Volume Title: Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance

Volume Author/Editor: Jacob A. Mincer, editor

Volume Publisher: NBER

Volume ISBN: 0-870-14202-X

Volume URL: http://www.nber.org/books/minc69-1

Publication Date: 1969

Chapter Title: Expectations in the Term Structure of Interest Rates

Chapter Author: Stanley Diller

Chapter URL: http://www.nber.org/chapters/c1217

Chapter pages in book: (p. 112 - 166)

FOUR

Expectations in the Term Structure of Interest Rates

STANLEY DILLER

I. INTRODUCTION

Since J. R. Hicks' *Value and Capital* [6] appeared, and even before, a controversy has persisted over what determines the yield differentials among securities identical except for the term left to maturity. Formerly a technical subject on the fringe of monetary affairs, the term structure of interest rates has recently become a policy issue, involving as it does the relationship between long- and short-term interest rates. "Operation Twist," for example, a widely publicized policy in the early 1960's, involved the government's attempt to keep the long-term rate low in order to encourage domestic investment and the short-term rate high to discourage capital outflow through adjusting the supply of different maturities.

The effectiveness of this policy depends, of course, on the determinants of the term structure. In this respect, there are essentially two points of view: One holds that the market for securities consists of a group of separate nonoverlapping markets defined for different maturities, with no tendency for the rates in the different markets to assume any particular relation to each other. This point of view, stated here in its extreme, says in effect: that there is no theory of the term structure. On the other side are those who regard the securities market as a collection of interrelated markets and the term structure of rates as subject to some unifying principle. Ordinarily, the principle involves some form of forecasting. The term structure, according to the second view, is determined, at least in part, by the market's forecasts of future rates. David Meiselman, who sparked the latest round of discussion, argues vigorously for the so-called pure expectations hypothesis [11]. More recently, Reuben Kessel [8] revitalized Hicks' idea that the term structure depends on some combination of market anticipations and liquidity preference.

For our own work we tentatively accept the expectations hypothesis that the yield differentials are explainable in terms of market forecasting (with or without the liquidity component) and consider how the forecasts are actually formed. We find that a substantial part of the variation of the forecasts inferred from the term structure is related to an extrapolation of past spot rates. In other words, a moving average of past spot rates can predict a substantial part of the implicit forecasts themselves. We find, however, that the forecasts are no more (and even a little less) effective in forecasting future spot rates than are *linear* extrapolations of past spot rates; that is, extrapolations based on moving averages whose weights are specifically selected to yield the best predictions of future spot rates.

Most of this chapter is devoted to an interpretation of these findings. In particular, one can interpret the linear extrapolative model as a summary of many explicit behavioral models in the same way that a resultant force can be said to summarize the effects of component forces. Since it is the resultant that we observe, the issue becomes: What can we infer about the behavioral models contributing to the final effect?

In Section III we investigate the relationship between the error-learning mechanism that Meiselman used to test the expectations hypothesis and the extrapolative model. From the statistical results of the error-learning model reported by Meiselman we are able to infer a particular method of extrapolation and show that this method fits the actual data better than some plausible alternative methods.

The significance of this experiment lies in our ability to generalize the work of Meiselman and others into a form from which one can draw additional implications about the way market forecasts are made.

One characteristic of the term structure that has received wide attention and evoked general agreement, from Keynes to Kessel, is the inverse relation that exists between the slope of the yield curve and the level of short rates. This observation is ordinarily explained by the so-called expected return to normalcy hypothesis: When current rates deviate from their normal level, future rates are expected to move in the direction of the normal level. Part of Section III demonstrates that this hypothesis is implied by the particular extrapolative model that fits Meiselman's data. As before, the significance of this experiment is the connection it provides between the apparently mechanical procedure of extrapolation and an actual behavioral model.

Finally, we consider the relationship between economic indicators and extrapolative forecasting. Most of the observed connection between the indicators and the forecasts is picked up in the extrapolative procedure. This result is due to the common variation between the indicators and spot rates over the course of the business cycle. The extrapolative procedure implicitly takes account of this common relation. There is a net relation between the indicators and the forecasts that is independent of their common cyclical variation. This relation, it appears, grows as the span of forecast increases.

When we compare the accuracy of the implied forecasts with an autoregressive model designed to exhaust the full extrapolative potential of the data, we find the implied forecasts are inferior. This result follows largely from our analysis of the apparent method of market forecasting. Since the forecasts are based primarily on extrapolations, they are no more effective than the autoregressive model, which is itself an extrapolation. The margin of inferiority of the forward rate forecasts may also be due to the presence of a liquidity premium or some other nonforecasting component of the forward rates that obscures the variation of the forecast component.

The accuracy analysis also uncovers a bias in the forecasts: They are shown to be consistently too high. Kessel, having obtained a similar result, attributed it to the presence of a liquidity premium.

In summary, a substantial part of the variation of the yield differentials can be explained by relating them to market forecasts; the variation of the forecasts, in turn, is in large measure determined by an extrapolative procedure; and the extrapolative procedure is consistent with and related to several behavioral models that have been separately proposed as determinants of the term structure.

II. THE TERM STRUCTURE OF INTEREST RATES

FUTURE PRICES

There are various sets of time series data on individual and group anticipations that would be amenable to the analysis on which this study is based; surveys of business and consumer expectations, forecasts of national income components, sales forecasts, and so forth. There are far less data available on forecasts attributable to a market consensus. Perhaps the best example of this type of forecast is to be found in price data of future transactions. After allowing for various business costs, such as storage and default risk, the current price at which a commodity to be delivered in the future is transacted should, in principle, be equal to the price that is currently expected to prevail at that point in the future. In the special case where expectations are unanimously held and the market responds only to expected values, ignoring the risk of capital loss that fluctuating prices entails (leaving aside transactions and other business costs), the market is at the margin indifferent as to whether it transacts a future commodity at the forward price or waits until the delivery date and transacts at the then spot price. Unanimity is assumed because otherwise we would have to explain why a market deviant does not continue to transact until his funds are exhausted, borrowing to transact further until the discrepancy between his and the market's views are obliterated. It would contribute, perhaps, to verisimilitude if we explained the absence of indefinite transactions by assuming that expectations are not held with certainty and that the uncertainty increased with the number of transactions; or alternatively, that the existing capital market precludes the availability of an unlimited amount of funds to the market deviant. In this and other respects we have abstracted from descriptive complexity. The assumption that dispersion of prices is not relevant will be discussed later.

A particular kind of future price that has recently evoked wide discussion is the price of forward loans. Unlike the prices of commodity futures, the prices of forward loans are not explicitly quoted but are rather implied in the term structure of interest rates. Since these prices are free of the business costs that are part of commodity futures, their expectations content is more immediate; although a possible market response to the gamble is still present. The price of forward loans, like other forward prices, may include a risk premium. In the following pages we review briefly the determinants of the term structure, how the forward rates of return are inferred from the rates of return of securities differing only in the term left to their maturity, and the evidence for equating the forward rates with the expected rates. The recent literature has discussed this part of the subject quite thoroughly; therefore, this study will give it minimum coverage.

A yield curve for a given year and a given type of security is a locus of points relating the rate of return of a security, on the vertical axis, to the term remaining until its maturity, or the number of years remaining before the security is paid off, on the horizontal axis. A typical point on a yield curve for high-grade bonds reveals, as of a given time, the year for which the curve is drawn, the yield to maturity of a bond with a given term to maturity. On the same curve, another point shows the yield on a bond with a different term, and so on, each point for a different term to maturity. For a security of a given term, the yield on the curve associated with that term is the discount factor that is used to equate a stream of fixed payments—that is, the annual coupon payments plus the par value when the bond matures—with the present value of the security. The well-known formula for this computation is:

$$PV = \frac{C_1}{1+R} + \frac{C_2}{(1+R)^2} + \cdots + \frac{C_n + P}{(1+R)^n},$$

where: PV = present value of current market price of security; C_i = coupon payments; P = principle; and R = market yield of bond. Since in any one year there may be several bonds on the market with the same term to maturity but with somewhat different characteristics, it is necessary to fit a line to the points to reduce the array to a unique yield for each term to maturity.[1]

[1] See Durand [4]. The complete set of Durand data is listed in the National Industrial Conference Board's *Economic Almanac, 1967–1968*, p. 416.

A long-term yield is essentially an average of shorter-term yields covering the same period, although the particular form of this average depends on the assumption made with respect to the method of payment.[2] From the yields of two securities, identical except for the terms left to their maturities, one may infer a forward rate of return r that would apply to loans beginning at the time the shorter of the two securities matured and ending when the longer one matured. The formula for this computation of an i period loan is as follows: [3]

$$r_n = \frac{(1 + R_n)^n}{(1 + R_{n-i})^{n-i}} - 1.$$

With this formula one can compute a table of forward rates from the term structure of long-term rates that would reveal the rates of interest on forward loans up to n periods in the future.

THE EXPECTATIONS HYPOTHESIS

The purchase of a security with, say, ten years to maturity is conceptually identical with and can be regarded as the purchase of ten one-year securities that materialize consecutively from the time they were all purchased until the end of the tenth year.[4] Leaving aside attitudes toward the risk of capital loss or of fluctuations of income, since each investor has the option of waiting until a particular security is available and purchasing it at the then spot price, any tendency for the forward yield to deviate from the spot yield expected to prevail should be countered by a change in the demand for the source. For example, if the forward yield of a one-period security available five periods hence exceeds the yield that the market currently expects will prevail on the spot market five years hence, there should be enough investors around who, instead of waiting to deal on the spot market five years later, will move to buy the apparently cheap security now. The increased demand for the security would raise its current price until its yield approaches the expected spot yield. This mechanism is, of course, symmetrical.

This simple idea underlies the hypothesis that the forward rates are equal to the spot rates expected in the future. Since this equation re-

[2] See Macaulay [9, p. 29].

[3] This idea is explained in Hicks [6, pp. 1 and 5] and, in greater generality, in Wallace [16, Chapter I].

[4] See Wallace [16].

quires the willingness of a sufficient number of investors to implicitly transact in the forward market by rearranging the maturity mixture of their portfolios, a priori evaluations of the hypothesis hinge on whether a sufficient number of investors are, in fact, willing to alter their portfolios in response to anticipations.[5] The hedging theory holds that the markets for short- and long-term securities are independent and, therefore, that the equilibrating mechanism between forward and expected rates is nonexistent.[6] The liquidity preference theory acknowledges a relationship among the markets for different maturities (or, what amounts to the same thing, the substitutability of the different securities) but stops short of recognizing a single market for all maturities. In this theory the variance of the prices of securities is a direct function of maturity (not because prices of longer-term securities change more often, but because they fluctuate over a wider range); and, other things the same, the greater the variance the less valuable the security. Given two securities, identical except for maturity, the longer-term security would have to yield more to make the average investor indifferent between them. Therefore, between the pure expectations and the pure hedging theories there is a continuum of degrees of substitutability assumed to exist among the different maturities. At the one extreme, the substitutability is infinite, and the term structure of rates depends solely on expectations; at the other, the substitutability is zero, and the term structure of rates is determined by the supply and demand for each maturity. In between these extremes, the assumed substitutability depends on the relative degrees of price fluctuations that investors are assumed to anticipate, as well as the assumed extent of investors' abhorrence of risk or, obversely, their preference for insurance. In this case, the yield differential necessary for investors' indifference is the insurance or liquidity premium.[7]

In this intermediate position, there are two separate factors affecting the degree of substitutability among the various maturities: (1) in-

[5] The hypothesis does not imply one market for all securities regardless of maturity; nor, in the case of several markets, does it imply that all investors be indifferent about the maturity structure of their portfolios. It is necessary only that the several markets overlap and contain in their overlapping sections a sufficient number of lenders and borrowers whose respective cross elasticity of demand for or supply of securities of different terms to maturity is infinite. See Meiselman [11, Chapter 1].

[6] See Culbertson [3, pp. 485–517].

[7] In this context the word "liquidity" refers strictly to the expected variance of security prices; the smaller the variance the greater the liquidity.

vestors' expectations of the variance of prices of different maturities; and (2) given the expected variances, investors' attitude with respect to the degree of risk of capital loss. Phillip Cagan, in his discussion of the first point, proposes that the differences in the expected variances are inversely related to the deviation of the actual level of rates from the normal level.[8] As for the second point, Reuben Kessel [8] finds that the required premium is greater the higher the level of rates. In both cases, since the liquidity premia will vary over time, there is no simple way to correct the forward rates for this factor in order to isolate the expectations component. The attempt to ascertain the basis for the forecast or the accuracy of the forecast is therefore marred by the presence of the varying liquidity premia.

The evidence that is typically offered for the existence of liquidity premia is the well-known tendency for yields on long-term securities to exceed those on short-term securities.[9] Any attempts, therefore, to measure the accuracy of forecasts by computing the mean square of the differences between the forward rates and the target spot rates will reveal a large error because of this bias. If, instead, the target rates are regressed on the forward rates,[10] the bias in the forward rates will fall out in the constant term. But the varying liquidity premia will lower the correlation and bias the regression coefficient toward zero, except for the special case when they are linearly related to the level of rates. In any case, the presence of a nonforecasting component, whether liquidity premia or some other random variable, hinders the evaluation of forecasts of interest rates.

III. LINEAR AUTOREGRESSIVE FORECASTING MODELS

LONG-TERM DATA

Error-learning model. The question of whether forward rates make accurate forecasts is, of course, independent of the question of whether

[8] See his "A Study of Liquidity Premiums on Federal and Municipal Securities" in *A Study of Interest Rates* [1].

[9] This phenomenon is often reversed in periods of high interest rates—a fact that is considered later in the study. For reasons not entirely clear, the reverse is also true in the mortgage market, where longer-term mortgages tend to yield less than shorter-term mortgages. On this point see Jack Guttentag's study of mortgage yields [1].

