Which Factors Matter to Investors? Evidence from Mutual Fund Flows

Brad M. Barber
Graduate School of Management, University of California, Davis

Xing Huang
Broad College of Business, Michigan State University

Terrance Odean
Haas School of Business, University of California, Berkeley

When assessing a fund manager's skill, sophisticated investors will consider all factors (priced and unpriced) that explain cross-sectional variation in fund performance. We investigate which factors investors attend to by analyzing mutual fund flows as a function of recent returns decomposed into alpha and factor-related returns. Surprisingly, investors attend most to market risk (beta) when evaluating funds and treat returns attributable to size, value, momentum, and industry factors as alpha. Using proxies for investor sophistication (wealth, distribution channels, and periods of high investor sentiment), we find that more sophisticated investors use more sophisticated benchmarks when evaluating fund performance. (*JEL* G11, G12, G23)

Received August 14, 2015; accepted May 3, 2016 by Editor Stefan Nagel.

Most mutual fund investors allocate their savings to actively managed mutual funds, which seek to beat the market through some combination of fundamental and/or technical analysis. In theory, when assessing a fund manager's skill, investors should consider all factors that explain cross-sectional variation in fund performance, regardless of whether the factors are priced or unpriced (Grinblatt and Titman 1989; Pástor, and Stambaugh 2002a).

In this paper, we investigate whether investors tend toward commonly used factors and industry tilts of mutual funds when assessing fund managers. The

We appreciate the comments of Joseph Chen, Susan Christoffersen, Richard Evans, Amit Goyal, David Hirshleifer, Petri Jylha, Jose Menchero, Stefan Nagel, Stephan Siegel, Jonathan Reuter, Zheng Sun, and seminar participants at Case Western University, Emory University, SAIF, Tulane University, UC Irvine, Vanderbilt University, the University of Stavanger, 2014 Helsinki Finance Summit, 2014 Rotterdam Behavioral Finance Conference, 2014 European Finance Association Meetings, 2014 Miami Behavioral Finance Conference, the Spring 2014 JOIM Conference, 2015 CICF, 2016 Citrus Conference, 2016 SFS Finance Cavalcade, and the Ben Graham Centre's 4th Intelligent Investing Symposium. All errors are our own. Supplementary data can be found on *The Review of Financial Studies* web site. Send correspondence to Brad M. Barber, Graduate School of Management, UC Davis, 1 Shields Avenues, Davis, CA 95616, telephone: 530-752-0512. E-mail: bmbarber@ucdavis.edu.

doi:10.1093/rfs/hhw054 Advance Access publication June 21, 2016

most surprising result to emerge from our analysis is the observation that flows do not respond as strongly to returns related to a fund's market risk (or beta). Consistent with this observation, we find that CAPM alphas are the best predictor of flows among competing performance evaluation models. As we discuss below, prior work has documented that fund flows respond to factor-related returns. We extend this work by documenting that the flow response to factor-related returns (size, value, momentum, and industry) is nearly as strong as the response to a fund's alpha. In auxiliary analyses, we document that investors differ in their response to performance; more sophisticated investors use more sophisticated benchmarks. Since investors should use all factors when assessing fund performance, these results suggest that investors with limited training or means are less equipped to evaluate fund managers.

To understand why investors should adjust for factor-related returns when assessing performance, consider size-related returns. Historically, the returns of small stocks have been correlated and small stocks have earned higher average returns compared with large stocks. A sophisticated investor will consider size when evaluating fund manager skill. In a year in which small stocks outperform large stocks, the investor will not conclude that all small cap fund managers are highly skilled. It does not matter whether the investor considers the return premium associated with size to result from risk, mispricing, or frictions. A sophisticated investor will not confuse skill with returns that could be earned through passive investments (e.g., in small cap index funds).[1]

Over the past twenty years, models such as the Fama-French three-factor model (Fama and French 1993) and its four-factor cousin (Carhart 1997) have become academic standards. The factors in these models have been shown to be empirically priced; stocks with higher market risk, smaller capitalization, higher book-to-market ratios, and recent momentum have earned greater returns on average. Though controversy exists within the profession as to whether the higher returns associated with small stocks, high book-to-market stocks, and positive momentum stocks are due to risk or mispricing, sophisticated investors should consider these factors when assessing fund manager skill.[2]

Because investors should consider all factors when assessing fund manager skill, one cannot infer that investors associate the factors they attend to with risk. However, investors are unlikely to ignore factors that they do associate with

---

[1] In this paper, we analyze fund flows in the U.S. equity mutual fund industry after 1997, when passive investment vehicles were available for broad market indexes, large cap, small cap, value, and growth. A sophisticated investor will also consider unpriced factors when evaluating mutual fund performance. Consider two industries that earn, on average, similar returns, but perform well in different periods. Funds concentrated in one of the two industries will perform better in some periods, but not in others. The investor will not attribute these periodic performance differentials to fund manager skill. Even factor returns that cannot be captured through passive investments should not be treated as alpha. Consider momentum. An investor might need to rely on active management to capture momentum returns. However, a sophisticated investor would not mistake recent positive or negative momentum returns as indicative of managerial skill. Rather, the investor wishing to capture long-term momentum returns will invest in funds with high momentum loadings (and low fees).

[2] For the two sides of this debate, see Fama and French (2004) and Hirshleifer (2001).

risk unless the costs of attending to those factors are higher than the benefits. Our empirical analysis proceeds in two steps. To set the stage, we estimate mutual fund alphas using six competing empirical models of managerial skill: market-adjusted returns, the capital asset pricing model, the Fama and French (1993) three-factor model, a four-factor model that adds momentum (Carhart 1997), a seven-factor model that adds the three industry factors of Pástor, and Stambaugh (2002a, 2002b), and a nine-factor model that adds profitability and investment factors (Fama and French 2015). In simple linear regressions of fund flows on the six performance measures, we find that the partial effect of CAPM alpha on fund flows is roughly double that of its nearest competitor (market-adjusted returns). To verify the robustness of this result, we then exploit cases in which a fund's ranking diverges across models to identify the model investors most commonly use to evaluate mutual fund performance. We use these cases to run a horse race of the six competing asset-pricing models. Our empirical tests involve pairwise comparisons of competing models, in which we regress monthly flows of new money on decile ranks of prior performance estimated from the competing models. In general, we find greater flows to mutual funds with higher ranks based on CAPM alpha than to funds with higher ranks based on competing models.

In our second series of tests, we decompose the returns of a fund into eight components: seven factor-related returns (market (beta), size, value, momentum, and three industry factors) and the fund's alpha, all estimated using a seven-factor model. Here, we find that returns related to a fund's beta do not generate the same flows as the fund's alpha or other factor-related returns. We find some evidence that investors attend to the value, size, and industry tilts of a fund when assessing managerial skill, but these effects are much weaker than those we observe for a fund's beta.

Sophisticated investors should attend to factor-related returns when assessing managerial skill. Viewed through this lens, our analysis is an investigation into how sophisticated investors are in assessing managerial skill. In aggregate, mutual fund investors do not attend to many aspects of fund performance. To test whether this lack of attention to factor returns is related to investor sophistication, we use three proxies for investor sophistication. First, we split our sample into direct versus broker sold. Del Guercio and Reuter (2013) document that broker-sold mutual funds, which tend to have a less sophisticated investors clientele, experience flows that are more responsive to a fund's market-adjusted return than to its four-factor alpha. Our analysis of how investors attend to factor-related returns across the two distribution channels echo their results; investors in the broker-sold channel respond more to factor-related returns than do investors in the direct-sold channel. Second, as suggested by Chiu and Kini (2014), we use periods of high mutual fund inflows as an indication of periods with high levels of investor sentiment and conjecture that less sophisticated investors trade more during these periods. Third, motivated by the evidence that correlates wealth with trading ability

(Barber and Odean 2000; Geng et al. 2014), diversification (Calvet, Campbell, and Sodini 2007), and the disposition effect (Dhar and Zhu 2006), we use wealth as a measure of investor sophistication in analyses that deploy data from a large discount broker over the 1991 to 1996 period. Consistent with the hypothesis that sophisticated investors use more sophisticated models for assessing fund performance, we consistently find that the flows of more sophisticated investors are less responsive to factor-related returns.

Prior work documents a strong positive relationship between mutual fund flows and a variety of past performance measures, including market-adjusted returns and alphas based on different factor models.[3] However, this literature does not address which of the performance measures is the best predictor of flows. For example, investors who respond solely to market-adjusted returns of mutual funds could drive the generally positive relation between flows and various factor-based measures of performance. But market-adjusted returns are highly correlated with the alternative performance measures. Thus, a researcher who regresses flows on a particular alpha measure may observe a positive relationship between flows and the measure, not because investors are paying attention to that measure but because the relationship is correlated with the measure to which they are tending—market-adjusted returns. We are able to examine which performance measures best predict flows by pitting measures based on competing models against one another and decomposing mutual fund returns into factor-related returns and alpha.

Barberis and Shleifer (2003) propose that rather than focusing on market-adjusted returns or on alphas, investors categorize assets into styles and do not distinguish between assets within a style. Using the nine Morningstar-style boxes as a proxy for styles, Teo and Woo (2004) confirm that flows into funds within a style category are correlated with the past returns of that style category. Their results are evidence that investors reward managers for returns attributable to size and value styles. However, an open question remains: do investors attend to factors or styles at all since investors who merely assess funds based on their market-adjusted returns could drive these results. Consistent with Teo and Woo (2004), we confirm that flows chase style category returns. However, in contrast with the style-investing story, we find that flows are as, or are more, responsive to deviations from style category returns as to the style category returns themselves.

[3] Researchers have used a variety of return benchmarks to study various mutual fund investor and managerial behaviors. Examples of studies using raw returns include Bergstresser and Poterba (2002) (tax-adjusted performance), Coval and Stafford (2007) (fund-flow price pressure relationship), Del Guercio and Tkac (2008) (Morningstar rating changes), and Ivkovic and Weisbenner (2009) (differential sensitivity of in- and outflows to relative and absolute performance). Some that use market-adjusted returns include Chevalier and Ellison (1997) (strategic alteration of fund risk), Karceski (2002) (overweighting of high beta stocks), Barber, Odean, and Zheng (2005) (retail investor sensitivity to fees), and Spiegel and Zhang (2013) (alternative flow measure). Some that use alpha estimates include Khorana (2001) (fund manager replacements), Del Guercio and Tkac (2002) (retail investor versus pension fund behavior), Lynch and Musto (2003) (discarded strategies), Nanda, Wang, and Zheng (2004) (star spillover for fund families), Keswani and Stolin (2008) (smart-money effect in the United Kingdom), Gil-Bazo and Ruiz-Verdu (2009) (performance fee relationship), and Sensoy (2009) (mismatched style indices).

Our results largely support the story that unsophisticated investors chase market-adjusted performance with one surprising exception: market risk exposure. Flows are much less responsive to returns due to a fund's market risk (beta) than to other components of return. For the most part, investors do not reward fund managers for returns attributable to a fund's beta. Furthermore, investors who are likely to be more sophisticated, such as those who pay lower fees, are the least likely to reward managers for positive returns attributable to beta.

The mechanism by which investors attend to a fund's market beta when assessing performance is a mystery, though we are able to reject several potential explanations. Style chasing does not explain this result for two reasons. First, when we include category-month fixed effects, which absorb variation in fund flows across Morningstar-style boxes, our main results are largely unaffected. Second, the average beta varies little across Morningstar-style boxes. Morningstar's ubiquitous star ratings of mutual funds have a large impact on fund flows, but do not explain the proportionately weak response to returns related to a fund's market risk. While the inclusion of star ratings in our regressions dampens the relation between flows and the components of a fund's returns (because star ratings are highly correlated with returns), the relative importance of the return components is similar to what we observe in our main results. Morningstar does provide information on a fund's beta and alpha with respect to various market indexes, but this information is not salient on Web sites and would require knowledge of both modern portfolio theory and Morningstar's detailed fund statistics to influence flows materially.

1. Literature Review

1.1 Literature on fund flows

Our results fit into the large literature on mutual fund flows. Early work establishes that fund flows respond to fund returns (Ippolito 1992; Chevalier and Ellison 1997; Sirri and Tufano 1998). Moreover, the relation between fund flows and returns tends to be convex; positive returns garner more new flows than those lost to negative returns (Chevalier and Ellison 1997; Sirri and Tufano 1998). Brown, Harlow, and Starks (1996) and Chevalier and Ellison (1997) argue that mutual funds respond to these implicit incentives by altering the riskiness of their funds so as to secure a favorable year-end ranking. As noted above, this stream of research uses various measures of mutual fund performance ranging from raw returns to multifactor alphas.

An emerging literature goes beyond simple flow-return relations. Clifford et al. (2013) focus on the impact of total risk (measured as a fund's trailing monthly standard deviation of returns) on fund flows and separately analyze inflows and outflows. They document that *both* inflows and outflows are positively related to total risk. In contrast, we investigate whether investors differentially respond to the components of a fund's return that are arguably a result of the risk associated with the fund. Huang, Wei, and Yan (2012) propose that investors should account for the precision of alpha estimates when allocating capital to mutual funds. They provide empirical support consistent with this hypothesis and argue that the impact of precision on flows is more pronounced for sophisticated investors. In a spirit more similar to our work, De Andrade (2009) infers from flows investors' differential sensitivity to risk in up and down markets. He finds investors prefer funds with low down-market betas and suggests that investors ". . . seek portfolio insurance, in addition to performance."

Mutual funds appear to pick benchmarks or adopt names that garner flows. Sensoy (2009) documents that one-third of the actively managed U.S. equity mutual funds specify a benchmark index in the fund prospectus that does not match the fund's actual style. Moreover, he documents that fund flows respond to these mismatched benchmarks. Cooper, Gulen, and Rau (2005) document that mutual funds that change names to reflect a hot investment style garner additional fund flows. In contrast to the inquiry into the self-selected benchmarks of mutual funds, we ask a more general question: how do investors adjust for return factors when evaluating fund performance? Note that the positive evidence in Sensoy (2009) and Cooper, Gulen, and Rau (2005) that investors pay attention to self-selected benchmarks and fund names does not address the more general question of what factors investors attend to when picking actively managed mutual funds.

