

Bond Fund Returns and Expenses: A Study of Bond Market Efficiency

WILLIAM REICHENSTEIN

WILLIAM REICHENSTEIN holds the Pat and Thomas R. Powers Chair in Investment Management at Baylor University.

Peter Lynch, Warren Buffett, and John Templeton, to name a few, are stock fund managers who have produced long-term records of superior performance. Can you name a bond fund manager who has produced a long-term record of superior performance?

Don't fret. You are not alone!

There are exceptionally bright bond managers. The problem is the nature of their quest. To consistently beat their peers — that is, the managers who invest in bonds with similar duration and quality — they must be able to consistently find undervalued bonds. The bond market appears to be too efficient to allow someone to consistently beat his or her peers. Perhaps a more accurate statement is that the margin of superiority for even the best of managers is too thin to offset the burden of a high expense ratio. If bond markets are efficient, investors should opt for a low-cost fund, because the burden of high costs will likely exceed the value a manager can add through superior security selection.

The goal of this article is to study 1994–1998 bond fund returns to see how closely they come to the predictions of the efficient markets hypothesis. The efficient markets hypothesis would predict:

1. There is a strong, negative relation between expense ratio and net return for bond funds with the same investment style.

2. The expense ratio is a deadweight loss. A 1% increase in the expense ratio decreases net return by 1%.
3. Loads are a deadweight loss.
4. Among funds with the same style, lower-cost bond funds consistently produce good relative returns compared to higher-cost funds.

MORNINGSTAR'S BOND STYLE BOXES

Morningstar Mutual Funds separates fixed-income funds into a three-by-three style grid based on the portfolio's duration and quality as of a recent date. Exhibit 1 illustrates the style box. The numbers denote the final numbers in each style-box sample.

Duration measures the portfolio's price sensitivity to changes in interest rates. It is

EXHIBIT 1 Morningstar Style Boxes

		Duration			Quality
Short	Intermediate	Long			
137	172	26	High		
19	23	13	Med		
	40		Low		

related to maturity. Prices of long-duration (maturity) bonds rise more when interest rates fall and fall more when interest rates rise than prices of short-duration (maturity) bonds. A duration of seven years means that the portfolio will lose about 7% of its value if interest rates rise 1 percentage point, and gain about 7% if rates fall 1 percentage point. The short, intermediate, and long portfolios have durations of three years or less, between 3.1 and 6.0 years, and more than six years.

Quality measures the average bond rating as assigned by Standard & Poor's or Moody's. Morningstar classifies funds with average ratings of AAA or AA as high-quality bond funds. Medium-quality and low-quality bond funds have average ratings of, respectively, A or BBB and BB or lower.

THE STUDY

A key issue examined in this study is whether differences in expense ratios can explain differences in net returns for same-style bond funds. To ensure similarity in style (or at least minimize differences in style), bond funds are separated into samples on the basis of their style as reported in the January 1999 version of "Morningstar Principal Pro Plus for Mutual Funds." Each sample consists of all funds within a given Morningstar style-box that has at least five years of returns as of January 1999. To assure that the funds are truly bond funds, a fund is removed if it has more than a minor amount of U.S. stocks, international stocks, or unspecified "other" assets.¹

In practice, there are style differences among funds within each Morningstar style-box sample. First, there is a range within each style dimension. For example, the final high-grade, long-term sample (henceforth, high-long) includes funds with durations of 7.1 and 14.5 years, yet both are classified as long-term funds. Similarly, Morningstar could classify two funds as medium-grade when one fund's average rating is about halfway between AA and A, and the other's is about halfway between BB and B.

Second, style differences sometimes exist between two funds with the same average rating. For example, as of year-end 1997, the Merrill Lynch Corporate Investment Grade and Putnam Income funds had average credit quality ratings of A. The Merrill Lynch fund focuses on A-rated securities, while the Putnam fund holds bonds rated AAA to BBB or lower.

