Volume Title: Essays on Interest Rates, Vol. 2

Volume Author/Editor: Jack M. Guttentag, ed.

Volume Publisher: NBER

Volume ISBN: 0-87014-224-0

Volume URL: http://www.nber.org/books/gutt71-2

Publication Date: 1971

Chapter Title: The Expectations Component of the Term Structure

Chapter Author: Stanley Diller

Chapter URL: http://www.nber.org/chapters/c4005

Chapter pages in book: (p. 413 - 433)

8

The Expectations Component of The Term Structure

Stanley Diller

THE TERM STRUCTURE

The term structure of interest rates—the variation in rates in relation to the term to maturity—involves one aspect of the general subject of yield differentials on financial assets.

The markets for short- and long-term securities are distinguished by the supposedly greater substitutability of short-term securities for money, the different economic roles associated with short- and long-term borrowing, the preponderance of short-term securities in Federal Reserve activities, and other factors as well.

Some writers are inclined to emphasize the differences between the short- and long-term markets largely on the basis of the differences among the lenders at various segments of the maturity spectrum, as well as differences in the purpose to which the loans are put. The markets are sometimes described as largely independent segments in which yields are determined by supply and demand conditions peculiar to each segment. Yet, while some lending institutions may reveal preponderant interests in certain maturity segments because of the timing of their liabilities, some part of their portfolios is typically permitted to seek the most favorable yield-risk combinations regardless of maturity. Many institutions, such as trust funds, are relatively free of maturity constraints and can serve as the medium through which the markets interact. Moreover, corporations often vary the maturity of their bor-

NOTE: This paper is based on a paper published in *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* Jacob Mincer, ed., New York, NBER, 1969.

rowings in response to market conditions. Recent writers have tended, therefore, to minimize institutional constraints and to emphasize instead the fluidity among markets not only by maturity or quality or type of financial assets but between financial and real assets as well.

This paper starts with a brief description of the expectations hypothesis, which relates the term structure at some time to the expectations at that time of the future interest rates. After describing an earlier test of this hypothesis the paper introduces a model that fits this earlier work into a more general context of autoregressive models and expectations. Particular behavioral models are derived from a general autoregressive model, and the relationship among them is explained. The forecasts implicit, by hypothesis, in the term structure are decomposed into a part attributable to autoregressive (or extrapolative) forecasting and a part not so attributable. The components are related to variables that might be reasonably used to forecast interest rates in an effort to explain the sources of the implied forecasts. Finally, the accuracy of the forecasts is measured and associated with the two components of the forecast.

THE EXPECTATIONS THEORY

Hicks was one of the first to consider the relationship between short- and long-term rates. His proposition is that under certain restrictive assumptions the long-term rate is the geometric mean of the short-term rates spanning the same term to maturity. The well-known formula for the price or present value of a bond is given by equation (1).

$$PV = \frac{C_1}{1+R} + \frac{C_2}{(1+R)_2} + \cdots + \frac{C_n + P}{(1+R)^n}, \tag{1}$$

where PV is the present value of the current market price of security, C_i is the coupon payments, P is the principal, and R is the market yield. Alternatively, a long-term bond is equivalent to the automatic reinvestment in a consecutive series of short-term bonds at rates current at the time of reinvestment. The formula in this case is:

$$PV' = \frac{C_1}{1+r_1} + \frac{C_2}{(1+r_1)(1+r_2)} + \frac{C_3}{(1+r_1)(1+r_2)(1+r_3)} + \cdots + \frac{C_n + P}{(1+r_1)\ \ldots\ (1+r_n)}, \tag{1a}$$

where r_i is the one-period rate for the ith period. The r_i are called forward rates. In the event there are no coupon payments (i.e., all C are 0), equating the two present value formulas is equivalent to equating the last terms of each:[1]

$$(1 + R)^n = (1 + r_1)(1 + r_2) \ . \ . \ . \ (1 + r_n).$$

From the formulas of two maturities (n and $n - 1$) we can derive

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

where R_n is the internal rate of return on a bond with n periods to maturity. In these formulas, the forward rates, r_i, are the short-period rates pertaining to a future period and are determined by the differential between currently observed long rates of appropriate maturity.[2]

The substances of Hicks' hypothesis is that a forward rate is in fact the short-period rate that is expected to prevail in that period and that the long rates of different maturities adjust in order to be consistent with these expected rates. In addition, he hypothesized that changes in the spot rate, i.e., r_1 or R_1, stimulate expectations of subsequent changes in r_i and, hence, in R_i. In this way variations in short-term rates, R_1, produce, via the expectation mechanism, variation in long-term rates R_i ($i > 1$). This hypothesis, often amended (as it was by Hicks) to include the effect of differences in liquidity between long- and short-term securities, has become known as the expectations hypothesis.

The expectations hypothesis is sometimes rejected on the grounds of its alleged implausibility: investors, it is sometimes said, do not attempt to forecast short-term rates far into the future. Properly viewed, however, the expectations hypothesis does not imply implausibile behavior.

The forward rates, which this study equates with forecasts, are not the numbers to which investors directly respond; nor is it correct to

[1] Neil Wallace has estimated the effects of ignoring these restrictions in computing the internal rate of return, R. He concluded that these effects are small. See "The Term Structure of Interest Rates and the Maturity Composition of the Federal Debt," unpublished Ph.D. dissertation, University of Chicago, 1964, pp. 10–12.