In a recent working paper, Agarwal, Green, and Ren (2015) adapt our methods to hedge funds and analyze the relation between hedge fund flows and returns. They also find that the CAPM alpha consistently wins a model horse race in predicting hedge fund flows. While factor-related returns garner flows for hedge funds, the relations are generally weaker than those we document for mutual fund investors. For example, when they decompose returns into alpha, traditional risks (e.g., market and size), and exotic risk (e.g., option factor risks), they find that traditional risks yield flow-return relations that are about half of that associated with alpha, while exotic risks yield flow-return relations that are generally greater. These results dovetail neatly with our interpretation that more sophisticated investors use more sophisticated methods in two ways. First, hedge fund investors, who are likely more sophisticated than mutual fund investors, seem to attend to a wider variety of factor-related returns when assessing performance compared with mutual fund investors. Second, hedge fund investors attend to traditional risks, which are easily measured and replicated in standard investments, more than to exotic risks.

1.2 Berk and Van Binsbergen (2016)

In independent work, Berk and Van Binsbergen (2016) also examine mutual fund performance and flow relationships.[4] As a starting point to their analysis,

[4] In September 2013, Berk and van Binsbergen and we became aware that both sets of authors had independently derived similar findings. Berk and van Binsbergen first posted their paper to SSRN in October 2013. We posted our paper to SSRN in March 2014.

they observe that managerial compensation, which is primarily determined by fund flows (Berk and Green 2004), predicts future fund returns (Berk and Van Binsbergen 2015). At various horizons, they measure the percent of time that the direction of a fund's flow is the same as the sign of its alpha as estimated using a variety of asset pricing models. At each horizon that they analyze, the sign of flows is more likely to have the same sign as the alpha from the CAPM model than from the alpha calculated using competing asset pricing models. While Berk and Van Binsbergen (2016) measure the correspondence between the sign of alpha under different risk models and the sign of flows, we primarily focus on the sensitivity of flows to components of returns attributable to market risk, size tilts, book-to-market tilts, momentum tilts, and industry tilts.

Our paper and theirs reach a common conclusion: fund flows are best explained by CAPM alphas than by competing models. However, the papers differ in motivation, methods, and interpretation. Berk and van Binsbergen motivate their analysis as a test of asset pricing models, while we are motivated to learn how sophisticated investors are in their evaluation of fund performance.

Regarding methods, both papers run a horse race of competing models. We extend this analysis by decomposing returns into alpha and factor-related returns. We also investigate whether these results differ, depending on the sophistication of investors using distribution channels as a proxy for investor sophistication. Skeptical that investors are estimating factor exposures and alphas using statistical analyses, we also explore potential mechanisms that investors use to attend to factor-related returns when evaluating fund performance and find evidence that Morningstar category assignments allow investors to attend to the size and value tilts of funds when assessing performance.

Regarding interpretation, Berk and van Binsbergen conclude the CAPM victory in the horse race indicates the CAPM is closest to the "true asset pricing model." In contrast, we argue that investors should consider all factor-related returns—priced and unpriced—when assessing the skill of a fund manager. The observation that investors attend to market risk (though as we show not completely) is both interesting and suggestive that market risk is a risk factor that many investors care about. However, this observation alone is not sufficient evidence to establish market risk as a priced risk factor.

Berk and Van Binsbergen (2016) write, "Because we implement our method using mutual fund data, one might be tempted to conclude that our tests only reveal the preferences of mutual fund investors, rather than all investors. But this is not the case . . . . if our test rejects a particular asset pricing model, we are not simply rejecting the hypothesis that mutual fund investors use the model, but rather, we are rejecting the hypothesis that any investor who could invest in mutual funds uses the model."

Berk and Van Binsbergen (2016) argue that nonmutual fund investors who have access to mutual funds will act to eliminate mispricing in the mutual fund market. However, the mispriced asset in this market is the skill of a fund manager, not necessarily the assets in a fund. Consider a hedge fund manager who identifies a mutual fund manager whose skill has been overvalued by the market. The mutual fund manager has garnered more assets under management than can be justified by the fund manager's ability. How can the hedge fund manager exploit this mispricing? When owning the mutual fund, the hedge fund manager can sell the shares. However, if shares are not owned, the hedge fund manager cannot short the mutual fund. And, though the mutual fund manager's skill is overpriced, this does not mean that the assets held by the fund are overpriced (e.g., imagine a market in which all assets are efficiently priced but active fund managers charged high fees). Thus, there may be no positive net present value opportunity available for the hedge fund manager to exploit. What if, instead, the hedge fund manager identifies a mutual fund manager whose skill has been undervalued by the market (i.e., a manager who is skilled and whose strategies can potentially support larger positions)? The hedge fund manager could directly invest in the mutual fund. More likely, though, the hedge fund manager will try to copy the mutual fund manager's strategies through trades in equity (or other) markets. These trades will not show up in mutual fund flow data, and thus mutual fund flow data will not provide information about the hedge fund manager's risk model. So mutual fund flow data do not inform us about the beliefs of nonmutual fund investors. In summary, we do not believe the results in either paper provide much evidence regarding the true asset pricing model. Instead, both papers provide evidence on how investors assess fund performance.

2. Data and Methods

2.1 Fund flows

Our dependent variable of interest is fund flows and is estimated using data from the CRSP mutual fund database. The CRSP database contains monthly data beginning in 1991. Since we use an estimation window of five years in our empirical analysis described below, our sample period covers the years 1996 to 2011 and includes about 4,000 equity funds. Because we are interested in investors who are attempting to identify managerial skill in their fund allocation decisions, we exclude from our analysis funds that CRSP identifies as index funds.

Following the majority of the prior literature on fund flows, we calculate flows for fund p in month t as the percentage growth of new assets, assuming that all flows take place at the end of the month:

$$F_{pt} = \frac{TNA_{pt}}{TNA_{p,t-1}} - (1 + R_{pt}), \tag{1}$$

where TNA_{pt} is the total net assets under management of fund p at the end of month t, and R_{pt} is the total return of fund p in month t.[5] We aggregate the

[5] In the rare cases in which two funds merge into a single fund during month t, beginning-of-period TNA is set equal to the combined assets of the two funds, while end-of-period TNA is set equal to the merged assets of the remaining fund.

flows and compute the value-weighted returns across multiple share classes within one fund portfolio. We restrict our analysis to funds with total net assets data (required to calculate fund flows), a minimum of $10 million in assets at the end of month $t-1$, and month t flows of more than -90% and less than 1,000%. We merge the CRSP data with the fund-style box from Morningstar equity fund universe by matching on fund CUSIPs. Our final sample consists of observations with successful merges.

2.2 Mutual fund performance

When selecting an equity mutual fund that actively manages its investments, an investor seeks to identify a mutual fund that is able to deliver an alpha, where the fund's alpha is estimated after stripping out any fund return that can be traced to the fund's exposure to factors known by the investor to affect cross-sectional equity returns (e.g., size).[6] What is less clear, and the focus of our research, is which factors mutual fund investors consider when estimating alpha. At one extreme, investors may simply rank funds based on their raw returns; at the other extreme, they may rank funds based on a multifactor model of returns, such as those commonly found in the academic literature on asset pricing.

We begin by running a horse race between six competing models that investors might reasonably employ when evaluating the performance of mutual funds: market-adjusted returns (MAR), the capital asset pricing model (CAPM), the Fama and French (1992) three-factor model (3F), which adds size and value factors, a four-factor model (4F) that adds momentum (Carhart 1997), a seven-factor model (7F) that adds the three industry factors of Pástor, and Stambaugh (2002a, 2002b), and a nine-factor model (9F) that adds profitability and investment factors (Fama and French 2015). In many cases, these models yield similar rankings of mutual funds (i.e., the six performance measures are highly correlated). However, we exploit the cases in which rankings differ across models to answer the question of which model best explains the choices that investors make when allocating capital to actively managed mutual funds.

We use monthly return and flow data on over 3,900 U.S. diversified equity mutual funds actively managed for the period 1996 to 2011.[7] We proceed in two steps. First, we estimate the abnormal return (alpha) for each mutual fund using each of the six competing models. Alpha estimates are updated monthly based on a rolling estimation window. Consider the seven-factor model, which includes factors related to market, size, value, momentum, and three industry

---

[6] Some caveats are worth acknowledging. Ferson and Lin (2014) argue that investors might have different alphas for the same fund if markets are incomplete and investors have different marginal rates of substitution. Cremers Petajisto, and Zitzewitz (2012) document that some indexes have positive alphas, suggesting alphas do not precisely measure fund manager skill. Berk and Van Binsbergen (2015) argue that the value-add of an active fund can be measured relative to the passive funds available to investors; this measurement is time varying.

[7] The relatively small number of funds in our sample is a result of data requirements. Most importantly, we require a five-year history of fund returns for inclusion in our sample; this requirement is necessary to estimate the factor tilts of a mutual fund.

factors in the estimation of a fund's return. In this case, for each fund in month t, we estimate the following time-series regression using sixty months of returns data from months $\tau = t-1, t-60$:

$$(R_{p\tau} - R_{f\tau}) = \alpha_{pt} + \beta_{pt}(R_{m\tau} - R_{f\tau}) + s_{pt}SMB_{\tau} + h_{pt}HML_{\tau} \\ + s_{pt}UMD_{\tau} + \sum_{k=1}^{3} i_{pt}^{k} IND_{\tau}^{k} + e_{p\tau}, \tag{2}$$

where $R_{p\tau}$ is the mutual fund return in month τ, $R_{f\tau}$ is the return on the risk-free rate, $R_{m\tau}$ is the return on a value-weighted market index, SMB_{τ} is the return on a size factor (small minus big stocks), HML_{τ} is the return on a value factor (high minus low book-to-market stocks), UMD_{τ} is the return on a momentum factor (up minus down stocks), and IND_{τ}^{k} is the return on the kth industry portfolios that measure the industry tilts of a mutual fund.[8] We construct the three industry portfolios by extracting the three main principal components of the Fama-French seventeen industry portfolios as in Pástor, and Stambaugh (2002a, 2002b), which we describe in detail in the Online Appendix. Readers can think of the industry portfolios as long-short portfolios constructed from the seventeen Fama-French industry portfolios that capture common industry returns orthogonal to the other factors we consider. The parameters $\beta_{pt}, s_{pt}, h_{pt}, m_{pt}$, and i_{pt}^{k} represent the market, size, value, momentum, and industry tilts (respectively) of fund p; α_{pt} is the mean return unrelated to the factor tilts; and $e_{p\tau}$ is a mean zero error term. (The subscript t denotes the parameter estimates used in month t, which are estimated over the sixty months prior to month t.) We then calculate the alpha for the fund in month t as its realized return less returns related to the fund's market, size, value, momentum, and industry exposures in month t:

$$\hat{\alpha}_{pt} = \left(R_{pt} - R_{ft}\right) - \left[\hat{\beta}_{pt}\left(R_{mt} - R_{ft}\right) + \hat{s}_{pt}SMB_{t} + \hat{h}_{pt}HML_{t} \right. \\ \left. + \hat{m}_{pt}UMD_{t} + \sum_{k=1}^{3} \hat{i}_{pt}^{k} IND_{t}^{k}\right]. \tag{3}$$

We repeat this procedure for all months (t) and all funds (p) to obtain a time series of monthly alphas and factor-related returns for each fund in our sample. Note that alpha captures returns due to stock selection, as well as those resulting from the timing of factor exposures, relative to average past exposures.

There is an analogous calculation of alphas and return components for the other factor models that we evaluate. For example, we estimate a fund's three-factor alpha using the regression of Equation (2), but drop UMD and IND^{k} as

[8] We obtain the market, size, book-to-market, and momentum factors from Kenneth French's online data library (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

independent variables. To estimate the CAPM alpha, we retain only the market excess return as an independent variable. To estimate the market-adjusted return (*MAR*), we simply subtract the market return from the fund return.

2.3 Horizon for performance evaluation

With rational expectations, investors respond to new information about the skill of fund managers, rewarding skilled managers with new deposits and penalizing poor managers with withdrawals. How investors should weight past returns when assessing fund manager skill is less clear; investors need to balance relevance (recent returns are likely more informative about the manager's current ability) versus the signal-to-noise ratio (short-term returns are mostly noise with very little signal about returns). In addition, numerous frictions (e.g., inertia, inattention, and transaction costs) would also create delays in the response of flows to fund performance. This creates an empirical complication in our analysis, as we must make a decision about what performance horizon to analyze when we compare models.

To address this issue, we empirically estimate the rate of decay in the flow-return relation using monthly fund returns. To set the stage, we estimate the following unrestricted model of the flow-return relation:

$$F_{pt} = a + \sum_{s=1}^{18} b_s MAR_{p,t-s} + cX_{pt} + \mu_t + e_{pt}, \tag{4}$$

where F_{pt} are flows for fund p in month t and $MAR_{p,t-s}$ represents the lagged market-adjusted return for the fund at lag s, where $s=1$ to 18 months. We settle on a lag length of 18 months based on the Akaike information criterion (AIC) of models, where we vary the number of lagged returns from 12 to 48. We include a matrix of control variable (X), which yields a vector of coefficient estimates (c). As controls, we include lagged fund flows from month t-19, lags of a fund's total expense ratio (TNA-weighted across share classes), a dummy variable for no-load funds (if all share classes are no-load funds), a fund's return standard deviation estimated over the prior twelve months, the log of fund size in month t-1, and the log of fund age in month t-1. We also include time fixed effects (μ_t).

This regression yields a series of coefficient estimates, b_s, that represent the relation between flows in month t and the fund's market-adjusted return lagged s months, $s=1,18$. In Figure 1, the red line graphs the estimated b coefficients (y axis) at various lags (x axis) and shows a clear decay in the relation between past returns and fund flows. Recent returns are more important than distant returns.

To capture this decay in the flow-return relation parsimoniously, we model the flow-return relation using an exponential decay model, with decay rate λ:

$$F_{pt} = a + b\sum_{s=1}^{18} e^{-\lambda(s-1)} MAR_{t-s} + cX_{pt} + \mu_t + e_{pt}. \tag{5}$$

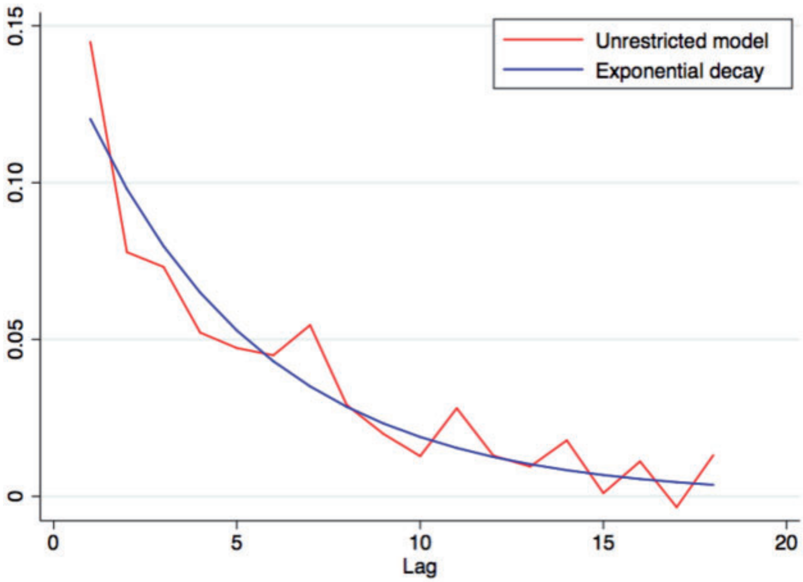

Figure 1
Decay in fund flow relation
The graph presents the regression coefficient estimates (y axis) by horizon (x axis) for two models of monthly fund flows (dependent variable): (1) an unrestricted model (with eighteen lags of monthly fund returns and individual coefficient estimates on each lagged return) and (2) an exponential decay model (for which the coefficient estimates on the lagged monthly returns are restricted to follow an exponential decay function with decay parameter lambda).

