Are Mutual Fund Performance Measures Created Equal? An Analysis of Mutual Fund Performance and Ranking

Cheryl J. Frohlich
University of North Florida

Anita Pennathur*
Florida Atlantic University
and

Oliver Schnusenberg
University of North Florida

Preliminary copy, please do not cite without permission

*Corresponding author. Florida Atlantic University, Department of Finance, College of Business, 2912 College Avenue, Davie, FL 33341, email pennathu@fau.edu, phone: 954-236-1272.

Are Mutual Fund Performance Measures Created Equal? An Analysis of Mutual Fund Performance and Ranking

Abstract

We investigate the performance measurement and ranking of mutual funds using five different performance measures identified previously in the literature. Using a sample of 2,861 mutual funds for the period 1993-2004, we test whether each of the performance measures results in significant excess returns, whether fund rankings according to the different measures are significantly correlated, and whether the performance measures identify the same funds as the best- and worst-performing funds over the sample period. Our results show that all models reveal significantly below-market performance over the entire sample period. Moreover, the more extreme performance deciles contain funds with higher return standard deviations. We find that the Kendall correlation coefficients indicate that only a few of the performance measures result in similar rankings. To further investigate the rank order of the funds, we identify the specific funds that were the best- and worst-performing funds; these specific funds, particularly the better-performing funds, differ depending on the model used.

Are Mutual Fund Performance Measures Created Equal? An Analysis of Mutual Fund Performance and Ranking

1. Introduction

Despite four decades of mutual fund performance research and a gamut of performance measures, empirical results on performance can be described, at best, as mixed. Measures range from the simple risk to reward ratios of Sharpe and Treynor to single-index and multi-factor models. Earlier research focused on benchmark specification and issues of misspecification. Gruber (1996) and Elton, Gruber, and Blake (1996) find that evidence of superior performance using net returns. Other studies (Grinblatt and Titman, 1989, Daniel, Grinblatt, Titman, and Wermers, 1997, and Wermers, 2000) examine gross returns to find that mutual funds outperform the benchmarks on a gross basis. Using a unique, survivor-bias free dataset, Carhart (1997) shows that most of the overperformance can be explained away by some common factors including a momentum factor. More recently, however, Bollen and Busse (2005) use daily data to find evidence of short-term performance persistence.

Using a sample of 2,861 funds with at least three years of data available from the CRSP mutual fund database for the period 1993-2004, we revisit the topic of mutual fund performance. While research shows that fund performance for a given model depends on the chosen benchmark and may vary over time, performance may also vary and result in different rank orders of mutual funds across different performance measures. We examine performance via the CAPM model (Jensen's alpha), the Sharpe reward to variability index, the Treynor reward to volatility index, the Fama-French three-factor model, and the Carhart four-factor model to assess consistency across the difference measures. Most of the studies mentioned examine datasets that did not extend beyond the mid-1990s and many of the datasets suffer from survivorship bias. Our updated,

survivor-bias free dataset spans the bull and bear markets of the late 1990s and 2000s, allowing for a broad spectrum and time-span in which to measure performance.

Our study makes two important contributions. First, while mutual fund prospectuses indicate that past performance is not indicative of future performance, mutual fund managers may find it beneficial to only report those measures that indicate superior stock selection abilities. Moreover, performance measures are important determinants of a mutual fund manager's compensation, and yet there is little disclosure on the details of manager compensation. Disclosure of performance measurement and manager incentives is increasingly important in an industry already plagued with trading scandals and increased SEC regulation. Indeed, a Business Week (2003) article argues that while fund managers' compensation is a tightly held secret, investors need to know what manager incentives are, and what performance measures are used. Furthermore, research on performance and flow-of-funds shows that investors chase high-performing funds (Ippolito, 1992, Chevaliar and Ellison, 1997, Sirri and Tufano, 1998), and therefore, performance measurement is crucial in markets where mutual fund investors chase returns. Thus, we ask the question: do we find some consistency across the measures?

Second, while academia debates the empirical validity of the CAPM and the demise of beta (Fama and French, 1992), there is evidence that the mutual fund industry relies on the more basic performance measures such as fund alpha to gauge performance. Indeed, Evans (2004) addresses the role of alpha in key decisions by fund management companies and finds that internal management decisions such as promotions and demotions of fund managers are based on fund alpha. Examining the labor market for mutual fund managers, Chevalier and Ellison (1999) show that manager termination is sensitive to fund performance as measured by fund alpha. Similarly, Hu, Hall, and Harvey (2000) document that fund managers' promotions and demotions

are positively and negatively related to past fund performance respectively. It appears that while academic literature focuses on measuring performance in a multi-factor framework, the mutual fund industry still employs the more traditional measures such as the Sharpe and Treynor ratios and Jensen's alpha. Thus, our analysis using a variety of measures expands this discussion on performance measurement measures.

We begin our analysis by analyzing the consistency of fund performance across the five measures. We examine whether a fund reporting over- or underperformance for one measure, records similar performance for other measures. Our premise is that if some models result in significant positive excess returns while others do not, fund managers may be motivated to choose the more favorable models as positive measures that are attributable to the fund manager. We then explore if the funds are ranked differently according to each measure in the aggregate. We rank the funds according to the appropriate criterion for each performance measure and examine the correlation of the rankings across five performance measures. Finally, we determine if the five performance measures identify the same funds as the best- and worst-performing funds over the sample period by investigating the rank order of the funds and identifying the specific best and worst funds according to each performance measure and the average rank of those funds according to the other four performance measures. A finding that funds are ranked differently depending on the performance measure used implies that fund managers' compensation may have been affected over the sample period if their compensation was based on the rank order of their fund compared to other funds.

The remainder of this paper is organized as follows. Section 2 provides a summary of the relevant literature on mutual fund performance measurement and a discussion of the five

performance measures. We present the data and methods in Section 3 while the empirical results are discussed in Section 4. Section 5 concludes.

2. Literature Review and Performance Measures

2.1 Literature Review

In general, the models used to measure performance in mutual funds are variants of the CAPM, including the multi-factor models and models involving higher moments of the return generating process.[1] Research documents that fund performance for the same period can vary significantly depending on the model used to measure performance. Earlier research focused on the Sharpe, Treynor, and Jensen's alpha measures. Treynor's (1965) reward-to-volatility ratio (RVOL) distinguishes between total risk and systematic risk, implicitly assuming that portfolios are well diversified and thus ignores any diversifiable risk. Therefore, if beta is the appropriate type of risk, a stock's (portfolio's) risk-adjusted returns can be determined by the Treynor Index. Alternatively, Sharpe's (1966) reward-to-variability ratio is used if total variability is thought to be the appropriate measure of risk, a stock's (portfolio's) risk-adjusted returns can be computed using the Sharpe Index. The Sharpe and Treynor Index eliminate the problem of only considering return as a measure of performance. However, neither ratio is independent of the time period over which it is measured. This means that the ratio can change from one period to another with different results. Moreover, both ratios also ignore the correlation of a fund with other assets, liabilities, or previous realizations of its own return (Hodges, Taylor, and Yoder, 1997). Furthermore, an analysis of these measures (Friend and Blume, 1970; Porter and Gaumnitz, 1972) reveals a systematic biased relationship with risk measures. Friend and Blume (1970) conclude

that this bias is due to the invalidity of the assumption that equal lending and borrowing opportunities for all investors exist. In addition, neither ratio always takes into consideration the transaction costs nor the expenses associated with the purchase and sale of assets (Murthi, Choi, and Desai, 1997). Since transaction costs are important in measuring a fund's performance, the ratio calculated without transaction costs may be biased (Pettengill, Sundaram, Mathur, 1995).

