First draft: October 2007
This draft: August 2008
Not for quotation: Comments welcome

Mutual Fund Performance

Eugene F. Fama and Kenneth R. French*

Abstract

In aggregate, mutual funds produce a portfolio close to the market portfolio but with high costs of active management that show up intact as lower returns. Persistence tests that sort funds on three-factor α estimates suggest information effects in the future returns of past winners and losers, but persistence is temporary, it is weak to nonexistent in sorts on average return, and it largely disappears after 1992. Bootstrap simulations that use entire histories of fund returns do not identify information effects in three-factor or four-factor α estimates.

* Graduate School of Business, University of Chicago (Fama) and Amos Tuck School of Business, Dartmouth College (French). We are grateful for the comments of John Cochrane and seminar participants at the University of Chicago.

Equilibrium accounting says that investors in U.S. equities in aggregate get the value-weight U.S. equity market return less their investment costs. Suppose that when returns are measured before costs, passive investors get passive returns; that is, they have zero α (abnormal expected return with respect to the true asset pricing model). Then if some active investors have positive α before costs, it is, dollar for dollar, at the expense of other active investors. Thus, like passive investment, active investment is in aggregate a zero sum game – aggregate α is zero before costs. After costs (that is, in terms of net returns to investors) active investment is a negative sum game. This is the argument in French (2008), and he gives references to earlier papers that make a similar point.[1]

We examine mutual fund performance from the perspective of equilibrium accounting. For example, at the aggregate level, if we find that the value-weight (VW) portfolio of all mutual funds produces a positive α before costs, we can infer that the VW portfolio held by investors outside mutual funds has a negative α. In other words, the mutual fund industry wins at the expense of investments held outside mutual funds. Our tests, however, do not produce this result. We find (Section I) that mutual funds in aggregate do not gain from or lose to other investors. The VW portfolio of mutual funds that invest primarily in U.S. equities is close to the market portfolio, and estimated before fees and expenses, its α is close to zero. Since the VW portfolio of funds produces an α close to zero in gross returns, the α estimated on the net returns to investors is negative by about the amount of fees and expenses. All this echoes equilibrium accounting, but for a subset of investment managers where the implications of equilibrium accounting for aggregate investor returns need not hold.

The aggregate results for mutual funds imply that if there are funds with positive α, they must be balanced by funds with negative α. Our main goal is to test for the existence of such winner and loser funds.

The challenge is to distinguish skill from luck. Given the multitude of funds, many have extreme returns by chance. A common approach to this problem is to test for persistence in fund returns, that is, whether past winners continue to produce high returns and losers continue to underperform (for example, Grinblatt and Titman 1992, Hendricks, Patel, and Zeckhauser 1993, Goetzmann and Ibbotson 1994, Carhart

[1] The implications of equilibrium accounting are often overlooked. For example, Berk and Green (2004) develop a model in which active investors in aggregate produce zero α after (rather than before) fees and expenses. Equilibrium accounting says this equilibrium is impossible unless passive investors lose massively to active investors. French (2008) provides evidence that this is not the case.

1997). If past performance is due to chance, it does not predict future performance. But if some managers have stock selection skills, persistence tests are a way to infer their existence. Persistence tests also have a downside. They typically rank funds based on short-term past performance. We may find little evidence of persistence because the allocation of funds to winner and loser portfolios is largely based on noise.

Section II extends and updates Carhart's (1997) evidence on return persistence. There are post-Carhart persistence studies, but they leave important open issues. For example, Carhart's (1997) informal evidence (plots of post-ranking raw returns) suggests that any persistence in performance is short-term. His results also suggest that performance is sensitive to the way funds are ranked. In formal tests that attack these issues, Bollen and Busse (2004) find no evidence of persistence when funds are ranked on past average return, but persistence shows up in rankings on α estimates from Carhart's (1997) four-factor model. They also find that the persistence uncovered in their tests is short-term, less than a year. Finally, Barras et al (2008) find that mutual fund performance is weaker in more recent time periods.

Bollen and Busse (2004) examine a small sample of 230 funds for the rather short 1985-1995 period. They sort funds using at most twelve months of past returns, which raises the possibility that the sorts are in large part based on noise. Our persistence tests use all mutual funds in the CRSP database that are available during 1984-2006 and invest primarily in U.S. equities, and our sorts use 12 to 60 months of past returns.

Our tests confirm that persistence results are sensitive to how funds are ranked. When we sort funds on past market-adjusted gross returns, we find little persistence in post-sort returns. Persistence is stronger when we sort on four-factor α or three-factor α estimates from the model of Fama and French (1993). These α sorts suggest that the loser decile contains some funds with bad information that lowers expected returns and the winner decile contains some funds with good information that enhances expected returns. The persistence results from sorts on α estimates are stronger for smaller funds, and they are weak after 1992.

The main contribution of this paper is a bootstrap simulation alternative to persistence tests that uses the entire return histories of all funds to infer the existence of superior and inferior managers. We compare the actual cross-section of fund α estimates to the results from 10,000 bootstrap simulations of the cross-section. The returns of the funds in a simulation run have all the properties of actual fund returns, except that in the simulations true α is set to zero. The simulations thus provide the distribution of the cross-section of α estimates when in truth there is no abnormal performance, good or bad, in fund returns. Comparing the distribution of α estimates from the simulations to the cross-section of α estimates for actual fund returns allows us to draw inferences about the existence of skilled managers.

For example, if the 10,000 simulation runs for gross returns (before fees and expenses) rarely produce a 10^{th} percentile α estimate below the 10^{th} percentile α estimate from the cross-section of actual gross fund returns, we can infer that there are fund managers with bad information that reduces expected returns. Likewise, if the 90^{th} percentile of α estimates from the simulation runs for gross returns is rarely above the 90^{th} percentile of the cross-section of actual fund α estimates from gross returns, we can infer that there are some managers with good information that enhances expected returns.

The simulations share a shortcoming of all performance tests. They only allow inferences about the existence of inferior or superior funds. Since a large cross-section of funds produces some extreme α estimates simply by chance, we cannot identify the specific managers that are skillful rather than lucky.

For fund investors the results from the simulations (Section III) are disheartening. The tests say clearly that the cross-section of precision-adjusted α estimates, $t(\alpha)$, for actual net fund returns to investors is dominated by funds that do not have information that produces expected returns sufficient to cover the costs the funds impose on investors. This result holds even for the extreme right tail of $t(\alpha)$ estimates for actual fund returns. Thus, if there are funds with sufficient information to cover costs, they are hidden among the mass of funds with insufficient information.

Mutual funds look better when returns are measured gross, that is, before fees and expenses. Comparing the cross-section of $t(\alpha)$ estimates from gross fund returns to the average cross-section of $t(\alpha)$ estimates from the simulation runs suggests that there are managers with bad information that lowers expected returns and there are managers with good information that enhances expected returns. Formal support for this inference is, however, lacking. Large fractions of the simulation runs (in which true α is zero) produce left and right tails for $t(\alpha)$ more extreme than the tails observed for actual gross fund returns. In the end, gross returns produce hints of information effects, but only hints. Formally, we cannot reject the hypothesis that the cross-section of α estimates from gross fund returns comes from a population of managers with no special stock selection skills, good or bad.

Persistence tests that examine future α for the winner and loser deciles from sorts of funds on past α (Section II) seem similar to simulation tests that examine the 10^{th} and 90^{th} percentiles of α estimates (Section III). There is, however, an important difference. If some funds consistently have private information,[2] the simulations may have more power to infer their existence because the tests use longer periods to estimate α. The weak performance evidence from the simulations then suggests that few if any funds have consistent access to information. On the other hand, if access to information is temporary, it can be masked in a long return history. Persistence tests (which sort funds on recent performance) might then be better for identifying when funds have information. Similar comments apply when funds by chance temporarily hold stocks that expose failures of the three- or four-factor models. Confirming Bollen and Busse (2004), our more extensive results suggest that any post-ranking performance identified in the persistence tests is indeed temporary.

Finally, the paper closest to ours is Kosowski et al (2006). They do persistence tests and bootstrap simulations that seem to produce stronger inferences about positive information effects in mutual fund returns. We discuss their tests after presenting our results.

I. Average Returns for EW and VW Portfolios of U.S. Equity Mutual Funds

Our mutual fund sample is from the CRSP (Center for Research in Security Prices) database. We include only funds that invest primarily in U.S. common stocks, and we combine, with value weights, different classes of the same fund into a single fund. (See French 2008 for details.) The data cover January 1962 to September 2006 (henceforth 1962-2006), but our central persistence and simulation tests use the period after 1983, when there are fewer data issues (Elton, Gruber, and Blake 2001). Our benchmarks for evaluating fund performance are the three-factor model of Fama and French (1993), and Carhart's (1997) four-factor model. To measure performance, these models use two variants of the time-series regression,

$$R_{it} - R_{ft} = a_i + b_i(R_{Mt} - R_{ft}) + s_iSMB_t + v_iVMG_t + m_iMOM_t + e_{it}. \quad (1)$$

[2] We define private information as any information (including publicly available information) that is not reflected in prices.

In this regression, R_{it} is the return on fund i for month t, R_{ft} is the riskfree rate (the one-month U.S. Treasury bill rate), R_{Mt} is the market return (the return on a value-weight portfolio of NYSE, Amex, and NASDAQ stocks), SMB_t and VMG_t are the size and value-growth returns of Fama and French (1993), MOM_t is our version of Carhart's (1997) momentum return, a_i is the average return left unexplained by the benchmark model (the estimate of α_i), and e_{it} is the regression residual. The full version of (1) is Carhart's four-factor model, and the regression without MOM_t is the Fama-French three-factor model. The construction of SMB_t and VMG_t (also known as HML_t) follows Fama and French (1993). The momentum return, MOM_t, is defined like VMG_t, except that we sort on prior return rather than the book-to-market equity ratio and the momentum sort is refreshed monthly rather than annually. (See Table 1 for details.)

Table 1 shows summary statistics for the explanatory returns in (1) for periods examined later in tests of mutual fund performance. The average value of the value-growth return, VMG_t, is large for 1962-2006 (0.48% per month, t = 3.82) and for each subperiod. The average value of the market premium, $R_{Mt} – R_{ft}$, for 1962-2006 (0.45% per month, t = 2.36) is close to the average VMG_t return. The size return, SMB_t, has the smallest average value for 1962-2006 (0.23% per month, t = 1.65). For every period in Table 1, the average momentum return is by far the largest factor premium (for example, 0.82% per month, t = 4.80, for 1962-2006). Large average MOM_t returns imply that the MOM_t slopes in (1) can have a big effect on intercept (α) estimates – a comment of some import for the persistence tests presented later.

