Financial Services Review 22 (2013) 349–365

FINANCIAL SERVICES REVIEW

A comparison of single factor and multiple factor alphas used in measuring mutual fund performance

Brian Betker,[a,*], Joseph Sheehan, CFA, CFP[b]

[a]Department of Finance, John Cook School of Business, Saint Louis University, 3674 Lindell, St. Louis, MO 63108, USA
[b]Principal, The Moneta Group, 100 S. Brentwood, St. Louis, MO 63105, USA

Abstract

We compare mutual fund alphas computed using single factor models commonly used by practitioners with the alphas computed using the multifactor models common in academic research. Although all methods seek to control for a fund's investing style, single-factor alphas exceed multifactor alphas by 50–75 basis points per year on average, with the biggest differences concentrated in small cap funds. Although good performance does not tend to persist, single-factor alphas are about twice as likely to show persistence compared to three- and four-factor alphas. Investors should add multifactor alphas to the set of information used to evaluate mutual funds. © 2013 Academy of Financial Services. All rights reserved.

JEL classification: G11

Keywords: Financial planning; Mutual funds; Performance evaluation

1. Introduction

The two authors of this article have spent many hours (one as a principal and one as a consultant) sitting around the conference table at a St. Louis, Missouri wealth manager looking at data on fund performance, with the goal of helping to decide which funds are suitable for investment by clients (and which funds should be removed from the firm's recommended list).[1] In our discussions of this data, we came to the realization that the academic and the financial planner had very different methods for computing alpha. The

* Corresponding author. Tel.: +1-314-977-7154; fax: +1-314-977-1479.
E-mail address: Betkerbl@slu.edu (B. Betker)

most common models used in the academic literature are the Fama and French (1993) three-factor model and Carhart (1997) four-factor model.[2] Fama and French's model includes factors for firm size and book to market, while Carhart adds a factor for momentum. Commonly used practitioner databases also provide alphas that adjust for investing style (size and value/growth), but do so in a single-factor framework. Thus, we wondered, simply, whether the method matters. In particular, we wondered if there were any consistent patterns in the relative magnitude of single and multifactor alphas and if these patterns were related to fund style.

Thus, the goal of this article is straightforward: we ask whether the "practitioner" (single-factor) alphas yield the same inferences about fund performance as the "academic" (multifactor) alphas. Or, does fund performance look better (or worse) using single-factor alphas? Furthermore, does good performance as measured by single-factor alphas persist over time? Essentially, we want to know whether a financial planner would invest client funds (or an individual would invest their own funds) differently depending on whether they were looking at single-factor or multifactor alphas for the same set of mutual funds. Years of academic research have shown that on average funds do not generate alpha, and (more importantly) there is no strong evidence of persistence in alpha. If the single-factor alphas are higher and more persistent, investors who rely on these measures may be mis-allocating funds—for example, choosing active managers based on single factor alphas, when the multifactor alphas suggest that indexing might make more sense.

Thus, our goal is not to advance the state of the art in performance measurement, but rather to provide financial planners and investors with some useful information about fund evaluation that can be easily applied using publicly available data. For example, Daniel, Grinblatt, Titman, and Wermers (1997) examine mutual fund performance using "characteristic based benchmarks" that are not publicly available and could not be easily reproduced by an analyst. In contrast, the Fama and French (1993) factors are made freely available on Ken French's Web site.

2. Alpha

Since Jensen's (1968) pioneering study of mutual fund performance, alpha has been a common metric used in assessing fund performance. Alpha is the return that remains after subtracting the "normal" return that the fund should have earned given its risk and/or style characteristics. Jensen's original study used a single-factor market model with the Standard and Poor's 500 (S&P 500) representing the market. Since then there have been hundreds of studies that assess fund performance, many of which use more complicated models to explain fund returns (recognizing, e.g., that the S&P 500 is not an appropriate benchmark for a small cap value fund). The vast majority of these studies conclude that a typical mutual fund does not beat its benchmark. Evidence on performance persistence is mixed, but in general those funds that do "beat the market" in any given period aren't especially likely to beat the market in the next period (i.e., positive alphas do not persist over time).[3] We consider five different models for estimating alpha: three single factor models, the three-factor model of Fama and French and the four-factor model of Carhart.

2.1. Single index alphas

A single-index model computes alpha from the following regression:

$$(r_{pt} - r_{ft}) = \alpha_p + \beta_p(r_{mt} - r_{ft}) + \varepsilon_{pt} \quad t = 1, 2, \ldots, T$$

Where r_p is the fund return, r_m is the return on the index and r_f is the return on one-month U.S. Treasury Bills. We use the following as indexes:

1. Standard and Poor's 500 (S&P 500). This is the index used by Jensen (1968) in his original study and is the default option in some commonly used databases.
2. Best fit index. This method chooses the index with the highest correlation to the fund's returns as the single index. We consider the Russell indexes for each of the nine equity style boxes and choose as the single index the one that has the highest R^2 in the index model regression.[4]
3. Style benchmark. This method derives a single index which is the weighted average of a set of indices. Given a set of indices $C_1, C_2, \ldots, C_N$, determine weights $x_1, x_2, \ldots, x_N$, so that the composite index S

$$S = x_1 C_1 + x_2 C_2 + \ldots + x_N C_N$$

is a "best fit" for the fund's returns. Sharpe (1992) proposes determining the weights so as to minimize the expression

$$\mathrm{Var}\,(r_p - S)$$

where Var is the variance. We use a set of five indexes to compute S: the Russell Large Growth, Large Value, Small Growth, and Small Value Indices, as well as the Citigroup three-month T-Bill index.[5]

2.2. Three- and four-factor models

Fama and French (1993) propose a three factor model:

$$r_{pt} - r_{ft} = \alpha_p + b_p(r_{mt} - r_{ft}) + h_p(HML_t) + s_p(SMB_t) + \varepsilon_{pt} \quad t = 1, 2, \ldots, T$$

Where r_f is the one-month T-Bill rate, r_m is the return on the CRSP value-weighted index, HML (High Minus Low) is the return on a portfolio of high book-to-market (value) stocks minus the return on a portfolio of low book-to-market (growth) stocks, and SMB (Small Minus Big) is the return on a portfolio of small-cap stocks minus the return on a portfolio of large-cap stocks.