The choice of an appropriate benchmark is essential in computing Jensen's alpha. Early research (e.g., Friend, Blume, and Crockett, 1970) also argues that an inaccurate benchmark can trick the Jensen's alpha calculation by over- or under-weighing small firm returns. Ippolito (1989) uses Jensen's alpha to examine the performance and determinants of performance of 143 funds for the period 1965-1984 and finds that positive fund performance is not related to either expense or turnover ratios. To this end, Elton, Gruber, Das, and Hlvaka (1993) demonstrate how index selection such as the S&P 500 is not suitable in instances where the small-firm effect is prevalent. Using the same mutual fund sample as Ippolito, the authors demonstrate how index selection can change the performance measurement from a positive (negative) alpha to a negative (positive) alpha. Similarly, Blake, Elton, and Gruber (1993) document that index selection is equally important for bond mutual funds and can impact the nature and the sign of fund performance. Thus, Gruber (1996) suggests the variable alpha based on multi-index models. Yet, as Lehman and Modest argue (1987), even multifactor models are subject to model misspecification and are very sensitive to benchmark specification. Later models such as Fama and French (1992) and Carhart (1997) add proxies for size and book-to-market and stock momentum, respectively.

---

[1] Examples of multiple moment models include the three moment model proposed by Prakash and Bear (1986) and the multiple moment model of Stephens and Proffitt (1991).

Studies examining gross returns such as Grinblatt and Titman (1993), Grinblatt, Titman and Wermers, (1995) and Daniel, Grinblatt, Titman, and Wermers (2000) find that funds overperform benchmarks on a gross basis. Using a theoretical framework, Berk and Green (2004) show that performance is related to fund size, and develop several empirical predictions for instances where mutual fund performance dissipates as the scale of operations increases. Other research (for e.g., Brown, Harlow, and Starks, 1996, Chevalier and Ellison, 1996, Sirri and Tufano, 1998, Del Guercio and Tkac, 2002) examines the relation between performance and flow of funds. Hendricks, Patel, and Zeckhouser (1993), Brown and Goetzmann (1995), and Wermers (1996) document the "hot hands" phenomenon and find evidence of performance persistence. Several recent papers, such as Pastor and Stambaugh (2002) and Jones and Shanken (2003) use Bayesian alphas to infer performance.

Along the lines of our study, Coles, Daniel, and Nardari (2004) compare the Jensen (1968) performance measure and two timing measures (Treynor and Mazuy, 1966, and Henricksson and Merton, 1981) to see if investors and researchers would make the wrong conclusions about fund manager performance due to benchmark or model misspecification. They find that benchmark misspecification is economically insignificant in making performance inferences, and that the power of deducting individual fund ability or distinguishing between a good fund and a poor fund is very low and that such power is not impacted by either model or benchmark misspecification.

Despite the mixed results in academic results on mutual fund performance measures, the mutual fund industry tends to rely on the simpler measures such as Jensen's alpha to measure and reward fund managers. To this end, our study provides a comprehensive and current evaluation of the consistency and rankings of these measures.

3. DATA AND METHODOLOGY

3.1 Data

We extract monthly returns from the CRSP 2004 Mutual Fund database. We select the funds in a manner similar to Carhart (1997) and Bollen and Busse (2005). We exclude funds with no stated objective, no expense, or load data and restrict the sample to U.S. equity funds with a stated objective of aggressive growth, growth and income, or long-term growth, as identified by ICDI objective. Our sample begins January 1, 1993 and ends December 31, 2004. Since our analysis involves regressions per fund, we require funds to have at least three years of monthly return data and information on total net assets, turnover, net asset value, and total loads. Our data are free of survivor-bias as CRSP reports all funds in existence during this period.

Parsing down by the above criteria, our final sample comprises of 2,861 mutual funds, with 207,952 total monthly observations. Thus, the average fund in our sample has about six years of available monthly data. Of the total sample, 881 funds are aggressive growth, 669 are growth and income, and 1,311 are long-term growth funds. Table 1 provides a description of the funds in our sample. Panel A of Table I presents the breakdown of the funds by year and fund objective. The number of sample funds with available data increases every year through 2002. This trend is especially pronounced for aggressive growth funds objective, which increases sevenfold from 1993 to 2002. Growth and income and long-term growth funds increase by approximately the same percentage over this time period. Between 2002 and 2004, the number of sample of funds with available data decline in every category, but the decline is most pronounced for long-term growth funds.

[INSERT TABLE I ABOUT HERE]

Panel B of Table I presents time-series cross-sectional averages for the funds in our sample. Total net assets (TNA), annual percentage turnover, net asset value (NAV), and the total load percentage are presented for the total sample and by fund objective. The 2,861 sample funds have average total net assets of $561.28 million over the sample period. Growth and income funds are the largest category, with total net assets of $791.25 million, while aggressive growth funds are the smallest, with total net assets of 317.08 million. The average fund in our sample has turnover of about 90 percent per year, where turnover is defined as the minimum of the aggregate purchases or sales of securities divided by the total net assets of the fund. Not surprisingly, aggressive growth funds had the highest turnover of about 108 percent, while growth and income funds had the lowest turnover of about 64 percent. The turnover percentages for our sample are somewhat higher than those reported by Carhart (1997), reflecting the hot markets of the late 1990s and early 2000. The average net asset value and total load of the sample funds over the time period investigated and across the sample funds is $16.89 and 4.14 percent, respectively. Both net asset value and total load percentage are fairly consistent across the different fund objectives.

Since our objective is to investigate whether the five performance measures result in a different rank order of funds over the sample period, we use the CRSP value-weighted index as the benchmark in all our performance measures. A consistent benchmark across the different performance models ensures that different rank orders are not due to different benchmarks. We extract the CRSP value-weighted index for NYSE, Amex, and Nasdaq stocks from CRSP and the monthly size, book-to-market, and momentum factors, as well as one-month T-bill rates from Ken French's website.

3.2 Methodology

Capital Asset Pricing Model

The traditional CAPM Jensen performance model is shown below:

$$R_{it} = \alpha_{iT} + \beta_{iT} RM_t + e_{it} \qquad t = 1,2,...,T \quad , \qquad (1)$$

where R_{it} is the return on a portfolio (mutual fund) in excess of the one-month T-bill return, and RM_t is the excess return on the CRSP value-weighted index of all NYSE, Amex, and Nasdaq stocks. If the portfolio manager has superior (inferior) ability, the intercept α_{iT} in equation (1) is positive (negative). The intercept represents the average incremental rate of return on the portfolio per unit of time. This incremental rate of return is due solely to the manager's ability to forecast future security prices. A naive random selection buy-and-hold policy should yield a zero intercept.