There is controversy about whether the average SMB_t, VMG_t, and MOM_t returns are rewards for risk or the result of mispricing. For our purposes, there is no need to take a stance on this issue. On an intuitive level, we can interpret SMB_t, VMG_t, and MOM_t as common factors in stock returns associated with patterns in average returns during our sample period. Abstracting from the variation in returns associated with SMB_t, VMG_t, and MOM_t then allows us to focus better on the effects of information about individual stocks (stock picking ability), which should show up in three- and four-factor intercepts.

Formally, our null hypothesis in the tests that use the three-factor model is that U.S equity mutual funds have no private information, and the riskfree security, the market, SMB, and VMG together span the ex ante mean-variance-efficient (MVE) portfolios that can be constructed from the portfolios of these funds. This implies that, measured before costs, the expected values of the three-factor intercepts for these funds are zero (Huberman and Kandel 1987). The alternative hypothesis is that some mutual funds have private information about individual stocks, so the riskfree security, the market, SMB, and VMG do not span the MVE portfolios that can be constructed from the portfolios of these funds. This implies that the expected values of the intercepts for these funds are positive or negative, depending on whether private information is good or bad. Similar statements apply when we use the four-factor model to evaluate performance.

Table 2 shows estimates of regression (1) for equal-weight (EW) and value-weight (VW) portfolios of the funds in our sample. The intercepts in (1) for EW fund returns tell us whether funds on average have information about stocks that allows them to produce expected returns different from those implied by their exposures to common factors in returns. In contrast, VW returns tell us about the fate of the average dollar (or aggregate wealth) invested in funds. Part A of Table 2 shows the regression intercepts (α estimates) for the three- and four-factor variants of (1) for returns measured gross and net of fund fees and expenses. Part B shows the regression slopes for the four-factor model. The market, SMB, and VMG slopes for the three-factor model are close to the slopes in the four-factor model. Only the slopes for net returns are shown. The slopes for gross returns are the same up to three decimal places.[3]

All the market slopes in Table 2 are close to 1.0, which is not surprising since our sample is restricted to funds that invest primarily in U.S. equities. The slopes on SMB_t are around 0.20 for EW fund returns and around 0.06 for VW returns. We can infer that smaller funds are more likely to invest in smaller stocks, but total dollars invested in funds (captured by VW returns) show little tilt toward smaller stocks. More interesting, the tilt toward smaller stocks in EW and VW returns, as captured by SMB_t slopes, is a bit lower after 1983. Thus, Banz' (1981) discovery of the size effect (higher average returns of small stocks) does not produce a shift in either funds or dollars invested in funds toward smaller stocks.

The slopes for VMG_t in fund returns for 1962-1992 are -0.10 (EW and VW returns), which suggests a slight tilt toward growth stocks. The VMG_t slopes in fund returns for 1993-2006 rise to 0.05 (EW) and 0.00

[3] Information about fees and expenses is sometimes missing on CRSP, especially early in the 1962-2006 period. When a fund is missing fees and expenses for a year, we assume its expense ratio is the same as other funds with similar AUM, with separate estimates for active and passive funds.

(VW), which suggests a slight move away from growth stocks toward a more neutral position. Again, however, academic trumpeting of the higher returns of value stocks relative to growth stocks (Fama and French 1992, Lakonishok, Shleifer, and Vishny 1994) does not result in much aggregate movement of funds or dollars invested in funds toward value stocks. Finally, the discovery of return momentum (Jegadeesh and Titman 1993) produces little aggregate movement of funds or dollars invested in funds toward positive momentum stocks. In fact, the momentum slopes for EW and VW fund returns move from slightly positive for 1962-1992 to values closer to zero for 1993-2006.

The intercepts in the estimates of (1) summarize the average performance of funds (EW returns) and the performance of aggregate wealth invested in funds (VW returns). In terms of net returns to investors, performance is poor. The three-factor intercepts for EW and VW net returns are negative for all the periods in Table 2, but the estimates for 1962-1983 and 1962-1992 are close to zero statistically and economically. The EW and VW fund returns of 1962-1983 and 1962-1992 have slightly positive exposure to momentum, however, and controlling for momentum leads to strongly negative four-factor intercepts for these periods, like the negative estimates for 1984-2006 and 1993-2006. The annualized four-factor intercepts for different periods range from -0.71% to -1.38%, and they are -1.75 to -3.68 standard errors from zero. These results are in line with previous work (for example, Jensen 1968, Malkiel 1995, Gruber 1996).

The intercepts in (1) for net returns tell us whether funds have sufficient private information to cover the costs they impose on investors. Gross returns are better for testing whether funds have any private information. For EW and VW gross fund returns, three-factor intercepts are positive for all periods except 1993-2006, but only the EW estimate for 1962-1992 (0.97% per year, $t = 2.14$) is more than 1.61 standard errors above zero. Again, however, the positive three-factor α estimates for periods before 1993 seem to be due to positive momentum exposure. For VW gross returns, four-factor intercepts for all periods in Table 2 are negative but close to zero. For EW gross returns, four-factor intercepts are randomly positive and negative, but again always close to zero.

We can offer an equilibrium accounting perspective on the results in Table 2. During 1984-2006, when there are fewer concerns about biases in the CRSP data, the annualized three- and four-factor α estimates for VW net fund returns are -0.80% and -0.93% (t = -2.62 and -2.98). Thus, for total wealth invested in funds, any benefits from active management are overwhelmed by fees and expenses. When we add back fees and expenses, there is no evidence that total wealth invested in funds gets benefits or suffers losses from active management. The annualized three- and four-factor α estimates for VW gross fund returns for 1984-2006 are close to zero, 0.10% and -0.03% (t = 0.32 and -0.09). VW fund returns also show little exposure to the size, value, and momentum factors during 1984-2006, and we can report that during this period the excess market return alone explains 99% of the variance of the monthly VW excess return for our universe of funds. Together these facts say that during 1984-2006, mutual fund investors in aggregate hold a portfolio that, before fees and expenses, mimics market portfolio returns. The aggregate portfolio of funds is, however, dominated by active funds, and the average return to investors is reduced by the high fees and expenses of these funds. As noted earlier, these results echo equilibrium accounting, but for a subset of investment managers where the implications of equilibrium accounting for aggregate investor returns need not hold.

Finally, our net returns ignore load fees, so they overstate the returns to investors in load funds. On the other hand, our gross returns are before fees and expenses, but they are net of trading costs. We do not attempt to add trading costs to gross returns because turnover is often missing on CRSP, and even when turnover data are available, differences across funds in investment styles and trading strategies make estimates of trading costs imprecise. We prefer to argue that (unlike management fees) trading costs are inherent in a fund's strategy, and actively managed funds should at a minimum provide expected returns that cover trading costs. Four-factor α estimates that are close to zero for EW and VW gross returns (Table 2) suggest that, at least on average, funds may achieve this "bare minimum" result.

II. Return Persistence

Table 2 says that on average mutual funds do not have information that produces gross returns above (or below) those predicted by the three- and four-factor benchmarks. But this result may just mean that funds with good information that allows them to outperform the benchmarks are balanced by funds with bad information that leads to underperformance. Persistence tests are one way to examine this possibility. If some managers have access to good or bad information, when we sort funds on past performance we should find that past winners are on average future winners and past losers are future losers.

We sort funds in several ways. Specifically, each month of 1984-2006 we allocate funds to deciles based on the t-statistic for the average monthly net-of-market return of the preceding 12 or 60 months or the t-statistic for the three-factor α estimate from 60 months of past returns. We sort on t-statistics since the precision of the sort variable should be a factor in identifying return persistence. To focus on stock picking ability, the persistence tests we show use gross returns, but we also comment on results for net returns. Finally, we can report that (i) sorts on net-of-market average returns and α estimates produce results similar to sorts on their t-statistics, (ii) sorts on 36-month past performance produce results similar to the 60-month sorts, and (iii) sorts on four-factor t(α) estimates produce results like the sorts on three-factor t(α).

If stock-picking talent is persistent, its tracks in future returns should be most apparent in the extremes of sorts on past performance. Thus, to enhance the power of the tests we focus on t(α) estimates for the post-ranking returns of the deciles with the best and worst past performance. We can report that when there is evidence of persistence, it decays quickly for less extreme deciles of the sorts. And when the extremes show no evidence of persistence, less extreme deciles support this inference.

We examine α estimates for monthly returns for each of the first three months, the second, third, and fourth quarters, and the first, second, and third years after portfolio formation. We use a calendar-based approach to compute the returns for multi-month intervals. We first compute the EW and VW single-month returns in month t for portfolios formed each of the preceding 36 months. Each EW single-month return for month t equally weights the funds assigned to a portfolio in month t-k. The VW return for month t weights each fund by its assets under management when the portfolio is formed at the end of month t-k times its compounded (gross or net) return for the next k-1 months, AUM(t-k) * [1 + R(t-k+1)] * … * [1 + R(t-1)]. The EW or VW month t return for a multi-month interval is the EW average of the month t returns for the portfolios formed during the interval. For example, the 1-to-12-month return in month t is an EW average of the (EW or VW) month t returns for the portfolios formed at the end of each of the preceding 12 months.

If fund managers consistently have private information, we do not expect much change in α estimates in the periods after portfolio formation. If there is evidence of performance but it decays toward zero, we infer either that (i) managers have only temporary access to information, or (ii) benchmark problems produce spurious but temporary α estimates, because, for example, funds by chance temporarily hold stocks whose expected returns do not conform well to the benchmark model.

We limit the persistence tests to 1984-2006 and to funds that reach the equivalent of five million 2006 dollars in assets under management (AUM) sometime before portfolio formation. For example, since the AUM minimum is in 2006 dollars, portfolios formed during 1984 include only funds that have reached about \$2.5 million in AUM sometime before portfolio formation. Once a fund passes the AUM minimum, it is included in all subsequent tests, so this requirement does not create survivor bias.

During 1962-1983 about 15% of the funds on CRSP report only annual returns, and the average annual EW return for these funds is 5.29% lower than for funds that report monthly returns. As a result, the EW average return on all funds is a nontrivial 0.65% per year lower than the EW return of funds that report monthly returns. Thus, during 1962-1983 there is a survival bias in tests (like our persistence tests) that use only funds that report monthly returns. (The problem is minor in VW returns because funds that report annual returns tend to be small.) After 1983 almost all funds report monthly returns. (See Elton, Gruber and Blake 2001 for a discussion of CRSP data problems for the period before 1984.)

There is another survival bias that leads us to drop funds that have not reached \$5 million AUM. Fund management companies commonly provide seed money to new funds to develop a return history. Funds are then opened to the public when their return histories turn out to be attractive. We have no reliable way to identify a fund's returns that are subject to this selection bias. The \$5 million AUM bound for admission to the tests alleviates the problem since AUM is likely to be low during the pre-release period.

