$ can be tested by using the statistic

$$F = \frac{\hat{s}(w_1)}{\hat{s}(w_2)}$$

which, being the ratio of two independent χ^2 distributions with n degrees of freedom, is distributed according to the F distribution with n, n degrees of freedom.

TABLE 2

Spectrum of Increments* of $({}_tF_{t-9}, {}_tF_{t-8}, \ldots, {}_tF_{t-1}, R_t)$

Period in Years	Spectral Density
∞	.249
16.	.237
8.	.218
5.33	.196
4.00	.159
3.20	.123
2.67	.101
2.29	.080
2.00	.070

F-ratio† = 3.57, $F_{20,20}(.05) = 2.12$, $F_{20,20}(.01) = 2.94$

†F-ratio equals highest value of spectrum divided by lowest value. Because the spectrum was calculated on the basis of 45 separate observations on the *entire* sequence $\{{}_tF_{t-9} \ldots R_t\}$, the degrees of freedom are probably greater than twenty, which is approximately 3.7 times 45 divided by the number of lags, eight. The spectrum above is obviously based on much more data than indicated by this calculation.

*Spectral density was calculated by using a Parzen window to smooth the Fourier transform of the correlogram of the increments of the sequence $\{{}_tF_{t-9}, {}_tF_{t-8}, \ldots, R_t\}$.

j increases, as the F-statistics show concisely. Yet the random-walk model must certainly be rejected for forward rates drawn from the short end of the yield curve. This is disconcerting, since it is for yields to maturity of less than five years maturity that most of the interesting variation of the yield curve occurs. In addition, as described by the standard term-structure formulas, these short-term forward rates are components of the longer-term yields to maturity, and thus the entire yield curve is affected by their misbehavior.

The general version of the model can be tested by estimating the spectrum of increments in the forward-rate sequence (4). That spectrum, which we have estimated by using Parzen window to smooth the Fourier transform of the correlogram, is reported in Table 2. As the F-statistic reveals, serial independence of the increments of the forward rate sequence can be accepted only at a very low significance level, i.e., low probability of type-1 error. Thus, the data tend to disconfirm the implications of even the broader version of the model.

B. Some Estimates of Two-Sided Distributed Lags

Here we report a test of the more restrictive version of Samuelson's model that we described in section IIb. While the test is less powerful than the ones described above, it is more convenient to apply given the nature of the data we are about to examine. The data are monthly observations on yields to maturity of three-month Treasury bills and one-year, two-year, three-year, four-year, and five-year U.S.

government bonds for the period January, 1950 to December 1966.[23] These data obviously do not permit the computation of monthly yields for forward one-month loans, which would be required to apply the techniques of section IIIa directly to the monthly data. However, we can estimate the two-sided distributed lags discussed in section IIb, which constitute the basis for a less powerful test of the expectations hypothesis. The test is less powerful for the reason that, while evidence of two-sided distributed lags permits us to reject the hypothesis, failure to find such evidence does not provide much reason for confidence in the hypothesis. That is because the expectations hypothesis implies not only that those distributed lags will

TABLE 3

x_t = 3-Month Bill Rate, R_{1t} = One-Year Bond Rate (Jan., 1950–Dec., 1966)

$$R_{1t} = \sum_{i=-10}^{10} h_i x_{t-i} + w_t$$

$\lvert i \rvert$	Coefficients on Future Values of x	Coefficients on Lagged Values of x
0	–	.8490
1	.0862	.0437
2	.0343	−.0236
3	.0747	.0264
4	−.0624	.0404
5	.0728	.0223
6	−.0444	.0146
7	.0472	.0212
8	−.0106	.0016
9	.0129	.0124
10	−.0040	−.0098

(Estimated standard error of coefficients = .0301.)

$$x_t = \sum_{i=-10}^{10} k_i R_{1t-i} + w_t'$$

$\lvert i \rvert$	Coefficients on Future Values of R_{1t}	Coefficients on Lagged Values of R_{1t}
0	–	.9319
1	.0319	.0785
2	−.0072	.0532
3	.0004	.0612
4	.0818	−.0346
5	−.0492	.0159
6	.0553	−.0194
7	.0099	.0435
8	.0286	.0002
9	.0378	.0583
10	.0102	−.0022

(Estimated standard error of coefficients = .0336.)

[23]The data are from the Salomon Brothers and Hutzler pamphlet *An Analytical Record of Yields and Yield Spreads*.

be one-sided, but also that they will assume the particular configurations given in (18).