Fabozzi, Francis, and Lee (1980) find that the use of monthly returns results in a fairly constant beta coefficient. Thus, since our data consist of monthly returns, the risk level of the fund under consideration should remain fairly constant. This constant risk level may alleviate problems that are common with market-timing techniques that may cause shifting in the risk level of the fund.

Treynor Index/Sharpe Index

One measure of risk is the beta, which measures the sensitivity of the stock's (portfolio's) returns to market returns. An alternative risk measure is the standard deviation of the stock's

(portfolio's) historical returns. Two commonly used techniques to measure risk-adjusted returns are the Treynor Index and the Sharpe Index, which are stated below:

$$TI_i = \frac{\overline{R}_{it}}{\beta_{iT}}, \quad \text{where } \overline{R}_{it} = \sum_{t=1}^{T} R_{it} / T \tag{2}$$

$$SI_i = \frac{\overline{R}_{it}}{\sigma_{iT}}, \quad \text{where } \overline{R}_{it} = \sum_{t=1}^{T} R_{it} / T \tag{3}$$

In Equation (2), β_{iT} is obtained from Equation (1) for each fund *i*. In Equation (3), σ_{iT} is the standard deviation of monthly returns for each fund *i* over the sample period from January 1, 1993 to December 31, 2004.

Fama-French Three-Factor Model

The Fama-French (1992) three-factor model is stated below:

$$R_{it} = \alpha_{iT} + \beta_{iT} RM_t + \tau_{iT} SMB_t + \rho_{iT} HML_t + e_{it} \qquad t = 1,2,...,T, \tag{4}$$

where *SMB* and *HML* are returns on value-weighted, zero-investment, factor-mimicking portfolios for size and book-to-market equity, respectively. The Fama/French factors are constructed using six value-weight portfolios formed on size and book-to-market. The portfolios are the intersections of two portfolios formed on size (market equity, ME) and three portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint is the median NYSE market equity at the end of the period. The BE/ME breakpoints are the 30th and 70th NYSE percentiles. SMB (Small Minus Big) is the average return on the three small portfolios minus the average return on the three big portfolios, *SMB = 1/3 (Small Value + Small Neutral + Small Growth) - 1/3 (Big Value + Big Neutral + Big Growth).* HML (High Minus Low) is the

average return on the two value portfolios minus the average return on the two growth portfolios, *HML = 1/2 (Small Value + Big Value) - 1/2 (Small Growth + Big Growth)*.[2]

Carhart Four-Factor Model

Carhart's (1997) four-factor model is:

$$R_{it} = \alpha_{iT} + \beta_{iT} RM_t + \tau_{iT} SMB_t + \rho_{iT} HML_t + \lambda_{iT} PR1YR_t + e_{it} \qquad t = 1,2,...,T\,, \quad (5)$$

where SMB, HML, and $PR1YR$ are returns on value-weighted, zero-investment, factor-mimicking portfolios for size, book-to-market equity, and one-year momentum in stocks returns, respectively.[3] The momentum factor is constructed from the six value-weighted portfolios formed using independent sorts on size and prior return of NYSE, AMEX, and NASDAQ stocks. It is the average of the returns on two (big and small) high prior return portfolios minus the average of the returns on two (big and small) low prior return portfolios. The portfolios are constructed monthly. Big means a firm is above the median market cap on the NYSE at the end of the previous month; small firms are below the median NYSE market cap. Prior return is measured from month -12 to -2. Firms in the low prior return portfolio are below the 30th NYSE percentile. Those in the high portfolio are above the 70th NYSE percentile.

Carhart (1997) reports that the four-factor model substantially improves on the average pricing errors of the CAPM and the three-factor model. In fact, the four-factor model eliminates almost all of the patterns in pricing errors, indicating that it well describes the cross-sectional variation in average stock returns. By capturing the size, book-to-market, and momentum factors, based on the studies by Fama and French (1993) and Jegadeesh and Titman (1993), the four-

[2] Descriptions are obtained from Ken French's website.
[3] *PR1YR* is obtained from Ken French's website.

factor model may result in a significantly different ranking from the Fama-French three-factor model.

4. Empirical Results

4.1 Performance measures

To investigate whether each of five different mutual fund performance measures results in significant overperformance over our sample period, we analyze the performance of our sample funds using each of the five measures. Furthermore, we place the sample funds into deciles according to each performance measure. This allows us to investigate the performance distribution of our sample funds across the different fund measures.

Table II presents the results for the CAPM model. The first row of Table II presents the average monthly return, standard deviation, alpha, and beta coefficient of Equation (1) for all funds over the entire sample period from 1993 to 2004. Equation (1) is estimated separately for each mutual fund. Then, the average intercept and coefficient are tested for significance.

As shown in Table II, the average monthly return between January 1, 1993 and December 31, 2004 is 0.47 percent, with a standard deviation of 0.71 percent. The average Jensen's alpha across all 2,861 sample funds is -0.04 percent, significant at the 1 percent level, with an average beta coefficient is 1.01. The negative Jensen's alpha confirms previous research that the average mutual fund underperforms the market.

[INSERT TABLE II ABOUT HERE]

When the funds are placed into deciles according to their alphas, the decile containing the funds with the highest alphas (decile 1) contains funds with an average alpha of 1.05 percent per month, indicating that the managers of these funds outperform the market. Moving down the

alpha deciles, the alpha turns a negative 0.04 percent per month in decile 5, confirming the fact that most funds in the sample underperform the market. Funds in decile 10 have an average alpha of negative 0.93 percent per month.

The average beta coefficient stays fairly constant between deciles 1 and 5, with average betas between 0.93 and 0.96. Beginning with funds in decile 6, however, the average beta increases to over 1, ending with a beta of 1.17 for funds in decile 10. Thus, lower performing funds appear to be riskier than the better-performing funds. The standard deviation of monthly returns is also increasing across the performance deciles, but it seems to be the most pronounced for extreme deciles. This suggests that total risk is highest for either extremely well-performing or extremely poorly-performing mutual funds.

The results for the Sharpe and Treynor Indexes are presented in Table III. Columns 1-3 of Table 3 present deciles for the Sharpe Index, while columns 4-6 present the deciles for the Treynor Index. Each index is computed for every fund for the history of monthly returns available using Equations (2) and (3). The funds are then ranked according to the index and placed into deciles, from best performing funds to worst performing funds. For both indices, the numerator is the excess return of the monthly CRSP value-weighted index over the one-month Treasury bill rate, which is the independent variable in Equation (1). For the Sharpe Index, the denominator is the standard deviation of monthly returns over the fund history; for the Treynor Index, the denominator is the beta from Equation (1).

