THE JOURNAL OF FINANCE • VOL. XLV, NO. 2 • JUNE 1990

The Structure of Spot Rates and Immunization

EDWIN J. ELTON, MARTIN J. GRUBER, and RONI MICHAELY*

ABSTRACT

Empirical studies of the modern theories of bond pricing typically choose proxies for the state variables in a rather arbitrary fashion. This paper empirically analyzes the question of the optimal spot rates to use as state variables. Our findings indicate that the four-year spot rate serves as the best proxy in the one-state-variable model. In the case of the two-state-variables model, the six-year rate and eight-month rate are identified as best. Tests of the out-of-sample prediction ability indicate that our model is superior to Macaulay's duration model and alternative proxies for state variables.

MODERN THEORIES OF BOND pricing hypothesize that the value of a default-free bond is a function of a small number of state variables that follow a diffusion process. These include the one-state models of Cox, Ingersoll, and Ross (1985), Brennan and Schwartz (1977), and Vasicek (1977) and the two-state models of Cox, Ingersoll, and Ross (1985), Richard (1978), Brennan and Schwartz (1979, 1980), and Nelson and Schaefer (1983).

When these models are approximated empirically, a change in one or more spot rates is almost always used as a proxy for the state variables. The one-month rate is sometimes chosen. (See Brennan and Schwartz (1977), Babbel (1983), and Nelson and Schaefer (1983).) The one-year rate is often a choice, and, finally, some researchers use a short and long rate (Brennan and Schwartz (1980) or Nelson and Schaefer (1983)).

The assumption behind the use of a small set of rates to describe bond prices or changes in bond prices (rate of return) is that all spot rates can be modeled as functions of this set. This is true for most bond pricing models and for all of the immunization literature from the simplest model of Macaulay (1938) to the more complex two-index models which have appeared in recent literature. Empirical tests of these sophisticated theoretical models have not demonstrated a superiority to much simpler models (see Ingersoll (1983), Nelson and Schaefer (1983), and Brennan and Schwartz (1983)).

One possible explanation is that no research has been done on the structure of spot rates. Rather, in almost all empirical research, the structure that was utilized and the proxy for the state variable (key rate) were specified on an ad hoc basis. The purpose of this paper is to analyze empirically the structure of spot rates. The structure is of interest in its own right. In addition, it is our hope that understanding this structure will allow more accurate models of bond pricing, immunization, and bond portfolio management to be developed.

* Stern School of Business, New York University. Elton and Gruber are Nomura Professors of Finance. We would like to thank Chris Blake, Ernest Bloch, Linda Canina, William Greene, and Bruce Tuckman for their helpful comments.

The remainder of the paper is organized as follows: Two alternative models describing the expected shift in the yield curve are presented in Section I. Section II describes the sample. In Section III we determine the optimal proxy for the state variable when the one-state-variable model is assumed. The analysis necessary for the two-state-variables model directly parallels that employed for the one-state-variable model and is presented in Section IV. In Section V we evaluate whether the proxies for the state variables outperform the commonly used proxies. It is shown that using the method developed here to select the optimal proxies to the state variables outperforms the proxies utilized elsewhere. Section VI concludes the paper.

I. Anticipated and Unanticipated Bond Returns

Bond returns can be divided into expected (anticipated) and unexpected (unanticipated) returns. A world with only anticipated returns implies that all bonds have the same rate of return over any time period but does not imply a world with a stationary yield curve. For example, if the expectations theory held and spot rates were different from each other, the yield curve would change over time in a deterministic manner. In such a world, pricing models would be irrelevant since all bonds and portfolios of bonds would have the same rate of return.

Of course, actual bond returns also include unanticipated components. These unanticipated components can be systematic (affect multiple bonds) or unsystematic (unique to a bond). For noncallable government bonds, the unanticipated systematic components arise from an unanticipated change in the shape or position of the yield curve. In this paper we will be concerned with unanticipated systematic returns. For equilibrium modeling and immunization, this is the element of return that is important. For active bond management models, this is the element of return that influences risk estimation. Finally, the systematic unanticipated return is an important component of models for generalized bond pricing.