Tables 3, 4, and 5 report estimates of pairs of distributed lags between the three month bill rate, which we denote by x, and yields to maturity R_{jt} on j-year bonds for j equals 1, 3, and 5. For each j the pair of estimated distributed lag functions is

$$R_{jt} = \sum_{i=-10}^{10} h_i x_{t-i} + w_t,$$

$$x_t = \sum_{i=-10}^{10} k_i R_{jt-i} + w'_t,$$

TABLE 4

x_t = 3-Month Bill Rate, R_{3t} = 3-Year Bond Rate (Jan., 1950–Dec., 1966)

$$R_{3t} = \sum_{i=-10}^{10} h_i x_{t-i} + w_t$$

$\|i\|$	Coefficients on Future Values of x	Coefficients of Lagged Values of x
0	–	.6948
1	.0850	.0270
2	.0444	.0084
3	.1255	–.0197
4	–.0747	.0523
5	.0670	.0423
6	–.0421	.0283
7	.0550	.0112
8	–.0140	.0113
9	.0061	.0004
10	.0170	–.0026

(Estimated standard error of coefficients = .0384.)

$$x_t = \sum_{i=-10}^{10} k_i R_{3t-i} + w'_t$$

$\|i\|$	Coefficients on Future Values of R_3	Coefficients on Lagged Values of R_3
0	–	.8631
1	.0484	.1240
2	.0102	.0718
3	–.0475	.1357
4	.0703	–.0781
5	–.0430	–.0092
6	.0693	–.0369
7	–.0130	.0489
8	.0318	–.0254
9	.0184	.0515
10	.0038	–.0008

(Estimated standard error of coefficients = .0483.)

TABLE 5

x_t = 3-Month Bill Rate, R_{5t} = 5-Year Bond Rate, (Jan. 1950–Dec., 1966)

$$R_{5t} = \sum_{i=-10}^{10} h_i x_{t-i} + w_t$$

$\|i\|$	Coefficients on Future Values of x	Coefficients on Lagged Values of x
0	–	.6145
1	.0854	.0113
2	.0245	.0334
3	.1262	−.0345
4	−.0416	.0349
5	.0488	.0481
6	−.0432	.0104
7	.0619	.0330
8	−.0214	.0112
9	.0067	.0005
10	.0153	−.0069

(Estimated standard error of coefficients = .0383.)

$$x_t = \sum_{i=-10}^{10} k_i R_{5t-i} + w'_t$$

$\|i\|$	Coefficients on Future Values of R_{5t}	Coefficients on Lagged Values of R_{5t}
0	–	.8801
1	.0448	.1623
2	.0380	.0579
3	−.0674	.1511
4	.0261	−.0543
5	−.0103	−.0193
6	.0570	−.0328
7	−.0071	.0444
8	.0395	−.0367
9	−.0030	.0235
10	−.0003	−.0003

(Estimated standard error of coefficients = .0566.)

where w and w' are statistical residuals and the observations are measured in deviations from their means. No prior constraints were placed on the distributed lag coefficients, which were estimated by the application of Hannan's "inefficient estimator." The spectral densities and cross spectra that went into calculating the distributed lag coefficients were estimated using standard covariance-cosine transformations together with a Parzen window. Thirty was the maximal lag used in calculating the spectral densities and cross spectra from the covariograms.[24]

The distributed lags with x on the right side appear to be two-sided. The largest coefficients (with the exception of the one on current x) always appears on a future

[24] For a description of the estimator, see Fishman [11].

value of x. In each regression two or more coefficients on future x's are more than twice their estimated standard error. On the other hand, the regressions with R_{jt} on the right side appear more one-sided. In these regressions, the larger coefficients tend to fall on lagged values of R_{jt} (the $j = 1$ regression is the exception), and some of the coefficients on lagged values of R_{jt} are sizable relative to their estimated asymptotic standard errors.