[10] The justification for this procedure is discussed more fully in Section IV.

forward rates are forecasts. Forecasting economic time series is notoriously difficult, even series that are far more stable than interest rates. Therefore, evidence of poor accuracy in no way impugns the expectations hypothesis or its liquidity preference variant. In his important study on the term structure, David Meiselman [11, Chapter II] devised an ingenious test to determine whether forward rates are, in fact, forecasts. His idea was to find some characteristics of known forecasts whose presence in a set of numbers would constitute evidence that this set behaved as though it consisted of forecasts. He formulated the characteristic he chose to isolate in terms of the error-learning model. Other studies in other areas have indicated a feedback mechanism in forecasting, whereby the error observed currently of previous forecasts inspires revisions of forecasts referring to a later period. Since the expectations hypothesis asserts that forward rates are, or contain, estimates of expected spot rates, a demonstration that forward rates follow a pattern similar to that of many known sets of forecast data is prima facie evidence that forward rates are or contain forecasts. The model that Meiselman tested is as follows:

$$ {}_{t+n}r_t - {}_{t+n}r_{t-1} = a + \gamma_n(R_t - {}_tr_{t-1}) + u, $$

where ${}_{t+n}r_t$ = the forecast made in period t of the rate expected in period $t + n$; ${}_{t+n}r_{t-1}$ = the forecast made in $t - 1$ of the rate expected in $t + n$; R_t = the spot rate in period t; and ${}_tr_{t-1}$ = the forecast made in $t - 1$ of the spot rate in t. All rates are one-year rates. The left side of the equation denotes the revision in period t of the forecast made originally in $t - 1$, the forecast referring in both cases to the spot rate in period $t + n$. The right-hand side denotes a constant term a and the proportion γ of the current error $R_t - {}_tr_{t-1}$ that is projected into the revision of subsequent forecasts, and a random term, u. Meiselman tested this model with the Durand data for eight sets of revisions representing eight spans of forecast (where the span n is the number of years between the time the forecast is made and the time to which it refers). He found that as the span increases, γ falls together with the ability of the model to explain the variation of the revisions.

Because the regression coefficient was significantly different from zero in all eight regressions, as well as because of other characteristics of the results, Meiselman concluded that the model adequately represented the data and therefore that the data behaved as though they were forecasts.

Now, if the data on forward rates can be said to contain forecasts, the questions we raise are: How are the forecasts generated, and what are their behavioral properties? In the following, Meiselman's and other data on the term structure are probed to extend our insights into expectational behavior in the capital market.

The extrapolative model. In Chapter 3 of this volume, Jacob Mincer establishes a relationship between the Meiselman-type error-learning models and linear extrapolative forecasting models. In this extrapolation, the forecast value for any period in the future, $t + n$, is computed from a weighted average of past values of the same series taken sequentially back from the target date $t + n$ through the current period t and back into the past. Since the values of the series between periods t and $t + n$ are not known, the values forecast for each of these periods are substituted for the actual values.[11] In other words, a forecast of the period $t + n$ implies forecasts of the preceding periods, $t + n - 1, t + n - 2, \ldots, t + 1$. In symbols:[12]

$$(1) \qquad {}_{t+n}F_t = C + B_1({}_{t+n-1}F_t) + B_2({}_{t+n-2}F_t) + \cdots + B_nA_t + B_{n+1}A_{t-1} + \cdots + {}_{t+n}E_t,$$

where ${}_{t+n-i}F_t$ = forecast made in period t referring to i periods prior to $t + n$; A_{t-i} = actual value of series i periods into the past; B_i = weights in linear combination; C = constant term, by hypothesis equal to zero; and E = autonomous component of forecast, i.e., the part not based on past values of the series. If the forecast were made in period $t - 1$ instead of in t, the fourth term on the right side of (1) would be $B_n({}_tF_{t-1})$ instead of B_nA_t, since A_t could not be known in period $t - 1$. Let us write out the equations for the forecasts referring to $t + 1$ made in both t and $t - 1$:

$$(2) \qquad {}_{t+1}F_t = C + B_1A_t + B_2A_{t-1} + \cdots + B_{n+1}A_{t-n} + {}_{t+1}E_t,$$

$$(3) \qquad {}_{t+1}F_{t-1} = C + B_1\,{}_tF_{t-1} + B_2\,A_{t-1} + \cdots + B_{n+1}\,A_{t-n} + {}_{t+1}E_{t-1}.$$

[11] This procedure is optimal, in the sense of minimizing mean square error of forecast, for a particular class of time series. See p. 87, above.

[12] The symbols F and A represent forecasts and target (or actual) values of any series. When the F refers to forward rates, the nonextrapolative component E includes not only the autonomous component but any nonforecasting component, such as errors of measurement and liquidity premia.

Subtracting (3) from (2):

$$(4) \qquad {}_{t+1}F_t - {}_{t+1}F_{t-1} = B_1(A_t - {}_tF_{t-1}) + ({}_{t+1}E_t - {}_{t+1}E_{t-1}).$$

Let us apply the above procedure to the forecasts made in periods t and $t-1$ referring to period $t+2$:

$$(5) \qquad {}_{t+2}F_t = C + B_1({}_{t+1}F_t) + B_2A_t + B_3A_{t-1} + \cdots + {}_{t+2}E_t.$$

$$(6) \quad {}_{t+2}F_{t-1} = C + B_1({}_{t+1}F_{t-1}) + B_2({}_tF_{t-1}) + B_3A_{t-1} + \cdots + {}_{t+2}E_{t-1}.$$

Subtracting (6) from (5):

$$(7) \qquad ({}_{t+2}F_t - {}_{t+2}F_{t-1}) = B_1({}_{t+1}F_t - {}_{t+1}F_{t-1}) + B_2(A_t - {}_tF_{t-1}) + ({}_{t+2}E_t - {}_{t+2}E_{t-1}),$$

and the difference equation for $t+3$:

$$(8) \qquad ({}_{t+3}F_t - {}_{t+3}F_{t-1}) = B_1({}_{t+2}F_t - {}_{t+2}F_{t-1}) + B_2({}_{t+1}F_t - {}_{t+1}F_{t-1}) + B_3(A_t - {}_tF_{t-1}) + ({}_{t+3}E_t - {}_{t+3}E_{t-1}).$$

By recursive substitution of the lower span revisions, each revision becomes a linear function of the current forecasting error alone, generating Meiselman's equations for each span. Thus, we find that linear extrapolations of type (1) and (2) are consistent with Meiselman's error-learning model.

We note that Meiselman's revision coefficients γ_i and the R_i^2 decline steadily with increasing span i. Mincer has derived an expression that relates the decline in the coefficients of determination to the decline in the coefficients of regression, when multispan forecasting is assumed to be recursive as in (1). According to Mincer the coefficient of determination of the ith regression is given by the following expression: [13]

$$(9) \qquad R_i^2 = \frac{1}{1 + \dfrac{(1 - R_1^2)\gamma_1^2}{R_1^2\gamma_i^2}(1 + \gamma_1^2 + \cdots + \gamma_{i-1}^2)},$$

where R_i^2 = coefficient of determination of ith regression; and γ_i = coefficient of regression of ith regression. This expression implies that the coefficients of determination will decline with increasing span of forecast whenever $\gamma_i \geqslant \gamma_{i+1}$; that is, when the coefficients of regression do not increase. Table 4-1 compares the coefficients of determination

[13] See Jacob Mincer, equation (36), p. 103 of this volume, for the derivation.

TABLE 4-1. Comparison Between the Estimated and Predicted Coefficients of Determinations for Regressions of the Error-Learning Model (Durand Data, 1901–55)

Span (1)	Estimated R^2 (adj) (2)	Predicted R^2 (adj) (3)	Span (1)	Estimated R^2 (adj) (2)	Predicted R^2 (adj) (3)
1	.9053	.9053	5	.4004	.4206
2	.7470	.7812	6	.3709	.3345
3	.5819	.6395	7	.4055	.3324
4	.4537	.5154	8	.3289	.2743

Note: Column 2 lists coefficients of determination computed by Meiselman; column 3 lists the ones predicted by equation (9).

that Meiselman found (column 2) with the ones predicted by equation (9). While the expressions for the standard errors of the two sets of statistics are not easily derived to permit an evaluation of the statistical significance of their differences, the figures in the two columns appear to be quite close.[14]

The inference sometimes drawn from Meiselman's declining correlations that the relevance of the expectations hypothesis diminishes with increasing term to maturity, therefore, does not follow.[15]

The consistency of Meiselman's whole set of error-learning equations with a hypothesis of linear extrapolative forecasting strengthens the interpretation of forward rates as expectational magnitudes. Now, we can go further and ask: What is the particular form of the

[14] Formula (9) based on Mincer's equation (33) assumes that the variances of the nonextrapolative components do not change with span. If they decrease with span, predicted R^2 will decline more rapidly than observed R_i^2. Alternatively, nonforecasting components might create the offsetting effect, according to Mincer's equation (33a), if the revision coefficients are declining. In Table 4-1, predicted R^2 do indeed decline somewhat more rapidly than the observed ones.

[15] For example, the following is quoted from Wood [18, p. 165].

> It is reasonable to suppose that investors will have fairly firm expectations regarding the level of rates one year from the present and will formulate their decisions on the basis of expectations; whereas expectations of what rates will be several years into the future are likely to be at best hazy and, as a consequence, investors are likely to determine their holdings of one-year relative to, say, eight-year securities to a large extent on the basis of considerations other than expectations of future short rates.

While the point may be correct, it cannot properly be concluded from Meiselman's findings.

extrapolative model that generates Meiselman's revision equations? This question is important because the form of the extrapolation contained in forward rates provides further insights into the expectational behavior in financial markets.

The relationship between the revision variables on the right side of equations (7) and (8) and the current error reveals the relationship between the extrapolative weights, B_i, of equation (1) and Meiselman's regression coefficients, γ_i.

For example, substituting (4) into (7) yields

$$(10) \qquad ({}_{t+2}F_t - {}_{t+2}F_{t-1}) = B_1[B_1(A_t - {}_tF_{t-1})] + B_2(A_t - {}_tF_{t-1}) + ({}_{t+2}E_t - {}_{t+2}E_{t-1})$$
$$= (B_1^2 + B_2)(A_t - {}_tF_{t-1}) + ({}_{t+2}E_t - {}_{t+2}E_{t-1});$$

and substituting (4) and (7) into (8):

$$(11) \qquad ({}_{t+3}F_t - {}_{t+3}F_{t-1}) = B_1B_1[B_1(A_t - {}_tF_{t-1})] + B_2(A_t - {}_tF_{t-1}) + B_3(A_t - {}_tF_{t-1}) + ({}_{t+3}E_t - {}_{t+3}E_{t-1})$$
$$= (B_1^3 + B_1B_2 + B_3)(A_t - {}_tF_{t-1}) + ({}_{t+3}E_t - {}_{t+3}E_{t-1}).$$

Similar expressions for forecasts that span more than three years are analogous to (1) and (11). These expressions produce a system of equations with only one exogenous variable, the current error of forecast. The dependent variable of each of the equations in the system is related to the one exogenous variable directly, as well as through its relationship with the dependent variables of the equations above its own. The expressions for the regression coefficient relating a given revision to the current error includes all the extrapolative weights B_i that are included in the expression for the coefficient of the equation above it plus one additional weight. One can therefore deduce the B_i recursively, in the following manner, denoting Meiselman's coefficients, $\gamma_i i = 1, 8$:

$$(12) \qquad \gamma_1 = B_1$$
$$\gamma_2 = B_1^2 + B_2$$
$$\gamma_3 = B_1^3 + 2B_1B_2 + B_3$$

and generally [16]

$$\gamma_i = \sum_{j=1}^{i} \gamma_{i-j} \cdot B_j, \quad \text{where } \gamma_0 = 1.$$

Using Meiselman's reported values of the γ_i we derived the set of B_i according to (12). For example, $B_2 = \gamma_2 - B_1^2$; $B_3 = \gamma_3 - B_1^3 - 2B_1B_2$; and so on for the other values of γ and B not shown here. Table 4-2, column 2, shows the γ_i coefficients, which were recomputed and changed slightly from those Meiselman reported; the B_i coefficients are in column 3. From the derivation of the B_i it is clear that if one of the weights, say B_2, is out of line, the succeeding weight, B_3, will be out of line in the other direction. For example, if B_2 were too high, we would subtract from γ_3 a larger number than we should in order to get B_3; therefore, a high B_2 would lead to a low B_3. If sampling fluctuation knocked one weight out of line, it would set in motion a wiggle that would reverberate down the column of weights. The minus sign attached to B_8 makes little sense in the present context, but it is preceded by a B_7 that is too large and by a B_6 that is too small.

Because of this sampling problem, the weights as raw data do not evince a coherent pattern; yet, when plotted, a pattern may be detected. To glean at least one estimate of the pattern, we plotted on Chart 4-1 the weights, B_i (computed from Meiselman's coefficients shown in Table 4-2, column 3), on semilog paper and drew a smooth curve through the points. The weights read from the smooth curve are in column 5 of Table 4-2. The scatter around the curve in Figure 4-1 is, of course, substantial, and the curve highly subjective. Some may contest the treatment of the extreme point, and, of course, the negative point is meaningless. A partial, though by no means conclusive, test for the validity of this procedure is illustrated in the following experiment.

Working with equations (4), (7), and (8), as well as the equations for the revisions of forecasts of spans 4 through 8, we formed a single independent variable for each equation by multiplying each variable on the right side by the relevant weight and summing the products. For example, in the case of equation (8) we multiplied the first independent variable by .7029, the second by .0318, and the third by .0114,

[16] See Mincer, p. 89.

