THE JOURNAL OF FINANCE • VOL. L, NO. 4 • SEPTEMBER 1995

Fundamental Economic Variables, Expected Returns, and Bond Fund Performance

EDWIN J. ELTON, MARTIN J. GRUBER, and CHRISTOPHER R. BLAKE*

ABSTRACT

In this article, we develop relative pricing (APT) models that are successful in explaining expected returns in the bond market. We utilize indexes as well as unanticipated changes in economic variables as factors driving security returns. An innovation in this article is the measurement of the economic factors as changes in forecasts. The return indexes are the most important variables in explaining the *time series* of returns. However, the addition of the economic variables leads to a large improvement in the explanation of the cross-section of *expected* returns. We utilize our relative pricing models to examine the performance of bond funds.

ONE OF THE MAJOR research topics in the recent literature of financial economics is the estimation and testing of relative pricing models. The purpose of this study is to develop and test relative pricing models for bonds utilizing a variety of different types of bonds. This study is important for several reasons. First, despite the economic importance of bond markets, very little attention (compared to that paid to common stocks) has been paid to relative bond pricing models. Second, recent developments in the methodology of the testing of relative pricing models (based on the Arbitrage Pricing Theory (APT) of Ross (1976)) have yet to be applied to the pricing of bonds. Third, this study is important because it uses fundamental economic variables as well as return indexes to explain both returns and expected returns on bonds. While the inclusion of fundamental variables along with return indexes has been used to explain relative prices for common stocks (see Chen, Roll, and Ross (1986) and Burmeister and McElroy (1988)), the importance of the variables has not been demonstrated for bonds. Fourth, this study is important because it is the first study to employ forecasts prepared by economists and investment professionals to measure unexpected changes in the fundamental economic influences that affect returns. Finally, this study is important because it is the first time relative pricing models parameterized on passive portfolios have been used to examine the performance of actively managed bond portfolios. This will both give us additional evidence on the performance of the models on a separate sample of returns and insight into the contribution of active management.

* Elton and Gruber are from the Leonard N. Stern School of Business, New York University; Blake is from the Graduate School of Business Administration, Fordham University. We would like to thank Stephen Brown, Bent Christensen, Jacob Boudoukh, René Stulz (the editor), an anonymous referee, and participants at the 1994 European Finance Association meetings (Brussels) for helpful comments.

Bonds are an important asset category, having an aggregate market value many times that of equities. In 1992, the Federal Reserve estimated that U.S. bond markets had an aggregate value of $9.1 trillion for publicly traded issues. This was almost twice the aggregate value of the stock market. Adding non-publicly traded issues would make this difference even larger.

There have been tests of equilibrium models based on bond data. However, the tests have either been restricted to government bonds or have utilized a very limited set of bond portfolios in conjunction with a much larger set of stock portfolios. The principal type of test of equilibrium models derived for the bond market involves estimating the parameters of a variant of the Cox, Ingersoll, and Ross (CIR) (1981) model (see Brown and Dybvig (1986), Pearson and Sun (1994), Longstaff and Schwartz (1992), and Gibbons and Ramaswamy (1993)). Tests of variants of the CIR model only utilize Treasury securities and involve estimating properties of the term structure of Treasury securities. Parameters to capture risk differentials of nonTreasury instruments have not been included in published models to date. The structure and testing of the CIR model are very different from the structure and types of tests that have been performed on relative pricing models for common stocks.[1] Tests of relative pricing models using stock return data have been two-pass (time-series and cross-sectional) tests popularized by Fama and MacBeth (1973) (e.g., Fama and French (1992)) or joint-estimation (constrained time-series) tests (e.g., Gibbons (1982), Stambaugh (1982), Burmeister and McElroy (1987 and 1988), Gibbons *et al.* (1989), Snow (1991), and Bansal and Viswanathan (1993)).

To the extent that these types of tests have been applied to bond data, the tests have been primarily limited to Treasury bonds. The reasons for this are the dearth of bond return data on non-Treasury instruments and the inability to use time-series data to estimate sensitivities because of the changing risk characteristics of most bonds as maturity changes. To get around the second difficulty, most researchers have used return data on pure-discount Treasury instruments that allow both maturity and duration to remain constant over time. Relative pricing models have been tested using data on Treasury bills with short maturities (see Fama and Bliss (1987) and Stambaugh (1988)). Similarly, Elton, Gruber, and Mei (1993) have tested a relative pricing model using McCulloch (1990) estimates of the prices of pure-discount instruments for longer maturities. Finally, the few researchers who have not limited their analyses to Treasury instruments have utilized a few bond portfolios in conjunction with a much larger set of stock portfolios (Fama and French (1992) and Gibbons (1986)).

In this study we fit a relative pricing model for bonds using test methodology developed in the common stock area and employing a large set of bond portfolios. Our basic data are the returns on a set of indices (portfolios) of Treasury,

[1] In this discussion we differentiate between relative pricing models and equilibrium models. Relative pricing models are not equilibrium models. The absence of arbitrage is a necessary condition for equilibrium but only guarantees that the assets studied have relatives prices (not absolute prices) consistent with equilibrium.

corporate, and mortgage securities compiled by Lehman Brothers, Merrill Lynch, and Ibbotson and Associates.[2] These indexes are calculated in a manner exactly parallel to the way the Center for Research in Security Prices (CRSP) stock indexes are calculated. The treatment of interest paid during a month is identical to the way dividends paid are treated by CRSP, and the prices used are the prices an investor would pay for purchasing a debt security (quoted price plus accrued interest).

Our technique for estimating a relative pricing model is constrained time-series regression. This technique has been popularized in a series of articles by Burmeister and McElroy and others. Its main advantage is that betas and relative prices are simultaneously estimated so that estimated parameters of the relative pricing model are not affected by estimation error in the estimated sensitivities as in the two-pass (Fama and MacBeth) procedure.

In constructing relative pricing models, we have a choice as to how factors are obtained. In the stock area, three methods have been used. Several authors (e.g., Roll and Ross (1980)) have employed factor analysis to develop underlying variables; several have used portfolios to represent hypothesized underlying influences (e.g., Fama and French (1992)); and several have used unexpected changes in hypothesized economic variables to represent the factors (e.g., Chen, Roll, and Ross (1986)).[3] The advantage of using the last approach is that it relates returns to fundamental influences in the economy rather than to other returns (which are themselves driven by economic influences). In this paper we not only employ a combination of returns on a set of portfolios that have been shown to affect bond returns, but we also employ unexpected changes in some economic variables to see if they add additional explanatory power. In our study we examine the ability of our model not only to explain the time series of returns, but also cross sectional differences in average returns. Because we examine cross sectional differences, we can determine whether the influences we hypothesize are priced by the market.

Two of the three methods that were used in the stock area to develop the factors driving returns have been employed in studying fixed-income returns. Gutelkin and Rogalski (1985), Litterman and Scheinkman (1991), and Knez, Litterman and Scheinkman (1994) use factor analysis to identify a set of indices to explain the time series behavior of the return on debt instruments. In addition, Litterman and Scheinkman form mimicking portfolios for their

[2] The Lehman Brothers and Merrill Lynch indices are value-weighted with the returns calculated using actively traded bonds. For example, Lehman Brothers and Merrill Lynch have a number of bonds that traders actively price on a daily basis. All other bonds are priced by formulas utilizing prices on the active bonds. The net effect is to calculate return using the actively traded bonds but weighting each bond's return by the proportion that bond type represents of the market. Returns are calculated from prices that take account of (include) accrued interest and are adjusted for coupon payments. The Ibbotson data were obtained from the Center for Research in Security Prices (CRSP) SBBI files.

[3] A discussion of the relative strengths and weaknesses of each is beyond the scope of this manuscript. The reader is referred to Elton and Gruber (1995) or Peazy (1994) for a detailed discussion of the appropriateness of each approach.

factors and examine how well linear combinations of these mimicking portfolios seem to be associated with a set of hypothesized economic influences.[4] The second approach is used by Blake, Elton, and Gruber (1993), who employ portfolios of debt instruments to explain the behavior of mutual funds over time. None of these studies (or, to the best of our knowledge, other studies) examines whether adding unexpected changes in economic variables to the types of factors (portfolios) described above produces a model that has greater ability to explain the time series behavior of bond returns. None of these articles examines the ability of the models tested to explain relative prices in an APT sense (the cross-sectional behavior of returns on alternative bonds), and thus none can supply information on whether any of the influences in the model are priced. That is, none of these studies tells us whether the investor requires a higher (or lower) return for taking on some of the systematic variability associated with a factor. While there are papers that accomplish these goals in the stock area, this is the first article to do so in the bond area.