To measure unanticipated shifts, we need to model what the expected yield curve should be one period hence. We employ two alternative models in this paper. In the first we assume that the yield curve is expected to remain unchanged. All changes in spot rates are assumed to be unanticipated. This model is the simplest model of expectations we can utilize and is the one which has been employed in most previous empirical work (see Nelson and Schaefer (1983) and Babbel (1983)). It is also consistent with some recent empirical work of Mankiw and Miron (1986) demonstrating that short-term interest rates follow a random walk. Finally, it is consistent with a liquidity preference or a preferred habitat theory of the term structure of interest rates.

As a second model of the anticipated term structure, we assume the expectation theory holds exactly. (Babbel (1983) also assumes this.) Thus, in each period t, we derived from the yield curve in t the forward rates for period $t + 1$.[1] These forward rates were assumed to be the expected spot rates in $t + 1$. In all future sections, we use these two different models of unanticipated returns: the change

[1] When we did not have monthly intervals, we interpolated.

in spot rates and the difference between the spot rate and the forward rate calculated in the prior period.

II. Sample

Our sample consists of McCulloch's estimates of spot rates for a series of maturities over the 30-year period, 1957–1986. McCulloch's estimates are derived in a consistent manner from yield curves over a long time span and have become an accepted standard for empirical work. For each month in the sample period, the data consist of spot rates for maturities of 1 to 18 months on a monthly interval, for maturities of 18 to 24 months on a quarterly basis, and on a yearly basis from 2 to 13 years.[2] We divided our data into six five-year subsamples for two reasons. This enables us to test the models in a forecast mode (have hold-out sample periods) and to replicate the results over several periods in order to study the robustness of our results.

III. Single-Factor Models

Empirical tests of one-state-variable models of bond pricing employ the unexpected change in a single spot rate as the systematic factor affecting the unexpected change in all spot rates. In some models, the unexpected change in the interest rate is used as a proxy for an unobserved factor affecting innovations in the term structure; in others, it is used as a proxy for the unobserved state variable.

The implications of such a model can easily be seen. Define the price of a bond as the discounted value of a series of cash flows C_i's at the appropriate spot rates r_{oi}'s. Assume that the unexpected change in all spot rates depends on the change in one factor F. Taking the derivative of the bond price with respect to the change in F, dividing by P, and rearranging yield an expression for the unanticipated rate of price change:

$$dP/P = -1/P \sum_{1}^{N} \frac{iC_i}{(1 + r_{oi})^i} \frac{dF}{(1 + r_{oi})} \frac{\partial r_{oi}}{\partial F}, \tag{1}$$

where dP/P is the return due to an unanticipated change in the yield curve.

Consistent with the literature, e.g., Babbel (1983) and Brennan and Schwartz (1983), we assume that the change in the factor can be proxied by the unanticipated change in a spot rate. However, in contrast to prior literature, we present and utilize a methodology for determining the best single-factor proxy.

A. Best One-Factor Proxy

In this section we first set forth the methodology for determining the best proxy. Then we present some empirical evidence on how well the proxy performs.

A.1. Determining the Best Factor Proxy

Assume that the unanticipated change in spot rates is linearly related to some

[2] For some years there are estimates of spot rates for maturities beyond those we used.

unknown factor; then,

$$dr_{it} = a_i + b_i dF_t + \varepsilon_{it}, \tag{2}$$

where dr_{it} is the unanticipated change in the ith spot rate at time t, and df_t is the unanticipated change in the unknown factor at time t.

The derivation of the optimal proxy for any maturity is contained in Appendix A. The derivation assumes stationarity of the factor structure and minimum squared forecast error as the criteria of choice between models. In Appendix A we show that under these assumptions the factors' proxies can be ranked by the variance of the spot times the R^2 (square of the correlation coefficient).

While this technique allows us to rank the proxies for each maturity, the problem of combining or weighting the performance across maturities still remains. We examined two alternative weighting schemes. The first weighting technique simply assumes that it is equally important to forecast unexpected changes at each yearly interval. The second weighting scheme employs a set of weights that might typically be used to reproduce the rate of price change on a portfolio of bonds. The effect of this weighting scheme (called cash flow weights) is to place more importance on forecasting accurately the unanticipated change in spot rates of intermediate maturities.[3]

While it is impossible to test the forecasting ability of every technique for all possible uses, if one technique performs well both for individual rates and for the rates on a portfolio of bonds, one can have some faith in the robustness of the technique.