CHART 4-1. Patterns of Smoothed Weights, B_i, for Extrapolation Equation (1)

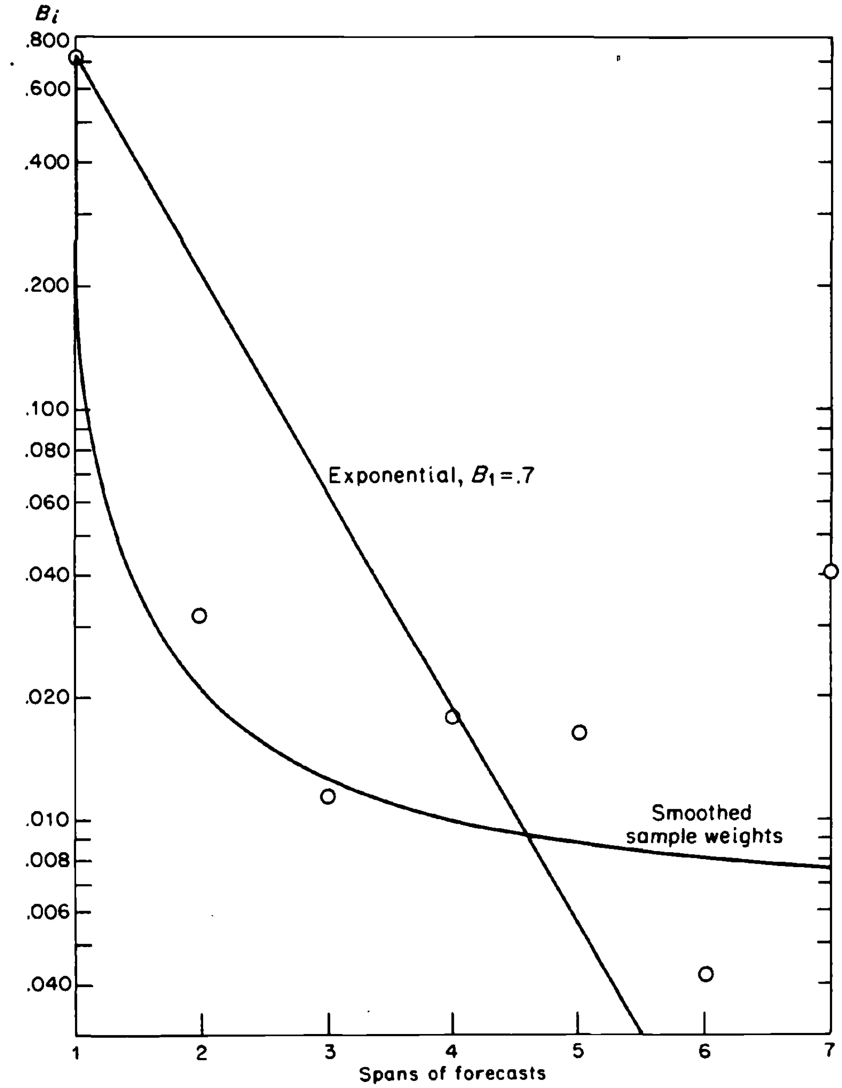

TABLE 4-2. Regression Statistics for Experiments With Estimated Extrapolative Weights

Span of Forecast (1)	γ_i (2)	B_i (Sample) (3)	$R^2_{B_i}$ (4)	$\bar{B}_i$ (Smooth) (5)	$R^2_{\bar{B}_i}$ (6)	B_i (Exponential Decline) (7)	$R^2_{B_i}(K)$ (8)
1	.7029	.7029	.90526	.7029	.90526	.7000	.90526
2	.5256	.0318	.93545	.0220	.93756	.2100	.90167
3	.4034	.0114	.93421	.0133	.93585	.0630	.89862
4	.3263	.0180	.86998	.0105	.87184	.0189	.86026
5	.2769	.0165	.86353	.0089	.86412	.0057	.87384
6	.2348	.0042	.92688	.0080	.92845	.0017	.92339
7	.2367	.0401	.92430	.0076	.92990	.0005	.91326
8	.2089	−.0016	.87279	.0072	.87801	.0001	.87044

Note: Each set of B_i is used with the same variables in the following set of regressions:

$$({}_{t+1}r_t - {}_{t+1}r_{t-1}) = A + B_1(R_t - {}_tr_{t-1})$$

$$({}_{t+2}r_t - {}_{t+2}r_{t-1}) = A + B_1({}_{t+1}r_t - {}_{t+1}r_{t-1}) + B_2(R_t - {}_tr_{t-1}).$$

In each case the products and variables are summed into one independent variable. There are eight simple regressions for each set of B_i. The R^2 are the coefficients of determination adjusted for degrees of freedom. They are computed for each of the regressions run, with the weights listed in the adjacent column to the left starting from row 1 and continuing down to the row in which the R^2 in question appears. For example, in column 4, .93421 is the R^2 computed from the simple regression the independent variable of which was computed by summing the products of .7029 $({}_{t+2}r_t - {}_{t+2}r_{t-1})$, .0318 $({}_{t+1}r_t - {}_{t+1}r_{t-1})$, and .0114 $(R_t - {}_tr_{t-1})$. The Durand data were used in this experiment.

the first three weights listed in Table 4-2, column 3. The independent variable of this regression is equal to the linear combination of the variables specified in (8), the constants in the linear combination being the weights. There are eight such regressions, each having one independent variable (formed by the linear combination of the variables listed in the equation) for the corresponding revision of forecast. In one set of eight regressions, we used the weights listed in column 3 and in the other set, the weights listed in column 5. In column 4 we list the adjusted R^2 coefficients of each of the eight sets of regressions, using the empirically estimated weights; column 6 shows the adjusted R^2 coefficients for the smoothed set of weights.

In all but one case, the R^2 coefficients are higher for the smoothed weights; although in no case could the differences appear statistically significant under standard test procedures. Such lack of statistical significance, however, is not too surprising when one considers that the first weight .7029 is almost ten times as large as the sum of all the other weights and, therefore, dominates the linear combination. It is nonetheless interesting to find that relatively small changes in small weights lead to consistently better results, even though only slightly so.

We may tentatively accept the hypothesis that the weights in the extrapolative equation (1) decline in accordance with a smooth curve relating the value of the weight to the period of the lag. The hypothesis allows us to infer the time perspective on the interest rate market. In Figure 4-1, while there is a sharp initial drop between B_1 and B_2, the pattern of decline is much more gradual thereafter; the curve becomes almost horizontal. This pattern implies that interest rates prevailing more than ten years before the time the forecast was made are considered in the forecasts of the future, although the weight attached to these rates is small. Column 7 of Table 4-2 lists the set of geometrically declining weights which produced the best fit in regressions described for the two other sets of weights. Column 8 lists the R^2 coefficients of each of these regressions, which serve as a standard for evaluating the smoothed weights. Here again, in all but one case, the smoothed weights produced a better fit than the geometrically declining weights, although, again, the differences would not appear significant under standard—but inapplicable—test procedures. The relevance of this comparison of weights resides in the fact that they both start from approximately the same place, while the geometrically declining weights indicate a shorter horizon. Therefore, the apparently better fit with smoothed weights provides some evidence that the more distant past is indeed considered in forecasting future rates. In Figure 4-1 the geometrically declining weights are represented by a straight line, since the curves are drawn on semilog paper.

In terms of equation (1), the weights B_i do not depend on the forecast span. As the span changes, the independent variables change, but the weights stay the same. The above analysis does not test whether the weights vary with forecast span. We can evaluate the consequences of this assumption. If the forecast span is not relevant there is no reason for having eight separate regressions. Since the estimate of B_1 is the same, it makes no difference in principle whether this weight is attached to, say, $(A_t - {}_tF_{t-1})$ or $({}_{t+1}F_t - {}_{t+1}F_{t-1})$ so long as the independent variable selected conforms to the appropriate dependent variable.[17] Therefore, we can estimate the B_i with one multiple regression whose first observation is the first observation of equation (4), whose second observation is the first observation of equation (7), whose third observation is the first observation of (8), and, finally, whose ninth observation is the second observation of (4). We can include a dummy variable indicating the span of forecast represented by the particular observation to test the hypothesis that the estimates of the B_i depend upon the independent variable to which they are attached. In other words, we can test the hypothesis that the weights in extrapolative equation (1) vary with the span of forecast by including a dummy variable representing each of the eight equations from which the observations are chosen.

In Table 4-3 we compare the B_i weights, estimated by multiple regression, with those computed from the coefficients γ_i of the error-learning model. The test for the significance of the difference between the two estimates shown there is only illustrative, since it was not possible to estimate the standard errors of both sets of weights. We used our estimates of the standard errors of the B_i from the multiple regression as admittedly poor substitutes for the standard errors of the difference. Unless the correlation between the two sets of estimates is very high and positive, our procedure understates the standard errors of the differences and therefore our t-values are too high. In spite of this, the t-values in column 5 of Table 4-3 show the differences to be statistically significant in only two out of eight cases.

Since there is an exact formula relating the B_i to the γ_i, we can infer either set from estimated values of the other. Having compared the set of B_i inferred from the estimated γ_i with the directly estimated B_i, we will now reverse the process by inferring a set of γ_i from the estimated B_i, reversing the procedure described in (12), and compare this set of γ_i with the set directly estimated from the error-learning model. The results of this comparison are shown in Table 4-4. Once

[17] In principle, it is possible to run these regressions directly on the forward and past spot rates, that is, on the eight forecasting equations of the form (1). Since adjacent interest rates are highly correlated, the estimated coefficients would be very unstable. The revisions were used to lessen the multicollinearity problem, on the assumption that the revisions are less intercorrelated than the rates themselves. This assumption is likely to be true since the revisions are less dependent on the over-all level of economic activity.

TABLE 4-3. Comparison Between the Weights Computed With a Multiple Regression and Those Computed From the Coefficients of the Error-Learning Model (Durand Data, 1900–54)

Number of Logs (1)	B_i (Mult. Reg.) (2)	S_{B_i} (3)	B_i (Computed) (4)	t-Value of Difference (5)
1	.7457	.0199	.7029	2.1507
2	.0548	.0248	.0318	0.9274
3	.0347	.0240	.0114	0.9708
4	−.0914	.0192	.0180	−5.6979
5	.0522	.0243	.0165	1.4691
6	.0051	.0258	.0042	0.0349
7	.0412	.0258	.0401	0.0426
8	−.0168	.0260	−.0116	−0.2000

Note: Column 2 lists the weights estimated in the multiple regressions, a few observations of which are as follows:

$$({}_{t+1}r_t - {}_{t+1}r_{t-1}) = A + B_1(R_t - {}_tr_{t-1}) + 7 \text{ zero values:}$$
$$({}_{t+2}r_t - {}_{t+2}r_{t-1}) = A + B_1({}_{t+1}r_t - {}_{t+1}r_{t-1}) + B_2(R_t - {}_tr_{t-1})$$
$$+ 6 \text{ zero values;}$$
$$({}_{t+8}r - {}_{t+8}r_{t-1}) = A + B_1({}_{t+7}r_t - {}_{t+7}r_{t-1}) + \cdots + B_8(R_t - {}_tr_{t-1})$$
$$+ \text{ no zero values.}$$

Finally, the ninth observation is of the same form as the first. There is only one multiple regression to compute all the weights. In the first observation the values of the variables 2 to 8 are zero. In the second observation the value of variables 3 to 8 are zero; and so on.

Column 3 lists the standard errors of the coefficients.

Column 4 is the weights implied by the γ_i of the error-learning model.

Column 5 divides the difference between the two estimates of the B_i (cols. 2 and 4) by the standard errors of those estimated by the multiple regression (col. 3). An explanation of this procedure is given in the text.

again, using the same procedure as in Table 4-3, the differences between the two estimates are statistically significant in only two out of eight cases. In this table, while we have estimates of the standard errors of the γ_i computed directly from the error-learning model, we have none for the estimates inferred from the set B_i estimated with the multiple regression. As before, it seems plausible to consider that the significance of the difference between the two estimates of γ_i is overstated in column 5.

The return to normalcy model. Perhaps the most widely recognized phenomenon in expectational economics is the so-called return to normalcy mechanism, whereby people expect a series to move in the direction of its normal level. Any extrapolative model that predicts the value of a series with a moving average of past values of the series,

TABLE 4-4. Comparison Between the Coefficients Inferred From the Directly Estimated Weights and the Coefficients Meiselman Estimated With the Error-Learning Model (Durand Data, 1900–54)

Span of Revision (1)	γ_j (Meiselman) (2)	S_{γ_i} (3)	γ_i (Computed) (4)	t-Value of Difference (5)
1	.7029	.0312	.7457	1.3718
2	.5259	.0419	.6109	2.0286
3	.4033	.0466	.5312	2.7446
4	.3262	.0486	.3641	0.7798
5	.2770	.0459	.2864	0.2048
6	.2349	.0414	.2546	0.4758
7	.2370	.0389	.2477	0.2751
8	.2082	.0401	.2209	0.3167

Note: Column 2 lists the coefficients Meiselman estimated with the error-learning model and column 3 their standard errors. Column 4 is computed from the weights estimated with a multiple regression and listed in column 2 of Table 4-3. The formula for this computation is the same as that shown by (12) in the text, except that now the B_i are known and the γ_i are inferred. Column 5 lists the ratio of the difference of the two estimates of the γ_i to a crude estimate of the standard error of this difference. The estimated standard error is too low because it fails to include the standard error of the M_i inferred from the B_i. Therefore, the t-values listed in column 5 are overestimates.

each past value weighted less than one, will produce forecasts that lie between the extremes of the series and its expected value.[18] Since the estimated weights of the extrapolative model introduced in this study are all less than one, the expectations inherent in the term structure conform with this behavior. This behavior alone, however, does not explain the widely observed manifestation of the expected return to normalcy [see 10] with respect to the term structure; that is, the tendency for yield curves to decline when current short-term rates are high and to increase when they are low. Nor is it clear what the meaning is of the normal rate as distinct from the expected and mean rate.