Third, a manager may adjust the portfolio's style. For example, Wasatch-Hoisington U.S. Treasury gener-

ally holds long-term bonds, but it moved to short-term bonds for most of 1995 in anticipation of rising interest rates. Its 1994-1998 return will differ from the returns on funds that consistently hold long-term bonds due to this interest rate bet.

In short, although Morningstar separates funds by style, style differences remain within each style-box sample. So, even if bond markets are perfectly efficient, differences in expenses should not be able to explain all differences in net returns.

To see how well expenses can explain returns within the style-box samples, we want to exclude funds with substantially different styles. A fund should be removed only if there is a clear, substantial style difference between the fund and other funds in the sample. This study removed three funds — less than 1% of all funds. It removed American Century-Benham Target Maturities Trust 2020 Portfolio from the high-long sample. This fund seeks to replicate the performance of the November 15, 2020, zero-coupon Treasury strip. During 1994-1998, its duration ranged from almost twenty-seven to twenty-two years. A duration of twenty-seven years is almost twice as high as the highest duration in the final high-long sample and three times as high as the average fund's duration. Other deletions are Fidelity Short-Term Bond and Fidelity Spartan Short-Term Bond from the medium-short sample. Both funds suffered large losses in 1994 when then-manager Donald G. Taylor held up to 45% of assets in foreign bonds, including emerging market bonds. The funds' dismal five-year records are almost completely due to this important style difference.

ANALYSIS OF EFFICIENT MARKETS PREDICTION

Prediction 1: *There is a strong, negative relation between expenses and net returns among funds with the same investment style.* Exhibit 2 presents the average expense ratio and the average net return for one-year, three-year, and five-year horizons ending 1998 for the low-cost, middle-cost, and high-cost groups of each sample.² Henceforth, this is called the averaged data.

The data strongly support the first prediction. For example, look at the five-year average returns in the high-short sample. Compared to the low-cost funds, the medium-cost funds have 0.31% higher expenses and 0.30% lower net returns. The high-cost funds have 0.94% higher expenses than low-cost funds and 0.95% lower returns.

EXHIBIT 2

Net Returns versus Expenses (%)

Fund Style	Cost	1998 Expense Ratio	5-Year Average Return	3-Year Average Return	1-Year Average Return
High-Grade Short-Term Funds	Low	0.40	5.77	5.98	6.40
	Medium	0.71	5.47	5.69	6.08
	High	1.34	4.82	5.18	5.57
High-Grade Intermediate-Term Funds	Low	0.51	6.78	6.89	8.32
	Medium	0.86	6.30	6.42	7.86
	High	1.34	5.81	5.88	7.38
High-Grade Long-Term Funds	Low	0.48	8.41	8.03	11.41
	Medium	0.82	7.54	7.74	10.57
	High	1.34	6.35	6.37	8.48
Medium-Grade Short-Term Funds	Low	0.55	5.55	5.90	6.15
	Medium	0.84	5.32	5.52	5.56
	High	1.20	5.16	5.31	5.40
Medium-Grade Intermediate-Term Funds	Low	0.62	6.69	6.61	7.26
	Medium	1.10	6.30	6.31	5.89
	High	1.81	5.67	5.97	5.42
Medium-Grade Long-Term Funds	Low	0.70	6.67	7.10	7.98
	Medium	1.20	6.43	6.75	4.03
	High	1.70	5.92	6.35	1.67
Low-Grade Intermediate-Term Funds	Low	0.73	8.04	8.71	0.92
	Medium	1.15	7.28	7.91	-0.18
	High	1.69	6.42	7.47	-1.82

Differences in expenses explain virtually all differences in average returns. For all seven samples and all three investment horizons, average net return decreases as we go from low-cost to middle-cost groups. For all seven samples and all three horizons, average net return decreases as we go from middle-cost to high-cost groups. In total, higher expenses correctly predict lower return forty-two out of forty-two times.