A.2. Empirical Evidence

Two time periods (1957–1961 and 1972–1976) were employed to select the best key rate. The entries in the body of Table I are the product of R^2 and the variance of the dependent spot rate averaged across all maturities using the models of unexpected change described earlier. Initial investigation showed that a key rate equal to the 4-year spot rate was the optimum if one rate was selected over both time periods and both weighting schemes. In Table I we compare the 4-year choice (which is used throughout the paper) with three other rates: the rate that works best for the period, model, and weighting scheme under investigation; a 1-year rate; and a 13-year rate. The latter two benchmarks were chosen because they were the key rates employed by Nelson and Schaefer (1983) and Babbel (1983) in the case of the one-year rate. Notice in Table I that in every case the key rate of four years significantly outperforms the 1- and 13-year benchmarks and that the differences in performance between this rate and the best are quite small even when the differences in maturity are large. This leads us to use the four-year-rate as the key rate in all further sections.

[3] Computing the weight on the forecast error for the portfolio of bonds involved estimating the coefficient of the term involving $(\partial r_{0i}/\partial F)$ from equation (1). This is obviously a function of coupons and spot rates. Simulations were run assuming constant yearly coupons and spot rates varying between 8% and 12%. The relative importance of any $(\partial r_{0i}/\partial F)$ was relatively stable over the different scenarios, and we utilized the average weights. These weights were 0.58, 0.79, 0.92, 0.98, 0.98, 1.0, 0.96, 0.89, 0.79, 0.71, 0.61, 0.51, and 0.40 for years 1 through 13, respectively. When we had multiple forecasts within a year, we divided the weight equally among them.

Table I

Determination of the Optimal Factor's Proxy for the One-State-Variable Model

Entries in the table are values of the modeling objective function which we define as the product of R^2 and the dependent spot rate's variance, averaged across all maturities, $\sum w_i R_{i,x}^2 \text{var}(dr_i)$. Two models are examined. Model 1 assumes that the innovation in the spot rate follows a random walk. (All changes are perceived as unexpected.) Model 2 derives the innovation in the spot rate under the pure expectation theory. (The unexpected change is the difference between the actual spot rate and the forward rate from the previous period.) Two weighting schemes are employed. In the first, equal weights are given to each yearly interval. For example, since for the first yearly interval there is information about 12 different spot rates, each of these spot rates receives a weight of (1/12)*(1/13). In the cash-flow-weighting method, weights are given according to the relative importance of the spot rate in a portfolio of 13 bonds, each maturing in a different year. Results are reported for four factor proxies: the one with the largest value for each period and model, the 4-year spot rate which is the overall optimal proxy, and the 1-year and 13-year spot rates which have been employed in previous studies.

	Model 1				Model 2			
	Largest Period Value	Maturity Employed (4 Yrs)	1 Yr	13 Yrs	Largest Period Value	Maturity Employed (4 Yrs)	1 Yr	13 Yrs
Panel A: 1957–1961								
Equal Weights	0.612 (4)[a]	0.612	0.512	0.531	0.550 (3)	0.549	0.462	0.491
Cash-Flow Weights	0.503 (4)	0.503	0.405	0.432	0.449 (4)	0.449	0.361	0.398
Panel B: 1972–1976								
Equal Weights	1.029 (2)	0.994	0.926	0.613	0.932 (1.75)	0.910	0.867	0.518
Cash-Flow Weights	0.827 (5)	0.823	0.712	0.478	0.746 (5)	0.744	0.657	0.398

[a] The number in parentheses is the maturity with the maximum value.

B. Pattern of Sensitivities

In Table II we present the average intercept (a_i) and slope (b_i) of the regression of the unexpected change in different spot rates against the unexpected change in the 4-year spot rate. This is done for both models of unexpected change, and their results are quite similar. Note first that the b's not only differ from one (the assumption in the Macaulay measure of duration), but they have a distinct pattern. They start below one, rise above one, and fall to well below one for long rates. While we report in this table only the average results over all six of our 5-year data periods, examination of each 5-year period shows a very similar pattern.