The key parameters of interest in this model are b, which measures the relation between a weighted sum of the previous eighteen monthly market-adjusted returns, and λ, which measures the decay in the return-flow relation over time. In Figure 1, the smooth blue line represents the estimated decay function, which closely tracks the unconstrained estimates from the regression of Equation (4).

We apply this decay function to the monthly alphas and factor-related returns for each fund-month observation. For example, when considering flows for funds in month t, we calculate the fund's alpha as a weighted average of the prior eighteen monthly alphas:

$$ALPHA_{pt} = \frac{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)} \hat{\alpha}_{t-s}}{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)}}, \tag{6}$$

where monthly alpha estimates are based on one of the six models that we evaluate and the exponential decay is based on the estimates from Equation (5).

2.4 Model horse race

We are interested in testing whether the mutual fund investment choices of investors are more sensitive to alphas calculated using one of six models. We begin by estimating a simple linear regression of fund flows on performance measures from the six competing models.

To address concerns about nonlinearities in the flow-return relation more robustly, we also consider pairwise comparisons of the competing models. To do so, we proceed as follows. In each month during our sample period we create deciles of mutual fund performance based on each of the six alpha estimates weighted by the eighteen-month exponential decay as described previously. Decile 10 contains the best performing funds, and decile 1 contains the worst funds. Thus, we ultimately have a time-series across months of six decile ranks (corresponding to the ranks based on the six competing models) for each mutual fund.

To be specific, consider the pairwise comparison of the CAPM and the three-factor model. We estimate the relation between flows and a fund's decile ranking based on the CAPM and three-factor models by estimating the following regression:

$$F_{pt} = a + \sum_i \sum_j b_{ij} D_{ijpt} + cX_{pt} + \mu_t + \varepsilon_{pt}, \tag{7}$$

where the dependent variable (F_{pt}) is the fund flow for mutual fund p in month t. D_{ijpt} is a dummy variable that takes on a value of one if fund p in month t is in decile i based on the CAPM and decile j based on the three-factor model. To estimate the model, we exclude the dummy variable for $j=5$ and $i=5$. The matrix X_{pt} represents control variables, and c represents a vector of associated coefficient estimates. The key coefficients of interest are b_{ij}, $i=1,...,10$, and $j=1,...,10$, which can be interpreted as the percentage flows received by a fund in decile i for the CAPM and decile j for the three-factor model relative to a mutual fund that ranks in the fifth decile on both performance measures.

Figure 2 provides a visual representation of the key dummy variables, D_{ijpt}. In the regression, the omitted dummy variable (regression constant) is identified by funds with a decile rank of 5 based on both models (black square). The gray and black cells represent funds with similar performance ranks based on both models. The empirical tests compare the coefficients corresponding to the forty-five lower off-diagonal cells (where funds have better performance based on the CAPM) to the forty-five upper off-diagonal cells (where funds have better performance based on the 3F Alpha). For example, we compare the coefficient estimate on the dummy variable for funds with a CAPM alpha in the ninth decile and 3F alpha in the third decile (red cell, $b_{9,3}$) to funds with a CAPM alpha in the third decile and 3F alpha in the ninth decile (green cell, $b_{3,9}$). To determine whether investors are more sensitive to the CAPM or three-factor alpha, we test the null hypothesis that $b_{ij} = b_{ji}$ for all $i \neq j$. For example, we test the null hypothesis that $b_{9,3} = b_{3,9}$ (i.e., whether funds in the green cell or red cell of

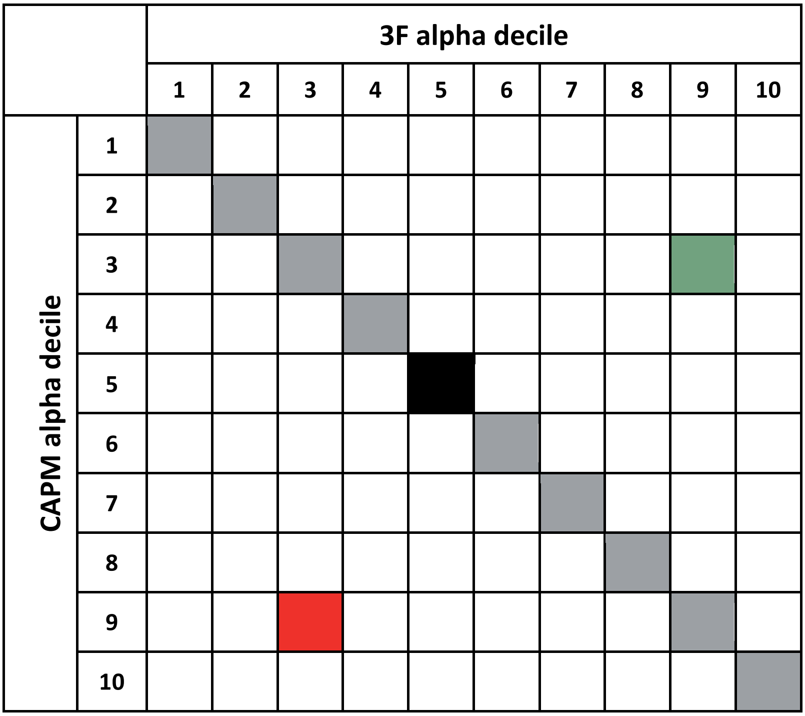

Figure 2
Horse race dummy variables
The figure shows the 100 possible dummy variables for the flow regression that compares relative fund flows based on a fund's CAPM alpha versus three-factor alpha, where ten is a top decile fund and one is a bottom decile fund. In the regression, the omitted dummy variable (regression constant) is funds with a decile rank of five based on both models (black square). The gray and black cells represent funds with similar ranks based on both models. The empirical tests compare the coefficients corresponding to the forty-five lower off-diagonal cells (where funds have better performance based on the CAPM) to the forty-five upper off-diagonal cells (where funds have better performance based on the 3F alpha). For example, we compare the coefficient estimate on the dummy variable for funds with a CAPM alpha in the ninth decile and 3F alpha in the third decile (red cell) to funds with a CAPM alpha in the third decile and 3F alpha in the ninth decile (green cell).

Figure 2 garner more flows). If investors place more weight on the CAPM alpha than on the three-factor alpha, we would expect to reject the null in favor of the alternative hypothesis, $b_{9,3} > b_{3,9}$; conversely, if investors place more weight on the three-factor alpha than on the CAPM alpha, we would reject in favor of the alternative hypothesis, $b_{9,3} < b_{3,9}$. Thus, we test the null hypothesis that the summed difference across all forty-five comparisons is equal to zero, and we calculate a binomial test statistic to test the null hypothesis that the proportion of differences equals 50%.

2.5 Return decomposition

To preview our empirical results, we generally find that CAPM performance ranks best predict fund flows compared with performance ranks based on competing models. This result implies that investors are sufficiently sophisticated to account for market factors when assessing managerial performance. The result does not imply that investors fully account for market-related returns, all investors use the CAPM, or mutual fund investors in aggregate completely ignore factors unrelated to market movements. Our second set of empirical tests addresses these issues by estimating the extent to which investors account for returns related to the factors we consider.

In our main tests, we rearrange Equation (3) to decompose the fund's return into its alpha and factor-related returns.

$$\left(R_{pt}-R_{ft}\right)=\hat{\alpha}_{pt}+\left[\hat{\beta}_{pt}\left(R_{mt}-R_{ft}\right)+\hat{s}_{pt}SMB_t+\hat{h}_{pt}HML_t\right.$$
$$\left.+\hat{m}_{pt}UMD_t+\sum_{k=1}^{3}\hat{i}_{pt}^{k}IND_t^k\right] \quad (8)$$

We base this return decomposition on the seven-factor model.[9] In this return decomposition, the fund's return consists of eight components: the fund's seven-factor alpha and returns related to the fund's market, size, value, momentum, and tilts with respect to three industry portfolios. In month t, we weight each of the return components over the prior eighteen months (t-1 to t-18) using the exponential decay function analogous to the weighting of alphas described previously. For example, consider the portion of the fund's return related to market risk (or beta). We calculate the portion of the fund's return related to market risk as

$$MKTRET_{pt}=\frac{\sum_{s=1}^{18}e^{-\hat{\lambda}(s-1)}\left[\hat{\beta}_{t-s}\left(R_{m,t-s}-R_{f,t-s}\right)\right]}{\sum_{s=1}^{18}e^{-\hat{\lambda}(s-1)}}. \quad (9)$$

There are similar calculations for returns related to the funds size, value, momentum, and three industry tilts, which we label SIZRET, VALRET, MOMRET, INDRET1, INDRET2, and INDRET3 (respectively).

With this return decomposition, we can determine whether investors respond differently to the components of returns by estimating the following panel regression across p funds and t months:

$$F_{pt}=b_0+b_1ALPHA_{pt}+b_2MKTRET_{pt}+b_3SIZRET_{pt}+b_4VALRET_{pt}$$
$$+b_5MOMRET_{pt}+\sum_{k=1}^{3}b_{5+k}INDRETk_{pt}+\gamma X_{pt}+\mu_t+e_{pt}, \quad (10)$$

where b_0 is the regression intercept, e_{pt} is the regression error term, γ is a coefficient vector associated with control variables (X_{pt}), and μ_t represents

[9] We also have estimated results based on the nine-factor model and find qualitatively similar results.

month fixed effects. The controls include the total expense ratio, a dummy variable for no-load, fund's return standard deviation, the log of fund size, the log of fund age, and lagged fund flows.

The parameter estimates of interest in Equation (10) are $b_i, i = 1,8$. For expositional ease, sophisticated investors are those who rely on the seven-factor model when evaluating fund performance, where "sophisticated" is used to describe their ability to account for the common return components when evaluating managerial skill. Sophisticated investors will direct capital based on the fund's alpha, but not returns related to known factors. Thus, we expect $b_1 > 0$, as investors *will* respond to a fund's seven-factor alpha, and $b_i = 0, i = 2,8$, as investors with this sophisticated benchmark *will not* respond to fund returns that can be traced to factor loadings and factor realizations. In contrast, for investors who only consider market risk when assessing fund performance, we expect $b_1 = b_3 = b_4 = b_5 = b_6 = b_7 = b_8 > 0$ and $b_2 = 0$ because investors who only adjust for market risk when assessing fund performance will discount returns that can be traced to market risk, but will treat returns that can be traced to the size, value, momentum, and industry tilts of a fund as alpha.

Because we are measuring these relations using fund-level fund flows, rather than investor-level fund flows, the coefficient estimates can be viewed as the weight placed on a particular factor by mutual fund investors in aggregate. The empirical question addressed by this approach is to *which* factors do investors attend when assessing the skill of a fund manager.[10]

2.6 Sample descriptive statistics

In Table 1, we provide descriptive statistics for our final sample, which consists of nearly 4,000 diversified, actively managed U.S. equity funds. Panel A presents descriptive statistics on fund characteristics across fund-month observations used in our main regression (Jan 1996 to Dec 2011). The average fund has a modestly negative monthly flow during our sample period (−0.53%), but with a standard deviation of 2.25% and interquartile range of more than 2%, the cross-sectional variation in fund flows is considerable. The average fund has total net assets of about \$1.4 billion, though the median fund is considerably smaller (\$396 million). The average age of the fund is 202 months (about 17 years), while the median fund age is 154 months (11.8 years). Our sample tends to be tilted toward larger and older funds since we require a five-year track record to estimate a fund's factor loadings. The average annual expense ratio for sample funds is 1.28%. A large proportion of funds (72%) has either a front-end or back-end load. (Recall that we categorize a fund as having a load if any of its share classes have a load attached to it). The mean monthly return standard deviation of sample funds is 4.92%.

10 As an alternative to decomposing the excess return of each fund into its components, we decompose the seven-factor alpha into components related to the fund's market-adjusted return and factor exposures by rearranging Equation (4). This approach yields qualitatively similar results (see Online Appendix for details).

Table 1
Descriptive statistics for mutual fund sample

	#Obs	Mean	SD	25th perc	Median	75th perc
A: Fund characteristics (fund-month obs.)						
Percentage fund flow	257053	−0.533%	2.254%	−1.620%	−0.609%	0.453%
Fund size ($mil)	257053	1443.500	2941.297	125.132	396.577	1240.571
Fund Age (months)	257053	202.450	148.946	111.000	154.000	225.000
Expense ratio	257053	1.276%	0.438%	0.995%	1.230%	1.517%
Load fund dummy	257053	0.723	0.448	0.000	1.000	1.000
Volatility (t-12 to t-1)	257053	4.926%	2.007%	3.358%	4.705%	6.137%
B. Fund alpha and factor exposures (fund-month obs.)						
Alpha	328705	−0.023%	2.478%	−1.134%	−0.034%	1.087%
Beta	328705	0.945	0.186	0.862	0.960	1.043
Size coefficient	328705	0.196	0.306	−0.049	0.141	0.398
Value coefficient	328705	0.032	0.350	−0.181	0.034	0.256
Momentum coefficient	328705	0.015	0.137	−0.066	0.005	0.085
Industry 1 coefficient	328705	0.041	0.101	−0.013	0.020	0.071
Industry 2 coefficient	328705	0.004	0.086	−0.043	0.003	0.051
Industry 3 coefficient	328705	−0.002	0.100	−0.050	−0.002	0.043
Adjusted R-squared	328705	0.831	0.151	0.784	0.878	0.932
C. Mean descriptive statistics on return components across 175 months (Jan 1996 to Nov 2011)						
ALPHA	175	−0.048%	0.815%	−0.484%	−0.056%	0.384%
MKTRET	175	0.320%	0.254%	0.190%	0.325%	0.461%
SIZRET	175	0.045%	0.365%	−0.193%	0.028%	0.292%
VALRET	175	0.015%	0.229%	−0.083%	0.005%	0.101%
MOMRET	175	−0.004%	0.126%	−0.077%	−0.003%	0.066%
INDRET1	175	−0.012%	0.177%	−0.102%	−0.016%	0.075%
INDRET2	175	0.053%	0.272%	−0.150%	0.040%	0.239%
INDRET3	175	0.021%	0.214%	−0.115%	0.016%	0.153%

(*continued*)

Table 1, panel B, presents descriptive statistics on the estimated alpha and factor loadings from the rolling window regressions, which include the 18-month period preceding this sample period (hence the greater number of observations for the regression statistics). The mean monthly alpha for the prior year is −2.3 bps per month (or about −28 bps per year), and this is consistent with the well-documented aggregate underperformance of mutual funds. The average fund has beta, size, value, and momentum coefficients of 0.95, 0.20, 0.03, and 0.02 (respectively), suggesting that the average fund has close to average market risk with a modest tilt toward small stocks and virtually no tilt toward value stocks and stocks with strong recent returns. The mean industry tilt of mutual funds is also generally close to zero. This is what would be expected if mutual funds in aggregate did not place large bets on particular industries. More importantly, there is considerable cross-sectional variation in factor loadings across funds. The standard deviations of beta, size, value, and momentum loadings are 0.19, 0.31, 0.35, and 0.14 (respectively), while industry loadings have a standard deviation of approximately 0.10.