Carhart (1997) proposes adding a fourth (momentum) factor to the Fama-French model:

$$r_{pt} - r_{ft} = \alpha_p + b_p(r_{mt} - r_{ft}) + h_p(HMLt) + s_p(SMB_t) + u_p(MOM_t) + \varepsilon_{pt}$$
$$t = 1, 2, \ldots, T$$

where MOM is the return on a portfolio of high momentum stocks minus the return on a portfolio of low momentum stocks.

3. Data and methods

We use the Zephyr database, which contains monthly returns data from 1991 to 2010 for all mutual funds which exist as of the end of 2010. The data therefore has a survivorship bias that makes it unsuitable for a study of mutual fund performance. However, since our only objective is to compare alphas computed using different methods, the data are suitable for our purpose.

We examine only actively managed U.S. equity funds. After eliminating index funds, bond funds and international funds, we have monthly returns, with dividends, for 2,703 funds. Our data also contains monthly index returns (with dividends) for the standard U.S. equity indexes (Russell, Wilshire, Dow Jones, S&P, etc.). Data needed to compute the three and four factor alphas comes from Ken French's Web site,[6] which also contains detailed descriptions of how the four factors are computed.

We compute 10, five, and three year alphas for all available mutual funds, using non-overlapping data periods, which means that we can compute up to two 10-year alphas, four five-year alphas and six three-year alphas for funds that have complete data. If a fund is missing any returns data for a given period we do not compute its alpha for that period. Funds enter the database over the two decade period, so, for example, the number of three-year alphas for 2006–2008 is larger than the number for 1991–1993.

4. Results

4.1. Summary statistics for alpha

We first examine the distribution of alphas generated by the five methods, which we will refer to as the S&P, best fit, style benchmark, three-factor and four-factor methods. Note that all reported alphas are annualized; if the regression intercepts are denoted as α_{monthly}, then we report:

$$\alpha_{\text{annual}} = (1 + \alpha_{\text{monthly}})^{12} - 1$$

Table 1 presents summary statistics for 10, five, and three year alphas in Panels A, B, and C, respectively. We discuss the results for five year alphas in Panel B; conclusions about 10 and three year alphas are qualitatively similar. For visual interest, Fig. 1 illustrates a comparison of the distributions for the five-year best-fit and four-factor alphas.[7]

The first thing to note is that the single-factor alphas are significantly higher than the multifactor alphas. For example, the mean (median) best fit alpha is 60 (57) basis points greater than the mean three-factor alpha, while the style alpha exceeds the three-factor alpha by a mean (median) 85 (73) basis points. Though not reported in the table, these means and medians are significantly different from each other at the 1% level, using a t test and Wilcoxon rank sum test, respectively. Alphas using the S&P 500 as the single index are even larger with an average (median) of 2.41% (1.50%) annualized.

Essentially, the single- versus multifactor approaches reverse the conclusion one would draw about the median fund manager in the database. Under the single-factor approach, one

Table 1 Summary statistics for alpha, 1991–2010

	S&P 500	Best fit	Style	Three factor	Four factor
A: 10-year alphas					
Average	2.79%	0.61%	1.01%	0.34%	0.42%
Median	1.84%	0.29%	0.48%	−0.18%	−0.12%
10th percentile	−1.70%	−2.97%	−2.56%	−3.36%	−3.22%
90th percentile	8.53%	4.12%	4.86%	4.02%	3.81%
Percent > 0	74.06%	54.58%	58.37%	47.02%	47.88%
Percent signif. > 0	20.77%	8.68%	12.18%	7.88%	8.03%
Percent < 0	25.94%	45.42%	41.63%	52.98%	52.12%
Percent signif. < 0	2.89%	7.35%	6.55%	9.78%	9.75%
N	3,250	3,250	3,250	3,250	3,250
B: Five-year alphas					
Average	2.41%	0.42%	0.67%	−0.18%	−0.11%
Median	1.50%	0.04%	0.20%	−0.53%	−0.50%
10th percentile	−2.95%	−3.98%	−3.60%	−4.77%	−4.46%
90th percentile	9.30%	5.23%	6.04%	4.68%	4.53%
Percent > 0	66.83%	50.72%	52.90%	41.78%	42.01%
Percent signif. > 0	12.55%	6.06%	7.19%	3.60%	3.07%
Percent < 0	33.17%	49.28%	47.10%	58.22%	57.99%
Percent signif. < 0	3.86%	7.69%	6.87%	8.62%	8.41%
N	7,356	7,356	7,356	7,356	7,356
C: Three-year alphas					
Average	1.74%	0.37%	1.10%	0.02%	−0.10%
Median	0.90%	0.01%	0.20%	−0.28%	−0.35%
10th percentile	−6.36%	−6.19%	−5.29%	−6.41%	−6.08%
90th percentile	11.30%	7.47%	8.25%	6.82%	6.21%
Percent > 0	58.14%	50.08%	52.38%	46.84%	45.95%
Percent signif. > 0	7.55%	5.70%	7.14%	3.28%	3.12%
Percent < 0	41.86%	49.92%	47.62%	53.16%	54.05%
Percent signif. < 0	2.05%	8.93%	4.73%	4.87%	5.10%
N	11,105	11,105	11,105	11,105	11,105

This table presents summary statistics for alpha, computed for nonoverlapping intervals over the period 1991–2010. Percent signif. > (<) 0 indicates alpha estimates that are significantly different from zero at the 5% level.

would conclude that the median fund manager adds value; under the multifactor approach they do not.