In most previous empirical work where the independent variable was selected to be the shortest rate in the sample, the sensitivities decline uniformly with the maturity of the spot rate. While it is more convenient to have such a pattern for developing duration measures, the cost, as seen in the previous section, is to have less explanatory power. Using the four-year spot as the independent variable increases explanatory power but creates a more complex pattern for sensitivities.[4]

There also appears to be a discernible but less distinct pattern in the intercepts which implies a nonzero mean in unexpected changes in spot rates. If either of our two models of estimating the unexpected change in spots were perfectly correct, all intercepts should be equal to zero. We can think of four reasons that can account for finding nonzero intercepts. The first is the error in variables problems. To the extent that the four-year rate is an imperfect proxy for the factor driving interest rates and that deviations of the four-year rate from the factor are random, the errors in the independent variable will cause the intercept to be positive and the slope to be lower. The second is the problem of omitted variables. If the process driving interest rates is really a two-factor model (and we present evidence later in the paper that it is), one would expect a pattern on the intercept similar to that found. The third reason is the presence of liquidity premiums. While this could impact both models, its impact is easiest to see (and should be larger) for model 2. The presence of the liquidity premium would mean that the changes in the expected spot rates we derived are overestimates of the true expectations. Liquidity premiums change more rapidly for shorter maturities than for long maturities and would result in the pattern of intercepts we found. The fourth reason applies to model 2. It is easy to show that an underestimate of the true thirty-day spot rate would cause the intercept to be negative when the dependent variable is a short-term spot rate and positive when it is a long-term spot rate.

We have not tested the statistical significance of the intercepts and slopes, for we believe that the importance of the presence of an intercept different from zero and of a slope different from one relates not to whether these differences

[4] The reason for the more complex pattern is the following. The b_i for any maturity spot is the product of the correlation between that spot and the four-year rate and the standard deviation of the spot divided by the standard deviation of the four-year rate. As we move to higher maturities, the standard deviation of the spot rate uniformly decreases. On the other hand, as we move from very low maturities to longer maturities, the correlation with the four-year rate increases (reaching a maximum when the maturity is four years) and then decreases. It is this pattern for the two elements affecting b_i which accounts for sensitivities at first increasing and then decreasing as we increase maturities.

Table II

Estimated Sensitivities for the One-Factor Model:

$$d(r_{it}) = a_i + b_i d(r_{4t}) + e_{it}$$

The factor proxy is the 4 year spot rate. The time period is 1957–1986, using monthly observations. Model 1 assumes that the innovation in the spot rate follows a random walk. (All changes are perceived as unexpected.) Model 2 derives the innovation in the spot rate under the pure expectation theory. (The unexpected change is the difference between the actual spot rate and the forward rate from the previous period.)

Time to Maturity (in yrs)	Model 1 a_i	Model 1 b_i	Model 2 a_i	Model 2 b_i
0.16	−0.0043	0.798	−0.2193	0.886
0.5	−0.0064	0.878	−0.1137	1.071
1	−0.0081	1.084	−0.0637	1.134
1.5	−0.0074	1.117	−0.0366	1.161
2	−0.0059	1.125	−0.0211	1.138
3	−0.0023	1.072	−0.0590	1.055
4	0.0000	1.000	0.0000	1.000
5	0.0023	0.930	0.0046	0.942
6	0.0042	0.863	0.0081	0.886
7	0.0058	0.803	0.0180	0.836
8	0.0072	0.753	0.0128	0.790
9	0.0084	0.710	0.0142	0.748
10	0.0093	0.671	0.0151	0.707
11	0.0100	0.636	0.0157	0.669
12	0.0105	0.604	0.0164	0.619
13	0.0109	0.573	0.0181	0.571

are statistically significant but to recognize whether these differences lead to better estimation (conditional forecasts).

IV. Parameterization of Two-Variable Models

The analysis necessary for the two-variable model directly parallels that employed for the one-variable model. Table III is directly analogous to Table I. The best pair of rates over the two time periods, weighting schemes, and models of unexpected return is the six-year rate and the difference between the six-year rate and the eight-month rate (referred to as the selcted key rates).[5] This model was compared against the two rates that work best for each individual time period weighting scheme and expectational model and against the two-rate model of Nelson and Schaefer (thirteen years and the difference between the thirteen-year and five-year rates). The selected key rates outperformed the Nelson and Schaefer two-rate model in all time periods and under the two weighting schemes employed. Like the case of the one-index model, the two-parameter model using the selected key rates works almost as well as the two-parameter model which is optimal for each time period, expectational model, and weighting scheme. We use the selected key rates as the basis for further analysis of the two-index model.