Like many current APT models, our models use unanticipated changes in economic variables to explain, in part, expected return. However, unlike others who estimate unanticipated changes in economic variables from time-series analysis, we utilize publicly available forecasts to measure unexpected changes in expectations. Changes in expectations have been shown to be the driving force in security prices (see Elton, Gruber, and Gultekin (1981)). Using expectational data directly has the dual advantage of focusing on the variable that directly affects returns and allowing us to easily define "unexpected changes."

After developing and parameterizing relative-pricing models on bond indexes, we apply the models with these parameters to analyze the performance of a sample of bond mutual funds with data over a different time period. We find that the model with out-of-sample parameters is useful in evaluating bond mutual fund behavior. Bond mutual funds have realized returns below our models' estimates of expected returns. This underperformance is about equal to transaction costs. These results provide additional support for our models, as they are consistent with previously published results that did not involve equilibrium models (see Blake, Elton, and Gruber (1993)). In addition, we find that the discrepancies in performance among different styles of managers disappear when we correctly account for economic variables.

The article is divided into four sections. In the first section, we discuss our general methodology, the sample, and the variables we use. In the second section, we present the alternative forms of the models we test, and present the estimation and tests of the models. In the third section we apply the models

[4] None of these articles examined changes in economic influences or portfolios that mimic economic influences to explain the return on debt instruments. In addition, our analysis also differs from the studies cited above in the data analyzed. Gultekin and Rogalski (1985) and Litterman and Scheinkman (1991) only examine government bonds and thus cannot detect factors related to issuer risk. Knez, Litterman, and Scheinkman (1994) only look at money market instruments and thus cannot detect influences (including types of credit risk) that are present in longer maturity instruments.

developed in the second section to the examination of the performance of bond mutual funds. The last section contains our conclusions.

I. Model Specifications

In this section of the article we will develop a plausible set of models to see how well they explain both the time series of returns and expected returns. The financial economics literature gives us some guidance as to the set of variables that might enter into these models and explain expected returns. Previous studies (e.g., Chen, Roll, and Ross (1986) and McElroy, Burmeister, and Wall (1986)) have shown that both the time series of individual returns and the cross section of expected returns used in common-stock studies are explained by:

1. market returns (the excess return on the stock market)
2. default risk (the difference in return between corporate bonds and government bonds)
3. term risk (the difference in returns between long-term and short-term government bonds)
4. unexpected changes in inflation
5. unexpected changes in a measure of economic performance.

It has been shown that these variables are significant in explaining expected and actual equity returns. Therefore, they should be important in explaining bond returns. The first variable, the excess return (net of the riskless rate) on the stock market, can be viewed as a measure of expectations about general economic conditions. Our second variable is a measure of default risk. Default risk should affect corporate bond returns. Several authors (e.g., Brennan and Schwartz (1983) and Nelson and Schaefer (1983)) have used a form of our third variable, a term premium, in constructing models of interest rates and bond prices. Our fourth and fifth variables are measures of macro-economic influences. Changes in both of these variables should affect the level of interest rates (and hence bond returns) and could affect the certainty of cash flows on different types of bonds (e.g., corporate bonds and mortgage debt). Evidence to support this is found in Ederington and Lee (1993), who show that similar variables impact Treasury bill futures prices and, hence, short-term interest rates.[5] We examine two additional variables. The first is simply an index of aggregate bond returns. If one were looking for the single factor that best explains individual bond returns, the single variable that would probably do the best job is a market index of bond returns. This is analogous to a market index for stock returns being the single index that best explains the performance of individual stocks. The aggregate index should also be included

[5] Evidence is also presented by Chen, Roll, and Ross (1986) and Burmeister *et al.* in Peazy (1994), who find that unexpected changes in these variables impact common stock returns. The unexpected changes do so because, as Chen *et al.* (1986) point out, they impact both future cash flows and discount rates. They should have a similar but less marked effect on discount rates and the uncertainty of realized cash flows for non-government bonds.

because it serves as the benchmark against which to measure the importance of other indices. We also add one more index, a measure of the return on mortgage securities relative to the return on government bonds. This index is added because options are an important element in bond returns. Mortgages are chosen to represent this influence because of the importance of the option effects on mortgage returns. In the remainder of this section, we discuss a more exact formulation of our variables, the sample to be used in our empirical work, and our statistical methodology.

A. Measuring The Variables

One of the innovations of this study is the use of survey data on expectations to measure unanticipated changes in expectations. Previous studies such as Chen, Roll, and Ross (1986) and Burmeister, Wall, and Hamilton (1986), have used changes in realizations of economic variables or differences between realizations and historical extrapolation of past data to measure unexpected changes. In another context, we have shown that expectations determine (are captured in) prices and that changes in expectations represent unexpected influences. Treating changes as unexpected is consistent with a rational expectations view of economic decision-making and is consistent with a large body of empirical evidence.[6] We use changes in expectations about an economic variable as a measure of unexpected changes in that variable. The use of survey data has the advantage of focusing directly on expectations rather than depending on an unspecified link between the measure used and expectations as justification for the measure.[7]

Our models of expected return utilize two fundamental expectational variables. The first fundamental expectational variable is a measure of unanticipated changes in inflation. The variable is based on monthly data provided by the Survey Research Center of the University of Michigan. The data are obtained by surveying consumers on inflation over the next year. We use the coincident change in the forecast as our first fundamental expectational variable.

The other fundamental expectational variable we utilize is a measure of general economic conditions. This variable is the unexpected change in the forecast of real Gross National Product (GNP) (nominal GNP with inflation removed). The variable is derived from monthly forecasts of nominal GNP

[6] See Elton *et al.* (1981).

[7] We utilize expectational inflation data from the Survey Research Center (SRC) of the University of Michigan and expectational GNP growth data from Eggert Enterprises. The SRC data have been extensively analyzed. Rich (1989) and Baghestani (1992) both analyze the SRC inflation series and conclude that the estimates are rational expectations. Unlike the SRC data, there is not the same lengthy time series for estimates from Eggert Enterprises, so these estimates have not been analyzed for rationality; however, given that the Eggert data are derived from surveys of well-known professional economists, while SRC estimates are derived from surveys of individual consumers, and that economists have been trained in ideas of rationality, the Eggert data are likely to be no worse than SRC estimates with respect to rationality.

growth rates provided by Eggert Enterprises.[8] The Eggert data are based on a survey of professional forecasters from financial intermediaries, companies, and brokerage firms, as well as on output from major econometric models. We regressed the coincident change in this series on our inflation variable and used the residuals with the sample mean added back (to leave the mean of the variable unchanged under the orthogonality transformation) as our second fundamental expectational variable.

Consistent with other studies, we use return series to measure default risk, term risk, and stock market return. The default-risk variable is the difference in return between the Blume/Keim (B/K) high-yield series and the Lehman Brothers (LB) intermediate government index. We chose this government return series because the coefficient in a regression between the returns on the B/K high-yield series and LB intermediate government series was close to one. This ensures that differences in the return between the two series are due to default risk rather than term structure risk. The term-risk variable is the difference in return between the Ibbotson and Associates long-term government bond index and their intermediate bond index. For the stock market variable, we use the return on the S&P 500 index with dividends as reported by Ibbotson and Associates. As mentioned earlier, we also use a mortgage return series to capture option elements in bond returns and an index of aggregate bond returns. The option series is the difference in return between the Lehman Brothers Government National Mortgage Association (GNMA) index and a government bond series with the same duration.[9] The aggregate return series we use is the return on a modified Lehman Brothers aggregate bond index. The modified aggregate index was created by combining the LB aggregate bond index with the Blume/Keim high-yield series to obtain a return series encompassing all bonds in the marketplace. The weights used in developing the index represent relative par value weights at the middle of each year. Monthly weights are approximated by linear interpolation of the annual weights. The modified Lehman Brothers aggregate series is used both to measure the impact of general bond returns and as a proxy for omitted variables.