Since investors evaluate the relative performance of funds at a particular point in time, we first want to verify that the product of factor loadings and factor realizations indeed generate economically meaningful cross-sectional variation in fund returns. To do so, we calculate descriptive statistics on each

Table 1
Continued

D Correlation between fund return components

	ALPHA	MKTRET	SIZRET	VALRET	MOMRET	INDRET1	INDRET2
(a) ALPHA	1						
(b) MKTRET	0.0479	1					
(c) SIZRET	−0.0821	0.0698	1				
(d) VALRET	0.0255	−0.1220	0.0114	1			
(e) MOMRET	−0.2240	−0.0791	−0.0134	0.0225	1		
(f) INDRET1	−0.0321	0.0233	−0.0221	−0.0311	0.0451	1	
(g) INDRET2	−0.1620	−0.0157	0.0115	0.0494	−0.0131	−0.1170	1
(h) INDRET3	−0.1690	0.0288	−0.0516	−0.1350	0.0798	0.0030	−0.0488

E Correlation between fund alphas

	MAR	CAPM	3F	4F	7F	9F
(a) MAR	1					
(b) CAPM	0.92	1				
(c) 3F	0.74	0.78	1			
(d) 4F	0.70	0.73	0.89	1		
(e) 7F	0.65	0.68	0.82	0.86	1	
(f) 9F	0.59	0.64	0.76	0.81	0.89	1

Panel A presents statistics across fund-month observations. Statistics on fund characteristics are across fund-month observations from May 1997 to November 2011, the period in which these data are used in subsequent regression analyses. Percentage fund flow is percentage change TNA from month t-1 to t adjusted for the fund return in month t. The *Load fund* dummy takes a value of one if any share class for the fund has a front- or back-end load.

Panel B presents statistics across fund-month observations from January 1996 to November 2011 of estimated coefficients from monthly rolling regressions using the seven-factor model.

Panel C presents time-series averages of cross-sectional descriptive statistics on monthly fund return components, which are the independent variables of interest in our fund-flow regressions. Returns due to factor tilts of a fund are estimated as the mean monthly factor return times the fund's estimated factor loading. In each month, return components represent an exponentially weighted average of the return component over the prior eighteen months (see the text for details). We first calculate descriptive statistics across funds in each month; the table presents the average of each statistic across months.

Panel D presents the correlation matrix between fund return components based on fund-month observations.

Panel E presents the correlation matrix between annual abnormal return measures calculated from six models: market-adjusted returns (MAR), the capital asset pricing model (CAPM), a three-factor model (3F) that adds size and value factors, a four-factor model (4F) that adds momentum, a seven-factor model (7F) that adds the three industry factors, and a nine-factor model (9F) that adds profitability and investment factors. Percentage fund flow is truncated below −90% and above 1,000%; all other variables are winsorized at the 1% and 99% level.

of the return components, which represent our key independent variables, in two steps. First, in each month during our sample period we calculate the mean, standard deviation, median, and 25th/75th percentile for each variable across funds. Second, we average the monthly statistics across months.

The results of this analysis are presented in Table 1, panel C. Not surprisingly, the seven-factor alpha generates the largest cross-sectional variation in performance (with a standard deviation of 0.815%). However, each of the factor loadings multiplied by the factor realizations for the eighteen months leading to month t generate large variation in the monthly returns earned on mutual funds. For example, the mean monthly return associated with market risk is 32 bps during our sample period, with a standard deviation of 25.4 bps. The average fund does not heavily load on the remaining return factors (size, value, momentum, and industry); thus, the mean return associated with these return factors is small (ranging from −0.4 bps for momentum to 5.3 bps for the second industry factor). More importantly, we observe considerable cross-sectional variation in the returns due to these non-market return factors across funds, with standard deviations ranging from 12.6 bps for momentum to 36.5 bps for value. This variation is the key to our empirical analysis, as we seek to estimate how sensitive investors are to fund returns reasonably attributed to factor returns when selecting actively managed mutual funds.

In panel D, we present the correlation matrix of return components based on overlapping fund-month observations. We are interested to learn whether there is a high degree of correlation among the components of return, as high correlation between the return components would potentially limit our ability to identify whether investors respond differently to the components of returns. The pairwise correlations are generally low (less than 25% in absolute value).

In panel E, we present the correlation matrix of alphas estimated based on the six models that we evaluate: MAR, CAPM, 3F, 4F, 7F, and 9F. In contrast to the correlation matrix of the return decomposition, the correlation across the various alpha estimates is quite high. The high correlations explain why prior studies generally find a positive relation between flows and a variety of performance benchmarks (see footnote 3).

To further assess the reasonableness of our estimated factor loadings and set the stage for the analysis of Morningstar category assignments in moderating fund flows, we present descriptive statistics on factor loadings across Morningstar-style boxes in Table 2. Morningstar categorizes diversified equity funds into one of nine style boxes. The style boxes have two dimensions: size (small, mid, and large) and fund investment style (value, blend, and growth). We expect our factor loadings to line up with a fund's style box assignment, and they do. Across the style boxes, modest variation in beta estimates (panel A) is clear, though growth funds tend to have higher betas than value firms. As expected, small funds have large relative loadings on the SMB factor, while there is modest variation in size loadings across the value dimension (panel B). Similarly, value funds have relatively large loadings on HML, while there is relatively modest variation in value loadings across size categories (panel C). Finally, growth (value) funds tend to have a modest tilt toward stocks with strong (poor) recent returns (panel D), while we observe little difference in the industry tilts across the size or value dimensions of funds.

More importantly, we observe considerable cross-sectional variation in factor loadings within each style box. For example, the cross-sectional standard deviation of beta within each of the nine style boxes (0.141 to 0.253) is similar in magnitude to the overall standard deviation (0.171). Similarly, the cross-sectional standard deviation of momentum loadings within each of the nine style boxes (0.107 to 0.157) is similar in magnitude to the overall standard deviation (0.139). The within category standard deviation in the size (0.167 to 0.220) and value loadings (0.236 to 0.381) are somewhat less than the overall standard deviation (0.303 for size and 0.321 for value). This is expected since the categories explicitly sort on funds' size and value tilts.

Table 2
Descriptive statistics by Morningstar-style box

	Large	Medium	Small	Agg by value
A. Beta				
Value	0.932	0.843	0.888	0.908
	(0.144)	(0.253)	(0.170)	(0.179)
Blend	0.929	0.916	0.945	0.929
	(0.143)	(0.189)	(0.141)	(0.152)
Growth	0.944	0.933	0.985	0.949
	(0.163)	(0.205)	(0.166)	(0.177)
Agg by size	0.936	0.908	0.955	0.932
	(0.152)	(0.217)	(0.164)	(0.171)
B. Size coefficient				
Value	−0.02	0.247	0.637	0.112
	(0.167)	(0.220)	(0.198)	(0.285)
Blend	0.004	0.319	0.652	0.154
	(0.179)	(0.202)	(0.189)	(0.301)
Growth	0.042	0.338	0.653	0.227
	(0.192)	(0.205)	(0.198)	(0.305)
Agg by size	0.013	0.312	0.650	0.176
	(0.183)	(0.211)	(0.195)	(0.303)
C. Value coefficient				
Value	0.186	0.271	0.311	0.218
	(0.254)	(0.351)	(0.236)	(0.279)
Blend	0.056	0.18	0.204	0.099
	(0.253)	(0.308)	(0.251)	(0.271)
Growth	−0.09	−0.068	−0.116	−0.089
	(0.295)	(0.369)	(0.310)	(0.319)
Agg by size	0.033	0.072	0.060	0.046
	(0.292)	(0.381)	(0.336)	(0.321)
D. Momentum coefficient				
Value	−0.055	−0.063	−0.051	−0.056
	(0.108)	(0.145)	(0.118)	(0.118)
Blend	−0.014	−0.028	−0.009	−0.015
	(0.107)	(0.138)	(0.116)	(0.115)
Growth	0.059	0.092	0.086	0.073
	(0.131)	(0.157)	(0.140)	(0.141)
Agg by size	0.004	0.027	0.032	0.013
	(0.127)	(0.165)	(0.142)	(0.139)
E. Industry 1 coefficient				
Value	0.048	0.055	0.027	0.047
	(0.076)	(0.101)	(0.070)	(0.082)
Blend	0.054	0.051	0.027	0.050
	(0.082)	(0.099)	(0.069)	(0.084)
Growth	0.035	0.062	0.034	0.042
	(0.097)	(0.149)	(0.120)	(0.117)
Agg by size	0.045	0.058	0.030	0.046
	(0.087)	(0.128)	(0.099)	(0.100)

(*continued*)

However, there is still considerable cross-sectional variation in the size and value loadings within a category, and we later exploit this to understand whether financial intermediaries, such as Morningstar, provide a mechanism by which investors can tend to factor-related returns when assessing fund performance.

Table 2
Continued

	Large	Medium	Small	Agg by value
F. Industry 2 coefficient				
Value	−0.029	−0.017	−0.011	−0.024
	(0.072)	(0.091)	(0.062)	(0.075)
Blend	−0.008	−0.001	−0.008	−0.007
	(0.076)	(0.082)	(0.059)	(0.075)
Growth	0.011	−0.002	−0.023	0.001
	(0.097)	(0.110)	(0.083)	(0.099)
Agg by size	−0.006	−0.005	−0.016	−0.008
	(0.086)	(0.099)	(0.073)	(0.087)
G. Industry 3 coefficient				
Value	−0.013	−0.009	−0.007	−0.012
	(0.081)	(0.100)	(0.082)	(0.086)
Blend	0.001	−0.002	−0.022	−0.003
	(0.086)	(0.102)	(0.087)	(0.090)
Growth	0.012	0.008	−0.004	0.008
	(0.109)	(0.132)	(0.093)	(0.113)
Agg by size	0.001	0.002	−0.010	0.000
	(0.095)	(0.119)	(0.090)	(0.100)

This table presents the mean and standard deviation of estimated factor coefficients (beta, size, value, momentum, and industry factors) across fund-month observations for each of the nine Morningstar-style boxes. Factor coefficients (beta, size, value, momentum, and industry factors) are estimated using a five-year rolling regression of fund excess return (market less risk-free return) on market, size, value, momentum, and industry factors.

3. Results

3.1 Model horse race

To set the stage, we begin by estimating a simple linear regression in which the dependent variable is the percentage of fund flow and the key independent variables are the six performance measures described in Table 1: market-adjusted returns and alphas from the CAPM, three-, four-, seven-, and nine-factor models. Controls include lagged fund flows from month t-19, lagged values of log of fund size, log of fund age, expense ratio, load fund dummy, return volatility, and month fixed effects.

The results of this analysis are presented in Table 3. In Column 1, we present results based on raw returns. In Column 2, we standardize each performance measure by its cross-sectional standard deviation in month t. In Column 3, we use percentile ranks based on each performance measure in month t. In all three versions, we find that the *partial* coefficient associated with the CAPM alpha is reliably greater than that observed on the other performance at conventional significance levels.[11] The differences are economically large. For example, consider the comparison of the coefficient estimate on the CAPM alpha versus that on market-adjusted returns. A one percentage point increase in a fund's CAPM alpha is associated with a 0.474 percentage point increase in monthly fund flow. A one percentage point increase in a fund's market-adjusted return

[11] Results are qualitatively similar to those using category-month fixed effects, but the statistical significance of the spread between the CAPM alpha and market-adjusted return hovers around 0.10. Multicollinearity may render this test less powerful than our subsequent tests (pairwise horse race and return decomposition).

Table 3
Fund flows and competing measures of fund performance

	Scale for independent variables		
	Raw returns returns	Standardized returns	Percentile rank rank
CAPM alpha	0.474***	0.415***	1.357***
	(0.061)	(0.043)	(0.130)
Market-adjusted	0.221***	0.277***	0.852***
	(0.056)	(0.042)	(0.126)
Three-factor alpha	0.186***	0.063	0.361***
	(0.063)	(0.049)	(0.139)
Four-factor alpha	−0.072	−0.028	0.039
	(0.045)	(0.049)	(0.139)
Seven-factor alpha	0.071	0.083**	0.285***
	(0.046)	(0.035)	(0.105)
Nine-factor alpha	0.005	−0.054*	−0.177*
	(0.036)	(0.030)	(0.092)
Controls	Yes	Yes	Yes
Month fixed effects	Yes	Yes	Yes
Obs	257,053	257,053	257,053
Adj. R-squared	0.174	0.174	0.172

This table presents regression coefficient estimates from panel regressions of percentage fund flow (dependent variable) on six different measures of mutual fund performance: market-adjusted returns and alphas estimated using CAPM and three-, four-, seven-, and nine-factor models. See Table 1 and the text for details. *Raw returns* are unscaled; *Standardized returns* are scaled by the cross-sectional standard deviation of the performance measure in month t; and *Percentile ranks* are based on percentile ranks in month t. Controls include lagged fund flows from month t-19, lagged values of log of fund size, log of fund age, expense ratio, load fund dummy, return volatility, and month fixed effects. Standard errors (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

has less than half that effect (0.221 vs. 0.474), and the remaining performances have an even smaller marginal impact on flows. The same general pattern emerges when we consider standardized performance ranks (Column 2) and percentile ranks (Column 3). The marginal effect of an increase in the CAPM alpha is statistically and economically a more important determinant of fund flows than those from competing models.