[INSERT TABLE III ABOUT HERE]

The first row in Table III presents the average monthly return, standard deviation, and indices for the 2,861 sample funds over the sample period. The average monthly returns and standard deviation are the same as in Table II. The average Sharpe Index for the sample funds is

0.05 over the sample period, indicating that mutual fund investors earn, on average, 0.05 percent excess return for every percentage point of standard deviation. The average Treynor Index for the sample funds is 0.21 percent, indicating that mutual fund investors earn 0.21 percent excess return for a fund with a beta of 1. These results are similar to those found in the single-index model, as the Sharpe Index for the monthly excess return of the CRSP value-weighted index is 0.08, and the average monthly excess return of the CRSP value-weighted index above the one-month T-bill rate between January 1993 and December 2004 is 0.39 percent per month.[4]

Decile rankings by monthly returns and standard deviations show that monthly returns decrease, from 1.41 percent per month in decile 1 for either index to just under -1 percent in decile 10 for both the Sharpe and Treynor indices. Likewise, the standard deviation of monthly returns decreases until approximately decile 7 and then increases in deciles 8 through 10 for either index. These results are similar to our findings for fund alpha deciles with more risk exhibited in the extreme deciles. The average Sharpe and Treynor ratios for funds in decile 1 are 0.27 and 1.38% respectively, while the rankings for the lowest decile are -0.17 and -1.35 respectively. Values for both indices drop below the average market values in decile 5, with a Sharpe Index of 0.07 and a Treynor Index of 0.38 percent. Overall, the results for the Sharpe and Treynor measures are similar to those reported for Jensen's alpha.

We present the for the Fama-French three-factor model and the Carhart four-factor model in Table IV. Carhart (1997) finds that the four-factor model substantially improves on the pricing errors of the CAPM and that the size, market-to-book, and momentum factors account for much of the cross-sectional variation in mean returns, while not substantially increasing

[4] For the Dow Jones Industrial Average, the average monthly excess return over the sample period was even more pronounced, with 0.57 percent. The S&P 500 index had an average excess monthly return of 0.45 percent over the sample period.

multicollinearity. In Table IV, both models employ the CRSP value-weighted index as the market proxy. Equations (4) and (5) are then estimated for the entire sample. Then, funds are sorted according to their alpha coefficients and placed into decreasing performance deciles. Next, alpha and each of the coefficients is tested for significance in each decile.

[INSERT TABLE IV ABOUT HERE]

For the entire sample of 2,861 funds, the alpha is a statistically negative and significant 0.19 percent for both models, suggesting that the models yield similar results to each other. The market factor has a coefficient very close to 1.00 (0.98 for the Fama-French model and 0.99 for the Carhart model) for the entire sample of funds. In both models, the overall coefficient for the factor-mimicking size portfolio is positive and statistically significant (0.03 for Fama-French and 0.01 for Carhart), suggesting that small stocks outperformed large stocks over the sample period. The overall coefficient for the factor-mimicking book-to-market portfolio is also positive and significant, with a coefficient of 0.15 in both models. The coefficient for the factor-mimicking momentum portfolio has a positive and statistically significant coefficient of 0.03 in the Carhart model. This is consistent with Carhart's (1997) performance persistence results of positive momentum coefficients in most of his deciles one year after the estimation period.

Mutual funds in decile 1 in the Fama-French (Carhart) model have a positive average alpha of 0.55 percent (0.50 percent) and an average monthly return of 1.03 percent (0.98 percent), while funds in decile 10 have a negative alpha of 1.04 percent (0.93 percent) and an average monthly return of negative 0.46 percent (0.21 percent) for the Fama-French (Carhart) model. Also similar to the results reported for the earlier measures, the standard deviation of monthly returns is more pronounced for funds in the extreme deciles.

Turning to the deciles, the size and book-to-market factors in the Fama-French and Carhart models are mostly positive and significant, as would be expected.[5] The momentum factor in the Carhart model is negative and significant for the lower deciles, but positive and significant for the higher deciles. This indicates that better-performing funds reverse their previous momentum, while worse-performing funds experience return momentum. These results are similar to the results reported by Carhart (1997), except that his reported SMB coefficients for the 1963 to 1993 sample period are larger than the ones reported here. Furthermore, Carhart's coefficients for the HML factor are negative, though only significant in some of his deciles. Also, Carhart's momentum factor is positive and significant in his deciles 1 through 7, but turns negative in his deciles 8 through 10.

Generally, the results reported in Table II, III, and IV for the Jensen, Sharpe and Treynor Indices, the Fama-French three-factor and Carhart four-factor models all indicate our sample of 2,861 funds, in aggregate, underperform the market. Moreover, all five measures show a higher standard deviation of monthly returns for the extreme deciles. Thus, at first glance, it appears that all five measures yield similar results and rank the mutual funds similarly. Consequently, it appears at first glance that mutual fund managers may be indifferent as to the performance measure used to evaluate their performance if their compensation is based on performance relative to other funds or their specific fund's performance according to a given measure.

4.2 Fund Rankings

We next investigate if the funds are ranked differently according to each measure in the aggregate. We compute two different correlation coefficients of rankings in Table V. Panel A of

[5] The exception is deciles 7 through 10 for the book-to-market factor in the Carhart model.

Table 5 presents the Spearman correlation coefficients for the rankings of the 2,861 mutual funds according to the five performance measures. Panel B of Table 5 presents Kendall correlation coefficients for the rankings. Spearman correlation coefficients are based on the differences in squares between ranks, while Kendall's τ is based on the consistency if relative ranks ever the entire sample. If the funds are ranked the same across the five different performance measures, then the correlations in Table V should be both high in magnitude and strongly positive and significant.

[INSERT TABLE V ABOUT HERE]

The correlation results in Panels A and B of Table V can be summarized succinctly. The only significant correlation coefficients between rankings for the different performance measures are between the CAPM and the Sharpe Index, the CAPM and the Fama-French three-factor model, and between the Sharpe Index and the Treynor Index. All other correlation coefficients, both Spearman and Kendall, are insignificant in both panels. While the significant correlation coefficients for rankings, particularly between the Sharpe and Treynor Index and between the CAPM and Fama-French model are not that surprising due to the similar nature of these models, we should expect other significant correlations, particularly between the Fama-French and Carhart models, since they differ only by the momentum factor, and between the CAPM model and the Treynor model, since the latter uses the all information used by the former.

The absence of significant ranking correlations for the five mutual fund performance measures suggests that mutual fund managers could indeed selectively choose the method that more favorably ranks their funds' performance, both in absolute terms and relative to other funds. While the overall correlation coefficients for the ranking of funds are informative, it is interesting

to see if the different performance measures at least identify the same funds as the best and worst funds.

4.3 Best-Performing and Worst-Performing Funds

In order to investigate whether the performance measures identify the same funds as the best- and worst-performing funds over the sample period, we next identify the specific best and worst funds according to each performance measure and the average rank of those funds according to the other four performance measures. The results from this analysis are displayed in Table VI. Panel A of Table VI presents the five best-performing funds according to each performance measure and the average rank assigned to those funds by the other four measures. For example, the Boston Partner Small Cap Value II/Institutional Fund is ranked the best performing fund over the January 1993 to December 2004 sample period according to the CAPM measure. The other four performance measures ranked that fund, on average, as number 11 out of 2,861. Panel B of Table VI presents the five worst-performing funds and the average rank assigned to those funds by the other four measures. The total rank possible is 2,861, the total number of sample funds.