Table 1 also reports the percentage of funds that generate a statistically significant (at the 5% level) alpha, defined as one with a t-statistic greater than two in absolute value. We can see that the vast majority of positive alphas computed using any method are in fact not reliably different from zero; a detail that is often not reported in databases that contain performance data. Even so, the single-factor alphas are about twice as likely to be significantly different from zero compared with the multifactor alphas (e.g., 6.06% of best-fit alphas compared to 3.07% of four-factor alphas).

4.2. Alphas in the style boxes

We further investigate the distribution of alpha in Table 2. Here we break funds out into Large cap growth (Panel A), Large cap value (Panel B), Small cap growth (Panel C), and

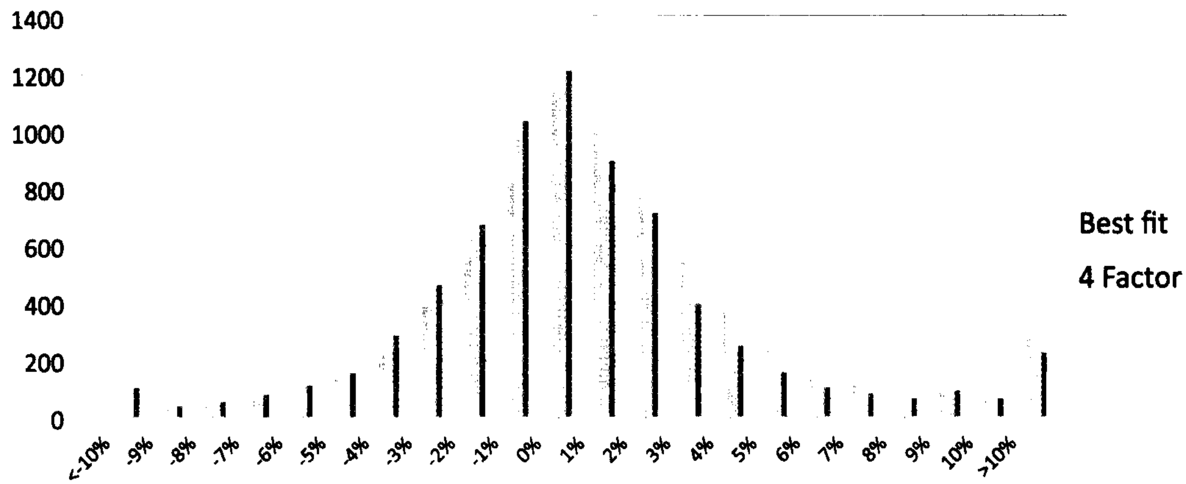

Fig. 1. Histograms for five-year best fit and four-factor alphas, January 1991–December 2010. This figure compares the histograms of the best-fit and four-factor alphas for nonoverlapping five-year periods between January 1991 and December 2010.

Small cap value (Panel D). For brevity's sake we only report results for these four style boxes. We use Morningstar's classification to sort funds into their style boxes.[8]

Again focusing on five-year alphas for purposes of discussion, we can see that for large cap funds (Panels A and B), the single- and multifactor approaches generate alphas that are very close to each other (save for the S&P alpha that is still much higher than the three and four factor alphas). In Panel A, median large cap growth alphas are all negative and no longer significantly different from each other (e.g., the median best fit alpha of −0.16% is not different from the median three-factor alpha of −0.28% using a Wilcoxon rank sum test). Median large-cap value alphas (Panel B) are again all negative, although single-factor medians are in this case significantly greater than multifactor medians.

Results for small cap funds in Panels C and D are quite different, again focusing on five-year alphas. As would be expected, S&P 500 alphas are very large, averaging 3.40% per year for small cap growth and 6.26% per year for small cap value. However, the S&P 500 is obviously a poor benchmark for small cap firms. More strikingly, for small cap growth funds (Panel C) median best fit and style alphas exceed the multifactor alphas by 228 to 246 basis points per year, and these differences are highly statistically significant. In Panel D, the corresponding differences range from 213 to 244 basis points per year. Looking at it another way, an analyst using the five-year style allocation alpha would conclude that 65.58% of small cap value funds generated positive alpha, whereas another analyst looking at the four-factor alpha would conclude that only 33.77% of funds did so. This could well convince the first analyst that active management of small cap value funds was worthwhile, while the second analyst might come to the opposite conclusion.

4.3. Fund rankings

We now turn from absolute performance to relative performance. Do the single-factor and multifactor alphas rank funds in the same way? If a fund appears to be a top-quartile performer using a single-factor alpha, will it also appear as a top-quartile performer using a

Table 2 Summary statistics for alpha, by fund objective, 1991–2010
Panel A: Large cap growth funds