[2] While the example is based on adjacent (i.e., a one-period difference in term to maturity) long rates, which together imply a one-period forward rate, the long rates can be spaced arbitrarily to imply forward rates whose maturity equals the difference in maturities between the long rates.

regard them as consciously pinpointed forecasts made by the public looking well into the future, though some individuals may form precise forecasts. Deducing the forward rates from the combinations of long rates and evaluating them as forecasts is an analytical device justified by the hypothesis and the arithmetic of interest rates; the efficacy of this method is independent of judgments of the plausibility of some hypothetical forecasting mechanisms.

Without actually specifying forecasts of rates ad infinitum, investors can react to yield differentials and adjust the maturity of their holdings in accordance with their expectations of rates. Regardless of the certainty of their convictions, investors are continually required to decide on combinations of yield and maturity on the basis of limited information, vague expectations, and publicized market attitudes. Deciding between the purchase of a long- and short-term security does not require a point forecast of a one-year rate twenty years out even though the aggregate of such decisions leads to a yield structure that corresponds with one that point forecasts could produce. In a period of high rates, for example, an investor may well decide to purchase a long-term bond—whether for the capital gain expected when rates ultimately fall or merely to receive a high yield for a long period. Summed over many investors, this thinking would depress long-term yields and produce apparent forecasts of declining rates. The investors who purchase the long-term bond implicitly forecast, at a minimum, that short-term rates will not rise or stay high. Among this group are those who think a decline in rates is imminent and others who think rates will fall only after an inflation subsides, perhaps ten years out. The weight of these opinions will mold the yield curve. While few if any investors will distinguish their 14-year forecast from their 15-year forecast, the availability of the two maturities forces a choice, and the resulting yields will reflect the frequency distribution of investors of various horizons. If by accident the yields become out of line, a sufficient number of investors, indifferent between the two maturities, will set them right. This arbitrage along the yield curve is facilitated by the investors' ability to borrow for the purpose of buying or short-selling securities, as well as by the ability of issuers to vary their maturities to correct imbalances among yields.

Investigations of implied market forecasts—whether interest rates, stock or commodity prices, foreign exchange rates, or personal incomes, profits, and any other economic variables—pertain to the weight of market forces, which are themselves the aggregate of individual and group decisions, rather than personal motivation or institutional anomalies. While these market forces are personified for expositional convenience, and motivations are established for hypothetical decision makers, the efficacy of the analysis is predicated on its ability to predict behavior and not on the plausibility of its expositional devices.

To test the expectations hypothesis it is necessary to treat the forward rates inferred from the term structure of long rates as forecasts. A well-known test by Meiselman[3] is based on an evaluation of the consequences of using the forward rates in a model that has on other occasions successfully described changes in forecasts. Meiselman reasoned that if forward rates are forecasts they would be reviewed as new information came to light. Thus, he supposed that errors in forecasting would lead to revisions of forecasts. He used the error-learning model, which describes the effect of the currently observed error of a prior forecast on the current revision of prior forecasts of some future period,

$$ {}_{t+n}r_t - {}_{t+n}r_{t-1} = a + B(R_t - {}_tr_{t-1}) + u, \tag{2} $$

where ${}_{t+n}r_t$ is the forecast made in t of the rate expected in period $t + n$; R_t is the spot rate in period t; and u is a random term.

Meiselman used Durand's *Basic Yields*[4] on corporate securities to compute the forward rates for the test. These data permit the computation of ten consecutive one-period forward rates for each observation period, which in turn permit eight regressions of equation (2), i.e., $n = 1, 8$. He found that while each of the eight regression coefficients were significantly different from zero, their magnitude and statistical significance declined as the span of forecast, n, increased. Meiselman regarded these results to be consistent with the hypothesis that the forward rates are forecasts. He argued that the coefficients should fall with increasing span of forecasts since forecasts of increasing span became increasingly remote from the current error, and revisions of this rate were therefore less dependent on the current error.

EXTRAPOLATIVE FORECASTING

Meiselman's error-learning model describes a specific technique of forecasting. Mincer has shown that it is actually a rearrangement of

[3] David Meiselman, *The Term Structure of Interest Rates,* Englewood Cliffs, 1962, Chapter II.

[4] David Durand, *Basic Yields of Corporate Bonds, 1900–1942,* New York, NBER, 1943. The complete set of data is listed in the National Industrial Conference Board's *Economic Almanac 1967–1968,* p. 416.

terms of an autoregressive model,[5] which allows us to view the forecast as dependent on prior values of the forecast and the actual variable. In this way we can interpret some characteristics of the expectations implied in Meiselman's results. Mincer relates the forecasts of a given span to the forecast or actual values of each prior period. For example, a one-period forecast is related to the current actual value and each of the prior actual values; a two-span forecast to the one-span forecast of the prior period, then to the current actual value, and finally to all prior actual values—as in the following equation:

$$ {}_{t+n}F_t = A + B_1({}_{t+n-1}F_t) + B_2({}_{t+n-2}F_t) + \cdots + B_nA_t + B_{n+1}A_{t-1} + \cdots + {}_{t+n}E_t, \quad (3)$$

where ${}_{t+n}F_t$ is the forecast made in t referring to $t+n$, A_{t-i} is the actual value of the series i periods in the past, and E is the random term. This model extrapolates a weighted average of current and past values, substituting extrapolated for actual values between t and $t+n-1$ for forecasts of $t+n$. In addition to the extrapolative, or autoregressive, component there is an autonomous component, E, of the forecast. Equations (4) and (5) describe the forecasts made in t and $t-1$, respectively, referring to $t+1$.

