The Performance of Bond Mutual Funds

Christopher R. Blake; Edwin J. Elton; Martin J. Gruber

Christopher R. Blake
Fordham University

Edwin J. Elton
New York University

Martin J. Gruber
New York University

The Performance of Bond Mutual Funds*

I. Introduction

The subject of mutual fund performance has received a great deal of attention in the literature of financial economics. Alternative metrics for performance have been proposed, and mutual fund performance has been examined extensively.[1] However, almost all of the empirical work on mutual fund performance has involved either common stock funds or funds that invest in both common stock and debt instruments (e.g., balanced funds). In this study, we will examine the performance of mutual funds that restrict their investments almost exclusively to debt instruments: bond funds.

There are several reasons why this study is important. First, bond funds constitute a major

Using linear and nonlinear models, we examine two samples of bond funds: one sample designed to eliminate survivorship bias, and a second much larger sample. Overall and for subcategories of bond funds, we find that bond funds underperform relevant indexes post-expenses. Our results are robust across a wide choice of models. We find that, on average, a percentage-point increase in expenses leads to a percentage-point decrease in performance. The nonlinear model weights closely match actual composition weights. We find no evidence of predictability using past performance to predict future performance for our unbiased sample.

part of the mutual fund industry, yet we have almost no information on their performance.[2] The present and potential importance of bond funds can be judged by looking at some historical data. At the end of 1978, there were 600 mutual funds registered in the United States, and they had total net assets of $58.1 billion. Of those funds, only 84 were bond and U.S. government security funds, and they held only 10.1% of the total net assets of mutual funds. By the end of 1990, there were 2,679 mutual funds with total net assets of $990.2 billion. Of those, 914 were bond funds representing 34.1% of the total number of funds, and those bond funds held 27.3% of the total net assets held by all mutual funds.[3] Given the importance of bond mutual funds in both size and growth, it is important to know more about how well they perform and what factors account for differences in their performance.

A second reason for studying bond funds is that most stock mutual fund performance studies include a large number of balanced funds. In those studies, all mutual funds, including balanced funds, are analyzed as if the portfolios consisted solely of common stock and Treasury bills. This study will aid in understanding whether the failure to explicitly incorporate one or more indexes for other types of debt instruments in the analysis can affect the interpretation of mutual fund performance results and will develop appropriate performance evaluation models for the bond portion of balanced funds.

A third reason for studying bond funds is that there are fewer influences affecting bond funds, and therefore the likelihood of measuring and understanding their performance is greater. Empirical evidence indicates that bond returns can be explained by no more than three and possibly two factors, while most studies of common stocks find that five to seven factors are necessary.[4] We know that the evaluation of stock and balanced mutual funds is extremely sensitive to the index or indexes that are used to measure performance.[5] Different indexes result in very different inferences about a fund's ability to outperform a passive strategy and result in very different rankings of funds. In this article, we will show that modeling is much simpler for bond mutual funds than for stock mutual funds and that bond fund performance is robust across a wide range of models.

A final reason for studying bond funds is to see if the multi-index

---

2. The exception to this is Cornell and Green (1991), who study high-yield bond funds. While theirs is an excellent study, the authors have appropriately, given the intention of their study, restricted themselves to a small segment of the bond mutual fund population.
3. See Wiesenberger Financial Services, *Investment Companies* (1979, 1991). The above figures understate the growing importance of bonds in the mutual fund industry since those figures exclude so-called balanced funds, flexible funds, and international bond funds, which hold large percentages of bonds in their portfolios.
4. See Elton, Gruber, and Nabar (1988) and Roll and Ross (1980).
5. See Lehmann and Modest (1987) and Elton, Gruber, Das, and Hlavka (1992).

models we use can explain the investment strategy followed by bond fund managers. That is, can we infer the types of securities held by a bond fund manager simply by studying the time-series return pattern? This type of analysis has been applied by both Elton, Gruber, Das, and Hlavka (1992) and Sharpe (1992) to bond-stock portfolios with some success; if it applies to bond portfolios, it will allow us to infer management style without having full access to composition data.

II. Samples

In this study we employ two very different samples of bond fund data. The first of these consists of the 46 nonmunicipal bond funds identified as having a "bond" or "specialized" investment policy in the 1979 edition of Wiesenberger Financial Services' *Investment Companies*.[6] This sample excludes money market funds. End-of-month net asset values (NAVs), distributions, and reinvestment NAVs were obtained from Interactive Data Corporation (IDC) for most of these funds. The IDC data were supplemented by data from the funds themselves for those funds not listed by IDC and for the many missing data points in the IDC data set.[7] The missing data were primarily reinvestment NAVs. The IDC data were then analyzed and compared to Wiesenberger data and data obtained from the funds themselves to check for any errors. Monthly returns were calculated by computing the value at the end of the month minus the value at the beginning of the month divided by the value at the beginning of the month. If the firm paid a distribution during the month, we assumed the investor used the distribution to purchase additional shares at the reinvestment NAV, and the ending value was calculated including these shares.[8]

---

6. Four of the funds listed as having a "bond" investment objective are excluded, either because they were variable-annuity backing funds or because investment was limited to a particular group of investors; not all "specialized" funds are included, only those that involved bonds (primarily U.S. government securities).

7. Interactive Data Corporation obtains its historical mutual fund data from NASDAQ. Missing mutual fund data in the IDC databases are primarily due to NASDAQ listing requirements or to funds not reporting to NASDAQ in time for NASDAQ listing.

8. We were unable to obtain all reinvestment NAVs. When we were unable to obtain a reinvestment NAV for a distribution, we estimated the value of the distribution at the end of the month. This value was determined by compounding the actual distribution to the end of the month using an estimate of the fund's monthly return. The estimated return for the fund was determined by assuming that the distribution came at the end of the month if the distribution occurred in the last half of the month or by deducting the distribution from the prior month-end NAV if the distribution occurred in the first half of the month. The resulting estimated return was then assumed uniform over the month for the purposes of compounding the distribution to the end of the month. Also, for so-called daily accrual funds, which pay daily distributions, only "accumulation factors" representing the total accumulation from one reinvestment date to the next were avail-

Our initial sample period is the 10-year period from the beginning of 1979 to the end of 1988. Ten of the 46 sample funds did not survive during the sample period, due to either liquidation, merger, or change in investment policy. For those funds that ceased to exist during a month, we assumed reinvestment to the end of the month at the Treasury-bill rate. For five of the 10 nonsurviving funds (two of which liquidated and three of which merged into other funds during our sample period), we were able to obtain only annual data.

In the final section of this article, we examine performance using a second, much larger, sample of funds. This sample consists of all bond funds that existed at the end of 1991 and was obtained from Investment Company Data, Incorporated (ICDI). This latter sample obviously has survivorship bias. Survivorship bias is less important for bond funds than it is for stock funds since bond fund performance is less variable and, consequently, fewer funds merge or dissolve. Furthermore, the smaller sample, which we analyze extensively, does not have this problem. Thus, we can obtain a rough measure of the magnitude of the problem by comparing the results for the funds that survive in that sample with the results for those that do not survive in that sample. As latter sections will show, survivorship bias causes bond fund performance to appear better than it actually is; thus, it cannot explain the underperformance we find for the larger sample.

We use several indexes to measure performance in this study. These include the return on 30-day Treasury-bills from Ibbotson and Associates; various government, mortgage, and investment-grade corporate bond indexes from Lehman Brothers; and the Blume/Keim high-yield bond index.[9] The Lehman Brothers indexes are comprehensive market-weighted indexes and were chosen because they are the most widely used indexes for bond portfolio evaluation and are widely used by passive bond portfolio managers.

III. Models

In the previous section of this article, we mentioned a set of bond indexes that will be used in this study. The return on any of those indexes can be considered the return on a corresponding passive portfolio. The simplest comparison to make for a fund is to compare its rate of return to that of a passive portfolio represented by one or more indexes. However, this comparison neglects the differentials in risk

---

able for distribution data. In a month where daily distributions had been declared and accrued but not yet reinvested or paid, we adjusted the actual month-end NAV by first calculating the average daily distribution for the accrual period and then adding the total amount accrued since the last reinvestment date to the month-end NAV.

