

Mutual Funds
by
Edwin J. Elton*
Martin J. Gruber**

April 14, 2011

* Nomura Professor of Finance, New York University

** Professor Emeritus and Scholar in Residence, New York University

1. Introduction

Mutual funds have existed for over 200 years. The first mutual fund was started in Holland in 1774, but the first mutual fund didn't appear in the U.S. for 50 years, until 1824. Since then the industry has grown in size to 23 trillion dollars worldwide and over 11.8 trillion dollars in the U.S. The importance of mutual funds to the U.S. economy can be seen by several simple metrics:¹

1. Mutual funds in terms of assets under management are one of the two largest financial intermediaries in the U.S.
2. Approximately 50% of American families own mutual funds.
3. Over 50% of the assets of defined contribution pension plans are invested in mutual funds.

In the U.S., mutual funds are governed by the Investment Company Act of 1940. Under law, mutual funds are legal entities which have no employees and are governed by a board of directors (or trustees) who are elected by the fund investors. Directors outsource all activities of the fund and are charged with acting in the best interests of the fund investors.

Mutual funds tend to exist as members of fund complexes or fund families. There are 16,120 funds in the U.S. Of these, 7,593 are open-end funds which are distributed by 685 fund families.² Funds differ from each other by the type of securities they hold, the services they provide, and the fees they charge. The sheer number of funds makes evaluation of performance important. Data, transparency, and analysis become important in selecting funds.

Usually when people talk about mutual funds they are referring to open-end mutual

¹ All descriptive statistics in this section as of the start of 2011 (or the last available data on that date) unless otherwise noted.

² The assets in fund families are highly concentrated, with the 10 largest families managing 53% of the assets in the industry and the top 25 families managing 74%. The number of mutual funds reported above excludes 6,099 Unit Investment Trusts.

funds, but there are three other types of mutual funds: closed-end funds, exchange-traded funds, and unit investment trusts. Examining each type as a percentage of the assets in the industry we find open-end mutual funds are 90.5%, closed-end funds 1.9%, exchange-traded funds 7.6%, and unit investment trusts less than .25%.

In this chapter we will discuss the three largest types of funds, with emphasis on the unique aspects of each. We will start with a brief discussion of each type of fund.

1.1 Open-End Mutual Funds

In terms of number of funds and assets under management, open-end mutual funds are by far the most important form of mutual funds. What distinguishes them from other forms is that the funds can be bought and sold anytime during the day, but the price of the transaction is set at the net asset value of a share at the end of the trading day, usually 4 PM. It is both the ability to buy and sell at a price (net asset value) which will be determined after the buy or sell decision, and the fact that the other side of a buy or sell is the fund itself, that differentiates this type of fund from other types.

Mutual funds are subject to a single set of tax rules. To avoid taxes, mutual funds must distribute by December 31st 98% of all ordinary income earned during the calendar year and 98% of all realized net capital gains earned during the previous 12 months ending October 31st. They rarely choose not to do so. They can lower their capital gains distributions by offsetting gains with losses and by occasionally paying large investors with a distribution of securities rather than cash.

Open-end mutual funds are categorized as follows: stock funds (48%), bond funds (22%), money market funds (24%), and hybrid funds, holding both bonds and stock, (7%). We will

concentrate our analysis on bond funds, stock funds, and hybrid funds, funds which hold long-term securities. These funds hold 76% of the assets of open-end funds.

Open-end mutual funds can be passive funds attempting to duplicate an index, or active funds which attempt to use analysis to outperform an index. Index funds represent 13% of the assets of open-end funds, with 40% of the index funds tracking the S&P 500 Index. These passive funds can offer low-cost diversification. In 2009 the median annual expense ratio for active funds was 144 basis points for stock funds and 96 basis points for bond funds. In general, index funds have a much lower expense ratio with expense ratios for individuals as low as 7 basis points.

1.2 Closed-End Mutual Funds

Closed-end mutual funds, like open-end mutual funds, hold securities as their assets and allow investors to buy and sell shares in the fund. The difference is that shares in a closed-end fund are traded on an exchange and have a price determined by supply and demand which (unlike open-end funds) can, and usually does, differ from the net asset value of the assets of the fund. Furthermore, shares can be bought or sold at any time the market is open at the prevailing market price, while open-end funds are priced only once a day. Perhaps the easiest way to think of closed-end funds is a company that owns securities rather than machines. The difference between the price at which a closed-end fund sells and its net asset value has been subject of a large amount of analysis, and will be reviewed in great detail later in this chapter. We will simply note here that closed-end stock funds tend to sell at prices often well below the net asset value of their holdings.

The composition of the 241 billion dollars in closed-end funds is different from the

composition of open-end funds. Bond funds constitute 58% of the assets in closed-end funds, and stock funds 42% of the assets. If we restrict the analysis to funds holding domestic assets, the percentages are 68% to bonds and 32% to equity.

1.3 Exchange-Traded Funds

Exchange-traded funds are a recent phenomenon, with the first fund (designed to duplicate the S&P 500 Index) starting in 1993. They are very much like closed-end funds with one exception. Like closed-end funds, they trade at a price determined by supply and demand and can be bought and sold at that price during the day. They differ in that at the close of the trading day investors can create more shares of ETFs by turning in a basket of securities which replicate the holdings of the ETF, or can turn in ETF shares for a basket of the underlying securities. This eliminates one of the major disadvantages of closed-end funds, the potential for large discounts. If the price of an ETF strays very far from its net asset value, arbitrageurs will create or destroy shares, driving the price very close to the net asset value. The liquidity which this provides to the market, together with the elimination of the risk of large deviations of price from net asset value, has helped account for the popularity of ETFs.

2. Issues with Open-End Funds

In this section we will discuss performance measurement, how well active funds have done, how well investors have done in selecting funds, other characteristics of good-performing funds, and influences affecting inflows.

2.1 Performance Measurement Techniques

No area has received greater attention in mutual fund research than how to measure performance. This section starts with a discussion of problems that a researcher must be aware of when using the standard data sources to measure performance. It is followed by a subsection that

discusses the principal techniques used in performance measurement of stock funds. The third subsection discusses performance measurement for bond funds. The fourth subsection discusses the measurement of timing.

2.1.1 Data Sources, Data Problems, and Biases

While many of the standard sources of financial data are used in mutual fund research, we will concentrate on discussing issues with the two types of data that have been primarily developed for mutual fund research. We will focus on the characteristics of and problems with data sets which contain data on mutual fund returns, and mutual fund holdings. Mutual fund return data is principally available from CRSP, Morningstar and LIPPER. Mutual fund holdings data is available on several Thompson and Morningstar databases.

There are problems with the returns data that a researcher must be aware of. First is the problem of backfill bias most often associated with incubator funds.³ Incubation is a process where a fund family starts a number of funds with limited capital, usually using fund family money. At the end of the incubator period the best-performing funds are open to the public and poor-performing funds are closed or merged. When the successful incubator fund is open to the public, it is included in standard databases with a history, while the unsuccessful incubator fund never appears in databases. This causes an upward bias in mutual fund return data. Evans (2010) estimated the risk-adjusted excess return on incubator funds that are reported in data sets as 3.5%. This bias can be controlled for in two ways. First, when the fund goes public it gets a ticker. Eliminating all data before the ticker creation date eliminates the bias. Second, eliminating the first three years of history for all funds also eliminates the bias at the cost of eliminating useful data for non-incubator funds.

³ This is developed and analyzed in Evans (2010). He employed a four-factor model (Fama-French and momentum) to estimate alpha or risk-adjusted excess return.

The second problem concerns the incompleteness of data for small funds. Funds under \$15 million in assets and 1,000 customers don't need to report net asset value daily. Funds under \$15 million either don't report data or report data at less frequent intervals than other funds in most databases. If they are successful they often enter standard databases with their history, another case of backfill bias. If they fail, they may never appear (see Elton, Gruber & Blake (2001)). This, again, causes an upward bias in return data. It can be eliminated by removing data on all funds with less than \$15 million in assets.

The third problem, which has never been studied, arises from the difference in the fund coverage across databases. When CRSP replaced Morningstar data with LIPPER data, over 1,000 funds disappeared from the database. What are the characteristics of these funds? Do the differences bias results in any way?

The fourth problem is that many databases have survivorship bias. In some databases, such as Morningstar, data on funds that don't exist at the time of a report are not included (dropped) from the database. Thus, using the January 2009 disk to obtain ten years of fund returns excludes funds that existed in 1999 but did not survive until 2009. Elton, Gruber & Blake (1996a) show that funds that don't survive have alphas below ones that survive, and excluding the failed funds, depending on the length of the return history examined, increases alpha by from 35 basis points to over 1%. The CRSP database includes all funds that both survive and fail, and thus is free of this bias. To use Morningstar data, one needs to start at some date in order to obtain funds that existed at that starting date and to follow the funds to the end of the time period studied or to when they disappear.

Holdings data can be found from Morningstar and from Thompson. The most widely used source of holdings data is the Thompson holdings database since it is easily available in

computer-readable form. The Thompson database lists only the holdings data for traded equity. It excludes non-traded equity, equity holdings that can't be identified, options, bonds, preferred, convertibles, and futures.

The Morningstar database is much more complete, including the largest 199 holdings in early years and all holdings in later years. Investigators using the Thompson database have the issue of what to do about the unrecorded assets. Usually, this problem is dealt with in one of two ways. Some investigators treat the traded equity as the full portfolio. Other authors treat the differences between the aggregate value of the traded equity and total net assets as cash. Either treatment can create mis-estimates of performance (by mis-estimating betas) that may well be correlated with other factors. Elton, Gruber and Blake (2010b) report that about 10% of funds in their sample use derivatives, usually futures. Futures can be used in several ways. Among them are to use futures with cash to manage inflows and outflows while keeping fully invested, as a timing mechanism, and as an investment in preference to holding the securities themselves. Investigators report numbers around 10% for the percentage of securities not captured by the Thompson database. However, there is wide variation across funds and types of funds. For funds that use futures sensitivities to an index will be poorly estimated. Likewise, for funds that have lower-rated bonds use options or convertibles or have non-traded equity, sensitivity to indexes can be poorly estimated. The problem is most acute when timing is studied. Elton, Gruber and Blake (2011b) analyze the problem of missing assets when alpha is being calculated, and find that the superior performing funds are very different depending on whether a complete set of assets or the Thompson database are used.

2.1.2 Performance Measurement of Index Funds

Index funds are the easiest type of fund to evaluate because generally there is a well-defined single index that the fund attempts to match. For example, when evaluating the “Wilshire 2000” index fund, the fund’s performance is judged relative to that index. We will concentrate on S&P 500 index funds in the discussion which follows, but the discussion holds for index funds following other indexes.

There are several issues of interest in studying the performance of index funds. These include:

1. Index construction
2. Tracking error
3. Performance
4. Enhanced return index funds

2.1.2.1 Index Construction

The principal issue here is how interest and dividends are treated. Some indexes are constructed assuming daily reinvestment, some monthly reinvestment, and some ignore dividends. Index funds can make reinvestment decisions that differ from the decisions assumed in the construction of the index. In addition, European index funds are subject to a withholding tax on dividends. The rules for the calculation of the withholding tax on the fund may be very different from the rules used in constructing the index. These different aspects of construction need to be taken into account in the conclusions one reaches about the performance of index funds versus the performance of an index.

2.1.2.1 Tracking Error

Tracking error is concerned with how closely the fund matches the index. This is usually

measured by the residuals from the following regression:⁴

$$R_{pt} = \alpha_p + \beta_p(I_t) + e_{pt}$$

Where

I_t is the return on the index fund at time t

α_p is the average return on the fund not related to the index

I_t is the return on the index at time t

e_{pt} is the return on portfolio p at time t unexplained by the index (mean zero)

β_p is the sensitivity of the fund to the index

R_{pt} is the return on the fund at time t

A good-performing index fund should exhibit a low variance of e_{pt} and low autocorrelation of e_{pt} over time so that the sum of the errors is small. Elton, Gruber and Busse (2004) found an average R^2 of 0.999 when analyzing the S&P 500 index funds indicating low tracking error. The β_p is a measure of how much of the portfolio is invested in index matching assets.

It is a partial indication of performance since it measures in part the efficiency with which the manager handles inflows and outflows and cash positions.

2.1.2.2 Performance of Index Funds

The α_p is a measure of performance. It depends in part on trading costs since the index fund pays trading costs where the index does not. Thus we would expect higher α_p for S&P 500 Index

⁴ Two variants of this equation have been used. One variant is to set the beta to one. This answers the question of the difference in return between the fund and the index. However, performance will then be a function of beta with low beta funds looking good when the market goes down. The other variant is to define returns as returns in excess of the risk-free rate. Failure to do this means that alpha will be partially related to one minus beta. However, beta is generally so close to one that these variants are unlikely to lead to different results.

funds where trading costs are low and index changes are small relative to small or mid-cap index funds where index changes are more frequent and trading costs are higher. Second, α_p depends on management fees. Elton, Gruber and Busse (2004) find that the correlation between future performance and fees is over -0.75 for S&P 500 index funds. Third, the value of α_p depends on management skill in portfolio construction. For index funds that are constructed using exact replication, management skill principally involves handling index changes and mergers although security lending, trading efficiency, and the use of futures are also important. For indexes that are matched with sampling techniques, portfolio construction also can have a major impact on performance. Problems with matching the index are especially severe if some securities in the index are almost completely illiquid, holding all securities in the index in market weights would involve fractional purchases, or because some securities constitute such a large percentage of the index that holding them in market weights is precluded by American law. Finally, European mutual funds are subject to a withholding tax on dividends which also affects performance and impacts alpha. Because of fees and the limited scope for improving performance, index funds almost always underperform the indexes they use as a target.