$$ {}_{t+1}F_t = A + B_1A_t + B_2A_{t-1} + \cdots + B_{n+1}A_{t-n} + {}_{t+1}E_t, \quad (4)$$

$$ {}_{t+1}F_{t-1} = A + B_1\,{}_tF_{t-1} + B_2A_{t-1} + \cdots + B_{n+1}A_{t-n} + {}_{t+1}E_{t-1}. \quad (5)$$

In (5), since A_t is unknown at the time of forecast, $t-1$, the extrapolated value, ${}_tF_{t-1}$, is substituted. Subtracting (5) from (4) yields:

$$ {}_{t+1}F_t - {}_{t+1}F_{t-1} = B_1(A_t - {}_tF_{t-1}) + ({}_{t+1}E_t - {}_{t+1}E_{t-1}), \quad (6)$$

where the last term on the right is a random term. Equation (6) is identical to Meiselman's error-learning model.

By extending this formula to later maturities we can derive the B's from the coefficients in Meiselman's regressions. For the target $t+2$ the equations are,

$$ {}_{t+2}F_t = A + B_1({}_{t+1}F_t) + B_2A_t + B_3A_{t-1} + \cdots + {}_{t+2}E_t. \quad (7)$$

$$ {}_{t+2}F_{t-1} = A + B_1({}_{t+1}F_{t-1}) + B_2({}_tF_{t-1}) + B_3A_{t-1} + \cdots + {}_{t+2}E_{t-1}. \quad (8)$$

Subtracting (8) from (7) yields:

$$ ({}_{t+2}F_t - {}_{t+2}F_{t-1}) = B_1({}_{t+1}F_t - {}_{t+1}F_{t-1}) + B_2(A_t - {}_tF_{t-1}) + ({}_{t+2}E_t - {}_{t+2}E_{t-1}). \quad (9)$$

[5] See Jacob Mincer, "Models of Adoptive Forecasting," in *Economic Forecasts and Expectations, op. cit.*

The difference equation for $t+3$ is

$$({}_{t+3}F_t - {}_{t+3}F_{t-1}) = B_1({}_{t+2}F_t - {}_{t+2}F_{t-1}) + B_2({}_{t+1}F_t - {}_{t+1}F_{t-1}) + B_3(A_t - {}_tF_{t-1}) + ({}_{t+3}E_t - {}_{t+3}E_{t-1}). \quad (10)$$

It is clear that the revision variables on the right of (9) and (10) are themselves functions of the current error of forecast. Substituting (6) into (9) yields:

$$\begin{aligned}({}_{t+2}F_t - {}_{t+2}F_{t-1}) &= B_1[B_1(A_t - {}_tF_{t-1})] + B_2(A_t - {}_tF_{t-1}) \\ &\quad + ({}_{t+2}E_t - {}_{t+2}E_{t-1}) \\ &= (B_1^2 + B_2)(A_t - {}_tF_{t-1}) + ({}_{t+2}E_t - {}_{t+2}E_{t-1}). \quad (11)\end{aligned}$$

Substituting (6) and (9) into (10) yields:

$$\begin{aligned}({}_{t+3}F_t - {}_{t+3}F_{t-1}) &= B_1B_1[B_1(A_t - {}_tF_{t-1})] + B_2(A_t - {}_tF_{t-1}) \\ &\quad + B_3(A_t - {}_tF_{t-1}) + ({}_{t+3}E_t - {}_{t+3}E_{t-1}) \\ &= (B_1^3 + B_1B_2 + B_3)(A_t - {}_tF_{t-1}) \\ &\qquad + ({}_{t+3}E_t - {}_{t+3}E_{t-1}). \quad (12)\end{aligned}$$

The weights, B_i, that appear in (6), (11), and (12) are identical to those in the extrapolative forecasting equation, (3). Therefore, each of the eight regression coefficients, M_i, that Meiselman estimated with the error-learning model (2) ($n = 1, 8$) are estimates of the corresponding combinations of B_i in the difference equations above. The M_i, estimated by simple regression (2), provide convenient estimates of the B_i, obtainable alternatively from multiple regression (3). The relationship between the two sets of weights is

$$\begin{aligned}M_1 &= B_1 \\ M_2 &= B_1^2 + B_2 \\ M_3 &= B_1^3 + 2B_1B_2 + B_3 \quad (13)\end{aligned}$$

and generally,

$$M_i = \sum_{j=1}^{i} M_{i-j}\, B_j, \text{ where } M_0 = 1.$$

It is enough to estimate either set of weights to obtain estimates of the other.

The study estimated both the M_i and B_i directly, the former by duplicating Meiselman's procedure and the latter by a method described below. In addition, using (13), it derived either set from the other and compared the direct and indirect estimates of both sets of weights. Since the sets are empirically estimated stochastic variables, their relationships do not correspond exactly with their algebraically derived relationships. The estimates and statistical tests of their equivalence are shown for illustrative purposes only.

To compare the direct and indirect estimates of either set of weights, a procedure is required to directly estimate the B_i. It is, of course, possible to directly estimate (3) for each value of n, the span of forecast. Alternatively, since the B_i, in principle, are identical for each n, it is convenient to pool the data into one regression involving all spans of forecast, each value of the dependent variable associated with the appropriate prior forecasts and actual values. Table 8-1 lists the direct estimates of M_i and B_i together with the indirect estimates of each from the other, as well as estimates of the significance of the differences between the two sets of estimates.