This strong relationship is not unique to the 1994-1998 period. Using averaged data, Clements [1991, 1999] and Bogle [1994] find similar results. Clements [1991] reports the results of a similar study conducted for the *Wall Street Journal* by Morningstar. Morningstar separated bond funds into five sectors: government-backed mort-

gage bonds, corporate bonds, Treasury bonds, general municipal bonds, and high-yield bonds. It then separated funds in each sector into low-cost, middle-cost, and high-cost groups. Finally, it calculated the groups' average returns for one-year, three-year, and five-year periods ending in February 1991.

Except in two cases, low-cost funds always beat middle-cost funds, which always beat high-cost funds; for the one- and three-year periods, middle-cost junk bonds edged out low-cost funds. Higher expenses predict relative returns twenty-eight of thirty times.

Bogle examines 1990-1992 returns across groups of government, corporate, and municipal bond funds separated by maturity and bond grade. He compares aver-

age net returns across funds with expense ratios of less than 0.5%, 0.5 to 1.0%, 1.01 to 1.5%, and more than 1.5%. In total, higher expenses predict relative returns twenty-four of twenty-four times.

Clements [1999] presents the results of an updated study conducted by Morningstar that includes five bond sectors: high-yield, short bond/government, intermediate bond/government, national municipal long, and national municipal intermediate. Morningstar calculated average five-year net return ending year-end 1998 for four cost groups within each sector. Relative returns always decreased systematically as average expenses increased. In total, higher expenses predict relative returns fifteen of fifteen times.

In total, across this study and the three others, average expenses predict relative returns 109 of 111 times.

The averaged data consistently imply that expenses almost perfectly predict relative returns. Averaged data, of course, may obscure the superior performance of individual funds. If we average Fidelity Magellan's superior performance under Peter Lynch with enough funds, Magellan's performance would appear decidedly average. To avoid this criticism, Exhibits 3 through 9 present the five-year performance of each fund in each sample.

Consider Exhibit 3, which shows the high-short sample. Each dot represents the 1994-1998 annualized net return and the 1998 expense ratio for one of the 137 bond funds. The three boxes denote average net return and average expense ratio for, respectively, the low-cost group, middle-cost group, and high-cost group of funds. The trend line reflects the best-fit line through the points. The "slope" indicates that, on average, net return decreases by 0.83% for each percentage point rise in expenses. The R^2 indicates that differences in expense ratios explain 38% of differences in net returns.

The graphs support the first prediction. In addition, the slope coefficients in Exhibit 10 support this prediction. The first three rows of Exhibit 10 present the slope coefficients from regressions that combine all seven samples for each horizon. Exhibit 10 also presents the slopes from regressions for all seven samples and all three horizons. The slope coefficient is negative in all twenty-four regressions. The slope is statistically negative in twenty-three of twenty-four regressions, meaning that the probability of the observed negative relation occurring by chance alone is less than 5% — the significance level we use throughout.³

Prediction 2: *Expenses are a deadweight loss.* This is a much stronger prediction than the first. It says the

slope coefficient must be statistically indistinguishable from -1.0. Not only must expenses reduce net returns, but also there must be a one-to-one relation between expenses and net returns.

In the five-year combined regression in Exhibit 10, the slope is -1.08, which is statistically indistinguishable from -1.0. The slopes for the one-year and three-year combined regressions are -1.42 and -0.92, respectively. The -0.92 is statistically indistinguishable from -1.0. Statistically, the -1.42 deviates too much from -1.0 to be attributed to chance.

So, although the -1.42 slope does not support the second prediction, it certainly does not suggest that investors benefit from higher expenses or that bond managers can overcome the burden of higher expenses. Rather it implies that, on average for 1998, a 1 percentage point higher expense ratio lowered net returns by more than 1%.

