Judging Fund Managers by the Company They Keep

Randolph Cohen

Joshua Coval

Ľuboš Pástor*

October 16, 2003

Abstract

We develop a performance evaluation approach in which a fund manager's skill is judged by the extent to which his investment decisions resemble the decisions of managers with distinguished performance records. The proposed performance measures use historical returns and holdings of many funds to evaluate the performance of a single fund. Simulations demonstrate that our measures are particularly useful in ranking managers. In an application that relies on such ranking, our measures reveal strong predictability in the returns of U.S. equity funds. Our measures provide information about future fund returns that is not contained in the standard measures.

---

*Cohen and Coval are from the Harvard Business School. Pástor is from the Graduate School of Business at the University of Chicago, NBER, and CEPR. Research support from the Center for Research in Security Prices and the James S. Kemper Faculty Research Fund at the Graduate School of Business, University of Chicago is gratefully acknowledged (Pástor). We thank Jonathan Berk, Jeff Busse, George Chacko, Aida Charoenrook, Mike Cooper, Roger Edelen, Gene Fama, Rick Green (editor), Owen Lamont, Rob Stambaugh, Pietro Veronesi, Tuomo Vuolteenaho, Russ Wermers, an anonymous referee, and seminar participants at the AIM Investment Center Conference on Mutual Funds at the University of Texas at Austin, Arizona State University, Berkeley Program in Finance, Dartmouth College, Emory University, Harvard University, Indiana University, Michigan State University, Morningstar, Northwestern University, Purdue University, Thomson Financial, University of Chicago, University of Pennsylvania, Vanderbilt University, and the 2003 Western Finance Association Meetings for helpful comments. Huong Trieu provided excellent research assistance.

Judging Fund Managers by the Company They Keep

Abstract

We develop a performance evaluation approach in which a fund manager's skill is judged by the extent to which his investment decisions resemble the decisions of managers with distinguished performance records. The proposed performance measures use historical returns and holdings of many funds to evaluate the performance of a single fund. Simulations demonstrate that our measures are particularly useful in ranking managers. In an application that relies on such ranking, our measures reveal strong predictability in the returns of U.S. equity funds. Our measures provide information about future fund returns that is not contained in the standard measures.

Introduction

People are often judged by the company they keep. Various human characteristics that are difficult to observe directly are commonly inferred from the characteristics of others who behave in a similar manner. For example, a bored air traveler eager to chat about Madonna's life story is more likely to turn to a fellow passenger reading the National Enquirer than to a passenger reading the Wall Street Journal because the readership of the former periodical is generally known to be more interested in celebrity gossip.

In the same spirit, when people compete, their success is often predicted by their techniques, and the quality of a given technique is frequently evaluated based on the track records of the technique's followers. As an example, suppose a group of basketball players, some of whom shoot with both hands and some with only one hand, have been taking 10 shots each at the basket. So far, the average score of the two-handers is 8/10, while the one-handers' average is only 4/10. Two players, one one-hander and one two-hander, have completed only half of their shots so far, and both have scored 4/5. Suppose you are to bet on which of the two players is going to achieve a higher score out of 10. Although the track records of both players are identical, it seems sensible to bet on the two-hander, because the track records of the other players show that two-handed shooters tend to score higher. The one-hander, who is employing what appears to be an inferior technique, is more likely to have been lucky in his first five shots.

Similar to basketball players, active mutual fund managers rely on a variety of techniques when trying to beat their benchmarks, and many of these techniques are common to groups of managers. For example, managers collect information from different sources and use different valuation methods, but there are clusters of managers who use similar sources and similar methods. Managers using similar techniques are likely to make similar decisions. The premise of this paper is that managers who make similar investment decisions should be expected to deliver similar future performance.

To further support this premise, suppose valuable private information (e.g. a new discovery by a biotech firm) is occasionally received by some managers but not by others. Skill is required to obtain this private information, so managers with more skill receive such information more often. Since they act on the same information, the skilled managers tend to make similar investment decisions (e.g. they all buy the stock of the biotech firm), even if they use different techniques to obtain the private information. As a result, we can tell whether a manager is skilled by comparing his investment decisions with the decisions of other skilled managers.

The premise seems reasonable even if all information is public, but this information is interpreted well by some managers and poorly by others. Managers with more skill are more likely to interpret public information well. In this case, skilled managers again tend to make similar investment decisions because they interpret information similarly. Building on our premise, we propose to judge a fund manager's skill by the extent to which his investment decisions resemble those of other successful managers.

One way to assess the similarity of the managers' investment decisions is to compare the compositions of their portfolios. For example, consider two managers with equally impressive past returns, where one manager currently keeps a big chunk of his portfolio in the stock of Intel, while the other manager holds mostly Microsoft. Suppose also that Intel is currently held especially by managers with good past performance, whereas Microsoft is held mostly by managers with undistinguished records. It seems reasonable to think that the first manager, whose decision to hold Intel is shared by a higher-caliber set of managers, has superior ability to select stocks, while the second manager, whose decisions coincide with those of subpar managers, has been merely fortunate.

The performance measure proposed in the paper is defined with respect to some simpler reference measure, for which we choose the traditional Jensen's alpha. We show that our measure of a manager's skill is a weighted average of the traditional skill measures across all managers, where the weights are essentially the covariances between the manager's portfolio weights and the weights of the other managers. Put differently, if two managers have highly similar portfolio weights, then one manager's skill contributes substantially to our measure of the other manager's skill.

Another way of comparing the managers' investment decisions is to compare their trades. A modified "trade-based" version of our performance measure judges a manager's skill by the extent to which recent changes in his holdings match those of managers with outstanding past performance. This measure is also a weighted average of the traditional skill measures, but now the weights are essentially the covariances between the concurrent changes in the manager's portfolio weights and those of the other managers. According to the trade-based measure, the manager is skilled if he tends to buy stocks that are concurrently purchased by other managers who have performed well, and sell stocks that are concurrently purchased by managers who have performed poorly.

Evaluating mutual fund performance is a topic of enormous relevance for the well-being of individual investors and for market efficiency in general. The traditional approach relies solely on historical fund returns to construct measures such as Jensen's alpha (Jensen, 1968) or Sharpe ratio (Sharpe, 1966). Since return histories of many funds are short, these traditional measures are often imprecise. To cope with the low precision, recent studies propose alternative performance measures that rely also on fund holdings.[1] Those measures employ the portfolio holdings of the fund whose performance is being evaluated, but they do not exploit the information about the fund's performance contained in the holdings and returns of other funds. Including that additional information and documenting its benefits is the contribution of this paper.

Our performance measures offer substantially higher precision than the traditional return-based measures. Since our measures are weighted averages of the traditional measures, precision is added by pooling information across funds. That is, instead of using just the historical returns of a given manager to estimate his performance, our measures use the return series of all managers whose holdings (or changes in holdings) overlap with those of the given manager. The biggest precision gains from using our measures are obtained for short-history funds. In fact, our measures have reasonably low standard errors even for funds with track records as short as one quarter.

A simulation analysis is conducted to examine the extent to which our skill measures are able to capture true skill. When managers are ranked by true skill and then separately by various performance estimators, our estimators produce higher rank correlations with true skill than standard estimators that do not exploit similarities in holdings or trades across managers. Our measures are particularly effective when the number of managers is large and when fund return history is relatively short. Our measures exhibit some bias in estimating the traditional measure, alpha. Due to the weight-averaging of skill across managers, our holdings-based measure is biased towards the average level of skill in the population of managers, similar to a Bayesian shrink-

[1] See, for example, Grinblatt and Titman (1993), Daniel, Grinblatt, Titman, and Wermers (1997), Wermers (2000), and Ferson and Khang (2002). With no reliance on holdings data, Pástor and Stambaugh (2002) and Busse and Irvine (2002) show that additional precision in the traditional performance measures can be achieved by incorporating returns on seemingly unrelated passive assets.

age estimator. The nature of the bias does not impair the rank-ordering of managers, though, and our estimators dominate the usual estimators of alpha even when the objective is to rank managers by their alphas. In sum, the simulations reveal that our measures are especially useful in applications that involve ranking managers.

One such application is our empirical examination of return predictability for U.S. equity funds. The literature is ambiguous as to whether mutual fund returns are predictable. Grinblatt and Titman (1992), Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994), Brown and Goetzmann (1995), Elton, Gruber, and Blake (1996), and Bollen and Busse (2002), among others, all find some persistence in fund performance, especially among the funds lagging their benchmarks. However, much of the persistence has been attributed to momentum in stock returns (Carhart, 1997), and some also to survivorship bias (Brown, Goetzmann, Ibbotson, and Ross, 1992). The mixed nature of the evidence begs for further research.

To analyze fund return predictability, at the beginning of each quarter between April 1982 and July 2002, funds are sorted into deciles according to our measures as well as alpha estimated over the past year. The deciles' returns are tracked over the subsequent quarter, and the performance of the decile portfolios is compared over the full sample. We find that fund returns exhibit significant persistence, even after adjusting for momentum in stock returns. The difference between the risk-adjusted returns of the top and bottom deciles sorted by alpha ranges from 3.7% to 5.2% per year across different benchmark models. Our holdings-based measure produces bigger differences, between 5.9% and 7.4% per year, and our trade-based measure yields results similar to alpha.

To see whether our measures contain information about future fund performance that is not contained in alpha, we conduct double-sorts, sorting first by alpha and then by our measures. Controlling for alpha, the average difference between the top and bottom quintiles of funds ranked by our holdings-based measure ranges from 2.4% to 4.4% per year, with t-statistics ranging from 1.94 to 3.21. The same difference computed using our trade-based measure ranges from 1.2% to 1.4% per year, with t-statistics of 2.31 to 2.73. These results indicate that both of our measures contain significant information about future fund returns that is not contained in alpha.

We also examine how much information is contained in alpha but not in our measures, sorting first by our measures and then by alpha. Controlling for our holdings-based measure, the average spread between the top and bottom alpha quintiles is always less than 1% per year and never significant. Controlling for our trade-based measure, the average spread is 1-2% per year, significant in half of the cases. Alpha therefore seems to contain some unique information as well, but it seems less informative than our measures.

The highest risk-adjusted returns in our tables are obtained by strategies that combine the information in alpha and in our measures. Consider the (5,5)-(1,1) portfolio, which buys the funds in the top quintiles according to both alpha and our holdings-based measure, and sells the funds in the bottom quintiles. The Fama-French alpha of this portfolio is 8.52% per year (t = 3.99), and the four-factor alpha (after adjusting for momentum) is 5.24% (t = 2.47). For the trade-based measure, the two alphas of the (5,5)-(1,1) portfolio are 6.54% (t = 3.72) and 4.73% per year (t = 2.48). These alphas are higher than those obtained by one-way quintile sorts, suggesting that combining alpha with our measures is useful in predicting future fund performance. Similar results are obtained when portfolios are formed with a one-quarter delay after portfolio disclosure dates, which allows portfolio holdings to be publicly observable.

We also explore whether investors respond to the information contained in our measures. The investors' response to alphas, in the form of fund flows, is well documented in the literature (e.g. Sirri and Tufano, 1998). To examine whether investors respond to our measures, we measure net flows into portfolios of funds double-sorted according to alpha and our measures. Fund flows respond strongly to fund alphas, confirming the earlier findings, but they do not respond significantly to our measures. Investors thus seem unaware of the information contained in our measures.

The paper proceeds as follows. Section I introduces our performance measures. Section II discusses a simulation exercise that evaluates the usefulness of these measures in capturing true skill. Section III implements the measures empirically to investigate predictability in the returns of U.S. equity funds. Section IV concludes.

I New Performance Measures

This section introduces two performance measures that judge a fund manager's ability by the extent to which his stock holdings (Section I.1) or stock trades (Section I.2) overlap with those of managers whose other investments have been successful.

I.1 A Measure Based on Levels of Holdings

Assume there are M managers, $m = 1, \ldots, M$, and N stocks, $n = 1, \ldots, N$, each of which is held by at least one manager. Let α_m denote the reference measure of skill of manager m (to be discussed in more detail), and let $w_{m,n}$ denote the current weight on stock n in manager m's portfolio. For each stock n, define its quality measure $\bar{\delta}_n$ as

$$\bar{\delta}_n = \sum_{m=1}^{M} v_{m,n} \alpha_m, \tag{1}$$

where

$$v_{m,n} = \frac{w_{m,n}}{\sum_{m=1}^{M} w_{m,n}}. \tag{2}$$

The quality of stock n is defined as the average skill of all managers who hold stock n in their portfolios, weighted by how much of the stock they hold. Stocks with high quality are those that are held mostly by highly-skilled managers. Managers who hold stocks of high quality are likely to be skilled, because their investment decisions are similar to those of other skilled managers (i.e. such managers are in "good company"). Since a larger position in a stock of given quality reveals more about the manager's ability, the population version of our performance measure is constructed as

$$\delta^*_m = \sum_{n=1}^{N} w_{m,n} \bar{\delta}_n. \tag{3}$$

In words, our measure of a manager's performance is the average quality of all stocks in the manager's portfolio, where each stock contributes according to its portfolio weight.

Our measure δ^*_m is defined in relation to some reference measure of skill, α_m. We choose the population value of Jensen's alpha, defined as the intercept from the regression of manager m's excess returns on the returns of the appropriate benchmark.[2] Note, however, that alpha is only one of many possible choices for α_m. Other sensible choices

[2] The benchmark is assumed to capture any style effects for which the manager should not be rewarded. This paper does not provide any new insights into the choice of the appropriate benchmark. Also note that replacing α_m in equation (1) by δ^*_m results in a different but unsatisfactory measure in equation (3) because such a circularly defined measure can be shown to be equal across all managers.

include the characteristic-based measure of Daniel, Grinblatt, Titman, and Wermers (1997), or the holding-based measure of Grinblatt and Titman (1993), for example. We opt for the traditional alpha measure solely in the interest of simplicity.