Columns 1 and 2 are reestimations of Meiselman's reported results with the error-learning model (equation 2). Column 3 combines the directly estimated B_i of equation (3) in accordance with equation (13). While the standard error for the direct estimates are reported, it is very difficult to estimate the standard errors of the indirect estimates of M_i since they involve extensive algebraic manipulation of a stochastic series. While inclusion of an estimate of this error would increase the standard error of the differences between the two estimates of M_i, this effect could be offset by a positive covariance between the two estimates of M_i.[6] Since there is no easy way to evaluate the relative strengths of these opposing influences on the standard error of the differences, it is difficult to estimate the direction of the bias of the results reported in column 4. These considerations aside, the two sets of estimates of M_i are fairly close.

The two estimates of B_i are similar, although the problem just alluded to prevents a conclusive evaluation of the significance of the differences in the two sets of estimates. In both sets, B_1 is large and there is a sharp decline to B_2 followed by a gradual decline. Sampling fluctuation in the direct estimates and the sensitivity of the indirect estimates to sampling fluctuation in the estimated M_i prevent a smooth pattern in the B_i, but the one just described is a reasonable approximation.

[6] The well-known formula for the standard error of the difference between two estimates, A and B, is $S_A + S_B - 2C_{A,B}$, where S signifies standard error and C is covariance. A positive covariance lowers the standard error of the difference.

TABLE 8-1. *Direct and Indirect Estimates of the Error-Learning Coefficients, M_i, and the Forecasting Weights, B_i, and Estimates of the Significance of the Differences Between the Direct and Indirect Estimates, Annual Data, 1901–54*

Span of Forecast	Direct Estimate of M_i (1)	Standard Error of Direct Estimate (2)	Indirect Estimate of M_i (3)	t Value of Difference (4)	Direct Estimate of B_i (5)	Standard Error of Direct Estimate (6)	Indirect Estimate of B_i (7)	t Value of Difference (8)
1	.7029	.0312	.7457	1.3718	.7457	.0199	.7029	2.1507
2	.5256	.0419	.6109	2.0286	.0548	.0248	.0318	0.9274
3	.4034	.0466	.5312	2.7446	.0347	.0240	.0114	0.9708
4	.3263	.0486	.3641	0.7798	–.0914	.0192	.0180	–5.6979
5	.2769	.0459	.2864	0.2048	.0522	.0243	.0165	1.4691
6	.2348	.0414	.2546	0.4758	.0051	.0258	.0042	0.0349
7	.2367	.0389	.2477	0.2751	.0412	.0258	.0401	0.0426
8	.2089	.0401	.2209	0.3167	–.0168	.0260	–.0116	–0.2000

NOTE: All estimates are based on the Durand data. Col. 1 duplicates Meiselman's estimates of equation (2); col. 2 lists the standard error of these estimates; col. 3 estimates M_i from B_i in col. 5 using equation (13); col. 4 approximates the significance of difference between cols. 1 and 3 using col. 2 as the estimate of the standard error of difference; col. 5 estimates the coefficients of equation (3) using pooled data for all spans of interest; col. 6 lists standard errors of these estimates; col. 7 estimates B_i from M_i in col. 1 using equation (13); col. 8 approximates the significance of the difference between cols. 5 and 7 using col. 6 as the estimate of the standard error of difference.

THE RETURN-TO-NORMALITY HYPOTHESIS

Since the forecasting equation (3) is formulated as a general autoregressive model, its weights, B_i, tell us a lot about how expectations are formed. We can compare various hypotheses with the empirical estimates. One widely used model hypothesizes a return-to-normality mechanism, whereby forecasts of a series move in the direction of the normal value of the series.[7] This hypothesis is one explanation for the often observed inverse relationship between the slope of a yield curve and the level of rates. Typically, yield curves incline at low levels of rates and decline at high levels.[8] According to the return-to-normality hypothesis, when short-term rates are high they are expected to decline; hence, long-term rates decrease with increasing maturity, and the yield curve declines. The reverse holds for low levels of short-term rates.

Algebraically the return-to-normality hypothesis amounts to the following:

$$ {}_{t+2}F_t - {}_{t+1}F_t = K(A_t - {}_nA_t),\ K < 0, \tag{14}$$

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, or spot, rate at t; ${}_nA_t$ is the normal rate at t; and K is negative to reflect the inverse relationship between the expected change and the deviation of the spot from the normal rate.

Assuming for the moment that the normal rate does not change, the following regression form gives estimates of K in (14),

$$ {}_{t+n}r_t - {}_{t+n-1}r_t = a + KA_t + V,\ n = 1,8. \tag{15}$$

[7] The phrase "move in the direction of" distinguishes the forecasts or expected values from the normal value, which generally connotes a long-term tendency rather than a particular point. It should, therefore, vary less than a forecast and, correspondingly, incorporate a greater span of the past variation of the series, to which it should attach more uniformly distributed weights in place of the decaying weights of a point forecast.

[8] This observation is true except at the very short end of the curves, which almost always inclines. This short-period incline combined with an over-all declining yield curve results in the familiar humped yield curve typical of periods of high interest rates. See Reuben Kessel, *The Cyclical Variation of the Term Structure of Interest Rates*, New York, NBER, 1964, p. 25; reprinted as Chapter 6 of this book.