One concern with these results is that we have assumed a linear relation between flows and performance. Our second approach, which relies on pairwise comparison of competing models and highly nonlinear estimation of fund-flow relations, addresses these concerns. We present the pairwise comparison of models in Table 4.[12] Recall that we compare the dummy variables that correspond to the upper and lower off-diagonals of the matrix depicted in Figure 2. To parsimoniously tabulate the results, the table presents the sum of the differences between the upper and lower off-diagonal elements and the percentage of coefficient differences that are greater than zero. Consider panel A, where we present the pairwise comparisons in which the CAPM emerges victorious. In all cases that we consider, the CAPM alpha is a better predictor of fund flows than the competing model. For example, the sum of the coefficient differences for the CAPM alpha versus market-adjusted returns is reliably

[12] See the Online Appendix for details of all forty-five comparisons for the CAPM tests. The CAPM is also victorious in a pairwise comparison with category-adjusted returns. Detailed results are reported in the Online Appendix.

Table 4
Results of pairwise model horse race

A CAPM victories

Winning model Losing model	CAPM MAR	CAPM 3-factor	CAPM 4-factor	CAPM 7-factor	CAPM 9-factor
Sum of coefficient differences	7.41***	22.94***	27.50***	28.52***	33.03***
t-stat	(3.46)	(10.55)	(13.29)	(14.68)	(17.56)
% of coefficient differences >0	77.78	100.00	100.00	100.00	100.00
Binomial p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

B Market-adjusted return (MAR) victories

Winning model Losing model	MAR 3-factor	MAR 4-factor	MAR 7-factor	MAR 9-factor
Sum of coefficient differences	15.93***	20.41***	25.11***	28.89***
t-stat	(7.39)	(9.28)	(13.37)	(15.08)
% of coefficient differences > 0	100.00	100.00	100.00	100.00
Binomial p-value	< 0.01***	< 0.01***	< 0.01***	< 0.01***

C 3-factor victories

Winning model Losing model	3-factor 4-factor	3-factor 7-factor	3-factor 9-factor
Sum of coefficient differences	19.62***	22.49***	24.79***
$t-$stat	(9.85)	(11.89)	(14.68)
% of coefficient differences >0	91.11	95.56	100.00
Binomial p-value	< 0.01***	< 0.01***	< 0.01***

(*continued*)

positive (7.41, $t=3.46$), and significantly more than half are positive (77.8% or 35 of the 45 differences). The CAPM and market-adjusted horse race is the closest contest for the CAPM. The CAPM comfortably beats the remaining models that consider returns related to size, value, momentum, investment, profitability, and industry.[13]

In Figure 3, we graph key results of four of the horse races from panel A of Table 4. For example, the top left graph shows forty-five differences in the key dummy variables that emerge when we compare the CAPM alpha and market-adjusted returns as a predictor of flows. The biggest differences in coefficient corresponds to the tallest bar, which is identified by comparing flows for funds with a CAPM decile of 9 and market-adjusted return decile rank of 1 to flows for funds with a CAPM decile of 1 and market-adjusted return decile rank of 9 (labeled "9v1" on the horizontal axis). Clearly, funds with the better performance based on the CAPM garner more flows in this comparison. The remaining forty-four bars represent the differences that we observe for all possible comparisons for funds with different decile ranks based on the two competing models. The three remaining graphs in Figure 3 present the forty-five comparisons for the horse races that depict the CAPM

[13] We also consider a horse race of the CAPM alpha versus a fund's Sharpe ratio, where the CAPM alpha again defeats the Sharpe ratio as a predictor of flows. The CAPM alpha also defeats category-adjusted fund returns (mutual fund return less the return of all sample funds in the same Morningstar-style box).

Table 4
Continued

D 4-factor victories		
Winning model	4-factor	4-factor
Losing model	7-factor	9-factor
Sum of coefficient differences	19.44***	22.72***
t-stat	(9.95)	(12.74)
% of coefficient differences > 0	93.33***	100.00***
Binomial p-value	<.01	<.01

E 7-factor victory	
Winning model	7-factor
Losing model	9-factor
Sum of coefficient differences	16.35***
t-stat	(7.97)
% of coefficient differences > 0	100.00
Binomial p-value	< 0.01***

This table presents the results of a pairwise comparison of competing asset pricing models ability to predict fund flows.
For example, we estimate the relation between flows and a fund's decile ranking based on the CAPM and three-factor models by estimating the following regression:

$$F_{pt} = a + \sum_i \sum_j b_{ij} D_{ijpt} + cX_{pt} + \mu_t + \varepsilon_{pt},$$

where the dependent variable (F_{pt}) is the fund flow for mutual fund p in month t. D_{ijpt} is a dummy variable that takes on a value of one if fund p in month t is in decile i based on the CAPM and decile j based on the three-factor model. To estimate the model, we exclude the dummy variable for j=5 and i=5. The matrix X_{pt} represents control variables, while the c contains a vector of associated coefficient estimates. As controls, we include lagged fund flows from month$t$-19, lags of a funds total expense ratio (TNA-weighted across share classes), a dummy variable for no-load funds (if all share classes are no-load funds), a funds return standard deviation estimated over the prior 12 months, the log of fund size in month t-1, and the log of fund age in month t-1. We also include time fixed effects (μ_t).
We compare the coefficients for which the decile ranks are the same magnitude, but for which the ordering is reversed. For example, we compare $b_{10,1}$ (mean flows for a top decile CAPM alpha fund and bottom decile three-factor alpha fund) to $b_{1,10}$ (mean flows for a bottom decile CAPM alpha funds and top decile three-factor alpha funds).
The table presents the results of two hypothesis tests for each horse race: (1) Ho: The sum of the differences in coefficient estimates is zero and (2) Ho: The proportion of positive differences is equal to 50%.
*, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

versus the three-factor, four-factor, or seven-factor model. The periodicity that becomes evident in these graphs emerges because the largest differences in coefficient estimates emerge when the decile ranks of the competing models are quite large (e.g., decile 9 vs. decile 1 as discussed in the example above).

Returning to Table 4, we present results of comparisons of competing models in which each panel presents results in which a particular model is victorious. Consistent with the results in panel A, we find that in results in panel B the market-adjusted returns are better able to predict fund flows than are the remaining four models we consider. (Note that using the market-adjusted return model is the equivalent of using raw returns to cross-sectionally rank funds, so we can also think of these results as comparing the responsiveness of flows to ranks based on raw returns to ranks based on the CAPM.) In panels C, D,

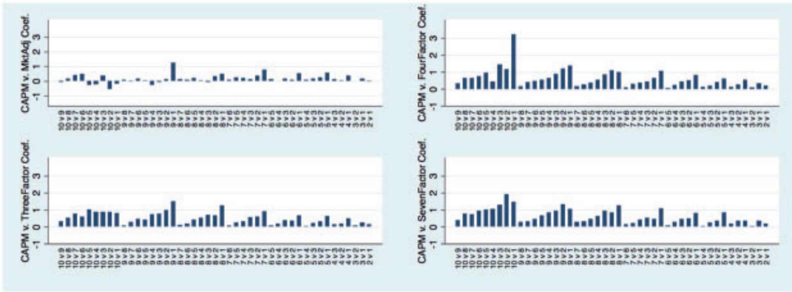

Figure 3
Flow differences for funds with different decile ranks
This figure shows the forty-five differences in coefficient estimates on dummy variables that compare funds with similar but opposite rankings based on the two models used to estimate performance. For example, the leftmost bar in each graph is the coefficient estimate on dummy variables for funds with a CAPM decile rank of ten and competing model decile rank of nine less the coefficient estimate on the dummy variable for funds with a CAPM decile rank of nine and competing model decile rank of ten. The four graphs compare the CAPM to the four competing models (market-adjusted returns, three-factor model, four-factor model, and seven-factor model).

and E, we see that the model with fewer factors consistently provides a better forecast of flows.

3.2 Do investors consider factor-related returns when evaluating fund performance?

The preceding analysis indicates that the CAPM does the best job of predicting fund-flow relations. This result implies that investors tend to consider the market risk of funds when evaluating fund performance, but tend to ignore other factor-related determinants of fund flows (i.e., size, value, momentum, or industry-related factors). We view these results as suggestive that investors in aggregate are more likely to consider market risk when evaluating a fund's performance than other factors. However, the horse race results do not imply that investors fully account for market-related factors in their fund investment decisions, and the results do not imply that investors completely ignore other factors that affect fund performance (e.g., size, value, momentum, or industry). In this section we analyze this issue in more detail. To preview our results, we generally find that investors attend most to market risk when evaluating fund performance, though fund returns related to a fund's market risk do positively affect fund flows. Put differently, investors in aggregate do not completely account for the market risk of a fund when allocating capital to mutual funds. When assessing performance, investors attend to the size, value, and industry tilts of mutual funds to a lesser extent than they do to market risk. We find no evidence that investors attend to momentum.

3.2.1 Main results and return decomposition. We regress fund flows on alphas and factor-related returns during the prior eighteen months using the seven-factor model to estimate the factor-related returns.[14] These results are presented in Table 5. To address issues of residual cross-sectional dependence within a month (a time effect) or residual serial dependence for a fund over time (a fund effect), we double-cluster standard errors by month and fund.[15]

In Column 1, we present results for all funds using our main specification. In this main specification, we include standard control variables and month fixed effects, but exclude Morningstar Star ratings. In Columns 2 and 3, we increase the granularity of the fixed effect control from month (Column 1) to month-category (Column 2) to month-category-star (Column 3). The remaining six columns present subsample results (Columns 4 through 9).

Consider first the results for our main specification in Column 1. Fund flows respond positively to the seven-factor alpha with an estimated sensitivity of 0.884, which is highly significant at conventional levels. The parameter estimate suggests that an 87-bps increase in alpha (roughly the interquartile range of estimated alphas observed in Table 1, panel C) is associated with an increase in fund flows of 0.77 percentage points. The sensitivity of flows to returns traced to market, size, value, momentum, and industry factor returns is reliably positive. These results suggest that, in aggregate, investors respond to fund returns that can be traced to a fund's investment style and do not fully discount returns that might be traced to these factors when assessing fund performance.

The magnitudes of the returns traced to factor loadings relative to the fund's alpha are of more interest. In the main specification of Column 1, fund returns related to a fund's market, size, and value factors do not generate the same flows as a fund's alpha. For example, the coefficient on the returns related to a fund's market risk (0.253) is 29% of the alpha coefficient (0.884), while the size and value coefficients are 86% and 75% (respectively) of those associated with a fund's alpha. In contrast, the coefficient on the momentum-related returns is not reliably different from those associated with a fund's alpha. We do find some evidence that investors consider the industry tilts of mutual funds in the second and third industry factors, which yield coefficient estimates that are 79% and 80% of those associated with a fund's alpha (respectively). However, the coefficient on returns associated with the first industry factor is not reliably different from those associated with a fund's alpha. When we formally test the null hypothesis that the coefficients on the returns traced to factor tilts differ from that for the fund's alpha, we can reject the null hypothesis of equality for the market, size, value, and two of the three industry coefficients. Thus, in aggregate, investors seem to tend most to the market risk (i.e., beta) of a fund

---

[14] In prior drafts of this paper, we considered past performance based on horizons ranging from one month to three years and a return decomposition based on the four-factor model. The results of this analysis are qualitatively similar to those presented here.

[15] In this and all subsequent analyses, we present results excluding outliers (defined as observations with a Cook's D statistic greater than $4/n$ in the full sample analysis in which n is the number of observations used to estimate the regression). The coefficient estimates including influential observations are qualitatively similar to those presented, though less precisely estimated.

Table 5
Return decomposition results: response of fund flows to components of fund returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fund sample:	All funds	All funds	All funds	Small funds	Big funds	Young funds	Old funds	Below-median ret	Above-median ret.
ALPHA	0.884***	0.789***	0.738***	0.843***	0.899***	0.918***	0.868***	0.701***	0.891***
	(0.027)	(0.030)	(0.027)	(0.033)	(0.030)	(0.036)	(0.028)	(0.036)	(0.036)
MKTRET	0.253***	0.207***	0.194***	0.236***	0.255***	0.252***	0.253***	0.163***	0.244***
	(0.056)	(0.050)	(0.041)	(0.056)	(0.055)	(0.057)	(0.055)	(0.058)	(0.056)
SIZRET	0.759***	0.685***	0.639***	0.529***	0.885***	0.725***	0.775***	0.714***	0.591***
	(0.054)	(0.053)	(0.048)	(0.061)	(0.062)	(0.062)	(0.062)	(0.063)	(0.075)
VALRET	0.665***	0.590***	0.568***	0.698***	0.647***	0.697***	0.653***	0.523***	0.684***
	(0.063)	(0.057)	(0.053)	(0.066)	(0.066)	(0.070)	(0.064)	(0.073)	(0.071)
MOMRET	1.059***	0.940***	0.851***	0.933***	1.106***	1.202***	0.996***	0.906***	0.999***
	(0.060)	(0.062)	(0.050)	(0.067)	(0.071)	(0.080)	(0.060)	(0.076)	(0.075)
INDRET1	0.920***	0.820***	0.838***	0.914***	0.915***	0.937***	0.918***	0.705***	0.945***
	(0.074)	(0.073)	(0.075)	(0.083)	(0.084)	(0.101)	(0.077)	(0.093)	(0.098)
INDRET2	0.701***	0.593***	0.630***	0.681***	0.714***	0.691***	0.701***	0.539***	0.741***
	(0.095)	(0.084)	(0.089)	(0.115)	(0.108)	(0.124)	(0.103)	(0.103)	(0.129)
INDRET3	0.692***	0.642***	0.479***	0.739***	0.664***	0.726***	0.684***	0.431***	0.776***
	(0.087)	(0.082)	(0.081)	(0.112)	(0.091)	(0.110)	(0.096)	(0.104)	(0.102)
Month fixed effects	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-cat. fixed effects	No	Yes	No	No	No	No	No	No	No
Month-cat.-rating FEs	No	No	Yes	No	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	257,053	257,053	248,463	257,053		257,053		257,053	
Adj. R-squared	0.173	0.190	0.216	0.175		0.173		0.175	

This table presents regressions coefficient estimates from panel regressions of percentage fund flow (dependent variable) on the components of a fund's return—a fund's alpha and seven factor-related returns. The seven factor-related returns are estimated based on the fund's factor exposure (e.g., tilt toward small versus large stocks) and the factor return (e.g., performance of small versus large stocks). The seven factors include the market (i.e., a fund's beta times the excess return on the market index), size, value, momentum, and three industry portfolios that capture the industry tilts of a mutual fund. Controls include lagged fund flows from month t-19, lagged values of log of fund size, log of fund age, expense ratio, load fund dummy, and return volatility. Standard errors (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

when assessing fund performance. Investors in aggregate do some accounting for the size, value, and industry tilts of a fund, but the responsiveness of flows to these return components is much stronger than that observed for returns related to a fund's market risk.

