Another Puzzle: The Growth in Actively Managed Mutual Funds

Martin J. Gruber

The Journal of Finance, Vol. 51, No. 3, Papers and Proceedings of the Fifty-Sixth Annual Meeting of the American Finance Association, San Francisco, California, January 5-7, 1996. (Jul., 1996), pp. 783-810.

Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199607%2951%3A3%3C783%3AAPTGIA%3E2.0.CO%3B2-N

The Journal of Finance is currently published by American Finance Association.

THE JOURNAL OF FINANCE • VOL. LI, NO. 3 • JULY 1996

Another Puzzle: The Growth in Actively Managed Mutual Funds

MARTIN J. GRUBER*

ABSTRACT

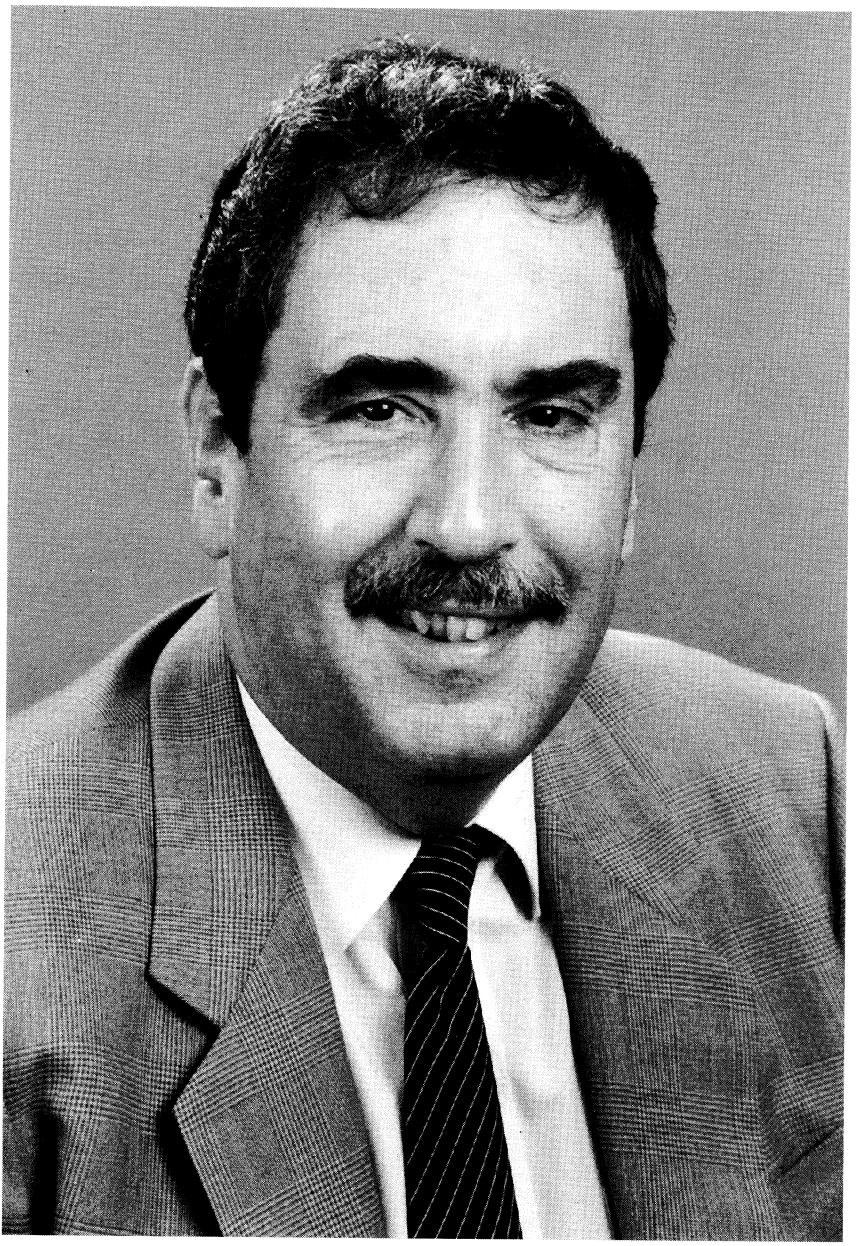
Mutual funds represent one of the fastest growing type of financial intermediary in the American economy. The question remains as to why mutual funds and in particular actively managed mutual funds have grown so fast, when their performance on average has been inferior to that of index funds. One possible explanation of why investors buy actively managed open end funds lies in the fact that they are bought and sold at net asset value, and thus management ability may not be priced. If management ability exists and it is not included in the price of open end funds, then performance should be predictable. If performance is predictable and at least some investors are aware of this, then cash flows into and out of funds should be predictable by the very same metrics that predict performance. Finally, if predictors exist and at least some investors act on these predictors in investing in mutual funds, the return on new cash flows should be better than the average return for all investors in these funds. This article presents empirical evidence on all of these issues and shows that investors in actively managed mutual funds may have been more rational than we have assumed.

I HAVE BEEN DOING research on various aspects of mutual fund performance for a number of years.[1] The more time I spent thinking about mutual funds, the more I was troubled by a question: why do investors buy actively managed mutual funds? Some parts of the answer are suggested in the literature of financial economics; some are not. In this article, I try to provide at least a partial answer by presenting empirical evidence on various pieces of this puzzle.

First, I briefly examine the importance of the open end mutual fund industry and the reasons for holding mutual funds. Next, I examine how well actively managed open end equity funds have performed relative to an appropriate set of indices. If active funds provided superior performance, there would be no puzzle. However, at the risk of giving away a punch line early, I will tell you that the average actively managed fund has negative performance compared to a set of indices. There is a possible explanation of why investors buy actively

* Stern School of Business, New York University. This article was prepared for the Presidential Address of the American Finance Association meetings in San Francisco, January 1996. I appreciate the research assistance of Jeffrey Busse and want to thank Yakov Amihud, Christopher Blake, Stephen Brown, Edwin Elton, and Gregory Udell for commenting on an earlier draft of this paper. I would also like to thank Bill Crawford and Lisa Simms of Micropal and Donald Cassidy of Lipper Analytical Services for providing me with some of the data used in this article.

[1] See for example, Elton, Gruber, Das, and Hlavka (1993), Blake, Elton, and Gruber (1993), and Elton, Gruber, and Blake (1996a and 1996b).



Martin J. Gruber
President of the American Finance Association
1996

managed funds even if they underperform indices. Holding the indices may be an unrealistic or costly alternative. I examine whether index funds exist that attempt to match the indices and how they perform relative to indices. If index funds exist that outperform actively managed funds, and if they provide most of the services offered by actively managed funds, the puzzle of why investors buy actively managed funds is intensified.

One possible explanation of why investors buy actively managed open end funds lies in their pricing. These funds sell at net asset value. This means that if management ability exists, it may not be incorporated in price. I first try to gain more insight into this by examining a type of mutual fund for which management ability should be included in price: closed end mutual funds.

I then return to the implications of the pricing of open end funds by examining a series of issues. If management ability exists and it is not included in the price of open end funds, then performance should be predictable. If performance is predictable and at least some investors are aware of this, then cash flows into and out of funds should be predictable by the very same metrics that predict performance. Finally, if predictors exist and at least some investors act on these predictors in investing in mutual funds, the return on new cash flows should be better than the average return for all investors in these funds.

I. Industry Perspective

Needless to say, this study deals with a significant phenomenon in financial markets. There has been a tremendous and persistent growth in the importance of the mutual fund industry over the past twenty years. This is true whether one measures growth by assets under management, number of mutual funds, or the number of academic articles concerned with some aspect of the mutual fund industry. Over the past twenty years, the compound annual growth rate in assets under management by mutual funds has been greater than 22 percent, while for the past 10 years it has been greater than 19 percent. At the end of 1994, there was more than $2.1 trillion invested in mutual funds, making mutual funds, measured by assets under management, the second largest type of financial intermediary in the United States, falling just short of commercial banks, but ahead of life insurance companies. Since I will be examining mutual funds investing in stock, I should point out that this is both the fastest growing and the largest category of mutual funds. Equity mutual funds accounted for 40.1 percent of the assets under management by all mutual funds in the United States. These funds held 12.2 percent of the total of all corporate equity.

A logical question to ask is what accounts for the appeal of mutual funds, and in particular actively managed mutual funds, as an investment vehicle. There have been a number of articles written about the reasons for holding mutual funds.[2] The list of reasons includes, but is not limited to:

[2] See for example, Sirri and Tuffano (1992). I have also included in the References a number of other articles on mutual fund performance that are not otherwise cited in this paper.

- Customer Services—including record keeping and the ability to move money around among funds
- Low Transaction Costs
- Diversification
- Professional Management (Security Selection)

The first three reasons, service, low transaction costs, and diversification, are provided by both active and passive (index) funds. What distinguishes the active fund is the fourth reason: professional management.