2.1.2.3 Enhanced Return Index Funds

A number of funds exist that attempt to outperform the indexes they declare as benchmarks. These are referred to as “enhanced return” index funds. There are several techniques used. First, if futures exist the fund can match the index using futures and short-term instruments rather than holding the securities directly. Holding futures and short-term instruments may lead to excess returns if futures generally deviate from their arbitrage value in a manner that means they offer more attractive returns. Some index funds have been organized on this premise. Second, if the fund invests in short-term assets that give a higher return than the short-term assets used in the

future spot arbitrage relationship, it can give a higher return. Finally, switching between futures and the replicating portfolio depending on the direction of the futures mispricing might enhance returns. Alternatively, a manager can attempt to construct an index fund from assets the manager views as mispriced. For example, the manager can construct a Government bonds index fund using what the manager believes are mispriced Government bonds. This strategy is more natural for index funds that can't use a replicating strategy because they have to hold securities in weights that differ from those of the index.

2.1.3 Performance Measurement of Active Equity Funds

The development of Performance Measurement for equity funds can be divided into two generations:

2.1.3.1 Early Models of Performance Measurement

Friend, Blume & Crockett (1970) was the first major study to consider both risk and return in examining equity mutual fund performance. They divided funds into low, medium and high risk categories where risk was defined alternately as standard deviation and beta on the S&P 500 Index. They then compared the return on funds in each risk category with a set of random portfolios of the same risk. Comparison portfolios were formed by randomly selecting securities until random portfolios containing the same number of securities as the active portfolios being evaluated. The random portfolios were divided into risk ranges similar to the active portfolios, and differences in return between the actual and random portfolios were observed. In forming random portfolios, individual stocks were first equally weighted and then market-weighted. The results were clear for one set of comparisons: mutual funds underperformed equally weighted random portfolios. The results were mixed for comparisons with market-weighted random portfolios, where funds in the high risk group appeared to outperform random portfolios. The advantage of this method over methods discussed below is that it makes no specific assumption

about equilibrium models or the ability to borrow or lend at a particular rate. On the other hand, results vary according to how the random portfolios are constructed and according to what risk ranges are examined, making results often difficult to interpret.

While this type of simulation study is an interesting way to measure performance, it is easier to judge performance if risk and return can be represented by a single number. The desire to do so led to the development of three measures that have been widely used in the academic literature and in industry. The first single index measure was developed by Sharpe (1966). Sharpe recognized that assuming riskless lending and borrowing the optimum portfolio in expected return standard deviation space is the one with the highest excess return (return minus riskless rate) over standard deviation. Sharpe called this the reward to variability ratio. It is now commonly referred to as the Sharpe ratio.

$$\frac{\bar{R}_p - R_f}{\sigma_p}$$

Where

\bar{R}_p is the average return on a portfolio

σ_p is the standard deviation of the return on a portfolio

R_f is the riskless rate of interest

This is probably the most widely used measure of portfolio performance employed by industry. This is true, though, as we discuss below, Sharpe now advocates a more general form of this model.

A second single index model which has been widely used is the Treynor (1965) measure, which is analogous to the Sharpe measure but replaces the standard deviation of the portfolio with the beta of the portfolio. Beta is defined as the slope of a regression of the return of the

portfolio with the return of the market. This measures performance as reward to market risk rather than reward to total risk.

The third single index model is due to Jensen (1968). This model can be written as

$$R_{pt} - R_{Ft} = \alpha_p + \beta_p (R_{Mt} - R_{Ft}) + e_{pt}$$

α_p is the excess return of the portfolio after adjusting for the market

R_{pt} is the return on portfolio p at time t

R_{Ft} is the return on a riskless asset at time t

R_{Mt} is the return on the market portfolio at time t

β_p is the sensitivity of the excess return on the portfolio t with the excess return on the market

e_{pt} is the excess return of portfolio p at time t not explained by the other terms in the equation

This measure has a lot of appeal because α_p represents deviations from the Capital Asset Pricing Model and as such has a theoretical basis. The Jensen measure can also be viewed as how much better or worse did the portfolio manager do than simply holding a combination of the market and a riskless asset (which this model assumes can be held in negative amounts) with the same market risk as the portfolio in question.

While these models remain the underpinning of most of the metrics that are used to measure mutual fund performance, new measures have been developed which lead to a more accurate measurement of mutual fund performance.

2.1.3.2 The New Generation of Measurement Model

The models discussed in the last section have been expanded in several directions. Single index models have been expanded to incorporate multiple sources of risk and more sophisticated models of measuring risk and expected return have been developed

2.1.3.2.1 Multi-Index Benchmarks Estimated Using Returns Data

Viewing a portfolio as a combination of the market and the riskless asset ignores other characteristics of the portfolio which affect performance. Merton (1973) suggests that an investor may be concerned with other influence such as inflation risk. Ross (1976) develops the arbitrage pricing model (APT) which shows how returns can depend on other systematic influences. These developments lead to researchers considering a generalization of Jensen's model:

$$R_{pt} - R_{Ft} = \alpha_p + \sum_{k=1}^K \beta_{pk} I_{kt} + e_{pt}$$

Where the I's represent influences that systematically affect returns and the β 's sensitivity to these influences.

What (I's) or systematic influences should be used in the model? The literature on performance measurement has employed several methods of determining the "I's." They include:

1. Indexes based on a set of securities that are hypothesized as spanning the major types of securities held by the mutual funds being examined.
2. Indexes based on a set of portfolios that have been shown to explain individual security returns.
3. Indexes extracted from historical returns using forms of statistical analysis (factor analysis or principal components analysis).

These approaches are described below.

Indexes based on the major types of securities held by a fund.

The first attempts to expand beyond the single index model were performed by Sharpe (1992) and Elton, Gruber, Das and Hlavka (1993). The motivation for EGD&H's development of a three-index model (the market, an index for small stocks and an index for bonds), was the work of Ippolito (1989). Unlike earlier studies, he found that mutual funds had, on average, large

positive alphas using Jensen's model. Furthermore, funds that had high fees tended to have higher alphas after fees. The period studied by Ippolito was a period when small stocks did extraordinarily well, and even after adjusting for risk, passive portfolios of small stocks had large positive alphas. Realizing that his sample included many funds that invested primarily in mid-cap or small stocks and small-cap stock funds tend to have bigger fees explains Ippolito's results. By including indexes for small stocks and bonds (Ippolito's sample included balanced funds), the surprising results reported by Ippolito were reversed. Funds on average tended to have negative alpha, and those funds with high fees tended to perform worse than funds with low fees.

Simultaneously with the EGD&H exploring the return on plain vanilla US stock funds, Sharpe (1992) was developing a multi-index model to explain the return on a much more diverse set of funds. He employed 16 indexes to capture the different types of securities that could be held by a wider set of funds.

The type of analysis performed by EGD&H and Sharpe not only produced better measurement of performance, but it allowed the user to infer, by observing the weights on each index, the type of securities held by the fund. This type of analysis has become known as return-based style analysis. It allows style to be inferred without access to individual fund holdings. EGD&H and Sharpe differ in the way they estimate their models. EGD&H use OLS, while Sharpe constrains each beta to be non-negative and the sum of the betas to add to one. Performance is estimated by Sharpe from a quadratic programming problem that minimizes the squared deviations from a regression surface given a set of linear constraints on the sign and the sum of betas. The advantage of Sharpe's approach is that the loading on each type of security can be thought of as a portfolio weight. The disadvantage is that by introducing additional constraints, the model does not fit the data as well.

Indexes based on influences that explain security characteristics.⁵

While authors have continued to use security-based models, often adding indexes to better capture the types of securities held (e.g., foreign holdings), a particular form of multi-index model has gained wide acceptance. This model is based on Fama and French's (1996) findings that a parsimonious set of variables can account for a large amount of the return movement of securities. The variables introduced by Fama and French include, in addition to the CRSP equally weighted market index minus the riskless rate, the return on small stocks minus the return of large stocks, and the return of high book-to-market stocks minus the return of low book-to-market stocks.

While the Fama-French model has remained a basic multi-index model used to measure portfolio performance, in many studies two additional variables have sometimes been added. The most often-used additional index was introduced by Carhart (1997). Drawing on the evidence of Jegadeesh and Titman (1993) that stock returns, in part, can be predicted by momentum, Carhart added a new variable to the three Fama French variables – momentum. Momentum is usually defined as follows: the difference in return on an equally weighted portfolio of the 30% of stocks with the highest returns over the previous 12 months and a portfolio of the 30% of stocks with the lowest return over the previous 12 months.

The idea behind incorporating this index is a belief that past return predicts future return and management should not be given credit for recognizing this. Later we will examine additional attempts to correct management performance for other types of publicly available information. Unlike indexes that represent sectors of the market such as large stocks, where index funds are readily available, the question remains as to whether management should be given credit for

⁵ One and two may seem similar. The difference is that one incorporates the types of securities held by a fund, while two incorporates influences (which may be portfolios of securities) but are used because they explain security returns.

incorporating publicly available information into portfolio decisions. To the extent that vehicles don't exist to take advantage of this and the correct way to incorporate this information is not clear, a case can be made for not incorporating these indexes.

Another addition to the Fama-French or Fama, French and Carhart models is to add a bond index to the model. The index is usually constructed as the return on a long-term bond index minus the return on the riskless rate. Its introduction is intended to adjust for the fact that many managers hold long-term bonds in their portfolio and that these securities have characteristics not fully captured by the other variables in the Fama-French model. Failure to include this index means that funds which have bonds other than one month T-bills will have the difference in performance between the bonds they hold and T-bills reflected in alpha. The effect of this on performance has been documented in Elton, Gruber and Blake (1996c).

Indexes extracted from historical returns.

Another approach to identifying the appropriate indexes to use in the performance model is to use a form of statistical analysis (factor analysis or principal component analysis) to define a set of indexes (portfolios) such that the return on this set of portfolios best explains the covariance structure of returns and reproduces the past returns on securities and portfolios. Connor and Korajczyk (1986 and 1988) present the methodology for extracting statistical factors from stock returns, and Lehman and Modest (1987) apply the statistical factors to evaluating mutual fund performance. This methodology continues to be used to evaluate mutual fund performance.

Performance Measurement Using Multi-Index Models

Most studies employing multi-index models and the Jensen measure use the α estimated from a multi-index model directly as a performance measure replacing the single index alpha.

Sharpe has suggested an alternative to the traditional Sharpe Measure called the Generalized Sharpe Measure that is an alternative to using alpha directly. In this measure a benchmark return replaced the riskless rate in the numerator of the traditional Sharpe Measure and is used to define the denominator. Define the benchmark as:

$$R_{Bt} = \sum_{k=1}^K B_{pk} I_{kt}$$

Sharpe (1992) formulated the generalized Sharpe measure as the average alpha over the standard deviation of the residuals or in equation form⁶

$$\frac{\frac{1}{T} \sum_{t=1}^T (R_{pt} - R_{Bt})}{\left[\frac{1}{T} \sum_{t=1}^T [R_{pt} - R_{Bt}]^2 \right]^{1/2}}$$

While the use of multi-index models estimated from a time series regression have been widely used to infer performance and style, several researchers have suggested using holdings data to correct potential weaknesses in time series estimation. .

Using portfolio composition to estimate portfolio betas

The models discussed to this point estimate betas from a time series regression of portfolio returns on a set of indexes. One difficulty with this approach is that it assumes that betas are stable over the estimation period. However, if management is active, the betas on a portfolio may shift over time as management changes the composition of the portfolio. Because portfolio weights changes as a function of management action, the estimates of portfolio betas from time series regression may not be well specified.⁷ Potentially better measure of the betas on

⁶ Despite Sharpe's article describing and defending the generalized Sharpe ratio, industry practice and much of the literature of financial economics continues to use the original Sharpe ratio in evaluating performance. Note that Sharpe has the riskless rate as a variable in his benchmark. If one used the normal regression procedure, portfolio returns and index returns would need to be in excess return form.

