

**USC FBE FINANCE SEMINAR**  
*presented by Diane Del Guercio*  
FRIDAY, Nov. 14, 2003  
10:30 am – 12:00 pm; **Room: JKP-204**

**Star Power:  
The Effect of Morningstar Ratings on Mutual Fund Flows**

Diane Del Guercio and Paula A. Tkac  
November 2003

Diane Del Guercio  
Lundquist College of Business  
University of Oregon  
Eugene, OR 97403-1208  
541-346-5179  
dianedg@oregon.uoregon.edu

Paula A. Tkac  
Research Department  
Federal Reserve Bank of Atlanta  
1000 Peachtree St. NE  
Atlanta, GA 30309-4470  
404-498-8813  
Paula.Tkac@atl.frb.org

We gratefully acknowledge research support from a grant from the Institute for Quantitative Research in Finance (Q-Group). We thank John Rekenthaler, Director of Research at Morningstar, Inc., for providing data, and Paul Gozali of Morningstar's research department for answering numerous data questions. We thank Robert Battalio, John Chalmers, Larry Dann, Harry and Linda DeAngelo, Jerry Dwyer, Lily Fang, Laura Field, Woodrow Johnson, Aditya Kaul, Elizabeth Odders-White, Cesare Robotti, Jonathan Taylor, Tracie Woidtke, and seminar participants at the 2002 WFA meetings, 2000 Pacific Northwest Finance Conference, University of Utah and the Atlanta Fed Brown Bag series for helpful comments. The views expressed here are the authors' and do not necessarily represent the views of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

**Star Power:  
The Effect of Morningstar Ratings on Mutual Fund Flows**

**Abstract**

Morningstar is the leading information intermediary in the mutual fund marketplace, and is widely believed to influence fund investors. The discreteness of Morningstar's one- to five-star rating system, and the known monthly frequency with which Morningstar updates fund ratings, imply that a rating *change* is an unambiguous and easily observable event. Exploiting these features, we revisit the predominantly cross-sectional literature on the relation between fund flow and performance, using event-study methods on a sample of 3,388 domestic equity mutual funds. We find that the Morningstar rating has independent power to influence the allocation decisions of fund investors. In particular, we find significant positive abnormal flow following rating upgrades, and negative abnormal flow following rating downgrades, typically ranging from 13-30% of normal flow. We show that this Morningstar effect is not explained by either the level of, or the concurrent change in, other performance measures (e.g., Jensen's alpha). We report new evidence on the timeliness of investor response to fund quality updates, and on the substantial asymmetry in the response to improving versus deteriorating performance. Armed with a more refined characterization of the implicit incentives facing mutual fund managers, we describe how our findings can inform more powerful tests of managerial behavior.

“[T]he brand that has emerged as dominant in the 1990s is not Fidelity, Putnam or even Merrill Lynch—but instead is Morningstar.”

- R. Pozen, *The Mutual Fund Business*, 1998, p. 75

## 1. Introduction

Information intermediaries, such as *Consumer Reports* with its product ratings or *U.S. News and World Report* with its college rankings, are commonplace in markets where consumers face important purchase decisions. In an environment involving both costly search for information and significant decision costs of evaluating various product quality and cost attributes, intermediaries create value by providing an accessible and easy-to-interpret quality rating from a reputable and unbiased source. While true automobile enthusiasts may relish poring over engine specifications before making a car purchase, many others would prefer instead to consult a summary quality rating from *Consumer Reports*. An information intermediary can help those car consumers with high decision costs make better choices than they would if the low-cost quality rating were not available (e.g., help them avoid purchasing a model with a history of poor reliability).

As first suggested by Sirri and Tufano (1998), choosing a mutual fund is similar to choosing a car in that there are numerous possible alternative products available and product quality is difficult to measure or predict. The mutual fund industry’s analog to *Consumer Reports* is Morningstar, Inc., the Chicago-based rating agency. A casual reading of the business press and mutual fund advertisements, as well as statements in several academic studies, suggest that Morningstar is the “most popular” and “best-known” ranking service among mutual fund investors.<sup>1</sup> Thus, Morningstar star ratings arguably are primary inputs to many investors’ decisions and as a consequence, are a likely influence on the behavior of mutual fund managers as they seek to obtain a higher star rating. In this study, we test whether the particular package of fund quality information embodied in the Morningstar star rating, over and above its correlation with other performance measures, has power to influence investors.

---

<sup>1</sup> Blume (1998), Sharpe (1998), and Morey (2000) scrutinize Morningstar’s rating algorithm, while Blake and Morey (2000) test the predictive power of star ratings.