Since the expectations hypothesis implies that both of these distributed lag functions will be one-sided, these results tend to disconfirm that hypothesis. It does not appear possible adequately to represent all yields to maturity by a set of one-sided moving sums of the same white noise.[25]

IV. COMPARISON WITH MEISELMAN'S MODEL

The negative results of the last section should be compared with those of David Meiselman who took his estimates of the set of revision equations (14) to be consistent with the expectations hypothesis. Meiselman's estimates are reproduced in Table 6. We have already shown that the version of our model presented in Section IIb is compatible with those revision equations, and that it implies that the correlations will be "high," strictly speaking, equal to unity. Thus, in our framework, the

TABLE 6

Meiselman's Regressions (1901–1954)

$$_{t+j}F_t - {}_{t+j}F_{t-1} = \hat{\alpha}_j + \hat{\beta}_j(R_t - {}_tF_{t-1})$$

j	$\hat{\alpha}_j$	$\hat{\beta}_j$	$R_A{}^{2\dagger}$	$DW^\ddagger$
1	.0004	.7030	.9051	1.6687
	(.0236)	(.0312)		
2	–.0024	.5256	.7472	1.7271
	(.0316)	(.0419)		
3	–.0069	.4035	.5825	1.8183
	(.0352)	(.0466)		
4	–.0304	.3263	.4544	1.8120
	(.0367)	(.0486)		
5	–.0146	.2768	.4005	2.1907
	(.0346)	(.0459)		
6	–.0074	.2338	.3771	2.3005
	(.0307)	(.0406)		
7	–.0132	.2387	.3885	2.1715
	(.0306)	(.0405)		
8	.0075	.2080	.3356	2.4292
	(.0298)	(.0395)		

*Estimated standard errors are in parentheses.

†R_A^2 denotes R^2 adjusted for loss of degrees of freedom.

‡DW denotes Durbin and Watson's statistic

[25] For what it is worth, we note that the results imply that in using the one-sided distributed lag model of deLeeuw [5], Malkiel [16], and Modigliani and Sutch [21], it is more appropriate to put the long rate on the right side, as De Leeuw and Malkiel do. Modigliani and Sutch employ the short rate as the regressor.

fact that Meiselman's R^2's are not close to one is to be interpreted as another aspect of the pessimistic evidence presented above. The fact that Meiselman found a "plausible" pattern of revision coefficients is in no way inconsistent with the findings of this study.

Our findings do imply that there are important non-forecasting determinants of the yield curve, and that this should be taken into account in estimating equations like Meiselman's. Thus, to account for the low R^2's in Meiselman's equations, let us replace (13) with

$$_{t+j}F_t - {}_{t+j}F_{t-1} = c_j u_t + v_{jt}, \qquad j = 0, 1, 2, \ldots \tag{21}$$

where v_{jt} is a random term with the properties

$$\begin{aligned} &E(v_{jt}) = 0 && j = 0, 1, 2, \ldots \\ &E(v_{jt} u_t) = 0 && j = 0, 1, 2, \ldots \\ &E(v_{jt} v_{kt}) = 0 && \text{for all } j \neq k. \end{aligned} \tag{22}$$

The random variable v_{jt} represents stochastic non-forecasting determinants of the increment in the forward rate maturing at $t + j$, and is assumed to be uncorrelated across maturities, the simplest assumption that can be used to illustrate the point under discussion. It is necessary to introduce the v_{jt}'s in (21) if we want to account for the low R^2's in Meiselman's equations and also maintain the optimal forecasting hypothesis that we have used throughout this paper. Thus positing (21) amounts to retaining the assumption of rational forecasting but permitting stochastic factors other than expectations to play a role in determining the yield curve.

In (15) we established that $\hat{\beta}_j$ could be used, if (13) were maintained, to obtain an estimate of the c_k's of (10). What happens if (13) is replaced by (21)? Then the least squares estimate $\hat{\beta}_j$ in the revision equation is given by

$$\hat{\beta}_j = \frac{\frac{1}{T}\sum_t (c_j u_t + v_{jt}) \cdot (c_o u_t + v_{ot})}{\frac{1}{T}\sum_t (c_o u_t + v_{ot})^2} .$$

Performing the indicated multiplication and taking probability limits yields

$$\begin{aligned} \text{plim } \hat{\beta}_j &= \frac{c_j c_o \text{ plim } T^{-1} \sum_t u_t^2}{c_o^2 \text{ plim } T^{-1} \sum_t u_t^2 + \text{plim } T^{-1} \sum_t v_{ot}^2} \\ &= \frac{\dfrac{c_j}{c_o}}{1 + \dfrac{\text{var}(v)_o}{c_o^2 \text{ var}(u)}} < \frac{c_j}{c_o} . \end{aligned} \tag{23}$$

Thus, under (21) and (22), least squares estimates of β_j provide estimates of c_j/c_o that are asymptotically biased downward.