⁷ Wermers (2002) documents a significant amount of style drift for mutual funds over time.

a portfolio at a moment of time can be estimated by combining the betas on individual securities with the weight of each security in the portfolio at that moment of time. This approach to estimating alphas has been examined by Elton and Gruber writing with others (2010b, 2011a, 2011b) in three contexts: forecasting future performance, to discern timing ability, and to study management reaction to external phenomena. The results indicate significant improvement is obtained by estimating betas from portfolio holdings.⁸

2.1.3.2.2 Using Holdings Data to Measure Performance Directly

A second approach to using holdings-based data was developed by Daniel, Grinblatt, Titman and Wermers (1997). Daniel et al formed 125 portfolios by first sorting all stocks into five groups based on market capitalization, then within each group forming five groups sorted by book-to-market ratios, and finally within these 25 groups five groups by momentum. Passive returns on each of the 125 portfolios are then calculated as an equally weighted average of the return on all stocks within each of the 125 groups. The benchmark return for any fund is found by taking each stock in a fund's portfolio and setting the benchmark return for each stock as the return on the matched cell out of the 125 cells described above. They then used the benchmark described above to measure security selection as follows.

$$\alpha_p = \sum_{i=1}^N w_{it} (R_{it} - R_{iB})$$

Here the weight w_{it} on each stock at the end of period is multiplied by the return on that stock in period t to $t+1$ (R_{it}) minus the return that would be earned on a portfolio of stocks with the same book-to-market, size, and momentum (R_{iB}), and the result summed over all stocks in the portfolio. This approach, like the Fama French Carhart approach, assumes we have identified

⁸ Two other studies have used this method of estimating betas in timing studies, but these will be reviewed later in this chapter under Timing.

the appropriate dimensions of return. It does not assume the linear relationship between characteristics and return inherent in a regression model. On the other hand, the cost of the approach in terms of data is great and the comparisons are discrete in the sense that comparison is made to the average return in one of 125 cells rather than as a continuous variable.

Another approach to using portfolio composition to measure performance has become known as the weight-based measure of portfolio performance. The basis of this measure is the research of Cornell(1979) and Grinblatt and Titman (1989a and 1989b). Many portfolio holdings measures are based on comparing performance to what it would have been if the manager hadn't changed the weights. The idea is simple and appealing. If the manager increases weight on securities that do well in the future and decreases weights on securities, and do poorly, he or she is adding value. Perhaps the most widely used measure here is the Grinblatt and Titman (1993) measure:

$$GT_{pt} = \sum_{i=1}^N (w_{i(t-1)} - w_{i(t-h-1)}) R_{it}$$

Where

GT_{pt} is the performance measure for a portfolio in month t

$w_{i(t-1)}$ is the portfolio weight on security i at the end of month $t-1$

$w_{i(t-h-1)}$ is the portfolio weight on security i at the end of h months before $t-1$

R_{it} is the return on stock i during month t

Summing the above equation over multiple periods gives a measure of performance for any fund. Note that the benchmark for the fund now becomes the return on the fund that would have been earned if the composition of the fund had been frozen at a point h periods before the current period. Note that the sum of the weights add up to zero so that the measure can be viewed

as the return on an arbitrage portfolio, and that the performance of securities held in the portfolio in unchanged weights is not captured. The holdings-based measures are all pre-expenses. Thus holdings-based metrics don't measure the performance an investor in the fund would achieve, but rather whether the manager adds value by his or her security selection.

2.1.3.2.3 Time Varying Betas

The regression techniques described earlier assume that the sensitivities of a fund to the relevant characteristics remain constant over time. Using holdings data to estimate betas is one way of dealing with changing betas.

An alternative to using holdings data to estimate changing betas is to fit some functional form for how betas change over time.

2.1.3.2.4 Conditional Models of Performance Measurement, Bayesian Analysis, and Stochastic Discount Factors

Three approaches have been set forth as a modification of the standard models of portfolio performance. The first recognizes that the risk sensitivity of any mutual fund can change over time due to publicly-available information, the second uses Bayesian techniques to introduce prior beliefs into the evaluation process and the third uses stochastic discount factors.

Conditional Models of Performance Measurement

The philosophy behind conditional models of performance measurement is that sensitivity to indexes should change over time since return on these indexes is partially predictable. Furthermore, management should not be given credit for performance which could be achieved by acting on publicly available information that can be used to predict return. We have already briefly discussed this philosophy when we examined the Carhart model.

In a broader sense, the extreme version of the conditional model says that superior performance occurs only if risk-adjusted returns are higher than they would be based on a

strategy of changing sensitivity to indexes by using public information in a mathematically defined manner.

Fearson and Schadt (1996) develop one of the best-known and often-used techniques for conditional beta estimation. Their version of the traditional CAPM specifies that risk exposure changes in response to a set of lagged economic variables which have been shown in the literature to forecast returns. The model they specify is

$$R_{pt} - R_{ft} = \alpha_p + \beta_p(Z_t)(R_{Mt} - R_{ft}) + e_{pt}$$

Where $\beta_p(Z_t)$ is the value of the conditional beta (conditional on a set of lagged economic variables) at a point in time. These conditional betas can be defined as

$$\beta_p(Z_t) = \beta_{p0} + \beta_{p1}Z_t$$

Where Z_t represents a set of conditioning variables. Fearson and Schadt define the conditioning variable in their empirical work as:

Z_1 is the lagged value of the one-month Treasury bill yield

Z_2 is the lagged dividend yield on the CRSP value weighted New York and American Stock Exchange Index

Z_3 is a lagged measure of the slope of the term structure

Z_4 is a lagged measure of the quality spread in the corporate bond market

Z_5 is a dummy variable for the month of January

The generalization of this approach to a multifactor return-generating model is straightforward. We replace the prior equation with a generalization to a K-factor model.⁹

$$R_{pt} - R_{ft} = \alpha_p + \sum_{k=1}^K \beta_{pk}(Z_t)I_{kt} + e_{pt}$$

⁹ The conditioning variables employed may change depending on the factors used.

Where I_{kt} are the factors in the return-generating process $\beta_{pk}(Z_t)$ is the sensitivity to factor K at time t and Z_t is as before

Elton, Gruber and Blake (2011a and 2011b) use holdings data to measure factor loading (betas) at monthly intervals and to test whether changes in these betas are related to the variables hypothesized by Ferson and Schadt. They find that the set of conditional variables hypothesized by Ferson and Schadt explains a high percentage of the movement in actual portfolio betas over time.

Christopherson et al (1998 a & b) propose that α as well as betas are conditional on a set of lagged variables. This involves one new relationship:

$$\alpha_p(Z_t) = \alpha_{p0} + \alpha_{p1} Z_t$$

Conditional models have also been developed for some of the weight-based measures discussed earlier. Ferson and Kong (2002) develop such a model where the manager gets no credit for changes in the weight and portfolio returns which are based on public information.

Mamaysky, Spiegel and Zhang (2007) take a different approach to measuring performance, with time varying coefficients. Rather than hypothesizing a set of lagged variables that help to determine betas at a period in time, they used Kalman filters to determine the time pattern of betas and performance over time. This allows the pattern to be determined by a set of variables that are statistically estimated rather than hypothesized by the researchers.

Bayesian Analysis¹⁰

A number of authors have used Bayesian analysis to continuously adjust the alpha resulting from a multi-index model. Baks, Metrick and Wachter (2001) assume that an investor

¹⁰ Stambaugh (1997) showed how movements of assets with long histories can add information about movements of assets with shorter histories, thus one reason to examine non-benchmark assets is that they may have a longer history.

has prior beliefs concerning whether any manager has skill. They use this prior and the history of returns to compute the posterior α using Bayesian analysis.

Pastor and Stambaugh (2000) assume a multi-index model. First they divide their indexes into those that an investor believes are in a pricing model and those that are not (labeled non-benchmark assets). Pastor and Stambaugh (2002) show that if non-benchmark assets are priced by benchmark assets exactly, then α s are completely unchanged by the choice of an asset pricing model. However, if they are not priced exactly, different models will produce different estimates of alpha and by incorporating a set of non-benchmark passive portfolios on the right-hand side of the return regression a better estimate of alpha is obtained. Pastor and Stambaugh assume investors have prior beliefs on how certain they are that they have correctly identified the correct asset pricing model and use Bayesian analysis to update these beliefs.¹¹

Stochastic Discount Factors

Several authors (Chen & Knez (1996), Farnsworth, Ferson, Jackson & Todd (2000), and Dahlquist & Soderlind (1999)) have tried to estimate stochastic discount factors and then have evaluated mutual funds as the difference between the funds' performance and the return on the fund if it earned the equilibrium return using the stochastic discount function. The idea is parallel to Jensen's alpha when the single factor model is interpreted as the CAPM model.

2.1.4 Measuring the Performance of Active Bond Funds

While there has been a vast literature on models for evaluating stock mutual funds, the literature dealing with the performance of bond funds is much less developed. This is true despite the fact, as was pointed out in the introduction, bond funds constitute a significant proportion of mutual fund assets.

¹¹ The Pastor-Stambaugh framework was applied by Busse and Irwin (2006) to daily data.

The first paper to present a detailed analysis of bond fund performance was Blake, Elton and Gruber (1994). In this paper the authors employ regression models of the type discussed earlier, as well as the QPS version of this model developed by Sharpe (1992). Blake, Elton and Gruber investigated a one-index model (either a general bond index or the submarket index that Morningstar identified as most like the bond fund), two three-index models, and a six-index model.

The six indexes were based on the major types of securities held by the fund and included an intermediate government bond index, a long-term government bond index, an intermediate corporate bond index, a long-term corporate bond index, a high-yield bond index, and a mortgage bond index. Unlike stocks, where performance seems extremely sensitive to the choice and definition of the indexes employed, the results for bond funds seem to be fairly robust across models as long as three indexes are used. The three indexes needed were a general bond index, a high yield index, and either a mortgage or term structure index.

Elton, Gruber and Blake (1995) built on their earlier work in bond mutual fund performance by developing a set of indexes that might be relevant for the pricing of individual bonds rather than indexes representing the major bond sectors. Following in the spirit of Chen, Roll and Ross (1986), Elton, Gruber and Blake (1995) employed both time series and cross-sectional tests on bond pricing and developed a new six-index model of bond pricing. The six variables included an aggregate index of stock returns, an aggregate index of bond returns, a measure of risk premium in the bond market (return on high yield bonds minus a government bond index), a series to represent option valuation (the return on mortgage bonds), and finally two variables to measure unanticipated changes in economic variables. The economic variables which were significant related to bond return were unanticipated changes in inflation and

unanticipated changes in GNP. While using unanticipated changes in economic variables is very much in the spirit of Chen, Ferson and Peters (2010), what makes this study stand out is the use of actual expectational data from consumer surveys and professional forecasters rather than derivations from historic data to estimate expectations and unanticipated changes in expectations.

Having developed and tested the model on bonds and passive bond portfolios, the six-index model is then applied to evaluating bond fund performance. The model not only produces reasonable estimates of performance; the estimates of performance (α) are not a function of the declared objective of the funds, a result that is not often found with alternative models.

Comer and Rodriguez (2006) continued the use of the major types of securities held by the fund to evaluate investment grade, corporate and government bond funds. In addition to a single index model, Comer and Rodriguez test a six-index model.¹² The six indexes they employ include three corporate government maturity return indexes (1 to 5 years, 5 to 10 years, and beyond 10 years), the return on high-yield bonds, the return on mortgages, and the return on Treasury bills. The models are first used to identify style, and then used to identify α and timing. They find negative alphas for bond funds, and while there is some difference between sectors, the rank correlation of bond funds across different models is very high. Another interesting finding of this paper is that net flows into bond funds follow risk-adjusted performance.

Chen, Ferson and Peters (2010) measure performance net of timing ability. This study differs from others in that it clearly differentiates timing from selectivity. Timing ability is the ability to use information to time the realization of factors in the performance model. Selectivity in performance is the use of information to select specific securities that will do well. Chen, Ferson and Peters chose indexes based on the term structure of interest rates, credit spreads,

¹² They also used a five-index model which includes the three government maturity variables and adds a corporate variable and a general government variable.

liquidity spreads, mortgage spreads, exchange rates, and a measure of dividend yield and equity volatility. This study adds liquidity, dividend yield, equity volatility and exchange rates to the set of variables employed in previous studies. The last three variables are added because the authors note that bond funds hold international bonds and a belief that bond returns are affected by stock market volatility.

One thing that distinguishes this study from previous studies is that the authors model and correct for non-linearities in the regression model that are unrelated to a bond fund manager's timing ability. They discuss and model four influences. The first is that the assets (e.g., options) held by the fund may have a non-linear relationship with the factors driving return. The second is the ability of management to generate fake timing ability by changing exposure between the interval over which returns are measured. The third can arise because of stale pricing. The fourth arises because portfolio betas may be correlated with market return because of their common reliance on public information.

Chen, Fearson and Peters (2010) correct for these influences and construct a model that measures corrected market timing and selectivity. The results report no real market timing and alphas negative but smaller in absolute value than expense ratios.

2.1.5 Measuring Timing

There are two different ways that investigators have measured timing. One is using return data and one is using holdings data.

2.1.5.1 Return Measures

Principally, there were two methods of evaluating mutual fund timing using returns data: the

Treynor and Mazuy (1966) method, and the Henriksson (1984) and Henriksson and Merton (1981) method. Treynor and Mazuy measured timing by putting a squared term in the equation.

If the single-model index was used to measure timing, then:

$$R_{pt} - R_{Ft} = a_p + \beta_p (R_{Mt} - R_{Ft}) + \beta_{pT} (R_{Mt} - R_{Ft})^2 + e_{pt}$$

Where the β_{pT} term measures timing and other terms as before. The basic logic behind the measure is that a manager with timing ability will increase the beta when high market returns are expected and decrease when low returns are expected, and this will induce curvature in the beta return relationship which can be seen by a positive β_{pT} . If a multi-index model is used, then one needs a squared term with each index where one wishes to measure timing. The coefficient on the squared term is the measurement of timing with respect to that factor.