B. Sample

Throughout most of this study we employ publicly available bond return indices (passive bond portfolios) as the independent variables utilized in fitting the equilibrium models. The Appendix lists the bond indices used. These

[8] We examined a number of other candidates for measuring general economic conditions and inflation. We selected the series that showed the greatest correlation with factors extracted from the bond return series.

[9] We use the return on a mixture of 75 percent of the intermediate and 25 percent of the long-term LB government bond indexes. The mixture we use was determined by equating the duration of the combined two government indices with the duration of the GNMA index and also by regressing the GNMA index on each of the two government bond indices and then finding the weights that made the combination of the regression coefficients equal to one.

include indexes of government bonds, corporate bonds, and mortgages. One of the advantages of using indices is that their durations remain fairly constant over time, unlike durations for individual bonds. We would have liked to use high-yield indices also, but a sufficient number of these indices were not available.[10] The sample period covers February 1980 through December 1992 (155 monthly observations).

We obtained data from three sources (Merrill Lynch, Lehman Brothers, and Ibbotson and Associates). The Ibbotson indices are return indices constructed from the return on a small number of bonds of a particular maturity. The basic price data for these indices comes from Salomon Brothers, and the composition of the indexes changes frequently to maintain an approximately constant maturity. The Merrill, Lynch, and Lehman indexes are comprehensive return indices covering a subset of all bonds that have certain characteristics (e.g., 1- to 3-year governments). Within each category, only the return on actively traded bonds is calculated from market prices. Returns on inactive bonds are calculated in each index as a function of the active bond prices. The net effect is a return series of actively traded bonds, but the portfolio return is computed not by using the weight of each of the actively traded bonds in the category, but rather the weight that bonds with the same characteristics as the actively traded bond represent of the population. In calculating return, both the Merrill and Lehman indices correctly incorporate accrued interest. The Lehman indices assume reinvestment of coupon income in the index at the end of the month. This is the same assumption made in calculating returns in the CRSP monthly stock files. Thus, the return calculations for the Lehman indices are exactly the same as the return calculations for the CRSP size indices that are widely used in testing equity asset pricing models. The Merrill indices assume coupons are reinvested in the index at the time of the disbursement of the interest. These indices are comparable to portfolios constructed from the CRSP daily stock files.

In the next section and the final section we utilize samples of bond funds. The data were supplied by Investment Company Data, Inc. (ICDI) and consist of the full history of monthly returns of all funds that existed as of December 1991. The sample obviously has survivorship bias. The impact of this will be discussed in a later section. In evaluating funds, we utilize data beginning in January 1986. This six-year period involved a tradeoff. We obviously would

[10] We know of only four indices that existed over our sample period (one Merrill Lynch, two Salomon Brothers, and the Blume/Keim index). Furthermore, the Blume/Keim index is identical to one of the Salomon indices over part of our sample period, so that we only have three independent indices. Since we use the return on one of the high-yield indices (the Blume/Keim index) as an independent variable, we are left with only two high-yield indices. Furthermore, the remaining Salomon index suffers from postselection bias over part of our sample period (see Blume and Keim (1987)), leaving us with only the Merrill Lynch index. If high-yield bonds were affected by the same influences as all other bonds, the inclusion of the Merrill Lynch index could improve our estimation. However, our concern is that influences other than those that affect government, corporate, or mortgages might affect high-yield bonds. In that case, the inclusion of a high-yield index would likely hurt the estimation, so we exclude that index.

like a long history, to be more confident in the overall results, and to make comparisons across subperiods. However, the number of bond funds grew dramatically in the 80s. Thus, each year we go back decreases the sample size considerably. The six-year period selected is a compromise between a large sample and a long history.

In the rest of the analysis we use three different samples of the returns on bond portfolios. The first is a sample of 20 bond indices that are used as passive portfolios. The second sample is a group of 36 bond funds stratified into three sub-groups (12 corporates, 12 governments, and 12 mortgages), which are used to examine alternative choices of variables. Our final sample is 123 bond mutual funds that are used to examine the performance of actively traded bond portfolios.

We now examine the general methodology used to estimate relative pricing models.

C. Methodology of Model Specifics

We hypothesize, in the tradition of Chen, Roll, and Ross (1986), that returns are generated by a mixture of tradeable portfolios and fundamental economic factors. In general, we can write the return generating process as

$$r_{it} = \mathrm{E}[r_i] + \sum_{j=1}^{J} \beta_{ij}(R_{jt} - \mathrm{E}[R_j]) + \sum_{k=1}^{K} \gamma_{ik} g_{kt} + \eta_{it} \qquad (1)$$

where

1. r_{it} is the return on asset i at time t ($i = 1, \cdots, N$);
2. R_{jt} is the return on tradeable portfolio j at time t;
3. g_{kt} is the unexpected change in the kth fundamental variable at time t;
4. β_{ij} is the sensitivity of asset i to the innovation of the jth tradeable portfolio;
5. γ_{ik} is the sensitivity of asset i to the innovation of the kth fundamental variable;
6. η_{it} is the time-t return of asset i that is unrelated to either tradeable portfolios or fundamental variables;
7. $\mathrm{E}[\cdot]$ denotes expectation;
8. $\mathrm{E}[g_k] = \mathrm{E}[\eta_i] = 0$.

Note that g_k represents unexpected changes in a fundamental economic variable, and, since the expected value of an unexpected change is by definition zero, $\mathrm{E}[g_k] = 0$.

From the Arbitrage Pricing Theory (APT) of Ross (1976), equation (1) leads to the following expression for the expected return of asset i:[11]

$$\mathrm{E}[r_i] = \lambda_0 + \sum_{j=1}^{J} \beta_{ij}\lambda_j^* + \sum_{k=1}^{K} \gamma_{ik}\lambda_k \tag{2}$$

where

1. λ_0 is the return on the riskless asset (R_F) if it exists. (We will assume it does.)
2. λ_j^* is the market price of sensitivity to the jth tradeable portfolio;
3. λ_k is the market price of sensitivity to the kth fundamental variable.

When variables in the return-generating process are tradeable portfolios, the APT market price of risk associated with such a portfolio is the portfolio's expected return minus λ_0. Thus, $\lambda_j^* = \mathrm{E}[R_j] - \lambda_0$ for $j = 1, \cdots, J$. Substituting this expression into equation (2) and recognizing that $\lambda_0 = R_F$ yields:

$$\mathrm{E}[r_i] = R_F + \sum_{j=1}^{J} \beta_{ij}(\mathrm{E}[R_j] - R_F) + \sum_{k=1}^{K} \gamma_{ik}\lambda_k \tag{3}$$

Substituting equation (3) into equation (1) and allowing R_F to vary over time yields:

$$r_{it} - R_{Ft} = \sum_{j=1}^{J} \beta_{ij}(R_{jt} - R_{Ft}) + \sum_{k=1}^{K} \gamma_{ik}(\lambda_k + g_{kt}) + \eta_{it} \tag{4}$$

Rearranging equation (4) yields:

$$r_{it} - R_{Ft} = \alpha_i + \sum_{j=1}^{J} \beta_{ij}(R_{jt} - R_{Ft}) + \sum_{k=1}^{K} \gamma_{ik}g_{kt} + \eta_{it} \tag{5}$$

where

$$\alpha_i = \sum_{k=1}^{K} \gamma_{ik}\lambda_k \tag{6}$$

[11] While we refer to this model throughout as an APT model, since we prespecify the factors we employ, the model can also be viewed as a linear factor model. In fact, some researchers prefer this terminology. See Connor (1984) and Hansen and Jagannathan (1991) for sufficient conditions for such models to hold.

This structure allows us to test the performance of any APT model.[12] Equation (6) will fail if we have misspecified the return-generating process or the APT does not hold. Thus, our tests are joint tests of the return-generating process and APT.

First, equation (5) can be estimated. Second, it can be estimated with restriction (6) imposed. Imposing (6) and comparing the results with the unconstrained equation (5) allows us to test whether imposing the APT restriction on a multi-factor model results in a statistically significant decrease in explanatory power. The constrained-form equation (4) can be estimated using iterated non-linear seemingly unrelated regressions (ITNLSUR) (see Gallant (1987)).