9. See Blume and Keim (1987) and Blume, Keim, and Patel (1991) for descriptions of the Blume/Keim high-yield index.

that may exist between a fund and an index. A more sophisticated approach is to compare a fund to a passive portfolio that is a mixture of the selected benchmarks and a risk-free asset and that has the same risk as the fund. This can most easily be done when the fund's excess return (the fund's return minus the riskless rate) is related linearly to the excess return on an appropriate index.

More specifically,

$$R_{it} = \alpha_i + \beta_i I_t + \epsilon_{it}, \tag{1}$$

where

R_{it} = the continuously compounded excess return on the ith fund during month t,

α_i = the average risk-adjusted excess return for the ith fund,

I_t = the continuously compounded excess return on an appropriate index during month t,

β_i = the sensitivity of the excess return on the ith fund to the excess return on the index, and

ϵ_{it} = the residual return of the ith fund during month t not accounted for by the model.

There are two ways to think about this model. First, this time-series regression model can be thought of as a single-index model similar to the market model used in the evaluation of common stocks. The β_i's in this characterization are the sensitivities (risk) of the funds to the selected market index. An alternative interpretation is that the manager of fund i has a choice of holding various combinations of a riskless asset and an active bond portfolio and that β_i is the fraction of the manager's money that, if invested in the *passive* portfolio with excess returns (above the riskless asset) I, along with the remainder in the riskless asset, would provide a total return that best matched the return on the manager's portfolio.[10]

To use this model, we have to define the appropriate index. One choice is the Lehman Brothers (LB) government/corporate bond index. The Lehman Brothers government/corporate bond index is a market-weighted index of government and investment-grade corporate issues that have more than 1 year remaining until maturity and is the one most often used as a basis of comparison for bond portfolios. Both because it is so widely used and because it is likely to be the index that most closely matches the aggregate of all funds in our sample of

10. Estimating this regression is equivalent to finding the pair of weights for the index and the riskless asset that minimizes the squared difference in return between this weighted combination of the passive portfolios and the portfolio that the manager actually holds.

mutual funds, we employ this index as one single-index model and call it Market-1.

For a second one-index model, we used the single index that seemed most appropriate given the objectives and policies of the fund as described by Morningstar in their winter 1989 *Sourcebook.*[11] We call this model Own-1. Examining Morningstar descriptions led us to divide our fund sample into six categories or groups. The intent was not to search for an optimal index for each fund but to see if an investor looking at a single issue of Morningstar could select an index that supplied more information about performance than did the overall index used in the first single-index model. To the extent that looking at only one date does not provide a description of a fund's typical behavior over the full period, or that the funds are not classified in a meaningful way, our results will tend to understate the real performance of Own-1 relative to Market-1.

The groups we used, the members of each group, and the index selected for each group for model Own-1 can be seen by examining table 1. (Each index is shown in parentheses.) The choices of the indexes, given the stated objectives of the funds and our groupings, are fairly obvious. The Lehman Brothers mortgage-backed securities index was used for funds that had as an objective government mortgage, the Blume/Keim high-yield index was used for funds that had high yield as an objective, and the Lehman Brothers government/corporate index was used for funds that had high quality as an objective. The Lehman Brothers corporate index was used for the three remaining groups that had general investment objectives involving corporate bonds. There was a group of funds that were not classified by Morningstar. The Lehman Brothers government/corporate index was used for this group.

The next set of models we employed was designed to see if introducing more indexes allowed us to explain more of the mutual fund returns or to gain more insight into their performance. These models can be described, in general, as

$$R_{it} = \alpha_i + \sum_{j=1}^{K} \beta_{ij}I_{jt} + \epsilon_{it}, \tag{2}$$

11. Alternatively, this model can be thought of as a multi- or four-index model where each fund is loaded on only one index. The use of the recent date was necessitated by the fact that Morningstar did not publish a description of bond mutual funds at the start of our sample. The result of taking a date outside of our sample period biases the results in favor of Market-1.

where

R_{it} = the continuously compounded excess return on the ith fund during month t,

α_i = the average risk-adjusted excess return for the ith fund,

I_{jt} = the continuously compounded excess return on an the jth index during month t,

β_{ij} = the sensitivity of the excess return on the ith fund to the excess return on the jth index,

K = the number of indexes employed (three or six in our analysis), and

ϵ_{it} = the residual return of the ith fund during month t not accounted for by the model.

In a manner analogous to the single-index model, we can consider equation (2) as a multiple-index model with the β_{ij}'s representing sensitivities or risks. Alternatively, we can think of the β_{ij}'s as the weights on sets of passive portfolios that best represent (reproduce the return series of) the fund under study.

We employed both three-index and six-index models in our analysis. Two different three-index models were employed. The first, called Risk-3, used the Lehman Brothers government/corporate bond index, the Lehman Brothers mortgage-backed securities index, and the Blume/Keim high-yield index. The government/corporate bond index was used to capture the influence of a large population of investment-grade bonds. The high-yield bond index was introduced to examine the effect of holding low-quality bonds. The mortgage index was introduced to capture the specific effect of a class of government securities that act differently because of the particular option features found in mortgages.[12]

A second three-index model, called Maturity-3, was also used. The Lehman Brothers government/corporate bond index was replaced by both the Lehman Brothers intermediate government bond index and the Lehman Brothers long-term government bond index, and the mortgage index was deleted.[13] This model was used to see if sensitivity to different portions of the maturity spectrum was more important than including the different attributes of mortgages.

A six-index linear model, called Reg-6, was introduced to allow us

12. The Lehman Brothers government/corporate bond index does not include mortgage-backed securities.
13. The Lehman Brothers intermediate government bond index is a market-weighted index of government bonds with maturities between 1 and 10 years. The long-term index contains bonds with maturities beyond 10 years.

TABLE 1   Mean Differences (Continuously Compounded Monthly Returns in %)

CUSIP*	Number of Months	Fund Return—T-Bill	Fund Return—Government/Corporate	Fund Return—Own Index
Group 1: Government mortgage backed (LB mortgage-backed securities index):				
00168210	120	.100	−.046	−.080
35349660	120	−.023	−.169	−.203
36079910	120	.019	−.126	−.161
Group 2: High quality (LB government/corporate index):				
09787310	120	.205	.060	.060
26188010	120	.148	.003	.003
31617Q10	120	.209	.063	.063
46623610	120	.088	−.057	−.057
49344010	120	.069	−.076	−.076
62830210	120	−.009	−.155	−.155
82657010	120	.044	−.101	−.101
9203110	120	.110	−.035	−.035
Group 3: Investment grade (LB corporate index):				
31614610	120	.030	−.115	−.128
44919810	120	.115	−.030	−.043
48841210	120	.109	−.037	−.050
77957010	120	.073	−.072	−.085
81429110	120	.057	−.089	−.102
91045020	120	.134	−.012	−.025
Group 4: General (LB corporate index):				
40975210	120	.054	−.092	−.105
49344020	120	.172	.027	.013
57564010	120	.126	−.020	−.033

CUSIP	No. of months			
63762370	120	−.037	−.182	−.196
81727040	120	.059	−.086	−.099
Group 5: General with significant high yield (LB corporate index):				
02490310	120	.092	−.053	−.067
31739010	120	.101	−.045	−.058
Group 6: High yield (Blume/Keim high-yield index):				
02501010	120	.137	−.008	−.151
24661710	120	.184	.039	−.103
31419610	120	.195	.050	−.092
32057310	120	.087	−.059	−.201
48841110	120	.347	.202	.059
49344030	120	.100	−.046	−.188
54400410	120	.189	.043	−.099
57564710	120	.294	.149	.006
68379610	120	.096	−.049	−.191
74678210	120	.260	.114	−.028
92203120	120	.228	.082	−.060
Group 7: Not listed in Morningstar (LB government/corporate index):				
01852830	120	−.061	−.207	−.207
02629810†	18	−.216	.077	.077
02630010†	107	.046	−.097	−.097
55057410†	59	−.139	−.053	−.053
57564810†	22	−.633	.074	.074
78643610†	73	−.061	−.067	−.067
Overall statistics:				
No. of positive mean differences		33	13	8
No. of negative mean differences		8	28	33
Average mean difference		.076	−.029	−.073

* CUSIP = Committee on Uniform Security Identification Procedures number.
† Indicates a fund that exited early for which we have monthly data. Returns are over the period before exit.

to capture both the differences in maturity range and the differences in risk premiums between securities. We used both the Lehman Brothers intermediate and long-term government bond indexes to capture maturity. The Lehman Brothers intermediate corporate index, long-term corporate index, and mortgage-backed index and the Blume/Keim high-yield index were used to capture risk and option differences.[14]