[INSERT TABLE VI ABOUT HERE]

In Panel A of Table VI, it is evident that the average performance rank assigned to the best-performing funds according to each method differs noticeably. The highest average rank assigned to any of the top five funds for any method is 10. Notice also that the funds listed in Panel A under each method are almost entirely different from each other,[6] which suggests that the different performance measures do not identify the same best five funds. The average rank assigned by the other four methods ranges from 10 for the Boston Partners Small Cap Value

[6] Exceptions of the Boston Partners and Fidelity funds are noted.

II/Institutional Fund, identified as second according to the Fama-French three-factor model, to 1,488 for the Style Manager: Large Cap Fund/A, identified as second according to the Sharpe ratio. The overall average of the average rankings displayed in Panel A is 242. Thus, a fund ranked in the top five according to any of the five mutual fund performance measures is ranked, on average, 242 according to any other measure.

Panel B of Table VI displays the average rank of the five worst-performing funds for each performance measure according to the other four performance measures. Interestingly, the results for the worst performing funds are more closely aligned across the different measures and some of the measures identify the same funds as the worst performers. For example, both the Fama-French and the Carhart model list the same five funds as the worst performers, and in the same order. The highest (best) rank assigned to any of the worst performers in Panel B by the other four methods is 225 (for the second worst performing fund according to the Treynor method, the Boston Partners Long Short Equity/Institutional fund). The lowest average rank, out of 2,861, is 2,859 for the INVESCO Advantage Fund/C, identified as the second worst-performing fund according to both the Fama-French and the Carhart model. The overall average of the average rankings displayed in Panel B is 2,631. While this average is similar in magnitude to the average of 242 for the best-performing funds in Panel A, it is mainly attributable to the two funds listed as the worst performers according to the Treynor Index and their relative high average ranking according to the other methods. Without these two funds, the overall average is 2,839. Thus, the five performance measures investigated here seem more successful at ranking the worst performers consistently than the better performers.

5. Conclusions and implications

Using a survivor-bias free sample of 2,861 funds with at least three years of available monthly returns on CRSP over the 1993 to 2004 period, we revisit the topic of mutual fund performance. Specifically, we examine the performance using the CAPM model, the Sharpe and Treynor Indices, the Fama-French three-factor and Carhart's four-factor models. We then rank the funds according to each performance measure and examine the correlation of the ranking across the five performance measures. Finally, for each performance measure, we compute the average rank according to the other four performance measures to determine the five best- and worst-performing funds. As performance measures indicate managerial abilities, our results have implications not only for mutual fund managers' compensation, but also for decisions regarding promotion, demotions, or terminations mutual fund managers.

The excess return results across the five performance measures are similar in several ways. First, all measures indicate that the average fund underperforms the market benchmark, which validates previous work (Jensen, 1969; Jensen, 1968; Sharpe, 1966; McDonald, 1974; Fama, 1972; Friend & Blume, 1970) when mean-variance models are used. It suggests that fund managers, despite fund type, are, on the average, not able to predict security prices well enough to outperform a buy-and-hold policy when the mean-variance models are used to measure performance. Second, when funds are sorted into decreasing performance deciles according to each performance measure, all measures indicate lower monthly returns for worse performers and higher standard deviations for extreme deciles. Thus, at first glance, it appears that the five performance measures rank funds similarly.

However, when computing Spearman and Kendall correlation coefficients for the rankings of sample mutual funds according to each performance measure, we find that only three measures are significantly related. The only significant correlation coefficients between rankings for the

different performance measures are between the Jensen alpha and the Sharpe Index, the Jensen alpha and the Fama-French three-factor model, and between the Sharpe Index and the Treynor Index. This finding suggests that the different performance measures rank funds differently according to their performance.

Last, we investigate whether the five best- and worst-performing funds according to each performance measure are ranked similarly across the performance measures. The results show that the five different performance measures result in similar rankings for worse-performing funds, but diverge more in identifying the better-performing funds. This is relevant from a mutual fund manager's perspective, as he/she may be able to pick and choose the performance measure that is used to evaluate the fund's performance and thereby influence both the absolute and relative rank of the fund; a fund manager could conceivably choose a method that identifies his/her fund as the top fund, while that fund is only ranked as 242 according to other evaluation measures. Nonetheless, the different evaluation measures result in similar fund rankings for poorly performing funds. This finding is relevant for an investment company that is attempting to replace the managers of its worst funds. Generally speaking, the different measures identify similar funds as the worst performers, and choosing a different performance measure to evaluate his or her performance may not be a successful strategy for the mutual fund manager trying to keep his or her job.

Further research is necessary into why performance models appear to rank mutual fund performance differently, especially for the better-performing funds. The ultimate question is what constitutes a well-performing fund according to each performance measure, and what aspects of mutual fund performance are captured by the different models. Perhaps some mutual fund managers should be evaluated according to one of the five measures based on the particular

fund's objective, but as long as the mutual fund manager has a choice in the performance model used to evaluate his or her performance, excessive mutual fund compensation will affect the returns to mutual fund shareholders. The divergent ranking of well-performing funds according to the five measures indicates that a broader spectrum of fund managers may receive higher, and potentially unreasonable, compensation.

REFERENCES

Ang, James S. and J.H. Chua. 1979. "Composite Measures for the Evaluation of Investment Performance." *Journal of Financial and Quantitative Analysis* 14: 361-384.

Arditti, F. D. 1971. "Another Look at Mutual Fund Performance." *Journal of Financial and Quantitative Analysis* 6: 909-912.

Banz, R.W. 1981. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics* 9: 3-18.

Basu, S. 1977. "Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Markets Hypothesis." *Journal of Finance* 32: 663-682.

Berk J.B., and R.C. Green, 2004. "Mutual Fund Flows and Performance in Rational Markets." *Journal of Political Economy* 112:1269-1295.

Black, F., M.C. Jensen, and M. Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests." in M.C. Jensen, *Studies in the Theory of Capital Markets*, New York: Praeger.

Blake, C., E. Elton, and M. Gruber. 1993. "The Performance of Bond Mutual Funds." *Journal of Business* 66: 371-403.

Blume, M.E. and I. Friend. 1973. "A New Look at the Capital Asset Pricing Model." *Journal of Finance* 28: 19-34.

Bollen N.P., and J.A. Busse, 2004. "Short-term Persistence in Mutual Fund Performance." *Review of Financial Studies*

Carhart, M.M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance* 52: 57-82.

Chan, Y.C. 1997. "Multivariate Testing of the Capital Asset Pricing Model in the Hong-Kong Stock Market." *Applied Financial Economics* 7: 311-316.

Chang, E. and W. Lewellen. 1985. "An Arbitrage Pricing Approach to Mutual Fund Performance." *The Journal of Financial Research* 8: 15-30.