	S&P 500	Best fit	Style	Three factor	Four factor
10-year alphas					
Average	1.50%	0.56%	0.84%	0.75%	0.60%
Median	0.62%	0.09%	0.02%	−0.18%	−0.30%
10th percentile	−2.76%	−3.65%	−3.03%	−3.21%	−3.20%
90th percentile	5.20%	4.29%	4.81%	5.16%	4.59%
Percent > 0	60.03%	52.29%	50.59%	46.79%	45.35%
Percent signif. > 0	10.62%	7.08%	9.31%	10.35%	8.26%
Percent < 0	39.97%	47.71%	49.41%	53.21%	54.65%
Percent signif. < 0	1.18%	6.16%	5.37%	6.42%	7.21%
Five-year alphas					
Average	1.44%	0.33%	0.56%	0.40%	0.22%
Median	0.74%	−0.05%	−0.16%	−0.28%	−0.38%
10th percentile	−3.97%	−4.33%	−4.04%	−4.27%	−4.25%
90th percentile	7.03%	5.14%	6.22%	6.09%	5.03%
Percent > 0	59.79%	49.00%	47.48%	45.55%	44.02%
Percent signif. > 0	5.86%	3.93%	4.98%	4.92%	2.99%
Percent < 0	40.21%	51.00%	52.52%	54.45%	55.98%
Percent signif. < 0	3.17%	7.85%	7.62%	7.15%	7.39%
Three-year alphas					
Average	0.99%	0.05%	0.90%	0.45%	−0.02%
Median	0.33%	−0.18%	−0.10%	−0.02%	−0.39%
10th percentile	−5.99%	−7.30%	−5.33%	−5.52%	−5.83%
90th percentile	8.79%	6.79%	7.53%	7.17%	6.18%
Percent > 0	52.86%	48.25%	48.99%	49.63%	45.96%
Percent signif. > 0	3.98%	2.92%	4.06%	3.11%	2.33%
Percent < 0	47.14%	51.75%	51.01%	50.37%	54.04%
Percent signif. < 0	3.07%	9.14%	6.35%	5.44%	6.31%

This table presents summary statistics for alpha, computed for nonoverlapping intervals over the period 1991–2010. Percent signif. > (<) 0 indicates alpha estimates that are significantly different from zero at the 5% level. Funds are categorized as large cap growth by Morningstar.

multifactor approach? If an investor has decided that hiring active managers best suits their investment philosophy, then as long as the various methods rank funds in the same way, it does not much matter which method is used to compute alpha.

Table 3 addresses this question. For ease of presentation, we compare rankings between the best fit and four-factor approaches, and then between the style allocation and four-factor approaches (comparisons between single-factor and three-factor alphas yield similar conclusions; we ignore the S&P 500 alpha in this exercise).

We place funds into quartiles based on their best-fit alpha and also their four-factor alpha, and then ask: looking at all funds in the (say) worst performance quartile using the best-fit alpha, what percentage of these funds appear in the worst performance quartile (or the second, third, or best quartile) using the four-factor alpha? If the two methods rank funds in a similar way, we would expect to see nearly all the "worst-quartile" best fit alphas also have "worst quartile" four-factor alphas, with almost no funds appearing in the second, third, or best four-factor quartiles. Hence each column in Table 3 adds up to 100%.

We then repeat the exercise for funds with second, third or best quartile performance using

Table 2 (Continued)
Panel B: Large cap value funds

	S&P 500	Best fit	Style	Three factor	Four factor
10-year alphas					
Average	2.43%	0.45%	0.63%	0.06%	0.39%
Median	1.85%	0.26%	0.23%	−0.24%	−0.10%
10th percentile	−0.99%	−2.52%	−2.29%	−3.04%	−2.56%
90th percentile	5.78%	3.60%	3.63%	3.22%	3.33%
Percent > 0	79.54%	53.56%	54.71%	43.91%	47.59%
Percent signif. > 0	17.93%	5.75%	7.59%	4.83%	5.29%
Percent < 0	20.46%	46.44%	45.29%	56.09%	52.41%
Percent signif. < 0	0.92%	6.21%	6.44%	8.74%	6.21%
Five-year alphas					
Average	1.78%	−0.18%	0.08%	−0.80%	−0.39%
Median	1.16%	−0.29%	−0.13%	−0.88%	−0.66%
10th percentile	−2.99%	−3.78%	−3.66%	−4.70%	−3.89%
90th percentile	7.75%	3.79%	3.87%	2.77%	3.20%
Percent > 0	62.79%	44.93%	47.14%	36.21%	39.12%
Percent signif. > 0	13.44%	4.61%	4.01%	2.21%	2.91%
Percent < 0	37.21%	55.07%	52.86%	63.79%	60.88%
Percent signif. < 0	2.91%	7.02%	6.12%	8.02%	6.52%
Three-year alphas					
Average	1.48%	−0.10%	0.14%	−0.71%	−0.48%
Median	0.95%	−0.22%	−0.21%	−0.77%	−0.48%
10th percenile	−7.12%	−7.05%	−6.33%	−7.37%	−6.31%
90th percentile	10.71%	6.72%	6.75%	6.15%	5.32%
Percent > 0	57.94%	48.17%	47.64%	41.99%	44.45%
Percent signif. > 0	9.70%	5.32%	5.78%	2.39%	2.66%
Percent < 0	42.06%	51.83%	52.36%	58.01%	55.55%
Percent signif. < 0	4.78%	9.24%	8.64%	8.70%	7.11%

This table presents summary statistics for alpha, computed for nonoverlapping intervals over the period 1991–2010. Percent signif. > (<) 0 indicates alpha estimates that are significantly different from zero at the 5% level. Funds are categorized as large cap value by Morningstar.

the best fit alpha, asking what four-factor quartile these funds fall into. Finally, we do the same thing in comparing style allocation alphas to four-factor alphas.

Turning to the results in Table 3, Panel A (10-year alphas), we see that there is considerable difference in the quartile rankings under the single- and four-factor approaches. Of the funds ranked worst quartile by the best fit alpha, 31% are ranked as second, third or best quartile by the four-factor alpha. Most of these (22.4%) are ranked in the second quartile. Likewise, of funds ranked top-quartile by the best fit alpha, 26% are ranked worst, second or third quartile by the four-factor alpha. Results are qualitatively similar when comparing style allocation alpha rankings to four-factor alpha rankings.