In the next section of this article we will employ this methodology to test descriptions of relative pricing models for bonds. Before doing so, we will use canonical correlation both to find out more about the structure of bond returns and to examine which of our variables are related to bond returns.

II. Empirical Results

A number of the variables we analyze have been used primarily in common-stock studies. Thus, before performing tests of index and relative pricing models, it is useful to get an idea of how much of the movement in bond returns is picked up by the seven explanatory variables considered. To do this, we formed a stratified sample of mutual funds. From our full sample of 123 bond funds, we select 12 funds that are classified by ICDI as government bond funds, 12 that are classified as corporate bond funds, and 12 that are classified as government mortgage funds. The funds are also selected to span the full spectrum of maturity ranges contained in our sample. For each bond fund we add expenses incurred to the monthly return series. This should approximate the return on a set of passive bond funds.[13]

The matrix of excess returns (the fund returns minus the 30-day T-bill rate from Ibbotson and Associates) on the sample (6 years of monthly returns for

[12] Huberman, Kandel, and Stambaugh (1987) point out that, in tests of a model such as the one under consideration, the market price of risk for a tradeable portfolio is the portfolio's excess expected return (over the riskless rate). We have incorporated this expression for the market price of risk directly in our model, so, to the extent that it does not hold (APT does not hold), equation (6) will be violated even if the fundamental economic variables are correctly priced. In addition, Huberman, Kandel, and Stambaugh show that, if mimicking portfolios can be found for the economic influences, stronger statements can be made about the size of the λs. This is an interesting avenue of research to pursue. However, as they point out in their conclusion, given that the weights in the mimicking portfolios must be estimated, it is an open question whether the additional restrictions result in increased power for the tests of the model.

[13] In a later section we show that after expenses are added back, bond funds have essentially a zero alpha. Thus, they can be used as proxies for passive portfolios. Alternatively, we could have performed the analysis in this section utilizing the passive indices over which we later fit the model. We choose not to do so for two reasons. First, we want to look at the relevancy of variables on a set of data that is different from the data over which we fit the model. Second, we want a larger spectrum of maturities within each bond type classification and more observations on mortgage bonds than we have in our passive-index sample.

each of the 36 funds) was correlated with the set of explanatory variables under study. Canonical correlation is used to examine the relationship between fund returns and the seven variables under study (an aggregate bond return index, an aggregate stock index, indices representing default, term, and option effects, and two observable expectational variables). Canonical correlation determines the linear combination of the seven variables and the linear combination of the 36 funds that are most highly correlated. It then determines the second linear combinations of the seven variables and the 36 funds that are most highly correlated, given that the effect of the first canonical correlate has been removed. The process continues until a number of linear combinations equal to the smaller of the number of funds or the number of variables has been found (in this case, seven). The F-test can be used to determine the number of correlations that are significantly different from zero. Only the first three linear combinations of the seven variables under examination are significantly different from zero at the 5 percent level. They are able to account for 89 percent of the variance in the fund return series. The variable measuring term spread does not appear to play an important role in the first three canonical variates. Consequently, the canonical correlation analysis was repeated, deleting the term-spread variable from the list of seven variables. When the analysis is done with six variables, once again, only three canonical combinations (this time of the six variables) show up as statistically significant at the 5 percent level. In addition, the proportion of the variance explained by the first three variates is unchanged. Because of this, subsequent analysis is restricted to a study of six variables: an aggregate bond index, an index of stock returns, a default index, an option index, a series of unexpected changes in real GNP growth, and a series of unexpected changes in inflation. A possible explanation for the term variable not having explanatory power is that its influence may in fact be captured by two other variables. The aggregate bond index minus the risk-free rate is in part a measure of the term premium. Likewise, it has been shown that the excess return on the S&P index is related to the term premium (see Fama and French (1992), and Chen, Roll, and Ross (1986)). Thus, term premium effects could have been captured by the other two variables.

Having selected the set of variables of interest, we investigate the performance of a model containing the six variables. We also investigate three models containing logical subsets of the six variables. By doing so we can judge the impact of including different types of variables on our ability to explain returns and expected returns. The four models we investigate are:[14]

1. A single-index model based on an aggregate bond index (index-1).

[14] A fifth model is also investigated to explore the reliability of the methodology used to detect differences in models. This model is a zero-factor model that allows us to examine whether our methodology can determine if the means of our twenty-bond index series are different at a statistically significant level.

2. A four-index model employing not only an aggregate bond index, but also an aggregate stock index and two indices representing default and option effects (index-4).
3. A four-index model that, in addition to an aggregate index of bond returns and an aggregate index of stock returns, incorporates unexpected changes in macro-economic measures of inflation and real GNP (fundamental-4).
4. A six-index model that incorporates all the influences contained in the second and third models (fundamental-6).

In all of the models, we treat expectational variables (unexpected changes in macro-economic measures of inflation and economic growth) as observable variables. These are denoted by g_{kt} in equation (5). We treat the return series for aggregate bonds, stocks, default risk, and option risk as proxies for unobservable variables. These are denoted by R_{jt} in equation (5). The first model (index-1), relating returns to an aggregate bond index, is included because it is the simplest model and because it is analogous to the single-index model most often used to examine stock returns. Index-1 is a naive model. A more complex model must, at the very least, outperform the single-index model to be considered for further use. Index-4 contains two aggregate return indices for stocks and bonds, plus two return indices for subsets of bond returns to capture default and option effects. The idea is to see if the inclusion of a stock index plus default and option indices leads to better results. The fundamental-4 and fundamental-6 models incorporate the two fundamental expectational variables.[15] The third model adds these two influences to the return index for bonds and the return index for stocks to see if the fundamental variables improve performance more than adding additional bond indices. The fourth model adds the fundamental indices to the four-variable return model to see if expectational data capture influences not captured by our set of return data. Tests in the remainder of this section are performed on the sample of 20 bond indices (passive portfolios) discussed in section I.*B*.

The time over which we fit the model involves a trade-off between wanting a long-term period to estimate the relative pricing models and losing additional bond portfolios. Both Merrill Lynch and Lehman introduced additional indices over time. We chose to go back to a point in time where we had twenty indices. Incorporating any earlier years would have involved a serious decrease in the number of indices that were available. (Eight of the twenty indices we employ were started in February 1980.) Our time frame for the APT tests is from February 1980 to December 1992. The first test we perform for each of the models is to test whether the cross-sectional (APT) restrictions can be rejected. This is equivalent to asking whether imposing the APT cross-

[15] An alternative way to view these models is as follows: The single-index model can be considered as analogous to a simple duration model with risk differentials between types of bonds ignored. The four-index model, because it contains proxies for wealth, can be considered as a form of Merton's (1973) intertemporal asset pricing model. The models fundamental-4 and fundamental-6 are in the spirit of the APT models hypothesized and tested by Chen, Roll, and Ross (1986) and Burmeister *et al.* in Peazy (1994).

sectional restrictions (6) on equation (5) results in an estimation of the residuals from the model that is inferior at a statistically significant level from that produced by equation (5) with no restrictions. To do this, we employ a test methodology described in Gallant (1987). Our tests utilize the following likelihood-ratio test statistic, recommended by Gallant for maximum-likelihood estimators:

$$L = n(\ln|\tilde{\Sigma}| - \ln|\hat{\Sigma}|) \tag{7}$$

where n is the number of time-series observations (155), $\tilde{\Sigma}$ is the estimate of the variance-covariance matrix of the residual errors from the restricted equation (4), and $\hat{\Sigma}$ is the estimate of the variance-covariance matrix of the residual errors from the unrestricted equation (5). L is asymptotically distributed as chi-squared with q degrees of freedom, where q is the number of parametric restrictions. For small samples, Gallant (1987) recommends the use of the F distribution with degree-of-freedom corrections instead of the chi-squared distribution. The small-sample adjustment is simply to compare L to qF_α, where F_α is the F statistic at significance level α with q numerator degrees of freedom and $nM - p$ denominator degrees of freedom, and where M is the number of equations estimated (20) and p is the number of parameters. If L is greater than qF_α, then the null hypothesis that the restrictions hold is rejected.