While the overall performance of professional management is frequently discussed, there is a more subtle question about performance related to the way open end mutual funds are priced. Open end mutual funds sell at net asset value. If "good management" exists, a fund that has superior management will sell at net asset value, just as a fund that has inferior management will sell at net asset value. Management is, per se, not priced. There is a counter argument that management is priced in the long run, because management raises the fees and expenses it charges customers to reflect "good management." However, the evidence I present shows that high fees are associated with inferior rather than superior management.

Clearly, one key aspect of why investors might buy mutual funds is performance. To study performance, one needs both a model to measure performance and a sample of funds to study.

II. Average Performance

In this section of the article and in many of the following sections, I will use several different measures of performance. These include:

1) A measure of return relative to the market
2) The excess return from a single index model
3) The excess return from a four index model.

More formally I will measure performance from the following equations:

$$R_{it} - R_{mt} \tag{1}$$

$$R_{it} - R_{ft} = \alpha_i^1 + \beta_{mi}^1(R_{mt} - R_{ft}) + e_i \tag{2}$$

$$R_{it} - R_{ft} = \alpha_i^4 + \beta_{mi}^4(R_{mt} - R_{ft}) + \beta_{si}(R_{st} - R_{lt}) + \beta_{gi}(R_{gt} - R_{ut}) + \beta_{di}(R_{dt} - R_{ft}) + e_i \tag{3}$$

where

R_{it} = the return on fund i in month t
R_{ft} = the return on a thirty day T-bill in month t
R_{mt} = the return on the S&P 500 index in month t

$R_{st} - R_{lt}$ = the difference in return between a small cap portfolio and a large cap portfolio based on Prudential-Bache indices in month t[3]

$R_{gt} - R_{vt}$ = the difference in return between a high growth portfolio and a value portfolio based on Prudential-Bache indices in month t

$R_{dt} - R_{ft}$ = the excess return on a bond index that represents an estimate of aggregate corporate and government bonds

β_{ki} = the sensitivity of the excess return on fund i to portfolio k where k can represent the market, a size factor, a growth factor, or a bond factor. When k represents the market, the superscript on beta indicates whether it came from the single index or the four index model.

$\alpha_i^1 \alpha_i^4$ = the risk adjusted excess return measured from the one index and the four index model respectively.

Note that all indices in equation (2) and (3) are computed as zero investment portfolios. This implies that the intercept (α) of a time series regression of a random portfolio against the indices should be zero.

In this article I will emphasize the results obtained using equation (3). I do so because as we shall see shortly, failure to include indices that span the major types of securities held by the funds under study can lead to incorrect conclusions about performance. The indices selected for the four index model span the major types of securities held by nonspecialized domestic stock funds, the type of fund examined in this study.

The sample of mutual funds used in this study consists of all common stock funds listed in Wiesenberger's *Mutual Funds Panorama* at the end of 1984, with certain exceptions. Mutual funds that were listed as foreign stock funds, specialized stock funds, or balanced funds were omitted from the sample. Including these funds would have involved adding additional indices to span the space covered by their investments. In addition, mutual funds with less than $15 million in assets at the end of 1984 were excluded because it was difficult to find consistent data on their performance and activities over time. The results in this article should be thought of as applying to a particular type of fund the characteristics of which are determined in advance.

The initial sample consisted of 270 mutual funds. The sample selected accounts for 77.2 percent of the assets held by all common stock funds in existence at the time the sample was constructed (end of 1984). Of these funds, 35 merged and 21 funds changed policy (no longer categorized as a common stock fund) during the period 1985–1994. To avoid survivorship bias in eval-

[3] The size index and growth index were constructed from the Prudential-Bache indices as follows: a) The small stock index is the average of the return on the small cap value index and the small cap growth index. Similarly, the large stock index is the average of the return on the large cap value and large cap growth index. The growth index is the average of the return on the large cap, mid cap, and small cap growth indices. The value index is the average of the return on the three size value indices.

uating the sample, I use a "follow the money" approach. If a policy change took place, I assume the investor placed his or her money in the average surviving fund. If a fund merged into another fund, I assume the investor placed his or her money in the fund that continued to exist after the merger.[4] The literature contains no evidence as to what investors do when a fund merges into another fund. I was able to obtain data on two cases where a fund merged into a fund managed by a separate organization. In both cases, over 90 percent of the money remained invested with the acquiring fund.

In order to understand why investors hold mutual funds, I start out by examining the average performance of mutual funds. Table I presents the monthly performance of the average mutual fund in my sample for the years 1985–1994. Whether one judges performance by returns relative to the market, risk adjusted returns from a single index model, or risk adjusted returns from a four index model, mutual funds underperformed the benchmark. Simply looking at unadjusted returns, we would conclude that mutual funds underperformed the market by 1.94 percent per year.[5] Since the average mutual fund in the sample had a beta smaller than one (0.96) and the market had a high rate of return during this period, this number understates performance. Using the single index model, the risk adjusted return is estimated to be −1.56 percent. Finally, the four index model suggests that mutual funds underperformed by 65 basis points per year. This ordering of performance results is very different from other studies that have used multi-index models adjusting for classes of securities.[6] What accounts for this difference in results? Examining the regression coefficients in Table I shows that mutual funds during this period tended to hold stocks that were smaller and more growth oriented than the combination of stocks in the S&P 500 index. In fact, this tendency is also present in earlier periods. During the 1985–1994 period, large stocks had a higher return than small stocks and value stocks had a higher return than growth stocks. In earlier periods small stocks had a higher rate of return than large stocks. This accounts for the different ordering of the models in this and earlier articles and shows that failure to include indices that span the type of securities a fund holds makes performance estimates more a matter of how the excluded categories of stocks did than how well management could select securities.

Table I shows that the four index model we employ does an excellent job of explaining mutual fund return behavior. The model explains 89 percent of the variability of return for the average fund in the sample. Also note that the

[4] The data presented shortly remain virtually unchanged under the alternative assumption that the investor reinvests in the average surviving fund in the sample. After the initial tables I employ this alternative reinvestment assumption.

[5] While all tables in the article report results on a monthly basis and all analysis was done on a monthly basis, I often report summary statistics in the text by multiplying the monthly data by 12 (an approximation to an annual basis).

[6] Most past studies, because of the performance of small stocks in the period studied, have found that the single index model overstates performance relative to a multi index model. See for example, Elton et al. (1993).

Table I

Average Monthly Performance

The table shows the average realized monthly percentage returns net of the S&P 500 and average coefficients from a time series regression of excess percentage returns against the single index and four index models. The sample consists of 270 open end funds and 9 closed end funds. The sample period is from January 1985 to December 1994.

Single Index Model:

$$R_i - R_f = \alpha_i^1 + B_{Mi}^1(R_M - R_f) + e_i$$

Four Index Model:

$$R_i - R_f = \alpha_i^4 + B_{Mi}^4(R_M - R_f) + B_{Si}(R_S - R_L) + B_{Gi}(R_G - R_V) + B_{Bi}(R_B - R_f) + e_i$$

where

R_i = the return on fund i.
R_f = the return on a thirty day T-Bill.
R_M = the return on the S&P 500 index.
$R_S - R_L$ = the difference in return between a small cap portfolio and a large cap portfolio.
$R_G - R_V$ = the difference in return between a high growth portfolio and a high value portfolio.
$R_B - R_f$ = the excess return on a bond index which represents an estimate of aggregate corporate and government bonds.

Fund Type	$R_i - R_M$	α^1	B_M^1	ρ^2	α^4	B_M^4	B_S	B_G	B_B	ρ^2
Open end funds										
All funds	−0.162	−0.130	0.963	0.810	−0.054	0.850	0.314	0.229	0.090	0.886
No load funds	−0.164	−0.127	0.964	0.800	−0.056	0.849	0.300	0.243	0.098	0.897
Load funds	−0.162	−0.134	0.966	0.830	−0.054	0.850	0.331	0.229	0.086	0.873
Maximum capital gains	−0.125	−0.182	1.073	0.780	−0.035	0.889	0.466	0.433	0.092	0.897
Growth	−0.154	−0.131	0.978	0.810	−0.045	0.860	0.306	0.264	0.077	0.886
Growth and income	−0.213	−0.073	0.829	0.840	−0.088	0.796	0.175	−0.031	0.108	0.878
Nonsurviving funds	−0.432	−0.350	0.929	0.710	−0.229	0.781	0.391	0.290	0.095	0.801
Closed end funds										
All funds, NAV return	−0.135	−0.031	0.864	0.780	0.014	0.787	0.193	0.153	0.077	0.820

sensitivity of mutual funds categorized by objective to each of our indices generally fits the way we expect such funds to behave. For example, funds that are categorized as seeking maximum capital gains have higher loadings on small stocks and growth stocks than do either growth funds or growth and income funds. Notice that growth and income funds have negative loadings on growth (positive on value), while more growth oriented funds have positive loadings on growth. The ability of the four index model to correctly capture the

investment policies we know are associated with funds with differing stated objectives is additional evidence that employing the multi-index model leads to more accurate performance evaluation. For these reasons I shall emphasize the use of the four index model throughout the rest of this article. However, I will also report some results based on unadjusted returns and risk-adjusted returns from a single index model.