The β's in all of the linear index models are estimated via regression analysis. While this optimizes the explanatory power of the models, it sometimes results in negative weights for (implied investment in) some asset categories. In the latter section of this article, after examining performance, we wish to examine our ability to infer investment policies of mutual funds. Since funds do not short sell, we would like to have a solution involving a replicating portfolio with no short sales. Sharpe (1992) has shown that such solutions can be obtained by formulating a quadratic programming problem. To find the best replicating portfolio, we solved the following quadratic programming problem for each fund, using the same indexes that we used for Reg-6:

$$\min_{\beta} \sum_{t=1}^{N} \left[\left(R_{it} - \overline{R}_i \right) - \sum_{j=1}^{6} \beta_{ij} \left(I_{jt} - \overline{I}_j \right) \right]^2 \tag{3}$$

$$\beta_{ij} \geq 0, \, \forall j$$

The β_{ij}'s in equation (3) are the appropriate sensitivities or portfolio weights given that all weights are greater than zero. $\overline{R}$ and $\overline{I}$ represent sample means. We will compare this model (called QPS-6) to the regression models primarily to see if it yields similar results and thus is a reasonable model for examining correspondence between sensitivities and managers' portfolio choices. This represents our final model.

IV. Overall Performance

In the prior section, we discussed the models we use to examine performance of bond mutual funds. In this section, we will apply these models to our sample. Initially, we will restrict our analysis to the 41 funds for which we have monthly data; we will end the section by examining the performance of the five funds that cease to exist early in our sample period and for which we have only annual data. Throughout this section, we present data grouped into categories as previously discussed.

A. Performance and Alternative Methods of Measurement

Table 1 shows the difference in return between each fund and the 30-day Treasury-bill rate, between each fund and the Lehman Brothers

---

14. A second six-index linear model was run employing the Lehman Brothers government/corporate index instead of the long-term government index. The results were virtually the same, so we will not report them here.

government/corporate index, and between each fund and the single index most closely matching its investment objective as described by Morningstar. For each fund for which we had monthly data and that exited early (indicated by a dagger in the table), the returns are over the period before it exited. Positive numbers indicate the fund outperformed the index. Examining the last two columns shows the average fund underperformed the overall index or the one that most closely matched its policy. This monthly underperformance was -0.029% and -0.073%, respectively. These numbers are presented to give a historical perspective. For a meaningful analysis of the contribution of active management, we need to utilize the models developed in the prior section. Those models take account of the fact that the results shown in table 1 could arise because fund investment policies did not match the composition of the index in either asset mix or maturity.

Table 2 shows the average monthly alpha (across all funds) for each of the models discussed in the last section.[15] The average alphas across models are very similar, varying from -0.023% to -0.069%. The only model that does not account for the high-yield effect (Market-1) is a possible outlier with an alpha of -0.023%, while the range of the remaining alphas is -0.046% to -0.069%. The number of funds with negative alphas is also similar across models, varying from 27 to 33. The importance of the number of negative alphas would be overstated if the residuals from a performance model were highly correlated across funds. They are not. For example, for the six-index models, the average correlation across funds is less than 0.10. Thus, correlation among residuals cannot account for the large number of negative alphas.

If we examine average performance for each of our subgroups, we see that group performance is negative across all models except for the high-yield group. For this group, alpha is positive for Market-1 and negative for all of the other models. We consider this to be evidence of the failure of Market-1 as a measure rather than evidence of superior performance of high-yield bond fund managers.

Another way to see if the alternative models lead to the same performance conclusions is to ask if the same funds are ranked as having the manager provide value across alternative models. Thirty-one funds out of 41 have the same sign for alpha (24 negative and seven positive) across all measures. Six additional funds (five of which are high yield) have the same sign for alpha except for Market-1. Thus, once we account for the effect of high yield, we would reach the same conclusions concerning performance across measures for 37 out of 41 firms.

---

15. The average alphas for this sample are not significantly different from zero, although they are for the larger samples discussed later. Standard deviations of average alphas across the 36 surviving funds were computed using the full variance/covariance matrix of the residuals in order to correct for both heteroscedasticity and cross correlations of the residuals.

TABLE 2   Fund Performance by Groups and Models: Average Monthly Alpha
                 (in %)

	Models					
	One-Index		Three-Index		Six-Index	
CUSIP*	Market-1	Own-1	Risk-3	Maturity-3	Reg-6	QPS-6
Group 1: Government mortgage backed:						
00168210	−.039	−.020	−.048	−.063	−.036	−.036
35349660	−.155	−.147	−.148	−.203	−.165	−.166
36079910	−.149	−.142	−.166	−.248	−.190	−.173
Group 1 average	−.114	−.103	−.121	−.171	−.130	−.125
Group 2: High quality:						
09787310	.068	.068	.040	.045	.067	.066
26188010	.040	.040	.034	.063	.078	.063
31617Q10	.120	.120	.120	.127	.131	.134
46623610	−.084	−.084	−.077	−.041	.003	−.040
49344010	−.069	−.069	−.088	−.076	−.060	−.058
62830210	−.090	−.090	−.093	−.077	−.084	−.077
82657010	−.115	−.115	−.149	−.169	−.138	−.139
92203110	−.039	−.039	−.045	−.054	−.024	−.025
Group 2 average	−.021	−.021	−.032	−.023	−.003	−.009
Group 3: Investment grade:						
31614610	−.121	−.099	−.126	−.136	−.112	−.112
44919810	−.020	−.002	−.019	−.042	−.015	−.030
48841210	−.038	−.019	−.050	−.068	−.043	−.045
77957010	−.001	.010	.008	.037	.037	.010
81429110	−.038	−.027	−.069	−.040	−.028	−.050
91045020	−.045	−.022	−.083	−.093	−.063	−.061
Group 3 average	−.044	−.027	−.056	−.057	−.037	−.048
Group 4: General						
40975210	−.109	−.088	−.131	−.148	−.115	−.109
49344020	.072	.084	.023	.082	.082	.045
57564010	−.030	−.009	−.045	−.063	−.036	−.036
63762370	−.171	−.155	−.243	−.237	−.223	−.223
81727040	−.074	−.055	−.077	−.060	−.042	−.039
Group 4 average	−.062	−.045	−.094	−.085	−.067	−.072
Group 5: General with significant high yield:						
02490310	−.049	−.034	−.116	−.110	−.078	−.092
31739010	−.017	−.002	−.031	−.002	.019	−.006
Group 5 average	−.033	−.018	−.073	−.056	−.030	−.049
Group 6: High yield:						
02501010	.013	−.106	−.098	−.094	−.067	−.091
24661710	.046	−.049	−.023	−.018	.020	−.009
31419610	.062	−.047	−.031	−.047	−.016	−.025
32057310	.015	−.112	−.134	−.116	−.135	−.112
48841110	.220	.114	.126	.117	.136	.125
49344030	.001	−.107	−.107	−.083	−.044	−.099
54400410	.078	−.042	−.044	−.031	−.033	−.042
57564710	.186	.079	.083	.103	.130	.087
68379610	−.032	−.158	−.153	−.163	−.149	−.159
74678210	.119	−.002	.012	−.009	.021	.007
92203120	.113	.017	.030	.029	.053	.041
Group 6 average	.075	−.038	−.031	−.028	−.008	−.025

TABLE 2   *(Continued)*

CUSIP*	Models					
	One-Index		Three-Index		Six-Index	
	Market-1	Own-1	Risk-3	Maturity-3	Reg-6	QPS-6
Group 7: Not listed in Morningstar:						
01852830	−.207	−.207	−.234	−.211	−.157	−.224
0262981O†	−.074	−.074	−.089	−.133	−.167	−.108
0263001O†	−.063	−.063	−.065	−.101	−.082	−.082
5505741O†	−.042	−.042	−.071	−.091	−.113	−.106
5756481O†	−.142	−.142	−.298	−.316	−.154	−.177
7864361O†	−.065	−.065	−.060	−.096	−.097	−.099
Group 7 average	−.099	−.099	−.136	−.158	−.128	−.133
All groups:						
Positive alphas	14	8	9	8	12	9
Negative alphas	27	33	32	33	29	32
Average alpha	−.023	−.046	−.067	−.069	−.046	−.055
t average alpha	−.214	−1.230	−1.280	−1.115	−.900	−1.175
Average adjusted R^2	.722	.786	.803	.792	.824	.817

NOTE.—*t*-values are for average alphas of the 36 surviving funds and reflect adjustment for heteroscedasticity and correlation across those funds; for all models, the average alphas of the five nonsurviving funds is more negative than the average alpha across all funds.