Algebraically, the return to normalcy hypothesis can be described in the following linear form:

$$(13) \qquad {}_{t+2}F_t - {}_{t+1}F_t = K(A_t - {}_N A_t), \quad K < 0,$$

where, $({}_{t+2}F_t - {}_{t+1}F_t)$ is the change expected at t of the target value, in this case the one period spot rate from $t + 1$ to $t + 2$; A_t is the target,

[18] This idea underlies Hicks' coefficient of expectation when that number is less than one [6, p. 205]. The relevance of the normal rate to the Keynesian liquidity preference function is considered in Section V below.

TABLE 4-5. Statistics Computed From the Regression of the Expected Change of Future Spot Rates on the Level of the Current One-Period Spot Rate (Durand Data, Annual Observations, 1900–54)

Span of Forecast (1)	K (2)	t-Value of K (3)	Constant Term (4)	t-Value of Const. Term (5)	R^2 (adj.) (6)
${}_{t+1}r_t - R_t$	−.1627	−7.2109	.6437	7.8584	.4904
${}_{t+2}r_t - {}_{t+1}r_t$	−.1264	−11.8510	.4909	12.6817	.7246
${}_{t+3}r_t - {}_{t+2}r_t$	−.0997	−17.0387	.3878	18.2505	.8452
${}_{t+4}r_t - {}_{t+3}r_t$	−.0741	−12.7946	.2948	14.0311	.7543
${}_{t+5}r_t - {}_{t+4}r_t$	−.0737	−8.2939	.3071	9.5246	.5612
${}_{t+6}r_t - {}_{t+5}r_t$	−.0475	−7.8774	.1964	8.9637	.5353
${}_{t+7}r_t - {}_{t+6}r_t$	−.0332	−6.1131	.1382	7.0177	.4070
${}_{t+8}r_t - {}_{t+7}r_t$	−.0361	−8.0801	.1511	9.3173	.5481
${}_{t+9}r_t - {}_{t+8}r_t$	−.0250	−4.1018	.0981	4.4308	.2299

Note: The regressions were of the following form ${}_{t+n}r_t - {}_{t+n-1}r_t = Q + k\,{}_nR_t + V_n$.

or spot rate, at t; ${}_NA_t$ is the putative normal value of the series, or normal rate, as of period t; and K is the proportion of the deviation expected to be offset; it is negative to reflect the inverse relation between the expected change and the deviation of the current spot rate from the normal rate.[19]

It is easy to confirm the inverse relation between the slopes of the yield curves and the levels of the one-period spot rates statistically, by running regressions analogous to equation (13).[20] The results of these regressions are listed in Table 4-5. In every case the relation

[19] The expected one-period change in the short-term rate is not quite the same as the average slope between adjacent points of a yield curve, which is a locus of long-term rates rather than forward rates. However, the long-term rates are simply averages of the forward rates. The difference between, say, R_3 and R_2 (using simple interest) is

$$R_3 = \frac{2}{6}(r_1 + r_2 + r_3) - \frac{3}{6}(r_1 + r_2) = \frac{2}{6} r_3 - \frac{(r_1 + r_2)}{6}.$$

$r_3 - r_2$ is therefore only approximately equal to the slope of the yield curve. The notation F and A (forecast and actual) in the following analysis is interchangeable with r and R, forward rate and spot rate, respectively, since the present analysis assumes the forward rates are forecasts.

[20] Since by any definition the normal rate will vary slowly, the use of R_t in place of $(R_t - {}_NR_t)$ will not seriously distort the estimated relationship. The empirical relevance of R_N will be shown presently.

is significantly negative and explains a substantial part of the variation of the expected changes in rates. Table 4-5, of course, is merely a statistical confirmation of the widely recognized relationship described above. Since regressing $({}_{t+n}F_t - {}_{t+n-1}F_t)$ on A_t in place of $(A_t - {}_nA_t)$ implies a constant normal rate, its magnitude can be inferred by dividing column 2 into column 4 of Table 4-5. The constant term (column 4) is equal to $K\,{}_nA_t$, which, when divided by K, will yield an estimate of the normal rate—approximately 4 per cent. At the expense of elegant phrasing, one may interpret this number as the average normal rate.

Since the normal rate, as yet undefined, is not observable, confirmations of the return to normalcy model's application to the term structure typically substitute A_t for $(A_t - {}_NA_t)$ in (13), or arbitrarily assign some value or limited number of values to ${}_NA_t$.[21] Our earlier analysis suggests a different method. Let us define ${}_NA_t$ in (13) as follows:

$$(14) \qquad {}_NA_t = B_2A_{t-1} + B_3A_{t-2} + \cdots + B_NA_{t-N-1}.$$

In other words

$$(15) \qquad {}_{t+2}F_t - {}_{t+1}F_t = KA_t - K\left(\sum_{i=1}^{N-1} B_{i+1}A_{t-i}\right).$$

On the hypothesis that (14) is true, we can estimate K as the partial regression coefficient in (15). But we already have an estimate of the partial regression coefficient, K. Recall that equations (2) and (5) above defined ${}_{t+1}F_t$ and ${}_{t+2}F_t$, respectively. Substituting (2) into (5)[22] we have:

$$(16) \qquad {}_{t+2}F_t = (B_1^2 + B_2)A_t + (B_1B_2 + B_3)A_{t-1} + \sum_{i=3}^{N} (B_1B_i + B_{i+1})A_{t-i-1},$$

[21] See [14], for example. Van Horne adds a variable he calls "deviation of actual from accustomed level" to Meiselman's formulation of the error-learning model. He divides his sample period into two subperiods. "For each . . . period . . . an arithmetic average of the beginning forward-rate levels is calculated. This average may be thought to represent the accustomed level for the period. The deviation is simply the difference of the actual forward-rate level from the accustomed level and is employed as the second independent variable [in the error-learning equation]." [14, p. 349.]

[22] This procedure is directly analogous to the one used earlier to deduce the extrapolative weights from Meiselman's error-learning model coefficients.

and subtracting (2) from (16) we get:

$$(17) \quad {}_{t+2}F_t - {}_{t+1}F_t = [(B_1^2 + B_2) - B_1]A_t + \sum_{i=2}^{N} [B_1B_i + B_{i+1}) - B_i]A_{t-i-1}.$$

The first coefficient on the right of (17) is, therefore, our estimate of K. It will clearly be negative when $B_1 > (B_1^2 + B_2)$. But, according to (12), $B_1 = \gamma_1$ and $(B_1^2 + B_2) = \gamma_2$, where the γ's are Meiselman's error-learning coefficients for the first and second spans, respectively. Therefore, K will be negative when $\gamma_i > \gamma_{i+1}$. In other words, the decline in Meiselman's coefficients as the span of forecast increases is algebraically identical to an inverse relationship between the expected change between any two spans of forecast and the deviation of the current spot rate from the normal rate.

To keep the algebra simple we presented the argument in terms of the first two spans of forecasts. The results for greater spans follow in an identical manner. The decline in the γ_i and, therefore, the fact that $B_1 > (B_1^2 + B_2) > (B_1^3 + 2B_1B_2 + B_3)$ and so forth, implies, as the analysis in the previous section showed, that the extrapolative weights, equation (1), observed for the term structure data, decline drastically from the first weight to the second one and then taper off. In Mincer's terminology, the β coefficients decline in a convex fashion.

It is the convex form of the extrapolation function that ties the decline in Meiselman's coefficients γ_i to the return to normalcy mechanism. For this pattern implies that the weight attached to the current value A_t (a weight, it turns out, that is equal in value to the error-learning coefficient for the same span of forecast) declines for successively greater spans of forecast.[23] Therefore, the subtraction of the shorter- from the longer-span forecast (that is, the expected change) is equivalent to subtracting a larger from a smaller weight attached to A_t—hence the inverse relation.

But if the weight attached to A_t declines with increasing span of forecast, the weights attached to the lagged actual values A_{t-i} must rise, if the sum of the weights for each span of forecast is to equal 1. In other words, as the span increases, the market gives increasing weight to the more distant past and distinguishes less between the more

[23] See Mincer's equation (17), p. 92 of this volume.

TABLE 4-6. Weights for Each Span of Forecast for Equation (12) Implied by the Single Set of Weights Estimated for Equation (1) (Durand Data, Annual Observations, 1916–54)

Span of Forecast (1)	A_t (2)	A_{t-1} (3)	A_{t-2} (4)	A_{t-3} (5)	A_{t-4} (6)	A_{t-5} (7)	A_{t-6} (8)	A_{t-7} (9)
1	.7029	.0220	.0133	.0105	.0089	.0080	.0076	.0072
2	.5161	.0288	.0198	.0163	.0143	.0132	.0125	
3	.3916	.0312	.0231	.0197	.0178	.0166		
4	.3065	.0317	.0247	.0219	.0201			
5	.2472	.0316	.0259	.0233				
6	.2054	.0314	.0265					
7	.1758	.0312						
8	.1547							

Note: Looking down any column we see the weight that a particular variable gets for a particular span of forecast. The weights are computed by taking the set shown in Table 4-1, column 5, which is an estimate of the first eight weights for equation (1) regardless of span of forecast. This single set implies a varying set of weights (one set for each span of forecast) attached only to the independent variables that constitute the current and lagged spot rates. For the second span, for example, equation (1) says ${}_{t+2}F_t = B_1({}_{t+1}F_1) + B_2A_t, \ldots B_1({}_{t+1}F_t)$ is implicitly equal to $B_1^2A_t + B_1B_2A_{t-1} \ldots$. To get the figures listed above, we add up all the coefficients attached to the particular variable: first A_t, then A_{t-1}, and so forth.

immediate and the more distant past. Long-span forecasts, therefore, take account of the full history of the series at the expense of the current value and, in this sense, approach what can be called a normal rate. Table 4-6 lists the estimated weights attached to the current and past rates for successive spans of forecast. (The table is triangular only because there was insufficient data to make it otherwise.)

TREASURY DATA

To ascertain whether the results described in the preceding sections were peculiar to the Durand data, we applied the same analysis to data read off the yield curves given in the *Treasury Bulletin*.[24]

Annual rates of return on government securities were read off the yield curves at quarterly intervals from March 1945 to December 1964. It was possible to get a continuous series only by restricting ourselves to not more than five-year maturities; each quarterly observa-

[24] Since Neil Wallace [16] had already read off these yield curves for his study, I simply brought forward the data he supplied.

tion includes the annual rates on one-year, two-year—up to five-year maturities; five rates in all. With five rates we are able to run only three spans of revisions in the context of the error-learning model and infer only three weights of the extrapolation equation (1). It was not possible to smooth so few weights to test for a systematic pattern.

Since the maturities along each yield curve are read at annual intervals, the implied forecasts are also annual. As such, they should not be related to the spot rates observed at quarterly intervals but rather to a four-term moving average of these spot rates. The coefficients obtained in this manner imply a system of weights that can be appropriately applied only to lagged spot rates observed annually instead of to the quarterly observations used in this report. Since the available data span is a period of only nineteen years, the use of annual observations is not feasible. In our work with the error-learning model we have therefore taken the annual forecasts as approximations of the quarterly forecasts. (An experiment justifying this procedure is described in a note to Table 4-10.) In the accuracy analysis described in Section VI we are able to relate the forecasts to a moving average of the spot rates.

Tables 4-7 through 4-10 list the results of applying the earlier analysis to the Treasury data.

Table 4-7, similar to Table 4-1, compares the estimated decline in the coefficients of determination in regressions of the error-learning model with the predicted decline in accordance with equation (9).[25]

TABLE 4-7. Comparison Between the Estimated and Predicted Coefficients of Determination for the Regressions of the Error-Learning Model (Treasury Data, 1946–64)

Span	Estimated R^2 (adj)	Predicted R^2 (adj)
1	.6313	.6313
2	.3838	.3438
3	.2013	.2542

Note: See note to Table 4-1.

[25] The decline in the empirically estimated R^2 is slower than in the predicted decline between the first and second span; this is consistent with the presence of a nonforecasting component in the forward rate. There are obviously too few observations (spans of forecast) to draw any conclusions; in fact, the last observation goes the other way.

Table 4-8, similar to Table 4-2, lists the relevant statistics for the error-learning model regressions. Again, the results are similar to those obtained with the Durand data. When the current error, the independent variable in the error-learning model, is replaced by a linear combination of the prior revisions, the coefficients of determination (column 5, Table 4-8, analogous to column 4 of Table 4-2) do not systematically decline. Table 4-9, like Table 4-3, compares the directly estimated extrapolated weights with those deduced from the error-learning model coefficients. Again, the differences do not appear to be significant. Finally, Table 4-10, like Table 4-4, records the comparison between the directly estimated error-learning model

TABLE 4-8. Error-Learning Model Applied to Interest Rates on Treasury Securities, Quarterly, 1946–64

Span (1)	b (2)	T_b (3)	R^2 adj. (4)	R^2 adj (modified model) (5)
1	.6805	11.3004	.63129	.63129
2	.4552	6.8617	.38376	.80800
3	.3923	4.4332	.20133	.61211

The general formula of the regression is ${}_{t+n}r_t - {}_{t+n}r_{t-1} = a + b(R_t - {}_tr_{t-1}) + \epsilon$. Column 3 is equal to the regression coefficients b divided by their standard errors.

Column 5 lists the coefficients of determination obtained by altering the independent variable to consist of a linear combination of past revisions plus current error. This column corresponds to column 4 of Table 4-2 for the Durand data.