The discrete nature of Morningstar's one- to five-star rating system, and the monthly frequency with which Morningstar updates fund ratings, create an opportunity to obtain insights on mutual fund investor decision-making that would be difficult to uncover using other performance measures. The discreteness and wide-availability of the rating at low cost to investors implies that a rating *change* is an unambiguous and easily observable event for fund investors. In addition, star ratings represent a well-defined packaging of information on fund quality that arrives at known monthly intervals. The importance of these features is that rating change events, or the monthly updates in investor information observable to the researcher, are identical to the signals observed by investors. This is not the case with traditional continuous performance measures, such as Jensen's alpha, which do not as easily lend themselves to a well-defined test of investor response to a change in performance. Such a test would require assumptions about the appropriate horizon of the relevant performance measure (e.g., monthly alpha changes, annual alpha changes, etc.), and about how large a change in performance prompts investors to reallocate assets. In contrast, isolating a response to discrete star rating changes amounts to a clean event-study test of whether, and how, investors respond to Morningstar's monthly quality updates.

We seek to determine whether the manner in which Morningstar re-packages the raw information matters to investors. Because all performance measures are ultimately functions of the same underlying series of returns, isolating the impact of Morningstar's particular packaging of return data requires a methodology that can disentangle its effect from that of other performance measures investors might use to evaluate funds. In short, a key task in identifying an independent Morningstar effect is to properly control for other factors that influence fund flow, a measure of the aggregate net effect of individual fund investor allocation decisions.

Our event-study method benchmarks the flow for a fund experiencing a rating change against its expected flow, conditional on both its more recent and its longer term past performance, in both raw and risk-adjusted terms. This isolates the influence of the star rating change apart from the underlying

return performance that led to it, and allows us to measure the association between the packaging of information in the Morningstar rating change and the flow response that follows it. In addition, estimating a fund-specific flow benchmark explicitly models the heterogeneity across funds in the flow-performance relation that is inherent in the data but ignored in the previous literature.

Previous studies of the flow-performance relation, such as Sirri and Tufano (1998) and Gruber (1996), have typically adopted a panel data approach that focuses squarely on the cross-section. The oft-cited result in this literature is that the funds at the top of the cross-sectional distribution garner a disproportionate amount of investor flow, while funds at the bottom do not suffer significant outflow. Chevalier and Ellison (1997) and Brown, Harlow, and Starks (1996) have used this empirical relation to predict that mutual fund managers have powerful tournament-like incentives to secure a position at the top of the distribution. These studies implicitly assume that the flow-performance relation is the same for all funds; i.e., that the average flow differentials in the cross-section carry over into the time-series for individual funds. We show that there is substantial heterogeneity across funds, and that if ignored, will lead to misleading flow estimates. This heterogeneity is not particularly surprising in light of the unobservable differences across funds in clienteles (e.g., percentage of clients in 401(k) plans), advertising expenditures, distribution channels, etc.

Fundamental questions remain unanswered regarding the empirical flow-performance relation that underlies managerial incentives to increase fund assets. For example, because most previous studies assume an annual tournament, we know little about whether investors revise their investment allocations and seek performance data at higher frequencies. Most previous studies have focused on the level of performance rather than performance changes. Thus, the question of whether investors react similarly to an improvement in performance as they do to a deterioration of the same magnitude has not yet been explored. Addressing these issues requires a well-defined performance change event, data on fund flow that matches the frequency of the relevant performance change, and a methodology that can properly link the flow response to the performance change that prompted it. An event-study

using monthly Morningstar star rating changes has these desired properties and thus furthers our understanding of investor decision-making in the mutual fund market. Armed with a more refined characterization of the implicit incentives facing mutual fund managers, our findings can then, in turn, inform more powerful tests of managerial behavior.

We use monthly data on 3,388 domestic equity mutual funds over the period November 1996 to October 1999, and identify over 12,000 rating change events. Our analysis of rating changes in seasoned funds indicates that investors value the packaging of fund quality information that a star rating provides, and view the rating change as “new” information. Specifically, we find significant positive abnormal flow following rating upgrades, and negative abnormal flow following rating downgrades, typically ranging from 13-30% of normal flow. Importantly, this Morningstar effect is not explained by either the level or changes in other performance measures (i.e., Jensen’s alpha and Sharpe measure) concurrent with rating changes.