A. Persistence Results for 12-Month Return Sorts

Sorts on 12-month average return seem unattractive since an average for 12 months is an imprecise estimate of expected return. Part A of Table 3 shows, however, that when funds are sorted on the t-statistic of the average net-of-market gross return of the last 12 months, post-sort three-factor α is reliably positive for the gross returns of the extreme decile of past winners and reliably negative for extreme losers, but only for six to nine months after portfolio formation. Much of this short-run persistence is likely due to momentum. As Carhart (1997) notes, the funds that perform best during the last 12 months are likely to hold more of the best performing stocks of that period, and the funds that have done poorly are likely to hold more of the stocks that have done poorly. Momentum (Jegadeesh and Titman 1993) then implies that the strong returns of recent winner funds and the poor returns of losers persist for a few months after portfolio formation.

Carhart (1997) suggests that the momentum exposures of winner and loser funds are inadvertent and temporary. There is a simple test. If momentum exposure is part of a long-term investment strategy, we do not expect much change in four-factor MOM_t slopes in the months after portfolio formation. But if momentum exposure is inadvertent and temporary, we expect that MOM_t slopes move toward zero after portfolio formation. The momentum slopes in Part A of Table 3 support the temporary exposure story. All the MOM_t slopes for the top decile of funds are positive in the first year after funds are sorted on past return, and all but one for the bottom decile are negative. But the MOM_t slopes decay to values close to zero within a year, and actually switch sign from months 1-12 to months 13-24 and 25-36.

Whether or not momentum tilts are intentional, controlling for momentum exposure with the four-factor model shrinks the α estimates for the first year of post-sort returns. Four-factor α produces no evidence of persistence in the post-sort EW or VW gross returns of the worst performing funds. The α estimates for EW loser returns for post-sort months 1, 2, and 3 are -0.01%, 0.02%, and -0.01% per month, with tiny t-statistics. The four-factor model also absorbs the persistence in the VW post-sort returns of the best performing funds, but EW returns suggest that for smaller funds some temporary persistence remains. Although they decay quickly, the four-factor α estimates for the EW returns of last year's winners are 0.29%, (t = 3.36), 0.14% (t = 1.76), and 0.12% (t = 1.58) for post-sort months 1, 2, and 3, and 0.14% per month (t = 2.01) for months 4-6.

B. Persistence Results for 60-Month Return Sorts

If some funds have consistent access to information, ranking funds on a longer period of past returns should produce stronger evidence of persistence since average returns for longer periods are more precise estimates of expected return. Part B of Table 3 shows, however, that sorts of funds on the t-statistic for the average net-of-market gross return of the last 60 months fail to produce stronger evidence of persistence than the 12-month sorts. Three-factor α estimates for post-ranking gross returns are almost always closer to zero in the 60-month sorts. The four-factor regressions explain why. Most of the persistence in three-factor α from sorts on 12-month return is due to transitory momentum exposure. The 60-month sorts also produce some post-sort momentum exposure, which again decays quickly. But initial post-sort exposure to momentum is weaker in the 60-month sorts. This is not surprising. Because momentum exposure is short-term, 12-month return sorts produce stronger exposures than 60-month sorts. All this is consistent with our inference that the momentum exposures of winner and loser funds are chance results rather than conscious strategy.

In the sorts on 60-month average return, the four-factor post-ranking α estimates for EW gross returns for 1984-2006 are positive for the winner decile and negative for the loser decile, but the α estimates are close to zero. For the winners, the four-factor α for the year after portfolio formation is 0.05% per month (t = 0.86), which is a faint hint that, after controlling for momentum, extreme winners have good information that enhances expected returns. For the EW loser decile, the four-factor α for the year after portfolio formation is -0.10% per month (t = -1.50), a hint that extreme losers have bad information that lowers expected returns. But even these weak hints apply primarily to smaller funds. Four-factor α for the year after portfolio formation is closer to zero in VW returns, 0.02% per month for winners and -0.05% for losers.

All hints of persistence in the post-ranking momentum-adjusted gross returns of the extreme winners and losers of the last 60 months seem to be special to the early years of 1984-2006. Part B of Table 3 shows that in EW and VW gross returns for 1993-2006, four-factor post-ranking α estimates for winners and losers are economically and statistically close to zero (for example, 0.00% per month, t = 0.05, for EW losers, and -0.03%, t = -0.36, for EW winners for the year after portfolio formation). Thus, if extreme winners and losers once had a bit of private information about stocks, access seems to have disappeared.

C. Sorts on 60-Month Three-Factor α Estimates

There is a case for sorting funds on α estimates, since (i) they are more precise than average returns and (ii) sorts on α estimates may focus better on the effects of information about individual stocks. But α sorts have a downside. They may in part focus on the benchmark model's problems in explaining the average returns on some stocks (Carhart 1997). Part C of Table 3 shows results for sorts of funds on t-statistics for three-factor α estimates from 60 months of past gross returns.

In the t(α) sorts for 1984-2006, post-ranking three-factor α estimates for the winner and loser deciles are typically more extreme than those from the sorts on 60 months of net-of-market returns. For example, the α for the EW decile of winners is 0.19% per month (t = 3.17) for the first year after funds are sorted on t(α) (Table 3 Part C), and only 0.09% (t = 1.40) in sorts on net-of-market average return (Table 3, Part B). Controlling for momentum with the four-factor model weakens the evidence of persistence a bit in the first year after portfolios are formed, but four-factor α estimates for EW returns remain reliably different from zero out to a year (longer for losers). Similar results are observed in unreported sorts on four-factor t(α).

The stronger evidence of persistence when we sort on three-factor t(α) may indicate that past winner and loser funds have information about individual stocks that is uncovered better by these sorts than by sorts on average return. But it is also possible that α sorts just home in better on shortcomings of the three-factor model. This explanation seems attractive given that the persistence in the three-factor t(α) sorts for 1984-2006 is largely due to smaller funds, which are likely to be less diversified. When we value-weight funds, the strong return persistence observed in post-ranking α estimates on EW loser returns disappears almost entirely, and persistence is also much weaker in VW winner returns. In any case, for current investment purposes, the issue may be moot. In tests for the more recent 1993-2006 period (Part C of Table 3), the three-factor post-ranking α estimates from sorts on three-factor t(α) are closer to zero, and when we add a control for momentum, the α estimates are indistinguishable from zero for EW and VW fund returns. Thus, the return persistence identified in the three-factor t(α) sorts for 1984-2006 is rather special to the early years of the period.

E. Synopsis

When we sort funds on t-statistics for market-adjusted average gross returns, there is persistence in returns, but it seems to be due to temporary and thus probably inadvertent momentum exposure. Controlling for momentum with the four-factor model eliminates most if not all persistence. Thus, sorts on past average gross returns suggest that funds do not have access to private information (good or bad) about stocks.

Average returns estimate expected returns without imposing a model for expected returns. In contrast, sorts on t(α) estimates from a benchmark model produce more precise estimates of expected returns left unexplained by the model, but they are more exposed to bad model problems. Our sorts on three-factor t(α) estimates from gross returns produce some evidence that, at least in the extremes, either (i) mutual fund managers have access to private information (bad information for losers and good for winners) or (ii) the sorts are subject to a bad model problem. The evidence of persistence is, however, special to the early years of 1984-2006 and to smaller funds. In the results for 1984-2006, sorts on three-factor (or four-factor) t(α) estimates produce little evidence of persistence in VW fund returns, and in the results for 1993-2006, t(α) sorts produce little evidence of persistence in EW or VW fund returns.

Gross returns (before fees and expenses) are more relevant than net returns for judging whether funds have any access to information. Net returns are better for judging whether investors get what they pay for, that is, whether managers produce expected returns that cover the costs they impose on investors. We have replicated the persistence tests on net returns. Skipping the details (available on request), we can report, not surprisingly, that net returns produce systematically lower post-sort α estimates than gross returns. When the four-factor model is used to evaluate performance, the persistence tests that use net returns produce no reliable evidence of funds with more than sufficient information to cover costs.

There is, however, sometimes a power problem that leads to ambiguity about whether on average funds have sufficient private information to cover costs. Specifically, in the 1993-2006 sorts on three-factor t(α) estimates on 60 months of past net returns, post-sort four-factor α estimates for the year after portfolio formation are -0.10% per month (t = -1.42) for the VW losers and -0.08% (t = -0.98) for the VW winners, both quite close to average VW fees and expenses (about 0.08% per month). Thus, in VW net returns for 1993-2006, winners and losers underperform by about the amount of fees and expenses, which suggests that they have no access to private information. But average fees and expenses are similar to the standard errors of the α estimates. As a result, the four-factor α estimates for the VW returns of 1993-2006 also do not allow us to reject the hypothesis that winners and losers have sufficient information to cover costs. For the winners this inference problem arises in the net return results for 1984-2006 and 1993-2006, whether we use EW or VW returns and whether we sort on t-statistics for market adjusted average returns or three-factor t(α) estimates. For the losers, the inference problem is specific to VW returns for 1993-2006 and t(α) sorts. Otherwise, we can always reject the hypothesis that the losers have sufficient information to cover costs.

If the VW net returns in the persistence sorts on t(α) for 1993-2006 lead us to entertain the hypothesis that during this period funds in general have sufficient information to cover costs, the more powerful tests on overall fund returns in Table 2 are more relevant. For 1993-2006 the EW and VW portfolios of all funds produce four-factor t(α) estimates for gross returns that are negative but within 1.21 standard errors of zero. In net returns the α estimates are lower by the amount of average fees and expenses, and they are more than 2.38 standard errors below zero. We can comfortably infer that during 1993-2006, indeed during all periods examined in Table 2, funds in aggregate do not have sufficient information to cover costs.

Kosowski et al (2006) draw stronger positive conclusions about persistence in mutual fund returns. They focus on EW net returns for a period beginning in 1975 and on four-factor post-ranking α estimates from sorts on four-factor α. Sorts on three-factor t(α) produce the strongest results in our reported tests for 1984-2006, and our unreported results from sorts on four-factor t(α) are similar. Our evidence suggests that the inferences of Kosowski et al (2006) are special to time period, EW returns, and sorts on α.

III. Bootstrap Simulations

Persistence tests rank winners and losers based on short-term performance. We turn now to a simulation approach that uses longer return histories to infer the existence of superior and inferior managers.

A. Setup

The period for the simulations is 1984-2006, and we include funds only after they pass $5 million AUM. For perspective on the performance of funds of different sizes, we also show results for funds after they pass $250 million and $1 billion. Since we estimate benchmark regressions for each fund, we limit the tests to funds that have at least eight months of returns after they pass an AUM bound, so there is a bit of survival bias. To avoid having lots of new funds with short return histories, we only use funds that appear on CRSP at least five years before the end of our sample period.

The population data for the simulations are monthly benchmark-adjusted fund returns. Specifically, a fund's returns are measured net of its α estimate for the part of 1984-2006 after the fund passes an AUM bound. For example, to compute three-factor-adjusted gross returns for a fund in the $5 million group, we estimate its three-factor α using its monthly gross returns for the part of 1984-2006 the fund is in the $5 million group. We then subtract the estimated α from the fund's monthly returns. Thus, the fund's three-factor α is zero in the population of three-factor-adjusted gross returns for the $5 million sample used in the simulations.