Panels B and C of Table 3 repeat the exercise for five and three year alphas, respectively, and yield similar conclusions. When a single-factor alpha ranks a fund into the best or worst performance quartile, about 20–25% of the time that fund is not ranked in the same performance quartile by the four-factor alpha. In the vast majority of those cases, the fund is ranked in the nearby performance quartile by the four-factor alpha (e.g., in the second quartile when the single-factor alpha performance is ranked as worst quartile).

Table 2 (Continued)
Panel C: Small cap growth funds

	S&P 500	Best fit	Style	Three factor	Four factor
10-year alphas					
Average	3.61%	0.51%	0.70%	−1.62%	−1.78%
Median	3.52%	0.20%	0.48%	−2.01%	−2.11%
10th percentile	−0.51%	−3.09%	−2.74%	−4.62%	−4.64%
90th percentile	7.42%	4.20%	4.28%	1.78%	1.53%
Percent > 0	85.95%	52.97%	57.84%	25.41%	23.24%
Percent signif. > 0	18.38%	11.35%	12.43%	2.16%	1.08%
Percent < 0	14.05%	47.03%	42.16%	74.59%	76.76%
Percent signif. < 0	1.08%	6.49%	5.41%	14.59%	19.46%
Five-year alphas					
Average	3.40%	1.36%	1.30%	−1.46%	−1.65%
Median	3.25%	0.61%	0.64%	−1.82%	−1.67%
10th percentile	−2.47%	−3.61%	−3.80%	−7.39%	−7.22%
90th percentile	9.31%	7.46%	7.13%	3.94%	3.33%
Percent > 0	78.14%	55.81%	57.21%	30.00%	30.93%
Percent signif. > 0	8.37%	11.16%	10.93%	2.79%	1.16%
Percent < 0	21.86%	44.19%	42.79%	70.00%	69.07%
Percent signif. < 0	0.93%	6.28%	6.28%	16.98%	18.37%
Three-year alphas					
Average	2.17%	1.90%	2.48%	−0.20%	−0.72%
Median	1.81%	0.88%	1.25%	−0.88%	−1.09%
10th percentile	−6.38%	−5.30%	−4.82%	−7.27%	−7.44%
90th percentile	10.57%	10.18%	11.34%	7.14%	6.25%
Percent > 0	62.60%	56.64%	59.24%	42.90%	39.08%
Percent signif. > 0	2.75%	12.37%	13.59%	3.82%	3.05%
Percent < 0	37.40%	43.36%	40.76%	57.10%	60.92%
Percent signif. < 0	0.92%	4.89%	3.66%	8.40%	9.01%

This table presents summary statistics for alpha, computed for nonoverlapping intervals over the period 1991–2010. Percent signif. > (<) 0 indicates alpha estimates that are significantly different from zero at the 5% level. Funds are categorized as small cap growth by Morningstar.

4.4. Persistence of alpha

Alphas, of course, can only be computed from historical data, so an important question is whether alphas persist. We investigate this question in two ways. First, we ask: if a fund generates a positive alpha in one period, is it likely to also generate a positive alpha in the following period? Then, we ask whether quartile rankings persist from one period to the next. We focus on persistence of positive alphas as these are likely to be most interesting to investors and financial planners.

Table 4 presents the results of the first investigation. Panels A–C report results using 10, five, and three-year alphas, respectively. Taking Panel C as an example, we compute alphas for the first three year period (1991–1993) and note that funds produced a positive alpha. Then in the subsequent three year period (1994–1996) we see how many of these funds produced a positive alpha again. We roll forward three years and repeat the experiment (i.e., if alpha was positive in 1994–1996, was it positive in 1997–1999?), continuing until the end of the sample period. Obviously, there will be many more comparisons of three-year alphas than of 10-year alphas.

Table 2 (Continued)
Panel D: Small cap value funds

	S&P 500	Best fit	Style	Three factor	Four factor
10-year alphas					
Average	7.11%	1.23%	1.74%	−0.55%	−0.14%
Median	8.03%	1.15%	1.56%	−0.66%	−0.48%
10th percentile	0.39%	−2.26%	−2.95%	−4.41%	−3.45%
90th percentile	11.55%	4.87%	5.64%	2.96%	2.96%
Percent > 0	89.86%	68.12%	71.01%	42.03%	46.38%
Percent signif. > 0	66.67%	11.59%	18.84%	1.45%	1.45%
Percent < 0	10.14%	31.88%	28.99%	57.97%	53.62%
Percent signif. < 0	0.00%	0.00%	1.45%	8.70%	5.80%
Five-year alphas					
Average	6.26%	1.27%	1.46%	−1.18%	−0.85%
Median	4.42%	0.74%	1.04%	−1.40%	−1.39%
10th percentile	−2.06%	−3.10%	−3.39%	−5.75%	−4.89%
90th percentile	16.69%	6.55%	6.87%	3.72%	4.53%
Percent > 0	80.52%	64.29%	65.58%	34.42%	33.77%
Percent signif. > 0	33.77%	11.04%	16.88%	4.55%	5.19%
Percent < 0	19.48%	35.71%	34.42%	65.58%	66.23%
Percent signif. < 0	0.65%	3.90%	4.55%	9.09%	9.09%
Three-year alphas					
Average	4.62%	1.27%	1.69%	−0.45%	−0.10%
Median	3.21%	0.40%	0.86%	−1.05%	−0.78%
10th percentile	−8.22%	−5.01%	−4.95%	−7.09%	−6.61%
90th percentile	20.04%	9.22%	8.60%	7.18%	7.59%
Percent > 0	69.79%	56.17%	58.30%	38.72%	42.55%
Percent signif. > 0	17.02%	11.06%	11.06%	4.68%	5.53%
Percent < 0	30.21%	43.83%	41.70%	61.28%	57.45%
Percent signif. < 0	3.40%	5.11%	4.68%	11.49%	8.94%

This table presents summary statistics for alpha, computed for nonoverlapping intervals over the period 1991–2010. Percent signif. > (<) 0 indicates alpha estimates that are significantly different from zero at the 5% level. Funds are categorized as small cap value by Morningstar.