We also find evidence of significant asymmetry in investor responses to improving versus deteriorating performance. An upgrade from four- to five-stars has the largest effect on flow in the six months following the upgrade, resulting in \$32 million in abnormal flow, or 25% above normal expected flow. While a downgrade from five- to four-stars does not result in negative abnormal flow, we do find a very large negative and significant response to a downgrade from four- to three-stars, on the order of twelve times larger than the normal flow for these funds. These results suggest that falling to a 4-star rating is a non-event for the typical fund, but that investors severely punish funds whose performance falls to a 3-star rating. Together, these event-study results paint a different picture of the flow-performance relation than the smooth convex relation of previous cross-sectional analyses.

Most of the observed flow responses are significant beginning in event month 0, indicating that investor response is detectable at the same monthly frequency that Morningstar releases new ratings. Abnormal flow remains significant in many cases for as many as seven months after the rating change, suggesting that some investors are slow to respond to the rating update. This immediate, but

persistent, response suggests the presence of both vigilant investors that closely monitor rating changes, and other investors that perhaps only respond after a rating appears in fund advertisements. However, because in an open-end fund there is no mispricing to exploit for acquiring the rating information before other investors, a prolonged abnormal flow response is not anomalous.

Surprisingly, we also find that rating changes occurring at year-end, when anecdotally investors reevaluate their portfolios, do not elicit a response different from those in other months. This lack of an end-of-the-year effect is consistent with the findings of Goetzmann and Massa (2002) for a large sample of individual accounts in an S&P500 index fund. Overall, our results suggest that the tournament among fund managers is not annual, as implicitly assumed in Brown, Harlow and Starks (1996) and Chevalier and Ellison (1997). We find it to be much more of an ongoing and high frequency tournament, suggesting a more complicated and dynamic strategic environment for fund managers than assumed in current tests of managerial incentives and behavior. Therefore, our results set the stage for future theoretical and empirical research into these questions. As we describe in section 4.4, our findings on monthly flow response suggest that more powerful tests of manager behavior are possible.

## **2. Data**

### *2.1 Background on Morningstar's Star Rating*

Morningstar first offered its star ratings in 1986 by paid subscription to its print product *Morningstar Mutual Funds*, but they are now available free on several popular websites including Morningstar.com, Yahoo! Finance, and MSN's Moneycentral. Morningstar covers virtually all funds in existence and calculates star ratings for all funds that are at least three years old. Star ratings are recomputed monthly and available on Morningstar's website approximately three business days after month-end. Below we briefly describe the rating process, but interested readers should see Blume (1998) and Sharpe (1997, 1998) for a thorough description of Morningstar's rating algorithm.

A fund's Morningstar star rating is based on its historical performance with respect to both return and risk relative to its peer group. Specifically, Morningstar uses a rolling 36 months of load-adjusted returns to compute a three-year risk-adjusted rating for each fund. For each of four mutually exclusive peer groups (domestic equity, international equity, taxable bond, and municipal bond), it then ranks funds according to this three-year risk-adjusted rating. A three-year star rating is assigned for each peer group based on this ranking. Specifically, funds with risk-adjusted ratings in the top 10% of their peer group are assigned five stars, the next 22.5% receive four stars, the next 35% receive three stars, the next 22.5% receive two stars, while the bottom 10% of funds in each peer group receive one star. Morningstar computes a five-year and ten-year star rating in the same way, with the only difference being the number of months of load-adjusted returns used in the calculations. The overall star rating is computed as a weighted average of these ratings.<sup>2</sup> For funds too young to have a five- and ten-year rating, the overall rating equals the three-year rating. This rule apparently applies to the majority of funds, as Sharpe (1997) reports that overall star ratings coincide with three-year star ratings for 72.2% of domestic equity funds at the end of 1996. In this study we analyze the overall star rating because this appears to be the most widely used and cited among publications geared to investors.<sup>3</sup>

## 2.2 Description of our Sample

We obtain the monthly time-series of Morningstar overall star ratings from November 1996 to October 1999 directly from Morningstar, Inc. These data include star rating histories of all funds in Morningstar's domestic equity star rating category as of October 1999. To be included in this database the fund must exist and have a star rating as of October 1999. Survivorship bias is a legitimate concern given the high rate of liquidation and mergers over this period. Therefore, through the use of monthly editions of Morningstar's CD-ROM *Principia* product, we fill in star ratings for all of the funds that

---

<sup>2</sup> For example, for a fund at least 10 years old, its overall star rating equals 0.2 times its three-year rating, plus 0.3 times its five-year rating, plus 0.5 times its ten-year rating.