Throughout this article, I employ a sample that is free of survivorship bias. The importance of this can be seen by examining the data on nonsurviving funds in Table I, which shows the performance of funds that disappeared from the sample because of mergers or policy change. These funds underperformed the market by 5.18 percent per year, had a one index adjusted annual return of −4.2 percent, and a four index adjusted return of −2.75 percent per year. Failure to include funds that disappeared would result in a serious overestimate of performance.

Finally, it is interesting to note that the overall performance of no load funds is virtually indistinguishable from that of load funds. While the performance is measured after deducting expenses, it is computed before load fees (as if load fees did not exist). Under any reasonable assumption about holding periods, investors in load funds had poorer performance than did investors in no load funds.

My results indicate that mutual funds underperform an appropriately weighted average of the indices by about 65 basis points per year. Expense ratios for my sample averaged 113 basis points per year. These numbers suggest that active management adds value, but that mutual funds charge the investors more than the value added.

III. Index Funds

Since mutual funds underperformed an appropriate set of indices by 65 basis points, why not simply hold the indices or at least the closest thing we can find to the indices: index funds? The answer is that over most of the period they didn't exist, or more correctly they did not exist with the low cost and the range of characteristics in which investors seemed interested. At the start of 1985, there were only three S&P 500 index funds available, two small stock funds, and no growth, value, or bond funds. Furthermore, the average expense ratio of the S&P 500 index funds was 1.24 percent per year, while for the small stock funds it was 0.875 percent per year. These expenses were higher than the amount by which the average mutual fund underperformed the risk adjusted benchmark. By 1989, the number of index mutual funds had grown to 26, and the dollar amount under management had grown from $527 million to $4.4 billion. The number of funds attempting to replicate the S&P 500 index had grown to 7, with a range of expense ratios from 21 basis points to 147 basis points per year. The average expense ratio was still high at 75 basis points.

By 1994, there were more than 100 index funds with $36.8 billion under management. Index funds existed that covered a wide variety of domestic and international securities. Expense ratios had come down but still varied. For

example, in 1994 there were 44 S&P 500 index funds with a range of expense ratios from 19 basis points (7 basis points for institutions) to 135 basis points. Forty-five basis points was the average expense ratio for all S&P 500 index funds. Bond index funds, small stock index funds, growth index funds, and value index funds were available with expense ratios below 30 basis points, although index funds also existed in these categories with much higher expense ratios.

It seems reasonable to assume that the intelligent investor who is interested in holding an index fund will choose one with a low expense ratio. Many index funds exist that have expense ratios below 30 basis points, provide all, or at least most, of the services provided by actively managed funds, and cover the types of securities in terms of size and growth implied by my model. There is still a question of how well these funds do versus the indices they hold. Two concerns are that they may underperform the indices because of transaction costs and that they may not track the indices very closely.

I used five years of data (1990–1994) to examine S&P 500 index funds, small stock index funds, and bond index funds. A shorter two-year period was used to examine value and growth index funds because more data was not available. For funds that prior to the start of the regression had reported expense ratios of 30 basis points or less, I regressed the monthly return on each index fund on the monthly return on the index the fund selected as a target to track.[7] The average R-squared was 0.997, and the minimum was 0.986. The average beta was 0.999, with a range of 0.991 to 1.004. The average annualized alpha was minus 20.2 basis points, compared to an average expense ratio of 21.9 basis points. Some of the index funds had alphas that were lower than their expense ratios; some were higher, but all differences were quite small.

While index funds were not available for the whole period, over the last five years of the period and certainly today, index funds exist that cover the range of securities held by the type of mutual funds examined in this study. It is also clear that low cost index funds exist that provide most of the services provided by actively managed funds and match the index they purport to match at a cost of about 22 basis points per year.

This brings us back to the initial puzzle. Why do individuals continue to buy actively managed mutual funds when such funds have lower risk adjusted returns than index funds? This behavior has to be based on either differences in service that appear small or else the belief that, because management is not priced, investors can select actively managed funds that will return superior performance.

We can gain added insight into this possibility by examining a type of actively managed fund where management is priced: closed end mutual funds.

[7] The exception to this was the growth and value funds. The time series of data was so small that I looked ahead at expenses in selecting these funds so that I could use data over their existence for the regression.

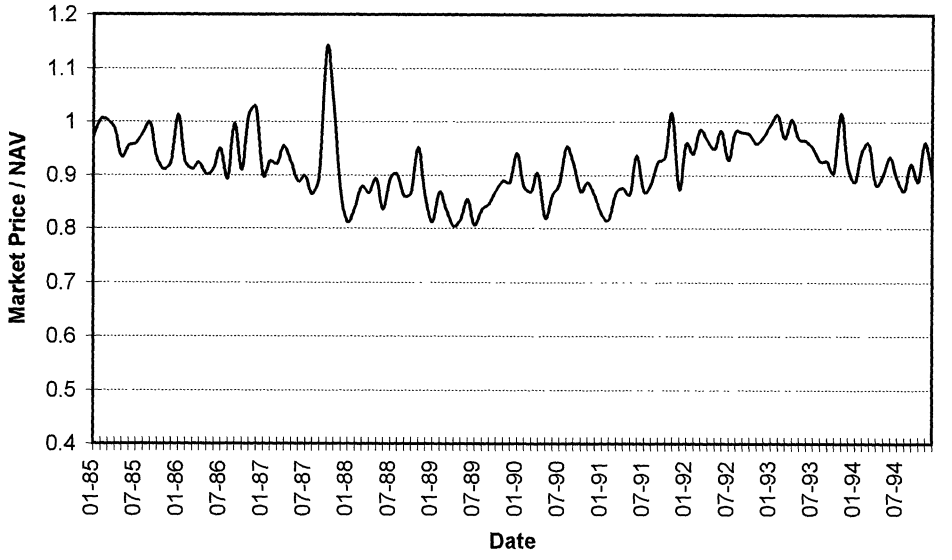


Figure 1. Average market price/net asset value. The figure shows the average market price/net asset value ratio for a sample of nine closed end funds from January 1985 to December 1994.

IV. Closed End Mutual Funds

A sample of closed end funds was constructed, using the same criteria as I used for open end funds, from among all funds listed in "Lipper Analytical and Closed End Survey" of March 1985. Recall that the funds selected (9 in total) are domestic nonspecialized common equity funds.

As shown in Figure 1, in almost each month in our sample period the average ratio of price to net asset value (Q ratio) is below one. Given that investors pay $1 for every dollar of net assets in open end funds, why do closed end funds sell at less than $1 for every dollar of net assets under management? There are several possible explanations. Three of these are that the management of closed end funds is inferior to that of open end funds, that the bundling and marketing of the underlying assets (securitization) in a closed end fund changes their characteristics, and the fact that management is priced and changes value.

The most apparent possible explanation is that management of closed end funds is not as good in security selection as that of open end funds. To examine this, I compute the net asset value return as the percentage change in net asset value, including dividends reinvested at net asset value. This is analogous to the way returns are calculated on open end funds. The results of calculating the estimates of performance for closed end funds are shown on the bottom line of Table I. Note that closed end funds have better performance than open end funds whether performance is measured by raw returns, alpha from the one index model, or alpha from the four index model. Focusing on the four index model, which I believe is the most appropriate measure, the difference in

alphas is 81.6 basis points per year. While this number may seem surprisingly high, it seems more realistic when we note that 62.5 percent of the difference in net performance is accounted for by the fact that the expenses charged by closed end funds were lower (by 51 basis points per year) than those charged by open end funds.

Table I also reveals that in terms of return on net asset value, the four index model does almost as good a job of explaining returns for closed end funds as it does for open end funds. Furthermore, closed end funds show the same tilt (although not as pronounced) in favor of growth stocks and small stocks, as was shown by open end funds.

It is possible that having a market price that differs from net asset value changes the characteristics of a share in a closed end fund. One way to see this is to examine the return on net asset value compared to the return earned by investors from holding the fund. For the sample under study, the average variability in the market return in funds was 17 percent greater than the variability of the return on the underlying assets, while the residual risk from the four index model was 140 percent greater. Market returns on shares were clearly acting differently than the return on the assets that underlie the shares. To get more insight into this, I studied the difference between the market return and the net asset return. When this difference in return is regressed against the four index model, the sensitivities (betas) on two variables are significantly different from zero at the 5 percent level: the growth-value variable, which has an average beta of -0.325, and the small versus large variable, which has a beta of $+0.142$ for the average fund. The fact that the size variable has a positive coefficient means that there is an additional systematic risk due to a size effect, and funds are priced as small stocks even after removing the size effect due to the capitalization of the companies owned by the funds. Furthermore, this additional small stock risk (beta) is cross-sectionally related to the size of the fund: the smaller the total net asset value of the fund, the higher its beta with the small minus large variable. The market prices a fund not only by the characteristics of the assets it holds, but also according to the characteristics of the fund itself.