When we perform this test, we find that the APT restrictions cannot be rejected at the 5 percent level for any of the four models under study.[16] Since we can not reject the cross-sectional restrictions for any model, it is appropriate to continue to study and to compare the *APT-restricted* form of the four models. Comparisons can be made between several pairings of the four models because some models are nested within others. For example, the index-1 model can be viewed as the index-4 model with the added constraints that the sensitivities of returns to the S&P return, the default index, and the option index are all equal to zero. Similarly, pairwise comparisons can be made between all other combinations of the models except between index-4 and fundamental-4 (because they cannot be nested). In all other cases (shown in Table I), the model that contains fewer variables can be considered a constrained form of the model with more variables, and the constraints (betas on additional indexes equal to zero) can be tested by a variation of the method described in the preceding paragraph. As an illustration, return to the comparison of the models index-1 and index-4. Recall that we are testing these models with expected return equal to the historical average. Thus, the index-4 model contains eighty parameters to estimate while the index-1 model constrains sixty of these parameters to be zero. The null hypothesis is that the sixty parameters are zero and rejection of the null implies that the index-4 model was significantly better at the 5 percent level in explaining returns.

[16] When we examine the zero-factor model, we find that equality of means can be rejected at the one percent level. The maximum-likelihood ratio associated with this hypothesis is 41.95, while the critical value at the one percent level is 37.69.

Table I

Likelihood Ratio Test Statistics

All models are in APT-constrained form (equation (4) in the text), and, for each pair of models shown below, the model that contains fewer variables is a restricted form of the other model. The number in each model's name (e.g., "index-1") refers to the number of variables used in the model. "Index" in a model's name means that the model utilizes only return indices; "fundamental" in a model's name means that the model utilizes both indices and fundamental expectational variables. The dependent sample consists of twenty passive bond indices (see Appendix). The sample period is from February 1980 through December 1992 (155 monthly observations). Null hypothesis: restrictions hold; alternative hypothesis: restrictions do not hold. Small-sample adjustment (see Gallant (1987)): Reject null when $L > qF_\alpha$, where F_α is F statistic at α level of significance with q degrees of freedom in numerator and $nM - p$ degrees of freedom in denominator. $n = 155$ observations, $M = 20$ equations, and $\alpha = 0.05$.

Pairwise Test (Involves APT-Restricted Models Only)						
Models	L[a]	p[b]	q[c]	$qF_{0.05}$	$L - qF_{0.05}$	Reject Null?
Fundamental-4 vs. Fundamental-6	519.36	122	40	55.92	463.43	yes
Index-4 vs. Fundamental-6	68.54	122	42	58.30	10.24	yes
Index-1 vs. Fundamental-6	659.54	122	102	127.14	532.53	yes
Index-1 vs. Fundamental-4	140.30	82	62	81.67	58.64	yes
Index-1 vs. Index-4	591.12	80	60	79.36	511.77	yes

[a] $L = n(\ln|\tilde{\Sigma}| - \ln|\hat{\Sigma}|)$, where $\tilde{\Sigma}$ is the variance-covariance matrix of restricted residuals and $\hat{\Sigma}$ is the variance-covariance matrix of unrestricted residuals.
[b] p = total number of estimated parameters.
[c] q = number of restrictions.

All pairwise tests that could be performed are shown in Table I. The results clearly show that adding return indexes (S&P, default, and option) improves the ability to explain the time-series pattern of returns and expected returns, and that adding the fundamental variables (GNP and inflation) also improves performance. It is clear that both the added return indexes and the fundamental variables are important influences in explaining the time series of returns and expected returns. The importance of unexpected changes in fundamental data in explaining differences in expected returns across different types of bonds has not previously been documented in the financial economics literature. Although Table I indicates that fundamental-6 is the preferred model, we will continue to examine the other models to understand better the reason for its dominance.

Another type of evidence that shows the importance of the fundamental variables is contained in Table II. In this table we present the λs (prices of risk) and statistical significance of the λs associated with each of the two fundamental variables. λ_1 is the price of risk for unexpected changes in real GNP, and λ_2 is the price of risk for unexpected changes in inflation. Note that in both the fundamental-6 model and in the fundamental-4 model, the market price of risk (λ) associated with each fundamental variable is significantly different

Table II

Market Price of Risk Associated with Fundamental Variables

λ_1 is the estimated market price of risk for unexpected changes in real gross national product (GNP); λ_2 is the estimated market price of risk for unexpected changes in inflation. The percentages are expressed in monthly terms. The dependent sample consists of twenty passive bond indices (see Appendix). The sample period is from February 1980 through December 1992 (155 monthly observations). The entries are derived from equation (4) using the returns on twenty passive portfolios.

Model	λ_1	$t_{\lambda 1}$	λ_2	$t_{\lambda 2}$
Fundamental-4	0.262%	2.68	−0.642%	−3.01
Fundamental-6	0.201%	2.64	−0.551%	−3.22

Table III

Adjusted R^2 for Restricted Regressions

Mean, maximum, and minimum adjusted R^2 for the explanatory power of the restricted model (equation (4)) for the twenty passive portfolios. These values indicate the ability of the model to explain the time series of monthly returns for the passive bond portfolios.

Model	Mean	Maximum	Minimum
Index-1	0.904	0.965	0.786
Index-4	0.921	0.979	0.810
Fundamental-4	0.909	0.966	0.788
Fundamental-6	0.923	0.980	0.808

from zero at the 1 percent level.[17] So far the evidence points to the importance of a model that includes fundamental variables. Tables III, IV, and V provide additional evidence on the importance of fundamental variables.

The numbers in Table I are affected by the ability of a model to explain both the time series and cross-section of returns. Tables III and IV attempt to decompose this into explanations of the time series and of the cross-section of returns. Table III presents the average of the time-series explanatory power of each model (averaged across all portfolios). It is clear that most of the adjusted R^2 is associated with the aggregate index of bond performance. While adding return indices for other securities or fundamental variables increases the

[17] The reader should note that, for non-linear models, t-tests on individual parameters are only asymptotically valid. Thus, Table II should be viewed with caution. The results in Table II are consistent with the results in Table I.

Table IV

Percentage of Cross-Sectional Variation Explained by Various Models

The "percentage explained" is calculated by taking 1 minus the sum of the squares of the differences between expected return and estimated expected return from equation (4) divided by the sum of the squares of the differences between expected return and the mean (across all passive portfolios) of expected return.

Model	Percentage Explained
Index-1	40.62%
Index-4	31.21%
Fundamental-4	87.05%
Fundamental-6	82.47%

Table V

Percent of Expected Return Explained by Various Indexes

The entries in this table are calculated by first taking the absolute value of the sensitivity times the risk premium for each index divided by the sum of the absolute values across all indices in the model. This number is then averaged across all passive portfolios. Sensitivities are estimated by time-series regressions of the passive portfolio returns on the indices; risk premiums are the lambdas shown in Table II for the fundamental variables and the average excess returns for the other indexes. "S&P" is the S&P 500 index; "Default" is the difference between the Blume/Keim high-yield bond index and the Lehman Brothers intermediate government bond index; "Option" is the difference between the Lehman Brothers GNMA index and a weighted average of the Lehman Brothers intermediate and long-term government bond indexes; "Aggregate" is a weighted average of the Lehman Brothers aggregate bond index and the Blume/Keim high-yield bond index; "GNP" is the change in expected real Gross National Product (GNP) growth; "Inflation" is the change in expected inflation.