* CUSIP = Committee on Uniform Security Identification Procedures number.

† Indicates a fund that exited early for which we have monthly data. Returns are over the period before exit.

The rankings across funds with positive alphas are also very similar. Table 3 provides the rankings for those funds with positive alphas on any measure. A dagger indicates that the alpha was negative for the fund. Except when using the first measure (the alpha from Market-1), an investor interested in examining which managers provided the most value added would reach similar conclusions no matter which measure was used, as long as the measure contained a high-yield index.

Studies that have analyzed common stock or balanced fund returns have increasingly found that utilizing multi-index models results in very different ranking than using single-index models (see Lehmann and Modest 1987; Grinblatt and Titman 1989*b*; Connor and Korajczyk 1991; and Elton, Gruber, Das, and Hlavka 1992). Further, different multi-index models result in very different rankings. As we have just seen, single-index models that include an index for high-yield bonds and mortgages (Own-1) and various multi-index models result in virtually the same ranking for bond funds.

Finally, returning to table 2, we see that there is an increase in explanatory power (adjusted R^2) with any model that includes a high-yield index and a mortgage index over any model that does not. However, even the simple one-index model explains over 70% of the return

TABLE 3   Performance Ranking by Models

CUSIP*	Market-1	Own-1	Risk-3	Maturity-3	Reg-6	QPS-6
48841110	1	2	1	2	1	2
31617Q10	3	1	2	1	2	1
57564710	2	4	3	3	3	3
49344020	7	3	7	4	4	6
09787310	8	5	4	6	6	4
26188010	11	6	5	5	5	5
92203120	5	7	6	8	7	7
77957010	†	8	9	7	8	8
74678210	4	†	8	†	9	9
24661710	10	†	†	†	10	†
31739010	†	†	†	†	11	†
46623610	†	†	†	†	12	†
54400410	6	†	†	†	†	†
31419610	9	†	†	†	†	†
32057310	12	†	†	†	†	†
02501010	13	†	†	†	†	†
49344030	14	†	†	†	†	†

* CUSIP = Committee on Uniform Security Identification Procedures number.
† The alpha was negative for the fund.

variation. Also, the reader should note that imposing the added restrictions of the QPS model leads to almost no change in results.

In this section, we have shown that the implied performance of bond funds is not very sensitive to the method used to measure it. Whether one is concerned with evaluating overall performance, ranking funds, or simply selecting the best funds, one reaches similar results with any one of several reasonable models. There is evidence that funds that hold a large percentage of their assets in high-yield securities are better measured by a model that explicitly accounts for these securities.

Before turning to other uses of our models, it is worth examining the effect of exiting funds on our analysis.

B.   Exiting Funds and Bias

A number of funds ceased to exist during our sample period. Those for which we had monthly return data are shown in table 2 under "Not Listed" and are marked with a dagger. The average alphas are all negative, where the alphas are calculated over the period before the fund ceased to exist.[16] The average alpha for these funds using Reg-6 is −0.123%. Annualized, this is about 1.02% below the average alpha for our sample of funds that existed through the entire sample period. Table 4 shows annual returns for the three funds that exited early for

16. We examined performance of other funds over these subperiods. The magnitude of the negative alphas are not due to all funds doing poorly in the subperiods.

TABLE 4   Annual Returns for Exiting Funds (in %)

Fund ID	1979	1980	1981
0186001	−3.5	−2.2	3.6
7123001	2.8	−2.6	1.5
9001001	−1.1	. . .	. . .
Average	−.6	−2.4	2.5
All sample funds	2.7	1.9	5.3

NOTE.—These funds exited early and only annual data are available.

which we have annual data but not monthly data.[17] These three funds underperformed the average fund in our sample in each year. The underperformance varied between 2.8% and 4.3% per year.[18] From the monthly and annual data, it is clear that failed funds underperform other funds.

Analyzing the total return on all funds (including those which dissolved) showed that the average fund underperformed a passive portfolio by 75–95 basis points annually.[19] The average expense ratio at the end of 1983, obtained from Wiesenberger's *Investment Companies*, for our sample funds still in existence at that time is 83 basis points, and this is a good estimate for the full 10-year sample period. Thus, the managers of funds that do not dissolve do better before transaction costs than random selection, while for the group as a whole, the performance before transaction costs is about equal to that achieved by random selection.

Most studies of mutual fund performance ignore in their analyses the part of a fund's portfolio invested in bonds. This is especially critical for those studies that include balanced funds. Unless bond returns are correlated with the indexes used for stocks, the presence of bonds will impact the fund's alpha by the difference between the return on the bond portion of the portfolio and the Treasury-bill rate

17. Two additional funds failed almost immediately on forming. We do not have any return data for those funds.

18. The annual data assume all interest income is received at the end of the year, whereas monthly data use reinvestment rates. This could overstate the underperformance. To examine the magnitude of this effect, consider the following. The average fund in this period had approximately 10% interest income. Assuming all interest was received at the beginning of the year and using a 10% short-term rate would add 1% (10% of 10%) to the return. However, given the generally poor returns of these funds, assuming instead that interest income was reinvested into the funds (as we did for the funds with monthly data) might well result in greater underperformance.

19. These numbers were computed assuming that the funds for which we do not have any return data were similar to the three funds for which we have only annual return data. Estimates of the impact of the 10 funds that exit our sample on overall results are made by calculating average return with and without those 10 funds included in the sample. Differentials in average return are translated directly into differences in average alphas.

times the percentage invested in bonds.[20] Table 1 shows that for most funds this difference is positive and on average about .076% per month for all funds and .123% per month for those that survived.

We regressed our indexes on the Standard and Poor's (S&P) 500 index, and the R^2, except for the high-yield index, ranged from .04 to .09. For the high-yield index, the R^2 was .25. Thus, studies that include balanced funds and do not include indexes that pick up the bond component produce upward-biased estimates of alphas.

V.   Forecasting

Up to this point, we have examined whether alternative index models tend to produce different estimates of performance. The purpose of this section is to examine whether past alphas are predictive of future alphas. To accomplish this, our 10-year sample was divided into two 5-year periods and three 3-year periods (based on the last 9 years of data). For the six models under study, alphas were computed for each subperiod. Rank correlations were calculated between the alphas computed from each model and the alphas computed for the same model in adjacent periods. The results are shown in table 5. The first thing to notice is that all of the rank correlations are small, and only one is statistically significant at the 2.5% level. Across the 5-year subperiods, and on average across the 3-year subperiods, the single-index model based on the government/corporate index (Market-1) produces the highest rank correlation. This is because the single-index model does not capture high-yield bond performance and high-yield funds had positive alphas with respect to Market-1 over much of the period. It should not be interpreted as indicating a superior index. The result we see here is similar to finding that stock funds that contain a large proportion of small stocks tend to produce consistently positive alphas using a single-index model based on the S&P index.[21]

As another test of the predictability of alpha, we examined the difference in squared forecast error from predicting alpha from each model using a fund's past alpha compared to a naive forecasting model that assumed that the future alpha was equal to the average historic alpha (across all funds) produced by that model. In all cases across all time periods, the naive model produced superior forecasts, indicating that there was no information produced by the ordering of historical alphas.

20. If a study has identified the correct equilibrium model for both bonds and stocks, then analyzing balanced funds does not introduce a bias.

21. See Elton, Gruber, Das, and Hlavka (1992). In table 5, we are implicitly assuming that residuals are uncorrelated. For the multi-index models, that is a reasonable assumption since the average correlation between residuals is less than 0.10. Furthermore, correlated residuals would bias results in favor of finding rank correlation, and thus that cannot explain the results shown in table 5 for the multi-index models.

TABLE 5   Rank Correlations across Adjacent Time Periods

	5-Year Subperiods	First Pair 3-Year Subperiods	Second Pair 3-Year Subperiods
Market-1	.234	.090	.378*
Own-1	−.111	−.046	−.037
Risk-3	.063	−.025	.059
Maturity-3	−.010	.145	.019
Reg-6	.010	.046	.074
QPS-6	−.001	−.068	.083

* Significant at the 2.5% level; the critical value at the 2.5% level (one tail) varies from .319 to .328 across the pairs of subperiods.