The same type of phenomenon occurs with respect to the growth variable. In interpreting the growth index, it is useful to note that for a cross section of individual stocks, the growth index is inversely correlated with the book-to-market ratio. I find that while the growth index is positively related to the net asset value (NAV) return of the fund (see Table I), the differential return (market return-NAV return) is related to the growth index at a negative and statistically significant level. Closed end funds are priced as if they have higher book-to-market ratios than exist for the underlying assets these funds own. In addition, there is a cross-sectional relationship, this time positive, between the ratio of price to net asset value at which a fund sells and its sensitivity (beta) with the growth variable.

It is clear that closed end funds have both systematic risk and nonsystematic risk that are different from that of the assets these funds hold. This can account for part of the reason why investors will not pay 1 dollar for a dollar

of assets in closed end funds, while they will pay 1 dollar for a dollar of assets in open end funds.

I would like to take an aside for a moment and point out that the above analysis has implications well beyond this article. We saw that selling a bundle of assets as an entity led to the bundle having different risks and a lower price than the assets in the bundle. Furthermore, the attributes of the bundle, e.g., its size, affected price and risk. This has implications for the securitization of any asset and suggests that there can be negative as well as positive aspects of securitization. Moreover, it has implications for pricing the claims on business firms. The firms may be priced as more than a bundle of assets and opportunities. Value additivity might not always hold. As we see in closed end funds, the way a set of underlying assets are aggregated affects their value. This should also be true for the aggregation of assets in a corporation.

There are other reasons why there can be different pricing for open end funds than for closed end funds. There can be differences in services provided. There can be differences in the tax liability imbedded in the price of the funds. There can be differences in the percentage of the funds held in restricted stock. I believe differences in services are relatively small; I see no reason why imbedded taxes should be higher for closed end than for open end funds; and while closed end funds tend to hold more restricted stock than open end funds, given the size and investment policy of many of the funds in my sample, this should be a minor effect and should at most account for a small part of the difference in pricing.

The final reason I believe investors pay more for net assets in open end funds than they pay in closed end funds is that the value of management in open end funds is not priced, while in closed end funds it is priced. This means that if investors believe that superior management exists and they can predict it, then they can buy it for no cost in open end funds and earn a superior return. They cannot do the same for closed end funds, for expectations of management performance should be incorporated in the price of the fund.

If management ability exists and it is not included in the price of open end funds, then past performance should be predictive of future performance. This is the subject to which I now turn.

V. The Persistence of Performance

The surprising thing about persistence is not that it exists, but rather how strong it appears to be.[8] I focus on predicting both returns and risk adjusted returns from the four index model. Expenses, raw returns, αs from the one index model and αs from the four index model all have predictive ability for

[8] A number of articles have appeared recently discussing persistence. See, for example, Brown and Goetzmann (1995), Elton, Gruber and Blake (1996), Malkiel (1995), Carhart (1994), and Sharpe (1996). Sharpe (1996) is particularly interesting, for it examines evidence of persistence across a sample that includes both U.S. and international bond and stock funds.

both raw returns and risk adjusted four index αs over both one year and three year intervals.

I judge how well each measure of performance forecasts a future measure of performance by using the following methodology. At the end of each year, funds are ranked and placed into deciles on the basis of a particular selection criterion, e.g., past returns. The performance of each decile measured in several ways is computed over a one and three year holding period. I call the first period the selection period and the second period the performance period. The statistical significance of the results is estimated in two ways: first, from the rank correlation of deciles in the selection and performance period; second, by computing the mean and standard deviation of the time series of differences in the excess return on deciles and testing whether these differences are statistically different from zero.

The methodology I use in this article is free of survivorship bias. To estimate alphas in the selection period, I eliminate all funds that do not have at least 30 months of data available. Since this is known before the forecast is made, an investor could follow this rule. I use a "follow the money" approach for all funds that disappeared during the performance period. Over the performance period alphas are computed for each fund separately, and the alpha for each decile is computed by treating the decile as an equally-weighted portfolio of the alphas of each fund. This approach differs from that usually taken of estimating the betas and alphas from a time series of the returns for a decile. This methodology is superior because the composition and the risks (betas) on each decile (portfolio) changes year to year.[9]

Some typical results using this methodology are shown in Table II. In Table II, Panel A, I present the monthly risk adjusted excess return from the four index model (α^4) that would be obtained by holding deciles of funds for one year where the deciles are formed on the criteria listed at the top of each column. The table also shows the difference between deciles and whether these differences are statistically significant.[10]

Note that each forecaster examined in the table supplies information about the future performance of funds calculated on the basis of the four index model. All of the rank correlations in Panel A are statistically significant at the 1 percent level except for expenses, which is statistically significant at the 5 percent level. Examination of Panel A shows that using the four index alpha does a better job of forecasting future risk adjusted performance than past returns and that expense ratios, while still containing information, do the worst job.

Starting with expense ratios, we see that the investor is better off by buying funds with low expense ratios. While there is information in using expense ratios and in most cases it is statistically significant, the magnitude of the

[9] For example, the average regression coefficient on the growth variable for the highest decile changes from 0.27 to 0.40 to 0.62 to 0.38 from 1985 to 1989.

[10] Statistical significance of Spearman's rank correlation is based on small sample critical values as reported in Siegel and Castellan (1988). t-tests are used for the difference in means between deciles.

Table II

Performance Measured by Monthly Four Index Alpha

The table shows the average realized monthly four index alphas in percentage for one and three year holding periods where the deciles were formed using the ranking criteria shown at the top of each column. The sample consists of the 227 funds that existed in January 1982. The sample period is from January 1985 to December 1994.

Decile	Exp. Ratio[a]	1 yr return	1 yr 4 Index alpha	Exp. Ratio[a]	3 yr return	3 yr 4 Index alpha
	Panel A. 1 Year Holding Period			Panel B. 3 Year Holding Period		
Worst 1	−0.154	−0.200	−0.216	−0.143	−0.159	−0.223
2	−0.111	−0.033	−0.090	−0.079	−0.054	−0.044
3	−0.063	−0.022	−0.037	−0.069	0.023	−0.026
4	−0.018	−0.050	−0.044	−0.001	−0.036	−0.057
5	−0.017	−0.021	−0.020	−0.002	0.003	−0.013
6	0.023	−0.054	−0.039	0.010	0.013	−0.031
7	0.041	0.009	0.044	0.047	−0.013	0.011
8	−0.016	0.032	0.005	0.013	−0.049	0.005
9	−0.018	−0.016	−0.013	−0.018	−0.007	0.058
Best 10	−0.002	0.015	0.068	−0.008	0.028	0.073
Spearman Rank Coefficient	0.697**	0.782*	0.891*	0.552	0.503	0.903*
Top decile–bottom decile	0.152*	0.214**	0.284*	0.135*	0.187*	0.297*
Top decile–average of bottom 9	0.035	0.054	0.114*	0.018	0.059**	0.109*
Top 5 deciles–bottom 5 deciles	0.078*	0.062***	0.094*	0.088*	0.039**	0.096*

* Significant at 1 percent level.
** Significant at 5 percent level.
*** Significant at 10 percent level.
[a] Expense ratios are reported from high to low.

differences between deciles (portfolios) reported in the table is not as large as it is for the other forecasters, and a positive risk adjusted return cannot be earned by purchasing the top decile of funds. When we compare returns and past four index alphas as a forecaster of future risk adjusted return, the four index risk adjusted returns yield stronger results. The top decile gives larger returns, the difference between combinations of deciles is larger, and the results are statistically significant more often.[11] The results are of economic

[11] Other interesting results not reported in the tables are that using the previous three year value for return or alpha leads to results that are often significant, but not as strong as using the one year prediction discussed above. In addition, if we assume that the investor only ranks funds which have an ρ^2 above 0.80 in the selection period (only consider funds for which the four index model works well) the results are even stronger than those reported in this article. For example, the α on the best group goes up from 0.068 to 0.094.

importance. For example, an investor buying the top decile of funds ranked on α^4 would have earned a risk adjusted return of more than 3/4 of a percent per year over the years 1985–1994.

Examination of Table II, Panel B, which assumes a three year selection period and a three year holding period, shows the same type of results as found in Panel A. Past four index alphas do a much better job of forecasting future risk adjusted returns than do either past unadjusted returns or past expenses. Here the superiority of using past alphas is even more striking, as seen by the size of the rank correlation coefficients, the differences in decile performance, and the statistical significance of the difference in these deciles. It is also worth noting that the size of risk adjusted returns (per month) is about the same over each year of the three-year period as it is over the one year holding period.