TABLE 8-2. *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	K (1)	t Value of K (2)	Constant Term (3)	t Value of Constant Term (4)	R^2(adj) (5)
${}_{t+1}r_t - R_t$	−.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 form ${}_{t+n}r_t - {}_{t+n-1}r_t = Q + k_n R_t + V_n$.

where V is a random term.[9] The results of estimating (15) are listed in Table 8-2. These estimates of K are in each case significantly negative, confirming the widely recognized relationship described above. The use of A_t in place of $(A_t - {}_nA_t)$ in (15) implies the following relation:

$$a = a' + K_n A_t,$$

where a is the constant term in (15) and a' is the constant term in the event $(A_t - {}_nA_t)$ is used in place of A_t. On the hypothesis that $a' = 0$, an estimate of the constant normal rate is obtained from the ratio a/K,

[9] Other writers have used the level of rates to estimate the normality effect or else have assigned some arbitrary value or small group of values to the normal value. See, for example, James Van Horne, "Interest Rate Risk and the Term Structure of Interest Rates," *Journal of Political Economy,* 1965, p. 349. Van Horne adds a variable he calls "deviation of actual from accustomed level" to Meiselman's formulation of the error-learning model and he divides his sample period into two subperiods. "For each . . . [sub] period . . . an arithmetic average of the beginning forward rate levels is calculated. This average may be thought to represent the accustomed level for the [sub] period. The deviation is simply the difference of the actual forward rate level from the accustomed level. . . ."

that is, the constant term divided by the regression coefficient. For each n, the estimate is approximately 4 per cent. This number is an estimate of the normal rate.

Just as the decline in the error-learning model coefficients, M_i, was equivalent to a particular pattern of weights in the extrapolative equation (3), so the negative K in (14), interpreted here as indicating an expected return to normality, also implies a particular pattern of weights in (3).[10]

To derive this pattern, define ${}_nA_t$ of (14) as

$$ {}_nA_t = B_2A_{t-1} + B_3A_{t-2} + \cdots + B_nA_{t-n-1}. \tag{16} $$

Substituting (16) into (14) yields

$$ {}_{t+2}F_t - {}_{t+1}F_t = KA_t - K\left(\sum_{i=1}^{n-1} B_{i+1}A_{t-i}\right), \tag{17} $$

where ${}_{t+1}F_t$ and ${}_{t+2}F_t$ were defined in equations (4) and (7) above, respectively, the first in terms of the A_{t-1} and the second in terms of ${}_{t+1}F_t$ and the A_{t-i}. Substituting (4) into (7) yields

$$ {}_{t+2}F_t = (B_1{}^2 + B_2)A_t + (B_1B_2 + B_3)A_{t-1} + \sum_{i=3}^{n} (B_1B_i + B_{i+1})A_{t-i-1}. \tag{18} $$

Subtracting (4) from (18) gives

$$ {}_{t+2}F_t - {}_{t+1}F_t = [(B_1{}^2 + B_2) - B_1]A_t + \sum_{i=2}^{n} [(B_1B_i + B_{i+1}) - B_i]A_{t-i-1}. \tag{19} $$

From (13) it is clear that the expression $[(B_1{}^2 + B_2) - B_1]$ in (19), equal to K in (17), is actually $M_2 - M_1$. In general, the proportion, K, of the deviation of the current rate from the normal rate, which is expected to be offset between $t + n - 1$ and $t + n$, is exactly equal to $M_N - M_{N-1}$. The more rapid the decline in M_i the sooner are future rates expected to overtake the normal rate. The difference in the M_i listed in column 1 of Table 8-1 show the rate of movement toward normality for the particular data used. According to these estimates

[10] Mincer, *op. cit.*, designates the pattern that produces a declining M_i and and a negative K as convex. He distinguishes this pattern, which is consistent with many different combinations of weights, from concave and exponential patterns.

approximately half the difference between current and normal rates is expected to be removed over the eight year period.

The rate of movement declines with span of forecast. Other convex patterns of B_i would imply different rates of movement toward normality. In the case of exponential weights, where $M_i = M_{i+1}$, there would be no movement toward normality, and in the case of a concave pattern of B_i, where $M_i < M_{i+1}$, expected rates would move away from normality.[11]

Another implication of the convex pattern of B_i, where $M_i > M_{i+1}$ (and actually an alternate exposition of the relations described above), is that as the span of forecast, n, increases, the weight attached to A_t declines, and the remaining weights both rise and approach equality. In other words, the longer the span of forecast, the lower the weight given to current experience and the greater the weight given to the past. In effect, the longest term forecast approaches the normal rate.

THE RELATIONSHIP AMONG EXTRAPOLATIVE MODELS

The principal conclusion of the above analysis is that three widely used forecasting models, the extrapolative, the error-learning, and the return-to-normality models, are actually three variants of a general extrapolative formula. There are, in principle, as many models as there are combinations of weights from an autoregression, although the word "model" is ordinarily used only when the particular combination of weights is consistent with a plausible behavioral hypothesis. There is a difference, however, 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 to a given set of data need not result in declining revision coefficients (as the span of forecast increases) and therefore in a particular pattern of implied extrapolative weights. Similarly, while the return-to-normality hypothesis is consistent with the extrapolative model, there is no logical necessity that K be negative. The model is a transformation of the extrapolative models, while the hypothesis that K is negative is subject to empirical test. A negative K is implied by error-learning coefficients that decline with forecast span.