While it appears that none of the models produce useful information about future performance for funds in general, it is possible that the models produce useful information about funds that are unusually good or bad performers. To check on this, we selected the four funds with the highest alphas and the four funds with the lowest alphas (from each model) and examined how well these funds performed in the subsequent period. The results for the 5-year subperiods are shown in table 6. The only measures in which the best performing funds continued to outperform the worst performing funds and the average fund were those obtained from Market-1 and Maturity-3. Even here the results are not very encouraging. The magnitude of the difference is small (less than .06% per month). Also, in the case of Maturity-3, the worst performing funds (in addition to the best performing funds) outperform the average fund. Furthermore, in the case of Market-1, even this very weak evidence of consistency is most likely due to the failure to consider relevant influences on bond returns. In addition, when 3-year subperiods were examined, there was no model where the best performing funds beat both the worst performing funds and the average fund in both of the 3-year subperiods.

All of these tests indicate strongly that, while all of the models produce broadly similar ranking of funds, none is useful in selecting funds that have higher alphas in subsequent periods.

TABLE 6   Persistence in Performance for Extremes (Average Monthly Alphas in Second Subperiod; 5-Year Subperiods; in %)

Model	Best Four	Worst Four	All Funds
Market-1	.032	−.027	−.006
Own-1	−.245	−.081	−.140
Risk-3	−.046	−.023	−.080
Maturity-3	−.035	−.068	−.077
Reg-6	−.103	−.115	−.092
QPS-6	−.147	−.129	−.085

VI.   Sensitivity Coefficients and Investment Policy

In this section, we will show that the historic return series for a fund can be used to infer its investment policy. The sensitivity of a fund's return to an index is determined by the proportion of the portfolio invested in the asset class represented by the index and the duration of bonds in that asset class relative to the duration of the index.[22] In equation form, this is

$$\beta_{ij} = \frac{D_{ij}}{D_j} X_{ij}, \tag{4}$$

where

β_{ij} = the sensitivity of fund i to the index used for asset class j,
D_{ij} = the duration of bonds in fund i that are classified as
belonging to asset class j,
D_j = the duration of the index used for asset class j, and
X_{ij} = the proportion of bonds in fund i invested in asset class j.

End-of-year (fourth-quarter) data on fund composition for the 6 even years between 1978 and 1988 were provided by Lipper Analytical Services.[23] The data we obtained were aggregated by type (government, corporate investment grade, corporate low grade) and by maturity (1–5 years, 5–10 years, etc.).[24] The classifications used by Lipper do not directly correspond to the indexes in QPS-6 in the sense that Lipper's maturity breakdown is for the portfolio as a whole and not for each type of bond. Thus, we reestimated the QPS model using indexes that conform to Lipper's classifications. Two divisions are possible. One is by bond type, divided into governments, investment-grade corporate, and high-yield corporate (represented by the LB government and corporate indexes and the Blume/Keim high-yield index). The second is by maturity. In this case, the indexes available (the LB intermediate and long-term government/corporate indexes) split maturity between 1 and 10 years and greater than 10 years, and we can arrange Lipper data to conform to this grouping.

We ran a QPS analysis with each of the two divisions for the 34 funds in our sample for which we could obtain Lipper data. We aggregated data within Morningstar groups. Aggregating the data should cause the aggregate duration to be closer to the index duration and thus make the slope in equation (4) closer to unity. Panel A of table 7 shows the QPS betas and average Lipper proportions (over the 6 years)

22. See Elton and Gruber (1991), pp. 557–58, for a proof.
23. For a number of firms, end-of-year composition data were missing. In those cases, we used adjacent quarters: third quarter of the prior year or first quarter of the next.
24. Lipper has finer classifications for corporate bonds.

TABLE 7 Lipper and QPS Proportions by Bond Type (in %; Proportions Scaled to Add to 100%)

| | Bond Type | | | | | |
| | Government Agency | | Corporate | | High Yield | |
Group	QPS	Lipper	QPS	Lipper	QPS	Lipper
A:						
Government mortgage backed	23.16	90.32	75.32	8.95	1.52	.73
High quality	43.72	39.30	48.92	54.62	7.36	6.08
Investment grade	24.42	30.04	68.65	66.42	6.92	3.54
General	18.89	14.43	61.74	66.36	19.36	19.21
General with significant high yield	12.56	23.84	61.37	51.82	26.07	24.34
High yield	.00	4.93	25.69	24.40	74.31	70.67

| | Bond Type | | | | | | | |
| | Government Agency | | Mortgage | | Corporate | | High Yield | |
	QPS	1988 Lipper	QPS	1988 Lipper	QPS	1988 Lipper	QPS	1988 Lipper
B. Government mortgage backed	14.03	22.55	54.54	77.45	25.97	.00	5.46	.00

for each of the six Morningstar groups (where the betas and proportions are scaled to add to one). An examination of panel A shows a very close correspondence between QPS betas and Lipper weights, except for the government mortgage funds. The lack of correspondence for the mortgage group is likely to be due to Lipper not disaggregating governments into mortgages and nonmortgage governments over the full period. However in 1988, Lipper split government bonds into mortgages and nonmortgage governments. For the government mortgage group, we reran the QPS as a four-index model by including the LB mortgage index and compared the resulting betas to the Lipper weights shown for December 1988.[25] The results are shown in panel B of table 7. There is much greater correspondence between QPS weights and Lipper proportions.

Table 8 shows the breakdown in proportions by maturity. The correspondence between QPS weights and Lipper weights is very good except for the government mortgage and high-yield categories. For both of these categories, the QPS weights show much shorter maturities. During this period, Lipper primarily classified mortgages and high-yield debt by their nominal maturities. Given expected prepayments, the *effective* maturity for mortgages is much shorter. Given default probability and the prepayment option almost always included in high-yield debt, it is not surprising that their return behavior is also more like a shorter term instrument.

Since we do not have individual fund duration measures for each of the various investment categories, we cannot test equation (4) directly. However, since we obtained Lipper proportions for two subdivisions of each maturity investment category, we can do an approximate test of equation (4). We formed the following two ratios:

$$\frac{X_{1-5i}}{X_{ij}} \tag{5}$$

and

$$\frac{X_{11-20i}}{X_{ik}}, \tag{6}$$

where X_{1-5i} is bond fund i's Lipper weight in the 1–5-year subcategory, X_{ij} is the fund's weight in the total intermediate-term category (1–10 years), X_{11-20i} is the fund's Lipper weight in the 11–20-year subcategory, and X_{ik} is the fund's weight in the total long-term category (greater than 11 years). The larger the equation (5) or (6), the more the fund invested in shorter duration bonds and the smaller should be the relative duration measure in equation (4). Thus, we should find an inverse relationship between (5) and β_{ij} divided by X_{ij}, and between

25. We also ran the above four-index QPS over just the last 2 years in our sample period (1987 and 1988) and obtained similar results.

TABLE 8   **Lipper and QPS Proportions by Bond Maturity (in %; Proportions Scaled to add to 100%)**

	Bond Maturity			
	1–10 Years		Greater than 10 Years	
Group	QPS	Lipper	QPS	Lipper
Government mortgage	80.1	23.5	19.9	76.5
High quality	32.9	37.2	67.1	62.8
Investment grade	46.1	43.5	53.9	56.5
General	40.1	34.5	59.9	65.5
General with significant high yield	21.1	33.1	78.9	66.9
High yield	59.5	25.4	40.5	74.6

(6) and β_{ik} divided by X_{ik}. To test this, we ran Spearman's rank correlation tests. For the intermediate-term category, the resulting rank correlation coefficient was -0.374, and for the long-term category it was -0.414, both significant at the .025 level.

In this section, we have shown that the past return series can be used to determine management's investment policy. While this is useful by itself, it should also increase the reader's confidence in the models employed in the prior analysis.