I find that alphas from the four index model do the best job of forecasting risk adjusted returns. This should be the relevant performance measure. However, some investors might be interested in forecasting raw return. In Table III, Panel A and Table III, Panel B, I present evidence on the ability of raw returns and the alpha from the four index model to forecast raw returns. The results indicate clearly that the four index model does a better job than past raw returns of forecasting future raw returns.

Before leaving this section, I want to examine whether the results obtained from forecasting risk adjusted returns by past alpha is really accounted for by differences in expenses in the performance period. Are we forecasting expenses and not risk adjusted performance? To see this, I examine the expenses in the performance period along with the risk adjusted performance (these are the same numbers as in Table II) when I forecast risk adjusted performance by past alpha from the four index model. It is obvious from looking at Table IV that the difference in the performance I obtain is due to successfully forecasting risk adjusted return, not forecasting expenses. While expenses provide information about future performance, past performance provides stronger and to some extent independent information about future performance.

VI. Expenses

I have presented evidence of persistence in mutual fund performance. Persistence can exist if superior management exists and it is not included in the price (costs) to fundholders. While it cannot be directly included in price, because funds are bought and sold at net asset value, it could be indirectly included in price through fees. It has been suggested that management prices excellent performance by charging higher fees. In fact, this is *not* the case. Table V shows the fees (computed on an annual basis) charged by funds that are ranked and separated into deciles by the four index alpha. The average fee is shown for each decile both in the year of ranking and in each of the following four years. It is obvious that the top ranked funds have fees that are about the size of the average fund and lower than the bottom deciles. The expense ratio for the top performing funds goes up more slowly over time than the expense ratio for the bottom performing funds. Expense ratios over time were examined

Table III

Performance Measured by Monthly Return

The table shows the average realized monthly percentage returns for one and three year holding periods where the deciles were formed using the ranking criteria shown at the top of each column. The sample consists of the 227 funds that existed in January 1982. The sample period is from January 1985 to December 1994.

Decile	Exp Ratio[a]	1 Year Return	1 Year 4 Index Alpha	Exp Ratio[a]	3 Year Return	3 Year 4 Index Alpha
	Panel A. 1 Year Holding Period			Panel B. 3 Year Holding Period		
Worst 1	0.977	0.936	0.879	1.003	0.944	0.874
2	0.995	1.080	1.066	1.035	1.086	1.083
3	1.073	1.112	1.060	1.056	1.164	1.093
4	1.096	1.064	1.103	1.122	1.095	1.081
5	1.092	1.104	1.098	1.123	1.135	1.110
6	1.155	1.092	1.065	1.124	1.116	1.137
7	1.120	1.146	1.184	1.147	1.117	1.142
8	1.126	1.162	1.175	1.148	1.104	1.132
9	1.124	1.087	1.072	1.124	1.113	1.189
Best 10	1.174	1.137	1.217	1.173	1.175	1.208
Spearman Rank Coefficient	0.903*	0.636	0.733**	0.927*	0.527	0.927*
Top decile–bottom decile	0.197**	0.201	0.338*	0.170*	0.231*	0.334*
Top decile–average of bottom 9	0.090***	0.050	0.139***	0.075**	0.077**	0.115**
Top 5 deciles–bottom 5 deciles	0.093*	0.066	0.101**	0.076*	0.040	0.113*

* Significant at 1 percent level.
** Significant at 5 percent level.
*** Significant at 10 percent level.
[a] Expense ratios are reported from high to low.

for all of the criteria used to forecast performance, and the results are the same as those presented in Table V. Expenses are not higher for top performing funds, nor do expenses increase more rapidly in the future for top performing funds.

VII. Predicting Cash Flows

We have just seen that performance is to some extent predictable. Furthermore, this performance is not included in price because 1) funds sell at net asset value and 2) good managers actually have lower expense ratios than bad managers, and they do not raise them as performance improves. If at least some investors are aware of this, then the same metrics that predict perfor-

Table IV

Annual Alpha and Annual Expenses in Performance Period when Ranking Four Index Alphas in Selection Period

The table shows the average annual four index alpha and expense ratio in percentage for a one year holding period where the deciles were formed using the one year four index alpha and for a three year holding period where the deciles were formed using the three year four index alpha. The sample consists of the 227 funds that existed in January 1982. The sample period is from January 1985 to December 1994.

	1 Year Analysis		3 Year Analysis	
Decile	α	Expenses	α	Expenses
Worst 1	−2.587	1.405	−2.679	1.539
2	−1.074	1.061	−0.524	1.044
3	−0.449	1.007	−0.310	1.034
4	−0.531	0.968	−0.679	1.035
5	−0.237	0.994	−0.159	0.924
6	−0.474	0.962	−0.366	0.976
7	0.531	0.986	0.134	0.949
8	0.062	1.016	0.055	0.996
9	−0.156	1.043	0.690	0.980
Best 10	0.819	1.052	0.879	1.029
SRC[a]	0.891*	0.164	0.903*	0.588

[a] SRC = Spearman Rank Coefficient.
* Significant at 1 percent level.
** Significant at 5 percent level.

mance should also predict cash flows. In fact, this is exactly what happens.[12] Furthermore, it happens with any return based ranking of funds during the selection period and any reasonable definition of cash flows.[13]

I employed two different definitions of new cash flows to a fund. They each start by defining new cash flows as the change in total net asset value minus the appreciation in the funds assets. Appreciation is calculated as the total net asset value at the beginning of the period times the rate of return the fund earned during the period. This is equivalent to defining new cash flows as the cash flow that is not due to dividends and capital gains. Put another way, it assumes all dividends and capital gains are left in the fund and measures cash flow into and out of the funds above that amount.

The first measure I employ is the dollar amount of the new cash flow into and out of the fund. The problem with this approach is that large funds tend to have larger absolute cash flows regardless of performance. A second and preferred measure (called normalized cash flow) divides each new cash flow by the net asset value of the fund at the beginning of the period. This measures

[12] As discussed in the conclusion discretionary flows are limited by the restricted choice offered by some retirement plans. The extent of this influence is worthy of further study.

[13] Other articles examining return and cash flows include Hendricks, Patel, and Zeckhauser (1993), Sirri and Tufano (1993), and Carhart (1994).

Table V

Expense Ratios for Deciles Formed on the Basis of Four Index Alphas

The table shows the annual expense ratios in percentage for deciles formed according to four index alpha. The expenses are shown for the period the deciles are formed and for each of the next four years. The sample consists of the 227 funds that existed in 1982. The sample period is from January 1985 to December 1994.

		Expenses			
Decile(t)	t	$t + 1$	$t + 2$	$t + 3$	$t + 4$
Worst 1	1.357	1.405	1.413	1.415	1.400
2	1.050	1.061	1.068	1.084	1.072
3	1.018	1.007	1.031	1.064	1.115
4	0.960	0.968	0.984	0.998	1.024
5	0.971	0.994	1.029	1.049	1.064
6	0.955	0.962	0.970	0.967	0.959
7	0.973	0.986	0.998	1.007	1.050
8	1.035	1.016	1.053	1.080	1.119
9	1.027	1.043	1.037	1.055	1.028
Best 10	1.042	1.052	1.048	1.058	1.057
Mean	1.039	1.049	1.063	1.078	1.089

new cash flow as a growth rate. This tends to magnify the reported cash flow for funds with very small net asset values.

Having described these measures, let me report that the results are exceedingly powerful whichever measure I use. Table VI reports the results of using the four index alphas as a basis for forming deciles when cash flows are examined in the following year. Note how strong the ranking is and that the differences between groups of deciles are all statistically significant at the 1 percent level. When ranking into deciles is done on the basis of 1 year returns, 3 year returns, 1 year alphas, and 3 year alphas, the results are also very strong. The lowest rank correlation across all measures is 0.94, and all of the differences between deciles like those reported in Table VI are statistically significant at the 1 percent level.

There is no doubt that investors chase past performance. It is interesting to see whether we can discern which measure of past performance they chase. Recall that the α from the four index model appears to give the best predictions of future performance. Thus one could argue that investors should use this to decide on where to invest cash flows. On the other hand, I don't seriously believe that the average investor runs the four index model to measure alphas. However, the investor may come close to this by subjectively correcting for size and growth.

In Table VII, I present some results from running in each year a cross-sectional regression of normalized cash flows against the lagged value of variables that have been shown to predict returns. The table reports the mean regression coefficient and the significance of the coefficients from the regres-

Table VI

Annual Cash Flows for Deciles Formed on the Basis of Four Index Alphas in the Year Following Formation

The table shows the average realized annual cash flow in millions of dollars and cash flow normalized by dividing by total net assets from the end of the previous year for deciles formed on the basis of one year four index alpha. The cash flows are for the year following decile formation. The sample consists of the 227 funds that existed in January 1982. The sample period is January 1985 to December 1994.