[11] See Mincer, *op. cit.*

AUTONOMOUS FORECASTING

There is more to forecasts, of course, than is implied in the extrapolative procedure. Variation in contemporaneous variables may also influence forecasts. Merely correlating these variables with the forecasts, however, would not reveal the extent of this influence. To the extent these contemporaneously correlated variables are themselves autoregressive, they impart an autoregressive component to the forecasts. For this reason part of the observed autoregressiveness of the forecasts may arise from the influence on them of variables that are themselves autoregressive. Whether there exists an autoregressive component of the forecasts independent of that which is imposed by the autoregressive component of functionally related variables is virtually impossible to determine.

It is possible, however, by partitioning the forecasts into autoregressive and random components, to divide the total relation between the forecasts and the other variable into the parts due to either component. One way to effect this partition is to regress the forward rates on the current and past spot rates and interpret the residuals of the regression as estimates of the random component of the forecasts.[12] An observed relationship between these residuals and other current economic variables would indicate that part of the forecast was based on current developments in the market, not entirely of the past.

The regression form used to distinguish the two components of the forecast is

$$_{t+n}F_t = a + b_1A_t + b_2A_{t-1} + \cdots + b_iA_{t-7} + {}_{t+n}E_t, \tag{20}$$

where $_{t+n}F_t$ is the forecast made at t referring to $t + n$, A_{t-i} is the spot lagged i periods, and $_{t+n}E_t$ is the residual term (the estimate of the random component of the forecasts). The lag terms are arbi-

[12] Another way is to compute a moving average of the current and past spot rate using the weights described above. In principle, a third way is to specify a model that predicts the autonomous (with respect to time, not to other variables) component and leaves a residual estimate of the autoregressive component. In the absence of a definitive method, it is most practical to exhaust one component and let the stated existence of the other component depend on rejecting the null hypothesis of its absence. The likelihood that forecasters utilized nonextrapolative information, given the rejection of the null hypothesis that they did not, is strengthened by the knowledge that part of the stated autoregressive component likely includes the autoregressive effects of contemporaneously related variables.

trarily limited to seven to conserve data. The computed value from the regression, ${}_{t+n}F_t{}^*$, is the estimate of the extrapolative component of the forecasts and the residual term, ${}_{t+n}E_t$, of the random or non-extrapolative component.

Equation (20) was fit for spans of one, two, three, and four quarters to Treasury bill yields.[13] The coefficients of determination are listed in column 2 of Table 8-3 and show a close relationship. In spite of

TABLE 8-3. *Partitioned Relationship Between the Index of Industrial Production and the Forward Rates, Treasury Data, Quarterly Observations, 1949–64*

Span of Forecast	$R^2_{F,I}$ (1)	R^2_{F,F^*} (2)	$R^2_{F,F^*,I}$ (3)	t Value of b_{F^*} (4)	R^2_{F,I,F^*} (5)	t Value of b_I (6)	$R^2_{F,F^*,I}$ (7)
1	.7159	.9543	.8866	21.8365	.0078	.6909	.9667
2	.7806	.9314	.8075	15.9938	.1049	2.6743	.9564
3	.8107	.9177	.7343	12.9808	.1947	3.8404	.9480
4	.8176	.8814	.4705	7.3615	.2394	4.3817	.9003

NOTE: R^2 is the coefficient of determination, I is the index of industrial production, F is the forward rate, and F^* is the extrapolative component.

this strong measure of autoregressiveness the question remains whether there is any relationship between the residuals and other variables that may influence forecasting of interest rates. For present purposes we use the FRB's index of industrial production as a proxy for such other influences.

Column 1 of Table 8-3 lists the coefficients of determination between the forward rates and the production index. Equation (21) measures the effect on the forecast of the extrapolative component (computed from equation 20), ${}_{t+n}F_t$, the concurrent production index, I, and the random component.

$$_{t+n}F_t = a + b_1({}_{t+n}F_t{}^*) + b_2I + u_n. \tag{21}$$

Columns 3 through 7 list the relevant statistics for these regressions.

Most of the correlation between the forward rates and the index of industrial production is captured by the extrapolative component

[13] The data are read from the yield curve for the middle month of each quarter. The yield curves appear in the *Treasury Bulletin*. The sample period used is 1946–64.

of the forecasts. That is, R^2_{F,F^*} includes most of $R^2_{F,I}$. But there is a net relation between F and I that is independent of F^*. Column 5 of Table 8-3 indicates that the correlation between I and the autonomous component of the forecasts increases with the span of forecast and, except for the first span, is statistically significant. Alternatively, when the relation between the forecasts and current and past spot rates is adjusted for the influence of I, the coefficients of determination (column 3) are smaller than the simple coefficients, listed in column 2. This result highlights the difficulty noted earlier of interpreting the estimated amount of autoregressiveness; the autoregressiveness of I contributes to that of F.

The forecasts of interest rates appear to rely progressively more on autonomous variables, like I, as the span of forecast increases. Consequently, the ability to explain the forecasts does not decline with increasing span of forecast as rapidly as the declining R^2s of equation (20) suggest.