[5] We chose the difference between the long and short rates, rather than the short rate, as the second key rate in order to minimize the problem of multicollinearity.

Table III

Determination of the Optimal Factor's Proxies for the Two-State-Variables Model

Entries in the table are values of the modeling objective function which we define as the product of R^2 and the dependent spot rate's variance, averaged across all maturities, $\sum w_i R_{i,x}^2 \operatorname{var}(dr_i)$. Two models are examined. Model 1 assumes that the innovation in the spot rate follows a random walk. (All changes are perceived as unexpected.) Model 2 derives the innovation in the spot rate under the pure expectation theory. (The unexpected change is the difference between the actual spot rate and the forward rate from the previous period.) Two weighting schemes are employed. In the first, equal weights are given to each yearly interval. For example, since for the first yearly interval there is information about 12 different spot rates, each of these spot rates receives a weight of (1/12)*(1/13). In the cash-flow-weighting method, weights are given according to the relative importance of the spot rate in a portfolio of 13 bonds, each maturing in a different year. Results are reported for three factor proxies: the one with the largest value for each period and model, the 6-year spot rate and the difference between the 6-year spot rate and the 8-month spot rate which are the overall optimal proxies, and the 13-year spot rate and the difference between the 13-year and the 5-year rates which have been frequently employed in previous studies.

	Model 1			Model 2		
	Largest Period Value	Maturities Employed[a] (6y, 8m)	Nelson & Schaefer Maturities	Largest Period Value	Maturities Employed (6y, 8m)	Nelson & Schaefer Maturities
Panel A: 1957–1961						
Equal weights	0.653 (7y, 8m)[b]	0.652	0.597	0.600 (6y, 6m)	0.598	0.539
Cash-Flow Weights	0.529 (11y, 7y)	0.509	0.487	0.465 (12y, 8m)	0.464	0.439
Panel B: 1972–1976						
Equal Weights	1.141 (6y, 6m)	1.137	0.722	1.056 (7y, 14m)	1.055	0.656
Cash-Flow Weights	0.880 (6y, 6m)	0.879	0.580	0.811 (7y, 16m)	0.806	0.516
Panel C: The Two Periods Combined						
Cash-Flow Weights	1.389 (6y, 8m)	1.389	1.067	1.271 (7y, 14m)	1.270	0.995

[a] The maturity employed is the innovation in the 6-year spot rate and the difference between the 6-year and the 8-month rates. Nelson and Schaefer maturities are the 13-year spot rate and the difference between the 13-year and the 5-year rates.

[b] The number in parentheses is the maturity with the maximum value. For example, (6y, 8m) is the 6-year innovation in the spot rate and the 6-year innovation in the spot rate minus the innovation in the 8-month rate.

When different maturity spots were regressed on the selected key rates, there was a clear pattern to the regression coefficients. While these are modeled later in the paper, we should note now that the slope coefficients on both rates are not monotonic with respect to the maturity of the dependent spot rate.

V. Selecting the Best Forecasting Model

In this section we evaluate whether the methods developed in the paper lead to better conditional forecasts of the unexpected change in spot rates than do commonly used benchmarks.

A. Alternative Forecast Methods

The analysis to this point has been concerned with determining the best one- and two-index methods to use in forecasting unexpected changes in spot rates. The one-index model (denoted by opt 1 in Table IV) bases forecasts on a univariate regression where the independent variable is the unexpected change in the four-year spot rate. The two-index model (denoted by opt 2) bases forecasts on a bivariate regression where the independent variables are the unexpected change in the six-year rate and the difference between the unexpected change in the six-year rate and the eight-month rate. For each of the models, sensitivities were estimated using data for the five years prior to the forecast period.

This way of estimating sensitivities is perfectly feasible as a way of getting estimates for bond pricing. However, it does not result in a closed-form specification. For many purposes, a closed-form model is desired. Hence, we constructed and tested two such models. The data analysis from previous sections suggests a form for sensitivities that is quadratic in the natural logarithm of time. Thus, we construct regressions of the following form:

$$b_i = c_0 + c_1 \ln(i) + c_2 (\ln(i))^2 + \varepsilon_i, \tag{3}$$

where, b_i is the sensitivity of a spot rate of maturity i to an unexpected change in the factor.