To construct our estimator of managerial skill, we replace α_m in equation (1) by $\hat{\alpha}_m$, the usual OLS estimator of alpha:

$$\hat{\delta}^*_m = \sum_{n=1}^{N} w_{m,n} \bar{\bar{\delta}}_n, \tag{4}$$

where

$$\bar{\bar{\delta}}_n = \sum_{m=1}^{M} v_{m,n} \hat{\alpha}_m. \tag{5}$$

Our performance measure has an interesting alternative interpretation. To obtain it, first let h_n denote the ratio of the dollar value of stock n held by all M managers to the total dollar value of all stocks held by these managers. Then

$$h_n = \frac{\sum_{m=1}^{M} w_{m,n}}{\sum_{n=1}^{N} \sum_{m=1}^{M} w_{m,n}} = \frac{\sum_{m=1}^{M} w_{m,n}}{M}, \tag{6}$$

a relation that will be useful shortly. Let W denote the $N \times M$ matrix whose (n,m) element is $w_{m,n}$, let V denote the $N \times M$ matrix whose (n,m) element is $v_{m,n}$, let $\hat{\alpha}$ denote the $M \times 1$ vector of $\{\hat{\alpha}_m\}_{m=1}^{M}$, let $\bar{\bar{\delta}}$ denote the $N \times 1$ vector of $\{\bar{\bar{\delta}}_n\}_{n=1}^{N}$, and let $\hat{\delta}^*$ denote the $M \times 1$ vector of $\{\hat{\delta}^*_m\}_{m=1}^{M}$. Since $\hat{\delta}^* = W'\bar{\bar{\delta}}$ and $\bar{\bar{\delta}} = V\hat{\alpha}$, the vector of our performance measures from equation (4) can be written as

$$\hat{\delta}^* = Z\hat{\alpha}, \tag{7}$$

where $Z = W'V$. The performance of manager m is thus

$$\hat{\delta}^*_m = \sum_{j=1}^{M} z_{m,j} \hat{\alpha}_j, \tag{8}$$

where $z_{m,j}$ is the (m,j) element of Z. Using equations (2) and (6), we have

$$z_{m,j} = \sum_{n=1}^{N} w_{m,n} v_{j,n} = \frac{1}{M} \sum_{n=1}^{N} w_{m,n} w_{j,n} \frac{1}{h_n}. \tag{9}$$

Our measure of manager m's skill, $\hat{\delta}^*_m$, is a weighted average of the usual skill measures across all managers.[3] The weight assigned to the performance of manager j, $z_{m,j}$, is a loose measure of covariance between the weights of managers m and j. This is sensible – if managers m and j own many of the same stocks, they may be using similar strategies, and we want to pay a lot of attention to the performance of manager j when evaluating manager m. The scaling factor $1/h_n$ in equation (9) downweights stocks that receive large weights on average in the portfolios of all managers. For example, if a certain stock occupied 20% of the market capitalization, many managers would have large weights in that stock, and its contribution to the performance measures would be overstated in the absence of the scaling factor. Similarly, if both managers hold a lot of a stock that few others hold, that is valuable information and the scaling factor emphasizes it.

Our approach adds value only if there is some commonality in the managers' investment decisions. Suppose manager m holds only stocks that are held by no other manager. In that case, $\hat{\delta}^*_m$ in equation (4) collapses to $\hat{\alpha}_m$, and our measure does not add any information beyond the traditional measure $\hat{\alpha}_m$.

Due to the symmetry of Z ($z_{i,j} = z_{j,i}$), which follows from equation (9), it is easy to show that the averages of $\hat{\delta}^*_m$ and $\hat{\alpha}_m$ across managers are equal:

$$\frac{1}{M}\sum_{i=1}^{M}\hat{\delta}^*_i = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{M} z_{i,j}\hat{\alpha}_j = \frac{1}{M}\sum_{j=1}^{M}\hat{\alpha}_j \sum_{i=1}^{M} z_{i,j} = \frac{1}{M}\sum_{j=1}^{M}\hat{\alpha}_j. \tag{10}$$

As a result, our skill measure provides only as much information as the usual measure about the performance of the mutual fund industry as a whole. Nevertheless, our measure can be useful in evaluating the funds' relative performance, as documented below.

The precision of our estimators can be assessed by their squared standard errors, which are displayed along the main diagonal of the $M \times M$ covariance matrix of $\hat{\delta}^*$. This matrix can be computed from equation (7) as

$$\text{Cov}(\hat{\delta}^*, \hat{\delta}^{*\prime}) = Z\Omega Z', \tag{11}$$

[3] The weights sum to one:

$$\sum_{j=1}^{M} z_{m,j} = \sum_{j=1}^{M}\sum_{n=1}^{N} w_{m,n} w_{j,n}\frac{1}{Mh_n} = \sum_{n=1}^{N}\frac{1}{Mh_n} w_{m,n}\sum_{j=1}^{M} w_{j,n} = \sum_{n=1}^{N} w_{m,n} = 1.$$

where $\Omega = \text{Cov}(\hat{\alpha}, \hat{\alpha}')$ is the $M \times M$ covariance matrix of $\hat{\alpha}$. If all funds had return histories spanning the same time period, Ω could in principle be calculated from one big multivariate regression of fund excess returns on benchmark returns. However, funds have histories of unequal lengths spanning different periods. The appendix describes how Ω can be calculated in such an environment.

To assess the gains in precision provided by our measure, write the squared standard error of our estimator for manager m as

$$\text{Var}(\hat{\delta}_m^*) = z_m \Omega z_m' = \sum_{i=1}^{M} \sum_{j=1}^{M} z_{m,i} z_{m,j} \omega_{i,j}, \tag{12}$$

where z_m is the m-th row of Z and $\omega_{i,j}$ is the (i,j) element of Ω. For simplicity, assume that all m elements of $\hat{\alpha}$ have the same standard error, so that $\omega_{i,i} = \omega^2$ for all $i = 1, \ldots, M$. Then $\omega_{i,j} = \omega^2 \rho_{i,j}$, where $\rho_{i,j}$ is the correlation between $\hat{\alpha}_i$ and $\hat{\alpha}_j$. Also assume that all $z_{m,i} > 0$, which is likely to hold in practice for most pairs of funds. Then

$$\begin{aligned} \text{Var}(\hat{\delta}_m^*) &= \sum_{i=1}^{M} z_{m,i}^2 \omega^2 + \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} z_{m,i} z_{m,j} \omega^2 \rho_{i,j} & (13) \\ &\leq \omega^2 \left(\sum_{i=1}^{M} z_{m,i}^2 + \sum_{i=1}^{M} \sum_{j=1, j\neq i}^{M} z_{m,i} z_{m,j} \right) = \omega^2 \left(\sum_{i=1}^{M} z_{m,i} \right)^2 = \omega^2, & (14) \end{aligned}$$

because we already showed that the rows of Z sum to one. This means that so long as the $\hat{\alpha}_m$'s are not perfectly correlated, $\hat{\delta}_m^*$ has a lower standard error than $\hat{\alpha}_m$.[4] Our gains in precision relative to $\hat{\alpha}$ therefore come from the imperfect correlations between the $\hat{\alpha}_m$'s. These correlations tend to be low when the cross-sectional correlation of the managers' residual returns is low. Thus, increasing the number of benchmarks that define $\hat{\alpha}$ can improve the precision of our estimator not only by increasing the precision of $\hat{\alpha}$, but also by reducing the residual correlations among funds.

The fact that $\hat{\delta}^*$'s tend to have substantially lower standard errors than $\hat{\alpha}$'s was confirmed empirically in the NBER version of this paper. In a large sample of U.S. equity funds, we found $\hat{\delta}^*$ to be about four to eight times more precise than $\hat{\alpha}$, on average, and to be more precise for 93 to 98 percent of all funds across three benchmark

[4]If all elements of $\hat{\alpha}$ do not have the same standard error, this relation holds only on average and there are likely to be some m's for which it does not hold. The calculation in equation (14) is analogous to computing the variance of a portfolio of stocks (with no short positions), which is typically smaller than the variance of any given stock in the portfolio.

models. The biggest precision gains from using our measures were obtained for short-history funds. For example, the correlations between the length of the fund's return history and the ratio of the squared standard error of $\hat{\delta}^*$ to the squared standard error of $\hat{\alpha}$ range from 0.45 to 0.56. The results for the changes estimator $\hat{\delta}^{**}$, which is described in the next section, are similar. We do not report these results in Section III simply to save space and to keep our empirical analysis focused on fund return predictability.

I.2 A Measure Based on Changes in Holdings

In the previous subsection, managers are inferred to be making similar investment decisions if they have similar holdings, regardless of the timing of their trades. In this subsection, we assume that managers make similar decisions if their trades are similar. Since trading involves transaction costs, decisions to trade are likely to reflect stronger views than decisions to passively hold. The performance measure developed here exploits similarities between changes in the managers' holdings, rather than their levels.

The return on the portfolio of manager m at time t can be written as

$$R_{m,t} = \sum_{n=1}^{N} w_{m,n} r_{n,t}, \tag{15}$$

where $r_{n,t}$ denotes the return on stock n. Define the change in the weights as

$$d_{m,n} = w_{m,n,t} - w_{m,n,t-1} \frac{1 + r_{n,t}}{1 + R_{m,t}}, \tag{16}$$

which is the difference between the current weight and the weight obtained if the manager neither bought nor sold any of this stock over the past period.[5] Let $\mathcal{N}_m^+ = \{n : d_{m,n} > 0\}$ denote the set of stocks purchased by manager m between $t-1$ and t and $\mathcal{N}_m^- = \{n : d_{m,n} < 0\}$ denote the set of stocks sold by manager m over the same time period. Analogously, let $\mathcal{M}_n^+ = \{m : d_{m,n} > 0\}$ denote the set of managers who made net purchases of stock n between $t-1$ and t, and let $\mathcal{M}_n^- = \{m : d_{m,n} < 0\}$ denote the set of managers with net sales. We normalize the changes in the weights as

$$x_{m,n}^+ = \frac{d_{m,n}}{\sum_{n \in \mathcal{N}_m^+} d_{m,n}}, \qquad x_{m,n}^- = \frac{d_{m,n}}{\sum_{n \in \mathcal{N}_m^-} d_{m,n}}, \tag{17}$$

[5] In our empirical analysis, one period equals one quarter. If the manager did not trade at all over the past quarter, so that $d_{m,n} = 0$ for all n, our measure is undefined for this manager. Also note that the time subscripts on d, w, and some related measures below are suppressed to simplify notation.

$$y^+_{m,n} \;=\; \frac{d_{m,n}}{\sum_{m \in \mathcal{M}^+_n} d_{m,n}}, \qquad y^-_{m,n} = \frac{d_{m,n}}{\sum_{m \in \mathcal{M}^-_n} d_{m,n}}, \tag{18}$$

so that $x^+_{m,n}$ ($x^-_{m,n}$) captures the fraction of manager m's purchases (sales) accounted for by stock n, and $y^+_{m,n}$ ($y^-_{m,n}$) captures the fraction of purchases (sales) of stock n accounted for by manager m.

For each stock n, define its quality measure $\bar{\delta}_n$ as

$$\bar{\delta}_n \;=\; \bar{\delta}^+_n - \bar{\delta}^-_n, \tag{19}$$

where

$$\bar{\delta}^+_n \;=\; \sum_{m \in \mathcal{M}^+_n} y^+_{m,n} \hat{\alpha}_m \tag{20}$$

$$\bar{\delta}^-_n \;=\; \sum_{m \in \mathcal{M}^-_n} y^-_{m,n} \hat{\alpha}_m, \tag{21}$$

and $\hat{\alpha}_m$ is the usual performance measure, acting as a reference measure again. (Using α_m in place of $\hat{\alpha}_m$ yields the population version of our trade-based skill measure, δ^{**}_m.) The quality of stock n is the difference between the average skill of all managers who bought stock n recently ($\bar{\delta}^+_n$) and the average skill of all managers who sold stock n recently ($\bar{\delta}^-_n$), where the averages are weighted by how much was bought or sold. Stocks of high quality are those that were recently bought mostly by high-skill managers and sold mostly by low-skill managers. Managers who recently bought high-quality stocks and sold low-quality stocks are likely to be skilled, because their decisions are similar to those of other skilled managers. Hence, our modified skill measure is

$$\hat{\delta}^{**}_m \;=\; \hat{\delta}^+_m - \hat{\delta}^-_m, \tag{22}$$

where

$$\hat{\delta}^+_m \;=\; \sum_{n \in \mathcal{N}^+_m} x^+_{m,n} \bar{\delta}_n \tag{23}$$

$$\hat{\delta}^-_m \;=\; \sum_{n \in \mathcal{N}^-_m} x^-_{m,n} \bar{\delta}_n. \tag{24}$$

This is the difference between the average quality of the stocks recently bought by manager m and the average quality of the stocks recently sold by this manager. Note that $\hat{\delta}^+_m$ captures to what extent the manager has been buying high-quality stocks and $\hat{\delta}^-_m$

captures to what extent the manager has been selling low-quality stocks. Our measure combines these two aspects of stock-picking skill.

To get more insight into the new measure, we write it out as

$$\begin{aligned}\hat{\delta}_m^{**} &= \sum_{n\in\mathcal{N}_m^+} x_{m,n}^+ \Big(\sum_{j\in\mathcal{M}_n^+} y_{j,n}^+\hat{\alpha}_j - \sum_{j\in\mathcal{M}_n^-} y_{j,n}^-\hat{\alpha}_j\Big) - \sum_{n\in\mathcal{N}_m^-} x_{m,n}^- \Big(\sum_{j\in\mathcal{M}_n^+} y_{j,n}^+\hat{\alpha}_j - \sum_{j\in\mathcal{M}_n^-} y_{j,n}^-\hat{\alpha}_j\Big) \\ &= \sum_{j=1}^{M} c_{m,j}\hat{\alpha}_j, \qquad (25)\end{aligned}$$

where

$$\begin{aligned}c_{m,j} &= \sum_{n=1}^{N}\Big[x_{m,n}^+ y_{j,n}^+ 1_{\{n\in\mathcal{N}_m^+\}}1_{\{j\in\mathcal{M}_n^+\}} - x_{m,n}^+ y_{j,n}^- 1_{\{n\in\mathcal{N}_m^+\}}1_{\{j\in\mathcal{M}_n^-\}} \cdots \\ &\quad - x_{m,n}^- y_{j,n}^+ 1_{\{n\in\mathcal{N}_m^-\}}1_{\{j\in\mathcal{M}_n^+\}} + x_{m,n}^- y_{j,n}^- 1_{\{n\in\mathcal{N}_m^-\}}1_{\{j\in\mathcal{M}_n^-\}}\Big], \qquad (26)\end{aligned}$$

and $1_{\{\}}$ denotes an indicator function equal to one or zero depending on whether the associated condition is true. The new measure is therefore a "weighted average" of the usual measures across all managers. The quotation marks are appropriate because the weights sum to zero, $\sum_{j=1}^{M} c_{m,j} = 0$, as shown in the appendix. The weight on manager j essentially reflects the covariance of the weight changes of managers m and j. This weight, $c_{m,j}$, is a sum of N terms, one for each stock, which capture the products of the managers' weight changes in that stock. If both managers traded that stock in the same direction (i.e. if both bought or sold it), this product is positive and increasing in the extent of the joint action. If one manager bought and the other sold, the product is negative. Hence the loose covariance interpretation. The higher the covariance of the weight changes of managers m and j, the more attention we want to pay to the skill of manager j when evaluating the skill of manager m.[6]

To summarize the modified measure, let C denote the $M \times M$ matrix whose (i,j) element is $c_{i,j}$. Also, let X^+ and X^- denote $M \times N$ matrices whose (i,j) elements are $x_{i,j}^+$ and $x_{i,j}^-$, respectively. Similarly, let Y^+ and Y^- denote $M \times N$ matrices whose (i,j) elements are $y_{i,j}^+$ and $y_{i,j}^-$, respectively. Zeros are substituted for all (i,j) elements of X^+ and Y^+ such that $d_{i,j} < 0$, as well as for all (i,j) elements of X^- and Y^- such that $d_{i,j} > 0$. With this notation, C can be expressed as

$$C = X^+(Y^+)' - X^+(Y^-)' - X^-(Y^+)' + X^-(Y^-)'. \qquad (27)$$

[6] The average of $\hat{\delta}_m^{**}$ across managers can be shown to equal zero in the special case when managers trade stocks only with each other. Since managers generally trade also with other entities such as individuals, the average $\hat{\delta}_m^{**}$ is likely to be close to zero but not exactly zero.