This finding provides an explanation for a somewhat curious feature of Meiselman's empirical results. Various writers (e.g., Pye [24], Diller [6], Mincer [20]) have noted that if (10) can be well approximated by a first-order Markov process, Meiselman's revision coefficients ought to decline geometrically as j increases, provided expectations are minimum-mean-squared-error forecasts. For suppose

$$R_t = \alpha R_{t-1} + u_t, \qquad \alpha < 1$$

or

$$R_t = \sum_{k=0}^{\infty} \alpha^k u_{t-k} .$$

Then according to (15), $E(\hat{\beta}_j) = \alpha^j$.

It has also been noted that Durand's one-year spot rates seem to be adequately approximated by a first-order autoregressive process. Thus, for the period 1905-1954,[26] the following regression was obtained by the method of least squares:

$$\underset{}{R_t} = \underset{(.1995)}{.2073} + \underset{(.0544)}{.9278}\, R_{t-1}, R_A^2 = .8554 .$$

Including additional lagged R's resulted in a drop in the adjusted R^2. According to (15), if the above equation is accepted as an adequate description of spot rates, then under the hypothesis of rational expectations $\hat{\beta}_1$ ought to have an expected value of about .93. Meiselman's estimate is only .703. There is thus a very sizable difference between these estimates that is consistent in sign with the prediction of equation (23).[27] Since the estimated standard errors of both of those estimates are small compared with the difference between the estimates, the indication is that the difference between them should be taken seriously.[28]

[26]Higher order autoregressions were also calculated with the result that they explained no more of the adjusted variance of R_t than did the first-order regression reported. Our holding back the few observations necessary to calculate the higher order autoregressions explains why the period over which the first-order regression was run, 1905-1954, excludes the first few observations included in Meiselman's regressions.

[27]Estimating a first-order autoregression by least-squares results in the famous "Hurwicz bias." For a positive autoregression parameter, least squares underestimates the parameter on the average. (See Marriott and Pope [18] and Kendall [14].) The bias is approximately $E(\hat{\alpha}) - \alpha = -(1 + 3\alpha)/N$ where α and $\hat{\alpha}$ are the autoregression parameters and its least-squares estimate and N is the sample size. Thus, accounting for that bias would only strengthen the argument in the text.

[28]An alternative reconciliation between these two estimates might be offered by appealing to the errors of measurement which most likely infest Durand's data. However, the assumptions of the classical errors-in-variable model, which might be invoked to explain a downward bias in Meiselman's estimate of β_1, are surely inappropriate here. Durand's smoothing procedures guarantee that measurement errors will be highly correlated along the yield curve. In the presence of measurement errors with such properties, the least-squares estimate of β_1 need not be biased downward and may even be biased upward if the measurement errors are sufficiently highly correlated along the yield curve.

V. CONCLUSIONS

The evidence summarized above implies that it is difficult to maintain both that only expectations determine the yield curve[29] *and* that expectations are rational in the sense of efficiently incorporating available information. The predictions of the random walk version of the model are fairly decisively rejected by the data, particularly for forward rates with less than five years term to maturity. This is important because that is the form of the model that provides a rationale for many formulations utilized in empirical work.

It is clear that our conclusions apply with equal force to the diluted form of the expectations hypothesis that allows forward rates to be determined by expectations plus time-invariant liquidity premiums. Incorporating such liquidity premiums would in no way change the covariances and the spectral densities on which our tests were based. On the other hand, it would clearly be possible to determine a set of time-dependent "liquidity premiums" that could be used to adjust the forward rates so that the required sequences would display "white" spectral densities.[30] Most of the literature on "liquidity premiums" can be interpreted as an attempt to "prewhiten" the data so that just this is accomplished. While this procedure has its merits in certain instances, it is essentially arbitrary, there being no adequate way to relate the "liquidity premiums" so derived to objective characteristics of markets, such as transactions costs. Their arbitrary nature probably explains the considerable disarray in which the literature on the subject stands.[31]

An alternative way to "save" the doctrine that expectations alone determine the yield curve in the face of empirical evidence like that presented above is to abandon the hypothesis that expectations are rational. Once that is done, the model becomes much freer, being capable of accommodating all sorts of ad hoc, plausible hypotheses about the formation of expectations. Yet salvaging the expectations theory in that way involves building a model of the term structure that, while requiring there be no room for profitable arbitrage on the basis of current expectations of the future, also permits expectations to be formed via a process that could utilize available information more efficiently and so enhance profits. That seems to be an extremely odd procedure.