This model performed well in fitting the cross-sectional pattern of sensitivities (b's). The average R^2 is higher than .97 for the regression of sensitivities on time in the one-index method and higher than .93 and .99 for the first and second sensitivities in the two-index model. Forecasts from these models are labeled opt 1 s for the one factor smoothed and opt 2 s for the two factors smoothed. While the motivation for this specification is to allow the derivation of a closed-form bond pricing model, its smoothing properties might result in better forecasts than using the historical sensitivities directly.

We have a number of benchmarks with which to compare our forecasts. The natural benchmark is the assumption inherent in Macaulay's duration measure. This measure simply assumes that the sensitivity of all rates to a change in any rate (including the four-year rate) is one. In Table IV these forecasts are labeled as naive. Our other benchmarks were the models used in Nelson and Schaefer (1983) and Babbel (1983). Both Babbel and Nelson and Schaefer use the one-year rate in their single-index models, and Nelson and Schaefer use the thirteen-

Table IV

The Out-of-Sample Performance of Competing Models

Model 1 assumes that the innovation in the spot rate follows a random walk. (All changes are perceived as unexpected.) Model 2 derives the innovation in the spot rate under the pure expectation theory. (The unexpected change is the difference between the actual spot rate and the forward rate from the previous period.) For each model the coefficients are estimated in the prior period. The Mean Square Error is then calculated as a basis of comparison. The models are ranked from best to worst for each subperiod (columns 2–5), and for the overall sample as well (first column).† "Opt 1" is the optimal one-factor proxy, the 4-year rate. "Opt 2" is the optimal two-factor proxies, the 6-year rate, and the difference between the 6-year and the 8-month rates. "Opt 1 s" is the smoothed one-factor proxy, estimated by equation (3). "Opt 2 s" is the smoothed two-factor proxies, estimated by equation (3). "Naive" assumes that the sensitivities of all rates to a change in any rate are one. "1 yr" is the model in which the 1-year spot rate is used as the factor proxy. "13 yrs" is the model in which the 13-year spot rate is used as the factor proxy. "13, 5 yr" is a two-state-variables model in which the 13-year spot rate and the difference between the 13-year and the 5-year rates are used as the factor proxies.

Panel A: Model 1

1 Overall	2 Period 1	3 Period 2	4 Period 3	5 Period 4
opt 2	opt 2]	opt 2	opt 2	opt 2
13, 5 yr	opt 2 s]	opt 2 s	13, 5 yr	13, 5 yr
opt 1	13, 5 yr	13, 5 yr	opt 1	naive
opt 1 s	opt 1 s]	opt 1]	opt 1 s	opt 1 s
naive	opt 1]	opt 1 s]	1 yr	opt 1
opt 2 s	13 yrs]	naive]	opt 2 s*	opt 2 s
1 yr	naive]	1 yr]	naive*	1 yr
13 yrs	1 yr	13 yrs	13 yrs	13 yrs

Panel B: Model 2

1 Overall	2 Period 1	3 Period 2	4 Period 3	5 Period 4
opt 2	opt 2]	opt 2	opt 2	opt 2
opt 1	opt 2 s]	opt 2 s	opt 1	naive
opt 2 s*	naive	opt 1	opt 1 s	opt 2 s
opt 1 s*	opt 1 s	opt 1 s	opt 2 s*	opt 1
naive	opt 1	naive	naive*	opt 1 s

* Indicates that the order of these two models is reversed for the cash flow weighting.

] Indicates a difference that is insignificant at the 5% level.

† The subperiods are 1962–1966, 1967–1971, 1977–1981, and 1982–1982. The two subperiods (1957–1961 and 1972–1976) that were used to determine the optimal factors' proxies are excluded since they do not represent true out-of-sample periods.

year rate and the difference between the thirteen-year and five-year rates in their two-index models. We adopt these rates as additional benchmarks in Table IV, calling them 1 yr, 13 yr, and 13, 5 yr. We use them only for tests of expectational model 1, for this is the framework in which they were proposed by Nelson and Schaefer.