The Henriksson and Merton Model estimate timing by assuming the manager has two betas: one in up markets and one in down markets. For the single index model timing is estimated with the following model:

$$R_{pt} - R_{Ft} = a_p + \beta_p (R_{Mt} - R_{Ft}) + \beta_{pT} C (R_{Mt} - R_{Ft}) + e_{pt}$$

Where C is a dummy that has value of one if $R_{Mt} > R_{Ft}$ and zero if $R_{Mt} \leq R_{Ft}$. Thus β_{pT} measures the differential beta in markets where the index outperforms the risk-free rate. If a multi-index model is used, then an additional term is used for each factor where timing is to be measured. Like in the measurement of performance there is an issue of whether conditional betas should be used. Becker, Ferson, Myers & Schill (1999), Ferson and Schadt (1996), and Ferson and Qian (2006) measure timing using conditional betas. By conditioning betas on a set of variables that are related to return (such as the dividend price ratio), the influence of these variables on timing is removed from the timing measure.

2.1.5.2 Holding Measures of Timing

Elton, Gruber and Blake (2011b), Daniel, Grinblatt, Titman and Wermers (1997), and Jiang, Yao and Yu (2007) use holdings data to estimate mutual fund betas and to measure timing. Since the betas on a portfolio are a weighted average of the betas on the securities that comprise the portfolio, there is an alternative way to estimate a mutual fund's beta. They can be estimated by first estimating each security's betas, then using holdings data to obtain security proportions, and finally using the product of security betas and proportions to get the mutual fund betas. The advantage of this approach is that it avoids the following problem: if management is changing the composition of a portfolio over time (e.g., because it is engaging in timing) the betas on the fund from a time series regression of fund returns will be poorly specified. Using holdings data at each point in time that holdings are observed provides a direct estimate of the betas on each of the relevant factors for the fund.

Elton, Gruber and Blake (2011b) measure timing using a method parallel to how alpha is measured. They measure timing as the difference in performance between the actual beta and the target beta (specified below) at the end of the period times the return in the next period. In equation form for any index:

$$Timing = \sum_{t=1}^T \frac{(\beta^* - \beta_t) R_{pt}}{T}$$

Where

1. β_{At} is the actual beta in period t
2. β_t^* is the target beta in period t
3. T is the number of time periods
4. Other terms standard

This measure captures whether the fund deviated from the target beta in the same direction as the return on the index deviated from its normal pattern. Does the fund increase its beta when index returns are high and decrease when index returns are low?

There are several possibilities about what to use for a target beta. For a plan sponsor trying to evaluate a fund that professes to be a timer and has an agreed-upon normal beta, the target beta might be the agreed-upon beta. For an outside observer, the average beta over time might be a reasonable choice. Finally, if one believes factors can be forecasted and the forecasting procedure is widely known, and if one also believes that the manager shouldn't get credit for using this public information, then the target beta could be the forecasted beta. For example, if one believes that the market can be forecasted by the dividend price ratio and that the manager should not be given credit for changing beta in response to changes in dividends over price, then beta forecasted by the dividend price ratio could be used as a target beta. Ferson and Schadt (1996) discuss how to capture changing beta from public information when timing is measured using historical returns. The same idea can be used here. The Elton, Gruber and Blake (2011b) measure is similar in concept to one developed earlier by Daniel, Grinblatt, Titman and Wermers (1997). As a target beta, Daniel, Grinblatt, Titman and Wermers (1997) use the actual beta from a prior period. The difference in beta is then the change in the beta from the prior period.¹³ Finally, Jian, Yao and Yu (2007) measure timing in a different manner. They show that the Treynor and Manzuy measure implies that:

$$\beta_t = a + \gamma R_{t+1} + e_t$$

Where

1. β_t is the beta at the end of period t estimated from portfolio holdings as described earlier.

¹³ Admati, Bhattacharya, Pfleiderer and Ross (1986) provide theoretical underpinning for the measure.

2. R_{t+1} is the return in a period subsequent to period t
3. Other terms standard

They used multiple lengths of the subsequent returns for R_{t+1} 1, 3 or 6 months to test timing.

Timing is then measured as the significance of Gamma.

2.2 How Well Have Active Funds Done?

As discussed earlier, the single index model can classify all funds as good or bad performers simply because a segment of the market did well or poorly. For example, as shown in Elton, Gruber, Das and Hlavka (1993) during a period studied by Ippolito (1989), passive small stock portfolios did spectacularly well with a yearly alpha of 10% when alpha is measured using the S&P Index. Since Ippolito's sample included many funds that invested heavily in small stocks, this leads to a large positive alpha on average over the funds he studied. When a multi-index model is used and a small stock index is included, the positive alpha found by Ippolito for the average fund becomes negative. Thus in analyzing performance we will primarily summarize results from multi-index models.

Table 1 presents summary results from a large sample of mutual fund studies. It is divided into five sections. Section A presents results for mutual fund performance using measures of performance based on betas estimated from running a time series regression of either mutual fund returns or the securities they hold on various indexes. Panel B summarizes studies using holdings-based measures of pre-expense performance. Panel C presents results on mutual funds' timing ability, and Panel D shows results on bond fund performance. Finally, Panel E summarizes results on the persistence of mutual fund performance.

The results in Panel A are consistent with one exception. Mutual funds underperform passive portfolios by from 65 basis points to 2% depending on the set of indexes chosen, the

methodology, and the time period chosen.¹⁴ These results are post-expenses. If expenses are added back, most of these studies would find positive pre-expense performance. Thus managers have selection ability, but not enough to cover expenses. Panel B tells the same story. Holdings-based performance measures are computed and reported pre-expenses. The pre-expense performance, in most cases, is less than expenses, thus net of expenses performance is negative.¹⁵

The results from timing studies are less uniform. Early studies found no evidence of timing. However, Bollen and Busse (2001) found significant positive timing using daily data and a time series regression, and Kaplin and Sensoy (2005) and Jiang, Yoo and Yu (2007) find positive timing using holdings data. All of these studies measure timing by looking at changes in the sensitivity to a single index. Elton, Gruber and Blake (2011b) and Ferson and Qian (2006) argue that changes in the sensitivity to the market often come about because of changes in sensitivity to other factors. For example, a fund moving into smaller stocks will usually increase its market sensitivity. When these latter two studies measure timing, taking into account not only changes in the sensitivity to the market but also changes in the sensitivity to other factors, they find no evidence of successful timing even though they find successful timing when timing is measured using only a market index.

The performance of bond funds (alpha) after expenses (Panel D) is also universally found to be negative. Like stock funds pre-expenses, most studies find performance is positive pre-

¹⁴ The exception is Ferson and Schadt (1996), who find positive alphas. This is somewhat surprising since their main methodology contribution is to remove the impact of management reacting to public information from performance measurement. Two studies justify the existence of actively managed mutual funds despite negative alphas. Moskowitz (2000) suggests and provides some evidence that active mutual funds perform better in recessions and are therefore potentially desirable relative to index funds. Glode (2011) provides a theoretical model that justifies average underperformance of active mutual funds because of superior performance in recessions as well as some evidence supporting this performance.

¹⁵ The exception is Grinblatt and Titman (1993) who find positive results larger than expenses. At least part of their results can be explained by the presence of selection bias in their sample.

expenses, again indicating some management skill but that management skill is smaller than expenses.

The average results for bond and stock mutual funds would suggest that randomly selecting funds is worse than indexing. From the investor standpoint, an important issue is whether there is persistence in performance, and of even more importance, whether there are some funds that outperform index funds, and can these be identified in advance.

The first problem is how to identify funds that will perform well in the future. Three metrics have often been used: past return, past alpha, and past alpha over standard deviation of residuals (the generalized Sharp ratio). Similarly, the evaluation criteria of the funds selected has been return, alpha from various models, and the generalized Sharpe ratio. If past return is used to rank funds, ranking is likely to be highly related to style. There are clearly long periods of time where small or large or value or growth funds have produced higher returns. For example, as mentioned earlier in Ippolito's (1989) sample period due to the performance of small stocks, small stock funds consistently outperform large stock funds with an alpha from the single index model of over 10%.¹⁶ Clearly, ranking and evaluating over this period using the single index model would show persistence. However, evaluation using a multi-index model that accounts for small stocks might not show persistence. Thus, ranking on either alpha or alpha over residual risk from a multi-index model is more likely to uncover real persistence in managerial ability if it exists.

Poor performance is easy to predict. Almost every study finds that poor performance in one period predicts poor performance in subsequent periods. One characteristic of the poor-performing group is high expenses. It seems that if you charge enough, you can do poorly in

¹⁶ Small stock alphas in the years after Ippolito's study were often negative when measured using the market model.

every period. While useful, the ability to predict poor performing funds does not suggest a trading opportunity since these funds can't be sold short. Thus studies that report the difference in predicted return between the top and bottom decile may not be supplying information that is useful.

The real issue from an investor's point of view is whether a group of good-performing funds can be identified and more importantly can funds be identified that will outperform index funds in the future.¹⁷ Outperformance should be judged by positive alpha from an appropriate multi-index model. Consistently, investors have found a positive alpha over subsequent periods when ranking is done by alpha or alpha over residual risk. These studies include Carhart (1997), who found when funds were ranked by alpha the top ranked group had positive alphas over the next five years; Busse and Irvine (2006), who found persistence and positive alphas using Bayesian estimates; Gruber (1996); Elton, Gruber and Blake (1996c), and Cohen, Coval and Pastor (2005), all of which find persistence for the top-ranked group and that the top group has a positive alpha.¹⁸

In addition, Baker, Litov, Wachter and Wurgler (2004) provide evidence that managers can select superior-performing stocks and that there is persistence in the ability of individual managers to do so.

The principal criticism of these studies is that if there were a missing factor in the ranking model and its performance were correlated over time, we could observe persistence when none exists. However, researchers have used so many different time periods and so many different factor models that it is unlikely that there is a missing factor in all models and that the factor is giving consistent alphas over all periods studied.

¹⁷ A second question is, if predictability exists, how long does the outperformance last?

¹⁸ Carhart (1997) is usually quoted as not finding persistence, and he doesn't for the top group when ranking is by return. He does find persistence in the top-ranked group using a multi-index model when ranking is by alpha.

Several articles have specifically questioned some multi-index models as capturing all relevant factors and whether the factors have been measured correctly. Chan, Dimmock and Lakonshok (2010) vary how the Fama-French factors are defined and find that the Russell 1000 growth index can have an alpha of -1.66% to +1.08% depending on how they define the factors. Cremers, Petajisto and Zitzewitz (2010) find the S&P index has a positive alpha using the Fama-French three-factor model. Elton, Gruber and Blake (1999) have suggested that growth is a more complex variable and performance might be better measured if portfolios of active funds were incorporated in an evaluation and ranking model.

A theoretical argument against predictability is presented in Berk and Green (2004). Berk and Green argue that performance decreases with size, either because of increased costs and/or the need to accept less profitable investments. Since fund flows follow performance, flows come into any fund until performance above indexes is eliminated. Whether fund flows eliminate persistence depends on how much and how quickly cost increases with size or how much and how quickly performance decreases with size, and the amount of new flows over any period. There have been four suggestions for why costs might increase or performance decrease: increasing fees, adding investments that are less promising to the portfolio (or indexing part of it), organization diseconomies, and transaction costs.

Expense ratios have two components: administrative costs (including sales costs) and management fees. For most funds the management fee schedule specifies that management fees will decrease with fund size in a particular manner. Changing the fee schedule is difficult and rarely done. Administrative costs have a large fixed component. Thus total fees as a percent of

assets decline with the size of the fund, and the relationship of expense ratios to size generally leads to performance increasing with size rather than decreasing.¹⁹

In addition, Pollet and Wilson (2008) have shown that as a fund grows larger the number of securities changes only slightly. Thus if there are diseconomies of scale it most likely involves transaction costs. These are being studied by a number of authors currently, and should shed light on how long persistence should last.²⁰ The most relevant is Christoffersen, Keim and Musto (2007). They studied Canadian mutual funds where trades have to be reported. They find that larger mutual funds have lower costs than smaller funds and that active funds have lower trading costs than passive funds. They argue the latter is likely due to bunching of trades around index changes being more costly than the trading costs caused by active managers trading on information.²¹

Fama and French (2010) provide the first direct test of Berk and Green. Fama and French (2010) point out that the Berk and Green prediction that most fund managers have sufficient skill to cover their costs is not supported by the data. They examine the cumulative distribution of net returns using bootstrap simulation and conclude that bad-performing funds have risk-adjusted returns that are extremely unlikely to have arisen by chance, while those funds that have done extremely well may have obtained these results by chance. They do find that in the upper tail of performance there may be some funds that exhibited superior performance at a statistically significant level. Chen, Hong, Huang and Kubik (2004) find that performance decreases with size, and attribute this to organization diseconomies. However, despite this, they find predictability of performance. While Green and Berk may well have a point that a fund can be so

¹⁹ Elton, Gruber & Blake (2011d) provide evidence that mutual funds that are successful decrease fees.

²⁰ See, for example, Edelen, Evans and Kadlec (2009) and Yan (2008).

²¹ Keim and Madhavan (1995, 1997) find that execution size increases transaction costs for institutional traders using plexus data. However, they cannot tell if the smaller orders were simply a bigger order being executed as a series of small orders or a small order.

large that superior performance can't be achieved, where that point is and how soon it is reached is an open question. Performance would seem to persist, but more research is needed on how long it persists.