In eighteen of twenty-one regressions, the slope is statistically indistinguishable from -1.0. In only one case is the slope less than 1.0 in absolute value, which suggests that a 1 percentage point higher expense ratio reduces net returns but by less than 1%.

Most of the evidence supports the second prediction. Twenty of twenty-four regressions support a one-to-one negative relation between expense ratio and net return. In three of the four contrary regressions, expenses drag down net return more than one-to-one. So, in twenty-three of twenty-four regressions expenses are, at best, a deadweight loss.

Prediction 3: *Sales load charges are a deadweight loss.* A sales load can be assessed at purchase (front loads) or at redemption (deferred load or redemption fees). To test this prediction, this study compares the average net return on load funds to the average net return on no-load funds. It makes this comparison for each of the seven samples and for each of the three horizons.

Net returns do not reflect sales loads. Brokers (or investment advisors, or account executives) share the loads with their firm. They like to claim that load funds produce higher net returns. If so, higher annual net returns can, through time, more than offset a one-time sales load. "You get what you pay for," is the broker's refrain.

Exhibit 11 presents the average net returns. They indicate that load funds do not offer higher net returns. Rather, *gross returns* — the sum of net return plus expense ratio — are essentially the same across load and no-load funds. Load funds usually have higher expense ratios than no-load funds, and higher expenses produce lower net returns.⁴