B. Tests

Each of the techniques discussed above is used to prepare forecasts for spot rates in each of four five-year periods. Parameters of all models are estimated over the previous five-year period. The two five-year periods 1957–1961 and 1972–1977 were excluded from the forecasting sample period because they would not constitute a true hold-out sample since they were used to select the indexes that worked best.[6]

The methodology used to compare and rank forecast techniques involves tests of differences in mean square forecast errors. More specifically, for each forecasting technique, squared forecast errors were computed for each maturity. This was repeated each month in the sample period, resulting in a matrix of 60 (months) by 31 (maturities) for each forecasting technique. For forecasting methods taken two at a time, differences between all pairwise entries were calculated along with a mean standard deviation. In computing the mean and standard deviation, we used the same two weighting schemes discussed earlier. From the central limit theorem, the differences in forecasts from any two techniques should be normally distributed, and we tested the significance of the differences.

We compare four sample periods as well as combining the four sample periods into one, using 240 months. Table IV shows the results when each maturity is weighted equally. The results when they are weighted by their importance in impacting the change in the bonds portfolio price are identical, except for those cases that are shown by an asterisk. All differences are significant at the 5% level except where there are brackets indicating insignificance.

C. Empirical Results

From Table IV we can see some definite differences in performance of the methods. The first striking result from Table IV is that, for the overall sample and in each subperiod, the optimum single-and two-factor proxies we developed outperformed the proxies utilized by others, including Macaulay's duration model. This would clearly hold in the period where we search for the optimum. However, the results shown in Table IV are for different periods. This consistent outperformance is evidence that the ranking of preferred proxies for the unknown factor is sufficiently stable from period to period, so that the procedure discussed in earlier sections leads to better results. The second striking characteristic of Table IV is the dominance of the two-factor models over their one-factor counterparts. This dominance holds in each period and for both definitions of unexpected change in spot rates. Elton, Gruber, and Nabar (1988) also found strong evidence of two factors affecting returns.

The third result shown in Table IV is that the smoothed sensitivities underperformed the sensitivities from the prior period. However, the smoothed sensitivities outperformed the naive or Macaulay model. Thus, for those who insist on a closed-form solution, there seems to be one that dominates Macaulay's.

[6] For the periods 1962–1966 and 1972–1981, we tried alternative variations of *opt 1* and *opt 2*. We explored the forecasting ability when an intercept was used or left out. We also tested forecasting ability when sensitivities and intercepts were calculated for the three 5-year periods and then averaged for forecasting purposes. These variations were not statistically different from one another, and their ranking varied considerably when we went across the two periods. Using an intercept and just one prior 5-year period to calculate values performed no worse than any other and was simpler.

However, for most purposes, calculating duration and storing a vector of sensitivities are not inconvenient.

The last result that needs to be highlighted is the overall performance of the naive or Macaulay measure. Previous literature has shown that this measure has not been easy to beat. However, in our tests, both the one-variable and two-variable optimum models always outperform Macaulay's measure. Moreover, for expectational model 2, it was the poorest performer, and, for expectational model 1, it outperformed the smoothed two-variable method and two benchmark one-variable models.[7]

In summary, our technique for choosing the optimal proxy worked well with both one- and two-variables models. Overall, the two-variables model outperforms the one-variable model, and all outperform benchmark models.

VI. Conclusion

Empirical work on bond pricing needs to assume something about the relationship between unexpected spot rate changes of different maturities. In the most common approach, unexpected changes in all spot rates are described by changes in one or more key spot rates. While previous studies have chosen these factors in an ad hoc manner, this paper shows how to select the factors and structure which best describe the correlations across unexpected changes in spot rates.

In fact, our findings indicate that the four-year rate serves as the best proxy in the one-state-variable model, while the eight-month and six-year rates do best in the two-state-variables model. Also, we find that the two-factor model outperforms the one-factor model. Our model produces out-of-sample predictions superior to those using Macaulay's duration model and those using other state variable proxies. These results may explain why models used in the area of immunization, which assume complex relationships between changes in spot rates, have not, in general, outperformed the simplest immunization model put forth by Macaulay: the inferior results may be due to an inappropriate choice of factors describing the unexpected changes in spot rates.