Chunhachinda, P., K. Dandapani, S. Hamid, and A. Prakash. 1997. "Portfolio Selection and Skewness: Evidence from International Stock Markets." *Journal of Banking and Finance* 21: 143-167.

Chunhachinda, P., K. Dandapani, S. Hamid, and A. Prakash. 1994. "Efficiency of Portfolio Performance Measures: An Evaluation." *Quarterly Journal of Business and- Economics* 33: 74-87.

Diacogiannis, G. 1994. "Three-Parameter Asset Pricing." *Quarterly Journal of Business and Economics* 15: 149-158.

Dybvig, P.H. and S.A. Ross. 1985a. "Differential Information and Performance Measurement Using a Security Market Line." *Journal of Finance* 40: 383-399.

Dybvig, P.H. and S.A. Ross. 1985b. "The Analytics of Performance Measurement Using a Security Market Line." *Journal of Finance* 40: 401-415.

Elton, E., M. Gruber, S. Das, and M. Hlavka. 1993. "Efficiency with costly Information: A Reinterpretation of Evidence from Managed Portfolios." *Review of Financial Studies* 6: 1-22.

Fabozzi, F.J., J.C. Francis, and C.F. Lee. 1980. "Generalized Functional Form for Mutual Fund Returns." *Journal of Financial and Quantitative Analysis* 15: 1107-1119.

Fama, E.F., and K. French. 1992. "Common Risk Factors in the Returns of Stocks and Bonds." *Journal of Financial Economics* 33: 3-36.

Fama, E.F. 1972. "Components of Investment Performance." *Journal of Finance* 27: 551-567.

Friend, I. and M. Blume. 1970. "Measurement of Portfolio Performance Under Uncertainty." *American Economic Review* 61: 561-575.

Friend, I., M. Blume, and J. Crockett. 1970. *Mutual Funds and Other Institutional Investors*. McGraw-Hill, New York.

Grinblatt, A. and S. Titman. 1994. "A Study of Monthly Mutual Fund Returns and Performance Evaluation Techniques." *Journal of Financial and Quantitative Analysis* 29: 419-444.

Gruber, M. 1996. "Another Puzzle: The Growth in Actively Managed Mutual Funds." *Journal of Finance* 51: 783-810.

Hendricks, D., Patel, J., and R. Zechhauser, 1993. "Hot hands in Mutual Funds: Short-run Persistence of Relative Performance, 1974-1988." *The Journal of Finance* 48: 93-130.

Hodges, C., W. Taylor, and J. Yoder. 1997. "Stocks, Bonds, the Sharpe Ratio, and the Investment Horizon." *Financial Analysts Journal* 53: 74-80.

Jean, W. H. 1971. "The Extension of Portfolio Analysis to Three or More Parameters." *Journal of Financial and Quantitative Analysis* 6: 505-515.

Jegadeesh, N. and S. Titman. 1993. "Returns to Buying Winners and Selling Losers: Implication for Stock Market Efficiency." *Journal of Finance* 48: 65-91.

Ippolito, R. 1989. "Efficiency with Costly Information: A Study of Mutual Fund Performance, 1965-1984." *Quarterly Journal of Economics* 104: 1-23.

Jensen, M.C. 1968. "The Performance of Mutual Fund in the Period 1945-1964." *Journal of Finance* 39: 389-416.

Jensen, M.C. 1969. "Risk the Pricing of Capital Assets and the Evaluation of Investment Portfolios." *Journal of Business* 42: 167-247.

Jensen, M.C. 1972. "Optimal Utilization of Market Forecasts and the Evaluation of Investment Performance." in *Mathematical Methods in Investment and Finance*, edited by G. P. Szego and K. Shell, Amsterdam: North-Holland.

Keim, D. B. 1983. "Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence." *Journal of Financial Economics* 11: 13-32.

Keim, D.B. 1985. "Dividend Yields and Stock Returns: Implications of Abnormal January Returns." *Journal of Financial Economics* 14: 473-489.

Kraus, A. and R.H. Litzenberger. 1976. "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance* 31: 1085-1100.

Lee, A., R.L. Moy, and C.F. Lee. 1996. "A Multivariant Test of Covariance- Co-Skewness Restriction for the Three-Moment CAPM." *Journal of Economics and Business* 48: 515-523.

Lehmann, B., and D. Modest. 1987. "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons." *Journal of Finance* 42: 233-265.

Leland, H.E. 1999. "Beyond Mean-Variance: Performance Measurement in a Nonsymetrical World." *The Financial Analyst* 55: 27-36.

Lintner, J. 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *The Review of Economics and Statistics* 47: 13-37.

Litzenberger, R.H. and K. Ramaswamy. 1979. "The Effect of the Personal Taxes and Dividends on Capital Asset Prices." *Journal of Financial Economics* 76: 163-195.

Markowitz, H. 1952. "Portfolio Selection." *Journal of Finance* 7: 77-91.

McDonald, J. 1974. "Objectives and Performance of Mutual Funds: 1960-1964." *Journal of Financial and Quantitative Analysis* 9: 311-333.

Murthi, B.P.S., Y. K. Choi, and P. Desai. 1997. "Efficiency of Mutual Funds and Portfolio Approach Measurement: A Non-Parametric Approach." *European Journal of Operational Research* 98: 408-418.

Pastor, L., and R. Stambaugh, 2002. "Mutual Fund Performance and Seemingly Unrelated Assets." *Journal of Financial Economics* 63:315-349.

Pettengill, G.N., S. Sundaram, and I. Mathur. 1995. "The Conditional Relation Between Beta and Returns." *Journal of Financial and Quantitative Analysis* 30: 101-116.

Porter, B.R. and J.E. Gaumnitz. 1972. "Stochastic Dominance vs. Mean-Variance Portfolio Analysis: An Empirical Evaluation." *American Economic Review* 62: 438-446.

Prakash, A.J. and Bear, R.M. 1986. "A Simplifying Performance Measure Recognizing Skewness." *The Financial Review* 21: 135-144.

Reinganum, M. R. 1981. "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings Yields and Market Values." *Journal of Financial Economics* 8: 19-46.

Roll, R. 1977. "A Critique of the Asset Pricing Theory's Test Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics* 4: 129-176.

Roll, R. 1978. "Ambiguity when Performance is Measured by the Securities Market Line." *Journal of Finance* 33: 1051-69.

Sharpe, W.F. 1966. "Mutual Fund Performance." *Journal of Finance* 39: 119-138.

Sirri, E.R., and P. tufano, 1998. "Costly Search and Mutual Fund Flows." *The Journal of Finance* 53:1589-1622.

Stephens, A. and D. Proffitt. 1991. "Performance Measurement When Return Distribution are Nonsymmetric." *Quarterly Journal of Business and Economics* 30: 23-39.

Tobin, J. 1958. "Liquidity Preference as Behavior Towards Risk." *Review of Economic Studies* 67: 65-85.

Treynor, J.L. 1965. "How to Rate Management of Investment Funds." *Harvard Business Review* 43: 63-75.