Turning to the results in Table 4, we can see that the single-factor alphas show much more evidence of positive persistence than the multifactor alphas. Taking five-year alphas in Panel B as an example, the style allocation alpha is positive in period 1 for 2,468 funds; of these, 1418 (57.5%) also have a positive alpha in the subsequent five-year period. If subsequent-period over/under performance was just a matter of chance we would expect to see 50% of the firms generate a positive alpha in the second period. A two-tailed t test indicates that 57.5% is statistically significantly larger than 50% (t = 7.41). In contrast, the multifactor alphas do not show evidence of positive persistence. Fewer than half of the funds that generate a positive alpha in the first period manage a positive alpha in the second period. Inferences about persistence of alpha are similar when looking at 10 and three year results in Panels A and C.[9]

Although for brevity's sake we do not report these results in any table, we find that negative alphas are more likely to persist when using the multifactor models. Taking five-year alphas as an illustration, we find that for the best-fit and style allocation alphas, a

Table 3 Quartile rankings: Single-factor alphas vs. four-factor alphas

A: 10-year alphas		Funds ranked in this quartile using the best fit alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	69.0%	22.8%	6.5%	1.9%
	2nd	22.4%	47.0%	27.1%	3.3%
	3rd	7.6%	25.5%	46.1%	20.9%
	Best	1.0%	4.7%	20.3%	74.0%
		Funds ranked in this quartile using the style alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	74.0%	19.5%	5.2%	1.5%
	2nd	21.4%	55.0%	20.7%	2.7%
	3rd	3.8%	23.2%	55.4%	17.7%
	Best	0.7%	2.3%	18.7%	78.1%
B: Five-year alphas		Funds ranked in this quartile using the best fit alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	74.3%	18.4%	6.0%	1.4%
	2nd	20.5%	51.8%	23.4%	4.2%
	3rd	4.4%	26.2%	53.1%	16.3%
	Best	0.9%	3.5%	17.5%	78.1%
		Funds ranked in this quartile using the style alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	74.8%	18.3%	5.8%	1.1%
	2nd	20.3%	54.6%	21.5%	3.5%
	3rd	3.6%	25.2%	54.8%	16.4%
	Best	1.3%	1.8%	17.9%	78.9%
C: Three-year alphas		Funds ranked in this quartile using the best fit alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	73.1%	19.0%	6.7%	1.4%
	2nd	20.0%	52.3%	22.8%	4.8%
	3rd	5.7%	25.1%	51.5%	17.7%
	Best	1.2%	3.6%	19.0%	76.1%
		Funds ranked in this quartile using the style alpha			
		Worst	2nd	3rd	Best
Percentage of funds ranked in this quartile using Four-factor alpha	Worst	76.5%	16.5%	5.1%	2.0%
	2nd	20.7%	55.3%	20.9%	3.0%
	3rd	2.4%	26.2%	56.8%	14.6%
	Best	0.4%	1.9%	17.2%	80.4%

This table compares the quartile rank of the best fit or style alpha to the quartile rank of the four factor alpha. The table shows, each quartile of single-factor alpha performance, what percentage of funds ranked in the worst, second, third or best performance quartiles using the four-factor alpha. Thus each column adds up to 100%.

negative alpha in period one is followed by a negative alpha in period two 51.9% and 51.8% of the time, respectively, whereas the figures for the three- and four-factor alphas are 62.2% and 62.6%, respectively.

Table 4 Persistence of alpha

	Number with α positive in first period	Of these, number with α positive in second period	Proportion with α positive in second period	t-statistic for H_o: proportion = 50%
Panel A: 10-year alphas				
S&P500	413	339	82.1%	13.04†
Best fit	362	189	52.2%	0.84
Style	361	212	58.7%	3.32†
Three-factor	318	133	41.8%	−2.92†
Four-factor	351	159	45.3%	−1.76
Panel B: Five-year alphas				
S&P500	2,999	2,098	70.0%	21.86†
Best fit	2,285	1,246	54.5%	4.33†
Style	2,468	1,418	57.5%	7.41†
Three-factor	1,892	901	47.6%	−2.07*
Four-factor	1,933	934	48.3%	−1.48
Panel C: Three-year alphas				
S&P500	4,731	2,819	59.6%	13.19†
Best fit	4,078	2,135	52.4%	3.01†
Style	4,443	2,438	54.9%	6.50†
Three-factor	3,830	1,749	45.7%	−5.36†
Four-factor	3,755	1,763	47.0%	−3.74†

This table shows the number of funds with a positive alpha in the first measurement period, and, of those funds, which also had a positive alpha in the second measurement period. A t-statistic tests the null hypothesis that the proportion of funds with two consecutive positive alphas is equal to 50%.

*Significantly different from 50% at the 5% level.

†Significantly different from 50% at the 1% level.

4.5. *Persistence of alpha quartile rankings*

Finally, we ask whether our various alpha measures differ in terms of how they rank funds from one period to the next—if a firm is ranked in the top quartile in period 1, is it likely to be ranked in the top quartile in the next period? Thus, we examine persistence of *relative* performance (ranking) rather than absolute performance as in the prior section.