In Column 2, we replace month fixed effects with month-category fixed effects, which absorb variation in average fund flows across the nine Morningstar-style boxes. The results are qualitatively similar to our main specification. Flows respond less to returns related to a fund's market risk than to alpha or other factor-related returns.

3.2.2 Morningstar fund ratings. Each month, Morningstar issues mutual fund ratings based on a fund's risk and return relative to its peer group over three-, five-, and ten-year horizons. Morningstar ranks funds within fund categories based on a risk-adjusted return, where the risk-adjustment is a modified measure of standard deviation that emphasizes downward variation. Ratings range from one star for low-performing funds to five stars for high-performing funds. The distribution of funds within stars is one (10%), two (22.5%), three (35%), four (22.5%), and five stars (10%). Moreover, Morningstar fund ratings have a causal impact on fund flows (Del Guercio and Tkac 2008). Given that Morningstar penalizes funds for volatility and star ratings influence fund flows, it is plausible that investors account for market risk and, to a lesser extent, other factor-related returns by following Morningstar fund ratings when allocating capital to mutual funds.

To investigate whether star ratings are a potential mechanism by which investors tend to factor-related returns (particularly returns traced to market risk), we replace the month fixed effects of our main specification with month-category-star fixed effects. The star component of the fixed effect is based on five categories, which we construct based on a fund's star rating. First, we calculate the TNA-weighted overall star rating across share classes for a fund.[16] (Generally, little variation is seen in star ratings across share classes.) We then create five categories of star ratings based on the following intervals: (1.0,1.5), [1.5,2.5), [2.5,3.5), [3.5,4.5), [4.5,5.0).

The results of this analysis are presented in Column 3, Table 5. Because star ratings are highly correlated with fund performance, the month-category-star fixed effects reduce the coefficient estimates relative to those in our main specification of Column 1. However, the relative importance of factor-related returns and alpha in explaining flows is once again qualitatively similar to that observed in our main specification. Thus, star ratings do not appear to explain the result that investors attend most to a fund's market risk when assessing performance and pay much less, but some, attention to the fund's size, value, and industry characteristics.

---

[16] Morningstar's overall star rating is a weighted average of the three-, five-, and ten-year star ratings for a fund with more weight given to the three-year rating.

3.2.3 Fund size and age. To further test the robustness of our findings, we partition our sample into small versus large funds and young versus old funds. To partition on fund size, in December of each year we split funds on the sample median of total net assets for funds and define below-median funds as small funds and above-median funds as large funds for the following year. Since young funds tend to have shorter track records, we anticipate that the sensitivity of flows to returns would be greater for young funds. However, given that we require a minimum of five years of performance data for funds, our sample omits the youngest funds when these effects are most dramatic (Chevalier and Ellison 1997). As a result, we define young (old) funds as those with less (more) than ten years of return history. We estimate subsample results by interacting a fund size (or fund age) dummy variable with the return components.

We present the results based on the fund size partition in Table 5, Columns 4 and 5.[17] In general, the results are quite similar between small and large funds with one exception: the importance of a fund's size-related returns. Among small funds, the coefficient estimate on returns related to a fund's size tilt is only 63% of that associated with the fund alpha (i.e., 0.529/0.843). In contrast, for large funds, the responsiveness of flows to size-related returns (0.885) is very similar to the flow response to a fund's alpha. We present the results based on the fund age partition in Table 5, Columns 6 and 7. We find very little difference in the flow response to return components between young and old funds.

3.2.4 Nonlinearities in the flow-return relationship. Our main regression imposes a linear relationship between fund flows and returns. Since prior research suggests that the relationship is convex (e.g., Chevalier and Ellison 1997), we test the robustness of our results by interacting a dummy variable that takes on a value of one for funds that are above the median fund return in a particular month with the fund return components. We summarize the results of this regression in the last two columns of Table 5. Consistent with prior work that documents a convex relation between flows and returns, we find that seven of the eight return components generate higher coefficient estimates in the flow regressions for funds that have above-median returns. However, we continue to find that returns related to a fund's market risk do not generate the same flow response as other return components.

We summarize the main message of these analyses in Figure 4, which presents eight graphs that each correspond to one of the eight return components that we analyze. Consider the graph that summarizes the responsiveness of flows to market-related returns (top left graph of Figure 4). Each bar in the graph represents the estimated coefficient on the fund's market-related return, divided by the coefficient on the fund's alpha, and the nine bars correspond

[17] Results with month-category fixed effects yield results qualitatively similar to those in Columns 4 to 9; see the Online Appendix for details.

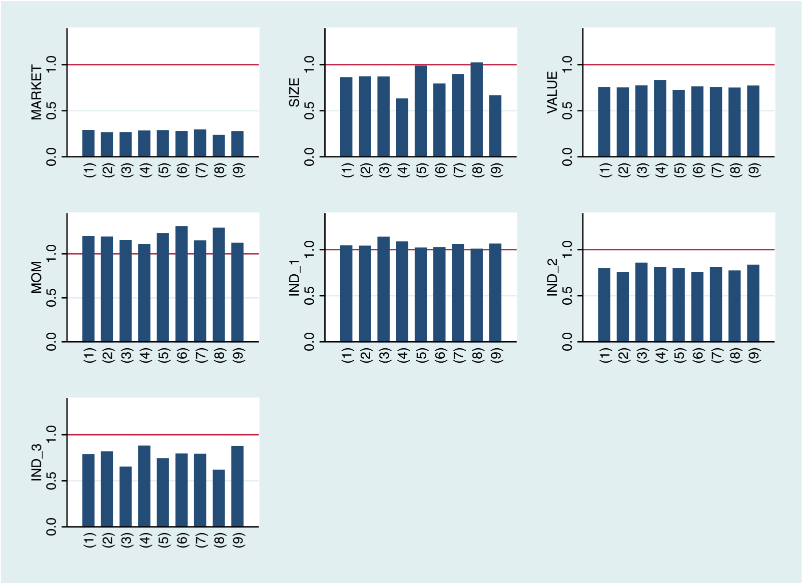

Figure 4
Relative importance of return components
Monthly fund flows are regressed on eight return components (alpha and seven factor-related returns). Each graph displays the ratio of the coefficient estimate for a return component to the estimated alpha, where the red line indicates the flows respond equally to the return component and alpha. Within each graph, the bars correspond to different models (see the text for details):

1. Main results: all funds, basic controls, month fixed effects
2. All funds, basic controls, and month-category fixed effects
3. All funds, basic controls, and month-category-star rating fixed effects
4. Small funds
5. Big funds
6. Young funds
7. Old funds
8. Below-median returns
9. Above-median returns

to the nine different sets of results that we present in Table 5. In this graph, regardless of the specification or subsample considered, we find that flows are much less responsive to a fund's market-related returns than to the fund's alpha. Scanning the remaining six graphs in the figure also reveals the main message of these analyses: fund flows have a very muted response to market-related returns of mutual funds, but tend to have a much stronger response to other return components. We find robust evidence that investors have a mildly muted response to value-related returns and somewhat weaker evidence of a muted response for size-related and industry-related returns (in that two of the three industry components yield ratios less than one). Fund flows respond to momentum-related returns as much (if not more) than to a fund's alpha.

3.3 Estimation error in factor loadings

One possible explanation for the muted response to market-related returns relative to other factor-related returns could be that investors believe that market betas are estimated more precisely than other factor loadings and that these other factor loadings tend to be not far from zero. With these beliefs, investors would rationally shrink market betas a bit toward one (the global mean) to account for estimation error, but would more aggressively shrink other estimated loadings toward their expected mean of zero because of the relatively large estimation error. As a result, when we regress flows on decomposed returns, we would obtain smaller coefficient estimates on the market-related returns versus the estimated coefficients on other factor-related returns even if investors respond equally to all factor-related returns.

If this explanation were true, we would expect to observe less estimation error in our estimates of market betas relative to other factor-related loadings. To see if this is, in fact, the case, we examine the persistence of in-sample versus out-of-sample factor loadings. Consider estimates of market betas for mutual funds. In month t, we sort all funds into quintiles based on the estimated beta for the five years ending in month t. To compute the in-sample beta estimates for each beta quintile for month t, we calculate the cross-sectional mean of the estimated beta within that quintile; we then calculate the average and standard error of the cross-sectional means across months. To compute the out-of-sample beta estimates, we first construct a time series of monthly fund returns during month $t+1$ for each beta quintile. We then estimate the out-of-sample market beta for each quintile using the out-of-sample return series. We repeat this analysis for each of the factor loadings.

The results of this analysis are presented in Table 6. In panel A, we present the in-sample and out-of-sample beta estimates for fund quintiles formed on the basis of in-sample betas. The in-sample betas range from a low of 0.677 for quintile 1 to a high of 1.154 for quintile 5, yielding a spread (Hi-Lo) of 0.477. As expected, the out-of-sample estimates tend to shrink toward the global beta for funds, which is slightly less than one, and the spread (Hi-Lo) thus declines to 0.254. The "Shrinkage ratio" for the beta estimates is calculated as the out-of-sample to in-sample Hi-Lo spread of the beta estimates: $0.254/0.477 = 53.2\%$. Panels B through G present similar results for each of the other estimated factor loadings. For each factor that we analyze, the rank ordering of the loadings across quintile portfolios is preserved out-of-sample, which indicates the estimated in-sample loadings are indeed informative (i.e., not pure noise). However, as expected, all out-of-sample parameter estimates shrink toward their global means.

Most importantly, this estimation error does not provide a likely explanation of the relative significance of the return components. The factor loadings on size, value, and the first industry factor are the most persistent, suggesting these factor loadings are more precisely estimated. The remaining factors (beta, momentum, and the other industry factors) tend to shrink more toward the

Table 6
In-sample versus out-of-sample factor loadings

	Quintile portfolios						
	1 (Lo)	2	3	4	5 (Hi)	Hi-lo	Shrinkage ratio
A. Beta estimates, fund quintiles based on market beta							
In-sample	0.677***	0.874***	0.951***	1.015***	1.154***	0.477	
	(0.0053)	(0.0036)	(0.0025)	(0.0024)	(0.0050)		
Out-of-sample	0.762***	0.915***	0.961***	0.986***	1.016***	0.254	53.2%
	(0.0202)	(0.0139)	(0.0106)	(0.0135)	(0.0266)		
B. Size coefficient, fund quintiles based on size coefficient							
In-sample	−0.175***	−0.020***	0.135***	0.337***	0.677***	0.852	
	(0.0015)	(0.0018)	(0.0028)	(0.0034)	(0.0063)		
Out-of-sample	−0.122***	−0.026	0.118***	0.273***	0.525***	0.647	75.9%
	(0.0122)	(0.0166)	(0.0227)	(0.0220)	(0.0240)		
C. Value coefficients, fund quintiles based on value coefficient							
In-sample	−0.450***	−0.117***	0.042***	0.207***	0.474***	0.924	
	(0.0064)	(0.0040)	(0.0038)	(0.0062)	(0.0078)		
Out-of-sample	−0.265***	−0.040*	0.089***	0.242***	0.451***	0.716	77.5%
	(0.0292)	(0.0211)	(0.0193)	(0.0222)	(0.0257)		
D. Momentum coefficients, fund quintiles based on momentum coefficient							
In-sample	−0.159***	−0.053***	0.007***	0.075***	0.212***	0.371	
	(0.0042)	(0.0025)	(0.0022)	(0.0032)	(0.0059)		
Out-of-sample	−0.080***	−0.032***	−0.014	0.014	0.071***	0.152	40.8%
	(0.0140)	(0.0103)	(0.0099)	(0.0118)	(0.0152)		
E. Industry coefficient 1, fund quintiles based on industry 1 coefficient							
In-sample	−0.050***	−0.006**	0.019***	0.056***	0.186***	0.236	
	(0.0010)	(0.0008)	(0.0010)	(0.0016)	(0.0034)		
Out-of-sample	−0.012**	0.004	0.020***	0.051***	0.167***	0.179	75.9%
	(0.0060)	(0.0045)	(0.0059)	(0.0082)	(0.0169)		
F. Industry coefficient 2, fund quintiles based on industry 2 coefficient							
In-sample	−0.110***	−0.030***	0.004***	0.041***	0.116***	0.226	
	(0.0023)	(0.0014)	(0.0013)	(0.0016)	(0.0025)		
Out-of-sample	−0.050***	−0.023***	−0.004	0.004	0.013	0.063	27.9%
	(0.0113)	(0.0083)	(0.0084)	(0.0113)	(0.0170)		
G. Industry coefficient 3, fund quintiles based on industry 3 coefficient							
In-sample	−0.125***	−0.039***	−0.002	0.033***	0.124***	0.249	
	(0.0031)	(0.0017)	(0.0014)	(0.0016)	(0.0020)		
Out-of-sample	−0.025	−0.009	−0.002	0.008	0.044***	0.070	27.9%
	(0.0228)	(0.013)	(0.0098)	(0.0111)	(0.0157)		

The table presents in-sample versus out-of-sample factor loadings for quintiles based on in-sample factor loadings. In month t, we sort all funds into quintiles based on the estimated beta for the five years ending in month t. For each beta quintile, we construct two time series of returns: in-sample and out-of-sample. The in-sample time series is based on monthly fund returns during the five-year estimation period (t-59 to t); the out-of-sample time series is based on the monthly fund returns during month t+1. We repeat this analysis for every month during our sample period. Armed with the in-sample and out-of-sample time-series of monthly fund returns for each beta quintile, we estimate the market beta for each quintile using the in-sample and out-of-sample return series yielding a total of ten beta estimates. We repeat this analysis for each of the factor loadings. The "Shrinkage ratio" is the ratio of the out-of-sample spread in extreme quintiles (Hi-Lo) to the in-sample spread. Standard errors are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

global mean. Thus, if investors care equally about all factor-related returns, but adjust flow response for greater estimation error in some factor loadings, we would observe smaller coefficient estimates in our flow-return regressions for the returns associated with size, value, and the first industry factor. We do not observe this. In fact, investors tend to respond less to returns related to a fund's market beta *despite* that the beta estimate is less precise.