Table I
Mutual Fund Distribution and Annual Statistics

Panel A reports the distribution of funds by category in each of the sample years 1993 to 2004. Total funds per year are funds with any available data in that year. Panel B reports time-series averages of annual cross-sectional averages from 1993 to 2004 for the total sample and for each fund objective (aggressive growth, growth and income, and long-term growth). TNA is total net assets at year end; turnover is the minimum of the aggregate purchases of securities or the aggregate sales of securities, divided by the total net assets of the fund; NAV is the net asset value of the fund; total load is the sum of front-end and rear-end loads, plus deferred sales charges.

Panel A – Distribution of Funds By Objective Over the Sample Period

	Fund Objective			
Year	Aggressive Growth	Growth and Income	Long-Term Growth	Total Funds Per Year
1993	113	110	209	432
1994	175	171	295	641
1995	210	206	324	740
1996	253	229	356	838
1997	358	320	531	1,209
1998	393	338	578	1,309
1999	493	403	725	1,621
2000	635	469	960	2,064
2001	787	551	1,066	2,404
2002	798	565	1,111	2,474
2003	694	541	901	2,136
2004	513	403	621	1,537

Panel B – Time-Series Averages of Cross-Sectional Average Annual Attributes, 1993-2004

		Fund Objective		
Variable	Total Sample (N = 2,861)	Aggressive Growth (N = 881)	Growth and Income (N = 669)	Long-Term Growth (N = 1,311)
TNA ($ millions)	$561.28	$317.08	$791.25	$592.96
Turnover (%)	89.94%	107.82%	64.37%	92.14%
NAV ($)	$16.89	$17.35	$17.42	$16.27
Total Load (%)	4.14%	4.01%	4.13%	4.22%

Table II
Jensen Model Mutual Fund Performance Results

All mutual funds with at least three years of monthly returns over the period January 1, 1993 to December 31, 2004 are included in the sample. For each mutual fund, the model intercept, alpha, is determined via regression analysis. Funds with the highest alpha are placed in decile 1; funds with the lowest alpha are placed in decile 10. RM is the CRSP value-weighted return on all NYSE, Amex, and Nasdaq stocks. t-statistics are in parentheses.

	Monthly Return	Std. Dev.	Alpha	RM
All Funds	0.47%	0.71%	-0.04%	1.01
			(-4.23)	(193.81)
Decile 1	1.06%	0.56%	1.05%	0.93
			(50.63)	(51.15)
2	0.74%	0.53%	0.43%	0.92
			(71.70)	(53.41)
3	0.67%	0.52%	0.20%	0.94
			(71.91)	(55.69)
4	0.63%	0.56%	0.07%	0.95
			(32.57)	(63.17)
5	0.63%	0.47%	-0.04%	0.96
			(-24.78)	(64.22)
6	0.51%	0.56%	-0.13%	1.01
			(-77.38)	(65.87)
7	0.37%	0.58%	-0.23%	1.04
			(-137.34)	(74.10)
8	0.28%	0.59%	-0.34	1.06
			(-141.07)	(84.61)
9	0.13%	0.69%	-0.50	1.08
			(-139.64)	(79.92)
Decile 10	-0.32%	1.00%	-0.93%	1.17
			(-41.94)	(60.87)

Table III
Sharpe Ratio and Treynor Ratio Mutual Fund Performance Results

All mutual funds with at least three years of monthly returns over the period January 1, 1993 to December 31, 2004 are included in the sample. For each mutual fund, the Sharpe and Treynor ratios are computed. For the Sharpe ratio, the average monthly excess return over the fund lifetime is divided by the standard deviation of monthly returns. For the Treynor ratio, beta is computed using the Jensen model. For each model, funds with the highest ratio are placed in decile 1; funds with the lowest ratio are placed in decile 10. t-statistics are in parentheses.

	Sharpe Ratio			Treynor Ratio		
	Monthly Return	Std. Dev.	Ratio	Monthly Return	Std. Dev.	Ratio
All Funds	0.47%	0.71%	0.05	0.47%	0.71%	0.21%
			(19.94)			(12.49)
Decile 1	1.41%	0.35%	0.27	1.41%	0.37%	1.38%
			(75.58)			(61.33)
2	1.09%	0.23%	0.16	1.09%	0.23%	0.87%
			(174.74)			(179.76)
3	0.95%	0.20%	0.12	0.94%	0.17%	0.66%
			(220.75)			(227.27)
4	0.83%	0.18%	0.09	0.83%	0.17%	0.51%
			(221.25)			(225.34)
5	0.68%	0.16%	0.07	0.69%	0.14%	0.38%
			(153.67)			(157.55)
6	0.53%	0.14%	0.04	0.53%	0.12%	0.24%
			(100.44)			(97.96)
7	0.35%	0.11%	0.01	0.34%	0.11%	0.07
			(20.68)			(19.25)
8	0.09%	0.12%	-0.03	0.09%	0.12%	-0.15%
			(-36.82)			(-38.37)
9	-0.25%	0.19%	-0.08	-0.25%	0.19%	-0.46%
			(-66.76)			(-70.84)
Decile 10	-0.96%	0.65%	-0.17	-0.94%	0.69%	-1.35%
			(-41.08)			(-13.01)

Table IV
Fama-French Model and Carhart Model Mutual Fund Performance Results

All mutual funds with at least three years of monthly returns over the period January 1, 1993 to December 31, 2004 are included in the sample. For each mutual fund, the model intercept, alpha, for the Fama-French and Carhart measures is determined via regression analysis. Funds with the highest alpha are placed in decile 1; funds with the lowest alphas are placed in decile 10. RM is the CRSP value-weighted return on all NYSE, Amex, and Nasdaq stocks. SMB, HML and PR1YR are the factor-mimicking portfolios for size, book-to-market, and one-year return momentum, respectively. t-statistics are in parentheses.