We calculate benchmark-adjusted returns for the three-factor and four-factor models, for gross and net returns, and for the three AUM bounds. The result is 12 sets of benchmark-adjusted returns (gross and net for two benchmarks and three AUM bounds). For each set, a fund's α is zero in the population of adjusted gross or net returns for that combination of benchmark model and AUM bound.

A simulation run is a random sample with replacement of the calendar months of 1984-2006. A simulation sample has 273 months (like January 1984 to September 2006). For each of the 12 sets of benchmark-adjusted returns, we estimate, fund by fund, the relevant (three-factor or four-factor) benchmark model on the simulation draw of months of adjusted returns for that benchmark, dropping funds that are in the simulation run for less than eight months. Each run thus produces 12 cross-sections of α estimates using the same random sample of months from 12 populations of adjusted fund returns that have the properties of actual fund returns, except that in each population of adjusted returns, the assumed benchmark model holds exactly (α is zero) for every fund.

We do 10,000 simulation runs to produce 12 distributions of t-statistics, t(α), for a world in which true α is zero. We focus on t(α), rather than raw estimates of α, to control for differences in precision due to differences in residual variance and in the number of months funds are in a simulation run.

A prime advantage of our simulation approach is that it mimics the joint distribution of fund returns. It thus captures all effects of the cross-correlation of fund returns on the distribution of t(α) estimates for funds. Because we jointly sample fund returns and explanatory returns we also capture all effects of (for example) correlated heteroscedasticity of the explanatory returns and disturbances of a benchmark model. Note, however, that except for funds that are in our tests for the entire 1984-2006 period, a fund is likely to show up in a simulation run for more or less than the number of months of 1984-2006 it is on CRSP. This is not serious since we focus on t(α) estimates and the distribution of t(α) depends on the number of months a fund is in a simulation run only through the degrees of freedom effect.

Note also that the information effects built into the simulations are different for gross and net returns. For gross returns, setting true α equal to zero simulates a world where no fund has private information. In contrast, setting true α equal to zero for net returns simulates a world where all funds have information sufficient to generate expected returns that cover the costs imposed on investors.

To develop perspective on the simulations, we first compare, in qualitative terms, the percentiles of the cross-section of t(α) estimates from actual fund returns and the average values from the simulations. We then turn to formal inferences about information effects in the tails of the cross-section of t(α).

B. First Impressions

Net Returns – When we estimate a benchmark model on the actual returns of each fund in an AUM group, we get a cross-section of t(α) estimates that can be ordered into a cumulative distribution function (CDF) of t(α) estimates for actual fund returns. A simulation run for the same combination of benchmark model and AUM group also produces a cross-section of t(α) estimates and its CDF for a world in which true α is zero. In our initial examination of the simulations we compare (i) the values of t(α) at selected percentiles of the CDF of t(α) estimates from actual fund returns and (ii) the averages across the 10,000 simulation runs of the t(α) estimates at the same percentiles. For example, the first percentile of three-factor t(α) estimates for the net returns of funds in the $5 million AUM group is -3.92, versus an average first percentile of -2.93 from the 10,000 three-factor simulation runs for the net returns of funds in this group (Table 4).

For each combination of gross or net returns, AUM group, and benchmark model, Table 4 shows the CDFs of t(α) estimates for actual returns and the average of the 10,000 simulation CDFs. The average simulation CDFs are quite similar for gross and net returns and for the two benchmark models. This is not surprising given that true α is always zero in the simulations. Note, however, that the dispersion of the average simulation CDFs decreases a bit from lower to higher AUM groups. This is at least in part a degrees of freedom effect. All $1 billion funds are also $5 million and $250 million funds, but lots of funds in lower

AUM groups die without reaching higher AUM groups. As a result, smaller funds have shorter average sample periods. Differences in the cross-correlations of fund returns may also play a role in the lower dispersion of the average simulation CDFs of the larger AUM groups.

The hypothesis that mutual funds have sufficient information to generate expected returns that cover costs fares poorly in Table 4. For every combination of benchmark model and AUM group, the CDF of the cross-section of $t(\alpha)$ estimates from actual net fund returns is entirely to the left of the average CDF from the simulations. In other words, the average percentile values of $t(\alpha)$ from the simulations of net fund returns (in which, by construction, funds have sufficient information to generate expected returns that cover costs) always beat (are larger than) the corresponding percentile values of $t(\alpha)$ for actual net fund returns. This evidence does not rule out the existence of some funds with information sufficient to cover costs, but it is a strong hint that formal tests (discussed later) are unlikely to produce this positive inference.

Gross Returns – It is possible that the fruits of information do not show up in net fund returns because they are absorbed by fees and expenses. Gross returns provide more direct evidence on whether fund managers have any information that is not reflected in prices.

Adding back fees and expenses (inevitably) pushes $t(\alpha)$ for actual fund returns toward higher values. But Table 4 shows that for all three AUM groups, the left tail of the CDF of three-factor $t(\alpha)$ estimates for actual gross returns is still to the left of the average from the simulations. For example, the simulations say that in the absence of private information, on average the fifth percentile of $t(\alpha)$ for gross returns for the \$5 million group is -1.85, but the actual fifth percentile is lower, -2.18. Thus, the left tails of the CDFs of $t(\alpha)$ suggest that there are misinformed active managers with poor information.

Conversely, the right tails of the CDFs of three-factor $t(\alpha)$ hint at the existence of active managers with good information. For the \$5 million AUM group, the CDF of $t(\alpha)$ estimates for actual gross fund returns moves to the right of the average from the simulations between the 60^{th} and 70^{th} percentiles. For example, on average the 95^{th} percentile of $t(\alpha)$ estimates for funds in the \$5 million group is 1.88 in the simulations, but the actual 95^{th} percentile is slightly higher, 1.98. For the two larger AUM groups the crossovers occur at higher percentiles, around the 80^{th} for the \$250 million group and the 95^{th} for the \$1 billion group.

In short, comparing the CDFs of three-factor t(α) estimates for actual gross fund returns with the average CDFs from the simulations suggests the existence of both the informed and the misinformed active investors of Fama and French (2007). The underperformance in the left tail of t(α) is, however, typically more extreme than the overperformance in the right tail. For example, the 5th percentile of the cross-section of three-factor t(α) estimates for the actual gross returns of funds in the $5 million group is 0.33 standard errors below the average from the simulations, but the 95th percentile for actual fund returns is only 0.10 standard errors above the simulation average.

The four-factor results for gross returns in Table 4 are similar to the three-factor results, with an interesting nuance. Adding a control for momentum exposure tends to shrink slightly the left and right tails of the cross-sections of t(α) for actual fund returns. This is consistent with what we saw in the persistence tests; that is, funds with negative three-factor α tend to have negative MOM_t exposure and those with positive three-factor α tend to have positive exposure. Controlling for momentum pulls the α estimates toward zero. This shrinkage effect is small but at least at the at the 99th percentile of t(α) estimates it suffices to kill the systematic advantage of the t(α) estimates for actual fund returns over the averages from the simulations.

Finally, the simulation distributions of t(α) are fat-tailed. The average simulation distribution of t(α) for the $5 million group (our full sample) is like a t distribution with eight degrees of freedom. Since every α estimate uses at least eight observations and most use many more, we can conclude that the simulation distributions of t(α) are more fat-tailed than can be explained by degrees of freedom. This suggests that properties of the joint distribution of fund returns and of fund and factor returns have important effects on the cross-section of α estimates – a comment of some import in our later discussion of Kosowski et al (2006).

C. Formal Tests

Comparing the percentiles of t(α) estimates for actual fund returns with the simulation averages gives hints about where there may be information effects. Table 4 also provides likelihood statements, in particular, the fractions of the 10,000 simulation runs that produce lower values of t(α) at selected percentiles than actual fund returns. These likelihoods allow us to judge more formally whether the tails of the cross-section of t(α) estimates for actual fund returns are extreme relative to what we observe when true α is zero.

Specifically, we infer that some funds lack information sufficient to cover costs if low fractions of the simulation runs produce left tail percentiles of t(α) below those from actual net fund returns; or equivalently, if large fractions of the simulation runs beat the left tail t(α) estimates from actual net fund returns. Similarly, we infer that some funds produce expected returns sufficient to cover costs if large fractions of the simulation runs produce right tail percentile t(α) estimates below those from actual net fund returns; that is, if low fractions of the simulation runs beat the right tail t(α) estimates from actual net fund returns.

The logic is the same for the simulation tests for gross returns, but the inferences center on whether there are funds that have access to any information. We infer that some funds act on bad information if large fractions of the simulation runs beat the left tail t(α) estimates from actual net fund returns. And we infer that there are funds with good information that enhances expected returns if small fractions of the simulation runs beat the right tail percentile t(α) estimates from actual gross fund returns.

Net Returns – Table 4 provides strong evidence that most funds lack sufficient information to cover costs. For every left tail percentile and all three AUM groups, less than 1.0% of the net return three-factor simulation runs (in which expected returns cover costs) produce three-factor t(α) estimates below the values observed for actual net fund returns. For example, the 10^{th} percentile of the cross-section of t(α) estimates from the net returns of \$5 million funds is -2.31, but only 0.11% of the simulation runs for this group have 10^{th} percentile t(α) estimates below -2.31. Clearly, the left tails of the cross-sections of three-factor t(α) estimates for actual net fund returns are dominated by funds with insufficient information to cover costs.

More interesting, the right tails of three-factor t(α) estimates for net returns also offer no evidence that there are funds with sufficient information to cover costs. The \$1 billion AUM group produces the strongest positive results, but even for these funds at most 24.41% of the simulation runs produce right tail t(α) estimates below the same percentile of the cross-section of t(α) for actual net fund returns. In other words, at every right tail percentile of t(α) more than 75% of the simulation runs beat the estimate of t(α) from actual net fund returns. The right tail performance of the \$5 million and \$250 million groups is weaker with typically less than 10% of the simulation runs producing percentiles values of t(α) below those from actual fund returns. For example, for the \$5 million group, the 95^{th} percentile t(α) estimate from the simulations is below the 95^{th}

percentile of actual net return t(α) estimates in only 5.2% of the simulations, so 94.8% of the 95th percentile t(α) estimates from the simulations beat the 95th percentile from actual net fund returns.

Switching from the three-factor to the four-factor model to estimate α confirms our inference that the left tail of t(α) estimates is dominated by funds with net returns that do not cover costs. If anything, the absence of right tail evidence that some funds have enough information to cover costs is stronger for the four-factor model; that is, the fractions of the simulation runs that beat right tail estimates of t(α) for actual net fund returns tend to be higher in tests that use the four-factor model, typically more than 90% for the $250 million and $1billion AUM groups and always more than 95% for the $5 million group (our full sample).