TABLE 4-9. Comparison Between the Weights, B_i, Estimated With a Multiple Regression and Those Computed From the Coefficients of the Error-Learning Model (Treasury Data, Quarterly, 1946–64)

Span (1)	B_i (mult. reg.) (2)	S_{B_i} (3)	B_i (comp. from γ_i) (4)	t-Value of Difference (5)
1	.6835	.0389	.6805	0.0771
2	.0579	.0451	−.0079	1.4590
3	.1006	.0480	.0826	0.0350

Note: See notes to Table 4-3.

TABLE 4-10. Comparison Between the Coefficients, M_i, Estimated With the Error-Learning Model and Those Computed From the Weights Estimated With a Multiple Regression (Treasury Data, Quarterly, 1946–64)

Span (1)	γ_i (err. learn. mod.) (2)	S_{γ_i} (3)	γ_i (comp. from B_i) (4)	t-Value of Difference (5)
1	.6805	.0602	.6835	0.0498
2	.4522	.0663	.5251	1.0543
3	.3923	.0885	.4991	1.2068

Note: See notes to Table 4-3.

Wallace [15, p. 25] has estimated the coefficients for the error-learning model on the basis of annual forecasts. We have calculated from his estimates the extrapolation weights comparable to those listed in column 4 of Table 4-9. In addition, we have estimated these weights directly with a multiple regression and from these estimates calculated the implied set of coefficients for the error-learning model. Each of these estimates are listed in the following table. At the bottom of the table are references to comparable figures for the quarterly model.

Span (1)	γ_i (err. learn. mod.) (2)	B_i (comp. from γ_i) (3)	B_i (mult. reg.) (4)	γ_i (comp. from B_i) (5)
1	.816	.816	.8116	.8116
2	.621	−.045	−.0480	.6107
3	.499	.029	.0822	.5389

Columns 2 and 5 are analogous to columns 2 and 4, respectively, in the above Table 4-10. Columns 3 and 4 are analogous to columns 4 and 2, respectively, of Table 4-9.

coefficients and those computed from the directly estimated extrapolative weights. As in the case of the Durand data, the differences are not statistically significant. The footnote to Table 4-10 records the tests shown in Tables 4-9 and 4-10, but this time treating the forward rates as annual forecasts. For this experiment, the current error is computed as the difference between the forward rate and a four-term average of the quarterly spot rates to ensure comparability. This more rigorous method strengthens the result.

Table 4-11, comparable to Table 4-5, lists the relevant statistics for the regressions of the expected change in rates on the current spot rate. The return to normalcy mechanism, while present, is not quite as important as in the case of the Durand data. The following are the weights attached to the spot rates for consecutive spans of forecast. (The number of weights shown is limited by the availability of relevant data.)

$$_{t+1}F_t = .6835A_t + .0579A_{t-1} + .1005A_{t-2}$$

$$_{t+2}F_t = .5251A_t + .1403A_{t-1}$$

$$_{t+3}F_t = .4991A_t$$

While limited, the data do indicate a similar pattern to the one described for the Durand data.

CONCLUSION

This section considered three separately conceived models, each of which can be described as a method of forecasting, and each of which is consistent with an expectational interpretation of forward rates. The principal conclusion of the analysis is that the three models are one model looked at from three points of view. The general form is the extrapolative model described by equation (1). This extrapolative model implies a relationship between revisions of forecasts and current errors of forecast, referred to as the error-learning model. It also implies a correlation between expected change in target values and the deviation of the current target value from the normal target value. The return to normalcy phenomenon is indicated by the negative sign of the correlation.

It is essential to distinguish between the specification of the model and the parameters that are estimated for it. While the error-learning model is a particular form of the extrapolative model, its application

TABLE 4-11. Statistics Computed From the Regression of the Expected Change of Future Spot Rates on the Level of the Current One-Period Spot Rate (Treasury Data, Quarterly Observations, 1946–64)

Span of Forecast (1)	K (2)	t-Value of K (3)	Constant Term (4)	t-Value of Const. Term (5)	R^2 (adj.) (6)
$_{t+1}r_t - R_t$	−.0286	−1.1820	.4471	7.1058	.0053
$_{t+2}r_t - {}_{t+1}r_t$	−.0653	−4.5013	.3207	8.5183	.2043
$_{t+3}r_t - {}_{t+2}r_t$	−.0622	−5.8360	.2441	8.8380	.3059
$_{t+4}r_t - {}_{t+3}r_t$	−.0422	−2.0732	.1829	3.4603	.0421

Note: See note to Table 4-7.

to a given set of data need not result in declining revision coefficients (as the span of forecast increase) and, therefore, in a particular pattern of implied extrapolative weights. It is a purely empirical result. Similarly, while the return to normalcy hypothesis is consistent with the extrapolative model, because it can be expressed as a linear transformation of it, there is no necessity in practice that K be negative. The model, equation (13), is a transformation of (1), but the hypothesis that K is negative is subject to an empirical test. This test, however, is redundant once the regression coefficients in the error-learning model are observed to decline; for then the extrapolative function is convex, which implies that the expected change will be inversely related to the current value. Certainly, the results in this study strengthen our confidence that the forward rates contain forecasts since the data were shown to be consistent with several reasonable descriptions of forecasting behavior: error-learning, extrapolation, and return to normalcy. The fact that the coefficients of determination in the regressions of the error-learning model decline with span is consistent with the presence of an autonomous component in the forecasts. This constitutes further evidence in favor of a forecasting interpretation of the term structure.

IV. DECOMPOSITION OF FORECASTS

Section III identified an extrapolative component in forward rates and argued that the term includes something more than mechanical projection. The analysis also pointed to the presence of a nonextrapolative component in forward rates. The present section is concerned with empirically isolating the two components of forward rates.

The following section attempts to isolate the extrapolative component by fitting multiple regressions of the forecasts on the current and past spot rates. The computed values of these regressions measure the part of the forecast that is directly or otherwise related to the current and past spot rates; it does *not* measure the market's actual use of a moving average. The residuals of these regressions measure the autonomous component of the forecasts, the component that is not even indirectly related to the historical pattern of the interest rate series. This estimate of the autonomous component holds, of course, only in the event the forward rates consist entirely of forecasts. Any nonforecasting component—i.e., liquidity premia, errors of measurement, the effects of market disequilibria, and so forth—would also fall out in the residuals. After separating the components we measure their contributions to the accuracy of the forecasts.

The next section utilizes two standard business indicators as possible methods of "autonomous" forecasting, providing a different decomposition of forward rates into the extrapolative and autonomous components. This method drops the assumption of statistical independence between extrapolation and autonomous forecasting. It, therefore, reveals the extent to which the previously estimated extrapolative component can subsume what may be autonomous forecasting. This illustration explains the relatively large extrapolative component found earlier. Indeed, if instead of isolating the autonomous component by first exhausting the extrapolative component of the forecasts, this study had first exhausted the autonomous potential and left the extrapolative component to the residual, there is some chance the proportion between induced and autonomous components would be reversed. The difficulty with this alternate method of decomposing the forecasts lies in this study's inability to specify the relevant autonomous variables. As it stands, the method of decomposition actually used is equivalent to an a priori specification of autonomous variables, each of them bereft of their autoregressive components.

EMPIRICAL DECOMPOSITION

To estimate the induced component of the forecasts we regressed the forward rates, one span at a time, on the current and past spot rates:

$$ {}_{t+n}F_t = b_1A_t + b_2A_{t-1} + \cdots + b_8A_{t-7} + {}_{t+n}E_t, \tag{18} $$

where, ${}_{t+n}F_t$ = forecasts made at t referring to $t + n$; ${}_{t+n}E_t$ = residual term; and A_{t-1} = spot rate lagged i periods. To conserve data we arbitrarily stopped the lagging process after seven lags. To more closely approximate a real forecasting situation, one would fit separate regressions not only for each span of forecast but also for each observation, including, for any given point of forecast, only those spot rates occurring prior to or coeval with the given period. This method, however, would severely limit the sample size, since many observations would be used up in the process of fitting the regressions. More-over, the method entails a separate weighting scheme for each forecast. We chose, therefore, to run the regressions once for the whole period.

Another method of estimating the extrapolative component is to use the weights estimated in our work with the error-learning model. The procedure used, however, has the advantage of estimating an autonomous component that is *uncorrelated* with the extrapolative component, a property required in our definition of autonomous forecasting. It is also a procedure that simulates behavior of forecasters who, by hypothesis, project the past values of the series into the future. It is true, of course, that this procedure maximizes the hypothesized *extrapolative* content of the forward rate. While this is a simplification of the expectational analysis, we are imposing an understatement of the importance and, possibly, a misspecification of the nonextrapolative component in the forward rate.

Each of the n regressions of (18) regresses the forward rate of a given span on the current and past spot rates. The computed value of (18) ${}_{t+n}F_t^*$ is our estimate of the induced component of the forecast and the residual term, ${}_{t+n}E_t$, of the autonomous component. We computed this regression for two sets of data: the Durand data, with n varying from 1 to 9, producing nine regressions; and the Treasury data, with four regressions. The results of these computations are shown in Table 4-12.

The figures presented in column 2 of the table indicate that the induced component accounts for a large proportion of the variation of the forecasts. The decline (as the span of forecast increases) of the coefficients of determination does not by itself imply that the variance of the autonomous component is increasing with span.[26] In this study's samples, however, the variance of the autonomous component increases with increases in the span of forecast.

To ascertain the relative effectiveness of either component in forecasting the future spot rates we have run a set of regressions of the following form:

$$A_{t+n} = B_1({}_{t+n}F_t^*) + B_2({}_{t+n}E_t) + U_{t+n},$$

[26] The well-known formula relating the three statistics is $S_E^2 = S_F^2(1 - R^2)$, where S_F^2 is the variance of the dependent variable, the forward rates in this case.

TABLE 4-12. The Per Cent Variation of the Forward Rates That Is Explained by the Current and Lagged Spot Rates (Durand and Treasury Data)

Span of Forecast (1)	R^2 Adjusted (2)	S_E^2 (3)
Durand: Annual Observations, 1916–54		
Year		
1	.9878	.0346
2	.9735	.0601
3	.9597	.0733
4	.9447	.0798
5	.9196	.0868
6	.9091	.0760
7	.8890	.0748
8	.8727	.0647
9	.8015	.0805
Treasury: Quarterly Observations, 1949–64		
Quarter		
1	.9543	.0377
2	.9314	.0615
3	.9177	.0724
4	.8814	.0839

Note: The general form of the regression is:

$$_{t+n}F_t = B_1A_t + B_2A_{t-1} + \cdots + B_8A_{t-7} + E_n.$$

Column 1 is the value of n; column 2, the adjusted coefficients of determination; and column 3, the squared standard errors of the estimate.

where, A_{t+n} = target rate in $t + n$; $_{t+n}F_t^*$ = induced component of forecast made at t of the spot rate (target) in $t + n$; $_{t+n}E_t$ = autonomous component of the same forecast; and U_{t+n} = residual in the regression. The residual, U_{t+n}, is not the error of forecast except in the special case that $B_1 = B_2 = 1$. The results of this regression allow us to apportion the contribution of the induced and autonomous components of the forecasts to the total accuracy of the forecasts. In particular, we can test the null hypothesis that the autonomous component contributes nothing to the accuracy of the forecasts. In Table 4-13, we show the relevant results of these regressions for the Durand data and the Treasury data.

In the case of the Treasury data, the autonomous component effectively adds to the forecasting accuracy of the induced component in three of the four spans of forecast. Except for the third span the importance of the autonomous component to the accuracy of the forecast increases with span; although it would be rash to generalize this outcome.

The Durand data, however, tell a different story. In column 3 of Table 4-13 we observe a significant relation, aside from the first two spans of forecast, between the autonomous component and the future spot rates, but the relation is inverse. Not only does the autonomous component not contribute to the accuracy of the forecast, it in fact detracts from it.

The reason for this perverse result is not at all clear. It is one thing to show that a component of the forecast does a bad job in the sense that it is unrelated to the target. In such a case we could perhaps pass it off as noise in the data or some other euphemism for our ignorance. But when this component is *systematically* perverse, when it varies inversely with the target, we should at least try to explain it. If the autonomous component consisted entirely or in part of a liquidity premium, would that fact explain its behavior? Why would a liquidity premium be related to a future spot rate except in so far as this spot rate were related to the current spot rate? If it were related to future spot rates, for reasons other than the relation between future and current spot rates, some form of forecasting would be implied. But the liquidity premium component of forward rates is isolated precisely because it is a nonforecasting component. Let us say, then, a liquidity premium is related to future spot rates because both it and future spot rates are related to current spot rates. But recall that the autonomous component is isolated by regressing the forward rates on the current and past spot rates, where the residuals of this regression are the estimates of the autonomous component. By the arithmetic of least squares, the residual term in the regression is necessarily uncorrelated with each of the independent variables, including the current spot rate. The liquidity premium cannot, therefore, be linearly related to the current spot rate and at the same time be included in the autonomous

TABLE 4-13. Selected Statistics From the Regression of the Future Spot Rates on the Induced and Autonomous Components of the Forecasts (Durand and Treasury Data)

Span of Forecast (1)	Partial Correl. Coef. Squared of $_{t+n}F_t^*$ (2)	t-Value of B $_{t+n}F_t^*$ (3)	Partial Correl. Coef. Squared of $_{t+n}E_t$ (4)	t-Value of B $_{t+n}E_t$ (5)	R^2 (adj.) (6)
			Durand: Annually, 1916–54		
Year					
1	.8582	14.3426	.0305	−1.0342	.8505
2	.7374	9.7700	.0746	−1.6556	.7277
3	.6362	7.7108	.1340	−2.2931	.6353
4	.5264	6.1469	.1713	−2.6510	.5432
5	.4563	5.3417	.1810	−2.7419	.4861
6	.3730	4.4973	.2011	−2.9248	.4266
7	.2912	3.7368	.1930	−2.8514	.3582
8	.2484	3.3519	.2014	−2.9285	.3310
9	.2003	2.9178	.2280	−3.1688	.3150
			Treasury: Quarterly, 1949–64 [a]		
Quarter					
1	.7898	13.7063	.0862	2.1715	.7856
2	.3165	4.8115	.0972	2.3195	.3378
3	.4555	6.4676	.0511	1.6410	.4499
4	.4143	5.9483	.1377	2.8256	.4431

Note: The general form of the regression is that shown in the equation on p. 142.
[a] The forecast components were related to a four-term moving average of the quarterly spot rates to make the forecasts and the actuals comparable.

component.[27] For these reasons, the liquidity premium cannot be used in any simple way to explain the perverse behavior of the autonomous component. To the extent that the liquidity premium is a linear function of the current spot rates, it will be included in the induced component and, perhaps for that reason, contribute to the bias in the forecasts that we will describe later.