	Fama-French Model						Carhart Model						
	Monthly Return	Std. Dev.	Alpha	RM	SMB	HML	Monthly Return	Std. Dev.	Alpha	RM	SMB	HML	PR1YR
All Funds	0.47%	0.71%	-0.19%	0.98	0.15	0.03	0.47%	0.71%	-0.19%	0.99	0.15	0.01	0.03
			(-23.51)	(293.16)	(24.91)	(5.53)			(-25.39)	(253.57)	(26.13)	(2.71)	(8.07)
Decile 1	1.03%	0.61%	0.55%	0.92	0.15	0.06	0.98%	0.69%	0.50%	0.89	0.17	0.11	-0.06
			(31.37)	(66.08)	(7.44)	(3.32)			(35.91)	(57.10)	(10.30)	(6.04)	(-4.44)
2	0.79%	0.49%	0.15%	0.93	0.14	0.03	0.87%	0.40%	0.16%	0.90	0.06	0.12	-0.06
			(45.47)	(69.98)	(6.79)	(2.09)			(54.25)	(62.13)	(3.58)	(6.94)	(-6.00)
3	0.74%	0.52%	0.01%	0.96	0.04	0.03	0.72%	0.46%	0.03%	0.93	0.02	0.09	-0.05
			(9.58)	(107.63)	(2.49)	(1.88)			(14.24)	(94.27)	(1.09)	(6.22)	(-5.12)
4	0.64%	0.50%	-0.06%	0.97	0.05	0.64	0.53%	0.48%	-0.06%	0.96	0.05	0.04	-0.00
			(-54.09)	(95.80)	(3.04)	(2.02)			(-48.79)	(85.81)	(3.32)	(2.44)	(-0.23)
5	0.55%	0.57%	-0.12%	0.97	0.08	0.04	0.55%	0.56%	-0.13%	0.97	0.07	0.03	0.00
			(-121.49)	(110.33)	(4.25)	(2.85)			(-102.30)	(120.63)	(4.37)	(2.15)	(0.50)
6	0.50%	0.59%	-0.19%	0.98	0.12	0.03	0.41%	0.66%	-0.21%	0.97	0.10	0.04	0.01
			(-161.54)	(117.14)	(6.85)	(2.30)			(-144.38)	(96.23)	(6.27)	(2.37)	(1.65)
7	0.42%	0.54%	-0.27%	0.99	0.14	0.04	0.34%	0.60%	-0.30%	1.00	0.14	0.02	0.34
			(-175.71)	(111.77)	(7.43)	(2.50)			(-206.05)	(108.02)	(7.45)	(1.37)	(3.29)
8	0.39%	0.58%	-0.37%	1.01	0.17	0.00	0.40%	0.64%	-0.39%	1.06	0.23	-0.07	0.10
			(-198.72)	(131.77)	(8.91)	(0.13)			(-223.00)	(106.42)	(12.99)	(-4.82)	(9.28)
9	0.12%	0.69%	-0.53	1.03	0.26	-0.00	0.12%	0.82%	-0.53%	1.11	0.28	-0.10	0.12
			(-132.42)	(108.03)	(13.03)	(-0.08)			(-153.91)	(91.60)	(15.46)	(-6.09)	(11.58)
Decile 10	-0.46%	0.87%	-1.04%	1.05	0.39	0.01	-0.21%	0.89%	-0.93%	1.13	0.39	-0.13	0.18
			(49.64)	(87.19)	(20.38)	(0.65)			(-49.07)	(91.92)	(20.05)	(-7.35)	(13.40)

Table V
Correlation Coefficients for Different Mutual Fund Performance Measures

For the five mutual fund performance measures in Tables II, III, and IV, funds are ranked and correlation coefficients between the five performance measures are computed. Spearman's ρ is based on the differences in squares between ranks. Kendall's τ is based upon the consistency of relative ranks over the entire sample. Both Spearman's ρ and Kendall's τ make only nonparametric assumptions about the associated distributions. p-values are in parentheses.

Panel A – Spearman Correlations Between Five Mutual Fund Performance Measures

	CAPM	Sharpe	Treynor	Fama-French	Carhart
CAPM	1.000	0.062	0.009	0.055	0.017
		(0.001)	(0.628)	(0.003)	(0.372)
Sharpe		1.000	0.059	0.001	0.016
			(0.002)	(0.946)	(0.401)
Treynor			1.000	0.028	0.020
				(0.135)	(0.284)
Fama-French				1.000	0.027
					(0.148)
Carhart					1.000

Panel B – Kendall Correlations Between Five Mutual Fund Performance Measures

	CAPM	Sharpe	Treynor	Fama-French	Carhart
CAPM	1.000	0.041	0.006	0.037	0.011
		(0.001)	(0.634)	(0.003)	(0.373)
Sharpe		1.000	0.040	0.001	0.011
			(0.001)	(0.946)	(0.399)
Treynor			1.000	0.019	0.013
				(0.133)	(0.286)
Fama-French				1.000	0.018
					(0.149)
Carhart					1.000

Table VI
Relative Rank Order of Five Best-Performing and Five Worst-Performing Funds

For the five mutual fund performance measures in Tables II, III, and IV, funds are ranked. All funds with at least three years of monthly returns and cross-sectional data over the sample period from January 1, 1990 to December 31, 2004 are included in the sample. For each measure, the five best-performing and the five worst-performing funds are listed and the average rank of that fund according to the other four methods is presented. Funds are listed in order of decreasing performance.

Panel A – Five Best Performing Funds

CAPM	Avg. Rank	Sharpe	Avg. Rank	Treynor	Avg. Rank	FF	Avg. Rank	Carhart	Avg. Rank
Boston Partners Small Cap Value II/Instl	11	Atlas Assets:Growth and Income Fund/B	462	Royce Fund:Special Equity Fund	13	Marshall Funds:Mid Cap Value Fund/A	419	Fidelity Advisor Leveraged Company Stock/A	51
Boston Partners Small Cap Value II/Invest	13	Style Manager:Large Cap Fund/A	1488	Federated Market Opportunity Fund/A	172	Boston Partners Small Cap Value II/Instl	10	Fidelity Advisor Leveraged Company Stock/T	54
Mercury HW Small Cap Value Fund/I	17	Sierra Trust:Growth and Income Fund/B	1432	Federated Market Opportunity Fund/C	221	Boston Partners Small Cap Value II/Invest	12	Fidelity Leveraged Company Stock	55
Provident Invmnt Counsel Mid Cap Fund/A	291	GE Funds:US Equity Fund/A	714	Federated Market Opportunity Fund/B	215	PIMCO Funds:Renaissance Fund/A	101	Fidelity Advisor Leveraged Company Stock/B	59
Fidelity Advisor Leveraged Company Stock/A	50	Royce Fund:Special Equity Fund	12	Boston Partners Small Cap Value II/Instl	10	PIMCO Funds:Renaissance Fund/B	116	Fidelity Advisor Leveraged Company Stock/C	60

Panel B – Five Worst-Performing Funds

Merrill Lynch Focus Twenty Fund/I	2829	Chase Vista Focus Fund/A	2820	INVESCO Advantage Fund/A	2857	Merrill Lynch Focus Twenty Fund/B	2834	Merrill Lynch Focus Twenty Fund/B	2822
Merrill Lynch Focus Twenty Fund/A	2831	Chase Vista Focus Fund/B	2829	INVESCO Advantage Fund/B	2858	INVESCO Advantage Fund/A	2857	INVESCO Advantage Fund/A	2857
Merrill Lynch Focus Twenty Fund/B	2834	INVESCO Advantage Fund/A	2857	INVESCO Advantage Fund/C	2859	INVESCO Advantage Fund/B	2858	INVESCO Advantage Fund/B	2858
Merrill Lynch Focus Twenty Fund/C	2834	INVESCO Advantage Fund/B	2858	Boston Partners Long Short Equity/Instl	225	INVESCO Advantage Fund/C	2859	INVESCO Advantage Fund/C	2859
Frontier Funds:Equity Fund Portfolio	2796	INVESCO Advantage Fund/C	2859	Boston Partners Market Neutral/Invest	247	Frontier Funds:Equity Fund Portfolio	2796	Frontier Funds:Equity Fund Portfolio	2776