The bottom line from the simulation tests for net returns is that if there are funds with enough information to cover costs, they are buried in the noise of funds with bad or insufficient good information.

Gross Returns – The simulations for gross returns in Table 4 allow inferences about the existence of funds with any information that affects expected returns. The $1 billion AUM group presents the strongest formal case for our earlier informal inference that some funds seem to trade on bad information. Below the 40th percentile, the left tail percentiles of the cross-section of three-factor t(α) estimates for the actual gross returns of these funds are above the simulation values in less than 10% of simulation runs. For the $5 million and $250 million groups, the actual three-factor t(α) estimates for percentiles up to the 20th exceed the corresponding simulation values in less than 20% of the three-factor simulation runs.

For percentiles below the 20th, the fractions of the simulation runs that produce three-factor t(α) estimates below those observed for actual gross fund returns are almost always greater than 5%. Thus, with a 5% threshold, the evidence in Table 4 does not reject the hypothesis that the extreme negative values of three-factor t(α) for actual gross fund returns are just bad luck. Even with a 10% threshold, we cannot infer that bad information contributes to the left tail of three-factor t(α) estimates from actual gross fund returns for the $5 million AUM group (the full sample) and the $250 group. Moreover, Table 4 shows that switching from the three-factor to the four-factor model to estimate α almost always increases the fraction of simulation runs in which left tail percentiles of t(α) for actual fund returns beat the estimates from the simulations. Controlling for momentum exposure thus produces higher likelihoods that the left tail of the cross-section of t(α) estimates for actual gross fund returns is due to bad luck, not bad information.

The absence of strong evidence that bad information contributes to the returns of poorly performing funds is good news for fund managers. The bad news is that the right tails of t(α) estimates for gross fund returns provide little evidence that there are funds with good information. The 95th and higher percentiles of the cross-section of three-factor t(α) estimates for actual gross fund returns are above the average values from simulations – but not so far above as to be unlikely for a world where funds have no private information. For the full ($5 million) sample, the 95th and higher percentiles of the cross-section of three-factor t(α) estimates from actual fund returns are above the same simulation percentiles in 54.79% to 76.03% of the simulation runs, which means 23.97% to 45.21% of the simulation runs beat the t(α) estimates from actual fund returns. The four-factor model uncovers even less evidence of good information. In the four-factor tests for the $5 million group, 41.31% to 71.02% of the simulations produce t(α) estimates for the 95th and higher percentiles below those for actual gross returns, so 28.98% to 58.69% of the simulations beat the extreme right tail t(α) estimates from actual returns. The results for $250 million and $1 billion funds are similar.

In the end, the simulation evidence that there are funds with good information that enhances expected gross returns is weaker than the weak evidence that there are funds with bad information. And our inability to identify information effects does not seem to be due to lack of power (an issue we intend to pursue in detail in a future paper). For example, the 95th percentile of gross return four-factor t(α) estimates for funds in the $5 million group exceeds the average from the simulations by just 0.13 standard errors (2.05 versus 1.92 in Table 4). Despite this small difference, 71.02% of the simulations produce 95th percentiles of t(α) below the 95th percentile for actual fund returns, which seems rather far above the numbers around 50% observed in Table 4 when t(α) percentiles for actual fund returns are closer to the averages from the simulations.

D. Kosowski et al (2006)

The paper closest to ours is Kosowski et al (2006). They use bootstrap simulations to draw inferences about information effects in the cross-section of four-factor t(α) estimates for net fund returns. Their main inference is more positive than ours. They find that the 95th and higher percentiles of four-factor t(α) for net fund returns are above the same simulation percentiles in more than 99% of simulation runs. This seems like strong evidence that among the best performing funds, some have more than sufficient information to cover costs. In contrast, our simulations uncover no evidence of information sufficient to cover costs. Two things account for their stronger results, (i) simulation approach and (ii) time period.

We jointly sample fund (and explanatory) returns, whereas Kosowski et al (2006) do simulations independently for each fund. Their simulations thus take no account of the correlation of the α estimates for different funds that arises because the explanatory returns of a benchmark model do not capture all common variation in fund returns (for example, industry effects). They summarize (but do not show) simulations that jointly sample the four-factor residuals of funds. But they never jointly sample fund and explanatory returns, which means (for example) they miss any effects of correlated movement in the volatilities of four-factor explanatory returns and residuals. In fact, in the results they show, the explanatory returns do not vary across simulation runs; the historical sequence of explanatory returns is used in every run.

Table 5 shows results for their 1975-2002 period using our simulation approach and the funds in our full ($5 million AUM) sample. We confirm their inference that the left tail of the cross-section of four-factor t(α) estimates for net returns is dominated by funds with insufficient information to cover costs. But our simulations do not confirm their inference that many right tail funds have more than sufficient information to cover costs. More than 99% of their simulation runs produce 95^{th} percentile t(α) estimates below the 95^{th} percentile from actual net fund returns. In our simulations the number is 53.09%, and the 95^{th} percentile t(α) for actual fund returns, 1.90, coincides almost exactly with the average value from the simulations, 1.91 . We infer that failure to account for the joint distribution of fund returns and the joint distribution of fund and explanatory returns substantially biases their tests toward suggesting positive information effects.[4]

Though less important, there is also a survival bias in their results. They require that funds exist for five years to appear in their tests. Skipping the details, we can report that when we impose this rule on the tests for 1975-2002, the fraction of simulation runs that produce 95^{th} percentile t(α) estimates below the 95^{th} percentile for actual net fund returns rises from 53.09% to 65.65%. We can also report that though our fund

[4] Cuthbertson et al (2008) apply the bootstrap simulation approach of Kosowski et al (2006) to UK mutual funds, with similar results and (we guess) similar problems.

inclusion rules differ somewhat from theirs, our sample sizes for 1975-2002, with and without the 60-month survival rule, are similar to those reported in their Table 1.

Time period is also a major source of differences in results. For 1984-2006, our simulation tests produce no hint of funds with sufficient information to cover costs. For example, in Table 4, the CDFs of t(α) for actual net fund returns are systematically to the left of the average CDFs from the simulations of net returns (in which funds have sufficient information to cover costs). But applied to the 1975-2002 period of Kosowski et al (2006), the 90^{th} and higher percentiles of t(α) for actual net fund returns are similar to the average values from the simulations (Table 5). And for 1975-2002 typically about half the simulation runs produce t(α) estimates at the 90^{th} and higher percentiles lower than those for net fund returns. This suggests that among the best performing funds of 1975-2002 are some that have information sufficient to cover costs. It is in contrast to the absence of such results for 1984-2006 (Table 4).

Likewise, in the tests on gross returns for 1984-2006, our simulations produce no reliable evidence of funds with any private information (Table 4). For 1984-2006, the distribution of t(α) for actual gross fund returns for the $5 million AUM group does have slight bulges in its left and right tails relative to the average distribution from the simulations of gross returns (in which no funds have private information). The simulations nevertheless say that the distributions of t(α) estimates for actual gross fund returns are not unusual for a world in which funds have no private information. In contrast, in the tests for 1975-2002, the 90^{th} through 98^{th} percentiles of t(α) for actual gross fund returns exceed the same percentiles from the simulations in more than 93% of simulation runs (Table 5). This suggests that among the top 10% of the funds of 1975-2002, some have good information.

What do we make of the stronger results for 1975-2002 versus 1984-2006? An intriguing story is that in olden times there were fewer funds, and it is possible, at least ex post, to infer that there were some funds with good information and perhaps some with sufficient information to cover costs. Over time, however, the knowledgeable managers lost their edge, perhaps because of the diseconomies of scale postulated by Berk and Green (2004). Or perhaps the entry of hordes of mediocre funds posturing as informed managers eventually makes it impossible to uncover the tracks of the truly informed. In either case, the evidence for 1975-2002 is interesting but probably irrelevant for the portfolio decisions of today's investors. And an alternative more pedestrian story is that the stronger results for 1975-2002 are to some extent due to subtle survival biases in the CRSP data that are more prevalent in earlier years (Elton, Gruber, and Blake 2001).

IV. Conclusions

For 1984-2006, when the CRSP mutual fund database is relatively free of survival bias, mutual funds on average and the average dollar invested in funds underperform three-factor and four-factor benchmarks by about the amount of fees and expenses. Thus, if there are funds with good information that enhances expected returns they are offset by funds with bad information.

We attempt to identify information effects via persistence tests and bootstrap simulations. In the persistence tests, sorting funds on t-statistics for market-adjusted average return does not point to the existence of funds with private information. Sorting funds on three-factor (or four-factor) $t(\alpha)$ estimates, however, produces some results consistent with information effects. These persistence results from sorts on $t(\alpha)$ are stronger for smaller funds, and they pretty much disappear after 1992.

The bootstrap simulations for 1984-2006 do not confirm the evidence of information effects from the persistence sorts on $t(\alpha)$. The simulation tests for net returns say that even in the extreme right tails of the cross-sections of three-factor and four-factor $t(\alpha)$ estimates there is no evidence of funds with information sufficient to cover costs. The simulation results for gross returns produce hints of the existence of funds with bad information and of others with good information. But they are just hints. Thus, rather large fractions of the simulation runs for gross returns (in which there are no information effects) produce more extreme left tails than the left tail of the cross-section of $t(\alpha)$ estimates for actual fund returns. And even larger fractions of the simulation runs produce stronger results than those observed in the right tail of the cross-section of $t(\alpha)$ estimates for actual fund returns. In the end, despite hints to the contrary, the simulation tests for 1984-2006 do not identify information effects in fund returns.

Persistence tests that examine post-ranking $t(\alpha)$ estimates for the winner and loser deciles from sorts of funds on pre-ranking $t(\alpha)$ seem similar to simulation tests that examine the 10^{th} and 90^{th} percentiles of $t(\alpha)$ estimates. Why do the two approaches produce somewhat different inferences? If some funds consistently have private information, the simulations probably have more power to infer their existence because the tests typically use longer periods to estimate α. But if access to information is temporary, its effects can be masked in a long return history. Persistence tests, which rank funds on short-term performance, are then probably better for identifying when funds have information. Our persistence results suggest that when there is post-ranking performance, it is indeed temporary – and so probably of little use to investors in choosing funds.

Similar comments apply when funds by chance temporarily hold stocks that do not conform to the three-factor or four-factor model. The problems of these models again show up better in persistence tests than in simulation tests that examine fund returns over longer periods. This explanation is attractive given our evidence that apparent information effects more often show up in persistence sorts on three-factor or four-factor t(α) than in sorts on past average return, which are less subject to bad model problems.