Model	S&P	Default	Option	Aggregate	GNP	Inflation
Index-4	4%	3%	4%	89%		
Fundamental-4	3%			73%	7%	17%
Fundamental-6	3%	2%	3%	73%	5%	14%

ability of a model to explain the *time-series* behavior of the average passive bond portfolio, the increase in explanatory power is small. This should not be surprising. Interest rate changes are the dominant cause of changes in returns on bond portfolios. Thus any bond index will explain a large percentage of the time series of bond returns. However, when we look at the ability of our alternative models to explain differences in the *expected return* of different passive bond portfolios (Table IV), a very different result emerges. Table IV reports the percentage of the total variance (across our sample) of average bond returns that is explained by each model. To compute this statistic, the squared error in the estimate of expected return (average return for each bond fund minus expected return from equation (4) squared), is computed, then summed over all firms. This sum is divided by the sum of squared deviations of the difference between the mean return for each fund and the average of the 20 means. This is the unexplained variation. This number is subtracted from one, and the result is multiplied by 100 to compute the percentage of cross-sectional variation explained by each model. We see from the table that adding the two fundamental variables to either a two-index or four-index return model markedly increases the power of the model to explain the cross-sectional differences in average returns. This evidence indicates that in explaining the *pattern* of returns over time, return indices are of key importance, while in explaining the cross-section of *expected* returns, fundamental variables play a key role. Table IV reinforces the more formal tests shown in Table I. In Table I we show that the index models are rejected when compared to the relative pricing fundamental models as models explaining expected return. Table IV shows that the differences in explaining expected return are large.[18]

The impact of these two effects can be seen in Table V. The amount any variable contributes to explaining mean return can be found for any security by multiplying the sensitivity of that security to that variable times the market price of that sensitivity. In order to aggregate these results across securities, the absolute value of each contribution is taken. To determine the percentage of expected return explained by each variable, we sum the absolute values of the amounts explained by each variable and then divide each absolute value by that sum and multiply by 100. The contribution of each variable is then averaged over all firms in the sample. The results in Table V can be interpreted as the average percent of explained mean return that is contributed by each variable.

The most striking point that can be seen from Table V is the influence of fundamental variables in explaining expected returns. When the aggregate bond index is employed with our three other security indices (no fundamental variables), it accounts, on average, for 89 percent of the estimate of expected return. When the fundamental variables are employed, the fraction of returns explained by fundamental variables is substantial and much greater than that explained by the non-aggregate return indices. Note that when the fundamental-6 model is used, the contribution of the two fundamental variables average

[18] Both Table I and Table IV demonstrate the importance of adding fundamental variables. However, the relative importance of adding the variables varies from Table I to Table IV. The test values shown in Table I are affected by the model's ability to explain both the time series and cross-section of returns. The values in Table IV are only affected by the model's ability to explain the cross-section of returns. We examined the individual observations. There were no significant outliers. Thus, our results are not due to the ability of the fundamental variables to explain the expected return on only one or two of our portfolios.

19 percent, while the contribution of the two security sub-indices average 5 percent. Similarly, when we compare the two four-variable models, we find that, in index-4, the two subindices used (default and option) account for 7 percent of return explained. In contrast, when the fundamental variables are used, the two fundamental variables account for 24 percent of expected return. Once again, we see evidence of the importance of using fundamental expectational variables to explain mean returns.

III. Evaluation of Mutual Funds

Blake, Elton, and Gruber (1993) present the first systematic evaluation of bond mutual funds. They perform the evaluation by comparing the performance of actively managed funds with the performance of index (passive) funds of comparable risk. This procedure is one approach to mutual fund evaluation. Its principal advantage is that the evaluation is based on a comparison of mutual fund performance with the performance from a feasible strategy that could be followed by an investor. The procedure does not make use of or depend on the accuracy of any particular equilibrium model. The alternative to evaluating mutual fund performance by comparing actively managed fund performance to index funds is to evaluate performance using an equilibrium or relative pricing model. While these results are interesting in terms of what they show about mutual fund performance, they are perhaps more interesting because they provide an additional test of the type of relative pricing model we are examining by seeing if the model leads to plausible conclusions. Keep in mind that we are parameterizing the model (estimating the prices of risk) from the passive portfolios and applying this to getting expected returns for bond mutual funds. The mutual funds can be thought of as a holdout sample over a partially nonoverlapping time period. These empirical results not only test the model; they test whether the parameters hold over a fresh sample.

The initial sample consists of all bond funds that are in the ICDI data set as of December 1991, have data available from January 1986, and are not classified as high-yield funds by ICDI as of December 1991. High-yield funds are not examined because, as explained earlier, there are not sufficient data to fit a relative pricing model to high-yield funds. Thus, the model might well be missing variables necessary to explain the expected return of high-yield funds. Four funds are subsequently deleted from the sample because management fees cannot be obtained and we want to study investment performance before and after fees.[19]

The sample of 123 mutual funds examined in this paper does suffer from survivorship bias. The ICDI data excludes funds that change from bonds to a nonbond investment policy, or went out of business (normally merger) between 1986 and 1991. The data reported in Blake, Elton, and Gruber (1993) allow an

[19] We examined the performance of these four funds. The inclusion of these funds in the after-fee sample makes no difference in our results.

estimate of the order of magnitude of this bias. It is of the order of magnitude of 27 basis points per year (see footnote 31 in Blake, Elton, and Gruber (1993)).

In this section we evaluate bond mutual funds using the relative pricing models developed in earlier sections of this paper. The sensitivity of each of the 123 bond funds to the explanatory variables is obtained by regressing excess returns against the appropriate variables for each of the four models discussed earlier (time-series regression). The return-index variables are all calculated in excess-return form, although the risk-free rate cancels out for the default risk and option indexes (which are calculated as differences in primary indexes). The risk-free rate is not subtracted from the fundamental economic variables. The λs or prices of risk are set at the values estimated in the previous section of this article.

Table VI shows the average sensitivity of the 123 funds for each of the four models. In addition, we examine average sensitivities of the funds grouped into corporate, mortgage, and government categories. The grouping is the one used by ICDI and reflects the investment policy of the fund. Insofar as there are inaccuracies in the classification, the differential results across groups will not be as clean. When we examine the results across alternative models we see that the sensitivities to the aggregate index and to both the default and option indexes remain reasonably constant and independent of the number of indices utilized in the regressions. There is more variation in the sensitivities to the fundamental variables. This variation is more extreme when comparing models that include the default and option variables with models that exclude these variables. This indicates that the fundamental variables act in part as proxies for default and option effects when these influences are not measured directly. For each model there are differences in the sensitivities of funds to the indexes across different objectives. However, the pattern of sensitivities is generally consistent with what is expected. For example, corporate bonds have greater positive sensitivity to the S&P index and to default risk than do other categories. Mortgage funds have greater sensitivity to the option index than do the other categories of funds. For almost all funds and models, the sensitivity coefficient for inflation surprises is negative. Thus, bond funds have lower returns when anticipated inflation unexpectedly increases. This result is plausible, since we would expect an increase in anticipated inflation to be associated with an increase in interest rates and, hence, a decrease in returns. However, researchers who have measured surprises from a time-series regression often find the opposite result (e.g., Burmeister and McElroy (1988)). These results provide some support for the use of expectational data instead of time-series regressions. The sign of the sensitivity to GNP is positive for corporate bond funds but negative for the other categories. Interest rates tend to rise in a recovery, which would hurt returns. However, the unexpected increase in GNP lowers the risk of corporate bonds. The positive sign for corporate bonds indicates that the lower risk, and hence lower spread, more than compensates for the general increase in interest rates.

Table VII shows the average percentage of expected return explained by each factor. For each firm, it is constructed exactly as in Table V, by taking the

Table VI

Average Sensitivities for the Mutual Fund Sample

The sensitivities are estimated from a time series regression of each fund's returns on the returns or changes in fundamental values of the indicated indices. The sensitivities are then averaged across all funds and within each category of fund: corporate bond, government mortgage, and government bond.