2.3 How Well Do Investors Do in Selecting Funds?

There is very little evidence on the wisdom of investor behavior in selecting mutual funds. The bulk of the evidence involves investor choice of mutual funds within 401(k) pension funds. 401(k) plans are studied because they are an important element of investor wealth and data on investor choices of mutual funds within plans is available.²²

Most of the research has called into question the rationality of investor decisions. For example, a common practice of investors is to allocate equal amounts to all the funds they are offered (see Benartzi and Thaler (2001), Liang and Weisbenner (2002), and Huberman and Jiang (2006)). There are clearly some utility functions where this is optimum, but it is unlikely to be the preferred choice across all market conditions and for all ages of participants. Another questionable aspect of decision-making is that investors increase their allocation to an asset category if more choices are offered in that category. Benartzi and Thaler, using primarily experimental data, find strong evidence of this. Huberman and Jiang, using data from Vanguard, find that while there is a tendency for this to occur, it is only marginally significant.

Yet another seemingly irrationality is that investors seldom change their allocation of their pension assets across options in a plan (see Agnew and Balduzzi (2010) and Ameriks and Zeldes (2004)). This means that as time passes their overall allocation is heavily dependent on the past returns of the various asset categories they hold. In addition, when they do transfer money it is primarily to the categories that have the highest returns in the recent past (Elton,

²² Elton, Gruber & Blake (2007) discuss that for most investors the 401(k) is their sole wealth outside of a savings account.

Gruber and Blake (2007) and Agnew and Balduzzi (2010)). Unless we believe past return predicts future return, and there is no evidence to support this, their actions are non-optimal.²³

Yet another irrationality is that investors' choices are heavily influenced by the default options they are offered (Madrian and Shea (2001)). Furthermore, they tend to stay with the default option over time. Thus, if the default choice is a money market account, they continue to place new contributions in this option.

There is some evidence on how well investors make decisions outside of pension funds. Elton, Gruber and Busse (2004) explore investors' choice of S&P 500 index funds. S&P 500 index funds are a commodity in that they hold the same assets. Furthermore, portfolio management actions such as portfolio lending have only a small impact on results. Thus the principal factor influencing return to investors is the expense ratio charged. Selecting a low-cost fund will result in selecting a better-performing fund in the future. In addition, selecting a fund that had a higher past return which captures both expenses and managerial skill leads to an even-better-performing fund in the future. It is easy to differentiate between index funds that will perform well in the future and those that will not: just buy low expense funds or funds with high past returns. Yet, EG&B find that more new money as a percent of assets under management flows into the poor-performing funds. This may well be because expense ratios are related to how much salespeople are paid to sell a fund.²⁴ If investors do not make rational decision in this simple context, it is hard to believe that they make good decisions in more complex contexts.

As discussed in the section on Performance, Gruber (1996) has some evidence that supports greater rationality of investors. High alpha funds in one period are, on average, high

²³ As discussed earlier, there is evidence that properly managed past alpha predicts future alpha, but this finding does not apply to returns.

²⁴ It has been argued that the high expenses may be associated with funds that provide better services. Service is hard to measure quantitatively. However, Elton, Gruber and Busse (2004) examined a common measure of this and found no relationship.

alpha funds in the next period. Gruber shows that funds with higher past alpha get greater flows and that funds with low alphas have large outflows. Gruber also provides some evidence of less rational behavior because he, like others, finds high past return which may not be predictive of future alpha also results in higher future flows.

2.4 Other Characteristics of Good-Performing Funds

There have been several articles that argue that characteristics of funds other than past alpha and expense ratios can be used to predict which funds will have positive alpha in the future. Kacpercykyk, Sialm and Zheng (2008) argue that the return gap can predict good performance in the future. They define the return gap as the difference in the actual return on the fund and the hypothetical return that would have been earned if the fund had continued to hold the assets reported the last time holdings were disclosed. The return gap is affected by trading costs and expense ratios since the hypothetical portfolio pays neither. However, it is also affected by the return on the changes in the securities held since the last report. They find persistence of the return gap for up to five years whether it is measured as raw returns or the difference in alpha using the four-factor (Carhart) model. They find that the return gap predicts future performance, whether performance is measure relative to the market, by the CAPM, the Fama-French three-factor model or the Carhart four-factor model.

Cremers and Petajisto (2009) compute a measure they call ‘active share.’ Active share is the absolute difference in holdings between a fund and its stated benchmark. In equation form

$$active\ share = 1/2 \sum_{i=1}^N |w_{fund,i} - w_{indexi}|$$

Where

1. $w_{fund,i}$ is the weight of stock i in the fund

2. $w_{index,i}$ is the weight of stock i in the benchmark index
3. The sum is over all assets in the universe

Active share has an intuitive explanation. Treynor Black (1973) argued that a fund's portfolio could be divided into an index fund and an active portfolio. Active share is the active portfolio described by Treynor Black (1973). Cremers and Petajisto find that active share predicts performance relative to a fund's benchmark. The funds in the quintile with the highest active share outperform their benchmarks by 1.13% to 1.15% per year, while funds with the lowest active share underperform by 1.42% to 1.82% per year. When they examine the differential return (return less benchmark) using the Carhart model, they find substantial alphas, 3.5% per year, for the top quintile with significant persistence. However, if active share is analyzed directly using the Carhart model, there is no relationship between active share and future alpha. This suggests that while active share forecasts ability to beat a benchmark, it does not predict alphas in future periods.

Cohen, Coval and Pastor (2005) argue that one can improve the prediction of future performance by selecting managers who are successful and whose portfolio decisions are similar to other successful managers. Their measure weights each stock by the average skill of the manager holding it. For stock n its value is²⁵

$$\delta_n = \sum_{m=1}^m U_{nm} \alpha_m$$

Where

1. α_m is the alpha of manager m

²⁵ They have an alternative measure that measures commonality in changes in a fund's holding. They find substantial predictability in performance one quarter ahead, with the greatest improvement in prediction for funds with short histories.

2. U_{mn} is the proportion of stock n manager m holds
3. δ_n is the quality of the stock n

A manager is then judged by the weighted average (by percent in each stock) of the δ_n in the manager's portfolio.

Chevalier and Ellison (1999) look at the relationship between performance and manager characteristics. In particular they look at the manager's age, where they graduated from college, the average SAT at the institution they graduated from, manager tenure, and whether the manager has an MBA. The one factor that seems to predict risk-adjusted performance is the average SAT scores at the schools they graduated from. As yet no one has explored all these suggestions to see if some are redundant and to explore what combination leads to the best forecasting technique.

2.5 What Affects Flows Into Funds?

Gruber (1996) was among the first studies to document the strong relationship between past performance and flows in and out of funds. He found that inflows were strongly related to performance across a broad range of performance measures. Sirri and Tufano (1998) expanded the analysis to include variables other than performance which might affect flows. They find a strong relationship between past return and inflows. For the top quintile of funds based on return, the inflow to funds is large and economically and statistically significant. For the bottom quintile of funds they found no relationship between performance and flows. Inflows are also negatively related to both expense ratios and, to a lesser extent, volatility. They find that the expense flow relationship is affected by return and that high expense funds with high prior returns get more rather than less inflow of funds. This might be due to the higher expense ratios being used in part to call attention to high returns.

Del Guercio and Tkac (2008) study the effect of Morningstar ratings on fund flows. Morningstar rates funds from one to five stars. Del Guercio and Tkac use event study methodology to study the impact of a discrete change in the ratings. In studying the effect of Morningstar rating changes they control for normal flows due to past performance and flows into the category to which the fund belongs. Thus they measure differential flows due to Morningstar rating change. They find that upgrading to 5 from 4 increases flows by 25% over normal flows, and flows continue to increase for seven months. A decrease from 5 to 4 has little impact on flows. However, a downgrade from 4 to 3 causes outflows 12 times more than expected and from 3 to 2 increased outflows much more than expected. Changes from 1 to 2 or 2 to 1 have little impact on flows. They speculate that this latter is due to all investors who are conscious of performance having already exited.²⁶ They attribute the lack of a response to a downgrade from 5 to 4 to financial planners using funds rated four stars or five, so that downgrading doesn't strongly affect financial planners' recommendations.

Massa, Goetzmann and Rouwenhorst (2001) examine flows at a macro level. They find that flows into equity funds (both domestic and international) are highly negatively correlated with flows into money market and precious metals funds. Furthermore, flows into international funds are highly correlated with flows into domestic equity funds (correlation .6 to .7), indicating that these are treated as economically similar by investors. Finally, flows into municipal bond funds are uncorrelated with flows into other bond classes (where flows are highly correlated), indicating that they are viewed as a separate category.

Chuprinin and Massa (2010) differentiate between flows to funds which cater to short-horizon investors versus funds that cater to long-horizon investors (proxied by variable annuity

²⁶ See Christofferson and Musto (2002) for similar arguments.

funds). They find that flows to long-horizon funds are more sensitive to economic cycles and less sensitive to performance than flows to short-horizon funds.

3. Closed End Funds

Closed end funds are started by a public offering in which shares of the fund are sold to the investing public and the proceeds are used to buy securities. After the initial offering, additional shares can be sold only through a new issue, and this occurs infrequently.²⁷ As discussed earlier in this chapter, the principal difference between closed end funds and open end funds is that while open end fund are sold at net asset value, closed end funds sell at a price that is almost always different and, for equity closed-end funds, usually below the market value of the assets held. The other principal difference between open and closed end funds is that closed end funds are traded on an exchange and can be traded at any time that the exchange is open, while open end funds are bought and sold during the day but investors receive (or pay) the net asset value of the underlying assets at 4:00pm on the day of the trade.²⁸

While there have been a huge number of interesting articles discussing closed end funds and the anomalies they present, we have decided to limit the discussion to two subjects: the discount or premium at which closed end funds sell, and the reasons for the existence of closed end funds.²⁹

3.1 Explaining the Discount

Explanations for the discount at which closed end funds sell include liquidity of

²⁷ The design and amount of new offerings is regulated. For example, the amount of new offering is restricted to no more than one-third of the value of the fund.

²⁸ Trades after 4:00pm receive or pay the net asset value at 4:00pm on the following day.

²⁹ The most cogent discussion of the major anomalies in the pricing of the closed end funds is presented by Lee, Shleifer and Thaler (1990). These include the premium for new funds, the cross-sectional and intertemporal behavior of discounts, and the price behavior when funds are terminated.

investments, management fees, management ability, tax liabilities, sentiment, greater risk of closed end fund returns compared to returns on their assets, and uncertainty about the size of future discounts.

Tax liability is perhaps the most straightforward explanation of closed end fund discounts. As early as 1977, Malkiel published an article with an extensive examination of the effect of tax overhang on discounts. When one buys a closed end fund that holds securities with a capital gain, one owns a share in the assets and a share in a future potential tax liability. However, given the high turnover of most domestic closed end funds, the tax overhang should be small.³⁰ Malkiel estimates that even with very high estimates of capital gains overhang, that overhang can account for only a small part of the discount at which closed end funds sell. In addition, as Lee, Shleifer and Thaler (1991) point out, a capital gain explanation for discounts predicts that discounts should increase when returns are high, but in fact there is no correlation between discounts and returns. However, in more recent articles Brennan and Jain (2007) examine the behavior of closed end funds around capital gains and dividend distributions, and find evidence that there is an effect of tax overhang. There is no doubt that tax overhang affects the pricing of closed end funds. However, it appears to account for only a small portion of the discount.

A number of authors have investigated explanations for the discount using expenses or the tradeoff between management ability and expenses. Malkiel (1977) looks at the relationships between discounts and management expenses, and finds no relationship. In a more recent article, Kumar and Noronha (1992) do find a positive relationship between expenses and discounts. Expenses should be examined in combination with performance. If management produces

³⁰ Tax overhang is only a deterrent to the extent the capital gains are realized while the investor is in the fund. Tax overhang could be correlated with turnover or performance which might impact the results.

superior performance before expenses, the question remains whether the net result of management ability and expenses can account for the discount. Returning to Malkiel's (1977) article, he finds a relationship between discounts and future performances net of expenses. Cherkes et al (2003) argue that the discount can be explained by the capitalized value of the services management adds less the capitalized value of the cost of such service. One of the services management provides is to supply lower-cost access to less liquid securities.

Berk and Stanton (2007) provide one of the more compelling explanations of the discount. Their argument is that if management is entrenched, poor management relative to expenses leads to a discount. However, if management is free to leave when performance is good, management will capture the extra performance in higher fees or leave for a different job. Thus the balance of expenses and performance means that an average fund sells at a discount.

Cherkes, Sagi and Stanton (2009) offer an alternative explanation for the discount on closed end funds. While they recognize the tradeoff between management ability and expenses, they study the problem as a tradeoff between the benefits of the closed end funds' ability to hold illiquid assets and the costs of this form of organization. We will discuss this further in the next subsection.

While a number of studies discuss the factors mentioned above, two explanations have been offered for the size and existence of the discount: one based on behavioral and one based on capital market characteristics. A well-known series of papers by Lee, Shleifer and Thaler (1990 and 1991), DeLong and Shleifer (1992) and Chopra, Shleifer and Thaler (1993) explains the discount on closed end funds by the irrational sentiment of retail investors.