4. Investor Sophistication and Fund-Flow Relations

Our primary analysis treats mutual fund investors as a homogenous group. However, different investors almost certainly use different methods to assess the performance of mutual funds. In this section, we test and find strong support for the conjecture that more sophisticated investors use more sophisticated benchmarks to evaluate mutual fund performance. We do so in three ways. First, we use direct-sold versus broker-sold distribution channels as a proxy for investor sophistication. Second, we compare the flow-return dynamics during periods of high sentiment, when less sophisticated investors arguably represent a higher proportion of fund investors, to periods of low sentiment. Third, we use a separate dataset of fund purchases and sales at a discount broker fund marketplace to compare the flow-return relations for wealthy and other investors.

4.1 Distribution channels

Chalmers and Reuter (2013) report that investors who purchase mutual funds through a broker tend to be younger, less well educated, and less wealthy than investors who buy funds directly sold from fund companies and that investors in broker-sold funds underperform investors in direct-sold funds. Christoffersen, Evans, and Musto (2013) report that flows to broker-sold funds are heavily influenced by payments made by fund companies to brokers. If investors in direct-sold funds are more knowledgeable than those in broker-sold funds, they likely would have more sophisticated models for benchmarking mutual fund performance. Consistent with this idea, Del Guercio and Reuter (2013) find that flows are more sensitive to alpha for direct-sold funds than for broker-sold funds, while broker-sold funds respond more to market-adjusted returns. Thus, we anticipate that investors in the direct-sold channel will respond less to factor-related returns than will investors in the broker-sold channel.

To test this conjecture, we analyze the impact of a fund's distribution channel on the flow-return relations. To do so, we first identify the primary distribution channel for each fund. As in Sun (2014), we classify a fund as broker-sold if 75% of its assets are held in a share class that meets any of the following three criteria: the fund charges a front-end load, a back-end load, or a 12b-1 fee greater than 25 bps. Bergstresser, Chalmers, and Tufano (2009) document that broker-sold funds tend to charge front-end loads, back-end loads, or 12b-1 fees as a means to provide compensation to brokers who sell funds to investors. Conversely, a fund is direct-sold if 75% of its assets are held in a share class that charges no front-end load, no back-end load, and no 12b-1 fee. In the average month during our sample period, 40% of funds are direct-sold, 53% are broker-sold, and the remaining 7% have an indeterminate distribution channel.

To test the hypothesis that flow-return relations differ across distribution channels, we modify the main return decomposition regression of Equation (10) by interacting each of the return components of a fund with a dummy variable that takes a value of one if the fund is primarily broker-sold.

We summarize the results of this single interaction regression for the full sample and main regression specification in the first three columns of Table 7. Column 1 presents the coefficient estimates for the direct-sold channel. Column 2 presents the corresponding estimates for the broker-sold channel (i.e., the sum of the coefficient estimate on the return component for the direct-sold channel and coefficient on the interaction of the return component with the broker-sold dummy). Column 3 presents the difference between the direct-sold and broker-sold channel (i.e., the estimated interaction terms). With the exception of momentum, we consistently find that investors in the broker-sold channel respond more to factor-related returns than do investors in the direct-sold channel. These results are consistent with the notion that investors in the broker-sold channel are less sophisticated in their assessment of fund performance than are investors in the direct-sold channel.[18]

These results provide strong support for the notion that more sophisticated investors use more sophisticated models to assess fund manager skill. Nonetheless, the *relative* importance of the various factors is generally similar for the two distribution channels. Perhaps most strikingly, the coefficient estimates on a fund's market-related return are smaller than other factor-related returns for both the direct-sold and broker-sold channels.

4.2 Periods of high versus low sentiment trading

Chiu and Kini (2014) argue that aggregate equity fund flows proxy for noise trader sentiment and document that firms time their equity issuance decisions to coincide with periods of positive sentiment.[19] Our second test builds on this observation to identify periods of extreme sentiment trading using a measure analogous to that in Chiu and Kini (2014). Specifically, for each month in our sample period, we create a sentiment measure ($SENT_t$) that captures temporal variation in aggregate flows in mutual funds:

$$SENT_t = \frac{\sum_{i=1}^{n} F_{it}}{\sum_{i=1}^{n} TNA_{i,t-1}}, \tag{11}$$

where the numerator sums the value of fund flows (F_{it}) across n funds and scales by the sum of lagged TNAs. Because our period is characterized by

[18] Investors in the direct-sold channel also respond less to a fund's alpha. When we calculate the ratio of the factor-related coefficient and the alpha coefficient (as in Figure 4), for all factor-related returns, except for momentum, we consistently find the ratio is less for direct-sold subsample than for the broker-sold subsample.

[19] Akbas et al. (2015), Ben-Rephael et al. (2012), and Brown et al. (2003) also discuss the use of mutual fund flows as a measure of investor sentiment.

Table 7
Investor sophistication and flow-return relations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	CRSP fund flows: Direct vs. broker sold funds			CRSP Fund flows: Periods of high vs. low sentiment			Broker data: Wealthy vs. other investors		
	Direct	Broker	Diff (brok.-dir.)	Low sentiment	High sentiment	Diff (high-low)	Wealthy	Others	Diff (oth.-wlthy)
ALPHA	0.816***	0.914***	0.0972***	0.717***	0.977***	0.260***	3.554***	4.127***	0.573**
	(0.032)	(0.031)	(0.032)	(0.040)	(0.034)	(0.049)	(0.676)	(0.649)	(0.269)
MKTRET	0.225***	0.291***	0.0662***	0.248***	0.200**	−0.0475	1.273	2.211	0.937**
	(0.051)	(0.055)	(0.013)	(0.064)	(0.079)	(0.100)	(1.619)	(1.821)	(0.462)
SIZRET	0.616***	0.839***	0.223***	0.483***	0.866***	0.383***	1.632	4.094***	2.462**
	(0.065)	(0.061)	(0.061)	(0.085)	(0.065)	(0.102)	(1.264)	(0.860)	(0.972)
VALRET	0.527***	0.750***	0.223***	0.380***	0.839***	0.459***	3.906***	4.873***	0.967
	(0.067)	(0.064)	(0.044)	(0.084)	(0.067)	(0.106)	(0.801)	(1.445)	(1.472)
MOMRET	1.095***	1.019***	−0.0762	0.868***	1.094***	0.226*	5.648**	5.872***	0.224
	(0.067)	(0.067)	(0.070)	(0.115)	(0.064)	(0.128)	(2.405)	(2.277)	(1.870)
INDRET1	0.770***	1.004***	0.234**	0.746***	0.986***	0.240*	2.132*	2.73	0.597
	(0.090)	(0.086)	(0.100)	(0.100)	(0.094)	(0.131)	(1.089)	(1.694)	(0.933)
INDRET2	0.581***	0.876***	0.295**	0.376***	0.896***	0.520***	2.691	0.223	−2.469
	(0.118)	(0.110)	(0.128)	(0.124)	(0.111)	(0.153)	(1.851)	(1.848)	(1.728)
INDRET3	0.561***	0.693***	0.132	0.639***	0.736***	0.0972	3.706***	7.196***	3.490***
	(0.095)	(0.096)	(0.094)	(0.139)	(0.093)	(0.162)	(1.399)	(1.077)	(0.741)
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations		234,690			256,292			24,629	
R-Squared		0.181			0.162			0.061	

This table presents regressions coefficient estimates from three interactive panel regressions of fund flows (dependent variable) on the components of a fund's return and an interaction dummy variable that proxies for investor sophistication. We consider three proxies: funds sold through broker-sold distribution channels (the model of Columns 1–3), periods of high investor sentiment as measured by (the quartile of) months with the greatest aggregate mutual fund flows (the model of Columns 4–6), and less wealthy investors (i.e., those with below median account sizes) at a large discount brokerage firm (the model of Columns 7–9). When we use broker data, fund flows are measured using trades and positions made at the broker over the period 1991 to 1996 (see the text for details). Controls include lagged fund flows from month t-19, lagged values of log of fund size, log of fund age, expense ratio, load fund dummy, and return volatility. Standard errors (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

aggregate inflows into mutual funds, this measure is positive in virtually all fund months during our sample period. As a result, we focus on periods of extreme positive sentiment, which we define as months in which the sentiment measure is in the top quartile during our sample period.[20]

We conjecture that more sophisticated investors will use more sophisticated benchmarks, and thus we would expect to observe flows relatively less responsive to factor-related returns during low-sentiment periods. We present the results of this analysis in Columns 4 to 6 of Table 7. Consistent with our conjecture, we find that flows are more responsive to factor-related returns during high-sentiment periods. These results support our conjecture that more sophisticated investors use more sophisticated benchmarks.

4.3 Wealthy versus other investor trades at broker

In our third analysis, we identify mutual fund purchases and sales of wealthy versus other investors using trades and position data for 78,000 households who have accounts with a large discount broker (LDB) over the period 1991 to 1996 (see Barber and Odean 2000 for details). In this analysis, we conjecture that wealthy investors generally will be more sophisticated than others, consistent with the evidence on trading ability (Barber and Odean 2000; Geng et al. 2014), diversification (Calvet, Campbell, and Sodini 2007), and the disposition effect (Dhar and Zhu 2006). We define a wealthy investor as a household with a total average account size (including stock, bond, cash, and mutual fund investments) above the median account size in the broker sample.[21]

The broker offers a mutual fund marketplace for buying and selling mutual funds. We use the trades in mutual funds to construct a measure of flows by summing the value buys (B) less the value of sells (S) of fund i across n wealthy households ($j=1,n$) in month t, which we scale by the positions (P) of fund i summed across these households:

$$F_{it}^{wealthy} = \frac{\sum_{j=1}^{n}\left(B_{ijt} - S_{ijt}\right)}{\sum_{j=1}^{n}\left(P_{ij,t-1}\right)}. \tag{12}$$

We make an analogous calculation for mutual fund trades executed by the less wealthy households. The results of this analysis are presented in Columns 7 to 9 of Table 7. Notably, while using this limited sample period and dataset, which at times yields imprecisely estimated flow-return relations, we again

[20] In earlier drafts, we also analyze a sentiment measure that uses the absolute value of flows in the numerator, $|F_{it}|$, and thus identifies periods with high levels of mutual fund trading (including switching across funds). Results are qualitatively similar using this alternative measure of sentiment.

[21] Our results are qualitatively similar when we split the LDB sample based on the top quartile versus below median account size. We also have self-reported income for a subset of the LDB sample and find qualitatively similar results when we split the sample on income.

generally find that flows are less responsive to returns related to a fund's market risk for *both* the wealthy and less wealthy households. Moreover, consistent with our conjecture that more sophisticated investors use more sophisticated benchmarks, we consistently find that flows that emanate from wealthy investors are less responsive to factor-related returns.

The tests in Table 7 focus on the difference in flow-return coefficients interacted with proxies for sophistication (broker distribution channel, periods of high sentiment, or wealth). In each of the three analyses the less sophisticated investors respond more to a fund's alpha, consistent with the evidence in Bailey, Kumar, and Ng (2011) that less sophisticated investors chase fund performance. When we calculate the ratio of the factor-related coefficient and the alpha coefficient (as in Figure 4), our results are qualitatively similar to the conclusions based on the unscaled differences in coefficient estimates (presented in Table 7). The Online Appendix also provides qualitatively similar results when we use category-month (rather than month) fixed effects. In sum, all three proxies for investor sophistication are consistent with the conjecture that more sophisticated investors use more sophisticated benchmarks.

5. Fund Categories and Flows

Our primary results indicate investors, in aggregate, place more weight on the CAPM than on other models when ranking mutual funds. Moreover, they partially adjust for returns related to fund's size and value tilts. We hypothesize that the muted response to size and value factors results from some investors using Morningstar-style categories when picking funds (e.g., treating all small cap funds as similar despite having different exposures to small cap stocks). If investors use Morningstar category boxes to assess mutual fund performance, then we would observe a muted response to the fund returns that can be traced to a fund's value or size tilts since Morningstar categories capture some of the variation in size and value tilts (see Table 2). These predictions dovetail neatly with our main results, where we indeed observe a muted response to returns that can be traced to a fund's size and value tilts (see Figure 4 and Table 5).

A secondary prediction, which we test in this section, is that investors will be less responsive to returns that can be traced to a fund's category characteristics (e.g., all Morningstar small value funds will likely have some tilt toward small value stocks) than to returns that can be traced to a fund's deviation from mean category characteristics (i.e., the relative tilt toward small/value cap for a fund identified as a Morningstar small value fund).

To test this second prediction, we decompose the size (and value) factor exposure of a fund into the average exposure of the Morningstar category to which it belongs and the fund's deviation from the mean category exposure. For example, the mean size category exposure for a small value fund is the mean size-related return across all funds categorized by Morningstar as small value funds. In general, we calculate the mean category return for the size factor as

$$CATSIZ_{ct} = \frac{1}{N_c} \sum_{p=1}^{N_c} SIZRET_{pt}, \tag{13}$$

where $SIZRET_{pt}$ is the size-related return for fund p and N_c is the number of sample funds in category c, where we consider the nine Morningstar categories. An analogous calculation is used for a fund's value exposure.

This return decomposition yields an augmented version of the regression from Equation (10), where the single independent variable for size-related returns (*SIZRET*) is now replaced with two independent variables associated the with size tilt of the fund's category (*CATSIZ*) and the deviation of the fund's size tilt from the category average (*FUNDSIZ* = *SIZRET*-*CATSIZ*). Similarly, the single independent variable for value-related returns (*VALRET*) is replaced with two independent variables that capture the fund's value category (*CATVAL*) and deviation from category (*FUNDVAL* = *VALRET*-*CATVAL*). If investors benchmark returns at the category level, then we should observe coefficients of zero on the *CATSIZ* and *CATVAL* variables; investors should not respond to returns that can be traced to the category-level exposure to size or value factors. However, if some investors treat category-level returns as alpha, we would expect to observe positive coefficients on these category-level coefficients. Note also that if investors do not distinguish between a fund's category-level size exposure and its fund-level size exposure, then we would observe equal coefficient estimates on *CATSIZ* and *FUNDSIZ* (or *CATVAL* and *FUNDVAL*). Thus, this framework also allows us to test whether investors treat the source of a fund's factor exposure (category assignment vs. deviation from category averages) equally.