VII. Large-Sample Tests

Investment Company Data, Incorporated, provided us with total monthly return data on all bond funds that existed as of December 1991. We formed two samples. The first sample consisted of all funds that had at least 2 years of data. It contained returns from each fund's inception (or first listing by ICDI), or from January 1977 if the fund existed prior to that date, until December 1991. The second sample consisted of return data for the full 5-year period from January 1987 through December 1991 for all funds that had such data. The latter sample was formed to control for any time-series pattern in alphas. We eliminated funds in the ICDI data set that were not listed by Wiesenberger as having either a "bond" or "U.S. government securities" investment policy during any year in the two sample periods (which eliminated 61 and 36, respectively), funds not open to investment by the public (which eliminated 4 and 2, respectively), funds not listed in Wiesenberger (which eliminated 8 and 4, respectively), and one duplicate fund (from the since-inception sample). This left us with 361 funds in the since-inception sample and 223 funds in the 5-year sample.[26]

---

26. Differences between these numbers and those discussed in the introduction are due primarily to the exclusion of municipal bond funds from our sample (there were more than 500 municipal bond funds in 1990) and the exclusion of funds that do not have a minimum of 2 years of data.

In any data set, there is a possibility of inaccurate data points. To examine this, we compared the alphas for Market-1 and Reg-6 for the 33 funds that existed in both the ICDI data set and our own data set (which has been carefully screened for errors and is described in earlier sections). The average difference in alphas (ICDI minus our own) was -0.004% for Market-1 and -0.003% for Reg-6. Given ICDI's belief that the accuracy of their data improves as the data become more recent and the small differences over the earlier period in average alphas, we feel comfortable with the accuracy of the ICDI data base.

All of the empirical results in this section are reported for our two samples (since inception and 5 year) aggregately and for each of four Wiesenberger investment objective subclassifications as of December 1990, as reported in the 1991 edition of Wiesenberger's *Investment Companies*. These four subgroups are corporate bond, high-yield bond, government mortgage, and government securities.[27]

Given what we have learned from the first part of this study, we chose to explore performance in this section by employing four of our earlier models. We continue to employ Market-1 because it is the most widely used and simplest measure of performance. We learned from our earlier results that incorporating an index for high-yield bonds improved the results. Therefore, we also employ Own-1. Because of the importance of high-yield bonds and because we could not clearly see the impact of mortgage securities in our earlier sample (given that we only had three mortgage funds in that sample), we employ Risk-3 as our third model. Finally, we employ Reg-6 to examine whether term and risk premium effects are jointly important in affecting fund performance measures.

The results from applying these four models to the ICDI returns data are shown in table 9 for the 5-year sample and table 10 for the since-inception sample. The results are striking. No matter which model is applied to either sample, the average alphas for the overall samples (appearing in the rows labeled "All bond funds") and for each Wiesenberger subgroup are negative. Furthermore, except for Market-1, average alphas in the 5-year sample are statistically different from zero at better than a 0.01 level of significance.[28] Bond funds as a

27. Beginning with the 1991 edition of *Investment Companies*, Wiesenberger no longer classifies funds by investment policy; instead, the investment objective categories have been greatly expanded. We repeated all of the above analyses using Morningstar and ICDI classifications and obtained very similar results.

28. Market-1 produces nonsignificant alphas partly because it results in slightly less negative alphas but primarily because it results in very large increases in the standard error of alpha. This is not surprising since Market-1 does not fit the data (explain bond fund performance) nearly as well as the other models. We do not report t-statistics for the since-inception sample. The problem with that sample is that the length of data (the time series) varies across funds. Normal corrections for heteroscedasticity and cross correlation require the same time series for each fund. We tried several approximations to the normal tests by utilizing estimates of the correlations of residuals, variances of

whole and subclassifications of bond funds underperform any reasonable set of passive portfolios.[29] While examining these results, keep in mind that our performance estimates are upward biased. The samples constructed from ICDI data do not contain funds that ceased to exist (due to either liquidation, merger, or change in investment policy) prior to December 1991. We know on the basis of logic and from our empirical work in an earlier section of this article that funds that cease to exist tend to have poorer performance than those that survive.

Panel A of table 11 shows the average annual expense ratios and average maximum annual 12b-1 fees for 209 of the funds in our 5-year sample as of December 1989.[30] The 12b-1 fees are maximum annual fees; thus, total expenses are overstated if these fees are added directly to the expense ratios. The mortgage and government securities groups had average negative performance approximately equal to their average expense ratios. The corporate bond group had average performance slightly less negative than its average expense ratio. If these three groups had actually charged the maximum 12b-1 fees, they would have underperformed our benchmark portfolios by less than the amount of total expense ratios. Thus, on average, the underperformance of our sample of bond funds is primarily due to expenses. For the funds that survived, there is some weak evidence of slight overperformance pre-expenses for these three groups.[31] The average expense ratio and average maximum 12b-1 fee for the high-yield group are higher than for the other three groups. If the maximum 12b-1 fees were charged, it could account for the much poorer performance of that group.[32]

---

residuals, and means and variances of the independent variables from the time-series data that were available. All approximations resulted in standard errors for alpha that were very similar to those found in the 5-year sample. Since the average alphas are basically the same in both samples, the *t*-values are essentially unchanged in the 5-year and since-inception samples.

29. The fund returns are after expenses and management fees, while index returns are not. The Vanguard Bond Market Fund (a bond index fund) reports total fees of 16 basis points in their 1992 prospectus. On a monthly basis, this is a little over 1 basis point. Thus, even after adjusting for fees, active funds significantly underperform indexes.

30. Elton, Gruber, Das, and Hlavka (1992) have shown that relative expense ratios are fairly constant across funds (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 the middle year in a sample period. Thus, we believe 1989 expense ratios provide good expense measures for our 5-year sample.

31. In an earlier section, we found that 10 of 46 funds failed in a 10-year period. A conservative estimate of the effect of failure is to assume that only half would fail in a 5-year period. Failed funds had an average alpha about .204 less than the surviving funds. Including failed funds should lower average alphas by about 5/46 of .204, or about 27 basis points per year. Including failed funds means that average performance was approximately equal to or slightly worse than expenses for the above three groups.

32. To make sure the performance of the high-yield group was not a function of the index we selected, the tests were rerun using the Salomon Brothers long-term high-yield index in place of the Blume/Keim index. The results were virtually unchanged.

TABLE 9 Regression Results for ICDI Funds: ICDI funds from January 1987 through December 1991 (5 years)—223 funds (Continuously Compounded Monthly Excess Returns)

A. Market-1 Model
Return = Constant + Government/Corporate

Fund Objective	Average Alpha (%)	Average Beta	Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
						Positive	Negative
Corporate	−.033	.89	.81	49	34	2	3
High yield	−.129	.45	.08	42	31	0	1
Government mortgage	−.020	.91	.88	30	19	0	2
Government securities	−.096	.94	.87	96	68	3	35
All bond funds*	−.077	.87	.71	223	157	6	43
	(−1.245)						

B. Own-1 Model
Corporate or Unlisted: Return = Constant + Government/Corporate
High Yield: Return = Constant + High Yield
Government Mortgage: Return = Constant + Mortgage
Government Securities: Return = Constant + Government

Fund Objective	Average Alpha (%)	Average Beta	Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
						Positive	Negative
Corporate	−.033	.89	.81	49	34	2	3
High yield	−.160	.73	.74	42	31	0	9
Government mortgage	−.123	.96	.90	30	30	0	17
Government securities	−.081	1.05	.86	96	63	3	29
All bond funds*	−.090	.93	.84	223	163	6	60
	(−3.740)						

C. Risk-3 Model

Return = Constant + Government/Corporate + Mortgage + High Yield

| Fund Objective | Average Alpha (%) | Average Beta | | | Average Adjusted R^2 | No. of Funds | No. of Negative Alphas | No. of Significant Alphas | |
		Government/ Corporate	Mortgage	High Yield				Positive	Negative
Corporate	−.045	.76	.10	.07	.85	49	34	2	7
High yield	−.153	−.13	.06	.74	.76	42	31	0	8
Government mortgage	−.090	.33	.62	.01	.92	30	29	0	15
Government securities	−.107	.95	.10	.00	.89	96	80	2	39
All bond funds*	−.099 (−3.859)	.61	.16	.16	.86	223	179	5	72

D. Reg-6 Model

Return = Constant + Intermediate Government + Long-Term Government + Mortgage + High Yield + Intermediate Corporate + Long-Term Corporate

| Fund Objective | Average Alpha (%) | Average Beta | | | | | | Average Adjusted R^2 | No. of Funds | No. of Negative Alphas | No. of Significant Alphas | |
		Intermediate Government	Long-Term Government	Mortgage	High Yield	Intermediate Corporate	Long-Term Corporate				Positive	Negative
Corporate	−.051	.24	.10	.05	.05	.26	.20	.87	49	38	1	14
High yield	−.225	−1.11	−.06	−.16	.63	1.81	−.20	.78	42	38	0	11
Government mortgage	−.098	.40	−.07	.53	.00	−.16	.28	.92	30	30	0	15
Government securities	−.089	.44	.22	.17	.01	−.03	.14	.90	96	80	2	42
All bond funds*	−.107 (−4.171)	.10	.10	.13	.13	.36	.11	.87	223	191	4	85

NOTE.—Significance level for alphas is .05 (two tail). t-values for average alphas across all bond funds in the sample appear in parentheses and reflect adjustment for both heteroscedasticity and cross correlation of residuals.