LS&T (1990) postulate two types of traders: rational traders (institutions) and noisy traders (retail investors). Rational traders have unbiased expectations, but noisy traders are

sometimes overly optimistic and at other times overly pessimistic. Rational investors are risk-averse, have finite horizons, and don't pursue aggressive trading activities to undo the influences of noisy traders. They argue that the added risk introduced by noisy traders, combined with the fact that closed end funds tend to be held by retail investors, means that closed end funds sell at a discount. Sentiment risk then becomes a systematic influence that affects not just closed end funds, but any investment (e.g., small stocks) held by retail as opposed to institutional investors.

Elton, Gruber and Busse (1996b) offered an alternative explanation for the discount on closed end domestic stock funds based on the market characteristics of these funds. They show that the loadings (betas) on the Fama French systematic factors: the market, the small versus large stock index, and a value minus growth index, are higher for the return on closed-end funds than they are for the returns on the securities these funds hold. Why do these differences in sensitivities arise?

Elton, Gruber and Busse (1996b) found the average market value of stocks held by closed end stock funds was \$5,572 million, while the average market value of the funds holding these stocks was \$343 million. Similarly, the average market-to-book ratio of the stocks held by funds was 3.9, while for the fund itself it was 0.9. This explains why the loadings on two recognized risk factors (small-large and value-growth) were so much larger for the funds than on the portfolio of securities they held. The higher loadings and positive factor prices mean more risk for the closed end funds than the portfolio they hold. The higher risk must be compensated for by higher expected return. The only way this can happen is for the average price for closed end funds to be lower than the NAV on these funds.

Either of the explanations (irrational sentiment as a systematic influence or the Fama-French model combined with the different risks associated with the fund and the portfolio they

hold) can be used to explain the persistent discount for closed end stock funds and, to a large extent, the movement of the discount over time.

3.2 Why Closed End Funds Exist

There is a second topic of great interest with respect to closed end funds: Why do they exist at all? The classic reason given for the existence of closed end funds is that their organizational form allows them to hold less liquid assets and to hold less cash. This reason has been explored both theoretically and empirically in a series of papers, perhaps most cogently in Cherkes, Sagi and Stanton (2009) and Deli and Varma (2002). Because closed end funds are not subject to inflows when investors buy a fund or of key importance outflows of cash when investors choose to sell a fund, they argue that closed end funds can hold more illiquid assets and less cash than open-end funds. This is, no doubt, an explanation for the creation of many types of closed end funds. Cherkes et al do a thorough job of exploring a liquidity-based theory of closed end funds. Deli and Varma test and find evidence that closed-end funds are more likely to hold securities in illiquid markets.

While the advantage of organization structure which allows for holding illiquid assets can account for some of the popularity of closed end funds, there is another advantage of organizational structure that has not received as much attention. Closed-end funds, unlike open-end funds, have the ability to use large amounts of leverage to finance their investments.

Elton, Gruber and Blake (2011c) design a study to more clearly show the impact of leverage. They study closed end bond funds because there are many more closed end bond funds than closed end stock funds. Furthermore, there are a number of closed end bond funds, each of which can be matched with an open end bond fund with the same portfolio manager, same objectives, and which are sponsored by the same fund family. By studying matched pairs of

funds, the effects of many of the influences affecting performance can be held constant. EG&B show that the characteristics of the assets and the returns on the assets earned by the open and closed end funds in the matched sample are almost identical. The difference between the open and closed end funds is the increased return to investors due to the use of leverage: leverage ratios for the closed-end funds averaged more than 50%. Leverage is advantageous to closed end funds because they borrow short-term, usually in the form of floating rate preferred stock and invest in longer-term bond funds. The advantage of fund leverage rather than investor leverage arises from at least three factors: interest paid on the preferred stocks issued by municipal closed end bond funds is not taxable to the holder of the preferred stock, limited liability to the holder of fund shares, and lower borrowing costs to the fund compared to investor borrowing costs. For example, the borrowing rate paid on preferred stock by municipal closed end bond funds is considerably lower than the federal fund rate.³¹

The research proceeds to show that the leveraged closed end bond funds are a more desirable asset to add to a portfolio of stocks or bonds than unlevered closed end funds or open-end funds. Furthermore, in a larger sample of closed end bond funds, differences in leverage account for more than 24% of the cross-sectional differences in discount, and discounts vary over time as a function of the difference between long rates and short rates, a measure of the desirability of leverage.

4. Exchange-Traded Funds (ETFs)

Exchange-traded funds are a fast-growing segment of the mutual fund industry. In 2010 there were over 1,800 exchange-traded funds with an aggregate investment of over \$900 billion.

³¹ The fact that the vast majority of closed end bond funds tend to employ leverage ratios close to the maximum allowed by law is evidence that managers of these funds believe that leverage is important.

Exchange-traded funds have been organized under three different sets of rules. The differences in organizational structure are important because they can affect what actions the ETF can take in managing the portfolio. The original ETF (spider) was organized as a trust. The trust structure requires exact replication of the index (rather than sampling). Furthermore, it doesn't allow security lending or the use of futures, and requires that dividends received from the securities the fund holds be placed in a non-interest-bearing account until they can be disbursed to shareholders. Most ETFs organized after the spiders were organized as managed funds. Managed funds have much greater flexibility, allowing sampling, the purchase and sale of futures, security lending, and the immediate reinvestment of dividends. The third possible organizational structure is a grantor trust. Investors in grantor trusts hold the shares directly, retaining their voting rights and receiving dividends and spinoffs directly. They can unbundle the trust, selling off some of the companies in the trust. There is no separate management fee, but rather there is a custodian fee for holding the shares. ETFs called "Holders" are grantor trusts.

ETFs are stocks and trade on exchanges like other stocks. ETF's assets are a basket of securities rather than physical assets, and as such they are similar to closed-end funds. They differ from closed-end funds in that new shares can be created or old shares can be deleted every day. For example, the largest ETF is the spider. The spider attempts to mimic the S&P 500 index with one share equal to approximately one-tenth of the price of the S&P 500 index. New or old shares are deleted or created in minimum orders of 50,000 shares for a payment of \$3,000, regardless of the number of units involved. At the end of the day the fund posts its holdings (including cash). An investor wishing to create shares turns in a bundle of stock holdings that match the S&P 500 index plus the appropriate amount of cash. There is more creation than deletion, and both are in large amounts. Creations and deletions occurred on approximately 15%

of the trading days. On these days Elton, Gruber, Comer and Li (2002) report that average creation and deletions average over \$100 million.

The system of creation and deletion and the ability to arbitrage price and NAV differences means that the price of a share in an exchange-traded fund has historically been close to NAV, unlike the price of closed-end funds. Most exchange-traded funds attempt to match an index and are passive in their investment strategy. The principal issues are:

1. Tracking error
2. The relationship of price to NAV
3. Their performance relative to other indexing vehicles
4. Their use in price formation
5. The effect of leverage
6. Active ETFs

Each of these will be discussed in turn.

4.1 Tracking error

Tracking error is the performance of the portfolio compared to the performance of the index.

It is usually measured by examining the residual of the following regression:³²

$$R_{pt} = \alpha_p + \beta_p I_t + e_{pt}$$

Where

1. R_{pt} is the return of the exchange-traded fund
2. I_t is the return on the index being matched
3. β_p is the sensitivity of the fund to the index

³² Once again an argument can be made to set $\beta_p = 1$ or to run the regression in excess return form.

4. e_{pt} is the residual
5. a_p is the average return on the fund not related to the index

Normally, β_p will be less than one with the difference representing the cash portion of the portfolio. There are three issues of interest. Is there a permanent difference from the underlying index? What's the average size of the error? Does the cumulative error converge rapidly to zero (errors are uncorrelated)? For large, well-diversified portfolios like those matching the S&P 500 index, tracking error is minimal and not very important. Using a sample of S&P index funds, Elton, Gruber and Busse (2004) report average R^2 in excess of 0.9999 for S&P 500 Index funds, which means the tracking error for S&P 500 Index funds is less than 0.0001%. This should also be true for exchange-traded funds tracking the S&P 500 Index.

However, many ETFs attempt to match indexes with sampling techniques rather than replication that exactly matches the index. Likewise, many ETFs attempt to match a country or sector index where a single security represents a large portion of the market and exact replication isn't possible because of rules prohibiting more than 5% of the portfolio being invested in a single security. These ETFs can have a serious problem in index replication. There has been very little research examining tracking error for these types of ETFs.

4.2 The Relationships of Price to NAV

The process of creation and deletion keeps price and NAV fairly close, particularly at the end of the day. However, there are deviations, and this is a potential cost to an investor who wishes to buy or sell and finds the price differs from NAV in an adverse way. It can, of course, also be a benefit if the investor buys when the price is below NAV and sells when it's above. For actively traded ETFs, prices and NAVs are very close and differences are transient. Engle and Sarkar (2002) examine differences for actively traded ETFs. The standard deviation of the premium and

discounts was around 15 basis points and was less than the bid-ask spread. For less actively traded funds (they used international funds to represent less actively traded funds) the standard deviation is much larger and deviations can persist over several days.³³ In examining these issues there are several difficulties. First, the NAV and price quotes are often not synchronized, and second, the data contains errors because prices are not accurately recorded or NAV does not accurately represent the portfolio value so that premiums and discounts are likely smaller than reported.

4.3 Performance Relative to Other Instruments

Passive ETFs often match the same index as an index fund. Also, there are sometimes futures on the index that can be used in conjunction with a bond to create a portfolio that matches the index.

How does the performance of these instruments compare? We will compare ETFs and index funds.³⁴ The difference in performance depends on the skill in matching the index, expenses, charter restrictions and tax considerations. Even for passive funds that construct their portfolio by exact replication, there can be differences in skill or differences in the actions allowed by the funds' charters that can affect relative performance. Probably the most important factor is how the fund handles changes in the index being matched. There are often large price changes around the time a security enters or leaves an index. The timing of the portfolio changes for the ETF which may represent management skill, or restrictions on the ETFs imposed by its charter, can affect return. Additional skill factors that affect relative performance include ability to lend securities, dealing with tender offers and mergers, policies involving cash, trading strategies, ability to reinvest dividends, transaction costs, and the ability to use (and skill in

³³ Chery (2004) also studies this phenomenon.

³⁴ See Elton, Gruber, Comer and Li (2002) for a comparison with futures.

using) futures. Depending on how the fund was organized, the ETF may or may not have flexibility on these issues. As pointed out by Elton, Gruber, Comer and Li (2002), ETFs organized as trusts such as spiders must hold the dividends received on underlying securities in a non-interest-bearing account where an index fund will reinvest the dividends or earn interest on them. If the market increases, this is a disadvantage to the ETF.³⁵ If the market decreases and the index funds reinvest dividends, then this is an advantage. Partly because of the disadvantage of holding dividends in a non-interest-bearing account and restrictions on lending, the use of futures and restrictions on rebalancing most ETFs issued after spiders choose a different organizational form. In addition to management skill affecting performance, expenses are a cost to investors and *ceteris paribus* hurt performance.³⁶

The final difference affecting performance is tax considerations. ETFs are considered tax efficient since they generally distribute fewer capital gains than index funds. Capital gains are generated when shares are sold and the price at which they were bought at is less than the selling price. The way ETF shares are created and deleted provides ETFs with a chance to maintain a high cost basis on shares in their portfolios. When ETF shares are redeemed, the trustee delivers in-kind securities that comprise the index. The trustee always delivers the lowest-cost shares, keeping the cost basis high. The IRS has ruled that the process of deletion is not a taxable exchange. Thus, if an investor turns in ETFs worth \$100 million and the trustee gives the investor securities with a cost basis of \$50 million, there are no capital gains taxes on the arbitrageur or the ETF. Poterba and Slovern (2002) studied the capital gain payment on the Vanguard S&P index fund and the ETF spider and found tax considerations gave the spider a tax

³⁵ Elton, Gruber, Comer and Li (2002) show how the shortfall of the spider compared to an index fund is a function of market movement and the size of the dividends.

³⁶ In comparing index funds, Elton, Gruber and Busse (1996) find that future performance is highly predicted by expense ratios. When they regressed the difference in returns between the fund and the index on the prior year's expense rates, they have a slope of -1.09% with an R^2 of 0.788.

advantage, but this was not nearly enough to overcome the other considerations that favored the index fund.

4.4 Their use of Price Formation

Hasbrouck (2003) and Schlusche (2009) examine the process of information incorporation when multiple contacts exist on the same index. For example, the S&P 500 index has the spider, an ETF, a floor-traded futures contract, and a small denomination electronically traded futures contract. Hasbrouck finds in this market information is first incorporated in the small denomination futures contract. In other markets the results can differ. For example, in the market for the S&P 400 mid-cap, which has an ETF and a futures contract, Hasbrouck (2003) finds information is reflected equally.

4.5 The Effect of Leverage

Several hundred ETFs have been developed that are levered, promising multiples of the daily returns on the index either positive or negative.³⁷ If a standard ETF return pattern can be expressed as a $1x$ where x is the index's return, then these products are expressed as $2x$, $3x$, $-2x$, and $-3x$. Unlike normal ETFs that hold the underlying securities, these products are constructed using derivatives. This means that the tax efficiency discussed earlier doesn't hold since realized gains from derivative contracts are taxed at ordinary income tax rates and creation and deletion is usually in cash, not in kind. Also, these products have much higher expense ratios than standard ETFs. These products are designed for short-term traders. Investors holding them over a long period need not get the promised multiple return ($2x$ or $3x$) over the longer period. This occurs as shown below because the products are re-levered every day to the stated objective.