We present the results of this analysis in Table 8. Consider first the results based on the decomposition of the size exposure. The coefficient on the mean category exposure of a fund (*CATSIZ*) is reliably positive, which indicates fund flows indeed respond to the category-level exposure of a fund as in Teo and Woo (2004). However, the response of flows to the fund's size category exposure is less than that associated with the fund's deviation from this category average (*FUNDSIZ* vs. *CATSIZ* coefficients, $0.849 > 0.681$, $p < 0.05$). The results are quite similar for the decomposition of a fund's value exposure, where the response of flows to the fund's value category exposure is less than that associated with the fund's deviation from this category average (*FUNDVAL* vs. *CATVAL* coefficients, 0.736 v. 0.542, $p < 0.01$). These results are quite consistent across the alternative specifications (Columns 2 and 3) and subsamples (Columns 4 to 9) that we consider. Taken together, these results suggest that some investors use a fund's category assignment to benchmark returns; these investors do not respond to style returns attributable to a fund's category, but do respond to style returns that deviate from the category mean. Other investors likely treat style returns as alpha, regardless of whether these

Table 8
Morningstar categories and fund flows

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Fund sample:	All funds	All funds	Small funds	Big funds	Young funds	Old funds	Below-median returns	Above-median returns
ALPHA	0.893***	0.743***	0.840***	0.910***	0.919***	0.881***	0.732***	0.915***
	(0.028)	(0.027)	(0.033)	(0.030)	(0.036)	(0.029)	(0.039)	(0.037)
MKTRET	0.253***	0.191***	0.233***	0.261***	0.240***	0.256***	0.213***	0.268***
	(0.056)	(0.042)	(0.056)	(0.055)	(0.057)	(0.055)	(0.059)	(0.055)
CATSIZ	**0.681*** **	**0.574*** **	**0.477*** **	**0.799*** **	**0.743*** **	**0.646*** **	**0.667*** **	**0.539*** **
	(0.069)	**(0.062)**	**(0.073)**	**(0.078)**	**(0.077)**	**(0.076)**	**(0.086)**	**(0.085)**
FUNDSIZ	**0.849*** **	**0.712*** **	**0.699*** **	**0.928*** **	**0.753*** **	**0.908*** **	**0.675*** **	**0.830*** **
	(0.070)	**(0.065)**	**(0.091)**	**(0.080)**	**(0.101)**	**(0.078)**	**(0.082)**	**(0.097)**
CATVAL	**0.542*** **	**0.529*** **	**0.613*** **	**0.519*** **	**0.551*** **	**0.546*** **	**0.413*** **	**0.541*** **
	(0.079)	**(0.068)**	**(0.078)**	**(0.085)**	**(0.091)**	**(0.080)**	**(0.081)**	**(0.094)**
FUNDVAL	**0.736*** **	**0.591*** **	**0.724*** **	**0.737*** **	**0.724*** **	**0.732*** **	**0.639*** **	**0.790*** **
	(0.062)	**(0.050)**	**(0.075)**	**(0.063)**	**(0.070)**	**(0.067)**	**(0.079)**	**(0.072)**
MOMRET	1.077***	0.858***	0.926***	1.126***	1.205***	1.008***	0.955***	1.062***
	(0.060)	(0.050)	(0.067)	(0.071)	(0.080)	(0.060)	(0.076)	(0.073)
INDRET1	0.923***	0.839***	0.930***	0.903***	0.972***	0.912***	0.726***	0.909***
	(0.074)	(0.074)	(0.083)	(0.084)	(0.101)	(0.076)	(0.093)	(0.097)
INDRET2	0.732***	0.642***	0.700***	0.740***	0.730***	0.727***	0.549***	0.893***
	(0.095)	(0.089)	(0.115)	(0.107)	(0.125)	(0.102)	(0.104)	(0.128)
INDRET3	0.720***	0.476***	0.720***	0.713***	0.710***	0.736***	0.604***	0.761***
	(0.085)	(0.079)	(0.114)	(0.087)	(0.112)	(0.094)	(0.104)	(0.095)
CATSIZ – FUNDSIZ	**−0.168****	**−0.138****	**−0.222*** **	**−0.129**	**−0.01**	**−0.261*** **	**−0.008**	**−0.291*** **
	(0.088)	**(0.083)**	**(0.104)**	**(0.101)**	**(0.122)**	**(0.094)**	**(0.112)**	**(0.109)**
CATVAL – FUNDVAL	**−0.194*** **	**−0.062**	**−0.111**	**−0.218*** **	**−0.173*** **	**−0.186*** **	**−0.226*** **	**−0.249*** **
	(0.064)	**(0.055)**	**(0.078)**	**(0.074)**	**(0.081)**	**(0.073)**	**(0.068)**	**(0.090)**
Month fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Star ratings	No	Yes	No	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	257,053	248,463	257,053		257,053		257,053	
R-squared	0.173	0.216	0.177		0.175		0.179	

This table presents a modified version of the main regression of Table 5. The single independent variable for size-related returns (SIZRET) is now replaced with two independent variables associated with the size tilt of the fund's Morningstar category (CATSIZ) and the deviation of the fund's size tilt from the Morningstar category average (FUNDSIZ). Similarly, VALRET is replaced with CATVAL and FUNDVAL. Key coefficient estimates are presented in bold. Tests for differences in the coefficient estimates on category-related returns and deviation from category are presented at the bottom of the table. Controls include lagged fund flows from month$t$-19, lagged values of log of fund size, log of fund age, expense ratio, load fund dummy, and return volatility. Standard errors (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

returns are attributable to the average style category return or deviation from it; effectively, such investors are simply chasing all returns not attributable to market risk.

6. Conclusion

What factors do investors consider when evaluating equity mutual fund performance? We addressed this question by analyzing the net flows into actively managed funds. Our key insight is that investors who attempt to identify a skilled active manager will strip out any fund-level returns that reasonably can be traced to a fund's exposure to factors known to affect cross-sectional equity returns. Fund flows should respond to alpha, but how do investors calculate a fund's alpha? At one extreme, unsophisticated investors may evaluate funds solely based on their market-adjusted returns. At another extreme, sophisticated investors will consider all available factors, both priced and unpriced, to assess a fund's performance.

Our empirical analysis has revealed that investors behave as if they are concerned about market risk, but are largely unaware of other factors that drive equity returns. Thus, when we ran a horse race between six asset-pricing models, the CAPM is able to best explain variation in flows across mutual funds. In additional analyses, we decomposed the returns of each mutual fund into eight components: a seven-factor alpha and flows associated with market, size, value, momentum factors, and three industry factors. We have found that flows respond to each of the eight return components, but to varying degrees. In general, the fund alpha generated the largest flow response. The response of flows to a fund's momentum-related return rivals that of the response to alpha. At the other extreme, flows are least sensitive to the fund returns that can be traced to market risk (beta). We have found some evidence that investors attend to the value, size, and industry tilts of a fund when assessing managerial skill, but these effects are much weaker than those we observed for a fund's beta. Moreover, we have found that investors strongly respond to the factor-related return associated with a fund's Morningstar-style category. Since the category-level return is not under the control of the manager, this result suggests some mutual fund investors confuse a fund's category-level performance and manager skill. However, in contrast to the style-investing story, we have found that flows are as, or more, responsive to deviations from style category returns as to the style category returns themselves. Thus, investors may not be focusing on style categories specifically, but may be simply responding to fund returns that can be attributed to style.

When assessing fund performance, investors will obtain the most precise estimates of managerial skill when they strip out all factor-related returns. Hence, we interpret these results as suggestive that investors vary in their sophistication level, and more sophisticated investors use more sophisticated benchmarks.

To test more directly the hypothesis that investor sophistication plays a role in the fund-flow relations that we document, we partitioned our sample based on funds' distribution channel and on periods of high and low sentiment. We separately analyzed return and fund flow relationships for wealthy and less wealthy investors at a large discount brokerage. We have found that the flows of investors who are likely more sophisticated—direct-sold fund investors, investors trading during low-sentiment periods, and wealthier investors—are generally less responsive to factor-related returns, suggesting that they are more aware that those returns are not indicative of the skills of the fund manager.

To adjust for factor-related returns when evaluating a fund, an investor needs to know the factor return. Sophisticated investors will seek out this information. But less sophisticated investors may not be aware of size, value, momentum, or industry returns. The market's performance, however, is ubiquitously reported. This may be one reason why investors do pay attention to market risk when evaluating mutual fund managers.

References

Akbas, F., W. J. Armstrong, S. Sorescu, and A. Subrahmanyam. 2015. Smart money, dumb money, and capital market anomalies. *Journal of Financial Economics* 118:355–82.

Agarwal, V., T. C. Green, and H. Ren. 2015. Alpha or beta in the eye of the beholder: What drives hedge fund flows? Working Paper.

Bailey, W., A. Kumar, and D. Ng. 2011. Behavioral biases of mutual fund investors. *Journal of Financial Economics* 102:1–27.

Barber, B. M., and T. Odean. 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55:773–806.

Barber, B. M., T. Odean, and L. Zheng. 2005. Out of sight, out of mind: The effects of expenses on mutual fund flows. *Journal of Business* 78:2095–119.

Barberis, N. and A. Shleifer. 2003. Style investing. *Journal of Financial Economics* 68:161–99.

Ben-Rephael, A., S. Kandel, and A. Wohl. 2012. Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics* 104:363–82.

Bergstresser, D., and J. Poterba. 2002. Do after-tax returns affect mutual fund inflows? *Journal of Financial Economics* 63:381–414.

Berk, J. B., and R. C. Green. 2004. Mutual fund flows and performance in rational markets. *Journal of Political Economy* 112:1269–95.

Berk, J. B., and J. H. Van Binsbergen. 2015. Measuring skill in the mutual fund industry. *Journal of Financial Economics*. 118:1–20.

———. 2016. Assessing asset pricing models using revealed preference. *Journal of Financial Economics* 119: 1–23.

Brown, K. C., W. V. Harlow, and L. T. Starks. 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance* 51:85–109.

Brown, S. J., W. N. Goetzmann, T. Hiraki, N. Shiraishi, and M. Watanabe. 2003. Investor sentiment in Japanese and US daily mutual fund flows. Working Paper.

Calvet, L. E., J. Campbell, and P. Sodini. 2007. Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy* 115:707–47.

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

Chalmers, J., and J. Reuter, 2013. What is the impact of financial advisors on retirement portfolio choices and outcomes? Working Paper.

Chevalier, J., and G. Ellison. 1997. Risk taking by mutual funds as a response to incentives. *Journal of Political Economy* 105:1167–200.

Chiu, H. H., and O. Kini. 2014. Equity issuance, equity mutual fund flows, and noise trader sentiment. *Review of Finance* 18:749–802.

Christoffersen, S. E., R. Evans, and D. K. Musto. 2013. What do consumers' fund flows maximize? Evidence from their brokers' incentives. *Journal of Finance* 118:201–35.

Clifford, C. P., J. A. Fulkerson, B. D. Jordon, S. R. Waldman. 2013. Risk and fund flows. Working Paper.

Cooper, M. J., H. Gulen, and P. R. Rau. 2005. Changing names with style: Mutual fund name changes and their effects on fund flows. *Journal of Finance* 60:2825–58.

Coval, J., and E. Stafford. 2007. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics* 86:479–512.

Cremers, M., A. Petajisto, and E. Zitzewitz. 2013. Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review* 2:1–48.

Dhar, R., and N. Zhu. 2006. Up close and personal: Investor sophistication and the disposition effect. *Management Science* 52:726–40.

De Andrade, F. Jr. 2009. Measures of downside risk and mutual fund flows. Working Paper.

Del Guercio, D., and J. Reuter. 2013. Mutual fund performance and the incentive to generate alpha. *Journal of Finance* 69:1673–704.

Del Guercio, D., and P. A. Tkac. 2002. The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds. *Journal of Financial and Quantitative Analysis* 37:523–57.

Del Guercio, D., and P. A. Tkac. 2008. Star power: The effect of Morningstar ratings on mutual fund flow. *Journal of Financial and Quantitative Analysis* 43:907–36.

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

———. 2004. The capital asset pricing model: Theory and evidence. *Journal of Economic Perspectives* 18: 25–46.

———. 2015. A five-factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Ferson, W. and J. Lin. 2014. Alpha and performance measurement: The effect of investor disagreement and heterogeneity. *Journal of Finance* 69:1565–96.

Geng, C., X. Li, A. Subrahmanyam, and H. Yu. 2015. Do wealthy investors have an informational advantage? Evidence based on account classifications of individual investors. Working Paper.

Gil-Bazo, J., and P. Ruiz-Verdu. 2009. The relation between price and performance in the mutual fund industry. *Journal of Finance* 64:2153–83.

Grinblatt, M., and S. Titman. 1989. Portfolio performance evaluation: Old issues and new insights. *Review of Financial Studies* 2:393–421.

Hirshleifer, D. 2001. Investor psychology and asset pricing. *Journal of Finance* 56:1533–97.

Huang, J. C., K. D. Wei, and H. Yan. 2012. Investor learning and mutual fund flows. Working Paper.

Ippolito, R. A. 1992. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry. *Journal of Law and Economics* 35:45–70.

Ivkovic, Z., and S. Weisbenner, 2009, Individual investor mutual fund flows. *Journal of Financial Economics* 92:223–37.

Karceski, J. 2002. Returns-chasing behavior, mutual funds, and beta's death. *Journal of Financial and Quantitative Analysis* 37:559–94.

Keswani, A., and D. Stolin. 2008. Which money is smart? Mutual fund buys and sells of individual and institutional investors. *Journal of Finance* 63:85–118.

Khorana, A. 2001. Performance changes following top management turnover: Evidence from open-end mutual funds. *Journal of Financial and Quantitative Analysis* 36:371–93.

Lynch, A. W., and D. K. Musto. 2003. How investors interpret past fund returns. *Journal of Finance* 58:2033–58.

Nanda, V., Z. J. Wang, and L. Zheng. 2004. Family values and the star phenomenon: Strategies of mutual fund families. *Review of Financial Studies* 17:667–98.

Pástor, L., and R. F. Stambaugh. 2002a. Investing in equity mutual funds. *Journal of Financial Economics* 63:351–80.

———. 2002b. Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics* 63:315–49.

Sensoy, B. A. 2009. Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics* 92:25–39.

Sirri, E. R., and P. Tufano. 1998. Costly search and mutual fund flows. *Journal of Finance* 53:1589–622.

Spiegel, M., and H. Zhang. 2013. Mutual fund risk and market share adjusted fund flows. *Journal of Financial Economics* 108:506–28.

Sun, Y. 2014. The effects of index funds and ETFs on money management fees. Working Paper, MIT.

Teo, M., and S.-J. Woo. 2004. Style effects in the cross-section of stock returns. *Journal of Financial Economics* 74:367–98.