Decile	Cash Flow	Normalized Cash Flow
Worst 1	−40.35	−0.154
2	−42.74	−0.112
3	−4.16	−0.051
4	14.32	−0.050
5	7.93	−0.037
6	20.79	−0.024
7	48.51	0.045
8	75.77	0.066
9	97.59	0.128
Best 10	152.44	0.290
Spearman Rank Coefficient	0.98*	1.00*
Top decile–bottom decile	192.79*	0.444*
Top decile–average of bottom nine	132.70*	0.311*
Top five deciles–bottom five deciles	92.02*	0.183*

* Significant at 1 percent level.
** Significant at 5 percent level.
*** Significant at 10 percent level.

sions. The first point to note is that the association between those variables that predict performance and subsequent cash flows is very high and statistically significant and of the sign we would expect if investors were acting on past performance. The past four index α as well as the four index α lagged one additional year are both statistically significant and have a ρ^2 with normalized cash flows of 0.13. There are a number of other variables besides past performance that might account for future cash flows. These range from marketing effort, general reputation, and any bias in our measure that may be due to a size effect.[14] I try to capture these by including a lagged cash flow variable in the regression. This variable does have large explanatory power, but even when past cash flows are included, the last two periods four index alphas are statistically significant predictors of cash flows.

When this set of regressions are rerun using a single index alpha or past returns in excess of the S&P 500 index as independent variables, very similar results are found and the explanatory power is almost identical, although

[14] This also should proxy in part for cash flows that arise because investors are locked into certain funds because of constraints on investment choices imposed by retirement accounts.

Table VII

Prediction of Cash Flows

The table shows the coefficients from a regression of the realized cash flows against the four index alpha from the prior two years, the cash flow from the prior year, the one index alpha from the prior year, and the return from the prior year. The sample consists of the 227 funds that existed in January 1982. The sample period is from January 1985 to December 1994.

$$CF_t = \gamma_0 + \gamma_1 \alpha^4_{t-1} + \gamma_2 \alpha^4_{t-2} + \gamma_3 CF_{t-1} + \gamma_4 \alpha^1_{t-1} + \gamma_5 R_{t-1} + e_t$$

γ_0	γ_1	γ_2	γ_3	γ_4	γ_5	ρ^2
0.018*	0.235*					0.106
0.014*	0.222*	0.112*				0.135
0.004	0.184*	0.054*	0.364*			0.244
0.015*	0.111**	0.059*	0.345*	0.081*		0.258
0.014*	0.127**	0.062*	0.355*		0.005**	0.258

* Significant at 1 percent level.
** Significant at 5 percent level.

slightly lower. To try to get a better idea of the significance of the return predictors, the relationships were rerun with past one index alphas and past returns in excess of the S&P 500 index entered along with past four index alphas. Note that adding either past return or past single index α to the estimated equation does lead to some slight improvement in explanatory power and that both the four index forecasters and the added forecaster remain statistically significant at the 5 percent level.[15] In an attempt to see whether size or expenses had any additional impact on cash flows, both variables were employed along with four index alphas as independent variables. Neither variable entered at a statistically significant level. These results indicate that investors do act on past performance in allocating money to mutual funds. Given the results of Section IV, investors act rationally. Furthermore, the best predictor of performance (the α from the four index model) is also marginally the best predictor of cash flows. Even though it is a complex measure, investors act as if they are paying attention to it. On the other hand, either the single index α or the excess return above the market adds marginal improvement in forecasting cash flows. At least some investors seem to be paying attention to these simpler measures of performance in allocating money among mutual funds.

VIII. Return on New Investment

I have established that there is persistence in mutual fund performance and that investors invest their money as if they are aware of this persistence. They

[15] Lagged values of the one index α and returns in excess of the market were added to this last set of equations. Neither entered at a statistically significant level. Past one index alphas and excess returns were not entered simultaneously in the regression because in several years the correlation between these two measures was close to one.

follow performance. The next step is to see if by following performance they enhance their performance. To do so, I examine the rewards that investors get on the new cash they move into and out of mutual funds each year. For this purpose, I define risk adjusted return as the alpha from my four index model which, as argued above, is the most relevant measure of the rewards an investor gets for the risk he or she takes.

The next issue to deal with is the issue of timing. I have shown that new cash flows follow good performance. If performance is good in year t, we expect high positive cash flows to follow. The question is: do these cash flows *subsequently* earn a positive risk adjusted return?

To answer this, we must make assumptions about both when the cash flows are invested and the holding period for investors. I obtained net cash flows (positive and negative) for each mutual fund in my sample on a quarterly basis.[16] I assume that new cash flows are invested at the end of the quarter in which they occurred. In fact, they are invested during the quarter. By assuming that cash flows are invested at the end of the quarter in which they arise rather than at the moment they occur, I am biasing the results against finding superior performance because, as we have already established, high returns occur during the period of time when cash flows occur. While I can't directly measure the impact of this, I try to gain an idea of its importance by also estimating excess risk adjusted return assuming investment at the beginning of the quarter rather than the end. I also examine the return to new cash flows assuming three alternative holding periods of one quarter, one year, and three years.

I compute the risk adjusted four index αs to investors in several ways. I first estimate the return on all positive new cash flows. The dollars into any fund in quarter t are multiplied by the monthly risk adjusted return *on that fund* in a subsequent period (e.g., 1 year). This is summed across all funds for all periods and divided by the total positive inflow to all funds in all periods. This is the average monthly α earned on positive new investments.

The same procedure is followed for cash outflows except that the sign is reversed. This is the risk adjusted return an investor got by disinvesting (taking money out of a fund).[17]

I measure the overall risk adjusted return on new cash flows (positive and negative) in two ways. The first is a simple weighted average of the risk adjusted return on positive cash flows and negative cash flows. The weights are respectively the value of new cash inflow and the absolute value of new cash outflow divided by the sum of the two.

[16] New cash flows are defined earlier in this article as cash additions (or deletions) from a fund after allowing for the reinvestment of dividends and capital gains.

[17] This is the same as the risk adjusted return on short sales, except that it represents an actual sale rather than a short sale. The implicit assumption here is that dollars removed from funds earn a zero alpha over the holding period being examined. If the investor had bought the fund which he actually sold, the risk adjusted returns would be negative. The reader should be careful in comparing this table to other tables (e.g., Table II), where returns are shown for buying rather than short selling poor performing portfolios.

Table VIII

Monthly Realized Cash Flow-Weighted Alpha

The table shows the monthly realized cash flow-weighted alpha in percentage computed by summing the product of the realized four index alpha and realized cash flows for holding periods of one quarter, one year, and three years. Results are shown for positive cash flows, negative cash flows, weighted average cash flows, and net cash flows as well as for realizations assuming the cash is invested at the end of the quarter (Panel A) and at the beginning of the quarter (Panel B) during which the cash flow is realized. Panel C shows realizations for load and no load funds assuming investment at the end of quarter. The sample consists of the 227 funds that existed in January 1982. The sample period is from January 1985 to December 1994.

	Holding Period		
Return on:	1 Quarter	1 Year	3 Year
	Panel A: Buy at End of Quarter		
Positive cash flows	0.0293	0.0241	0.0033
Negative cash flows	0.0712	0.0187	0.0082
Weighted average	0.0440	0.0221	0.0055
Net cash flow	0.1470	0.0824	0.0492
	Panel B: Buy at Beginning of Quarter		
Positive cash flows	0.1032	0.0466	0.0128
Negative cash flows	0.0522	0.0214	0.0139
Weighted average	0.0852	0.0376	0.0133
Net cash flow	0.2885	0.1318	0.0813

	Panel C: Load vs. No Load			
	1 Quarter		1 Year	
	No Load	Load	No Load	Load
Positive cash flows	0.0433	0.0199	0.0183	0.0281
Negative cash flows	0.0520	0.0834	0.0139	0.0217
Weighted average	0.0463	0.0425	0.0168	0.0257
Net cash flow	0.1456	0.1481	0.0584	0.1003

The second method looks at new cash flows as an overall portfolio. The dollars invested is equal to the new cash inflows minus the absolute value of the new cash outflows. The procedure is identical to that presented above, except that the denominator of the weights on cash inflows and outflows is set equal to the net amount of dollars invested (i.e., inflows minus outflows). This method produces a measure of the return on the new portfolio held by the aggregate of all investors. It assumes that dollars which are disinvested in funds are used to finance part of the new investment in funds.

All of the results are shown in Table VIII. Let us examine one case in detail. Look at the cash involving purchases at the end of the quarter assuming a one year holding period. The risk-adjusted excess returns earned on new cash inflows to funds was 2.4 basis points per month or 28.9 basis points per year. The investors who removed money from funds save 1.87 basis points per month or 22.4 basis points per year. The average dollar moved saved 26.5 basis points

per year, the weighted average of the previous two numbers. Finally, if we assume that the cash flow out of funds was reinvested in new funds, the investor saved a negative risk adjusted return on the cash disinvested and earned a positive return on the reinvestment of those funds to give a yearly risk adjusted return of 99 basis points per year.