Then the $M \times 1$ vector of our performance measures in equation (22) can be written as

$$\hat{\delta}^{**} = C\hat{\alpha}, \tag{28}$$

and the precision of the changes measure can be calculated as

$$\text{Cov}(\hat{\delta}^{**}, \hat{\delta}^{**\prime}) = C\Omega C', \tag{29}$$

where $\Omega = \text{Cov}(\hat{\alpha}, \hat{\alpha}')$ as before. Henceforth, we often refer to the measure $\hat{\delta}^{*}$ as the "levels" measure and to the measure $\hat{\delta}^{**}$ as the "changes" measure.

Our changes measure in equation (22) weighs stock qualities by the relative magnitudes of the weight changes across stocks. There is an interesting alternative definition, $\hat{\delta}_m^{**A} = \sum_{n=1}^{N} d_{m,n}\bar{\delta}_n$, which instead relies on the absolute magnitudes of these changes. This alternative changes measure is slightly inferior to $\hat{\delta}^{**}$ in our simulations, but its predictive power for fund returns in one-way sorts is similar to that of $\hat{\delta}^{**}$. We focus on $\hat{\delta}^{**}$ mostly due to its appealing interpretation as the difference between the average qualities of stocks bought and sold.

I.3 Some Issues Related to Our Measures

First, note that our performance measures are not "optimal" in the sense of being solutions to an optimization problem. It seems impossible to design the relevant optimization problem without putting some structure on the nature of commonality in skill across managers. Since the true structure of this commonality is unknown, we do not pursue optimality. Instead, we propose measures that capture the underlying intuition in a way that should be reasonably robust to various forms of this commonality.

Our changes measure exploits commonality in trades across managers. The extent to which funds buy and sell stocks at the same time is analyzed by Lakonishok, Shleifer, and Vishny (1992), Grinblatt, Titman, and Wermers (1995), and Wermers (1999), among others. However, this literature on "herding" does not propose using the degree of similarity in the managers' trades to evaluate their performance. Hong, Kubik, and Stein (2002) show that fund managers from the same city are more likely to hold, buy, or sell the same stock at the same time than are managers from different cities. Given this evidence, managers from the same city are likely to be assigned more similar skill according to our measures than managers from different cities. This seems appropriate. Hong, Kubik, and Stein attribute their finding to word-of-mouth information diffusion, and it seems sensible to assign similar skill to managers using similar information.

Given the design of the proposed measures, it seems worthwhile to relate this paper to the literature on "window dressing" (e.g. Haugen and Lakonishok, 1988, Lakonishok et al., 1991, Musto, 1997, 1999, and Carhart et al., 2002). Managers who reshuffle their portfolios around disclosure dates presumably believe that they are judged by the quality of their disclosed portfolios. For example, holdings of unusually risky assets or assets with poor prior performance might lead investors to infer weak managerial ability. In this paper, managers are essentially also judged by the quality of their portfolios, but this quality is assessed differently – by the extent to which these portfolios resemble the portfolios of other well-performing managers. In our approach, a fund manager may be viewed favorably even if he holds stocks that performed poorly recently. This approach seems valuable to rational investors, as discussed in the introduction. In contrast, the literature does not provide a definitive explanation as to why window dressing per se should affect the beliefs of rational investors who can observe the funds' track records.

One possible reason behind window dressing is that managers think they are evaluated by the company they keep. Suppose the stock of Cisco did well recently. A manager might buy Cisco before disclosure date because she wants her disclosed portfolio to look similar to the portfolios of the successful managers who held Cisco while its stock went up. This tactic may not work, however, because the other managers may have sold Cisco in the meantime. Moreover, this tactic does not confuse our changes measure because the purchases of Cisco are not concurrent across managers. Also note that the window dressing literature has focused on various asset pricing issues (e.g. the January effect), but it has not addressed performance evaluation. In short, while there are some similarities between this paper and the window dressing literature, there are also important differences in motivation, focus, and implications for rational investors.

II Simulations

The purpose of the simulation analysis is to assess the usefulness of our estimators in capturing true managerial skill, which is conveniently known in our simulated environment.

Managers receive private signals about each stock's expected return, and a manager's true skill is measured by the probability that the signals he receives are useful. When managers are ranked by their true skill and separately by various performance estimators, useful information about estimator quality is provided by the resulting rank order correlations. As shown below, our estimators contain important information about true skill that is not captured by standard estimators that make no use of similarities in holdings or trades across managers. The biggest benefits from using our estimators are obtained for funds with relatively short return histories.

II.1 Simulation Design

The simulations consider a simple setting in which M managers receive signals about expected excess returns of N stocks. Stock n's excess return at time t is

$$r_{n,t} = \mu_{n,t} + e_{n,t}, \quad n = 1, \ldots, N; \quad t = 1, \ldots, T \tag{30}$$

where $\mu_{n,t}$ is the stock's expected excess return and $e_{n,t}$ is an error term. We simulate $\mu_{n,t}$ from $N(0, \sigma_\mu^2)$ and $e_{n,t}$ from $N(0, \sigma_e^2)$, independently across time and stocks.

In every period t, each manager m receives a signal $s_{m,n,t}$ about each stock n. With probability γ_m, this signal is equal to the stock's true expected excess return, and otherwise it is equal to a noise term drawn from an identical distribution:

$$s_{m,n,t} = \begin{cases} \mu_{n,t} & \text{with probability } \gamma_m \\ u_{n,t} & \text{with probability } 1 - \gamma_m \end{cases} \tag{31}$$

where $u_{n,t} \sim N(0, \sigma_\mu^2)$. Higher γ_m means higher signal quality, so γ_m captures manager m's true skill. The γ_m's are drawn independently from the standard uniform distribution, and they are held constant across the T periods.

The managers know the signal structure in equation (31) as well as their own skill γ_m and the error volatility σ_e. They have no information about $\mu_{n,t}$ other than the signal, and they learn about $\mu_{n,t}$ from the signal using the Bayes rule. As shown in the appendix, the expected excess return on stock n perceived by manager m at time t is

$$E_{m,n,t} = \gamma_m s_{m,n,t}, \tag{32}$$

the perceived variance of stock n's returns is

$$V_{m,n,t} \quad = \quad \sigma_e^2 + \sigma_\mu^2 + \gamma_m (s_{m,n,t}^2 - \sigma_\mu^2) - s_{m,n,t}^2 \gamma_m^2, \tag{33}$$

and the perceived return covariance matrix is diagonal. The managers are assumed to maximize the Sharpe ratios of their portfolios, so that manager m's weight in stock n is

$$w_{m,n,t} \propto V_{m,n,t}^{-1} E_{m,n,t}. \tag{34}$$

As mutual funds are typically not allowed to short, we also require no short sales. This constraint can be imposed simply by putting a zero weight on any stock with a negative signal, thanks to our assumption of uncorrelated stock returns.

At each period t, we record each manager's realized excess return $r_t^m = w_{m,t}' r_t$ as well as expected excess return $\alpha_{m,t} = w_{m,t}' \mu_t$, where $w_{m,t}$ is the $N \times 1$ vector of manager m's portfolio weights, r_t is the vector of $r_{n,t}$'s, and μ_t is the vector of $\mu_{n,t}$'s.[7] Then

$$\alpha_m = \frac{1}{T} \sum_{t=1}^{T} \alpha_{m,t} = \frac{1}{T} \sum_{t=1}^{T} w_{m,t}' \mu_t, \tag{35}$$

denotes manager m's average expected excess return, which we refer to as the traditional measure of true performance. We calculate four performance estimators. The traditional estimator of α, $\hat{\alpha}$, is simply the manager's average realized excess return:

$$\hat{\alpha}_m = \frac{1}{T} \sum_{t=1}^{T} r_t^m = \frac{1}{T} \sum_{t=1}^{T} w_{m,t}' r_t. \tag{36}$$

Our performance measure based on levels of holdings, $\hat{\delta}^*$, is calculated as

$$\hat{\delta}_m^* = Z_m \hat{\alpha}, \tag{37}$$

where Z_m is defined as row m of matrix Z in equation (7), and the weight matrix W contains the managers' weights in the last (T-th) period.[8] Our performance measure based on changes in holdings, $\hat{\delta}^{**}$, is calculated as

$$\hat{\delta}_m^{**} = C_m \hat{\alpha}, \tag{38}$$

where C_m is defined as row m of matrix C in equation (28), and the weight changes are calculated between periods $T-1$ and T. For comparison purposes, we also construct a simple Bayesian estimator that shrinks $\hat{\alpha}$ towards a common mean:

$$\hat{\alpha}_m^B = \frac{1}{2} \hat{\alpha}_m + \frac{1}{2} \bar{\hat{\alpha}}, \tag{39}$$

[7]For simplicity, we use no benchmark in the simulations, so that excess returns coincide with abnormal returns. Simulating benchmark returns seems like an unnecessary complication that would unlikely make much difference, other than to make our assumption of uncorrelated e_n more realistic.

[8]If there happens to be a stock that is not held by any manager in a given simulated sample, the corresponding row of W is deleted in calculating $\hat{\delta}^*$, to prevent division by zero in equation (2).

where $\bar{\hat{\alpha}}$ is the average of $\hat{\alpha}_m$'s across managers. The four estimators rely only on holdings and return information. We also compute the population versions of δ^* and δ^{**}:

$$\delta_m^* = Z_m \alpha \tag{40}$$

$$\delta_m^{**} = C_m \alpha, \tag{41}$$

where α is the $N \times 1$ vector of the true α_m's. These measures possess an unfair advantage over the estimators, as they reflect information that is unknown outside the simulated world, but they help us assess the maximum potential gains from using our estimators.

We conduct 10,000 simulations for each set of parameter values. The number of managers $M = 30, 100$, and 300, the number of stocks $N = 30$ and 100, and the number of annualized time periods $T = 1, 5, 10, 20$, and 30. Throughout, $\sigma_\mu = 0.1$ and $\sigma_e = 0.5$.[9] In each simulated sample, we calculate α, $\hat{\alpha}$, $\hat{\alpha}^B$, δ^*, δ^{**}, $\hat{\delta}^*$, and $\hat{\delta}^{**}$ for each manager. The managers are ranked according to each of these measures as well as according to their true skill γ for the purpose of computing the rank order correlations.

II.2 Simulation Results

Table 1 judges the ability of various performance measures to imitate the ranking of managers by their true skill γ. The table reports the rank order correlations of each estimator with γ, averaged across 10,000 simulations. For $T = 1$ year, both of our estimators, $\hat{\delta}^*$ and $\hat{\delta}^{**}$, deliver higher rank correlations with γ than the traditional estimator $\hat{\alpha}$. For example, with 300 managers and 30 stocks, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ deliver rank correlations of 0.44 and 0.45, respectively, whereas $\hat{\alpha}$ delivers only 0.27. The population measures δ^* and δ^{**} also outperform α. Our measures are particularly effective when the number of managers is large, as there is more cross-sectional information to pool, but they beat $\hat{\alpha}$ even for $M = 30$. As the number of stocks N increases, the managers' portfolios become better diversified and it becomes easier for all measures to detect skill. $\hat{\delta}^{**}$, which uses holdings information not only from period T but also $T - 1$, outperforms $\hat{\delta}^*$.

The length of the sample period matters in judging the estimators' relative success. For $T \in \{5, 10, 20\}$ years, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ still outperform $\hat{\alpha}$, despite the fact that the population measures δ^* and δ^{**} are outperformed by α; this is due to the higher precision

[9]The average annual return volatility across all 14,149 firms with at least 60 months of contiguous returns on CRSP between January 1926 and December 2002 is 0.5546. The median is 0.4957.

in estimating δ^* and δ^{**}. For $T = 30$, $\hat{\alpha}$ already beats $\hat{\delta}^*$ but it still underperforms $\hat{\delta}^{**}$. For $T > 30$ years, $\hat{\alpha}$ beats both of our estimators. This is not surprising – as $T \to \infty$, a manager's skill is best inferred from his average realized return, since true skill is held constant over time in the simulations. Therefore, the biggest benefits from using our estimators as opposed to $\hat{\alpha}$ are obtained for funds with relatively short return histories. But even for funds with return histories of up to 30 years or so, the noise in $\hat{\alpha}$ is so large that true skill is better captured by $\hat{\delta}^*$ and $\hat{\delta}^{**}$.

The main goal of this section, accomplished in Table 1, is to explore how various skill measures capture true skill γ. In addition, Table 2 evaluates the ability of these measures to capture α, which is commonly estimated in the literature. Panel A of Table 2 reports the average rank correlations of each estimator with α. The conclusions are similar to those from Table 1. For $T \leq 30$ years or so, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ achieve higher rank correlations with α than the traditional estimator $\hat{\alpha}$, despite the fact that $\hat{\alpha}$ is designed to capture α but $\hat{\delta}^*$ and $\hat{\delta}^{**}$ are designed to capture their own population quantities, δ^* and δ^{**}. For example, with $T = 10$ and $M = N = 30$, both $\hat{\delta}^*$ and $\hat{\delta}^{**}$ deliver rank correlations of 0.77 with α, whereas $\hat{\alpha}$ delivers only 0.68.[10] For $T > 30$, $\hat{\alpha}$ wins, as in Table 1. $\hat{\delta}^{**}$ beats $\hat{\delta}^*$ in most cases. The population measures δ^* and δ^{**} attain high rank correlations (sometimes as high as 0.99), which present an upper bound on how close we can get to α with our estimators. To sum up, our estimators $\hat{\delta}^*$ and $\hat{\delta}^{**}$ seem successful at imitating the rankings based not only on γ but also on α.

Panel B of Table 2 reports the mean squared errors (MSEs) in estimating α for all performance measures.[11] The MSEs are calculated as averages across managers and simulations of the squared differences between α and its estimate. As expected, the MSE of the Bayesian estimator $\hat{\alpha}^B$ is lower than the MSE of $\hat{\alpha}$ unless T is large. More interestingly, for $N = 30$ and $T < 30$, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ both have lower MSE than $\hat{\alpha}$, and $\hat{\delta}^{**}$ beats even $\hat{\alpha}^B$, despite the fact that it is not designed to capture α but δ^{**}. As T and N increase, the lowest MSE is ultimately achieved by $\hat{\alpha}$, of course. The MSEs of δ^* and δ^{**} reveal the extent to which our population measures depart from the true α. Our estimators have high precision, which makes them preferred for low T's, but they also exhibit some bias with respect to α, so the overall effect on the MSE is unclear.

[10] Note that the Bayesian estimator $\hat{\alpha}^B$ delivers the same rank correlations as $\hat{\alpha}$, by construction.

[11] Calculating the MSEs in Table 1 would not be sensible, as the units of γ and α are different.