Group	S&P	Default	Option	Aggregate	GNP	Inflation
Panel A: Index-1						
Corp				0.9143		
Morg				0.8423		
Gov				1.0526		
All				0.9649		
Panel B: Index-4						
Corp	0.0111	0.0039	0.0386	0.9056		
Morg	−0.0071	−0.0679	0.6045	0.9639		
Gov	−0.0001	−0.0704	0.0234	1.0691		
All	0.0021	−0.0457	0.1465	0.9946		
Panel C: Fundamental-4						
Corp	0.0100			0.9048	0.1004	−0.0585
Morg	−0.0198			0.8652	−0.0852	0.0538
Gov	−0.0109			1.0492	−0.2386	−0.0018
All	−0.0059			0.9649	−0.0972	−0.0089
Panel D: Fundamental-6						
Corp	0.0101	0.0011	0.0425	0.9102	0.0847	−0.0642
Morg	−0.0047	−0.0661	0.6104	0.9496	−0.1240	−0.0445
Gov	0.0016	−0.0689	0.0266	1.0594	−0.0905	−0.0196
All	0.0031	−0.0456	0.1504	0.9885	−0.0403	−0.0391

absolute value of the sensitivity times the risk premium for each factor and dividing by the aggregate of the absolute values across all factors. Note that for the fundamental variables, these risk premiums are those presented in Table II and are estimated from a sample of returns on passive portfolios. The results for each fund are then averaged within each category and across all firms. Consider the fundamental-6 model. About 73 percent of the expected return is explained by the aggregate index. However, a substantial proportion is explained by the two fundamental variables, 7.5 percent from the GNP variable

Table VII

Percentage of Mean Return Explained by Each Variable

The entries in this table are calculated by first taking the absolute value of the sensitivity times the risk premium for each index divided by the sum of the absolute values across all indices in the model. This number is then averaged across all funds. Sensitivities are estimated by time-series regressions of the bond fund returns on the indices; risk premiums are the lambdas shown in Table II for the fundamental variables and the average excess returns for the other indexes. "S&P" is the S&P 500 index; "Default" is the difference between the Blume/Keim high-yield bond index and the Lehman Brothers intermediate government bond index; "Option" is the difference between the Lehman Brothers GNMA index and a weighted average of the Lehman Brothers intermediate and long-term government bond indexes; "Aggregate" is a weighted average of the Lehman Brothers aggregate bond index and the Blume/Keim high-yield bond index; "GNP" is the change in expected real Gross National Product (GNP) growth; "Inflation" is the change in expected inflation.

Panel A: Index-4

Group	S&P	Default	Option	Aggregate
Corp	5.30%	0.49%	3.16%	91.05%
Morg	3.01%	0.48%	9.72%	86.78%
Gov	3.59%	0.54%	5.83%	90.04%
Avg	4.03%	0.51%	5.75%	89.70%

Panel B: Fundamental-4

Group	S&P	Aggregate	GNP	Inflation
Corp	4.24%	70.48%	12.09%	13.19%
Morg	4.73%	73.36%	7.31%	14.61%
Gov	3.53%	67.65%	13.84%	14.99%
Avg	4.00%	69.73%	11.94%	14.32%

Panel C: Fundamental-6

Group	S&P	Default	Option	Aggregate	GNP	Inflation
Corp	3.98%	0.37%	2.54%	73.18%	7.43%	12.49%
Morg	2.36%	0.41%	8.58%	74.70%	4.68%	9.26%
Gov	3.27%	0.41%	4.69%	71.64%	8.66%	11.33%
Avg	3.31%	0.40%	4.78%	72.77%	7.45%	11.29%

and 12 percent from the inflation variable. The remaining 8 percent is explained by the other three indices. Thus fundamental variables are important determinants of expected returns for active as well as passive funds.[20] While the importance of variables in the fundamental-6 model does not seem to change much across types of bond funds, there are some interesting differ-

[20] Like Burmeister and McElroy (1988), we do not make aggregate return orthogonal to the other indexes. Thus we are likely to be understating the performance of the fundamental variables in explaining expected return.

Table VIII

Performance (Alphas) of Mutual Funds After All Expenses except Load Charges

The column labeled "Avg" contains the average alpha (model intercept) of the funds in the indicated group; the column labeled "*t*" contains *t*-values that are adjusted to correct for both heteroscedasticity and cross correlation of the residuals. Funds are grouped by investment policy: "Corp" is corporate bond; "Morg" is government mortgage; "Gov" is government bond.

Panel A: Corp (40 Funds)			Panel C: Gov (58 Funds)		
Model	Avg	*t*	Model	Avg	*t*
Fundamental-6	−0.1184%	−2.30	Fundamental-6	−0.1124%	−3.57
Fundamental-4	−0.1241%	−2.18	Fundamental-4	−0.0582%	−1.19
Index-4	−0.0724%	−2.27	Index-4	−0.1129%	−5.81
Index-1	−0.0635%	−2.02	Index-1	−0.1081%	−3.79
Panel B: Morg (25 Funds)			**Panel D: All (123 Funds)**		
Model	Avg	*t*	Model	Avg	*t*
Fundamental-6	−0.1196%	−3.29	Fundamental-6	−0.1158%	−3.72
Fundamental-4	0.0154%	0.16	Fundamental-4	−0.0647%	−1.47
Index-4	−0.1105%	−4.89	Index-4	−0.0992%	−5.17
Index-1	−0.0433%	−0.83	Index-1	−0.0804%	−3.35

ences. For example, the S&P index explains the greatest amount of expected return for corporates, and the option index is most important in explaining expected returns for mortgages.

Table VIII shows the performance of our sample of mutual funds after all expenses are paid (except load charges).[21] This is the performance of the return series an investor would receive if all funds were no load funds. The average intercept (alpha) *across all funds* is negative for each of the models under study. Depending on the model used, the average intercept across all funds varies between a negative 6.4 basis points to a negative 11.6 basis points per month. Annualized, this is on the order of ¾ percent to 1⅓ percent per annum. There is no reason to believe managers that manage a certain type of fund (such as a mortgage fund) are superior to managers of other types of funds. When fundamental-6 is used to evaluate performance, the same average performance is computed for managers of corporate, government, or mortgage bond funds. When the other models are used, very different estimates are obtained. This gives support to the use of fundamental-6 (and in particular the

[21] For each average alpha shown in Tables VIII and IX, the reported *t* value has been adjusted using the full variance/covariance matrix of residuals across the funds in a given group in order to correct for both heteroscedasticity and cross correlations of the residuals.

Table IX

Performance (Alphas) of Mutual Funds with Expenses Added Back

The column labeled "Avg" contains the average alpha (model intercept) of the funds in the indicated group; the column labeled "*t*" contains *t*-values that are adjusted to correct for both heteroscedasticity and cross correlation of the residuals. Funds are grouped by investment policy: "Corp" is corporate bond; "Morg" is government mortgage; "Gov" is government bond. Monthly fund expenses are added to fund returns before alphas are calculated.

Panel A: Corp (40 Funds)			Panel C: Gov (58 Funds)		
Model	Avg	*t*	Model	Avg	*t*
Fundamental-6	−0.0436%	−0.85	Fundamental-6	−0.0259%	−0.82
Fundamental-4	−0.0493%	−0.87	Fundamental-4	0.0283%	0.58
Index-4	0.0024%	0.08	Index-4	−0.0264%	−1.36
Index-1	0.0113%	0.36	Index-1	−0.0216%	−0.76
Panel B: Morg (25 Funds)			**Panel D: All (123 Funds)**		
Model	Avg	*t*	Model	Avg	*t*
Fundamental-6	−0.0443%	−1.22	Fundamental-6	−0.0354%	−1.14
Fundamental-4	0.0907%	0.96	Fundamental-4	0.0158%	0.36
Index-4	−0.0352%	−1.56	Index-4	−0.0188%	−0.98
Index-1	0.0320%	0.62	Index-1	0.0000%	0.00

use of fundamental economic variables) as an appropriate relative pricing model.[22]

Table IX shows the same analysis where the return series have expenses added back. Expenses are annual expenses as reported in Wiesenberger's *Investment Companies* for December 1989 divided by twelve and added to each month's return.[23] The alphas pre-expenses are very small, varying from minus 3.5 basis points to plus 1 basis point on average. There is no evidence that managers, on average, can provide superior returns on the portfolios they manage, even if they provide their services free of cost.

When we regress alphas (computed after expenses) for the fundamental-6 model on expenses, we get an insignificant intercept (−0.06 percent) and a slope of −0.784 with a *t* statistic of −1.686. Thus, on average for every $1

[22] In our prior study of bond mutual funds, we compare bond funds to a set of passive portfolios. The results of this earlier study concerning aggregate performance are very similar to those reported above. In that study, overall performance varies from minus 8 basis points to minus 11 basis points depending on the measure used, which is almost identical to the results we find using our equilibrium models.