The index of industrial production is merely one of many possible indicators likely to affect the forecasts of interest rates. The observed relationship does not imply that investors actually consulted this particular indicator. The determination of which indicators were actually consulted is a statistical question only insofar as alternate hypotheses are tested. The Dow Jones index of industrial stock prices (denoted by S), for example, yields somewhat stronger results than the index of industrial production. The results of this experiment are shown in Table 8-4. As in the case of the earlier experiment the relation between F and S grows with increasing span. Without the contribution of S the extrapolative component of F deteriorates much more rapidly than when its

TABLE 8-4. *The Partitioned Relationship Between the Dow Jones Index of Industrial Stock Prices and the Forward Rates, Treasury Data, Quarterly Observations, 1949–64*

Span of Forecast	$R^2_{F,S}$ (1)	R^2_{F,F^*} (2)	$R^2_{F,F^*.S}$ (3)	t Value of b_{F^*} (4)	$R^2_{F,S.F^*}$ (5)	t Value of b_S (6)	$R^2_{F,F^*,S}$ (7)
1	.7861	.9543	.8532	18.8278	.0327	1.4359	.9675
2	.8595	.9314	.7215	12.5709	.1704	3.5399	.9596
3	.8892	.9177	.5921	9.4116	.2770	4.8339	.9533
4	.8889	.8814	.2356	4.3359	.3309	5.4918	.9123

NOTE: The form of this table is identical to that of Table 8-3 except for the substitution of industrial stock prices, S, for industrial production, I.

contribution is not isolated (column 3 compared with column 2). But the strengthening relationship shown in column 1 as span of forecast increases is due not only to the effect of S on F^* but also to its autonomous effect, which grows to 33 per cent of the forecast by the fourth span. There is little question but that a more elaborate attempt to specify a model of interest rate forecasting would succeed in reducing further the putative effect of extrapolative forecasting, which in the preceding results appears dominant.

THE ACCURACY OF THE FORECASTS

To measure the accuracy of the forecasts, it is useful to separate the bias from the random error of the forecasts.[14] The distinction is particularly important in the case of the forward rates since they may contain a nonforecasting component that on the average makes the forward rate, when viewed entirely as a forecast, too high. Many writers[15] think that long-term rates are on the average higher than short-term rates because holders require a premium to compensate for the lower liquidity of long-term bonds. If so, and if this nonforecasting component were not isolated, the forward rates, which include the premia, would appear to be less accurate forecasts than they are. However, only the mean of the premia would contribute to the measured bias; any variation in the premia, as a result of variation in their determinants, would accentuate the computed error of forecast.

It is useful to compare the mean square error of the forecasts with a set of benchmark forecasts as well. A convenient benchmark forecast is found by autoregressing the target value A_t on its own past values, which is the general form of the various naive models that are often used in this connection. In the general form,

$$A_t = \sum_{i=1}^{n} b_i A_{t-i} + E. \tag{22}$$

The so-called "no change" and "same change" models can be derived by setting b_1 and $b_i = 0$ for $i > 1$ in the former case and $b_1 = 2$, $b_2 = -1$, and $b_i = 0$ for $i > 2$ in the latter case. Since in accordance

[14] See Mincer and Zarnowitz, "The Evaluation of Economic Forecasts," *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* Jacob Mincer (ed.), New York, NBER, 1969.

[15] For example, see Reuben Kessel's article, Chapter 6 of this book.

with (22) the degree of fit of the regression is the measure of accuracy, the benchmark forecasts are the most stringent to use, since they are specifically chosen to maximize the fit.

The total mean square error and the random error[16] of the forecasts, as a ratio to the benchmark, are given in Table 8-5. The ratios in column 1 include the bias and are therefore higher than the corresponding ratios in column 2. While for each span the total mean square error is higher for the forecasts than for the benchmark, the former approach the latter as the forecast span increases. The relative improvement of the forecasts as the span increases is indicated in colunm 2 as well. These numbers are lower than the corresponding numbers in column 1 because the bias term is removed. The relative improvement with span lowers the random error of the forecasts below that of the benchmark by the fourth span.

The improved relative accuracy of the forecasts as span increases

TABLE 8-5. *Ratios of Mean Square Errors and of Their Random Component of Forecasts Relative to Benchmark Forecasts, Treasury Data, Quarterly Observations, 1949–64*

Span of Forecasts	Ratio of Mean Square Errors (1)	Ratio of Random Errors (2)
1	1.1481	1.0854
2	1.4078	1.0701
3	1.2716	1.0341
4	1.0607	0.9607

NOTE: Column 1 gives the ratio of $E(A-F)$ for the forecasts relative to the same term for the autoregressive benchmark. A number greater than 1.0 implies the forecasts have a higher total error than the benchmark forecasts have.

Column 2 gives the ratio of $(1 - r_{AF}^2)\ S_A^2$ for the two sets of forecasts. This term measures the random error, as distinct from the bias. The symbol r_{AF}^2 refers to the coefficient of determination in the regression of the target values, A, on the forecasts, F; S_A^2 is the variance of A. Since the autoregressive benchmark is computed to exclude a bias, the denominators of the ratios in columns 1 and 2 are identical; for the benchmark the total and random errors are the same.