Figure 1 plots the bias in estimating α for four performance estimators, $\hat{\alpha}$, $\hat{\alpha}^B$, $\hat{\delta}^*$, and $\hat{\delta}^{**}$. In each simulated sample, managers are ranked in ascending order by their α. Let α_m denote the true alpha of the m-th ranked manager, and $\tilde{\delta}_m$ denote that manager's estimated performance (i.e. $\tilde{\delta}_m = \hat{\alpha}_m$, $\hat{\alpha}_m^B$, $\hat{\delta}_m^*$, or $\hat{\delta}_m^{**}$). For each rank m ($m = 1, \ldots, M$), the bias is computed by averaging $\tilde{\delta}_m - \alpha_m$ across 10,000 simulated samples. The figure plots the bias against the rank m, with four panels based on different values of M and T. All panels show that the traditional estimator $\hat{\alpha}$ is unbiased, but both the Bayesian estimator $\hat{\alpha}^B$ and our levels estimator $\hat{\delta}^*$ are biased. For the managers with above-(below-)average skill, both estimators are biased downwards (upwards), which is entirely expected as a result of the shrinkage of the individual $\hat{\alpha}_m$'s toward their common mean. This shrinkage is explicit in $\hat{\alpha}^B$ and implicit in $\hat{\delta}^*$, which is a weighted average of the $\hat{\alpha}_m$'s across all managers (equation 8). The bias of $\hat{\delta}^*$ is slightly bigger than that of $\hat{\alpha}^B$. The changes estimator $\hat{\delta}^{**}$ is also biased, but unlike $\hat{\delta}^*$, it is biased even on average. This follows by design – while both $\hat{\delta}^*$ and $\hat{\delta}^{**}$ are weighted averages of the $\hat{\alpha}_m$'s, the weights used in $\hat{\delta}^*$ sum to one, but those in $\hat{\delta}^{**}$ sum to zero, as shown earlier.

The bias with respect to α suggests the need for caution when using our estimators to estimate α. It seems reasonable to use $\hat{\delta}^*$ for such a task, with similar costs and benefits as when using the Bayesian estimator, but $\hat{\delta}^{**}$ is not designed to capture α even on average. More importantly, the bias does not compromise the success of our estimators in ranking managers by their α or γ, as we saw in Tables 1 and 2. Our performance measures should thus be particularly useful in applications that involve such ranking, and one such application is presented in the following section.

The simulation results are robust to numerous modifications. First, the results are robust to reasonable changes in σ_μ and σ_e. All conclusions remain the same when the simulations are rerun with $\sigma_\mu = 0.05$ instead of 0.1 (i.e. 5% instead of 10% dispersion in annual expected excess stock returns); in fact, the results become stronger, as both of our estimators outperform $\hat{\alpha}$ even for $T = 30$. With $\sigma_e = 0.3$ instead of the data-implied value of 0.5, the conclusions are again unaffected except that $\hat{\alpha}$ outperforms $\hat{\delta}^*$ and $\hat{\delta}^{**}$ for $T > 10$ years. Less volatile returns improve the relative performance of $\hat{\alpha}$, and more volatile returns hurt $\hat{\alpha}$. Drawing γ_m's with a common component (as in the previous draft) also leads to very similar results.

Allowing stock returns to be correlated leads to the same conclusions, as we discovered by repeating the simulations with 0.1366 correlation between all pairs of stocks.[12] One complication in working with correlated returns is that inverting the covariance matrix of returns in equation (34) and especially ruling out short sales is computationally infeasible for our desired number of simulations and parameter values. We conduct two sets of simulations to get around this problem. First, we let each manager rank stocks by their signals and invest equally in the top 25% of all stocks, instead of using a mean-variance criterion. The results are quite similar to those reported; all rank correlations are slightly lower because managers using this suboptimal strategy do not take full advantage of their skill, but the comparative results are in fact stronger, as $\hat{\delta}^*$ and $\hat{\delta}^{**}$ outperform $\hat{\alpha}$ throughout, even for $T = 30$. (The results are very similar also when stock returns are uncorrelated.) Second, we conduct mean-variance simulations for $N = 30$ where we rule out short sales only approximately (by zeroing out negative weights), and the results are again very similar to those in Tables 1 and 2.

Our measures have one other useful property that is easy to verify by simulation. Imagine a world with a large number of managers, all of whom exhibit no skill ($\gamma = 0$, $\alpha = 0$) and whose holdings overlap sufficiently. The fact that there is no skill in this world would be hard to detect by $\hat{\alpha}$, because the managers' $\hat{\alpha}$'s would be dispersed around zero due to return noise. However, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ should be zero for all managers, because the manager-specific noise is diversified away by averaging across many $\hat{\alpha}$'s.

III Empirical Analysis

The simulation results suggest that our skill measures $\hat{\delta}^*$ and $\hat{\delta}^{**}$ can capture true skill better than the standard measure $\hat{\alpha}$. If that is the case, $\hat{\delta}^*$ and $\hat{\delta}^{**}$ should be at least as successful as $\hat{\alpha}$ in predicting fund returns. In this section, the relative predictive ability of all three measures is analyzed empirically in a large sample of equity mutual funds.

[12] 0.1366 is the average correlation across all 44,418,986 pairs of firms with at least 60 months of overlapping contiguous returns on CRSP between January 1926 and December 2002.

III.1 Data

Mutual fund returns are obtained from the Center for Research in Securities Prices (CRSP).[13] These returns are net of fees. Since we are interested in detecting true before-cost managerial ability, we add fees back to obtain gross fund returns. Specifically, we take the annual expense ratio and 12(b)1 fees given by CRSP, add them together, divide by 12, and add the resulting number to each monthly return in a given year. The stock data for all NYSE, AMEX, and NASDAQ stocks is obtained from the CRSP monthly stock file. A firm's simple stock return is computed as an annual value-weighted return on the firm's common stock issues (typically one). Delisting returns are included when available on CRSP. If a firm is delisted but the delisting return is missing, we investigate the reason for disappearance. If the delisting is performance-related, we assume a -30% delisting return, following Shumway (1997); otherwise a zero delisting return is assumed. Quarterly stock returns are compounded from monthly returns.[14] Market equity, measured as the combined value of all common stock classes outstanding, is taken from CRSP at the end of each month.

Mutual funds are required to file holdings reports with the SEC twice a year, but most funds publicly disclose their portfolio holdings on a quarterly basis. Thomson Financial collects and sells this data. We access the data through Wharton Research Data Services, which currently makes the mutual fund holdings data available for the period from the end of 1980Q1 through the end of 2002Q2. The data is commonly known as the Spectrum data, since it used to be collected by CDA/Spectrum prior to their purchase by Thomson.[15] The Spectrum mutual fund holdings file contains four columns: date, stock identifier CUSIP, fund identifier, and the number of shares of the given stock held by the given fund on the given date. All dates are quarter-ends (3/31, 6/30, 9/30, or 12/31). Firms with no Spectrum data are recorded as having zero mutual

[13] In the NBER version of this paper, fund returns were computed from the monthly returns on the stocks held by the fund at the most recent quarter-end, assuming that the fund follows a buy-and-hold strategy between the quarter-ends and that the quarterly rebalancing into the reported holdings is costless. Such "hypothetical" fund returns (Grinblatt and Titman, 1989) ignore any intraquarter trading by the fund, including potential window-dressing activity, as well as any ability of the fund to time the market by cleverly adjusting the fraction of its cash holdings. Using the hypothetical fund returns, we found return predictability that is significantly weaker than reported here.

[14] If any of the two monthly stock returns following a valid return at the beginning of the quarter is missing, it is assumed to be zero when computing the cumulative stock return for this quarter.

[15] See Wermers (1999) for detailed information regarding the construction of the database. This data is free of survival bias, as noted for example by Daniel, Grinblatt, Titman, and Wermers (1997).

fund ownership. We match each CUSIP to a CRSP PERMNO, the permanent number CRSP assigns to that security. Holdings associated with CUSIPs for which we found no associated PERMNO are ignored; these account for a very small fraction of holdings.

We merge the Spectrum holdings data with the CRSP mutual fund data via a hand-matching of the funds by name. To produce a reliable matching, we also compare total net assets invested in equities as reported in CRSP to the market value of the stock holdings reported in Spectrum. In matching funds, we often find that separate CRSP listings are merely different share classes of a single fund. In such cases, we select the share class with the greatest number of months of valid data in a given year; if there is no difference by this measure, we use the fund with the lower ICDI identifier.

One limitation of our fund holdings data is that changes in holdings are observed only on a quarterly basis. If Fund 2 mimics the trades of Fund 1 with a short delay, we are unable to discern such a pattern unless the delay appears in the funds' quarterly holdings. Both funds are likely to be assigned similar performance by our measures, which seems appropriate only if Fund 1 is unable to profit on its trades before those trades are mimicked by Fund 2. Useful evidence on the importance of this issue is provided by Chen, Jegadeesh, and Wermers (2000), who find that stocks that experience net purchases by mutual funds have higher subsequent returns than stocks that experience net sales. The authors find that the abnormal performance following the funds' aggregate trades lasts for about a year. Most of the profits from fund trades are therefore not short-lived enough to confuse our performance measures.

Another complication is that the quarter-end dates reported by Thomson do not always correspond to the actual dates when the portfolio snapshots are taken. This leads to a time mismatch for funds whose fiscal quarters differ from calendar quarters. Nonetheless, Wermers (1999, 2000) reports that a vast majority of mutual funds use a fiscal year whose quarters coincide with calendar quarters. Like Wermers, we use the approximation that all holdings reported within a given calendar quarter are valid for the end of that calendar quarter. Another problem is that the disclosed portfolios may differ from undisclosed portfolios due to window dressing (Musto, 1999). Since all of these complications add noise to our measures, they should make it more difficult for our measures to predict fund returns. Also note that we do not separate the identity of fund managers from the funds they manage. Baks (2002) constructs a database of fund manager returns and characteristics by tracking the managers' career moves from one fund to another, and uses it to examine fund manager performance. He concludes that the fund typically has a bigger effect on performance than the manager, which helps motivate our focus on funds rather than managers in our empirical work.

III.2 Predicting Mutual Fund Returns

At the beginning of each quarter, we compute the traditional measure alpha ($\hat{\alpha}$) for each fund with at least 12 monthly returns by regressing fund returns in excess of the risk-free rate on benchmark returns. If some returns are missing, we run the regression across the months in which returns are available. Using $\hat{\alpha}$ as a reference measure, we then calculate the estimators of our measures, $\hat{\delta}^*$ and $\hat{\delta}^{**}$, from equations (4) and (22).

Nine versions of each measure are computed, using three benchmark models and three "lookback" periods over which the above measures are calculated. The CAPM alpha is calculated with respect to the market benchmark, as implied by the capital asset pricing model (CAPM), the Fama-French alpha is calculated with respect to the market, size, and value benchmarks, following Fama and French (1993), and the four-factor alpha is calculated with respect to the market, size, value, and momentum benchmarks, following Carhart (1997).[16] The lookback periods are 12 months, 24 months, and the entire life of the fund. In the interest of space, we only report the results for the Fama-French and four-factor models and only the 12-month and entire-life lookback periods, but the results for the CAPM alphas and the 24-month lookback period are similar.

At the beginning of each quarter, funds are sorted into decile portfolios according to each performance measure. The returns on the decile portfolios are calculated over the three months following the portfolio formation, equal-weighting the funds within each decile. The three-month return series are linked across quarters to form a monthly series of returns on each decile portfolio, covering the period April 1982 through September 2002. (The holdings data begin at the end of 1980Q1 and we require two years of data to obtain the initial estimates.) Table 3 reports the post-ranking alphas of the decile portfolios as well as of the "10-1" portfolio, constructed by going long the top decile of the best past performers and short the bottom decile. Alpha is defined with respect to

[16] All benchmark returns together with the risk-free rate were obtained from Kenneth French's website.

the same benchmarks as the reference measure, e.g. the table reports the post-ranking Fama-French alphas of the portfolios sorted on the past Fama-French $\hat{\alpha}$'s as well as on our measures that use the Fama-French alpha as a reference measure.

All three measures seem capable of predicting fund returns. Consider Panel A of Table 3, in which the lookback period is 12 months. The Fama-French alpha of the 10-1 portfolio produced by $\hat{\alpha}$ is 5.2% per year, which is comparable to the 4.5% alpha produced by $\hat{\delta}^{**}$ but substantially smaller than the 7.4% alpha produced by $\hat{\delta}^{*}$. Not surprisingly, the persistence in performance weakens when the momentum benchmark is included.[17] Nonetheless, the four-factor alphas remain statistically significant for all three measures. As before, the highest alpha is produced by $\hat{\delta}^{*}$ (5.9%), while $\hat{\delta}^{**}$ and $\hat{\alpha}$ deliver alphas of 3.5% and 3.7%, respectively. Similar results are obtained in Panel B, in which the lookback period is the entire life of the fund. Note that the observed persistence is not restricted to poorly performing funds. In fact, the positive alphas of the top deciles are roughly 2-3 times larger in absolute value than the negative alphas of the bottom deciles. To summarize, all three measures seem able to predict fund returns, and the most predictive power is achieved by $\hat{\delta}^{*}$.[18]

Since our performance measures are correlated with $\hat{\alpha}$, it seems interesting to ask whether our measures contain information not contained in $\hat{\alpha}$ that is useful in forecasting fund returns. To address this question, we perform a series of conditional sorts. Each quarter, we first sort funds into quintiles based on their $\hat{\alpha}$, and then within each $\hat{\alpha}$ quintile, we sort funds into quintiles based on $\hat{\delta}^{*}$.[19] Panels A of Tables 4 and 5 report the benchmark-adjusted returns for the resulting 25 portfolios, as well as for five "5-1"

[17]Carhart (1997) argues that some apparent persistence in fund performance is due simply to momentum in stock returns. The argument is that, due to momentum, managers who happen to hold mostly stocks that performed well (poorly) over the past year are likely to do well (poorly) also over the following year, even in the absence of any rebalancing on their part. It seems hard to argue that sitting on one's laurels and doing nothing is a managerial skill that should be given credit. Note, however, that Carhart's tables reveal significant persistence in $\hat{\alpha}$ even after adjusting for momentum.

[18]Berk and Green (2002) argue that net fund returns should be unpredictable if investors compete for returns and if managerial ability exhibits decreasing returns to scale. We analyze gross returns.

[19]Sorting into deciles rather than quintiles leads to similar though slightly noisier results. Note that we conduct conditional rather than independent sorts because some cells in independent sorts contain few or no funds due to the correlation between $\hat{\alpha}$ and $\hat{\delta}^{*}$. In conditional sorts, this correlation leads to the potential concern that the second-step improvement we report may simply be obtained from a finer sort on $\hat{\alpha}$ (i.e. sorting each $\hat{\alpha}$ quintile again by $\hat{\alpha}$ into five subquintiles). However, the spreads in returns seem too large to be attributed to within-quintile spreads in $\hat{\alpha}$. Moreover, this concern is dismissed in the reverse sorts reported below.

portfolios that buy funds with high $\hat{\delta}^*$ and short funds with low $\hat{\delta}^*$ within a given $\hat{\alpha}$ decile. Our cleanest measure of whether $\hat{\delta}^*$ provides information beyond $\hat{\alpha}$ is provided by the "average" portfolio that invests equally in the five 5-1 portfolios.

These results suggest that $\hat{\delta}^*$ contains significant information about future fund returns above and beyond $\hat{\alpha}$. Controlling for $\hat{\alpha}$, the average difference between the top and bottom quintiles of funds ranked by $\hat{\delta}^*$ ranges from 2.4% to 4.4% per year, across both benchmark models and lookback periods. These quantities are significant not only economically but also statistically, with t-statistics ranging from 1.94 to 3.21.