Table 5 contains the results, with Panels A, B, and C showing data for 10, five, and three year alphas, respectively. For each method of computing alpha, we ask, given all the funds that appeared in (say) the lowest performance quartile in period one, how were these funds distributed across the four performance quartiles in period two? If relative performance persisted, we would expect to see nearly 100% of these funds appear in the worst-performing quartile in period two. On the other hand, if period two relative performance is independent of period one performance, we'd expect to see about 25% of the funds appear in each period two performance quartile. Thus, each row in the table adds up to 100%. Under the null hypothesis that period two performance is not related to period one performance, we test whether the proportions in the table are statistically different than 25%.

Focusing again on five-year alphas (Panel B) for discussion purposes, we can see that there is clearly not perfect quartile ranking persistence—the entries off the diagonal are all substantially larger than 0%. Furthermore, with a few exceptions, most of the table entries are not significantly different from 25%, indicating that for the most part subsequent period

Table 5 Persistence of alpha quartile rankings.
Panel A: 10-year alphas

S&P 500 alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	25.6%	26.7%	15.3%	32.4%
	2nd	36.0%*	22.9%	25.7%	15.4%
	3rd	21.1%	32.0%	26.3%	20.6%
	Best	17.7%	18.3%	32.6%	31.4%
Best fit alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	19.9%	17.0%	23.3%	39.8%*
	2nd	26.3%	29.1%	28.0%	16.6%
	3rd	24.0%	29.1%	22.3%	24.6%
	Best	30.3%	24.6%	26.3%	18.9%
Style alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	19.9%	23.3%	18.8%	38.1%*
	2nd	28.0%	27.4%	30.9%	13.7%
	3rd	21.7%	30.3%	23.4%	24.6%
	Best	30.9%	18.9%	26.9%	23.4%
Three-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	17.6%	20.5%	21.6%	40.3%*
	2nd	24.6%	25.1%	30.9%	19.4%
	3rd	24.6%	32.6%	19.4%	23.4%
	Best	33.7%	21.7%	28.0%	16.6%
Four-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	22.2%	18.2%	20.5%	39.2%*
	2nd	26.3%	27.4%	25.7%	20.6%
	3rd	25.1%	30.9%	25.7%	18.3%
	Best	26.9%	23.4%	28.0%	21.7%

For each method of computing alpha, this table shows the distribution of fund quartile rankings from one period to the next. Each row in the table adds up to 100%.

*Proportion is significantly different from 25% at the 5% level of significance.

performance is not strongly related to prior period performance. In Panel B, the most persistence of alpha quartile rankings is actually shown by the multifactor alphas. For both the three- and four-factor alphas, the entries down the diagonal are all significantly larger than 25%, indicating that the most likely quartile rank for the second period is the same as the quartile rank for the first period. However, with proportions of only about 31%, the tendency for quartile rank to persist is not overwhelming.

Table 5 (Continued)
Panel B: 5-year alphas

S&P 500 alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	28.2%	25.5%	19.7%	26.6%
	2nd	26.1%	32.8%*	26.7%	14.4%*
	3rd	25.2%	25.6%	29.1%	20.1%
	Best	20.9%	16.0%*	24.4%	38.7%*
Best fit alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	27.5%	23.0%	21.9%	27.6%
	2nd	22.6%	34.0%*	26.9%	16.5%*
	3rd	24.9%	25.2%	28.7%	21.3%
	Best	25.4%	17.8%*	22.4%	34.4%*
Style alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	28.7%	22.5%	22.2%	26.6%
	2nd	23.4%	36.4%*	26.1%	14.2%*
	3rd	23.8%	24.3%	29.4%	22.5%
	Best	24.6%	16.8%*	22.2%	36.4%*
Three-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	30.4%*	23.6%	19.9%	26.1%
	2nd	23.4%	31.8%*	26.6%	18.2%*
	3rd	20.2%	27.0%	31.3%*	21.4%
	Best	26.4%	17.5%*	22.2%	34.0%*
Four-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	31.8%*	23.5%	19.4%	25.2%
	2nd	23.9%	31.1%*	27.8%	17.2%*
	3rd	21.5%	27.1%	30.6%*	20.7%
	Best	23.1%	18.2%*	22.1%	36.6%*

For each method of computing alpha, this table shows the distribution of fund quartile rankings from one period to the next. Each row in the table adds up to 100%.

*Proportion is significantly different from 25% at the 5% level of significance.

Conclusions regarding quartile ranking persistence in Panel C (three-year alphas) are largely consistent with those from Panel B. Turning briefly to Panel A (10 year alphas), it is interesting to note that the only pattern that emerges is one of worst to first—a fund that tallied worst-quartile performance in the first 10 years of the sample actually has the greatest chance of winding up in the top performance quartile in the second 10 years.