[16] The random error is the component of the mean square error whose expected value is zero; that is, it excludes any bias. A formula for the random error is

$$\text{random error} = (1 - r_{AF}{}^2)\ S_A{}^2$$

where $r_{AF}{}^2$ is the coefficient of determination in the regression of F on A, and $S_A{}^2$ is the variance of A. See Jacob Mincer, *op. cit.*

reveals the increasing importance of the autonomous component of the forecasts with increasing span. Since the gross correlation between A and F is a measure of the accuracy of the forecasts, the decomposition of this correlation into partials for the extrapolative and autonomous components of the forecasts helps show the sources of the forecasts' accuracy and the changes in these sources with increasing span of forecast. The partial correlations for equation (23), shown in Table 8-6, reveal the increasing importance of the autonomous component with increasing span of forecast.

$$A_{t+n} = b_1({}_{t+n}F_t^*) + b_2({}_{t+n}E_t) + {}_{t+n}U_t. \tag{23}$$

TABLE 8-6. *Selected Statistics From the Regression of Target Spot Rates on the Extrapolative and Autonomous Components of the Forecasts, Treasury Data, Quarterly Observations, 1949–64*

Span of Forecast	Partial Correlation Coefficient Squared of ${}_{t+n}F_t^*$ (1)	t-Value of $b_{{}_{t+n}F_t^*}$ (2)	Partial Correlation Coefficient Squared of ${}_{t+n}E_t$ (3)	t-Value of $b_{{}_{t+n}E_t}$ (4)	R^2 (multiple) (5)
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 given in equation (23). The forecast components were related to a four-term moving average of the quarterly spot rates to make the forecasts and the targets comparable.

SUMMARY AND CONCLUSIONS

The evidence revealed in this study is consistent with the hypothesis that expectations influence the term structure of rates. While the importance of extrapolation in business forecasting is well known, statistical verification of this fact tends to exaggerate the importance of extrapolation. Even where the market or the individual utilizes knowledge of contemporaneous relationships, the autoregressiveness of these related variables redounds on the forecasts themselves. In spite of this exaggeration the data reveal some amount of nonextrapolative forecasting, which contributes to the accuracy of the forecasts.

It is convenient to summarize the extrapolative component of the forecasts, regardless of its source, with an equation that describes each forecast as a linear combination of past forecasts and observed values (that is, targets of earlier forecasts). The convenience stems from the inferences that can be drawn from the pattern of weights in the linear combination. The most obvious inferences concern the relative importance that attaches to earlier forecasts and observed values in determining current forecasts. This study found, for example, that the weight given to the last previous forecast (or observed value, in the case of the one-span forecast) is relatively high but that subsequent weights are much lower, although the rate of decline is small after the decline from the first to the second. Further inferences stem from the parameters of other extrapolative behavioral models that are implied by the weights in the linear combination. So, for example, the error-learning coefficients—the proportions of past errors that are used to correct forecasts of later periods—fall off with increasing span of forecast. This decline does not imply (but is consistent with) a decline in the extrapolative component of longer span forecasts and less still a lower forecasting content of more distant forward rates. On the contrary, the decline in the learning coefficients is implied by the widely observed inverse relationship that exists between the direction of expected change and the difference between the current level and its long-run expectation. By algebraically relating the three models (extrapolative, error-learning, and return-to-normality), the study was able to estimate their parameters. While the three models imply different motivations for the forecasts, they are mutually consistent and statistically indistinguishable.

Apart from analyzing the extrapolative component, the study estimated its importance as the span of forecast changes. It found that the autonomous (or nonextrapolative) component both explained a larger fraction of the variation of longer-span forecasts and accounted for a larger fraction of their accuracy. The second finding, concerning accuracy, is important because it excludes the inference that the falling-off of the extrapolative component signifies a falling-off of the forecasting component of the forward rates. It is consistent also with the idea that autonomous forecasting is more useful for longer-term forecasts—or the obverse, that extrapolative forecasting becomes less useful with increasing span of forecast.

The expectational mechanism described in this report, even if it were the only factor governing yield differentials of varying terms to maturity, does not ensure that as of any given time the yield curve depicts the market's forecasts. A sudden increase in the supply of a particular maturity or a decision by one or more large financial institutions to alter the maturity of their portfolio would alter the shape of the yield curve. But if the market's forecasts have not changed, relative bargains among certain maturities would emerge, and the demand for these maturities would reinstate the original yield curve. Investors may differ in their preference for different maturities for reasons such as liquidity or hedging liabilities and will respond with different degrees of alacrity to the temporary departure of the yield curve from the one reflecting current expectations. Transaction costs may discourage or delay total reinstatement.

This study did not investigate the influence of liquidity preference. Short-term securities are generally thought to be more liquid because of their broader markets, lower transaction costs, greater collateral value, and less volatile price fluctuation. Not all investors, however, value these qualities, and some prefer long-term bonds because they stabilize income, obviate reinvestment, require less management, etc. Most writers think the former group predominates in the market and thus that, apart from the influence of expectations, short-term securities yield less because of their (on balance) greater desirability. The resulting yield curve would incline unless the expectation of a decline in rates dominated the liquidity effect. According to the liquidity hypothesis, there is a spectrum of degrees of liquidity, inversely related to term-to-maturity. A change in the structure of available maturities, as a result, for example, of a major government refinancing, would alter the aggregate liquidity of financial assets. If, for example, the debt were lengthened and as a result over-all liquidity fell, the ensuing demand for liquidity would lower the yields (bid up the prices) on shorter-term securities and alter the term structure that is consistent with a given pattern of expectations. This hypothesis, which includes the effect of changes in the supply of money, is supplementary rather than competitive with the expectations hypothesis. A more exhaustive theory would include the effects of changes in the yield differentials of debt and equity, changes in expected inflation, and many other phenomena as well. These ideas have in common the proposition that the market for securities and indeed for all assets may be usefully viewed as a unit in which different investment concerns—yield, liquidity, risk, taxes, etc.—are synthesized into an over-all price structure.