We also examine how much information about future fund returns is contained in $\hat{\alpha}$ but not in $\hat{\delta}^*$. To do that, we sort funds into quintiles in reverse, first by $\hat{\delta}^*$ and then by $\hat{\alpha}$, and report the results in Panels B of Tables 4 and 5. Controlling for $\hat{\delta}^*$, the average 5-1 spread produced by $\hat{\alpha}$ is always less than 1% per year and never significant. This suggests that most of the information contained in $\hat{\alpha}$ is already in $\hat{\delta}^*$.

Tables 6 and 7 report the results for $\hat{\delta}^{**}$. The results are only slightly weaker than those for $\hat{\delta}^*$. Controlling for $\hat{\alpha}$, the average difference between the top and bottom quintiles of funds ranked by $\hat{\delta}^{**}$ ranges from 1.2% to 1.4% per year, with t-statistics ranging from 2.31 to 2.73. This means that $\hat{\delta}^{**}$ adds incremental information about future fund returns over and above $\hat{\alpha}$. Controlling for $\hat{\delta}^{**}$, the average 5-1 spread produced by $\hat{\alpha}$ ranges from 1.3% to 2.1% per year, significant in half of the cases. $\hat{\alpha}$ thus seems to contain some incremental information beyond $\hat{\delta}^{**}$.

Since our measures and $\hat{\alpha}$ seem to contain some incremental information relative to each other, it appears that mutual fund portfolio strategies would benefit from combining the information in these measures. Indeed, the highest risk-adjusted returns in our tables are typically offered by the (5,5)-(1,1) portfolio, which buys the funds in the top quintiles according to both $\hat{\alpha}$ and one of our measures, and sells the funds in the bottom quintiles. For example, such a portfolio for $\hat{\delta}^*$ in Panel B of Table 4 has a Fama-French alpha of 8.52% per year ($t = 3.99$), and a four-factor alpha of 5.24% ($t = 2.47$). For $\hat{\delta}^{**}$, the two alphas of the (5,5)-(1,1) portfolio in Panel B of Table 6 are 6.54% ($t = 3.72$) and 4.73% per year ($t = 2.48$). All of these alphas are higher than those obtained by one-way quintile sorts, which suggests that mutual fund investors would benefit from combining the information in our measures with the information in $\hat{\alpha}$.

To assess the benefits of our measures to investors, we must examine only portfolio strategies that are feasible. The holdings data used to compute $\hat{\delta}^*$ and $\hat{\delta}^{**}$ becomes publicly available with a lag, since mutual funds can report their holdings to the SEC with a lag of up to two months (e.g. Myers et al., 2001). To design realistic investment strategies, we assume that fund portfolios are formed with a one-quarter delay after the portfolio disclosure date. Specifically, we use returns through month t and holdings as of month t to predict fund returns in months $t+4$ through $t+6$ (as opposed to $t+1$ through $t+3$). The sample period is shifted by one quarter to July 1982 through December 2002. We conduct double sorts as in Tables 4 through 7, dropping the reverse sorts to save space, and report the results in Tables 8 and 9.

These results show that our measures, especially $\hat{\delta}^*$, help predict fund returns even after allowing for the reporting delays associated with fund holdings. Across all quintiles of $\hat{\alpha}$, sorting on $\hat{\delta}^*$ creates a positive 5-1 spread in benchmark-adjusted future fund returns. The average spread between the top and bottom $\hat{\delta}^*$ portfolios is between 2.6% and 4.1% per year, with t-statistics ranging from 1.90 to 2.67 across four model-lookback combinations. Combining $\hat{\delta}^*$ with $\hat{\alpha}$ helps further increase the spread – the alphas of the (5,5)-(1,1) portfolio range from 7.69% to 8.89% per year! The results for $\hat{\delta}^{**}$ are more affected by the reporting delays, as the average 5-1 portfolio alphas are marginally insignificant at just under one percent. These results indicate that mutual fund investors can benefit from the information in our measures, especially $\hat{\delta}^*$.

A natural question raised by these results is whether mutual fund investors are aware of the valuable information contained in our measures, and whether they respond to it by adjusting their fund allocations. It is well known that mutual fund flows respond strongly to $\hat{\alpha}$ (e.g. Sirri and Tufano, 1998). To see whether they also respond to our measures over and above their response to $\hat{\alpha}$, we conduct similar conditional sorts as before and report the results in Table 10. At each quarter-end, we sort funds into deciles by $\hat{\alpha}$ and then within each decile by $\hat{\delta}^*$ (Panel A) or $\hat{\delta}^{**}$ (Panel B). For each of the resulting one hundred portfolios, we report the average quarterly net fund inflows across the whole sample. These inflows are computed with a one-quarter delay, for reasons explained earlier. In the interest of space, only the results using a 12-month lookback period and the four-factor model are reported, but the results using the funds' full history and the Fama-French model are very similar.

Looking across the columns confirms that net fund flows increase reliably with $\hat{\alpha}$. The differences between the net inflows into the tenth and first porfolios range from 4.4% to 10.5% per quarter and are uniformly statistically significant. Looking across the rows, however, reveals no reliable relation between net flows and our performance measures. These results suggest that while investors pay close attention to $\hat{\alpha}$ in allocating their capital across funds, they pay little if any attention to $\hat{\delta}^*$ and $\hat{\delta}^{**}$.

This ignorance is costly, given the evidence in Tables 4 through 9. For example, suppose investor A sorts funds into quintiles by $\hat{\alpha}$ using a 12-month lookback period and a one-quarter delay, and invests in the 5-1 portfolio.[20] This investor would have earned a four-factor alpha of 3.4% per year ($t = 3.6$) over the whole sample period. Investor B performs a conditional sort into quintiles first by $\hat{\alpha}$ and then by $\hat{\delta}^*$, and invests in the (5,5)-(1,1) portfolio, with the same lookback and delay as investor A. Investor B would have earned a four-factor alpha of 8.1% per year ($t = 3.9$). The difference between the alphas of investors A and B, 4.7% per year ($t = 3.6$) over a 20.5-year period, is highly economically significant, and is likely to exceed any reasonable rebalancing costs. We conclude that mutual fund investors would benefit from using not only $\hat{\alpha}$ but also our measures to predict future fund performance.

IV Conclusion

This paper proposes new performance measures that exploit the information contained in the similarity of a manager's holdings (or changes in holdings) to those of managers who have performed well, and in their distinctiveness from those of managers who have performed poorly. These performance measures use historical returns and holdings of many funds to evaluate the performance of a single fund.[21] As a result, these measures are typically more precise than the traditional return-based measures. In simulations, our measures are found to be particularly well suited for empirical applications that

[20] For simplicity, we assume that it is possible to short funds, but our conclusions do not depend on the ability to short because much of the profits of our long-short strategies comes from the long position.

[21] Some information pooling across funds takes place also in the Bayesian frameworks developed by Jones and Shanken (2002) and Stambaugh (2002), in which a fund's alpha is related to the alphas of other funds through a link in the prior. The techniques as well as the objectives of these papers are quite different from ours. Jones and Shanken compare inferences about alphas with and without imposing their prior dependence, while Stambaugh focuses on inference about surviving funds. Neither study considers similarities in fund managers' decisions and their relation to performance evaluation.

involve ranking managers. This suitability is confirmed in our empirical analysis. Our measures are shown to contain a significant amount of information about future fund returns, information that is not contained in the standard alpha measure. Our evidence suggests that mutual fund investors could benefit significantly from investing in funds selected by combining the information contained in alpha and in our measures, at least before costs and fees. Despite this potential benefit, mutual fund flows do not appear to respond to our measures nearly as strongly as they respond to alpha.

The basic idea in the paper is that managers who make similar investment decisions have similar skill. The proposed skill measures are designed to capture this idea in a simple way, but future research can extend these measures in various dimensions. For example, the measure of stock quality can be modified to give more weight to the $\hat{\alpha}$'s with lower standard errors. As a broader example, useful information about similarities in investment decisions may also be contained in the funds' residual return correlations, if one is willing to assume that such similarities are stable over time. (No such assumption is needed when using holdings data because similarity in holdings can be evaluated across stocks at any given point in time.) Finally, our measures rely on recent holdings or changes in holdings. The earlier holdings are relevant only to the extent that they determine the manager's own track record, which is only one of many track records reflected in our measures. It seems reasonable to forecast the manager's future performance from his current rather than past decisions, but it might also be interesting to design performance measures that incorporate similarities among the managers' historical holdings. Future research can also pursue a variety of applications of our measures.

Table 1
Simulation Evidence on Various Skill Measures Relative to True Skill γ

Random samples of returns on N stocks and signals of M managers over T years are simulated as described in Section II. In each sample, we calculate six different performance measures for each manager. Three of these measures are population versions of the traditional measure α and of our levels and changes estimators, δ^* and δ^{**}. The remaining three measures are sample measures: $\hat{\alpha}$ denotes the average realized excess return on the manager's portfolio, $\hat{\delta}^*$ denotes our levels estimator, which exploits similarities in holdings across managers, and $\hat{\delta}^{**}$ denotes our changes estimator, which exploits similarities in changes in holdings across managers. All managers are ranked according to their performance estimated using the six measures, and rank correlations with the ranking based on true skill γ are reported. All numbers are averaged across 10,000 simulations.

Rank Correlations with True Skill (γ)

	N=30						N=100					
M	$\hat{\alpha}$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	α	δ^*	δ^{**}	$\hat{\alpha}$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	α	δ^*	δ^{**}
T=1												
30	0.26	0.34	0.35	0.80	0.80	0.82	0.46	0.64	0.65	0.92	0.92	0.93
100	0.27	0.40	0.42	0.81	0.82	0.85	0.47	0.76	0.77	0.93	0.94	0.94
300	0.27	0.44	0.45	0.82	0.83	0.85	0.47	0.80	0.81	0.93	0.94	0.95
T=5												
30	0.53	0.64	0.63	0.94	0.86	0.87	0.77	0.89	0.90	0.98	0.94	0.95
100	0.54	0.72	0.74	0.95	0.85	0.88	0.78	0.93	0.95	0.98	0.95	0.96
300	0.54	0.76	0.78	0.95	0.84	0.88	0.79	0.94	0.96	0.98	0.94	0.96
T=10												
30	0.66	0.75	0.75	0.96	0.86	0.88	0.86	0.93	0.93	0.99	0.95	0.96
100	0.68	0.81	0.83	0.97	0.85	0.89	0.88	0.94	0.96	0.99	0.95	0.96
300	0.68	0.82	0.86	0.97	0.85	0.88	0.88	0.94	0.96	0.99	0.95	0.96
T=20												
30	0.79	0.82	0.83	0.98	0.87	0.89	0.92	0.94	0.95	0.99	0.95	0.96
100	0.80	0.84	0.87	0.98	0.85	0.89	0.93	0.95	0.96	0.99	0.95	0.96
300	0.80	0.84	0.88	0.99	0.85	0.89	0.93	0.95	0.96	1.00	0.95	0.96
T=30												
30	0.84	0.84	0.85	0.98	0.87	0.89	0.94	0.94	0.95	0.99	0.95	0.96
100	0.86	0.84	0.88	0.99	0.85	0.89	0.95	0.95	0.96	1.00	0.95	0.96
300	0.86	0.84	0.88	0.99	0.85	0.89	0.96	0.95	0.96	1.00	0.95	0.96

Table 2
Simulation Evidence on Various Skill Measures Relative to Traditional Skill α

Random samples of returns on N stocks and signals of M managers over T years are simulated as described in Section II. In each sample, we calculate six different performance measures for each manager. Two of these measures are population versions of our levels and changes estimators, δ^* and δ^{**}. The remaining four measures are sample measures: $\hat{\alpha}$ denotes the average realized excess return on the manager's portfolio, $\hat{\alpha}^B$ denotes the Bayesian estimator that shrinks $\hat{\alpha}$ halfway toward the average of $\hat{\alpha}$ across managers, $\hat{\delta}^*$ denotes our levels estimator, which exploits similarities in holdings across managers, and $\hat{\delta}^{**}$ denotes our changes estimator, which exploits similarities in changes in holdings across managers. All managers are ranked according to their performance estimated using the six measures, and rank correlations with the ranking based on the traditional skill measure α are reported in Panel A. Panel B reports averages across managers of the squared differences between the estimator and the true α (times 100). All numbers are averaged across 10,000 simulations.

Panel A. Rank Correlations with Traditional Skill (α)

	N=30						N=100					
M	$\hat{\alpha}$	$\hat{\alpha}^B$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	δ^*	δ^{**}	$\hat{\alpha}$	$\hat{\alpha}^B$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	δ^*	δ^{**}
T=1												
30	0.32	0.32	0.41	0.40	0.96	0.95	0.49	0.49	0.67	0.68	0.98	0.98
100	0.33	0.33	0.48	0.47	0.98	0.97	0.50	0.50	0.80	0.80	0.99	0.98
300	0.33	0.33	0.52	0.51	0.99	0.97	0.50	0.50	0.84	0.84	0.99	0.99
T=5												
30	0.55	0.55	0.66	0.65	0.89	0.91	0.78	0.78	0.90	0.91	0.96	0.96
100	0.56	0.56	0.75	0.76	0.88	0.91	0.79	0.79	0.94	0.95	0.96	0.97
300	0.57	0.57	0.78	0.80	0.87	0.91	0.80	0.80	0.95	0.96	0.95	0.97
T=10												
30	0.68	0.68	0.77	0.77	0.88	0.90	0.87	0.87	0.93	0.94	0.95	0.96
100	0.69	0.69	0.82	0.84	0.87	0.90	0.88	0.88	0.95	0.96	0.95	0.96
300	0.70	0.70	0.83	0.87	0.86	0.90	0.89	0.89	0.95	0.96	0.95	0.96
T=30												
30	0.85	0.85	0.85	0.86	0.88	0.90	0.95	0.95	0.95	0.95	0.95	0.96
100	0.86	0.86	0.85	0.88	0.86	0.89	0.95	0.95	0.95	0.96	0.95	0.96
300	0.87	0.87	0.85	0.89	0.85	0.89	0.96	0.96	0.95	0.96	0.95	0.96

Panel B. Mean Squared Errors

	N=30						N=100					
M	$\hat{\alpha}$	$\hat{\alpha}^B$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	δ^*	δ^{**}	$\hat{\alpha}$	$\hat{\alpha}^B$	$\hat{\delta}^*$	$\hat{\delta}^{**}$	δ^*	δ^{**}
T=1												
30	2.65	1.71	1.48	0.75	0.11	0.11	0.78	0.53	0.49	0.23	0.09	0.12
100	2.61	1.65	1.40	0.52	0.12	0.10	0.78	0.52	0.49	0.19	0.10	0.13
300	2.62	1.65	1.40	0.47	0.12	0.10	0.78	0.52	0.49	0.19	0.10	0.13
T=5												
30	0.53	0.36	0.36	0.23	0.09	0.09	0.16	0.13	0.17	0.11	0.08	0.08
100	0.53	0.36	0.35	0.17	0.09	0.09	0.16	0.13	0.17	0.10	0.09	0.08
300	0.52	0.35	0.35	0.15	0.10	0.09	0.16	0.13	0.17	0.10	0.10	0.08
T=10												
30	0.26	0.20	0.22	0.16	0.08	0.10	0.08	0.08	0.12	0.10	0.08	0.08
100	0.26	0.19	0.22	0.13	0.09	0.09	0.08	0.08	0.13	0.09	0.09	0.08
300	0.26	0.19	0.22	0.12	0.09	0.09	0.08	0.08	0.13	0.09	0.09	0.08
T=30												
30	0.09	0.09	0.13	0.12	0.08	0.10	0.03	0.05	0.10	0.09	0.08	0.08
100	0.09	0.08	0.13	0.10	0.09	0.09	0.03	0.05	0.10	0.09	0.09	0.09
300	0.09	0.09	0.13	0.10	0.09	0.09	0.03	0.05	0.11	0.09	0.09	0.09

Table 3
Return Predictability for Funds Sorted by Various Measures of Past Performance

At the end of each quarter, funds are sorted into decile portfolios by various measures of past performance: $\hat{\alpha}$, the OLS estimate of the fund's alpha, $\hat{\delta}^*$, our levels estimator, and $\hat{\delta}^{**}$, our changes estimator. The returns on the decile portfolios are calculated over the three months after portfolio formation, equal-weighting the funds within each decile. The three-month return series are linked across quarters to form a monthly series of returns on each decile portfolio. All performance measures are calculated using the data over the past 12 months (Panel A) and all available past data (Panel B). The table reports the OLS estimates of the deciles' full-period alphas (in percent per year), as well as their t-statistics (in parentheses). These alphas are defined in the same way as the reference measures $\hat{\alpha}$. The Fama-French alpha with respect to the market, size, and value benchmarks, following Fama and French (1993), and the four-factor alpha with respect to the Fama-French and momentum benchmarks, following Carhart (1997). The sample period is April 1982 through September 2002.