Table 5 (Continued)
Panel C: 5-year alphas

S&P 500 alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	25.2%	22.1%	23.1%	29.7%*
	2nd	27.0%	29.3%*	26.8%	16.9%*
	3rd	24.2%	31.0%*	27.8%	17.0%*
	Best	27.7%	16.3%*	21.1%	34.9%*
Best fit alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	25.3%	22.2%	23.9%	28.6%*
	2nd	23.1%	34.2%*	27.6%	15.2%*
	3rd	22.3%	28.3%	28.7%*	20.7%*
	Best	29.8%*	15.3%*	19.7%*	35.2%*
Style alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	23.1%	24.5%	23.7%	28.6%*
	2nd	26.6%	31.4%*	26.5%	15.5%*
	3rd	24.1%	28.1%	27.1%	20.8%*
	Best	26.7%	16.0%*	22.6%	34.8%*
Three-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	24.7%	21.0%	22.8%	31.5%*
	2nd	24.1%	30.2%*	27.7%	18.1%*
	3rd	22.0%	29.6%*	29.7%*	18.7%*
	Best	29.8%*	19.1%*	19.8%*	31.4%*
Four-factor alphas		The table shows the percentage of funds ranked in this quartile in period 2:			
		Worst	2nd	3rd	Best
Given that a fund ranked in this quartile in period 1	Worst	26.5%	22.2%	22.8%	28.4%*
	2nd	23.3%	31.2%*	27.7%	17.9%*
	3rd	23.2%	27.9%	29.9%	19.0%*
	Best	27.4%	18.7%*	19.5%	34.4%*

For each method of computing alpha, this table shows the distribution of fund quartile rankings from one period to the next. Each row in the table adds up to 100%.

*Proportion is significantly different from 25% at the 5% level of significance.

5. Conclusion

Conversations between this article's authors about measuring mutual fund performance led them to realize that the financial planner and the academic were looking at alpha in potentially rather different ways. The standard practitioner-oriented databases report single-factor alphas (S&P 500, best fit or style allocation) whereas standard academic practice is to use the three- or four-factor models of Fama and French (1993) and Carhart (1997) to estimate alpha. It seemed logical to ask whether this made any difference when evaluating funds. We look at absolute performance as well as relative performance (quartile rankings).

We find that single-factor alphas are larger than multifactor alphas, by 50–75 basis points on average. This result is concentrated in small cap funds; large cap fund alphas are not meaningfully different using best-fit or style allocation alphas compared with three- or four-factor alphas. However, for small cap funds the single-factor methods produce alphas that are more than 200 basis points larger than the multifactor methods. Single factors alphas are also more likely to show positive persistence than multifactor alphas.

When it comes to relative performance, funds that are ranked as top-quartile by the single-factor methods are not ranked as top-quartile 20–25% of the time when using multifactor methods.

We conclude that a financial planner or investor looking at single-factor alphas might come to a different conclusion about small cap mutual fund performance, compared with someone examining multifactor alphas. The case for active management of small cap funds certainly looks stronger when viewed through the lens of single-factor alphas. The data needed to compute the three- and four-factor alphas is readily available, and the computation requires nothing more than a spreadsheet. We recommend that investors and financial planners add multifactor alphas to the set of information used to evaluate mutual funds.

Notes

1 Of course, past performance is not the only, or even the most important, criterion used to select funds. Fees, management's tenure, comfort with the fund manager's investment philosophy, etc. are also important. Nonetheless the performance data is readily available and is part of the decision process.

2 Amenc, Goltz, and Lioui (2011) report (their Table 8) that most practitioners in their survey analyze alpha using single factor models, returns-based style analysis or peer-group analysis. Only 21% used multifactor models to analyze alpha.

3 The literature on mutual fund performance is voluminous and will not be reviewed here. Recent articles on the topic include Brooks and Porter (2012), Hsu, Kalesnik, and Myers (2010), and Duan, Hu and McLean (2009), all of which contain many references to the literature. Performance persistence has been studied by, among many others, Grinblatt and Titman (1992), Goetzmann and Ibbotson (1994), and Carhart (1997).

4 We also computed single-factor alphas using Morningstar's classification of the fund's style box (large cap growth, mid cap blend, etc.) to determine which index to use as the single factor, but these results were similar to the best fit approach and so are not reported. Detzel (2006) compares Morningstar's style box classifications to a factor-loadings approach to fund classification.

5 Computing S using these five indexes is the default option in the Zephyr database. We also computed the style benchmark alpha using seven indexes, adding the Russell

Midcap Growth and Midcap Value, but results using this method were similar and are not reported.

6 http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

7 Figures plotting the style allocation alpha versus the three- or four-factor alpha look very similar to Figure 1 and so for brevity's sake are not reported.

8 We also looked at defining a fund's style according to its best fit index, but the results are quite similar to those in Table 2 and are not reported.

9 Note that in Table 4 we are only asking whether alpha is positive, but in Table 1 we noted that most positive alphas are not statistically significantly different from zero. We also examined persistence of statistically significant alphas. Although for brevity's sake we do not report these results in any table, we also found that in this exercise significantly positive alphas were in general about twice as likely to persist when using single factor alphas compared to multi-factor alphas.

References

Amenc, N., Goltz, F., & Lioui, A. (2011). Practitioner portfolio construction and performance measurement: Evidence from Europe. *Financial Analyst's Journal, 67*, 39–50.

Brooks, L. D., & Porter, G. E. (2012). Mutual fund performance attribution: 1994–2005. *Financial Services Review, 21*, 259–273.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.

Daniel K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*, 1035–1058.

Detzel, F. L. (2006). Determining a mutual fund's equity class. *Financial Services Review, 15*, 199–212.

Duan, Y., Hu, G., & McLean, R. D. (2009). When is stock picking likely to be successful? Evidence from mutual funds. *Financial Analyst's Journal, 65*, 55–66.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Goetzmann, W. N., & Ibbotson, R. G. (1994). Do winners repeat? Patterns in mutual fund performance. *Journal of Portfolio Management, 20*, 9–18.

Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance, 47*, 57–82.

Hsu, J. C., Kalesnik, V., & Myers, B. W. (2010). Performance attribution: Measuring dynamic allocation skill. *Financial Analyst's Journal, 66*, 17–26.

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

Sharpe, W. F. (1992). Asset allocation: Management style and performance measurement. *Journal of Portfolio Management, 18*, 7–19.