The effect of daily re-levering on multi-day returns is easy to see with a two-period example. Assume an investor has one dollar, borrows $(m-1)$ dollars and invests m dollars in a $1x$

³⁷ See Cheng and Madhavan (2009) for an excellent discussion.

ETF holding the borrowing at $m-1$ for both periods. The ending value (ignoring interest on the borrowing and recognizing that $(m-1)$ is paid back) is:

$$m(1+r_1)(1+r_2)-(m-1) \quad (1)$$

If the investor invests one dollar in an m x levered ETF, the return is:

$$(1+r_1m)(1+r_2m) \quad (2)$$

For one period the payoff is the same, but because rebalancing occurs, the two-period payoff is different. The difference (return on the levered ETF minus return on “homemade” leverage) is $(m^2 - m)r_1r_2 > 0$. If $r_1r_2 > 0$ then the daily rebalancing gives a higher return. If $r_1r_2 < 0$ then daily rebalancing gives a lower return. Cheng and Madhavan (2009) show that with high volatility and little trend, an investor invested in an m x ETF will get less than m x in return. Given the high fees and that income is mostly ordinary income rather than capital gains even with an upward trend, an investor is likely to get less than expected over longer time frames. However, an investor may still chose this form of index fund, for it allows higher level of debt than the investor can get on personal accounts.

4.6 Active ETFs

Active ETFs have only recently been introduced, and so have not yet been subject to serious academic study. ETFs require daily posting of the portfolio to facilitate creation and deletion. Many trades for mutual funds are executed over several days to mitigate price impacts. Daily reporting of positions can cause front running. This has slowed their introduction.

Conclusion

In this Chapter we have attempted to review both relevant topics related to mutual funds and the literature that will allow the reader to delve deeper into any of the subjects which we have covered. The subject of mutual funds is so broad that we have had to use personal interests

in deciding what to cover. We apologize for our sins of omission both with respect to the subjects covered and the papers cited. The vast literature on mutual funds is a testimony to both the importance of this form of financial intermediary and the interest in it. No essay could possibly cover in entirety the immense scope of the research that has been, and is being, done on mutual funds.

Table 1

Mutual Fund Performance Results (Annualized)

A. Articles Using Mutual Fund Returns (Post Expenses)

		Average Performance
1.	Jensen (1968)	-1.1
2.	Lehman & Modest (1987)	Negative
3.	Elton, Gruber, Das & Hlavka (1993)	-1.59
4.	Gruber (1996)	-.65
5.	Elton, Gruber, Blake (1996c)	-.91
6.	Ferson and Schadt (1996)	+.24
7.	Carhart (1997)	-1.98
8.	Pastor, Stambaugh (2002)	-.86 to -1.25
9.	Elton, Gruber & Blake (2003)	-.91
10.	Fama & French (2010)	-.83
11.	Elton, Gruber & Blake (2011a)	Negative

B. Using Holdings Data (Pre-Expenses)

1.	Grinblatt & Titman (1989a)	(slight positive)
2.	Grinblatt & Titman (1993)	2.00%
3.	Daniel, Grinblatt, Titman & Wermers (1997)	.77
4.	Wermers (2002)	.71

C. Timing

1.	Daniel, Grinblatt, Titman & Wermers (1997)	Timing ability
2.	Busse (1999)	Timing ability
3.	Becker, Ferson, Myers & Schill (1999)	No timing ability
4.	Bollen & Busse (2001)	Timing ability
5.	Kaplin & Sensoy (2005)	Timing ability
6.	Jiang, Yao & Yu (2007)	Timing ability
7.	Elton, Gruber & Blake (2011b)	No timing ability
8.	Ferson & Qian (2006)	No timing ability

D. Bond Funds

1.	Blake, Elton & Gruber (1994)	-.51%
2.	Elton, Gruber & Blake (1995)	-.75% to -1.3%
3.	Comer & Rodriguez (2006)	-1.00 to -1.14%
4.	Chen, Ferson & Peters (2010)	-.70%

Panel E
Persistence

		Measure Used		Result	Positive alpha for top group
		Ranking Measure	Evaluation Measure		
1.	Grinblatt & Titman (1992)	G&T Measure ³⁸	G&T Measure	Persistence	NR
2.	Hendricks, Patel & Zeckhauser (1993)	Returns	Returns	Persistence	NR
3.	Brown & Goetzmann (1995)	Returns	Returns Capm 3-factor	Persistence Primarily worst group	NR
4.	Carhart (1997)	Returns	4-factor Alpha	Lowest Decile	No
5.	Carhart (1997)	Alpha	4-factor Alpha	Lowest & highest decile	Yes
6.	Elton, Gruber & Blake (1996c)	Alpha	Alpha	Persistence	Yes
7.	Gruber (1996)	Alpha	Alpha	Persistence	Yes
8.	Cohen, Coval & Pastor (2005)	Alpha	Alpha	Persistence	Yes
9.	Busse & Irvine (2006)	Baysian Alpha	Alpha	Prediction	Yes
10.	Elton, Gruber & Blake (2011a)	Alpha	Alpha	Persistence	Yes
11.	Elton, Gruber & Blake (2011d)	Alpha	Alpha	Persistence	Yes

NR means not relevant since the authors don't measure performance relative to index or set of indexes.

Bibliography

References:

Admati, Anatoly; Pfleiderer, Paul; Pers, Stephen, and Bhattacharya, Suddipto (1986). On timing and selectivity. *Journal of Finance*, 41:715-730.

Agnew, Julie and Balduzzi, Pierluigi (2010). The response of aggregate 401(k) trading to asset returns. Unpublished manuscript, Boston College.

Ameriks, John and Zeldes, Stephen (2004). How do household portfolio shares vary with age? Working paper, Columbia University.

Baks, Klaas P.; Metrick, Andrew; and Wachter, Jessica (2001). Should investors avoid all actively managed mutual funds: a study in Bayesian performance evaluation. *Journal of Finance* 56:45-85.

Balduzzi, Pierluigi; Agnew, Julie; and Sunden, Annika (2003). Portfolio choice and trading in a large 401(k) plan. *American Economic Review* 93:193-205.

Baker, Malcolm; Litov, Lubomir; Wachter, Jessica; and Wurgler, Jeffrey (2004). Can mutual fund managers pick stocks? Evidence from their trades prior to earnings announcements. *Journal of Financial and Quantitative Analysis* 45:1111-1131.

Barber, Brad; Odean, Terrence; and Zheng, Lu (2005). Out of sight out of mind: effects of expense on mutual fund flows. *Journal of Business* 78:2095-2119.

Becker, Connie; Ferson, Wayne; Myers, David; and Schill, Michael (1999). Conditional market timing with benchmark investors. *Journal of Financial Economics* 52:119-148.

Benartzi, Shlomo and Thaler, Richard (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review* 91:79-98.

Bergstresser, Daniel and Poterba, James (2002). Do after-tax returns affect mutual fund inflows? *Journal of Financial Economics* 63:381-414.

Berk, Jonathan and Green, Richard (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy* 112:1269-1295.

Berk, Jonathan and Stanton, Richard (2007). Managerial ability, compensation and the closed end fund discount. *Journal of Finance* 62:529-556.

Blake, Christopher R.; Elton, Edwin J.; and Gruber, Martin J. (1994). The performance of bond mutual funds. *The Journal of Business*, 66:371-403.

- Bollen, Nicolas P.B. and Busse, Jeffrey A. (2001). On the timing ability of mutual fund managers. *Journal of Finance* 56:1075-1094.
- Brennan, Michael and Jain, Ravi (2007). Capital gains taxes, agency costs and closed end fund discounts. Unpublished manuscript, UCLA.
- Brickley, James; Manaster, Steven; and Schallheim, James (1991). The tax timing option and the discounts on closed end investments companies. *The Journal of Business* 64:287-312.
- Brown, Keith; Harlow, Van; and Starks, Laura (1996). Incentives in the mutual fund industry. *Journal of Finance* 5:85-110.
- Brown, Stephen J.; Goetzmann, William N.; Ibbotson, Roger G.; and Ross, Stephen A. (1992). Survivorship bias in performance studies. *Review of Financial Studies* 5:553-580.
- Brown, Stephen J. and Goetzmann, William N. (1995). Performance persistence. *Journal of Finance* 50:679-698.
- Brown, Stephen J. and Goetzmann, William N. (1997). Mutual fund styles. *Journal of Financial Economics* 43:373-399.
- Busse, Jeffrey (1999). Volatility timing in mutual funds: evidence from daily returns. *Review of Financial Studies* 12: 1009-1041.
- Busse, Jeffrey A. and Irvine, Paul J. (2006). Bayesian alphas and mutual fund persistence. *Journal of Finance* 61:2251-2288.
- Carhart, Mark M. (1997). On persistence in mutual fund performance. *Journal of Finance* 52:57-82.
- Carpenter, Jennifer and Lynch, Anthony (1999). Survivorship bias and attrition effects in measures of performance persistence. *Journal of Financial Economics* 54:337-374.
- Chan, Louis; Chen, Hsui; and Lakonishok, Josef (2002). On mutual fund investment styles. *Review of Financial Studies* 15:1407-1437.
- Chan, Louis; Dimmock, Stephen; and Lakinishok, Josef (2010). Benchmarking money manager performance: issues and evidence. *Review of Financial Studies* 22:4553-4599.
- Chay, J.B. and Tryzinka, Charles (1999). Managerial performance and the cross sectional pricing of closed-end bond funds. *Journal of Financial Economics* 52:379-308.
- Chen, Joseph; Hong, Harrison; Huang, Ming; and Kubik, Jeffrey (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review* 94: 1276-1307.

Chen, Wai-Fu; Roll, Richard; and Ross, Steve (1986). Economic forces and the stock market. *Journal of Business* 56:383-403.

Chen, Yong; Ferson, Wayne; and Peters, Helen (2010). Measuring the timing ability and performance of bond mutual funds. *Journal of Financial Economics* 98:72-89.

Chen, Zhiwu and Knez, Peter (1996). Portfolio performance measurement: theory and application. *Review of Financial Studies* 9:511-555.

Cheng, Minder and Madhavan, Ananth (2009). The dynamics of levered and inverse exchange traded funds. *Journal of Investment Management*

Cherkes, Martin (2003). A practical theory of closed end funds as an investment vehicle. Working paper, Princeton University.

Cherkes, Martin; Sagi, Jacob; and Stanton, Richard (2009). A liquidity based theory of closed-end funds. *Review of Financial Studies* 22:257-297.

Cherry, Josh (2004). The limits of arbitrage: evidence from exchange traded funds. Unpublished manuscript, University of Michigan.

Chevalier, Judith and Ellison, Glenn (1997). Risk-taking by mutual funds as a response to incentives. *Journal of Political Economy* 105:1167-1200.

Chevalier, Judith and Ellison, Glenn (1999). Are some mutual fund managers better than others: Cross-sectional patterns in behavior and performance. *Journal of Finance* 54:875-900.

Chopra, Navin; Lee, Charles; Shleifer, Andrei; and Thaler, Richard (1993). Yes, discounts on closed-end funds are a sentiment index. *Journal of Finance* 48:801-8.

Christoffersen, Susan and Musto, David (2002). Demand curves and the pricing of money management. *Review of Financial Studies* 15:1499-1524.

Christoffersen, Susan; Keim, Donald; and Musto, David (2007). Valuable information and costly liquidity: evidence from individual mutual fund trades. Unpublished manuscript, University of Pennsylvania.

Christopherson, Jon; Ferson, Wayne; and Glassman, Debra (1998a). Conditional measures of performance and persistence for pension funds. *Research in Finance* 16:1-46.

Christopherson, Jon; Ferson, Wayne; and Glassman, Debra (1998b). Conditioning manager alpha on economic information: another look at the persistence of performance. *Review of Financial Studies* 11:111-142.

Chuprinin, Oleg and Massa, Massimo (2010). Investor horizon and portfolio management: evidence from variable annuity funds. Insead working paper.