EXHIBIT 3 High-Grade Short-Term Bond Funds

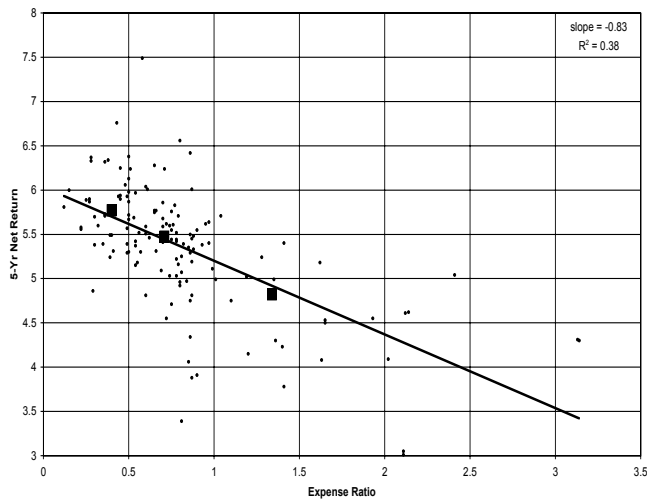


EXHIBIT 4 High-Grade Intermediate-Term Bond Funds

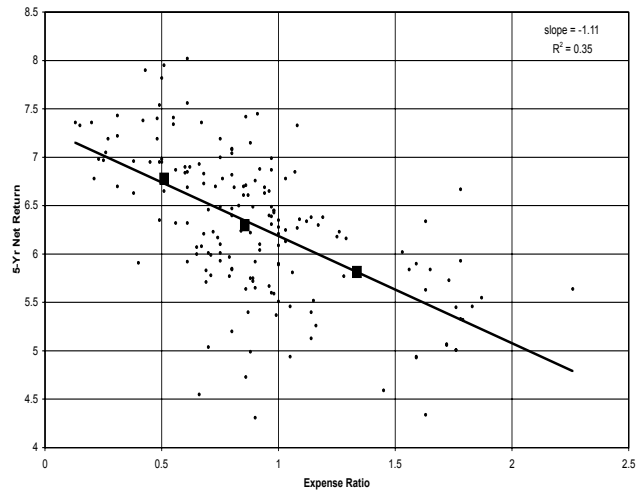


EXHIBIT 5 High-Grade Long-Term Bond Funds

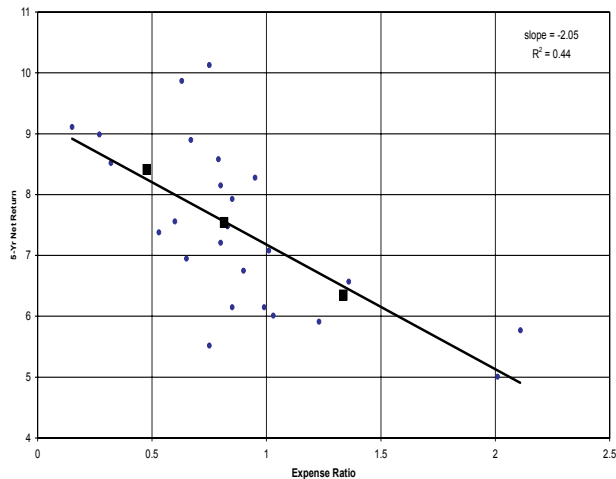
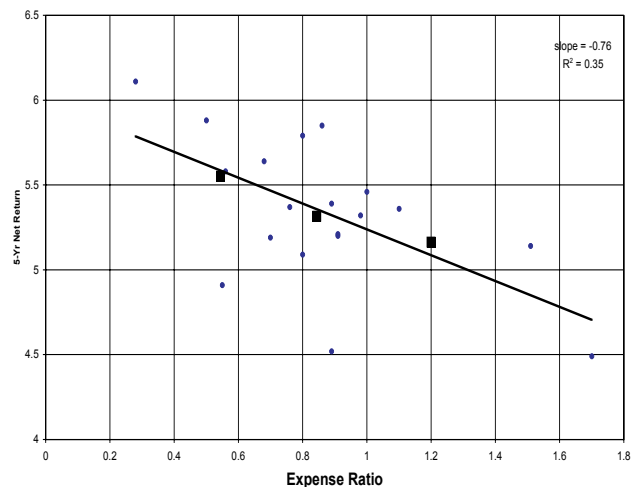


EXHIBIT 6 Medium-Grade Short-Term Bond Funds



Nothing in Exhibit 11 supports the refrain, “You get what you pay for.” Rather, the results suggest that investors who are not cost-conscious will likely get nicked for a load and further nicked each year with higher expenses. For example, load funds in the high-short sample have an average front-plus-redemption load of 2.99%, and they charge an average 66 percentage points more in annual expenses. Expenses, whether in the form of a load or annual expense ratio, are deadweight losses.

Prediction 4: *Among funds with the same style, low-cost bond funds consistently produce good relative returns.* There is a difference between the two statements: “Among funds with the *same style*, low-cost bond funds consistently produce good relative returns” and “Among funds *within the same style-based sample*, low-cost bond funds consistently produce good relative returns.” This study examines the second statement, and the evidence supports it for a five-year horizon but not for short horizons.