It is clear from examining Table VIII that assuming a one quarter holding period leads to higher risk adjusted returns for investors. However, note that the improvement in results comes almost exclusively from disinvestment in funds rather than new investment in funds. Finally, if a three year holding period is assumed, the excess risk adjusted returns become much smaller, although they are still positive.[18]

Panel B of Table VIII shows that the risk adjusted returns we discussed above probably understate the return actually earned on new cash flows. If the money were invested at the start of the quarter (an impossibility since the cash flow had not yet occurred), returns would be very much larger. Assuming none of the cash flow that occurred during the quarter was reinvested until the quarter was over (Panel A) understates the returns earned on new cash flows.

Finally, Panel C presents the results (under the timing assumption of Panel A) reported separately for no load and load funds. Note that the return on marginal money in load funds is higher than that on no load funds. However, the return on no load funds is greater than zero. The excess return earned by the new cash flow into load funds during the period is not large enough to compensate investors for the load.

Throughout this article I have analyzed returns before taxes. This seems appropriate for many of the investors in mutual funds. Estimates of the percentage of new cash flow to stock funds that came from tax deferred sources (e.g., IRA, Keogh, and 401K plans) exceeded 50 percent in 1993 and by the end of 1994, over half of the assets held in these funds were from tax deferred sources. To the extent new cash flow was supplied by investors with tax-deferred accounts, the risk adjusted return on investing in actively managed funds was clearly advantageous. But is this the case for an investor subject to taxes?

By buying an actively managed portfolio rather than an index fund, the investor gains extra risk adjusted return as measured by α^4, but pays taxes on capital gains earlier. Whether buying active funds is the preferred strategy depends on the return expected on index funds, the dividend yield on index funds, the expense ratio of index funds, the α^4 on actively managed funds, the investor's tax rates (on dividends and capital gains), and the investor's time horizon (point in time at which the investor wants to maximize wealth).[19]

[18] It is worthwhile mentioning that if annual cash flow data are used rather than quarterly data and cash flows that occur during the year are assumed to be invested at the end of the year, the risk adjusted return on new cash inflows are slightly negative, while the risk adjusted return on outflows and overall cash flows are very slightly positive.

[19] The reader should note that to obtain the extra return from active management I assume the investor turns over his or her portfolio once a year. I describe as time horizon the period over which the investor expects to pursue an active or passive strategy.

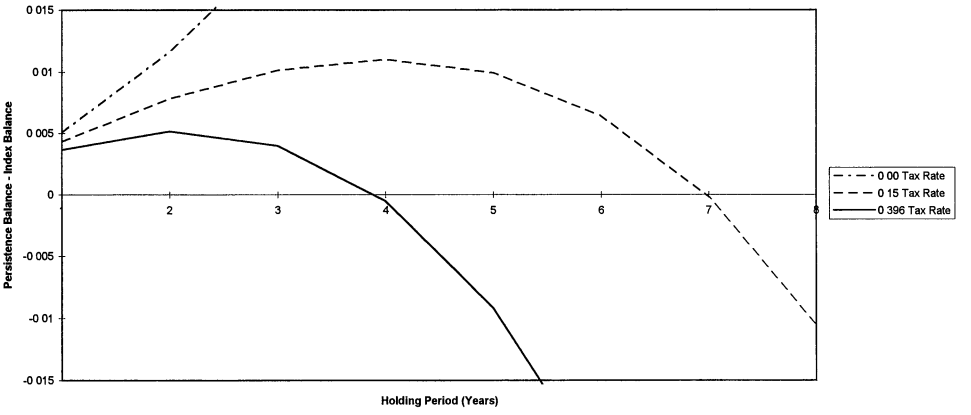


Figure 2. Net advantage of persistence strategy over index strategy 1985–1994 market environment. The figure shows the net advantage of investing according to the persistence strategy over investing in index funds for various federal tax rates. Historical averages from January 1985 to December 1994 are used for the market return and dividend yield. Expenses reflect present day averages.

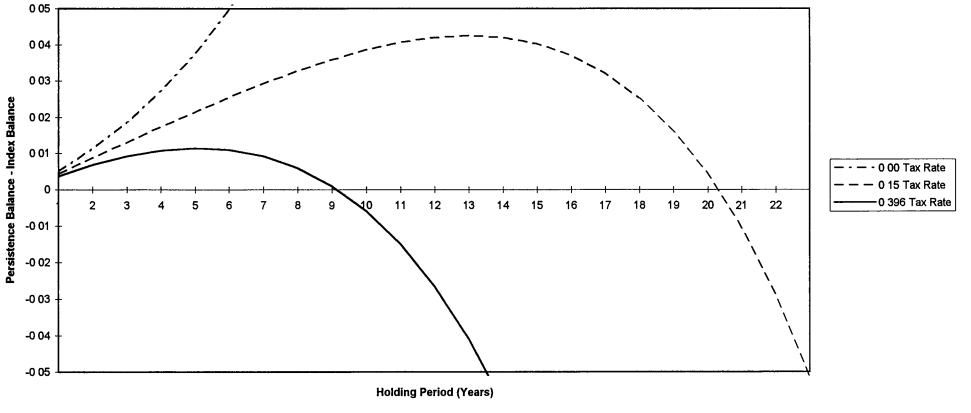


Figure 3. Net advantage of persistence strategy over index strategy 1926–1994 market environment. The figure shows the net advantage of investing according to the persistence strategy over investing in index funds for various federal tax rates. Historical averages from January 1926 to December 1994 are used for the market return and dividend yield. Expenses reflect present day averages.

In the appendix, I present the formulas for determining the after tax wealth N years in the future that will arise from investing in index funds or actively managed portfolios assuming the investor sells his or her portfolio and thus pays taxes on the portfolio's realized and unrealized capital gains once a year. In Figures 2 and 3, I present the added advantage of investing in active rather than passive funds for different holding periods and different tax rates under particular assumptions about the investment environment. The reader can use

the formulas in the appendix to examine the after tax advantage or disadvantage of active management under different assumptions.

Figure 2 uses data representative of market and fund behavior over the period 1985–1994. The return on the market was assumed to be 14.4 percent, the dividend yield on the market was assumed to be 3.6 percent, the expense ratio for index funds was assumed to be 22 basis points, and the added excess risk adjusted return from active management was assumed to be 28.9 basis points per year. Clearly, under this assumption an investor who can defer taxes (e.g., a 401K plan investor) is better off holding active portfolios. An investor who is subject to a 15 percent marginal tax rate is better off holding an active portfolio as long as he or she plans to invest for less than seven years. Finally, an investor subject to the highest marginal tax rate 39.6 percent is better off in the active portfolio as long as the relevant time horizon is smaller than 3.8 years.[20]

Figure 3 is constructed under the same assumptions as Figure 2, except that the market return is set to 10.2 percent and the dividend yield to 4.2 percent. These numbers better reflect the long run market environment. In this case, the investor with a 15 percent marginal tax rate should prefer active management as long as the time horizon was less than 20 years, while for the investor in the marginal 39.6 percent bracket, the breakeven horizon is 9.1 years.[21]

Why do these graphs appear as they do? Let us look at Figure 2. For any tax rate there is an initial advantage for the investor in active management, for it is worth postponing the tax on a 10.8 percent capital gain by one year to pick up an additional return of 50.9 basis points. The advantage disappears for longer holding periods because of the added advantage (of passive management) of compounding return on unrealized capital gains.

IX. Summary and Conclusion

In conclusion, let me return to my initial mutual fund puzzle: why do investors buy actively managed open end mutual funds and pay one dollar for each dollar under management when

[20] In constructing this figure current tax rates were used. The fact that the largest federal tax rate on capital gains was 28 percent while on dividends it was 39.6 percent was recognized. State taxes vary widely and were ignored in this analysis. The breakeven time horizon for other marginal tax rates, 28 percent, 31 percent, and 36 percent are very close (within 0.15 years) to that for the 39.6 percent tax rate. The breakeven time horizon should be larger than the numbers shown in these figures because I assumed zero turnover on the part of index funds in constructing the figures.

[21] All of this analysis was repeated assuming a more conservative alpha of 22 basis points rather than the 28.9 basis points assumed above. The 22 basis points represent the annual return on positive cash flows for investors in no load funds who have a one year holding period, while the 28.9 basis points represent this return for holders of both load and no load funds. Assuming an alpha of 22 basis points and the 1985–1994 assumptions described in the text, the breakeven time horizon for the 15 percent tax rate and the 39.6 tax rate become 6 years and 3.4 years respectively. For the long term assumptions, these breakeven times become 16.3 years and 7.9 years respectively.

a) Mutual funds, on average, offer a negative risk adjusted return
b) The investor can get a better deal by buying index funds
c) In the closest investment to open end funds, closed end funds, the investor will not pay as much as one dollar for every dollar under management.