	Decile										
	1	2	3	4	5	6	7	8	9	10	10-1
Panel A. Sorting funds by past 12 months of performance											
Fama-French Alpha as reference measure											
$\hat{\alpha}$	-1.62	-0.39	0.00	0.15	0.43	0.75	0.94	1.19	1.62	3.57	5.19
	(-1.62)	(-0.57)	(0.00)	(0.30)	(0.87)	(1.44)	(1.84)	(2.13)	(2.31)	(3.57)	(3.67)
$\hat{\delta}^*$	-1.87	-0.91	-0.75	-0.24	-0.01	-0.01	0.18	2.00	2.72	5.48	7.36
	(-1.30)	(-0.87)	(-1.03)	(-0.42)	(-0.02)	(-0.01)	(0.33)	(2.81)	(2.86)	(4.11)	(3.23)
$\hat{\delta}^{**}$	-1.13	-0.27	-0.12	0.37	0.53	0.07	0.97	0.75	1.51	3.32	4.45
	(-1.23)	(-0.45)	(-0.21)	(0.67)	(1.08)	(0.17)	(1.77)	(1.34)	(2.23)	(3.63)	(4.53)
Four-Factor Alpha as reference measure											
$\hat{\alpha}$	-1.21	-0.63	0.19	1.13	0.89	0.29	0.65	1.05	1.81	2.48	3.69
	(-1.20)	(-0.80)	(0.31)	(2.13)	(1.81)	(0.54)	(1.29)	(1.68)	(2.63)	(2.60)	(2.64)
$\hat{\delta}^*$	-1.58	-0.89	-0.29	-0.11	0.51	0.72	0.67	1.97	1.33	4.30	5.88
	(-1.14)	(-0.81)	(-0.38)	(-0.17)	(0.91)	(1.32)	(1.25)	(2.56)	(1.37)	(3.46)	(2.73)
$\hat{\delta}^{**}$	-0.60	-0.20	0.30	0.38	0.54	0.76	0.18	0.86	1.15	2.92	3.52
	(-0.62)	(-0.31)	(0.47)	(0.81)	(1.10)	(1.56)	(0.32)	(1.55)	(1.66)	(3.11)	(3.25)
Panel B. Sorting funds by entire past record											
Fama-French Alpha as reference measure											
$\hat{\alpha}$	-1.26	-0.05	-0.46	-0.18	0.17	0.24	0.87	1.11	1.77	4.38	5.64
	(-1.42)	(-0.08)	(-0.68)	(-0.35)	(0.39)	(0.48)	(1.82)	(2.04)	(2.79)	(4.41)	(4.49)
$\hat{\delta}^*$	-1.82	-1.36	-0.54	-0.55	0.27	0.21	1.08	1.51	2.47	5.32	7.15
	(-1.67)	(-1.70)	(-0.90)	(-1.12)	(0.52)	(0.40)	(1.95)	(2.34)	(2.66)	(4.12)	(3.84)
$\hat{\delta}^{**}$	-1.08	-0.09	0.31	-0.27	0.53	0.52	1.12	0.26	1.34	3.31	4.39
	(-1.35)	(-0.14)	(0.59)	(-0.49)	(1.05)	(1.05)	(2.05)	(0.46)	(2.05)	(3.61)	(4.67)
Four-Factor Alpha as reference measure											
$\hat{\alpha}$	-0.74	-0.14	0.36	0.24	0.88	0.66	0.36	1.02	1.47	2.51	3.25
	(-1.03)	(-0.24)	(0.69)	(0.43)	(1.79)	(1.22)	(0.74)	(1.62)	(2.37)	(2.95)	(3.89)
$\hat{\delta}^*$	-1.62	-0.69	-0.27	-0.26	0.40	1.32	1.01	1.54	2.11	3.08	4.70
	(-1.51)	(-0.88)	(-0.43)	(-0.48)	(0.77)	(2.49)	(1.83)	(2.31)	(2.27)	(2.68)	(2.78)
$\hat{\delta}^{**}$	-0.93	0.32	0.15	0.61	0.66	0.47	0.52	0.89	0.39	3.14	4.06
	(-1.14)	(0.52)	(0.28)	(1.07)	(1.23)	(0.98)	(0.95)	(1.40)	(0.56)	(3.43)	(4.90)

Table 4
Double Sorts Comparing δ^* to Alpha Using Past 12 Months of Performance

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to $\hat{\delta}^*$, our levels estimator. These estimators are constructed using the past 12 months performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 1, 2, and 3 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. Panel B reverses the order, sorting first by $\hat{\delta}^*$ and then, within each quintile, by $\hat{\alpha}$. The sample period is April 1982 through September 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^*$

Quintile	Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$					
of $\hat{\delta}^*$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-2.89	-0.58	-0.12	0.12	-0.18	-0.73	-1.55	-0.95	0.05	-0.76	0.40	-0.56
2	-1.61	-1.79	-0.53	0.46	0.54	-0.59	-1.28	-0.23	-1.08	-0.01	1.18	-0.28
3	-1.73	0.21	0.14	-0.55	2.18	0.05	-2.05	1.21	0.58	0.81	1.28	0.37
4	-1.05	0.22	0.61	0.77	3.77	0.86	-0.85	1.20	0.95	1.81	2.29	1.08
5	2.34	2.40	2.60	4.65	6.58	3.71	1.22	2.27	1.91	2.78	5.41	2.72
5-1	5.22	2.98	2.72	4.53	6.76	4.44	2.77	3.22	1.86	3.54	5.01	3.28
t-stat	(2.68)	(1.66)	(1.74)	(2.58)	(3.38)	(2.77)	(1.57)	(1.66)	(1.09)	(2.00)	(2.66)	(2.06)

Panel B. Sorting funds by $\hat{\delta}^*$ and then $\hat{\alpha}$

Quintile	Quintile of $\hat{\delta}^*$						Quintile of $\hat{\delta}^*$					
of $\hat{\alpha}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-2.24	-0.05	0.24	2.00	3.83	0.76	-1.37	-0.90	1.42	1.01	2.36	0.51
2	-2.51	-0.21	0.21	1.52	3.64	0.53	-1.94	0.08	1.05	1.47	2.16	0.56
3	-0.87	-1.26	0.44	0.80	3.66	0.56	-1.19	0.67	0.70	1.63	2.96	0.95
4	-0.94	-0.82	-0.35	0.08	3.22	0.24	-1.12	-0.01	-0.08	1.37	2.83	0.60
5	-0.41	-0.08	-0.36	0.95	6.28	1.28	-0.54	-0.93	0.06	1.09	3.87	0.71
5-1	1.84	-0.04	-0.60	-1.04	2.45	0.52	0.83	-0.04	-1.36	0.08	1.51	0.20
t-stat	(1.49)	(-0.04)	(-0.67)	(-1.07)	(2.61)	(0.82)	(0.68)	(-0.05)	(-1.53)	(0.08)	(1.51)	(0.33)

Table 5
Double Sorts Comparing δ^* to Alpha Using Entire Fund Performance Record

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to $\hat{\delta}^*$, our levels estimator. These estimators are constructed using the entire performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 1, 2, and 3 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. Panel B reverses the order, sorting first by $\hat{\delta}^*$ and then, within each quintile, by $\hat{\alpha}$. The sample period is April 1982 through September 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^*$

Quintile			Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$			
of $\hat{\delta}^*$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-1.66	-1.96	-0.84	0.19	0.42	-0.77	-1.92	-0.65	0.24	0.01	0.51	-0.36
2	-1.41	-0.84	0.06	0.33	2.29	0.09	-1.49	-0.78	-0.65	-0.34	1.46	-0.36
3	-1.46	-0.78	-0.17	0.65	2.41	0.13	-0.58	0.62	0.69	0.73	2.35	0.76
4	-0.14	0.31	-0.07	0.84	3.18	0.83	0.12	1.27	1.79	1.39	1.50	1.21
5	1.56	1.51	2.18	2.76	7.12	3.03	1.74	1.13	1.73	1.52	4.15	2.05
5-1	3.22	3.47	3.02	2.57	6.71	3.80	3.65	1.78	1.49	1.50	3.63	2.41
t-stat	(2.09)	(2.52)	(2.49)	(2.02)	(4.02)	(3.21)	(2.16)	(1.23)	(1.10)	(1.11)	(2.36)	(1.94)

Panel B. Sorting funds by $\hat{\delta}^*$ and then $\hat{\alpha}$

Quintile			Quintile of $\hat{\delta}^*$						Quintile of $\hat{\delta}^*$			
of $\hat{\alpha}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-2.18	-1.30	0.31	1.55	2.82	0.24	-1.41	-1.06	0.69	1.09	2.07	0.28
2	-1.89	-0.02	-0.07	1.42	2.76	0.44	-2.19	-0.62	1.12	1.00	1.25	0.11
3	-1.39	-0.36	0.27	0.79	3.62	0.59	-0.81	0.97	1.13	1.37	2.95	1.12
4	-0.94	-0.19	0.48	1.48	4.85	1.14	-0.17	-0.29	-0.09	1.64	2.24	0.66
5	-1.62	-0.78	0.32	1.35	5.40	0.93	-1.04	-0.42	1.41	1.55	4.23	1.15
5-1	0.56	0.51	0.01	-0.20	2.58	0.69	0.37	0.65	0.72	0.46	2.16	0.87
t-stat	(0.62)	(0.90)	(0.01)	(-0.31)	(3.18)	(1.61)	(0.42)	(1.02)	(0.98)	(0.53)	(2.69)	(1.79)

Table 6
Double Sorts Comparing δ^{**} to Alpha Using Past 12 Months of Performance

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to $\hat{\delta}^{**}$, our changes estimator. These estimators are constructed using the past 12 months performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 1, 2, and 3 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. Panel B reverses the order, sorting first by $\hat{\delta}^{**}$ and then, within each quintile, by $\hat{\alpha}$. The sample period is April 1982 through September 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^{**}$

Quintile	Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$					
of $\hat{\delta}^{**}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-2.67	0.08	0.73	0.75	1.56	0.09	-1.04	-0.23	0.44	1.18	1.57	0.38
2	-1.73	-0.58	0.72	0.28	2.03	0.14	-1.64	0.73	1.12	1.18	0.35	0.35
3	-1.29	0.54	-0.93	0.56	2.29	0.23	-0.80	0.53	-0.35	0.06	1.78	0.24
4	-0.62	0.60	1.41	0.84	3.33	1.11	-0.53	0.67	-0.12	1.06	2.30	0.68
5	0.74	0.65	1.00	0.98	3.98	1.47	-0.20	1.35	1.38	0.90	4.26	1.54
5-1	3.41	0.57	0.27	0.24	2.42	1.38	0.84	1.58	0.94	-0.28	2.69	1.15
t-stat	(2.97)	(0.72)	(0.35)	(0.25)	(2.15)	(2.73)	(0.76)	(1.84)	(1.09)	(-0.29)	(2.40)	(2.36)

Panel B. Sorting funds by $\hat{\delta}^{**}$ and then $\hat{\alpha}$

Quintile	Quintile of $\hat{\delta}^{**}$						Quintile of $\hat{\delta}^{**}$					
of $\hat{\alpha}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-1.90	0.53	0.43	0.97	1.44	0.29	-0.78	-0.56	0.46	2.02	0.80	0.39
2	-1.47	0.32	0.57	0.34	1.46	0.25	-1.20	-0.09	0.15	-0.48	1.90	0.06
3	-1.91	-0.54	-0.39	1.14	2.61	0.18	-1.30	1.01	0.61	0.76	2.04	0.62
4	-0.87	0.06	0.75	0.50	3.14	0.72	-0.79	0.29	1.18	-0.15	1.69	0.44
5	0.88	0.50	0.57	1.52	4.64	1.62	1.66	0.52	1.66	0.75	3.96	1.71
5-1	2.78	-0.03	0.14	0.55	3.21	1.33	2.43	1.08	1.20	-1.27	3.16	1.32
t-stat	(2.13)	(-0.03)	(0.15)	(0.55)	(2.32)	(1.70)	(1.88)	(1.14)	(1.35)	(-1.29)	(2.44)	(1.78)

Table 7
Double Sorts Comparing δ^{**} to Alpha Using Entire Fund Performance Record

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to $\hat{\delta}^{**}$, our changes estimator. These estimators are constructed using the entire performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 1, 2, and 3 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. Panel B reverses the order, sorting first by $\hat{\delta}^{**}$ and then, within each quintile, by $\hat{\alpha}$. The sample period is April 1982 through September 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^{**}$

Quintile	Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$					
of $\hat{\delta}^{**}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-1.54	0.23	-0.55	1.12	2.50	0.35	-2.04	0.10	1.34	0.69	1.85	0.39
2	-1.26	-0.64	0.26	0.52	1.29	0.04	-0.10	-0.38	0.79	0.92	0.49	0.34
3	-0.07	-1.19	-0.99	2.08	2.89	0.54	-0.66	1.38	-0.86	0.21	1.11	0.24
4	-0.81	0.44	0.98	-0.39	2.84	0.61	-0.05	0.27	0.16	0.08	2.03	0.50
5	0.72	0.62	0.32	0.57	5.29	1.50	0.77	1.16	1.59	0.30	4.53	1.67
5-1	2.26	0.40	0.87	-0.55	2.79	1.15	2.82	1.06	0.25	-0.39	2.68	1.28
t-stat	(2.12)	(0.51)	(1.11)	(-0.67)	(2.49)	(2.31)	(2.47)	(1.19)	(0.30)	(-0.41)	(2.34)	(2.34)