[27] For a discussion of the view that liquidity premia are linearly related to the current spot rates, see Kessel [8, p. 26].

A possible explanation for the unusual result lies in the sharp decline in rates in the 1930's. The greater the span of forecast, the more effect this decline will have on the results, since there will be more observations in which a forecast made in the 1920's is related to a spot rate in the 1930's. For example, in the case of a one-span forecast, there is a sharp difference between the forecast made in 1929 and the spot prevailing in 1930, while, in the case of a nine-span forecast, the forecasts made from 1921 through 1929 will be matched with spot rates that are unusually low. To the extent that the reversion to normalcy hypothesis is working, the forecasts will move in the same direction though not to the same extent as future spot rates. Since, as we have shown, this hypothesis works through the induced component, only the autonomous component displays the perverse result.

ESTIMATING AUTONOMOUS INDICATORS IN FORWARD RATES

There are many indicators of economic activity that largely share a common historical process, particularly with the interest rate series. In all likelihood, a substantial part of the relationship between the forecasts and the indicators stems from this common historical process. Hence, if we view the indicators as variables used in autonomous forecasting, we must drop the notion of independence between extrapolation and autonomous components. This method of decomposition of forecasts can partition the observed relationship between the forecasts, that is, the forward rates, and the indicators into the part due to a shared historical process and the part that is autonomous.

The coefficients of determination of the forward rates inferred from the Treasury data regressed by span of forecast on the Federal Reserve Board's Index of Industrial Production are .7159, .7806, .8107, and .8176, respectively, for the first four spans of forecast.

To accomplish the partition we have run a set of multiple regressions for each of the four spans of forecast. The forecasts are regressed on their induced component and on the Index of Industrial Production. In other words, we have estimated the following regressions:

$$ {}_{t+n}F_t = B_1({}_{t+n}F_t^*) + B_2 I + u_n, \tag{19} $$

where ${}_{t+n}F_t^*$ is the column of computed values of equation (18), and I is the Index of Industrial Production. The object is to determine the net contribution of either independent variable (given the other) to the explained variation of ${}_{t+n}F_t$. Table 4-14 gives the relevant statistics for this experiment.

Comparing columns 2 and 4 of this table we conclude that the relation between the Index of Industrial Production and the forecasts of future rates stems largely from the common historical pattern in the variation of the spot rates and the index. There is, however, a net relation between the index and the forecasts after allowing for the common historical relationship, and this net relation, as seen in column 4, grows with increases in the span of forecast.

The squared partial correlation coefficients between ${}_{t+n}F_t$ and ${}_{t+n}F_t^*$, given I, do not reveal the full importance of the induced component in the total forecasts. The figures in column 2 of Table 4-14 are smaller, and necessarily not larger, than the coefficients of determination between ${}_{t+n}F_t$ and ${}_{t+n}F_t^*$ listed in column 2 of Table 4-12. The partials associated with ${}_{t+n}F_t^*$ reveal the net explanatory power of the induced component, given that the common variation between ${}_{t+n}F_t$ and I is already accounted for. The extent to which I shares in the common variation of ${}_{t+n}F_t$ and ${}_{t+n}F_t^*$ is, in effect, deducted from the net contribution of ${}_{t+n}F_t^*$. This point illustrates an important characteristic of our method of partitioning the forecasts into the induced and autonomous components. Even though the induced component appears to explain a large part of the variation of the forecasts, it may not reveal the extent of the market's reliance on past rates, whether

TABLE 4-14. Regression of Forward Rates on Estimated Induced Component and Index of Industrial Production (Treasury Data, Quarterly, 1949–64)

Span of Forecast (quarters) (1)	Squared Partial Correl. Coef. of F^* (2)	t-Value of Reg. Coef. Attached to F^* (3)	Squared Partial Correl. Coef. of I (4)	t-Value of Reg. Coef. Attached to I (5)	Gross Coef. of Determination (adj.) (6)
1	.8866	21.8365	.0078	.6909	.9667
2	.8075	15.9938	.1049	2.6743	.9564
3	.7343	12.9808	.1947	3.8404	.9480
4	.4705	7.3615	.2394	4.3817	.9003

Note: This table is based on equation (19) in the text.

by extrapolation or other autoregressive models, in making its forecasts. The indicator varies over time in relation to the business cycle in a manner similar to that of the interest rates. Through this common relation, part of the correlation between the forward rates and the indicator will show up in the relation between the forward rates and the past spot rates.

The partials in column 4 of Table 4-14 tell a less ambiguous story. Here we see the influence of the index on the forecasts that is independent of the variation of the past spot rates. The crucial difference between the two sets of partials that underlies the apparent asymmetry of our method resides in the relative ease with which we and others can more or less exhaust the induced variation of the forecasts compared with the difficulty of exhausting the autonomous variation.[28]

To say that the Index of Industrial Production is related to the autonomous component of the forecast does not imply that investors actually consulted this particular indicator. By eliminating variables not related to the forecasts we can reduce the set of possible indicators; but the proper selection among the set of variables that cannot be excluded is not a statistical problem.

We experimented with one other indicator we considered likely to influence the forecasts of future spot rates, namely, an index of industrial stock prices.[29] As before, we regressed the forecasts on their induced component and the indicator. The coefficients of determination of the forecasts regressed on the Index of Industrial Stock Prices are, respectively, .7861, .8595, .8892, and .8889 for the first four spans of forecast. The Index of Industrial Stock Prices is related somewhat more to the forecasts than was the Index of Industrial Production. For the same purpose as before we have run regressions similar to (19), replacing the Index of Industrial Production with the Index of Industrial Stock Prices. The results are shown in Table 4-15.

[28] We say more or less because we had to arbitrarily limit the number of lags in the computation of the induced component to seven; we have a limited sample, and we have fit the entire period instead of only the spot rates prevailing in the period prior to the forecast period.

[29] We did not experiment with different indexes for this purpose (or with any other variables) but settled on the Dow Jones Index of Industrial Stock Prices recorded in the *Survey of Current Business*. Whether one indicator is more correlated with the forecasts than another is not crucial in the present context, since our main purpose in this section is to illustrate empirically the proposition that the decomposition of forecasts is a method of analysing market behavior, not a literal description of it.

TABLE 4-15. Regression of Forward Rates on Estimated Induced Component and Index of Industrial Stock Prices (Treasury Data, Quarterly, 1949–64)

Span of Forecast (quarters) (1)	Squared Partial Correl. Coef. of F^* (2)	t-Value of Reg. Coef. Attached to F^* (3)	Squared Partial Correl. Coef. of S (4)	t-Value of Reg. Coef. Attached to S (5)	Gross Coef. of Determination (adj.) (6)
1	.8532	18.8278	.0327	1.4359	.9674
2	.7215	12.5709	.1704	3.5399	.9596
3	.5921	9.4116	.2770	4.8339	.9533
4	.2356	4.3359	.3309	5.4918	.9123

Note: See equation (19) for the general regression underlying this table.

When both indicators are included in the regression as well as the induced component of the forecasts, with the forward rates again serving as the dependent variable, the results are similar to those shown for the indicators taken separately. In Table 4-16, we see that the presence of stock prices in the regression reduces the net contribution of the Index of Industrial Production. The adjusted coefficients of determination shown in Table 4-16 are, in fact, lower than those for

TABLE 4-16. Regression of Forward Rates on Estimated Induced Component, Index of Industrial Production, and Index of Industrial Stock Prices (Treasury Data, Quarterly, 1949–64)

Span of Forecast (quarters) (1)	Squared Partial Correl. Coef. of F^* (2)	t-Value of Reg. Coef. Attached to F^* (3)	Squared Partial Correl. Coef. of S (4)	t-Value of Reg. Coef. Attached to S (5)	Squared Partial Correl. Coef. of I (6)	t-Value of Reg. Coef. Attached to S (7)	Gross Coef. of Determination (adj.) (8)
1	.8545	18.7676	.0343	1.4604	.0094	−.7561	.9673
2	.7211	12.4550	.0733	2.1792	.0002	.1120	.9589
3	.5929	9.3479	.1091	2.7104	.0077	.6830	.9529
4	.2359	4.3044	.1345	3.0541	.0163	.9960	.9122

Note: The general regression underlying this table is as follows:

$$_{t+n}F_t = B_1({}_{t+n}F_t^*) + B_2S + B_3I + V.$$

the simpler regression in Table 4-15. This experiment concludes the analysis of the relationship between the indicators of economic activity and the decomposition of forecasts.[30]

V. NONFORECASTING COMPONENTS OF FORWARD RATES

LIQUIDITY PREFERENCE

The expression "liquidity preference" appears in the interest rate literature in two distinct contexts, a fact that has led to a certain amount of confusion. One concept of liquidity preference describes the liquidity premium as a component of the forward rates that is added to the expected rate in accordance with the degree of uncertainty with which the expectation is held. The less certain the expectation, the higher the premium. Since the prices of longer-term securities typically vary more than those of shorter maturities, the premium would ordinarily be higher on longer maturities; the greater variation of prices implies a greater variation in the possible holding period yields and, therefore, a greater uncertainty attached to any particular yield. This analysis is applicable, although with varying degrees of importance, whether the security is held to maturity (in which case there is no selling price variation) or is sold in advance of maturity. Even when the security is held to maturity, its value will fluctuate with the price, and the measured wealth of its holder along with it. In many situations — in principle, in all situations — a decline in measured or paper wealth is as significant as a decline in realized or cash equivalent wealth. Where it is not, the price fluctuations are less important. Since long-term securities provide stable income streams and obviate the expense of continual reinvestment, they are preferable to short-term securities in cases in which unrealized capital changes are not important. A positive liquidity premium implies that on balance the reverse is true. Since there are arguments for and against the likelihood of a liquidity premium, its presence is ultimately an empirical question.

In Keynesian literature, the term "liquidity preference" means

[30] For an interesting analysis of the relationship between economic indicators and the forward rates see Wendel [17].

something else. There, the liquidity preference function relates the demand for money to the *level,* not the dispersion, of interest rates. According to Keynesian theory, when this level is low relative to some normal or typical rate, investors expect future rates to be higher, and vice versa. If the interest rate were regarded simply as the opportunity cost of holding money, implying a negatively inclined demand curve for money (as a function of the interest rate), the introduction of this expectational theory justifies hypotheses about its shape [13]. Liquidity preference in this context refers to *expectations* of a change in the level of rates, not to the *risk* of a change in rates. For a given expectation, the liquidity premium due to risk will vary with the certainty with which the expectation is held. When, for example, the current rate is at its normal level, and therefore, according to the expectational theory under review, rates are not expected to change, liquidity premia may still exist and may differ in accordance with the *risk* that the expectation of no change will prove to be wrong. The determinants of this risk (and, therefore, of the liquidity premia) are difficult to specify; but certainly the dispersion of rates in the near past will be a factor in the expected dispersion of future rates.[31]

This notion of an expected return to normalcy is the basis for the controversy some years back over the so-called liquidity trap. In principle, a liquidity trap exists when the level of rates falls so far below normal that investors, expecting an ultimate rise in rates, and

[31] It may be that this dispersion is itself a function of the level of rates, a fact that would empirically obscure the distinction between the two concepts of liquidity preference. But, if there is such a relation, it should hold not for the *arithmetic* difference between the level of the current rate and the mean level but for the *absolute* difference, since it is the probability of a turn, whether up or down, that justifies the relation. Therefore, to predict liquidity premia with knowledge of the level of rates alone requires an additional hypothesis that investors' concern over dispersion is asymmetrical. When the level of rates is high and expected to fall, the investor risks a double beating: He accepts a lower rate on longer term securities, whose prices will fall drastically in the event future rates rise instead of fall. This situation could produce a large liquidity premium. However, when the level of the current rate is low and expected to rise, the investor requires a higher rate on longer-term securities in compensation for the expected rise in rates and capital loss. In the event he is wrong and rates actually fall further he is twice blessed—he obtains the higher initial yield plus the unexpected capital gain. Since the investor, in effect, hopes he is wrong, he does not require a liquidity premium to cover this possibility. This asymmetry may account for the often hypothesized positive relation between liquidity premia and the level of rates. On this point see Cagan [1].

therefore decline in prices, are unwilling to hold bonds out of fear that the expected capital loss will wipe out any income they earned from holding bonds instead of money. At this level of rates, the demand for money is said to be infinite. The use of the term "liquidity" in this context implies market forecasting of future rates. In the case of the liquidity trap, the term "liquidity" refers to the investor's preference for holding money rather than bonds because the *expected return* from holding bonds (consisting of interest payment plus the difference between the buying and expected selling prices of the bonds) is too low to compensate him for losing the convenience of storing his wealth in the form of money. Although originally expressed in terms of bonds and money, the concept of liquidity preference can be generalized to mean equalization of *expected* rates of return on all assets. When the level of rates is high, for example, investors, expecting rates to fall (and therefore prices to rise), will accept a lower nominal yield on longer-term securities in expectation of a favorable capital change when the rates fall. This situation is embodied in a declining yield curve. At the same time that the declining yield curve implies the market's forecast of lower rates, it can be said to reflect the market's attempt to equalize expected yields on different maturities. In either case, the shape of the yield curve is a measure of expectations.