EXHIBIT 7 Medium-Grade Intermediate-Term Bond Funds

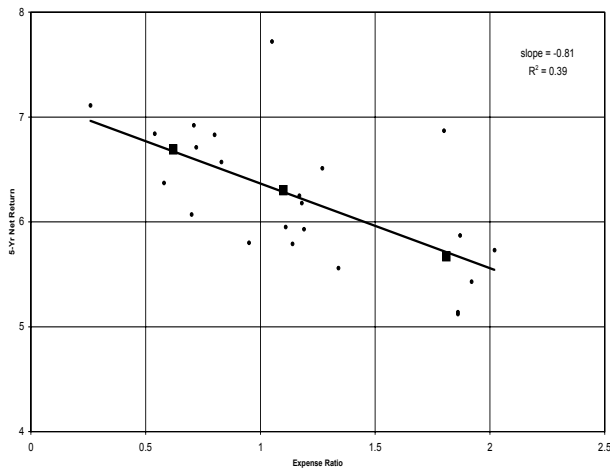


EXHIBIT 9 Low-Grade Intermediate-Term Bond Funds

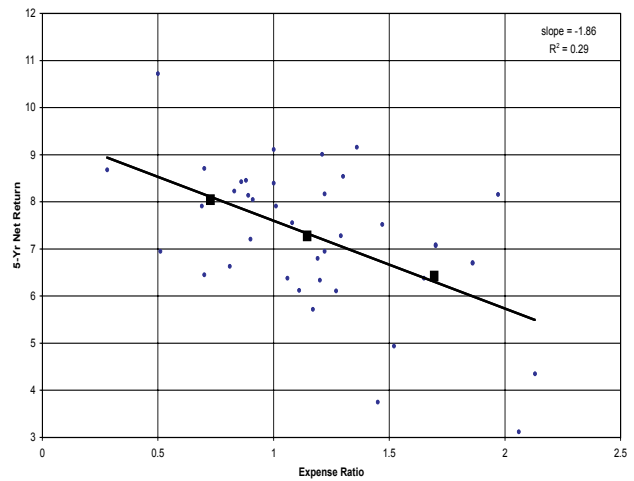
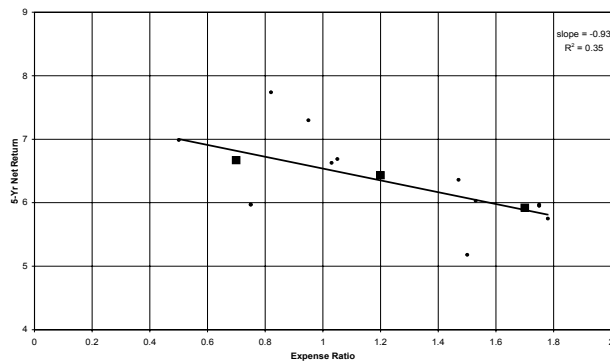


EXHIBIT 8 Medium-Grade Long-Term Bond Funds



Exhibits 3 through 9 provide the bird's-eye evidence supporting the second statement for a five-year horizon. Exhibit 12 presents additional supporting evidence. It presents the average rank of the lowest-cost and three lowest-cost funds for one-, three-, and five-year investment horizons.

For example, the lowest-cost fund in the high-short sample produced the thirty-second best five-year return of 137 funds. The three lowest-cost funds produced an average rank of thirty-fifth for the five-year horizon. For the five-year horizon, the lowest-cost fund produced a top-quartile return in all seven samples. This demonstrates the long-term advantage of buying low-cost bond funds.

Among funds *within the same style-based sample*, low-cost funds do not always produce good relative returns for short horizons. For example, the Vanguard High-Yield Corporate fund has the lowest expense ratio in the low-intermediate sample. Its net return ranked it thirty-eighth of forty funds in 1996 and twenty-fourth in 1997. Yet, it ranked sixth for the five-year period. Vanguard Admiral Long-Term Treasury has the lowest costs in the high-long sample. Its net return ranked it twenty-third of twenty-six funds in 1996, but third for the five-year period.

For short horizons, differences in within-sample styles can more than offset differences in expenses. When interest rates fall, a higher-cost fund with a relatively long duration within that sample will likely beat a lower-cost fund with a shorter duration; the additional return from the longer duration more than offsets the burden of higher expenses. Similarly, for a short horizon, within-sample differences in quality can more than offset differences in expenses.

Differences in return due to differences in style are positive in some years and negative in others, while a low-cost advantage provides a return enhancement year after year. Thus the longer the horizon, the harder it is for a higher-cost fund to beat a lower-cost fund with a similar style. Among funds with similar styles, differences in longer-term returns are primarily attributable to differences in costs.