A deeper understanding of the structure of unexpected changes in spot rates is of fundamental importance as an aid to understanding risk in the bond markets. Thus, our results may be used in the pricing of debt instruments, the pricing of options on debt instruments, the construction of bond portfolios, and performance evaluation of bond portfolios.

[7] We performed one other set of tests. We added the errors at each point in time across the maturities, weighting each maturity equally or weighting according to their importance in price sensitivity. This allows cancellation of errors. If errors are random, then allowing overestimates for the maturity to cancel with underestimates for second maturity will produce a better measure of the sensitivity of price to an unexpected change in interest rates. However, there is a second pattern. In this pattern the sign is very much a function of maturity: short maturities having positive residuals and long maturities having negative residuals, or vice versa. Hence, this technique is very sensitive to the choice of the longest maturity. A bond with a maturity different from 13 years would perform very poorly because the errors would not tend to cancel out. Thus, the results of Table IV are more reflective of true relative performance.

Appendix A

Assume that the model determining the unanticipated change in a spot rate is a linear model with an unknown factor. Let the subscript x represent the optimal spot rate to choose as a proxy for the unknown factor, and let subscript i represent any factor. Thus,

$$dr_{i,t} = a_i dr_{x,t} + \varepsilon_{i,t}, \tag{A1}$$

where $\varepsilon_{i,t}$ is a normally distributed random variable with mean zero and variance σ^2, and

$$R_{i,x}^2 = 1 - \frac{\mathrm{var}(\varepsilon_{i,t})}{\mathrm{var}(dr_{i,t})} \tag{A2}$$

or

$$R^2 \mathrm{var}(dr_{i,t}) = \mathrm{var}(dr_{i,t}) - \mathrm{var}(\varepsilon_{i,t}). \tag{A3}$$

Minimizing the residuals' variance is equivalent to maximizing the right-hand side of equation (A3). Hence, in order to find the best proxy for the state variable, we maximize the weighted average of $R_{i,x}^2 \mathrm{var}(dr_{i,t})$ over the choice of the appropriate proxy, x:

$$\max \sum w_i R_{i,x}^2 (\mathrm{var}(dr_i)). \tag{A4}$$

REFERENCES

Babbel, D., 1983, Duration and the term structure of interest rate volatility, in G. Bierwag, G. Kaufman, and A. Toevs, eds.: *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (JAI Press, Greenwich, CT).

Brennan, M. and E. Schwartz, 1977. Savings bonds, retractable bonds, and callable bonds, *Journal of Financial Economics* 5, 67–88.

——— and E. Schwartz, 1979, Continuous time approach to the pricing of bonds, *Journal of Banking and Finance* 3, 133–155.

——— and E. Schwartz, 1980, Conditional predictions of bond prices and returns, *Journal of Finance* 35, 405–419.

——— and E. Schwartz, 1983, Duration, bond pricing, and portfolio management, in G. Bierwag, G. Kaufman, and A. Toevs, eds.: *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (JAI Press, Greenwich, CT).

Cox, J., J. Ingersoll, and S. Ross, 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–408.

Elton, E., M. Gruber, and P. Nabar, 1988, Bond returns, immunization and the return generating process, *Studies in Banking and Finance* 5, 125–154.

Ingersoll, J., 1983, Is immunization feasible, in G. Bierwag, G. Kaufman, and A. Toevs, eds.: *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (JAI Press, Greenwich, CT).

Macaulay, F. R., 1938, *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856* (Columbia University Press, New York).

Mankiw, N. G. and J. A. Miron, 1986, The changing behavior of the term structure of interest rates, *Quarterly Journal of Economics*, 211–228.

McCulloch, J. H., 1971, Measuring the term structure of interest rates, *Journal of Business* 44, 19–31.

———, 1975, The tax adjusted yield curve, *Journal of Finance* 30, 811–830.

Nelson, J. and S. Schaefer, 1983, The dynamics of the term structure and alternative portfolio immunization strategies, in G. Bierwag, G. Kaufman, and A. Toevs, eds.: *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (JAI Press, Greenwich, CT).

Richard, S., 1978, An arbitrage model of the term structure of interest rates, *Journal of Financial Economics* 6, 33–37.

Vasicek, O., 1977, An equilibrium characterization of the term structure, *Journal of Financial Economics* 5, 177–188.