Recent writers on the term structure have been aware of the distinction between the two concepts. Wendel [17], for example, distinguishes "regressive forecasting" from "liquidity-hedging," his terms for the expected return to normalcy and the attempt to avoid the dispersion of possible yields around the expected yield, respectively. Kessel [7] ingeniously used the distinction to explain the so-called humped yield curve. After demonstrating to his satisfaction[32] that liquidity premia are positively related to the level of one-period spot rates, Kessel argued that these premia dominate the short part of the yield curve. When the level of rates is high relative to some normal level, however,

[32] According to Kessel, if the bill rate is taken as compensation for eschewing the services of money, then liquidity premia may be viewed as compensation for eschewing the services of money substitutes. Since a three-week bill is closer to a two-week bill than a two-week bill is to money, the premium for holding a three-week bill instead of a two-week bill is less than the premium for holding a two-week bill instead of money. To complete the analogy, Kessel reasoned that liquidity premia should rise with interest rates: The cost of holding two- instead of three-week bills should rise with the cost of holding money instead of two-week bills [8, pp. 25 ff].

future rates are expected to revert toward their normal level, causing the yield curve to decline. This effect, according to Kessel, is most apparent beyond the short part of the curve. The resulting yield curve when the level of rates is high has the familiar humped shape, similar in appearance to a right-skewed frequency distribution. Thus, Kessel's explanation of the humped yield curve clearly requires both a liquidity premium positively related to the level of rates and a convex distribution of extrapolative weights.

ERRORS OF MEASUREMENT

There is little question but that the smoothing of the term structure data into yield curves enhances the autoregressiveness of the data. The use of the raw data in place of readings from the yield curves would result in lower coefficients of determination in the extrapolative regressions, as well as in the error-learning model regressions. Apart from the closeness of fit, the question remains whether the smoothing of data is also responsible for the convex pattern of weights observed earlier and, therefore, for the conclusions that were based on this pattern of weights.

One conclusion of the earlier analysis was that the parameters observed in connection with any one of the three autoregressive models implied the parameters of the other two. It is enough, therefore, to show that the parameters of at least one of the models do not depend on the smoothing of the data in order to release the entire analysis from this constraint.

A study of the actual yield curves reveals that, regardless how faithful the yield curves are to the raw data, there is no tendency for them to alter the direction of rates along a given curve. In October 1959, for example, the level of government rates was relatively high, and the yield curve, as well as the yields themselves, declined with maturity; in April 1963 the level of rates was relatively low, and both the curve and the yields inclined.[33] The same idea materializes in the case of Durand's data [4, "Basic Charts"]. This characteristic of the data motivated the expected return to normalcy model and constitutes evidence of the appropriateness of the model.

Since the inverse relationship between the slope of the yield curve

[33] *Treasury Bulletin*, December 1959, p. 54, and June 1963, p. 71, respectively.

and the level of rates is characteristic of the raw data rather than an artifact of the yield curves, the test of the expected return to normalcy model is sound. While the use of the raw data would result in poorer fits and less stable coefficients, it could not abrogate the inverse relationship, which is observable even without regression analysis. The main results of this study, therefore, do not depend on the smoothing of the data. In the case of the Durand data, however, the importance of the extrapolative component is likely to be overstated.

VI. ACCURACY OF THE FORECASTS

THE MEAN SQUARE ERRORS

Several commentators have rejected the expectations hypothesis on the grounds it was implausible for market forecasts to be so bad for so long ([5] and [3]). Meiselman, however, regarded the plausibility criterion as irrelevant, alleging that the expectations hypothesis asserts only that the market *attempts* to forecast future rates, not that it is successful. Between these extremes are those writers who regard the bad forecasting record implied by the term structure as an indication that expectations, while perhaps an important determinant of the term structure, is not the only relevant factor. Within this group, Kessel emphasized the importance of liquidity premia, while Wendel proposed some combination of liquidity preference and hedging.

The presence of liquidity premia would cause an upward bias in the forecasts. However, it does not follow that an upward bias implies the presence of liquidity premia, since the forecasts themselves may be biased. Whatever the cause of the commonly observed tendency for long-term rates to exceed short-term rates, it is useful to separate this effect from the purely random error of forecast. Even when non-forecasting components are not an issue, for example, when the data are explicit forecasts, isolation of a bias in prior forecasts may suggest adjustments of current forecasts to offset the bias. Therefore, instead of simply reporting the mean square error of forecasts, we shall separate the contributions to the error of the bias and the random term.

The mean square error of forecasts, $M = E(A_i - F_i)^2$, can be broken down into the mean error squared $(\bar{A} - \bar{F})^2$, and the mean of the squared deviations of the individual errors from the mean error, or the variance of the errors, $E[(A_i - F_i) - (\bar{A} - \bar{F})]^2$. The purpose of this decomposition is to allow us to distinguish between the systematic errors of forecast, or the bias, and the random errors.[34] A perfect correlation between the actuals and forecasts implies the absence of a random error but does not preclude a consistent bias in the forecasts. The random error, for example, would be unaffected by the existence of a constant liquidity premium that caused the forward rate to consistently overstate the true forecast. In the case where the correlation coefficient was 1, the mean square error of forecast would be zero only in the case where $a = 0$ and $b = 1$ in the regression:

$$A_{t+n} = a + b({}_{t+n}F_t) + u. \tag{20}$$

Since $a = \bar{A} - b\bar{F}$, a would equal the mean bias only in the case where $b = 1$. There are, therefore, two sources of bias: mean bias and bias in the slope, the remaining error being random. It is useful, for this reason, to further decompose the mean square error.

$$M = E(A_i - F_i)^2 = (\bar{A} - \bar{F})^2 + (1 - b_{AF})^2 S_F^2 + (1 - r_{AF}^2) S_A^2, \tag{21}$$

where b is the regression coefficient in (18) and r^2 the coefficient of determination. Following Mincer and Zarnowitz, we combine the first two terms on the right into U_F and call the residual term M_F^1, so that $M_F = U_F + M_F^1$. If the combined term, U_F, is not significantly different from zero, it is not necessary to show the two components of error separately. Since the significance of both the constant term and the regression coefficient of (20) enters into the significance of U_F, it is desirable to test the two terms together.[35]

To evaluate the effectiveness of the forecast it is useful to establish as a criterion a hypothetical set of forecasts that can be generated mechanically, and the effectiveness of which varies with the ease of forecasting any particular series. In general, the ease in forecasting a series is a function not only of the variance of the series but also of the degree of autocorrelation in the series. The more systematically the series varies over time the easier it is to forecast, for reasons that should be apparent from our earlier analysis. However, a series whose

[34] See Jacob Mincer and Victor Zarnowitz, "The Evaluation of Economic Forecasts," and the references therein, in Chapter 1 of this volume.

[35] The formula for this test is given in [7, p. 24].

variation is random with respect to time, unless its range of movement or its variance is very small, is difficult to forecast (assuming no knowledge about the relation between this series and other series).

Many researchers have used the so-called naive models to generate their hypothetical or "benchmark" forecasts. The naive models are simply a class of moving averages relating the current value of some series X to its own lagged values, or

$$X_t = \sum_{i=1}^{n} B_i X_{t-i}. \tag{22}$$

The two most common variants of these models are the "no change" and the "same change" models. In the former, B_1 is set to 1 and B_i to 0 for $i > 1$. In the latter, $B_i = 2$, $B_2 = -1$, and $B_i = 0$ for $i > 2$. Other, more complicated, systems of weights have been used. But these methods have in common the fact that the B_i coefficients are chosen without regard to achieving the maximum fit between X_t and ΣX_{t-i}. Since the degree of fit is, in this context, the measure of the effectiveness of the hypothetical forecasts, these methods do not provide the strongest criterion possible for evaluating them.[36]

For reasons analogous to those used to support our method of estimating the induced component of the forecast, an estimation form is required that will exhaust the extrapolative potential of the data. The naive model must provide the sternest criterion possible in evaluating the forecast. To aim for less is to tie the conclusion to the particular naive model chosen and therefore to invite questions of its importance. No set of a priori weights will guarantee the best extrapolation.

In principle, an optimal extrapolation depends on the structure of the series and must be estimated from the data.[37] In practice, this is difficult to accomplish but, as an approximation, we fit one regression for the whole period and use the (squared) standard error of the estimate as the measure of the mean square error of forecast. By regressing the forward rates on the appropriate target values within the sample period for which the benchmark was computed, we computed comparable measures of the errors of forecast of the forward rates and of the

[36] While we have no desire to magnify this fine point, it is a fact that some researchers test their models against excessively weak naive models and congratulate themselves unduly on their success.

[37] See Mincer and Zarnowitz, p. 32 ff of this volume.

optimal benchmark. We can say optimal benchmark because the method equates the squared standard error of the estimate, necessarily a minimum, with the mean square error of forecast.

To compute this benchmark we ran, for each span of forecast n, the following multiple regression

$$A_{t+n} = B_1 A_t + B_2 A_{t-1} + \cdots + B_{i+1} A_{t-i} + W_n, \tag{23}$$

where A_{t+n} is the spot rate at $t + n$ and W_n a random term. Equation (23) measures the maximum amount of the variation of the spot rate that it is possible to explain with the variation of its own lagged values.[38]

We computed the summary statistics M_F, U_F, and M_F^1, described earlier, for both the forward rates and the hypothetical forecasts based on the naive model. We then took ratios of the former statistics to the latter to determine the forecasting effectiveness of the forward rates relative to the moving average forecasts. The presumption is that one forecaster, with knowledge of current and past economic activity and of the relations among the available economic time series, should forecast better than another forecaster with knowledge only of the past behavior of the series being forecast. Table 4-17 shows the relevant results of the simple regressions of the actual values (that is, the forecast targets) on the relevant forward rates for the Durand and the Treasury data. The general form of the regressions is shown in equation (20) above. Table 4-18 decomposes the errors of forecast into three components, as described in equation (21), and relates the components to those of the hypothetical forecasts. If the forward rate had forecast as well as the hypothetical forecasts, the figures in column (8) of Table 4-18 would center on unity. In the case of the Durand data, the ratios are between 1.5 and 2.0; although the corresponding figures for the Treasury data are not far from unity. Comparing columns (7) and (8), especially for the Treasury data of Table 4-18, one can see the extent to which the bias obscures the relation between the forecasts and the future rates. In conclusion, while the forecasts

[38] One should take account of the stringency of this criterion in evaluating the empirical results below. However, given the low level of accuracy of the autonomous component of the forecast, described above, it is unlikely the conclusion of the present analysis would change with respect to the Durand data, although perhaps it would with respect to the Treasury data, if a weaker naive model were used.

TABLE 4-17. Relation Between Forecasts and Future Spot Rates

Span of Forecast (1)	b_{AF} (2)	t-Value (b_{AF}) (3)	Intercept (4)	R^2 (adj.) (5)	S_u^2 (6)
		Durand: Annually, 1916–54			
1	1.0397	13.6119	−.4013	.8328	.6170
2	1.0456	8.6671	−.6383	.6670	1.2318
3	1.0318	6.3368	−.7592	.5141	1.7930
4	.8811	4.7555	−.7523	.3688	2.3291
5	.9938	3.8369	−.9016	.2705	2.6919
6	.9473	3.0377	−.8118	.1819	3.0188
7	.8608	2.3785	−.5509	.1118	3.2775
8	.8330	1.9541	−.5341	.0709	3.4287
9	.6167	1.2561	.2368	.0154	3.6332
		Treasury: Quarterly, 1949–64			
1	.8603	13.9589	.0666	.7885	.1627
2	.5467	5.2188	1.0583	.3353	.5114
3	.6416	6.7367	.8176	.4605	.4150
4	.7056	6.5607	.6960	.4471	.4253

Note: Each of the regressions is of the following form: $A_{t+n} = a + b({}_{t+n}F_t) + u_n$, $n = 1, 9$.

implied by the Durand data clearly do not evince effective forecasting ability, those implied by the Treasury data do much better, and they deteriorate less rapidly than the naive model.