LITERATURE CITED

1. Bierwag, G. and M. Grove. "A Model of the Term Structure of Interest Rates," *Review of Economics and Statistics*, (February, 1967).
2. Cagan, Phillip. "A Study of Liquidity Premiums on Federal and Municipal Government Securities,"in *Essays on Interest Rates*, *Volume 1* (NBER), ed. J. Guttentag and P. Cagan. New York: Columbia University Press, 1969.

[29] Needless to say, the use of Durand's data, which are subject to substantial error, constitutes an important limitation on the confidence with which the empirical results of this study, and the host of other studies that have used those data, can be viewed.

[30] See the appendix for an example of what we have in mind.

[31] Compare Wood [34, 36], Van Horne [30, 31], Kessel [15], Roll [25], Cagan [2] and Diller [6].

3. Conard, Joseph W. *The Behavior of Interest Rates.* National Bureau of Economic Research. Columbia University Press: New York, 1966.
4. Cootner, Paul (ed.). *The Random Character of Stock Market Prices.* Cambridge: M.I.T. Press, 1964.
5. deLeeuw, Frank. "A Model of Financial Behavior," in *The Brookings Quarterly Econometric Model of the United States*, ed. Duesenberry et. al. Chicago: Rand McNally, 1965.
6. Diller, Stanley. "Expectations in the Term Structure of Interest Rates," in *Economic Forecasts and Expectations* (NBER), ed. J. Mincer. New York: Columbia University Press, 1969.
7. Doob, J. L. *Stochastic Processes.* New York: John Wiley & Sons, 1953. Chap. 2, esp. pp. 91–94.
8. Durand, David. *Basic Yields of Corporate Bonds, 1900–1943*, Technical Paper No. 3. New York: NBER, 1943.
9. Fama, Eugene F. "Efficient Capital Markets: A Review of the Theory and Empirical Evidence," *Journal of Finance*, 25 (May, 1970).
10. Feller, William. *An Introduction to Probability Theory and Its Applications, Volume II*, New York: Wiley, 1966.
11. Fishman, George S. *Spectral Methods in Econometrics*, Harvard University, Cambridge, Mass.: Harvard University Press, 1969. Chap. 4.
12. Gordon, Donald F. and Allan Hynes. "On the Theory of Price Dynamics," in *Microeconomic Foundations of Employment and Inflation Theory*, ed. Phelps et al.
13. Granger, C.W.J. and H.J.B. Rees. "Spectral Analysis of the Term Structure of Interest Rates," *Review of Economic Studies*, 35 (January, 1968).
14. Kendall, M.G. "Note on the Bias in Estimation of Autocorrelation," *Biometrika*, 41 (1954).
15. Kessel, Ruben. *The Cyclical Behavior of the Term Structure of Interest Rates*, NBER Occasional Paper 91. New York: Columbia University Press, 1965.
16. Malkiel, Burton. *The Term Structure of Interest Rates*, Princeton: Princeton University Press, 1966.
17. Mandelbrot, Benoit. "Forecasts of Future Prices, Unbiased Markets and Martingale Models," *Journal of Business*, 39 (Special Supplement, January, 1966).
18. Marriott, F.H.C. and J.A. Pope. "Bias in the Estimation of Autocorrelations," *Biometrika*, 41 (1954).
19. Meiselman, David. *The Term Structure of Interest Rates*. Englewood Cliffs, N. J.: Prentice-Hall, 1962.
20. Mincer, Jacob. "Models of Adaptive Expectations," in *Economic Forecasts and Expectations*, (NBER), ed. J. Mincer. New York: Columbia University Press, 1969.
21. Modigliani, Franco and Richard Sutch. "Innovations in Interest Rate Policy," *American Economic Review, Papers and Proceedings*, (May, 1966).
22. Muth, John F. "Rational Expectations and the Theory of Price Movements," *Econometrica* (July, 1961).
23. Phelps, Edmund. "The New Microeconomics in Employment and Inflation Theory," in *Microeconomic Foundations of Employment and Inflation Theory*, ed. E. Phelps, et al. New York: Norton, 1970.
24. Pye, Gordon. "A Markov Model of the Term Structure of Interest Rates," *Quarterly Journal of Economics* (February, 1966).
25. Roll, Richard. "Interest Rate Risk and the Term Structure of Interest Rates: Comment," *Journal of Political Economy*, (1966).
26. ——. "The Efficient Market Model Applied to U.S. Treasury Bill Rates" (forthcoming).
27. Samuelson, Paul A. "Proof that Properly Anticipated Prices Fluctuate Randomly," *Industrial Management Review*, 6 (Spring, 1965).
28. Sargent, Thomas. "Interest Rates in the Nineteen-Fifties," *Review of Economics and Statistics* (1968).