Panel B. Sorting funds by $\hat{\delta}^{**}$ and then $\hat{\alpha}$

Quintile	Quintile of $\hat{\delta}^{**}$						Quintile of $\hat{\delta}^{**}$					
of $\hat{\alpha}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-1.17	-0.88	0.09	-0.48	1.98	-0.09	-0.86	0.13	0.51	0.71	1.46	0.39
2	-1.91	0.46	0.76	0.45	1.32	0.22	-1.05	0.26	0.46	1.02	0.24	0.18
3	-1.03	-0.49	-0.21	-0.17	1.09	-0.17	-0.49	1.11	-0.12	-0.33	1.40	0.31
4	0.26	-0.12	0.25	0.93	4.06	1.08	-0.48	0.65	-0.35	0.54	2.01	0.47
5	-0.32	1.46	1.44	2.25	5.02	1.97	0.78	0.80	1.40	1.74	4.22	1.79
5-1	0.85	2.34	1.35	2.72	3.04	2.06	1.65	0.67	0.89	1.03	2.76	1.40
t-stat	(0.70)	(2.61)	(1.53)	(2.64)	(2.46)	(2.73)	(1.69)	(0.70)	(1.23)	(1.19)	(2.91)	(2.81)

Table 8
Double Sorts Using Past 12 Months of Performance with One-Quarter Delay

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to either $\hat{\delta}^*$ (Panel A) or $\hat{\delta}^{**}$ (Panel B). These estimators are constructed using the past 12 months performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 4, 5, and 6 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. The sample period is July 1982 through December 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^*$

Quintile			Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$			
of $\hat{\delta}^*$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-3.29	-1.51	-0.79	-1.05	0.10	-1.31	-3.87	-1.36	-0.70	-0.33	1.14	-1.02
2	-2.33	-0.47	0.30	-0.21	1.74	-0.19	-1.67	-0.32	1.28	0.04	2.03	0.27
3	-2.24	-0.35	1.10	0.26	2.40	0.24	-0.78	-0.69	0.40	0.81	2.18	0.38
4	-0.40	0.05	0.72	0.43	3.27	0.81	0.14	-0.54	0.75	1.40	2.27	0.81
5	1.78	2.41	3.19	2.31	4.44	2.83	0.77	1.59	0.45	2.26	4.26	1.87
5-1	5.06	3.92	3.98	3.36	4.35	4.13	4.64	2.95	1.15	2.59	3.12	2.89
t-stat	(2.64)	(2.35)	(2.45)	(1.93)	(2.46)	(2.66)	(2.65)	(1.61)	(0.65)	(1.52)	(1.88)	(1.90)

Panel B. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^{**}$

Quintile			Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$			
of $\hat{\delta}^{**}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-2.34	0.23	0.26	-0.25	2.84	0.15	-2.84	-0.32	0.33	0.71	2.36	0.05
2	-1.26	-1.02	1.58	0.56	2.23	0.42	-0.79	-0.79	0.53	1.59	1.76	0.46
3	-1.28	0.54	0.28	1.43	0.61	0.31	-0.48	-0.14	0.38	0.62	0.74	0.22
4	-1.01	1.16	1.10	-0.65	2.02	0.52	-0.22	0.27	1.14	0.98	1.87	0.81
5	0.60	0.31	0.80	0.76	3.07	1.11	0.11	0.23	0.15	0.31	3.82	0.92
5-1	2.95	0.08	0.54	1.01	0.23	0.96	2.95	0.55	-0.18	-0.41	1.46	0.88
t-stat	(2.42)	(0.10)	(0.67)	(1.18)	(0.22)	(1.97)	(2.40)	(0.62)	(-0.22)	(-0.42)	(1.30)	(1.68)

Table 9
Double Sorts Using Entire Fund Performance Record with One-Quarter Delay

At the end of each quarter, funds are sorted into quintile portfolios according to $\hat{\alpha}$, the OLS estimate of the fund's alpha and then sorted within the quintile according to either $\hat{\delta}^*$ (Panel A) or $\hat{\delta}^{**}$ (Panel B). These estimators are constructed using the entire performance record of each fund. Fund returns are then averaged within each of the 25 portfolios over months 4, 5, and 6 following portfolio formation. The three-month return series are linked across quarters to form a monthly series of returns on each portfolio. The average of this return series is reported for each of the 25 portfolios in Panel A. The final two rows report the difference between each of the 5th and 1st portfolios and the associated t-statistic. The sample period is July 1982 through December 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^*$

Quintile	Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$					
of $\hat{\delta}^*$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-3.62	-1.22	-0.88	0.16	0.72	-0.97	-2.90	-1.01	-0.08	-0.47	-0.08	-0.91
2	-1.02	-0.54	0.27	0.63	0.94	0.06	-1.55	0.08	-0.10	0.62	1.35	0.08
3	-0.45	0.53	0.21	0.01	2.12	0.48	0.33	0.53	0.27	-0.01	1.72	0.57
4	-0.32	-0.21	-0.55	0.42	3.31	0.53	0.01	0.98	0.31	1.34	1.52	0.83
5	1.52	1.11	1.77	1.90	5.27	2.31	0.46	0.30	1.63	1.36	4.79	1.71
5-1	5.14	2.32	2.64	1.74	4.54	3.28	3.36	1.31	1.71	1.83	4.87	2.62
t-stat	(3.29)	(1.75)	(2.06)	(1.31)	(2.55)	(2.67)	(2.02)	(0.90)	(1.35)	(1.45)	(3.06)	(2.13)

Panel B. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^{**}$

Quintile	Quintile of $\hat{\alpha}$						Quintile of $\hat{\alpha}$					
of $\hat{\delta}^{**}$	1	2	3	4	5	Avg.	1	2	3	4	5	Avg.
	Fama-French Alpha as reference measure						Four-Factor Alpha as reference measure					
1	-1.79	-0.67	-0.40	1.30	1.68	0.03	-1.62	0.75	-0.10	0.21	1.59	0.16
2	-0.62	0.85	-0.67	0.92	0.63	0.22	-0.80	0.59	0.76	-0.50	1.40	0.29
3	-0.37	0.27	0.98	1.03	3.12	1.01	-0.18	0.31	0.84	0.25	1.50	0.55
4	-0.01	-0.36	0.51	0.51	2.58	0.65	0.35	0.36	0.26	1.27	0.76	0.60
5	0.57	0.41	-0.74	-0.82	3.50	0.58	-0.03	0.53	0.04	0.15	3.55	0.85
5-1	2.36	1.08	-0.34	-2.12	1.81	0.56	1.59	-0.22	0.14	-0.06	1.96	0.68
t-stat	(2.13)	(1.31)	(-0.44)	(-2.80)	(1.55)	(1.12)	(1.38)	(-0.25)	(0.17)	(-0.08)	(1.82)	(1.37)

Table 10
Average Net Fund Flows into Funds Sorted by Past Performance

At the end of each quarter, funds are sorted into decile portfolios by $\hat{\alpha}$, the OLS estimate of the fund's alpha, and then within deciles by either $\hat{\delta}^*$ (Panel A) or $\hat{\delta}^{**}$ (Panel B). These estimators are constructed using the past 12 months performance record of each fund. Average percentage quarterly net fund inflows are calculated for each portfolio for the quarter corresponding to months 4 through 6 following portfolio formation. The final two rows and columns report the differences between the 10th and 1st portfolios and their associated t-statistics (in parenthesis). The sample period is July 1982 through December 2002.

Panel A. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^*$ over past 12 months of performance

Decile	Decile of $\hat{\alpha}$											
of $\hat{\delta}^*$	1	2	3	4	5	6	7	8	9	10	10-1	t-stat
					Four Factor Alpha as reference measure							
1	-2.48	0.19	0.58	-1.09	0.14	1.18	1.74	2.48	3.29	5.62	8.10	(5.12)
2	-0.03	0.79	0.91	-0.94	1.02	1.05	1.97	1.50	4.28	5.10	5.13	(3.18)
3	-0.90	-0.35	-0.41	0.90	0.75	0.65	0.57	2.16	3.76	5.13	6.02	(4.36)
4	-0.06	0.83	-0.57	0.60	0.19	0.10	2.26	3.23	4.56	4.74	4.81	(3.14)
5	-0.32	-0.30	-0.15	0.39	1.61	1.44	1.10	1.53	2.65	4.11	4.43	(2.69)
6	1.07	-0.35	0.43	1.18	1.41	0.88	2.65	1.55	3.27	7.27	6.20	(3.68)
7	-0.57	-0.10	0.70	-0.56	2.02	2.02	1.61	2.59	3.47	6.80	7.37	(4.58)
8	-0.15	0.72	-1.45	-0.35	0.08	1.53	1.80	2.79	3.98	7.30	7.45	(4.78)
9	2.87	0.93	0.66	1.34	1.60	2.57	3.31	3.85	3.33	7.23	4.36	(2.36)
10	0.28	1.02	1.52	3.00	0.87	0.34	5.26	3.77	5.65	8.52	8.24	(4.13)
10-1	2.77	0.83	0.94	4.09	0.72	-0.83	3.52	1.29	2.36	2.90		
t-stat	(1.90)	(0.54)	(0.60)	(2.60)	(0.50)	(-0.58)	(2.13)	(0.79)	(1.44)	(1.39)		

Panel B. Sorting funds by $\hat{\alpha}$ and then $\hat{\delta}^{**}$ over past 12 months of performance

Decile	Decile of $\hat{\alpha}$											
of $\hat{\delta}^{**}$	1	2	3	4	5	6	7	8	9	10	10-1	t-stat
					Four Factor Alpha as reference measure							
1	-1.31	0.07	-0.56	-0.43	-0.10	0.72	0.67	1.72	3.70	5.01	6.32	(3.85)
2	0.83	0.07	0.34	1.62	1.05	-0.33	3.33	2.49	4.42	5.65	4.81	(2.84)
3	-0.22	-0.12	-0.03	-0.23	0.80	0.75	2.01	2.17	2.97	4.64	4.87	(3.31)
4	0.77	-0.62	-0.33	-0.27	0.88	2.10	2.49	2.93	5.19	5.02	4.25	(2.44)
5	-0.89	0.99	0.10	2.10	1.99	1.50	1.61	2.67	1.43	6.72	7.62	(4.75)
6	-0.01	0.92	0.96	1.37	0.56	1.15	2.83	1.73	4.08	4.56	4.57	(3.12)
7	0.69	0.13	1.59	-1.17	1.90	1.75	2.37	2.99	4.56	6.43	5.74	(3.68)
8	-0.07	1.29	0.10	0.72	1.26	0.92	2.39	1.73	4.97	6.60	6.67	(4.03)
9	1.45	0.06	-0.59	0.43	0.68	1.14	2.19	2.82	3.28	8.40	6.95	(3.84)
10	-1.49	0.53	0.21	0.24	0.62	2.32	1.95	4.21	3.59	8.96	10.45	(5.68)
10-1	-0.18	0.46	0.77	0.67	0.72	1.60	1.28	2.49	-0.10	3.95		
t-stat	(-0.14)	(0.29)	(0.57)	(0.47)	(0.50)	(1.07)	(1.02)	(1.54)	(-0.07)	(1.91)		

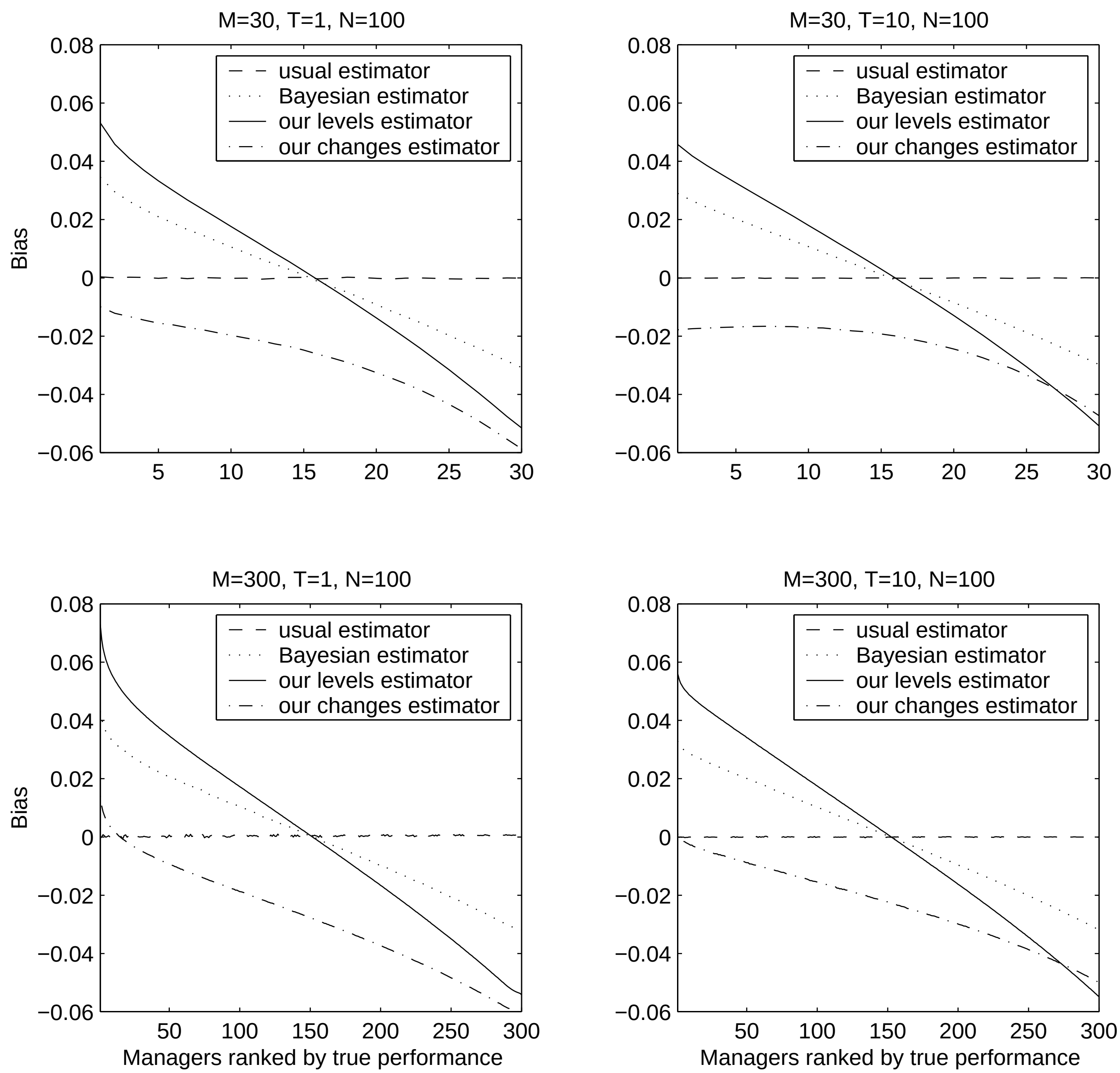

Figure 1. Bias of alternative performance estimators with respect to the traditional performance measure.