- Cohen, R., Coval, J. and Pastor, L. (2005), Judging fund managers by the company they keep. *Journal of Finance* 60:1057-1096.
- Comer, George (2005). Hybrid mutual funds and market timing performance. *Journal of Business* 79:771-797.
- Comer, George and Rodriguez, Javier (2006). Corporate and government bond funds: an analysis of investment style, performance and cash flow. Working paper: Georgetown University.
- Connor, Gregory and Korajczyk, Robert A. (1986). Performance measurement with the arbitrage pricing theory: a new framework for analysis. *Journal of Financial Economics* 15:373-394.
- Connor, Gregory and Korajczyk, Robert A. (1988). Risk and return in an equilibrium APT: applications of a new test methodology. *Journal of Financial Economics* 21: 255-289.
- Connor, Gregory and Korajczyk, Robert (1991). The attributes behavior and performance of U.S. Mutual Funds. *Review of Quantitative Finance and Accounting* 1:4-25.
- Cooper, Michael; Gulen, Huseyin; and Rau, Raghavendra (2005). Changing names with style: mutual fund name changes and their effects on fund flows. *Journal of Finance* 60:2825-2858.
- Cornell, Brad (1979). Asymmetric information and portfolio performance measurement. *Journal of Financial Economics* 7:381-390.
- Cremers, Martijn and Petajisto, Antti (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies* 22:3329-3365.
- Cremers, Martijn, Petajisto, Antti and Zitzewitz, Eric (2010). Should benchmark indices have alpha? Revisited performance evaluation. Working Paper, Yale University.
- Dahlquist, Magnus and Soderlind, Paul (1999). Evaluating portfolio performance with stochastic discount factors. *Journal of Business* 72:347-84.
- Daniel, Kent; Grinblatt, Mark; Titman, Sheridan; and Wermers, Russ (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance* 52:1035-1058.
- Del Guercio, Diane and Tkac, Paula A. (2002). The determinants of the flow of funds of managed portfolios: mutual funds v. pension funds. *Journal of Financial and Quantitative Analysis* 37:523-557.
- Del Guercio, Diane and Tkac, Paula A. (2008). Star power: the effect of Morningstar ratings on mutual fund flow. *Journal of Financial and Quantitative Analysis* 43:907-936.
- Deli, Daniel and Varma, Raj (2002). Closed-end versus open: the choice of organizational form. *Journal of Corporate Finance* 8:1-27.

DeLong, J. Bradford and Shleifer, Andrei (1992). Closed-end fund discounts: a yardstick of small investor sentiment. *Journal of Portfolio Management* 18:46-53.

Duffee, Gregory R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance* 57:405-43.

Edelen, Roger; Evans, Richard; and Kadlec, Gregory (2009). Scale effects in mutual fund performance: the role of trading costs. Unpublished manuscript, University of Virginia.

Edwards, Amy; Harris, Lawrence; and Piwowar, Michael S. (2006). Corporate bond market transaction costs and transparency. *Journal of Finance* 61, 1361-1397.

Elton, Edwin J.; Gruber, Martin J.; Das, Sanjiv; and Hlavka, Matthew (1993). Efficiency with costly information: a reinterpretation of evidence from manager portfolios. *Review of Financial Studies* 6:1-23.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (1995). Fundamental variables, APT, and bond fund performance. *Journal of Finance* 50:1229-1256.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher (1996a). Survivorship bias and mutual fund performance. *Review of Financial Studies* 9:1097-1120.

Elton, Edwin J.; Gruber, Martin J.; and Busse, Jeffrey A. (1996b). Do investors care about sentiment? *Journal of Business* 71:475-500.

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

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (1999). Common factors in active and passive portfolios. *European Finance Review* 3:53-78.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2001). A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *Journal of Finance* 56:2415-2450.

Elton, Edwin J.; Gruber, Martin J.; Comer, George; and Li, Kai (2002). Spiders: where are the bugs? *Journal of Business* 75:453-472.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2003). Incentive fees and mutual funds. *Journal of Finance* 58:779-804.

Elton, Edwin J.; Gruber, Martin J.; and Busse, Jeff. (2004). Are investors rational? Choices among index funds. *Journal of Finance* 58: 427-465.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2006). The adequacy of investment choices offered by 401(k) plans. *Journal of Public Economics* 90:1299-1314.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2007). Participant reaction and the performance of funds offered by 401(k) plans. *Journal of Financial Intermediation* 42:240-271.

Elton, Edwin J.; Gruber, Martin J.; Brown, Stephen J.; and Goetzmann, William (2010a). *Modern portfolio theory and investment analysis*, 8th edition. John Wiley, Hoboken NJ.

Elton, Edwin J.; Gruber, Martin J.; Blake, Christopher R.; Krasny, Yoel; and Ozelge, Sadi (2010b). The effect of the frequency of holdings data on conclusions about mutual fund management behavior. *Journal of Banking and Finance* 34:912-922.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2011a). Holdings data, security returns, and the selection of superior mutual funds. *Journal of Financial and Quantitative Analysis*: April 2011.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2011b). An examination of mutual fund timing using monthly holdings data. *Journal of Banking and Finance* 4:912-922.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2011c). Why do closed end funds exist? An additional explanation for the growth in domestic closed end bond funds. Unpublished manuscript, New York University.

Elton, Edwin J.; Gruber, Martin J.; and Blake, Christopher R. (2011d). Does size matter? The relationship between size and performance. Unpublished manuscript, New York University.

Engle, Robert and Sarkar, Debojyoti (2002). Pricing exchange, traded funds. Unpublished manuscript, New York University.

Evans, Richard (2010). Mutual fund incubation. *Journal of Finance* 65 1581-1611

Fama, Eugene F. and French, Ken R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance* 51:55-87.

Fama, Eugene F. and French, Ken R. (2010). Luck versus skill in the cross-section of mutual fund return. *Journal of Finance* 65: 1915-1947.

Farnsworth, Heber; Ferson, Wayne; Jackson, David; and Todd, Steven (2000). Performance evaluation with stochastic discount factors. *Journal of Business* 75:473-504.

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

Ferson, Wayne E. and Khang, Kenneth (2002). Conditional performance measurement using portfolio weights; evidence for pension funds. *Journal of Financial Economics* 65:249-282.

Ferson, Wayne E. and Merijun, Qian (2006). When can market timers time? Unpublished manuscript, University of Southern California.

- Friend, Irwin; Blume, Marshall, and Crocket, Gene (1970). Measurement of portfolio performance under uncertainty. *American Economic Review* 60:561-575.
- Glode, Vincent (2011). Why mutual funds underperform. *Journal of Financial Economics* 99: 546-559.
- Goetzmann, William N. and Ibbotson, Roger G. (1994). Do winners repeat: predicting mutual fund performance. *The Journal of Portfolio Management* 20:9-18.
- Grinblatt, Mark and Titman, Sheridan (1989a). Mutual fund performance: an analysis of quarterly portfolio holdings. *Journal of Business* 62:393-416.
- Grinblatt, Mark and Titman, Sheridan (1989b). Portfolio performance evaluation: old issues and new insights. *Review of Financial Studies* 2:393-422.
- Grinblatt, Mark and Titman, Sheridan (1992). The persistence of mutual fund performance. *Journal of Finance* 47:1977-1984.
- Grinblatt, Mark and Titman, Sheridan (1993). Performance measurement without benchmarks: an examination of mutual fund returns. *Journal of Business* 66:47-68.
- Gruber, Martin J. (1996). Another puzzle: the growth in actively managed mutual funds. *Journal of Finance* 51:783-810.
- Han, Son and Li, Dan (2009). Liquidity crisis runs and security design. Working paper, Federal Reserve Board.
- Harris, Lawrence and Piwowar, Michael (2006). Municipal bond liquidity. *Journal of Finance* 61:1361-1397.
- Hasbrouck, Joel (2003). Intraday price formation in the market for U.S. equity markets. *Journal of Finance* 58:2375-2400.
- Hendricks, Darryll; Patel, Jayendu; and Zeckhauser, Richard (1993). Hot hands in mutual funds: short-run persistence of relative performance, 1974-1988. *Journal of Finance* 48:93-130.
- Henriksson, Roy D. and Merton, Robert C. (1981). On market timing and investment performance II: statistical procedures for evaluating forecasting skills. *Journal of Business* 54:513-534.
- Henriksson, Roy D. (1984). Market timing and mutual funds performance: an empirical investigation. *Journal of Business* 57:73-96.
- Huberman, Gur and Jiang, Wei (2006). Offering versus choices in 401(k) plans: equity exposure and number of funds. *The Journal of Finance* 61:763-801.

Huberman, Gur; Iyengar, Sheena S.; and Jiang, Wei (2007). Defined contribution pension plans: determinants of participation and contribution rates. *Journal of Financial Services Research* 31:1-32.

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

Jagannathan, Ravi and Korajczyk, Robert A. (1986). Assessing the market timing performance of managed portfolios. *Journal of Business* 59:217-236.

Jegadeesh, Narasimham and Titman, Sheridan (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance* 48:93-130.

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

Jiang, George; Yao, Tong; and Yu, Tong (2007). Do mutual funds time the market? Evidence from holdings data. *Journal of Financial Economics* 88:119-145.

Kacperczyk, Marcin T.; Sialm, Clemens; and Zheng, Lu (2008). Unobserved actions of mutual funds. *Review of Financial Studies* 21:2379-2416.

Kaplan, Steven N. and Sensoy, Berk A. (2005). Do mutual funds time their benchmarks? Working paper, University of Chicago.

Keim, Donald B. and Madhavan, Ananth (1995). Empirical evidence on the behavior of institutional traders. *Journal of Financial Economics* 37:371-399.

Keim, Donald B. and Madhavan, Ananth (1997). Transaction costs and investment style: an interexchange analysis of institutional equity trades. *Journal of Financial Economics* 46:265-292.

Kothari, S., and Warner, Jerold (2001). Evaluating mutual fund performance. *Journal of Finance* 56:1985-2010.

Kumar, Raman. and Noronha, G. M. (1992). A re-examination of the relationship between closed-end fund discounts and expenses. *Journal of Financial Research* 15:139-147.

Lee, Charles; Shleifer, Andrei; and Thaler, Richard (1990). Closed end mutual funds. *Journal of Economic Perspectives* 4:153-166.

Lee, Charles; Shleifer, Andrei; and Thaler, Richard (1991). Investor sentiment and the closed end fund puzzle. *Journal of Finance* 46:76-110.

Lehmann, Bruce N. and Modest, David M. (1987). Mutual fund performance evaluation: a comparison of benchmarks and benchmark comparisons. *Journal of Finance* 42:233-266.

- Liang, Nellie and Weisbenner, Scott (2002). Investor behavior and the purchase of company stock in 401(k) plan design. Unpublished manuscript, University of Illinois.
- Lynch, Anthony and Musto, David (2003). How investors interpret past fund returns. *Journal of Finance* 58:2033-2058.
- Madrian, Brigitte and Shea, Dennis (2001). The power of suggestion: inertia in 401(k) participants' savings. *Quarterly Journal of Economics* 116;1149-1187.
- Malkiel, Burton (1977). The valuation of closed end investment company shares. *Journal of Finance* 32, 847-858.
- Mamaysky, Harry; Spiegel, Matthew; and Zhang, Hong (2007). Improved forecasting of mutual fund alphas and betas. *Review of Finance* 11:359-400.
- Manzler, David (1990). Liquidity, liquidity risk and the closed-end fund discount. Unpublished manuscript, University of Cincinnati.
- Massa, Massimo; Goetzman, William; and Rouwenhorst, Geert (2001). Behavioral factors in mutual fund flows. Unpublished manuscript, Yale University.
- McConnell, John and Saretto, Alessio (2010). Auction failure and the market for auction rate securities. *Journal of Financial Economics* 97:451-469.
- Merton, Robert C. (1973). An intertemporal capital asset pricing model. *Econometrica* 41:867-887.
- Moskowitz, Tobias (2000). Discussion. *Journal of Finance* 55: 1695-1703.
- Pastor, Lubos and Stambaugh, Robert (2000). Investing in equity mutual funds. *Journal of Financial Economics* 63:351-380.
- Pastor, Lubos and Stambaugh, Robert (2002). Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics* 63:315-349.
- Pollet, Joshua and Wilson, Mungo (2008). How does size affect mutual fund behavior? *Journal of Finance* 63:2941-2969.
- Pontiff, Jeffrey (1996). Costly arbitrage, evidence from closed-end funds. *Quarterly Journal of Economics* 111:1135-1151.
- Pontiff, Jeffrey (1997). Excess volatility and closed-end funds. *American Economic Review* 86:155-169.
- Poterba, James and Slovern, J. (2002). A new investment option for taxable investors. *American Economic Review* 92:422-427.

- Ross, Steven.A. (1976). The arbitrage pricing theory of capital asset pricing. *Journal of Economic Theory* 13:341-360.
- Schlusche, Bernd (2009). Price formation in spot and futures markets: exchange traded funds vs. index funds. *Journal of Derivatives* 17:26-40.
- Sharpe, William (1966). Mutual fund performance. *Journal of Business* 39:119-138.
- Sharpe, William (1992). Asset allocation management style and performance measurement. *Journal of Portfolio Management* 18:7-19.
- Sharpe, William (1994). The Sharpe ratio. *The Journal of Portfolio Management* 21:49-58.
- Sirri, Eric and Tufano, Peter (1998). Costly search and mutual fund flows. *Journal of Finance* 53:1589-1622.
- Stambaugh, Robert (1997). Analyzing investments whose histories differ in length. *Journal of Financial Economics* 45:285-33.
- Treynor, Jack (1965). How to rate management of investment funds. *Harvard Business Review* 43:63-75.
- Treynor, Jack and Mazuy, M. (1966). Can mutual funds outguess the market? *Harvard Business Review* 44:131-136.
- Treynor, Jack and Black, Fisher (1973). How to use security analysis to improve portfolio selection. *Journal of Business* 45:68-86.
- Wermers, Russ (2002). A matter of style: the causes and consequences of style drift in institutional portfolios. Working paper.
- Yan, Xuemin (2008). Liquidity, investment style and the relation between fund size and fund performance. *Journal of Financial and Quantitative Analysis* 43:741-768.
- Zheng, Lu (1999). Is money Smart? -- A study of mutual fund investors' fund selection ability. *The Journal of Finance* 54: 901-933.