FORECASTING AND THE TERM STRUCTURE OF INTEREST RATES

This study has explored the importance of extrapolative forecasting as a determinant of the term structure. It has shown not only that an extrapolative model accounts for a large part of the variation of the forward rates but that this model implies a method of making forecasts that is both plausible and consistent with other proposed models. If extrapolation of past spot rates were in fact the *only* method the market used to forecast future rates, then, at best, its forecasts would equal the performance of the naive model. (In this context recall that the naive model used in this study is optimal in view of the method used to evaluate the forecasts.) While the Treasury data reveal some

TABLE 4-18. Comparison Between the Errors of Forecast Due to the Forward Rate Forecasts and the Moving Average Forecasts (Durand Data, Annual Observations, 1916–54)

Span of Fore- cast (1)	$(\bar{A}-\bar{F})^2$ (2)	$(1-b_{AF})^2 S_F^2$ (3)	M_F^1 $(1-r_{AF}^2)S_A^2$ (4)	M_F (cols. 2 + 3 + 4) (5)	M_A^* (6)	RM (7)	RM^1 (8)
			Durand: Annually, 1916–54				
1	.0626	−.0152	.6170	.6644	.3935	1.6884	1.5679
2	.2087	−.0372	1.2318	1.4033	.8131	1.7258	1.5115
3	.3765	−.0467	1.7930	2.1228	1.1297	1.8790	1.5871
4	.5488	−.0645	2.3291	2.8134	1.5182	1.8531	1.5341
5	.7590	−.0695	2.6919	3.3814	1.7190	1.9670	1.5659
6	.9356	−.0953	3.0188	3.8591	1.8729	2.0604	1.6118
7	1.0396	−.0883	3.2775	4.2288	1.9904	2.1245	1.6466
8	1.1661	−.0747	3.4287	4.5201	1.8147	2.4908	1.8894
9	1.2510	−.0407	3.6332	4.8435	1.8151	2.6684	2.0016
			Treasury: Quarterly, 1949–64				
1	.1341	−.1247	.1627	.1721	.1499	1.1481	1.0854
2	.1009	.0605	.5114	.6728	.4779	1.4078	1.0701
3	.0574	.0379	.4150	.5103	.4013	1.2716	1.0341
4	.0157	.0286	.4253	.4696	.4427	1.0607	0.9607

Note: Columns 2 through 5 are based on equation (19). See text for explanation of symbols. Column 6 is computed from equation (21), where $A_{t+n} = A^*_{t+n} + W_n$. Since there is no bias in the moving average, $M = M^1$.

autonomous forecasting, its extent is too small to compensate for the market's inability to take maximum advantage of the extrapolation. If it is true that extrapolation is the dominant method of forecasting, then the ratios in column 8 of Table 4-18, again referring to the Treasury data, are actually surprisingly close to unity. The fact that the performance of the forward rates improves relative to the naive model as span increases, and actually exceeds the naive model in the fourth span, is consistent with the evidence in Table 4-12, which shows that the importance of the autonomous component increases with span.

Nevertheless, the skimpy evidence of autonomous forecasting is disappointing to those who support a simple expectations hypothesis. It is not possible to ascertain whether the ineffectiveness of the autonomous component is due to a varying liquidity premium, in the absence of a method of accounting separately for the variation of this component.

Another possible source of trouble lies in the greater vulnerability of the forward rates to errors of measuring the long-term rates as the maturity of the latter increases.[39] Such an effect can be visualized as follows: Assume simple interest. Say the one-period spot rate is 3.00 and the two-period, long-term rate 3.01; the implied forward rate is $2(3.01) - 3.00 = 3.02$. Now consider the same numbers for higher maturities: The nine-period long-term rate is 3.00 and the ten-period long-term rate 3.01. In this case, the implied forward rate is $10(3.01) - 9(3.00)$ or 3.10. If either of the two rates in each comparison included a measurement error of one basis point, it would throw off the ten-span forward rate by ten points, but the two-period rate by only two points. We have no way of evaluating the importance of this effect.

In the case of the Durand data, however, the markedly inferior performance of the forward rates compared with that of the naive model is troublesome. The perverse behavior of the autonomous component explains part of this result, but there still remains the inability to make maximum use of the extrapolative potential.

The persistent bias in the forward rate forecasts is another matter. This study has not isolated the source of this bias. There is an undeniable tendency for yield curves to incline; an undeniable tendency for longer maturities to yield more than shorter maturities. Kessel persuasively argues that the source of this bias lies in the presence of a liquidity premium. Since interest rates for most of the postwar period have been low relative to historic levels, the continuing expectation that they would rise also imparts a bias to the forecasts. Whatever the source of the bias, there is little question but that a method of evaluating the forecasts that failed to separate the bias from the random error would cast a shadow over the relevance of forecasting in the determination of the term structure.

VII. SUMMARY AND CONCLUSIONS

This report is an extension of recent attempts to evaluate the hypothesis that yield differentials of securities, differing only by their term to maturity, are determined by market forecasts of future spot rates of

[39] This point was made by Van Horne [15].

interest. Since its inception, this hypothesis has appealed to the theorist's urge to summarize in a simple way the vagaries of many groups acting from diverse motives over the whole range of securities. Opponents of the hypothesis, aside from those who argue that the markets for different maturities are independent of each other, think it implausible to suppose that investors make some point estimate of a short-period rate expected to prevail several years later. This criticism, however, is based on an invalid analogy between a market and an individual. The market, like a mean, is an abstraction that may differ from any of the elements participating in it. Horizons differ among the many investor groups, and few may extend over a long period. But the fact remains that yields on nine-year and ten-year securities differ by varying amounts that are unlikely to be explained by variations in the supply of securities. These variations in the yield curve will inspire speculative action whenever their implied forward rates are unusual. The process of ensuring reasonable forward rates is arbitrage rather than speculation, and it lends a certain smoothness to the term structure. The practice of arbitraging along the yield curve implies no unusual horizon, although some investors will look further than others and define reasonableness in a narrower range, at which point arbitrage shades into speculation. There is no great risk in eliminating an implied negative forward rate, but exchanges along the yield curve become more and more speculative as investors attempt to replace one reasonable forward rate with another. How far the observed forward rate must differ from the expected rate before it inspires arbitrage will differ for different investors: Some, obviously, never arbitrage along the yield curve; others, only when the implied forward rates are, perhaps, negative; others will come in for fine adjustments. There is in this process nothing to offend one's sense of reality, even if the more abstract analysis temporarily ignores the market mechanism.

At some point there may be two alternative investments available—a security with n periods left to maturity and one with $n + 1$. The yield differential of the longer security depends on what one-period yield is *expected* in the $(n + 1)$st period. A rise in the one-period yield in $n + 1$ would drive down the price on the longer security at the start of period $n + 1$, in order that buyers of this security in $n + 1$ could get the same yield as that available on a one-period security during the $(n + 1)$st period. The expectation in period t of a fall in the price of the $n + 1$

term security in period $n + 1$ is equivalent to an expectation of a rise in the one-period rate in period $n + 1$. If, on balance, investors act to equalize the yields of the different maturities, then any observed differences in these yields at some point implies particular expectations of future rates (or prices). The link between the equalization of expected holding-period yields and forecasts of future rates, therefore, resides in the fact that a holding-period yield is equal to the sum of the coupon return (if any) and the difference between the known buying and the *expected* selling prices of the security.

The present study is not directly concerned with an evaluation of the expectations hypothesis. Instead, the hypothesis is assumed correct, and the data are evaluated for their implications of how the forecasts are made and for the effectiveness of the forecasts. To consider how the forecasts are made, a dichotomy—one of many possible frames of reference—is proposed that separates the forward rates, regarded for this purpose as forecasts, into the part that is related to the current and past spot rates and the part that is not—the induced and autonomous components, respectively. The induced (or extrapolative) component is most generally expressed as a weighted average of current and past spot rates but is not limited to an extrapolative procedure as the term is generally used. Instead, the induced components collects all possible arrangements of current and past spot rates, as well as the autoregressive components of other variables that are contemporaneously related to the forecasts. Most of the text is devoted to a discussion of this point. The optimal arrangement of the current and past spot rates, as distinct from the arrangement the forecasts are observed to utilize, is used as a standard for evaluating the accuracy of the forecasts.

The study considered the following three variants of the induced component:

(1) the extrapolation model $\quad {}_{t+n}F_t = B_1A_t + B_2A_{t-1} + \cdots + B_nA_{t-n-1};$

(2) the error-learning model $\quad {}_{t+n}F_t - {}_{t+n}F_{t-1} = a + B(A_t - {}_tF_{t-1});$

(3) the return to normalcy model $\quad {}_{t+n}F_t - {}_{t+n-1}F_t = K(A_t - {}_nA_t),$
$k < 0;$

where, ${}_{t+n}F_t$ = forecast made at t of a target at $t + n$; A_t = actual value at t; and ${}_nA_t$ = normal value at t. The first model says forecasters project some weight of averaged current and past actual values. The second model says forecasters revise their forecasts in accordance with revealed errors of earlier forecasts. The third model says forecasters project a change in target values in the direction of a normal value. All three models are, of course, linear.

Since one linear autoregressive model is naturally a linear transformation of any other linear autoregressive model, the parameters estimated for one such model imply corresponding weights for the others. However, for a given span of forecast, the models need not utilize the same number of past values; more generally, the models may differ in the number of their zero coefficients. This study circumvents the problem of comparability by considering several spans of forecasts, each of which utilizes different values of the series. Therefore, the error-learning model, which utilizes only one lagged value of the series for a given span, will, over several spans, generate a sufficient number of parameters for comparison with an extrapolative model of one span. Since the term structure data incorporate several spans of forecast, they are well suited to the present study.

By establishing algebraically the transformations bridging the extrapolative, error-learning, and return to normalcy models, the study is able to demonstrate the comparability between the parameters directly estimated for any one of the models and those implied by the estimated parameters of the other two. Because of the common derivation of the three models, or of any other linear autoregressive models, it is clearly not possible to select from among them one that can be said to describe the data most adequately. The models are simply alternate methods of describing the induced component.

The algebraic equivalence among the three models, however, does not nullify the value of distinguishing among them, since they each imply different motivations for behavior. It is entirely possible, for example, to obtain a set of parameters for the extrapolative and error-learning models that are plausible but that yield parameters for the return to normality model that contradict the hypothesized expected return to normalcy. While it is not possible to ascertain the precise motivation underlying the forecasts, it is possible to distinguish between the models that do and do not yield parameters that are plausible with respect to their behavioral implications.

Since interest rates are related on a contemporaneous basis with other variables that could conceivably influence the forecasts, the autoregressive structure of interest rates will include part of the autoregressive structure of these other series—the extent of the inclusion being determined by the extent of the contemporaneous correlation. The relative influence of the induced and autonomous components on the forecasts depends, therefore, on which of the two components is measured directly. This study estimated the autonomous component only after exhausting the extrapolative potential of the forecasts, thereby excluding from the autonomous component the autoregressive components of the other variables. It is important to interpret the relative importance of the two components in the light of this method, which clearly exaggerates forecasters' actual reliance on the autoregressive structure of the target series.

With respect to the accuracy of the forecasts, the study showed that the forecasts did not perform as well as the autoregression model. While this result reveals a bad forecasting record, it does not imply that the market is not trying to forecast. This report has shown that moving averages of past rates account for the major part of the variation of the forecasts; the last result implies that these moving averages can themselves stand improvement.

An important part of the accuracy analysis lies in the breakdown of the total mean square error into the part due to bias and the part that is random. The importance of the bias relative to the total mean square error increases, in the case of the Durand data, from about 10 per cent for the one-span forecast to about 25 per cent for the nine-span forecast. For the Treasury data, the ratio is about one-third throughout. The extent of this bias has evoked considerable comment in the literature on the term structure, and its source has been variously explained. For most of the period in which the bias is observed in both sets of data, the one-period spot rates were abnormally low. The mechanism described earlier—the apparent expectation that rates will return toward some central tendency—would also produce a bias in the errors of forecast, especially in the Durand data where the rise in rates expected in the early 1930's failed to materialize for two decades.

The effect of liquidity preference on the term structure is a complex subject that is outside the scope of this report. It is likely that the magnitude of the liquidity premium changes with shifts in the degree of an investor's certainty about future rate movements. If so, the liquidity premia would obscure part of the effectiveness of the true forecasting component of the forward rates. A sequel to this study could very reasonably be based on an analysis of this subject. A clean test for the efficacy of this analysis could be based on the determination whether the removal of properly isolated liquidity premia, so isolated, showed the true forecasting component of the forward rates more clearly.

REFERENCES

[1] Cagan, Phillip, and Guttentag, Jack, eds., *A Study of Interest Rates*, National Bureau of Economic Research, forthcoming.
[2] Conard, Joseph, *Introduction to the Theory of Interest*, California University Press, 1959.
[3] Culbertson, John M., "The Term Structure of Interest Rates," *Quarterly Journal of Economics*, November 1957.
[4] Durand, David, *Basic Yields of Corporate Bonds, 1900–1943*, Technical Paper No. 3, National Bureau of Economic Research, New York, 1942.
[5] Hickman, W. Braddock, "The Term Structure of Interest Rates: An Exploratory Analysis," National Bureau of Economic Research, New York, 1942 (mimeographed).
[6] Hicks, John R., *Value and Capital*, London, 1946.
[7] Johnston, J., *Econometric Methods*, New York, 1963.
[8] Kessel, Reuben, *The Cyclical Movement of the Term Structure of Interest Rates*, National Bureau of Economic Research, 1965.
[9] Macaulay, Frederick R., *The Movement of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*, National Bureau of Economic Research, New York, 1938.
[10] Malkiel, Burton, "Expectations, Bond Prices, and the Term Structure of Interest Rates," *Quarterly Journal of Economics*, 1957, pp. 197–218.
[11] Meiselman, David, *The Term Structure of Interest Rates*, Englewood Cliffs, New Jersey, 1963.
[12] Theil, Henri, *Economic Forecasts and Policy*, Amsterdam, 1961.
[13] Tobin, James, "Liquidity Preference as Behavior Toward Risk," *Review of Economic Studies*, 1958.
[14] Van Horne, James C., "Interest Rate Risk and the Term Structure of Interest Rates," *Journal of Political Economy*, August 1965, pp. 344–351.
[15] ———, "Liquidity Preference, Interest Rate Risk, and the Term Structure of Interest Rates," doctoral dissertation, Northwestern University, 1965, unpublished.
[16] Wallace, Neil, "The Term Structure of Interest Rates and the Maturity

Composition of the Federal Debt," doctoral dissertation, The University of Chicago, 1964, unpublished.
[17] Wendel, Helmut, "Short Run Interest Rate Expectations in the Government Securities Market," doctoral dissertation, Columbia University, 1966, to be published.
[18] Wood, John H., "Expectations, Errors, and the Term Structure of Interest Rates," *Journal of Political Economy*, February 1963.