The figure plots the biases of four alternative performance estimators with respect to the traditional performance measure α, when managers are ranked by the true value of α. In each simulated sample with M (30 or 300) managers, $N = 100$ stocks, and T (1 or 10) years of simulated data, managers are ranked in ascending order by their α. Denote the true performance of the m-th ranked manager by α_m. For each estimator $\tilde{\delta}$ and each rank m ($m = 1, \ldots, M$), the bias is computed by averaging $\tilde{\delta}_m - \alpha_m$ across 10,000 simulated samples. Four estimators are considered: the usual estimator $\hat{\alpha}$, the Bayesian estimator $\hat{\alpha}^B$, our levels estimator $\hat{\delta}^*$, and our changes estimator $\hat{\delta}^{**}$.

Appendix

Computing $\Omega = \text{Cov}(\hat{\alpha}, \hat{\alpha}')$ for funds with misaligned return histories.

Let $\mathcal{S} = \{1, \ldots, T\}$ denote the set of dates in the whole sample period in which fund returns may be available. Suppose that returns on fund 1 are available in the subset $\mathcal{S}_1$ of the whole sample period, $\mathcal{S}_1 \subset \mathcal{S}$. These returns are stacked in the vector R_1, whose dimension is $N_1 \times 1$, where N_1 is the number of observations for fund 1. Analogously, suppose that returns on fund 2 are available in $\mathcal{S}_2 \subset \mathcal{S}$, and they are stacked in the $N_2 \times 1$ vector R_2. Also let R_B denote the vector of returns on K benchmark portfolios, available for the whole period $\mathcal{S}$. All returns are in excess of the risk-free rate. Define

$$R_{1,t} = \alpha_1 + R_{B,t}\beta_1 + \epsilon_{1,t}, \qquad t \in \mathcal{S}_1 \tag{A1}$$

$$R_{2,t} = \alpha_2 + R_{B,t}\beta_2 + \epsilon_{2,t}, \qquad t \in \mathcal{S}_2 \tag{A2}$$

The estimated intercepts $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are obtained by running separate OLS regressions, one using the data from $\mathcal{S}_1$ and the other using the data from $\mathcal{S}_2$.[22] The standard errors of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are computed accordingly. To calculate Ω, we also need $\text{Cov}(\hat{\alpha}_1, \hat{\alpha}_2)$ for each pair of funds. Rewrite equations (A1) and (A2) as

$$R_1 = X_1\theta_1 + \epsilon_1 \tag{A3}$$

$$R_2 = X_2\theta_2 + \epsilon_2, \tag{A4}$$

where $\theta_j = (\alpha_j \ \beta_j')'$, X_j is the subset of $X = [\iota_T \ R_B]$ corresponding to $\mathcal{S}_j$ (i.e. we take only the rows that correspond to the dates in $\mathcal{S}_j$), $j = 1, 2$, and ι_T is a T-vector of ones. Then

$$\hat{\theta}_1 = (X_1'X_1)^{-1}X_1'R_1 = (X_1'X_1)^{-1}X_1'(X_1\theta_1 + \epsilon_1) = \theta_1 + (X_1'X_1)^{-1}X_1'\epsilon_1$$

$$\hat{\theta}_2 = (X_2'X_2)^{-1}X_2'R_2 = (X_2'X_2)^{-1}X_2'(X_2\theta_2 + \epsilon_2) = \theta_2 + (X_2'X_2)^{-1}X_2'\epsilon_2$$

Hence,

$$\text{Cov}(\hat{\theta}_1, \hat{\theta}_2) = \text{E}[(\hat{\theta}_1 - \theta_1)(\hat{\theta}_2 - \theta_2)'] = (X_1'X_1)^{-1}X_1'\text{E}(\epsilon_1\epsilon_2')X_2(X_2'X_2)^{-1} \tag{A5}$$

[22]Using the techniques of Pástor and Stambaugh (2002), the intercepts of short-history funds can be estimated more precisely by incorporating the returns on longer-history funds. The most efficient use of all information involves regressing fund returns on the returns of all longer-history funds and the benchmarks. To ensure feasibility, we would have to choose a subset of longer-history funds, and the estimates would depend on that subset, which seems undesirable. (Pástor and Stambaugh use a small set of carefully-selected longer-history passive assets instead of longer-history funds.) We opt for the simplicity of estimating the regression interecepts fund by fund.

Note that ϵ_1 is $N_1 \times 1$ and ϵ_2 is $N_2 \times 1$, so that $\mathrm{E}(\epsilon_1 \epsilon_2')$ is $N_1 \times N_2$. Let σ_{12} denote the contemporaneous covariance between ϵ_1 and ϵ_2. Then $\mathrm{E}(\epsilon_1 \epsilon_2')$ is a matrix whose (i, j) element is σ_{12} if $\mathcal{S}_1(i) = \mathcal{S}_2(j)$ and zero otherwise, since the epsilons are assumed to be uncorrelated over time. Let $\mathcal{O}$ denote the overlap of the funds' sample periods, $\mathcal{O} = \mathcal{S}_1 \cap \mathcal{S}_2$, let X_O denote the row subset of X corresponding to $\mathcal{O}$, let N_O denote the number of elements in $\mathcal{O}$, and let I_{N_O} denote the identity matrix of dimension N_O. Equation (A5) can be rewritten as

$$\begin{aligned} \mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2) &= (X_1'X_1)^{-1} X_O' (\sigma_{12} I_{N_O}) X_O (X_2'X_2)^{-1} && \text{(A6)} \\ &= \sigma_{12} (X_1'X_1)^{-1} (X_O'X_O)(X_2'X_2)^{-1}. && \text{(A7)} \end{aligned}$$

Our estimate of $\mathrm{Cov}(\hat{\alpha}_1, \hat{\alpha}_2)$ is given by the $(1, 1)$ element of (A7). (As an example, if the history of fund 2 is subsumed by the history of fund 1, $\mathcal{S}_2 \subset \mathcal{S}_1$, so that $\mathcal{O} = \mathcal{S}_2$, then $\mathrm{Cov}(\hat{\theta}_1, \hat{\theta}_2) = \sigma_{12}(X_1'X_1)^{-1}$.) To estimate σ_{12}, we run the regressions (A1) and (A2) on the overlapping data $\mathcal{O}$ and take the sample covariance of the resulting residuals.[23]

Proof of the statement immediately following equation (26).

$$\begin{aligned} \sum_{j=1}^{M} c_{m,j} &= \sum_{j=1}^{M} \sum_{n=1}^{N} \Big[x^+_{m,n} y^+_{j,n} 1_{\{n \in \mathcal{N}^+_m\}} 1_{\{j \in \mathcal{M}^+_n\}} - x^+_{m,n} y^-_{j,n} 1_{\{n \in \mathcal{N}^+_m\}} 1_{\{j \in \mathcal{M}^-_n\}} \cdots \\ &\quad - x^-_{m,n} y^+_{j,n} 1_{\{n \in \mathcal{N}^-_m\}} 1_{\{j \in \mathcal{M}^+_n\}} + x^-_{m,n} y^-_{j,n} 1_{\{n \in \mathcal{N}^-_m\}} 1_{\{j \in \mathcal{M}^-_n\}} \Big] \\ &= \sum_{n=1}^{N} x^+_{m,n} 1_{\{n \in \mathcal{N}^+_m\}} \sum_{j=1}^{M} y^+_{j,n} 1_{\{j \in \mathcal{M}^+_n\}} - \sum_{n=1}^{N} x^+_{m,n} 1_{\{n \in \mathcal{N}^+_m\}} \sum_{j=1}^{M} y^-_{j,n} 1_{\{j \in \mathcal{M}^-_n\}} \cdots \\ &\quad - \sum_{n=1}^{N} x^-_{m,n} 1_{\{n \in \mathcal{N}^-_m\}} \sum_{j=1}^{M} y^+_{j,n} 1_{\{j \in \mathcal{M}^+_n\}} + \sum_{n=1}^{N} x^-_{m,n} 1_{\{n \in \mathcal{N}^-_m\}} \sum_{j=1}^{M} y^-_{j,n} 1_{\{j \in \mathcal{M}^-_n\}} \\ &= \sum_{n=1}^{N} x^+_{m,n} 1_{\{n \in \mathcal{N}^+_m\}} - \sum_{n=1}^{N} x^+_{m,n} 1_{\{n \in \mathcal{N}^+_m\}} - \sum_{n=1}^{N} x^-_{m,n} 1_{\{n \in \mathcal{N}^-_m\}} + \sum_{n=1}^{N} x^-_{m,n} 1_{\{n \in \mathcal{N}^-_m\}} \\ &= 0. && \text{(A8)} \end{aligned}$$

[23] There is a more sophisticated approach to estimating σ_{12} when $\mathcal{S}_2 \subset \mathcal{S}_1$. The results in Stambaugh (1997) can be used to estimate σ_{12} using also the data in $\mathcal{S}_1$ that is not in $\mathcal{S}_2$. However, since $\mathcal{S}_2 \subset \mathcal{S}_1$ is unlikely to happen for all pairs of funds and since we want the same procedure for all pairs, we simply take the estimate of σ_{12} from the overlapping data.

Inference and investment of our simulated investors.

Here we calculate the expected return and the covariance matrix of returns as perceived by the simulated mean-variance investors considered in Section II. The subscripts m and t, which denote investor and time, are dropped throughout for convenience.

First note that with a diffuse prior on μ_n, the Bayes rule implies that

$$\mu_n | s_n, \gamma = \begin{cases} s_n & \text{with probability } \gamma \\ u_n \sim N(0, \sigma_\mu^2) & \text{with probability } 1-\gamma, \end{cases} \tag{A9}$$

so that

$$\begin{aligned} \mathrm{E}(\mu_n | s_n, \gamma) &= \gamma s_n \\ \mathrm{Var}(\mu_n | s_n, \gamma) &= \mathrm{E}(\mu_n^2 | s_n, \gamma) - [\mathrm{E}(\mu_n | s_n, \gamma)]^2 = (\gamma - \gamma^2) s_n^2 + (1-\gamma)\sigma_\mu^2. \end{aligned}$$

Let $\mathrm{E}(\gamma)$ denote the investor's expected perception of his own skill, and let $\mathrm{Var}(\gamma)$ denote the variance associated with this perception. Using the law of iterated expectations, the expected return on stock n after observing the signals is equal to

$$\mathrm{E}(r_n | S) = \mathrm{E}[\mathrm{E}(\mu_n | s_n, \gamma)] = \mathrm{E}(\gamma) s_n. \tag{A10}$$

Using the variance decomposition rule, the perceived variance of stock n's return is

$$\begin{aligned} \mathrm{Var}(r_n | S) &= \sigma_e^2 + \mathrm{E}[\mathrm{Var}(\mu_n | s_n, \gamma)] + \mathrm{Var}[\mathrm{E}(\mu_n | s_n, \gamma)] \\ &= \sigma_e^2 + \sigma_\mu^2 + \mathrm{E}(\gamma)(s_n^2 - \sigma_\mu^2) - s_n^2 [\mathrm{E}(\gamma)]^2 + s_n \mathrm{Var}(\gamma). \end{aligned} \tag{A11}$$

The perceived covariance between returns on stocks i and j is

$$\begin{aligned} \mathrm{Cov}(r_i, r_j | S) &= \mathrm{E}[\mathrm{Cov}(\mu_i, \mu_j | S, \gamma)] + \mathrm{Cov}[\mathrm{E}(\mu_i | S, \gamma), \mathrm{E}(\mu_j | S, \gamma)] + \mathrm{Cov}(e_i, e_j | S) \\ &= \mathrm{E}[\mathrm{E}((\mu_i - \gamma s_i)(\mu_j - \gamma s_j) | S, \gamma)] + \mathrm{Cov}[\gamma s_i, \gamma s_j] + \rho_e \sigma_e^2 \\ &= \mathrm{E}[\gamma^2 s_i s_j - \gamma^2 s_i s_j - \gamma^2 s_i s_j + \gamma^2 s_i s_j] + s_i s_j \mathrm{Var}(\gamma) + \rho_e \sigma_e^2 \\ &= s_i s_j \mathrm{Var}(\gamma) + \rho_e \sigma_e^2. \end{aligned} \tag{A12}$$

For simplicity, we assume that each manager knows his own γ, so that $\mathrm{E}(\gamma) = \gamma$ and $\mathrm{Var}(\gamma) = 0$. The expression for the optimal weights in equation (34) follows immediately.

References

Baks, Klaas P., 2002, On the performance of mutual fund managers, Working paper, Emory University.

Berk, Jonathan B., and Richard C. Green, 2002, Mutual fund flows and performance in rational markets, Working paper, University of California, Berkeley.

Bollen, Nicolas P. B., and Jeffrey A. Busse, 2002, Short-term persistence in mutual fund performance, Working paper, Vanderbilt University.

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679–698.

Busse, Jeffrey A., and Paul J. Irvine, 2002, Bayesian alphas and mutual fund persistence, Working paper, Emory University.

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

Carhart, Mark M., Ron Kaniel, David K. Musto, and Adam V. Reed, 2002, Learning for the tape: Evidence of gaming behavior in equity mutual funds, *Journal of Finance* 57, 661-693.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers, 2000, The value of active mutual fund management: An examination of the stockholdings and trades of fund managers, *Journal of Financial and Quantitative Analysis* 35, 343-368.

Daniel, Kent D., Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

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

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.

Ferson, Wayne E., and Kenneth Khang, 2002, Conditional performance measurement using portfolio weights: Evidence for pension funds, *Journal of Financial Economics* 65, 249–282.

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

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: An analysis of quarterly portfolio holdings, *Journal of Business* 62, 394–415.

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

Grinblatt, Mark, and Sheridan Titman, 1993, Performance measurement without benchmarks: An examination of mutual fund returns, *Journal of Business* 66, 47–68.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088–1105.

Haugen, Robert A., and Josef Lakonishok, 1988, *The incredible January effect: The stock market's unsolved mystery*, Homewood, IL, Dow Jones - Irwin.

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

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2002, "Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers", Working paper, Stanford University.

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

Jones, Christopher S., and Jay Shanken, 2002, Mutual fund performance with learning across funds, Working paper, University of Southern California.

Lakonishok, Josef, Andrei Shleifer, Richard Thaler, and Robert W. Vishny, 1991, Window dressing by pension fund managers, *American Economic Review* 81, 227–231.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23-44.

Musto, David K., 1997, Portfolio disclosures and year-end price shifts, *Journal of Finance* 52, 1563–1588.

Musto, David K., 1999, Investment decisions depend on portfolio disclosures, *Journal of Finance* 54, 935–952.

Myers, Mary M., James M. Poterba, John B. Shoven, and Douglas A. Shackelford, 2001, Copycat funds: Information disclosure regulation and the returns to active management in the mutual fund industry, Working paper, University of Chicago.

Pástor, Ľuboš, and Robert Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.

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

Shumway, Tyler, 1997, The delisting bias in CRSP data, *Journal of Finance* 52, 327–340.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Stambaugh, Robert F., 1997, Analyzing investments whose histories differ in length, *Journal of Financial Economics* 45, 285–331.

Stambaugh, Robert F., 2002, Inference about survivors, Working paper, Wharton.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *Journal of Finance* 54, 581-622.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses, *Journal of Finance* 55, 1655-95.