* Figures in rows labeled "All bond funds" include funds in the Investment Company Data, Incorporated (ICDI) sample that are not classified by Wiesenberger; six funds are not classified by Wiesenberger in the 5-year sample.

TABLE 10   Regression Results for ICDI Funds: ICDI funds since inception or January 1977—361 funds (Continuously Compounded Monthly Excess Returns)

A. Market-1 Model
Return = Constant + Government/Corporate

Fund Objective	Average Alpha (%)	Average Beta	Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
						Positive	Negative
Corporate	−.033	.85	.82	92	60	3	10
High yield	−.053	.59	.21	54	30	0	1
Government mortgage	−.011	.77	.78	40	21	2	3
Government securities	−.095	.94	.82	162	113	6	43
All bond funds*	−.064	.84	.72	361	234	13	58

B. Own-1 Model
Corporate or Unlisted: Return = Constant + Government/Corporate
High Yield: Return = Constant + High Yield
Government Mortgage: Return = Constant + Mortgage
Government Securities: Return = Constant + Government

Fund Objective	Average Alpha (%)	Average Beta	Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
						Positive	Negative
Corporate	−.033	.85	.82	92	60	3	10
High yield	−.065	.73	.74	54	34	4	3
Government mortgage	−.104	.84	.84	40	36	1	20
Government securities	−.079	.94	.81	162	104	8	35
All bond funds*	−.069	.87	.81	361	244	18	69

C. Risk-3 Model

Return = Constant + Government/Corporate + Mortgage + High Yield

Fund Objective	Average Alpha (%)	Average Beta			Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
		Government/ Corporate	Mortgage	High Yield				Positive	Negative
Corporate	-.036	.77	.04	.05	.84	92	60	3	18
High yield	-.066	-.04	.05	.71	.75	54	38	2	3
Government mortgage	-.088	.19	.65	.01	.85	40	35	1	17
Government securities	-.110	.82	.14	.00	.84	162	131	3	61
All bond funds*	-.083	.60	.16	.12	.83	361	275	12	102

D. Reg-6 Model

Return = Constant + Intermediate Government + Long-Term Government + Mortgage + High Yield + Intermediate Corporate + Long-Term Corporate

Fund Objective	Average Alpha (%)	Average Beta						Average Adjusted R^2	No. of Funds	No. of Negative Alphas	No. of Significant Alphas	
		Intermediate Government	Long-Term Government	Mortgage	High Yield	Intermediate Corporate	Long-Term Corporate				Positive	Negative
Corporate	-.041	.27	.14	-.01	.03	.27	.14	.86	92	65	3	17
High yield	-.116	-.86	-.08	-.15	.61	1.26	.06	.78	54	41	1	11
Government mortgage	-.102	.29	-.07	.53	.00	-.01	.18	.85	40	37	1	22
Government securities	-.093	.46	.17	.16	.00	-.03	.12	.86	162	134	0	63
All bond funds*	-.101	.20	.04	.10	.08	.21	.23	.83	361	288	6	118

NOTE.—Significance level for alphas is .05 (two tail).

* Figures in rows labeled "All bond funds" include funds in the Investment Company Data, Incorporated (ICDI) sample that are not classified by Wiesenberger; 13 funds are not classified by Wiesenberger in the since-inception sample.

TABLE 11   Management Fees (5-Year Sample—209 Funds)
                  A. Annual Management Fees as of December 1989 (in 1990
                  Wiesenberger)

Fund Objective	Average Expense Ratio (%)	Average Maximum 12b-1 Fee (%)
Corporate	.882	.130
High yield	1.147	.257
Govt. mortgage	1.062	.166
Govt. securities	1.018	.165
All bond funds	1.027	.179

B. Alpha Regressed on Expense Ratio

Model	Average Intercept (%)	t Average Intercept	Average Slope	t Average Slope
Market-1	−.005	−.124	−.834	−1.953
Own-1	−.012	−.497	−.891	−3.063
Risk-3	−.024	−1.213	−.859	−2.952
Reg-6	−.014	−.843	−1.086	−4.160

Note.—Only 209 of 223 funds in the 5-year Investment Company Data, Incorporated (ICDI) sample were listed by Wiesenberger in 1990. Averages are calculated from 60 monthly time-series regressions.

Panel B of table 11 shows the results of regressing alphas on expense ratios. The coefficients of the expense ratio for all four models are close to negative one, indicating that a percentage-point increase in expenses reduces performance by about 1 percentage point. Thus, on average, an investor is better off selecting a low-expense fund. After accounting for expenses but not 12b-1 fees, the intercepts are negative but very close to zero. Expenses account for the major portion of the amount by which mutual funds underperform passive bond portfolios.[33]

Before leaving this discussion on overall performance, a few additional remarks are in order. First, examining the average betas in each of the four groups (in table 9 and table 10) shows that the weightings

33. Note in table 11, panel B, that the intercepts are not significantly different from zero while the slope coefficients are all significantly different from zero at the 1% level, except for the slope for Market-1 (which is almost significant at the 5% level). The statistical results reported in this table were derived using a series of monthly cross-sectional regressions, similar to the methodology employed by Fama and MacBeth (1973). The relationship between alphas and expenses was also examined using ordinary least squares and weighted least squares. The regression coefficients were substantially unchanged, but the t-values of the slope coefficients were larger negative numbers under either alternative methodology. We also ran the regression including maximum 12b-1 fees. The slopes are flatter when those 12b-1 fees are added to expense ratios, while the intercepts are lower. This is probably due to the fact that those 12b-1 fees are maximums that are not always charged.

on the indexes strongly reflect the group membership. For example, consider the three-index model in table 9. For the mortgage group, the beta on the mortgage index is .62, the beta on the government/corporate index is .33, and the beta on the high-yield index is .01. For the government securities group, the beta is .95 on the government/corporate index, .10 on the mortgage index, and 0 on the high-yield index. The same reasonableness in pattern is true for the other groups and the six-index model.

In earlier sections of this article, we have shown that performance measurement was robust across models, after we accounted for high-yield bonds. We now repeat this analysis recognizing that, with the larger samples, it is also necessary to account for mortgages. (There were only three mortgage funds in our smaller sample.) Table 12 presents comparisons of relative performance results for the four models used in this section applied to the 5-year sample. Compare the performance classifications of Risk-3 and Reg-6. Reg-6 classifies 32 funds as having positive alphas, and Risk-3 has 28 of those same 32 funds classified as positive. In addition, seven of the top 10 are common to the two models. Thus, an investor using either model would come to similar conclusions regarding which funds are the top performers.

As shown in table 12, Market-1 and to a lesser extent Own-1 classify many more funds as having positive alphas than do the multi-index models. The vast majority of the funds classified as positive by Risk-3 and Reg-6 are also classified as positive by Market-1 and Own-1. Market-1 and Own-1 miss four out of 44 compared to Risk-3, and they miss five or six out of 32 compared to Reg-6. Ten of the top 15 ranked funds according to Reg-6 are also in the top 15 according to Own-1. Finally, Market-1 has nine of the top 15 according to Reg-6 also in the top 15. The similarity in ranking between Reg-6 and Risk-3 is greater than that between Reg-6 and Own-1, which is in turn greater than that between Reg-6 and Market-1.