Finally, the mutual fund industry as a whole holds a portfolio much like the market portfolio, and realizes returns close to market returns, before fees and expenses. Our tests also support the inference that the cross-section of average fund returns is consistent with a world where individual fund performance, good and bad, is due to chance rather than skill. But we have only examined mutual funds. It is possible there are other segments of the investment industry, for example, hedge funds, where the segment as a whole or players within the segment produce returns that point to the existence of skill. Again, however, we face the constraints of equilibrium accounting. If, for example, the hedge fund industry produces positive α due to skill, there must be other investors that pay for its winnings, dollar for dollar, with negative α. And equilibrium accounting implies that the balancing of the dollar losses of losers against the gains of winners occurs before costs. After costs, that is, in terms of returns to investors, active management must in aggregate be a negative sum game, by the amount of the aggregate costs that active managers impose on investors (French 2008).

References

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, *Journal of Financial Economics* 9, 3-18.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows in rational markets, *Journal of Political Economy* 1269-1295.

Bollen, Nicolas, and Jeffrey Busse, 2004, Short-term persistence in mutual fund performance, *Review of Financial Studies* 18, 569-597.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Cuthbertson, Keith, Dirk Nitzche, and Niall O'Sullivan, 2008, UK Mutual fund performance: Skill or luck? *Journal of Empirical Finance*, forthcoming.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56, 2415-2430.

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

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

Fama, Eugene F., and Kenneth R. French, 2007, Disagreement, tastes, and asset prices, *Journal of Financial Economics* 83, 667-689.

French, Kenneth R., 2008, The cost of active investing, *Journal of Finance*, 63, forthcoming.

Goetzman, William, and Roger Ibbotson, 1994, Do winners repeat? Patterns in mutual fund performance, *Journal of Portfolio Management* 20, 9-18.

Grinblatt, Mark, and Sheridan Titman, 1992, Performance persistence in mutual funds, Journal of Finance 47, 1977-1984.

Gruber, Martin J. 1996, Another puzzle: The growth of actively managed mutual funds, *Journal of Finance* 51, 783-810.

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

Huberman, Gur, and Shmuel Kandel, 1987, Mean-variance spanning, *Journal of Finance* 42, 873-888.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 2033-2058.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551-2595.

Lakonishok, Joseph, Andrei Shleifer, and Robert W. Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541-1578.

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

Table 1 — Summary statistics for monthly explanatory returns of the three- and four-factor models

R_M, is the return on a value-weight market portfolio of NYSE, Amex (after 1962), and NASDAQ (after 1972) stocks, and R_f, is the one-month Treasury bill rate. The construction of SMB_t and VMG_t follows Fama and French (1993). At the end of June of each year k, we sort stocks into two size groups. Small includes NYSE, Amex (after 1962), and (after 1972) NASDAQ stocks with June market capitalization below the NYSE median and Big includes stocks with market cap above the NYSE median. We also sorts stocks into three book-to-market equity (B/P) groups, Growth (NYSE, Amex, and NASDAQ stocks in the bottom 30% of NYSE B/P), Neutral (middle 40% of NYSE B/P), and Value (top 30% of NYSE B/P). Book equity B is for the fiscal year ending in calendar year k-1, and the market cap P in B/P is for the end of December of k-1. The intersection of the (independent) size and B/P sorts produces six value-weight portfolios, refreshed at the end of June each year. The size return, SMB_t, is the simple average of the month t returns on the three Small stock portfolios minus the average of the returns on the three Big stock portfolios. The value-growth return, VMG_t, is the simple average of the returns on the two Value portfolios minus the average of the returns on the two Growth portfolios. The momentum return, MOM_t, is defined like VMG_t, except that we sort on prior return rather than B/P and the momentum sort is refreshed monthly rather than annually. At the end of each month t-1 we sort NYSE stocks on the average of the eleven months of returns to the end of month t-2. (Dropping the return for month t-1 is common in the momentum literature.) We use the 30th and 70th NYSE percentiles to assign NYSE, Amex, and NASDAQ stocks to Low, Medium, and High momentum groups. The intersection of the size sort for the most recent June and the independent momentum sort produces six value-weight portfolios, refreshed monthly. The momentum return, MOM_t, is the simple average of the month t returns on the two High momentum portfolios minus the average of the returns on the two Low momentum portfolios. The table shows average monthly return, the standard deviation of monthly returns and the t-statistic for the average monthly return. Time periods are as indicated, except that 2006 is September 2006.

	Average Return				Standard Deviation				t-statistic			
	R_M-R_f	SMB	VMG	MOM	R_M-R_f	SMB	VMG	MOM	R_M-R_f	SMB	VMG	MOM
1962-2006	0.45	0.23	0.48	0.82	4.42	3.22	2.90	3.98	2.36	1.65	3.82	4.80
1962-1983	0.25	0.44	0.56	0.86	4.48	3.03	2.59	3.57	0.92	2.36	3.48	3.91
1984-2006	0.64	0.03	0.40	0.79	4.36	3.38	3.17	4.35	2.42	0.13	2.10	3.01
1962-1992	0.38	0.24	0.47	0.82	4.53	2.87	2.54	3.41	1.60	1.62	3.59	4.65
1993-2006	0.61	0.20	0.49	0.83	4.15	3.90	3.57	5.04	1.90	0.67	1.76	2.10

Table 2 – Intercepts and slopes in variants of regression (1) for equal-weight (EW) and value-weight (VW) portfolios of mutual funds

The first two lines of the table show the average number of funds (Funds) and the average across funds of average monthly assets under management (AUM) in millions of dollars. Part A shows the annualized intercepts (12*α) and t-statistics for the intercepts (t(α)) for the three- and four-factor versions of regression (1) estimated on equal-weight (EW) and value-weight (VW) returns on the portfolio of mutual funds in our sample. Part B shows the regression slopes, t-statistics for the slopes (t(b-1) for the market slope) and the regression R^2 from the estimates of the four-factor version of (1). Part A shows results for gross and net fund returns. Part B shows regression slopes for only net returns. The time periods are as indicated, except that 2006 is September 2006.

	1962-2006		1962-1983		1984-2006		1962-1992		1993-2006	
	EW	VW	EW	VW	EW	VW	EW	VW	EW	VW
Funds	819		246		1374		331		1921	
AUM	410.7		136.7		675.8		179.0		933.2	
Part A: Three-factor and four-factor α estimates (monthly values multiplied by 12)										
Net Returns										
Three factor										
12 * α	-0.52	-0.49	-0.23	-0.08	-0.94	-0.80	-0.08	-0.16	-1.44	-1.16
t(α)	-1.38	-1.67	-0.37	-0.17	-2.21	-2.62	-0.17	-0.41	-2.56	-3.09
Four factor										
12 * α	-0.85	-0.86	-1.25	-0.87	-0.91	-0.93	-0.86	-0.71	-1.38	-1.38
t(α)	-2.19	-2.90	-2.08	-1.75	-2.08	-2.98	-1.92	-1.84	-2.39	-3.68
Gross Returns										
Three factor										
12 * α	0.61	0.26	0.74	0.53	0.36	0.10	0.97	0.52	-0.10	-0.23
t(α)	1.61	0.89	1.21	1.04	0.85	0.32	2.14	1.38	-0.18	-0.61
Four factor										
12 * α	0.29	-0.11	-0.29	-0.27	0.39	-0.03	0.18	-0.02	-0.03	-0.45
t(α)	0.74	-0.35	-0.49	-0.53	0.90	-0.09	0.41	-0.06	-0.06	-1.20
Part B: Four-factor regression slopes and R^2										
b	0.97	0.97	0.94	0.97	0.98	0.97	0.94	0.95	1.00	0.99
t(b-1)	-4.14	-5.31	-5.08	-3.42	-1.81	-5.20	-6.75	-6.03	0.33	-1.53
s	0.20	0.05	0.26	0.08	0.18	0.05	0.26	0.08	0.19	0.05
t(s)	19.71	6.86	15.41	5.38	15.81	6.26	19.44	7.21	13.86	5.77
v	-0.03	-0.05	-0.08	-0.09	-0.00	-0.03	-0.10	-0.10	0.05	0.00
t(v)	-2.85	-5.86	-4.12	-5.65	-0.23	-2.93	-6.66	-7.45	3.21	0.16
m	0.03	0.03	0.08	0.06	-0.00	0.01	0.06	0.05	-0.01	0.02
t(m)	3.50	5.17	5.81	5.38	-0.34	1.92	6.17	5.07	-0.56	2.89
R^2	0.98	0.99	0.97	0.98	0.98	0.99	0.98	0.98	0.98	0.99

Table 3 – Persistence tests for sorts on gross returns

Each month from December 1983 to August 2006 funds with at least $5 million AUM for this or any preceding month are sorted into deciles based on their prior gross returns. Equal-weight (EW) and value-weight (VW) monthly gross returns on the decile portfolios are calculated for the 36 months (if available) after portfolio formation. The table summarizes three- and four-factor α estimates for monthly gross returns for each of the first three months, the second, third, and fourth quarters, and the first, second, and third years after portfolio formation. The month t return for a k-month post-formation interval is an EW average of the month t EW or VW returns for the portfolios formed at the end of the preceding k months. The table shows three- and four-factor α estimates and their t-statistics, t(α), and four-factor momentum slopes, m, and their t-statistics, t(m), for the bottom and top past performance deciles. Funds are sorted on the t-statistics of their average net-of-market gross returns for the preceding 12 months (8 month minimum, in Part A) and the preceding 60 months (24 month minimum, in Part B), and on the t-statistics for three-factor α estimates from 60 months of past gross returns (Part C). The periods for post-sort returns are January 1984 to September 2006 (Parts A, B, and C) and January 1993 to September 2006 (Parts B and C) for the first month after portfolio formation.