29. Sims, Christopher A. "Money, Income, and Causality" (manuscript, National Bureau of Economic Research, February, 1971).
30. Van Horne, James. "The Expectations Hypothesis, the Yield Curve, and Monetary Policy: Comment," *Quarterly Journal of Economics*, (1965).
31. ——. "Interest Rate Risk and the Term Structure of Interest Rates," *Journal of Political Economy*, (1965).
32. Whittle, P. *Prediction and Regulation by Linear Least-Square Methods*, London: English University Press, 1963.
33. Wold, H. *A Study in the Analysis of Stationary Time Series*, Stockholm: Almqvist and Wiksell, 1938.
34. Wood, John. "The Expectations Hypothesis, the Yield Curve, and Monetary Policy," *Quarterly Journal of Economics* (1964).
35. ——. "Expectations, Errors and the Term Structure of Interest Rates," *Journal of Political Economy* (1963).
36. ——. "The Expectations Hypothesis, the Yield Curve, and Monetary Policy: Reply," *Quarterly Journal of Economics* (1965).

APPENDIX ON KESSEL'S METHOD OF ESTIMATING LIQUIDITY PREMIUMS

In the light of the approach taken in this paper, Kessel's [15] method of estimating liquidity premiums deserves a few comments. Kessel posited that forward rates are the sum of the appropriate expected rate and a liquidity premium. In particular,

$$_{t+1}F_t = {}_{t+1}E_t + {}_{t+1}L_t$$

where $_{t+1}E_t$ is the market's expectation at time t of the one-period rate to prevail at time $t + 1$, and $_{t+1}L_t$ is the one-period forward liquidity premium. Following Kessel, we posit that the systematic part of the liquidity premium is linearly related to the one-period spot rate

$$_{t+1}L_t = k_o + k_1 R_t + w_t \tag{a}$$

where k_o and k_1 are parameters and the w_t's are independent, identically distributed random variables with mean zero and which are distributed independently of R_t. Further, we posit that the one-period spot rate is governed by a first-order autoregressive process

$$R_{t+1} = h_o + h_1 R_t + u_{t+1}, \qquad |h_1| \leqslant 1 \tag{b}$$

where u is a "white-noise" process that is independent of the w process and of R_t. Durand's data appear to be well described by such a process. Now Kessel's approach to estimating the dependency of the liquidity premium on the spot rate was to regress $({}_{t+1}F_t - R_{t+1})$ against R_t; the coefficient on R_t was then interpreted as an estimate of k_1 in (a). On the assumption that (b) describes the R process, and on the assumption that expectations are rational so that $_{t+1}E_t = h_0 + h_1 R_t$, Kessel's method provides a sensible means of estimating (a). For on these assumptions ${}_{t+1}F_t$ is given by

$$ {}_{t+1}F_t = (h_0 + k_0) + (h_1 + k_1)R_t + w_t $$

and Kessel's dependent variable, ${}_{t+1}F_t - R_{t+1}$, is given by

$$ {}_{t+1}F_t - R_{t+1} = k_0 + k_1 R_t + w_t - u_{t+1} \,. \tag{c} $$

Since w_t and u_{t+1} are statistically independent of R_t, estimating (c) provides a means of recovering the parameters of the liquidity-premium function (a). Thus, on the assumptions listed above, Kessel's method seems sensible and immune to the kind of criticism made of it by Conard [3], who claimed that the method was deficient because it failed to control adequately for "inertia" in the formation of expectations. On the other hand, the method's validity is predicated on the R-process being first-order autoregressive. Conard's criticism might be rationalized as reflecting his belief that for the data Kessel studied, a higher order autoregressive process would be required to describe the evolution of the spot rate adequately.

Notice that, assuming (a) and (b), the variable $R_{t+1} - {}_{t+1}F_t$, which occupied an important role in our tests, will in general be serially correlated. For by (c), $R_{t+1} - {}_{t+1}F_t$ will be serially correlated as long as k_1 does not equal zero and R_t is serially correlated. It is obvious that $R_{t+1} - {}_{t+1}F_t$ can be "pre-whitened" by adding to it the systematic part of the liquidity premium, in an attempt better to approximate $R_{t+1} - {}_{t+1}E_t$. For we know that,

$$ R_{t+1} - {}_{t+1}F_t + k_0 + k_1 R_t = u_{t+1} + w_t \,, $$

which by hypothesis is serially uncorrelated.