Table 13 examines how well the alpha calculated over the first 3 years of the 5-year sample predicts the alpha calculated over the subsequent 2 years, and it shows significantly positive rank correlations for most of our models and subclassifications. However, Brown, Goetzmann, Ibbotson, and Ross (1992) have shown how selection bias induces a correlation between past and future performance measures, and this could explain our results. Table 14 shows how the best and worst performing funds did in the subsequent period, where the best or worst funds were the top or bottom 10% or five funds, which ever was larger. The higher performance of the upper tail could be explained by high-performing funds in the first subperiod that did poorly in the second subperiod not being included in the sample due to survivorship bias. However, the persistence of underperformance in the lower tail

TABLE 12   Comparison of Performance Classification across Models: Number of Positive and Negative Alphas (5-Year ICDI Sample)

| | Joint Ranking | | | | | | | |
| | Market-1 | | Own-1 | | Risk-3 | | Reg-6 | |
	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
Market-1:								
Positive	66							
Negative		157						
Own-1:								
Positive	55	5	60					
Negative	11	152		163				
Risk-3:								
Positive	40	4	40	4	44			
Negative	26	153	20	159		179		
Reg-6:								
Positive	26	6	27	5	28	4	32	
Negative	40	151	33	158	16	175		191

NOTE.—Each number in the table indicates how many of the funds that are classified as positive or negative by the model in the corresponding column are classified as having the same or opposite sign by the model in the corresponding row. ICDI = Investment Company Data, Incorporated.

TABLE 13   Spearman Rank Correlations for ICDI Alphas: 3-Year Subperiod 1 and 2-year Subperiod 2

Group	N	Cutoff	Market-1 Correlation	Own-1 Correlation	Risk-3 Correlation	Reg-6 Correlation
All bond funds*	223	.132	.451	.445	.212	.403
Corporate	49	.283	−.085	−.085	.186	−.099
High yield	42	.306	.387	.449	.391	.444
Government securities	96	.201	.618	.582	.181	.401
Government mortgage	30	.362	.582	.547	.326	.297

NOTE.—Cutoff is for rejecting null of zero correlation at significance level of .025 (one tail). ICDI = Investment Company Data, Incorporated.

* Figures in the row labeled "All bond funds" include the six funds not classified by Wiesenberger in the 5-year sample.

TABLE 14   5-Year ICDI Sample: Alphas Ranked in First 3-Year Subperiod; Averages Calculated in Second 2-Year Subperiod: Best and Worst 10%

	Best	Worst	Average
Overall (22 and 22):			
Market-1	.024	−.384	−.068
Own-1	.034	−.371	−.091
Risk-3	−.145	−.247	−.093
Reg-6	−.011	−.428	−.117
Corporate (5 and 5):			
Market-1	−.091	−.146	−.035
Own-1	−.091	−.146	−.035
Risk-3	−.105	−.134	−.037
Reg-6	−.026	−.032	−.047
High yield (5 and 5):			
Market-1	.127	−.422	−.112
Own-1	.081	−.540	−.163
Risk-3	.034	−.452	−.102
Reg-6	−.010	−1.029	−.386
Government (10 and 10):			
Market-1	.056	−.505	−.104
Own-1	.031	−.480	−.089
Risk-3	−.292	−.256	−.136
Reg-6	.001	−.090	−.068
Mortgage (5 and 5):			
Market-1	.116	.006	.049
Own-1	.003	−.133	−.098
Risk-3	.017	−.090	−.040
Reg-6	.060	−.082	−.026

NOTE.—ICDI = Investment Company Data, Incorporated.

cannot be explained by survivorship bias.[34] Given the lack of persistence in our unbiased sample in the earlier sections of this article and the Brown et al. (1992) results, we must be cautious in drawing conclusions. A more detailed analysis must await a subsequent study.

VIII. Conclusion

Bond mutual funds are a major investment vehicle, and yet there has been no serious analysis of their performance heretofore. This study represents the first major analysis of the performance of this important investment vehicle.

Overall and for subcategories of bond funds, we found that bond funds underperformed relevant indexes. For most models and fund subgroups, this underperformance was approximately equal to the average management fees, indicating that, pre-expenses, the surviving funds performed about on par with the indexes.

The regressions of alphas on expense ratios for the ICDI sample indicated that, on average, a percentage-point increase in expense leads to a percentage-point decrease in returns, so that investors without forecasting ability should select low-expense funds.

There appear to be fewer available bond index mutual funds than stock index mutual funds. The results in our article strongly indicate that the introduction of more bond index funds would be very useful to individual investors. Individual investors are likely to pay large bid-ask spreads for purchases of small numbers of bonds. Thus, the lack of available index funds for individual investors, coupled with the high transactions costs that accompany small purchases, might account for the past appeal of actively managed bond funds, despite substantial underperformance.

Our results were robust across a wide choice of models. In the small sample, the inclusion of the high-yield index was sufficient to result in models with similar results. In the larger samples of the previous section, it was necessary to include a separate mortgage index as well as a high-yield index. It did not affect ranking to allow for a term or corporate-risk (other than high-yield) effect. This is in contrast to the common stock area, where the choice of factors crucially affects performance.

We could find no evidence of predictability using past performance to predict future performance for our unbiased sample. When we examined this issue for the larger biased samples, we found some evi-

---

34. It might be due to either funds with higher fees not performing significantly better or worse (before fees) than funds with low fees or bad managers continuing to remain bad managers.

dence of predictability. However, whether this is due to bias or to the larger sample sizes must await further study.

Finally, we showed that our QPS procedure produced weights on the passive portfolios that closely matched the actual fund holdings in their respective investment categories. This provides both confidence in our procedure and a method of determining management policy from the fund returns themselves.

References

Blume, M. E., and Keim, D. B. 1987. Lower grade bonds: Their risks and returns. *Financial Analysts Review* (July–August), pp. 26–33.

Blume, M. E.; Keim, D. B.; and Patel, S. A. 1991. Returns and volatility of low-grade bonds, 1977–1989. *Journal of Finance* 46, no. 1 (March): 49–74.

Brown, S.; Goetzmann, W.; Ibbotson, R.; and Ross, S. 1992. Survivorship bias in performance studies. *Review of Financial Studies* 5, no. 4:553–80.

Connor, G., and Korajczyk, R. A. 1991. The attributes, behavior, and performance of U.S. mutual funds. *Review of Quantitative Finance and Accounting* 1:5–26.

Cornell, B., and Green, K. 1991. The investment performance of low-grade bond funds. *Journal of Finance* 46, no. 1 (March): 29–48.

Cumby, R. E., and Glen, J. D. 1990. Evaluating the performance of international mutual funds. *Journal of Finance* 45, no. 2 (June): 497–521.

Elton, E. J., and Gruber, M. J. 1991. *Modern Portfolio Theory and Investment Analysis*. 4th ed. New York: John Wiley & Sons.

Elton, E. J.; Gruber, M. J.; Das, S.; and Hlavka, M. 1992, in press. Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*.

Elton, E. J.; Gruber, M. J.; and Nabar, P. 1988. Bond returns, immunization, and the return-generating process. In M. Sarnat and G. Szego (eds.), *Studies in Banking and Finance, Essays in Memory of Irvin Friend*. New York: North Holland.

Fama, E., and MacBeth, J. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 71 (May/June): 607–36.

Grinblatt, M., and Titman, S. 1989a. Mutual fund performance: An analysis of quarterly portfolio holdings. *Journal of Business* 62, no. 3:393–416.

Grinblatt, M., and Titman, S. 1989b. Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies* 2:393–421.

Jensen, M. 1968. The performance of mutual funds in the period 1945–1964. *Journal of Finance* 23, no. 2 (May): 389–416.

Lehmann, B. N., and Modest, D. 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance* 42, no. 2 (June): 233–65.

Lipper Analytical Securities Corporation. 1978–1988. *Lipper—Portfolio Analysis Report on Fixed Income Funds*. New York: Lipper Analytical Services, Inc.

Morningstar, Inc. 1989. *Mutual Fund Sourcebook, Winter 1989*. Chicago: Morningstar, Inc.

Roll, R., and Ross, S. 1980. An empirical investigation of the arbitrage pricing theory. *Journal of Finance* 35, no. 5 (December): 1073–1103.

Sharpe, W. F. 1966. Mutual fund performance. *Journal of Business* 39, no. 1 (January): 119–38.

Sharpe, W. F. 1992. Asset allocation, management style, and performance measurement. *Journal of Portfolio Management* (Winter), pp. 7–19.

Treynor, J. 1965. How to rate management investment funds. *Harvard Business Review* 44, no. 1 (January/February): 63–75.

Wiesenberger Financial Service. 1979–1991. *Investment Companies*. Boston: Warren, Gorham & Lamont, Inc.