There is a widely held belief that mutual fund returns show little persistence. That is, a mutual fund with an above-average return in one period is almost as likely as not to produce a below-average return in the next period. Among others, Malkiel [1995] and Bogle [1994]

EXHIBIT 10

Summary of Regressions of Net Returns on Expense Ratio

Sample	Investment Horizon	Slope	Slope Different from Zero?	Slope Different from -1.0?	(Closeness of Fit) R ²
Combined	5 yr	-1.08	yes	no	0.57
	3 yr	-0.92	yes	no	0.59
	1 yr	-1.42	yes	yes	0.72
High-Short	5 yr	-0.83	yes	no	0.38
	3 yr	-0.57	yes	yes	0.23
	1 yr	-0.73	yes	no	0.12
High-Intermediate	5 yr	-1.11	yes	no	0.35
	3 yr	-1.18	yes	no	0.42
	1 yr	-1.26	yes	no	0.20
High-Long	5 yr	-2.05	yes	yes	0.44
	3 yr	-1.91	yes	no	0.29
	1 yr	-2.46	yes	no	0.23
Medium-Short	5 yr	-0.76	yes	no	0.35
	3 yr	-1.09	yes	no	0.59
	1 yr	-1.21	yes	no	0.17
Medium-Intermediate	5 yr	-0.81	yes	no	0.39
	3 yr	-0.43	no	no	0.09
	1 yr	-1.56	yes	no	0.25
Medium-Long	5 yr	-0.93	yes	no	0.35
	3 yr	-0.72	yes	no	0.46
	1 yr	-5.58	yes	yes	0.43
Low-Intermediate	5 yr	-1.86	yes	no	0.29
	3 yr	-1.40	yes	no	0.23
	1 yr	-3.33	yes	no	0.17

argue that *stock fund* returns show little persistence. Most of this evidence relates to one-year stock-fund returns, but Malkiel presents evidence on a lack of persistence over ten-year horizons. There has been little work on the persistence of fixed-income funds, but Domian and Reichenstein [1997] demonstrate that money market funds show extreme persistence, even for one-year horizons. Expense ratios virtually dictate one-year relative returns among money market funds.

This study and the others mentioned here suggest that low-cost bond funds will produce substantial persistence, especially over longer investment horizons.

SUMMARY

Historical bond returns strongly support the prediction of a strong, negative relation between bond expenses and net returns. Most evidence suggests that expenses, including loads, are a deadweight loss to the investor. Finally, low-cost funds consistently rank among the best over a five-year investment horizon. The evidence suggests bond fund returns will show substantial persistence over five-year and longer horizons. Among a sample of similar-style funds, consistently low-cost bond funds will persistently rank among the top-performing funds,

EXHIBIT 11

Average Net Returns on Load and No-Load Funds (%)

Sample		Investment Horizon			Expense Ratio	Combined Load
		5 Year	3 Year	1 Year		
High-Short	Load	4.92	5.34	5.67	1.27	2.99
	No-Load	5.55	5.74	6.18	0.61	0.00
High-Intermediate	Load	6.38	6.46	7.86	0.89	3.97
	No-Load	6.21	6.33	7.84	0.91	0.00
High-Long	Load	6.76	6.80	9.11	1.13	4.74
	No-Load	7.84	7.79	10.74	0.71	0.00
Medium-Short	Load	5.19	5.38	5.60	1.08	2.35
	No-Load	5.51	5.79	5.80	0.63	0.00
Medium-Intermediate	Load	6.02	6.19	5.73	1.43	4.25
	No-Load	6.56	6.47	6.83	0.76	0.00
Medium-Long	Load	6.18	6.61	3.58	1.36	3.60
	No-Load	6.90	7.14	7.65	0.69	0.00
Low-Intermediate	Load	6.96	7.71	-1.39	1.32	4.17
	No-Load	8.11	8.96	2.77	0.80	0.00

Combined load is the sum of front-end load and redemption fee.

and the longer the horizon, the stronger will be the returns' persistence.

ENDNOTES

I would like to thank Patrick Reinkemeyer of Morningstar for his assistance

¹Criteria for the high-short sample are: Fixed-Inc Style Box = High-Short; Morningstar category = Short Government, General Bond-Ultrashort, or General Bond-Short-Term Bond; % U.S. Stocks, % Non-U.S. Stocks, and % Other assets < 2% each; Fund Name ≠ Obj, Cat, Idx; and Total Annualized Return 5-Yr < 100%.