Part A: Funds sorted on t-statistics for average net-of-market gross returns for the preceding 12 months

	EW Gross Returns									VW Gross Returns								
	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36
									1984-2006									
Three-Factor α Estimates for Bottom Decile																		
α	-0.29	-0.25	-0.24	-0.26	-0.17	-0.04	-0.19	0.02	0.08	-0.19	-0.13	-0.14	-0.19	-0.08	0.01	-0.11	0.07	0.18
t(α)	-2.59	-2.33	-2.39	-2.79	-1.99	-0.52	-2.55	0.40	1.24	-1.89	-1.29	-1.44	-2.26	-1.04	0.12	-1.70	1.19	2.44
Three-Factor α Estimates for Top Decile																		
α	0.55	0.42	0.39	0.36	0.14	-0.01	0.25	-0.05	-0.05	0.40	0.29	0.25	0.22	0.07	-0.05	0.15	-0.01	-0.01
t(α)	4.99	3.82	3.74	3.92	1.68	-0.11	3.31	-0.69	-0.68	3.71	2.68	2.48	2.58	0.97	-0.58	2.15	-0.15	-0.18
Four-Factor Estimates of α and momentum slope m for Bottom Decile																		
α	-0.01	0.02	-0.01	-0.09	-0.09	-0.02	-0.06	-0.01	-0.01	0.06	0.11	0.08	-0.06	-0.03	-0.00	-0.01	0.01	0.06
t(α)	-0.08	0.22	-0.07	-1.12	-1.05	-0.29	-0.86	-0.14	-0.17	0.74	1.37	0.96	-0.75	-0.40	-0.06	-0.15	0.20	0.90
m	-0.28	-0.26	-0.23	-0.16	-0.08	-0.02	-0.13	0.03	0.09	-0.24	-0.24	-0.21	-0.13	-0.05	0.01	-0.10	0.06	0.12
t(m)	-15.56	-14.99	-13.41	-9.13	-4.43	-0.98	-8.99	2.41	7.04	-14.58	-13.65	-12.29	-7.76	-2.88	0.82	-7.57	4.54	8.63
Four-Factor Estimates of α and momentum slope m for Top Decile																		
α	0.29	0.14	0.12	0.14	0.01	-0.06	0.07	0.03	0.04	0.15	0.03	0.01	0.05	-0.01	-0.06	0.01	0.06	0.05
t(α)	3.36	1.76	1.58	2.01	0.10	-0.68	1.26	0.44	0.65	1.72	0.35	0.09	0.65	-0.16	-0.66	0.19	0.98	0.77
m	0.26	0.27	0.27	0.21	0.13	0.05	0.17	-0.08	-0.09	0.25	0.25	0.24	0.17	0.08	0.01	0.13	-0.07	-0.06
t(m)	13.98	15.31	16.29	14.01	7.97	2.64	13.24	-5.19	-6.48	13.73	14.11	14.10	10.86	5.31	0.46	10.49	-5.06	-4.51

Table 3, Part B – Funds sorted on t-statistics for average net-of-market gross returns for the preceding 60 months

	EW Gross Returns									VW Gross Returns								
	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36
1984-2006																		
Three-Factor α Estimates for Bottom Decile																		
α	-0.25	-0.22	-0.20	-0.19	-0.14	-0.08	-0.17	-0.06	-0.10	-0.15	-0.12	-0.11	-0.13	-0.09	-0.03	-0.10	0.04	0.02
t(α)	-3.09	-2.82	-2.57	-2.56	-2.06	-1.30	-2.44	-1.06	-1.75	-2.09	-1.61	-1.46	-1.93	-1.30	-0.41	-1.51	0.59	0.34
Three-Factor α Estimates for Top Decile																		
α	0.25	0.19	0.16	0.12	0.04	-0.04	0.09	-0.04	-0.01	0.12	0.10	0.07	0.05	0.01	-0.05	0.03	-0.02	0.05
t(α)	3.37	2.66	2.44	1.81	0.52	-0.48	1.40	-0.61	-0.11	1.82	1.60	1.16	0.79	0.08	-0.80	0.45	-0.33	0.98
Four-Factor Estimates of α and momentum slopes m for Bottom Decile																		
α	-0.11	-0.09	-0.08	-0.11	-0.11	-0.08	-0.10	-0.10	-0.12	-0.06	-0.02	-0.02	-0.07	-0.06	-0.02	-0.05	0.01	0.00
t(α)	-1.49	-1.24	-1.14	-1.50	-1.57	-1.23	-1.50	-1.68	-1.94	-0.79	-0.26	-0.29	-1.01	-0.82	-0.26	-0.71	0.09	0.02
m	-0.14	-0.13	-0.11	-0.08	-0.03	-0.00	-0.06	0.03	0.01	-0.10	-0.10	-0.09	-0.06	-0.03	-0.01	-0.05	0.03	0.02
t(m)	-8.78	-8.36	-7.25	-5.11	-2.07	-0.15	-4.33	2.82	1.07	-6.31	-6.46	-5.47	-4.22	-2.11	-0.64	-3.57	2.22	1.45
Four-Factor Estimates of α and momentum slopes m for Top Decile																		
α	0.10	0.06	0.05	0.07	0.05	0.02	0.05	0.06	0.05	0.02	0.02	0.02	0.03	0.04	0.01	0.02	0.06	0.05
t(α)	1.61	0.89	0.88	0.98	0.63	0.25	0.86	0.90	0.79	0.39	0.25	0.28	0.54	0.57	0.14	0.39	0.98	1.11
m	0.14	0.13	0.11	0.06	-0.01	-0.05	0.03	-0.10	-0.06	0.10	0.09	0.05	0.02	-0.03	-0.06	0.00	-0.07	-0.01
t(m)	10.44	9.57	8.05	3.80	-0.56	-3.35	2.35	-6.96	-4.23	6.99	6.46	4.00	1.08	-2.25	-4.33	0.19	-5.97	-0.72
1993-2006																		
Three-Factor α Estimates for Bottom Decile																		
α	-0.20	-0.18	-0.15	-0.13	-0.04	0.03	-0.09	0.02	-0.04	-0.12	-0.12	-0.13	-0.14	-0.06	0.03	-0.08	0.05	0.05
t(α)	-1.70	-1.62	-1.41	-1.29	-0.46	0.38	-0.95	0.26	-0.67	-1.50	-1.44	-1.59	-1.87	-0.79	0.48	-1.16	0.88	0.83
Three-Factor α Estimates for Top Decile																		
α	0.24	0.16	0.12	0.06	-0.07	-0.18	0.01	-0.17	-0.09	0.08	0.03	-0.02	-0.03	-0.10	-0.18	-0.06	-0.11	0.02
t(α)	2.17	1.57	1.30	0.64	-0.63	-1.61	0.06	-1.59	-0.98	0.85	0.30	-0.19	-0.28	-1.00	-1.92	-0.76	-1.29	0.37
Four-Factor α Estimates for Bottom Decile																		
α	-0.02	-0.01	-0.01	-0.02	0.01	0.05	0.00	-0.01	-0.04	-0.01	-0.01	-0.03	-0.07	-0.02	0.05	-0.02	0.02	0.02
t(α)	-0.17	-0.12	-0.07	-0.22	0.16	0.65	0.05	-0.11	-0.57	-0.19	-0.08	-0.41	-0.92	-0.25	0.69	-0.28	0.31	0.43
Four-Factor α Estimates for Top Decile																		
α	0.07	0.01	0.00	-0.00	-0.06	-0.12	-0.03	-0.05	-0.03	-0.03	-0.08	-0.08	-0.05	-0.06	-0.12	-0.07	-0.03	0.02
t(α)	0.75	0.12	0.02	-0.03	-0.52	-1.07	-0.36	-0.51	-0.29	-0.30	-0.89	-0.92	-0.51	-0.66	-1.29	-0.82	-0.40	0.26

Table 3, Part C – Funds sorted on t-statistics for three-factor α estimates from 60 months of past gross returns

	EW Gross Returns									VW Gross Returns								
	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36	1	2	3	4-6	7-9	10-12	1-12	13-24	25-36
1984-2006																		
Three-Factor α Estimates for Bottom Decile																		
α	-0.25	-0.24	-0.22	-0.23	-0.18	-0.14	-0.21	-0.09	-0.14	-0.10	-0.12	-0.13	-0.12	-0.09	-0.03	-0.10	-0.00	-0.02
t(α)	-3.80	-3.55	-3.49	-3.63	-2.98	-2.40	-3.48	-1.87	-3.16	-1.47	-1.69	-1.91	-1.84	-1.31	-0.41	-1.67	-0.09	-0.42
Three-Factor α Estimates for Top Decile																		
α	0.30	0.26	0.24	0.23	0.16	0.09	0.19	0.01	0.05	0.21	0.16	0.14	0.14	0.11	0.06	0.12	0.02	0.06
t(α)	4.37	3.93	3.72	3.68	2.44	1.48	3.17	0.21	0.82	3.22	2.65	2.48	2.49	1.85	1.09	2.18	0.42	1.20
Four-Factor Estimates of α and momentum slopes m for Bottom Decile																		
α	-0.18	-0.17	-0.17	-0.19	-0.17	-0.14	-0.18	-0.12	-0.17	-0.01	-0.03	-0.05	-0.07	-0.05	0.00	-0.05	-0.01	-0.04
t(α)	-2.82	-2.61	-2.66	-3.01	-2.74	-2.40	-2.96	-2.38	-3.85	-0.14	-0.46	-0.76	-0.98	-0.71	0.04	-0.81	-0.14	-0.91
m	-0.07	-0.06	-0.05	-0.04	-0.01	0.00	-0.03	0.03	0.03	-0.09	-0.09	-0.08	-0.06	-0.04	-0.03	-0.05	0.00	0.02
t(m)	-4.66	-4.40	-3.78	-2.59	-0.75	0.26	-2.06	2.43	3.30	-6.35	-5.77	-5.41	-3.94	-2.64	-2.05	-3.90	0.25	2.35
Four-Factor Estimates of α and momentum slopes m for Top Decile																		
α	0.22	0.19	0.18	0.20	0.15	0.10	0.17	0.08	0.10	0.13	0.09	0.09	0.11	0.10	0.07	0.09	0.07	0.06
t(α)	3.33	2.88	2.87	3.20	2.27	1.60	2.72	1.26	1.53	2.05	1.48	1.57	1.92	1.73	1.25	1.70	1.33	1.26
m	0.07	0.07	0.05	0.03	0.01	-0.01	0.02	-0.06	-0.05	0.08	0.07	0.05	0.03	0.00	-0.01	0.02	-0.05	-0.00
t(m)	5.13	5.12	3.86	1.86	0.52	-0.72	1.73	-4.72	-3.37	5.83	5.72	4.16	2.42	0.33	-0.82	1.96	-4.06	-0.44
1993-2006																		
Three-Factor α Estimates for Bottom Decile																		
α	-0.18	-0.15	-0.14	-0.15	-0.08	-0.04	-0.11	-0.01	-0.08	-0.11	-0.11	-0.12	-0.08	-0.03	0.02	-0.06	0.03	0.01
t(α)	-2.18	-1.87	-1.86	-2.05	-1.16	-0.62	-1.61	-0.16	-1.67	-1.29	-1.21	-1.46	-0.94	-0.42	0.26	-0.76	0.48	0.17
Three-Factor α Estimates for Top Decile																		
α	0.24	0.19	0.15	0.14	0.05	-0.04	0.09	-0.14	-0.07	0.15	0.09	0.06	0.04	0.01	-0.06	0.02	-0.10	-0.01
t(α)	2.37	1.92	1.66	1.62	0.60	-0.51	1.04	-1.56	-0.78	1.51	0.92	0.69	0.49	0.12	-0.73	0.31	-1.30	-0.13
Four-Factor α Estimates for Bottom Decile																		
α	-0.09	-0.07	-0.08	-0.11	-0.07	-0.04	-0.08	-0.03	-0.11	-0.00	0.01	-0.03	-0.00	0.01	0.05	0.01	0.02	-0.02
t(α)	-1.22	-0.94	-1.11	-1.47	-0.91	-0.58	-1.10	-0.53	-2.20	-0.03	0.12	-0.34	-0.03	0.19	0.70	0.15	0.39	-0.38
Four-Factor α Estimates for Top Decile																		
α	0.14	0.08	0.07	0.10	0.03	-0.04	0.05	-0.06	-0.01	0.04	-0.02	-0.01	-0.01	-0.01	-0.06	-0.02	-0.05	-0.00
t(α)	1.41	0.92	0.84	1.08	0.35	-0.45	0.57	-0.74	-0.11	0.42	-0.18	-0.15	-0.18	-0.13	-0.72	-0.23	-0.66	-0.04

Table 4 - Percentiles of t(α) estimates for actual and simulated fund returns: January 1984 to September 2006

The table shows values of t(α) at selected percentiles (Pct) of the distribution of t(α) estimates for actual (Act) net and gross fund returns. The table also shows the percent of the 10,000 simulation runs that produce lower values of t(α) at the selected percentiles than those observed for actual fund returns (% < Act). Sim is the average value of t(α) at the selected percentiles from the simulations. The period is January 1984 to September 2006 and results are shown for the three- and four-factor models for the $5 million, $250 million, and $1 billion AUM fund groups.