[23] Elton *et al.* (1993) show that relative expense ratios are fairly constant across years; i.e., funds with higher expense ratios in one year tend to have the higher ratios in subsequent years. Also, any linear time trend would be mitigated by using expense ratios from the middle of the sample period. Thus, we believe 1989 expense ratios are good measures for our six-year sample.

of expenses, performance decreases by $0.78. The small and insignificant intercept indicates that the major cause of underperformance is expenses and that an investor with no forecasting ability should select a low-expense bond fund.[24]

In this section we have seen that the relative pricing models examined earlier seem to do a good job of explaining mutual fund performance. The reasonableness of the results supplies additional evidence on the reasonableness of the models we have developed, and the market prices of risk we have estimated are applicable to sets of fixed-income securities not included in our original sample.

IV. Conclusion

In this article, we have developed relative pricing (APT) models that are successful in explaining expected returns in the bond market. In value, the bond market is several times larger than the equity market, yet it has received relatively little attention heretofore from researchers developing models of expected return.

Modern relative pricing models utilize unanticipated changes in economic variables as factors driving security returns. An innovation in this article is the measurement of those factors as changes in forecasts. When we compare the four alternative APT models, those that do not contain the fundamental expectational variables are rejected at the 5 percent level in favor of models that do contain those variables. The return indices are the most important variables in explaining the *time series* of returns. However, the addition of the fundamental variables leads to a large improvement in the explanation of *expected* returns. Furthermore, when we examine the percentage of expected returns explained by each of the variables, the fundamental variables are much more significant than all indices with the exception of the aggregate index.

We utilize our fundamental relative pricing models to examine the performance of bond funds. Bond funds underperform the returns predicted by the relative pricing models by the amount of expenses on average, and the models using fundamental variables do a better job than other models in accounting for the difference in performance between types of bond funds.

[24] This is consistent with Blake *et al.* (1993). The intercepts and slopes (with t statistics in parentheses) for the other three models are as follows. Index-1: −0.02% (−0.849) and −0.733 (−2.507); index-4: −0.03% (−1.199) and −0.922 (−3.806); fundamental-4: −0.02% (−0.396) and −0.530 (−0.822).

APPENDIX

Bond Indices Used As Passive Portfolios

Lehman Brothers	Ibbotson and Associates	Merrill Lynch
1–3 Govt.	Int. govt.	M0A0 (mortgage master)
Int. treasury	Long govt.	C8F0 (15+ yr. med. qual. industrials)
Int. agency		C8I0 (15+ yr. med. qual. utilities)
Long agency		C8K0 (15+ yr. high qual. finance)
Industrial		G802 (15+ yr. treasuries)
Finance		G8P0 (15+ yr. govt. agencies)
Int. Aaa		GVPO (1–5 yr. govt. agencies)
Long Aaa		
Int. Baa		
Long Baa		
FHLMC*		

* Federal Home Loan Mortgage Corporation.

REFERENCES

Baghestani, Hamid, 1992, On the formation of expected inflation under various conditions: Some survey evidence, *Journal of Business* 65, 281–293.

Bansal, R. and S. Viswanathan, 1993, No-arbitrage and arbitrage pricing: A new approach, *Journal of Finance* 48, 1231–1262.

Blake, Christopher R., Edwin J. Elton, and Martin J. Gruber, 1993, The performance of bond mutual funds, *Journal of Business* 66, 371–403.

Blume, Marshall E., and Donald B. Keim, 1987, Lower grade bonds: Their risks and returns, *Financial Analysts Journal* 43, 26–33.

Brennan, Michael, and Eduardo Schwartz, 1983, Duration Bond Pricing and Portfolio Performance, in Bierwag, Kaufman, and Toevs, eds., *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (JAI Press, Greenwich, CT).

Brown, Stephen J., and Philip H. Dybvig, 1986, The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates, *Journal of Finance* 41, 617–630.

Burmeister, Edwin, and Marjorie B. McElroy, 1987, APT and multifactor asset pricing models with measured and unobserved factors: Theoretical and econometric issues, paper presented at the November meetings of the Southern Finance Association, Washington, D.C.

Burmeister, Edwin, and Marjorie B. McElroy, 1988, Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory, *Journal of Finance* 43, 721–733.

Burmeister, Edwin, and Kent D. Wall, 1986, The arbitrage pricing theory and macroeconomic factor measures, *The Financial Review* 21, 1–20.

Burmeister, Edwin, Kent D. Wall, and James D. Hamilton, 1986, Estimation of unobserved expected monthly inflation using Kalman filtering, *Journal of Business and Economic Statistics* 4, 147–160.

Chen, Nai-Fu, Richard Roll, and Stephen Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–404.

Connor, Gregory, 1984, A unified beta pricing theory, *Journal of Economic Theory* 34, 13–31.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross, 1981, A reexamination of traditional hypotheses about the term structure of interest rates, *Journal of Finance* 36, 769–799.

Ederington, Louis H., and Jae Ha Lee, 1993, How markets process information: News releases and volatility, *Journal of Finance* 48, 1161–1191.

Elton, Edwin J., and Martin J. Gruber, 1995. *Modern Portfolio Theory and Investment Analysis*, 5th ed., (John Wiley & Sons, New York).

Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Matt Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies* 5, 1–22.

Elton, Edwin J., Martin J. Gruber, and Mustafa Gultekin, 1981, Expectations and share prices, *Management Science* 27, 43–76.

Elton, Edwin J., Martin J. Gruber, and Jianping Mei, 1993, Return generating processes, expected returns, and term premiums, Working paper, New York University.

Fama, Eugene F., and Robert Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680–692.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607–636.

Gallant, Ronald A., 1987, *Non-linear Statistical Models* (John Wiley & Sons, New York).

Gibbons, Michael R., 1982, Multivariate tests of financial models: A new approach, *Journal of Financial Economics* 10, 3–27.

Gibbons, Michael R., 1986, Empirical examination of return generating processes of arbitrage pricing theory, Unpublished Manuscript no. 881, Stanford University.

Gibbons, Michael R., and Krishna Ramaswamy, 1993, A test of the Cox, Ingersoll, and Ross model of the term structure, *Review of Financial Studies* 6, 619–659.

Gibbons, Michael R., Stephen A. Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Gultekin, N. Bulent, and Richard J. Rogalski, 1985, Government bond returns, measurement of interest rate risk, and the arbitrage pricing theory, *Journal of Finance* 40, 43–61.

Hansen, Lars P., and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Huberman, Gur, Shmuel Kandel, and Robert Stambaugh, 1987, Mimicking portfolios and exact arbitrage pricing, *Journal of Finance* 42, 1–9.

Knez, Peter J., Robert Litterman, and Jose Scheinkman, 1994, Explorations into factors explaining money market returns, *Journal of Finance* 49, 1861–1882.

Litterman, Robert, and Jose Scheinkman, 1991, Common factors affecting bond returns, *Journal of Fixed Income* 1, 54–61.

Longstaff, Francis, and Eduardo Schwartz, 1992, Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance* 47, 1259–83.

McCulloch, J. Huston, 1990, U.S. term structure data, 1946–1987, in *Handbook of Monetary Economics*, Vol. I, 672–715.

McElroy, Marjorie B., Edwin Burmeister, and Kent D. Wall, 1985, Two estimators for the APT model when factors are measured, *Economics Letters* 19, 271–275.

Merton, Robert C., 1973, An intertemporal asset pricing model, *Econometrica* 41, 867–880.

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

Pearson, Neil D., and Tong-Sheng Sun, 1994, Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance* 49, 1279–1304.

Peazy, John, ed., 1994, *Practitioner's Guide to Factor Models* (Research Foundation of the Institute of Chartered Financial Analysts, Charlottesville, Va).

Rich, Robert, 1989, Testing the rationality of inflation forecasts from survey data: Another look at the SRC expected price change data, *Review of Economics and Statistics*, 71, 682–686.

Roll, Richard, and Stephen Ross, 1980, An empirical investigation of the arbitrage pricing theory, *Journal of Finance* 35, 1073–1103.

Ross, Stephen A., 1976, The arbitrage pricing theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Snow, Karl, 1991, Diagnosing asset pricing models using the distribution of asset returns, *Journal of Finance* 46, 955–984.

Stambaugh, Robert, 1982, On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis, *Journal of Financial Economics* 10, 237–268.

Stambaugh, Robert, 1988, The information in forward rates: Implications for models of the term structure, *Journal of Financial Economics* 21, 41–70.

Wiesenberger Financial Service, 1986–1991. *Investment Companies*, (Warren, Gorham & Lamont, Inc., Boston).