I believe I have solved the puzzle. Future performance is in part predictable from past performance. This can occur because the price at which funds are bought and sold is equal to net asset value and does not change to reflect superior management. A group of sophisticated investors seems to recognize this, as evidenced by the fact that the flow of new money into and out of mutual funds follows the predictors of future performance. Furthermore, investors who supplied new cash flow benefit from this, for the risk adjusted returns earned on new cash flows (both into and out of funds) over the ten years of this study are positive and above the return earned by both the average active and the average passive fund.

This raises another question. Why do we see any money remain in funds that are predicted to do poorly and in fact do perform poorly? I propose a possible explanation for this phenomenon: The existence of two clienteles, a sophisticated clientele and a second clientele that I will refer to as a disadvantaged clientele.

The sophisticated clientele directs its money to funds based on performance. The disadvantaged clientele consists of three groups:

1) Unsophisticated investors—a group that directs its money to funds based at least in part on other influences such as advertising and advice from brokers.
2) Institutionally disadvantaged investors—a group primarily represented by pension accounts that are restricted by the plan they are part of to a set of funds that underperforms the best active funds.
3) Tax disadvantaged investors—a group that has held one or more funds for enough time so that capital gains taxes make it inefficient to remove money from these funds. This group can still act as sophisticated investors in placing new money.

All of the evidence in this article is consistent with this hypothesis.

a) The stock of money underperforms appropriate benchmarks. The stock of money is likely to contain a large percentage of the funds invested by the disadvantaged clientele.
b) The flow of money performs better than appropriate benchmarks. Sophisticated investors are likely to constitute a larger percentage of new cash flows into and out of mutual funds. The investor who moves cash into and out of funds earns a positive risk adjusted return and gets the services provided by mutual funds at no net cost.[22]

[22] I should point out that even some sophisticated investors will buy index funds. They may give up the extra return from selecting a superior manager to avoid the nonsystematic risk of selecting an inferior one. This is not a problem for the sophisticated investor who diversifies across funds and across time.

c) The flow of new money into the best performing funds is much larger than the flow of money out of the poorer performing funds. Tax disadvantaged investors and to some extent institutionally disadvantaged investors will not or cannot move money out of bad funds, but tax disadvantaged investors and to some extent institutionally disadvantaged investors can place new money in good performing funds.

d) The flow of new money into and out of mutual funds underperforms naive rules for selecting mutual funds as examined in Table II. Again I explain this as due to the presence of both a disadvantaged clientele and a sophisticated clientele. The sophisticated investor earns more than the average positive return on marginal cash flows; the disadvantaged investor earns less.

e) Because sophisticated investors can't short sell funds, they cannot eliminate inefficient funds.[23] However, by disinvesting (or not investing) in these funds they eliminate the worst performing funds in the sample over time. This is the reason that 28 percent of the poorer performing funds in my sample merged or changed to noncommon stock funds over the ten year sample period.

I believe that I have shown that the continued investment in actively managed funds as well as index funds is much more rational than we have assumed in the past.

Appendix A

Formulas Used to Compute the Horizon Value of Money Invested in Index Funds or Active Management

1. Index Funds

Assumptions:

Taxes are paid on dividends each year.

Taxes are paid on capital gains at the end of the horizon.

$$IV_N = V_0(1 + R_M - R_E - R_D T_I)^N(1 - T_{CG}) + \left[\sum_{t=1}^{N} R_D V_{t-1}(1 - T_I) + V_0\right] T_{CG} \quad (1)$$

where

$$V_t = V_0(1 + R_M - R_E - R_D T_I)^t.$$

[23] A small number of mutual funds have recently begun to allow short sales.

2. Active Management

Assumptions:

Portfolio is turned over once a year at which time taxes on dividends and capital gains are paid.

$$AV_N = V_0[(R_M + R_A - R_D)(1 - T_{CG}) + R_D(1 - T_I) + 1]^N \qquad (2)$$

where

V_0 = starting investment at time 0,
IV_N = end of horizon after tax wealth if invested in index funds,
AV_N = end of horizon after tax wealth from active management,
R_M = return on the market including dividends,
R_E = the expense ratio for index funds,
R_D = dividend yield,
R_A = extra return from active management,
T_I = tax on dividends,
T_{CG} = tax on capital gains,
N = number of years in investment horizon.

REFERENCES

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

Brown, Stephen J., and William N. Goetzmann, 1995, Attrition and mutual fund performance, *Journal of Finance* 50, 679–698.

Brown, Stephen J., William N. Goetzmann, Roger G. Ibbotson, and Stephen A. Ross, 1992, Survivorship bias in performance studies, *Review of Financial Studies* 5, 553–580.

Carhart, Mark M., 1992, Persistence in mutual fund performance re-examined, Working paper, University of Chicago.

Carhart, Mark M., 1994, On persistence in mutual fund performance, Working paper, University of Chicago.

Connor, Gregory, and Robert A. Korajczyk, 1991, The Attributes, Behavior, and Performance of U.S. Mutual Funds, *Review of Quantitative Finance and Accounting* 1, 5–26.

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

Elton, Edwin J., and Martin J. Gruber, 1992, Portfolio analysis with a nonnormal multi-index return generating process, *Review of Quantitative Finance and Accounting* 2, 5–16.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, Survivorship bias and mutual fund performance, *Review of Financial Studies*, forthcoming.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 1995, Fundamental variables, expected returns, and bond fund performance, *Journal of Finance* 50, 1229–1256.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 1996, The persistence of risk-adjusted mutual fund performance, *Journal of Business* 69, 133–157.

Elton, Edwin J., Martin J. Gruber, and Manfred W. Padberg, 1976, Simple criteria for optimal portfolio selection, *Journal of Finance* 31, 1341–1357.

Elton, Edwin J., Martin J. Gruber, Sanjiv Das, and Matthew Hlavka, 1993, Efficiency with costly information: A reinterpretation of evidence from managed portfolios, *Review of Financial Studies* 6, 1–22.

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

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1994, Size and book-to-market factors in earnings and returns, Working paper, Graduate School of Business, University of Chicago.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Grinblatt, Mark, and Sheridan Titman, 1988, The evaluation of mutual fund performance: An analysis of monthly returns, Working paper, The John E. Anderson Graduate School of Management, UCLA.

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62, 393–416.

Grinblatt, Mark, and Sheridan Titman, 1992, The persistence of mutual fund performance, *Journal of Finance* 47, 1977–1984.

Hendricks, Darryll, Jayendu Patel, and Richard Zeckhauser, 1993, Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988, *Journal of Finance* 48, 93–130.

Henriksson, Roy D., 1984, Market timing and mutual fund performance: An empirical investigation, *Journal of Business* 57, 73–96.

Henriksson, Roy D., and Robert Merton, 1981, On market timing and investment performance II: Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–534.

Ippolito, Richard A., 1989, Efficiency with costly information: A study of mutual fund performance, 1965–1984, *Quarterly Journal of Economics* 104, 1–24.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 389–416.

Lehmann, Bruce N., and David M. Modest, 1987, Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons, *Journal of Finance* 42, 233–265.

Lipper Analytical Securities Corporation, *Lipper Analytical & Closed End Survey*, 1985–1990 (Lipper Analytical Services Inc., New York).

Lipper Analytical Securities Corporation, *Lipper Closed-End Fund Performance Analysis Service*, 1991–1994 (Lipper Analytical Services Inc., New York).

Lipper Analytical Securities Corporation, *Lipper Mutual Fund Performance Analysis*, 1985–1994 (Lipper Analytical Services Inc., New York).

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds 1971–1991, *Journal of Finance* 50, 549–572.

Sharpe, William F., 1966, Mutual fund performance, *Journal of Business* 39, 119–138.

Sharpe, William F., 1988, Determining a fund's effective asset mix, *Investment Management Review* 2, 59–69.

Sharpe, William F., 1996, The styles and performance of large seasoned U.S. mutual funds, *World Wide Web* http://gsb-www.stanford.edu/~wfsharpe/ls100.htm.

Siegel, Sidney, and N. John Castellan, Jr., 1988, *Nonparametric Statistics for the Behavioral Sciences*, 2nd ed. (McGraw-Hill, New York).

Sirri, Erik R., and Peter Tufano, 1992, The demand for mutual fund services by individual investors, Working paper, Harvard Business School.

Sirri, Erik R., and Peter Tufano, 1993, Buying and selling mutual funds: Flows, performance, fees, and services, Working paper, Harvard Business School.

Treynor, Jack L., 1965, How to rate management of investment funds, *Harvard Business Review* 43, 63–75.

Treynor, Jack L., and Fisher Black, 1973, How to use security analysis to improve portfolio selection, *Journal of Business* 46, 66–86.

Wiesenberger Inc., *Mutual Funds Panorama*, 1985–1994, Wiesenberger Investment Companies, New York.