	5 Million			250 Million			1 Billion		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
3-Factor Net Returns									
1	-2.93	-3.92	0.32	-2.69	-3.79	0.48	-2.47	-4.23	0.02
2	-2.44	-3.45	0.24	-2.28	-3.32	0.50	-2.14	-3.46	0.29
3	-2.18	-3.19	0.18	-2.05	-3.11	0.43	-1.95	-3.27	0.19
4	-1.99	-3.00	0.13	-1.90	-2.97	0.27	-1.81	-3.13	0.20
5	-1.86	-2.85	0.13	-1.77	-2.85	0.20	-1.70	-2.95	0.25
10	-1.41	-2.31	0.11	-1.36	-2.35	0.16	-1.32	-2.45	0.13
20	-0.91	-1.73	0.09	-0.89	-1.84	0.08	-0.87	-1.94	0.06
30	-0.55	-1.24	0.10	-0.56	-1.38	0.12	-0.54	-1.57	0.01
40	-0.26	-0.91	0.09	-0.27	-1.00	0.12	-0.26	-1.17	0.03
50	0.02	-0.60	0.08	-0.01	-0.70	0.09	0.00	-0.81	0.05
60	0.29	-0.28	0.11	0.26	-0.36	0.27	0.26	-0.52	0.05
70	0.58	0.07	0.32	0.54	-0.06	0.33	0.54	-0.23	0.06
80	0.94	0.48	1.03	0.87	0.36	1.31	0.87	0.20	0.38
90	1.44	1.00	1.89	1.34	0.87	3.28	1.32	0.68	0.90
95	1.88	1.51	5.20	1.73	1.30	5.75	1.71	1.26	7.28
96	2.02	1.69	7.88	1.85	1.44	7.09	1.82	1.43	11.41
97	2.20	1.89	9.73	2.00	1.59	7.98	1.96	1.67	20.42
98	2.47	2.10	6.13	2.20	1.82	9.97	2.16	1.91	24.41
99	2.95	2.41	0.63	2.56	2.27	18.87	2.51	2.21	20.97
4-Factor Net Returns									
1	-2.92	-3.93	0.33	-2.68	-3.89	0.29	-2.46	-4.03	0.14
2	-2.43	-3.45	0.26	-2.28	-3.36	0.39	-2.13	-3.35	0.43
3	-2.17	-3.10	0.36	-2.05	-3.05	0.60	-1.93	-3.04	0.59
4	-1.99	-2.89	0.37	-1.89	-2.85	0.63	-1.79	-2.92	0.53
5	-1.85	-2.75	0.30	-1.76	-2.77	0.35	-1.68	-2.82	0.38
10	-1.40	-2.22	0.28	-1.35	-2.32	0.25	-1.30	-2.45	0.15
20	-0.89	-1.66	0.23	-0.87	-1.76	0.21	-0.84	-1.90	0.14
30	-0.54	-1.22	0.30	-0.54	-1.37	0.22	-0.52	-1.50	0.10
40	-0.24	-0.88	0.34	-0.25	-1.03	0.24	-0.24	-1.20	0.09
50	0.03	-0.58	0.27	0.01	-0.69	0.33	0.02	-0.85	0.14
60	0.31	-0.27	0.33	0.28	-0.39	0.40	0.28	-0.46	0.31
70	0.61	0.05	0.42	0.56	-0.08	0.49	0.56	-0.16	0.40
80	0.96	0.47	1.19	0.89	0.32	1.41	0.88	0.20	0.68
90	1.47	1.06	4.23	1.36	0.90	5.05	1.34	0.83	4.80
95	1.92	1.51	4.54	1.76	1.39	11.38	1.73	1.28	8.52
96	2.06	1.65	4.46	1.88	1.50	10.75	1.84	1.41	9.72
97	2.25	1.80	2.96	2.03	1.61	8.08	1.99	1.50	7.41
98	2.51	2.07	3.41	2.23	1.78	6.29	2.19	1.65	5.36
99	3.00	2.36	0.30	2.58	2.10	6.28	2.54	2.22	21.22

Table 4 (continued)

	5 Million			250 Million			1 Billion		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
3-Factor Gross Returns									
1	-2.92	-3.17	18.80	-2.68	-2.97	18.51	-2.47	-3.19	4.03
2	-2.43	-2.73	14.35	-2.28	-2.66	12.28	-2.14	-2.63	9.55
3	-2.17	-2.52	10.60	-2.05	-2.45	10.82	-1.95	-2.46	8.09
4	-1.99	-2.32	11.45	-1.90	-2.33	9.07	-1.81	-2.31	8.13
5	-1.85	-2.18	11.15	-1.77	-2.12	12.81	-1.70	-2.17	9.01
10	-1.41	-1.70	12.53	-1.37	-1.74	10.05	-1.32	-1.81	6.96
20	-0.91	-1.08	21.85	-0.89	-1.19	13.11	-0.87	-1.37	4.63
30	-0.56	-0.68	27.58	-0.56	-0.75	21.35	-0.55	-0.96	6.33
40	-0.26	-0.35	33.28	-0.27	-0.43	25.82	-0.26	-0.55	13.50
50	0.01	-0.05	38.32	-0.01	-0.12	32.73	0.00	-0.27	14.30
60	0.29	0.28	49.55	0.26	0.17	35.74	0.26	0.02	17.07
70	0.58	0.63	60.48	0.54	0.49	44.01	0.54	0.32	20.33
80	0.93	1.03	68.21	0.87	0.88	53.61	0.87	0.72	30.48
90	1.43	1.53	67.44	1.34	1.38	58.45	1.32	1.27	45.35
95	1.88	1.98	68.45	1.73	1.77	58.76	1.71	1.73	55.48
96	2.02	2.13	68.68	1.85	1.88	56.57	1.82	1.92	64.96
97	2.20	2.34	72.68	2.00	2.05	60.13	1.96	2.20	77.82
98	2.46	2.63	76.03	2.21	2.35	69.66	2.17	2.30	67.86
99	2.95	2.96	54.79	2.57	2.80	77.49	2.52	2.92	86.17
4-Factor Gross Returns									
1	-2.92	-3.11	23.31	-2.67	-2.86	26.72	-2.46	-3.15	5.05
2	-2.43	-2.72	14.86	-2.27	-2.62	14.12	-2.13	-2.44	19.05
3	-2.17	-2.47	13.52	-2.05	-2.41	13.50	-1.93	-2.31	14.48
4	-1.99	-2.25	16.44	-1.89	-2.30	10.74	-1.79	-2.22	11.98
5	-1.85	-2.11	16.09	-1.76	-2.12	13.44	-1.68	-2.11	11.46
10	-1.40	-1.61	19.69	-1.35	-1.65	15.93	-1.30	-1.82	6.91
20	-0.89	-1.05	24.38	-0.87	-1.15	16.92	-0.84	-1.30	7.59
30	-0.54	-0.65	31.70	-0.54	-0.71	26.49	-0.52	-0.86	12.95
40	-0.24	-0.32	36.44	-0.25	-0.41	27.76	-0.24	-0.55	14.28
50	0.03	0.00	45.36	0.01	-0.13	30.74	0.02	-0.24	18.33
60	0.31	0.30	49.15	0.27	0.14	32.58	0.28	0.09	25.96
70	0.60	0.64	56.76	0.56	0.51	45.13	0.56	0.41	31.57
80	0.96	1.02	62.63	0.89	0.90	53.28	0.88	0.82	43.91
90	1.46	1.61	72.54	1.36	1.44	62.67	1.34	1.39	58.35
95	1.92	2.05	71.02	1.76	1.93	72.53	1.73	1.82	62.90
96	2.06	2.18	69.64	1.88	2.00	66.81	1.84	1.96	65.69
97	2.24	2.32	63.52	2.03	2.18	70.84	1.99	2.08	61.95
98	2.51	2.54	58.23	2.24	2.35	66.32	2.19	2.28	62.18
99	3.00	2.92	41.31	2.59	2.64	57.71	2.55	2.50	47.28

Table 5 – Percentiles of four-factor t(α) for actual and simulated fund returns: 1975 to 2002

The table shows values of four-factor t(α) at selected percentiles (Pct) of the distribution of t(α) for actual (Act) net and gross fund returns. The period is 1975 to 2002 (as in Kosowski et al 2006) and results are shown for the $5 million AUM group (our full sample). The table also shows the fractions (% < Act) of the 10,000 simulation runs that produce lower values of t(α) at the selected percentiles than those observed for actual fund returns. Sim is the average value of t(α) at the selected percentiles from the simulations.

	Net Returns			Gross Returns		
Pct	Sim	Act	%<Act	Sim	Act	%<Act
1	-2.85	-3.75	0.64	-2.86	-2.94	35.68
2	-2.37	-3.25	0.63	-2.38	-2.63	17.82
3	-2.12	-2.85	1.23	-2.13	-2.32	21.90
4	-1.94	-2.67	1.13	-1.95	-2.13	23.71
5	-1.81	-2.48	1.50	-1.81	-1.95	28.88
10	-1.37	-1.95	2.22	-1.37	-1.48	31.89
20	-0.87	-1.43	1.97	-0.87	-0.93	39.12
30	-0.52	-1.05	2.04	-0.53	-0.53	47.92
40	-0.23	-0.70	2.77	-0.23	-0.17	59.93
50	0.04	-0.38	4.16	0.04	0.13	65.52
60	0.31	-0.05	6.31	0.31	0.48	76.00
70	0.61	0.28	9.00	0.60	0.82	80.61
80	0.96	0.72	19.28	0.95	1.26	86.96
90	1.46	1.38	42.58	1.45	1.90	93.56
95	1.91	1.90	53.09	1.90	2.52	97.00
96	2.05	2.07	56.97	2.04	2.63	96.39
97	2.24	2.28	60.51	2.23	2.82	96.35
98	2.51	2.50	53.36	2.49	3.03	95.06
99	3.03	2.84	27.84	3.01	3.30	83.98