Asset composition criteria — e.g., % U.S. stocks < 2% — eliminate many low-rated “bond funds” and balanced funds. Fund name criteria eliminate objective averages, category averages, and bond indexes; bond index *funds* are not eliminated. The five-year return criterion eliminates funds with less than five years of returns. The medium-short sample is subject to the same category criteria.

High-intermediate and medium-intermediate samples are funds with categories of Intermediate Government or General Bond — Intermediate-Term Bond. The low-intermediate sample is funds with a High-Yield Bond category. High-long and medium-long samples contain funds in the Long Government or General Bond — Long-Term Bond categories. Category criteria eliminate international and municipal bond funds.

For the low-intermediate sample, we relax the asset composition criteria for “other” assets to less than 5%. This increases the sample size from fourteen to forty, but increases concerns about style differences within this sample.

²When a sample cannot be split into even groups, the extra observations go in the middle-cost group. For example, the middle-cost group in the low-intermediate sample has fourteen funds, while the low-cost and high-cost groups each have thirteen funds.

³For the statistically inclined, the test rejects at the 5% level the null hypothesis that the slope is zero in twenty-three of twenty-four regressions.

⁴Another test of the hypothesis that sales loads are a dead-weight loss is to compare the gross returns on load and no-load

EXHIBIT 12

Average Rank of Low-Cost Bond Funds

Sample	Investment Horizon	Investment Horizon			No. of Funds
		5 Year	3 Year	1 Year	
High-Short	Lowest Cost	32	60	30	137
	Lowest 3	35	57	14	137
High-Intermediate	Lowest Cost	14	14	46	172
	Lowest 3	15	17	33	172
High-Long	Lowest Cost	3	4	4	26
	Lowest 3	5	6	8	26
Medium-Short	Lowest Cost	1	2	5	19
	Lowest 3	7	2	7	19
Medium-Intermediate	Lowest Cost	2	9	2	23
	Lowest 3	6	10	3	23
Medium-Long	Lowest Cost	3	6	4	13
	Lowest 3	7	4	2	13
Low-Intermediate	Lowest Cost	6	10	3	40
	Lowest 3	11	15	12	40

funds. Gross returns do not reflect loads. So, if their gross returns are equal, the loads are a deadweight loss. The test's null hypothesis is the average gross returns on load funds equal the average gross return on no-load funds. There are separate tests for each of the seven samples and for each of three time horizons. The test supports the null hypothesis at the 5% level seventeen of twenty-one times. In three of the four rejections, the average gross return on no-load funds exceeds the average on load funds. These rejections occur for the one-year horizon for the medium-long sample and one- and three-year horizons for the low-intermediate sample. For the three-year horizon and the high-short sample, the test rejects the null, and the average gross return on load funds (6.62%) exceeds the average gross return on no-load funds (6.34%). Even in this case, no-load funds produce higher net returns than no-load funds, 5.74% to 5.34%; no-load funds' lower expense ratio more than makes up for their slightly lower gross return.

REFERENCES

- Bogle, John, C. *Bogle on Mutual Funds: New Perspectives for the Intelligent Investor*. Burr Ridge, IL: Irwin, 1994.
- Clements, Jonathan. "In Picking Bond Fund, Expense Factor Remains the Key." *Wall Street Journal*, April 4, 1991, p. C1.
- . "If You're Manager Is So Smart, Why Are His Expenses So High?" *Wall Street Journal*, July 6, 1999, p. R1.
- Domian, Dale L., and William Reichenstein. "Performance and Persistence in Money Market Fund Returns." *Financial Services Review*, 6, no. 3 (1997), pp. 169-183.
- Malkiel, Burton G. "Returns from Investing in Mutual Funds, 1971 to 1991." *Journal of Finance*, 50 (1995), pp. 549-571.