<sup>3</sup> In support of this, we collected information on all mutual fund advertisements that appeared during the month of June 1999 in either *Barron's Money Magazine*, or the Money & Investing section of the *Wall Street Journal*. Of 43 ads that advertised a Morningstar rating, 43, or 100% of them included the overall rating. Of these 43, 37 were for domestic equity funds. In addition, the overall star rating is the only Morningstar rating provided on Yahoo! Finance and MSN Money Central.

disappeared over this period, totaling 4,040 fund-months. To match the rest of the sample, we only include dead funds if they were old enough to have a star rating by October 1999. We supplement fund ratings with data on returns, total net assets, and other fund characteristics from the *1999 Survivor-bias Free Mutual Fund Database* from the Center for Research in Security Prices (CRSP). We perform numerous checks to ensure that the data from the two sources are properly linked.<sup>4</sup> Finally, we delete all fund-months that coincide with a fund merger since the fund flow numbers are likely to be distorted. Our final dataset contains 111,715 fund-months from 3,388 distinct funds.<sup>5</sup>

Since our goal is to measure the response of mutual fund investors to a change in a fund's Morningstar rating, it is important that the ratings we study reflect information available to investors at that time. Morningstar's master data file contains historical star ratings for each fund that reflect the rating algorithm currently in place, rather than the one in place at that time. Consequently, our analysis begins in November 1996 because Morningstar rewrote their master file when they changed their algorithm in October to exclude international funds from the domestic equity rating category peer group.

We confirm that the ratings data we use match those available to investors by comparing our star ratings to those on the original CD-ROM at two points in time, July 1997 and October 1999. Only 2.3% of the funds with a star rating for July 1997 in our data could not be found on the original CD-ROM, indicating that backfilling is not a concern. We also find that only 103, or 4.5% of the funds, have different star ratings listed in the two sources in July 1997. In October 1999, we find no discrepancies in star ratings between the two sources.

---

<sup>4</sup> We are grateful to Morningstar for providing a spreadsheet with fund name and a monthly time-series of total net assets. These data were invaluable in properly linking the two data sources. There are 173 funds on the spreadsheet from Morningstar that we could not identify as existing on the CRSP database. An additional 4,406 fund-months dropped out because CRSP lacked data on an important variable for our analysis (e.g., total net assets).

<sup>5</sup> We consider a share class (e.g., class A) of a fund to be a distinct fund for two reasons. First, Morningstar treats each share class independently for the purpose of assigning a star rating. In fact, we commonly observe that a star rating differs between classes of the same underlying fund. Second, since our main interest is in flow measurement, it is not obvious that flows of different classes are closely related. For this reason, we are not double-counting observations as would happen in a study of fund performance.

### 2.3 *Timing and Measurement of Fund Flows and Summary Statistics*

We analyze monthly net dollar flows in or out of a fund, using the standard definition in the literature of the change in total net assets minus appreciation:

$$\text{Flow}_{it} = \text{TNA}_{it} - \text{TNA}_{it-1} (1 + R_{it}).$$

This definition of fund flow assumes that flow occurs at the end of the month. Updated star ratings reflect the most recent month's performance and are available to investors at the beginning of the following month. Thus, if a rating change occurs using return data through October 1998, we record November 1998 as the month of the rating change and designate the flow in November as the first month possibly affected by the change.

Panel A of Table 1 provides summary statistics on monthly flow aggregated over the entire sample period for each Morningstar star rating group. The flow is measured and summed across all funds in the same rating group. Consistent with widely-cited statistics on the popularity of four- and five-star funds, these funds attract large dollar flows totaling \$349 billion, while funds with three-stars or less experience net outflows of \$133 billion. Thus, four- and five-star funds collectively attract 124.4% of all flow to domestic equity funds. Notice that unrated funds in our sample collectively receive more flow than funds rated with three stars or less.

Figure 1 provides similar information relating flow and star rating, but on a disaggregated basis. This figure depicts the familiar convex relation between fund flow and past performance as documented by Sirri and Tufano (1998) and others, and highlights the importance of identifying a methodology that can disentangle the Morningstar effect from other influences on fund flow.

## 3. **Empirical Approach and Methodology**

The typical approach in the literature for assessing the influence of performance measures on mutual fund flow relies on panel, or Fama-MacBeth-type, regressions. To investigate the question of whether Morningstar ratings drive fund flows we could employ this method by adding Morningstar

star rating dummies as explanatory variables in addition to other commonly included performance measures and control variables such as size, fund age and lagged flow. However, as we argue below, event study methods have several advantages over the cross-sectional panel approach.

For the purpose of comparison with previous studies, Table 1 Panel B contains the results of a pooled panel regression of fund flow on fund performance and non-performance characteristics. The results are very similar to those reported in the literature. Coefficients on all performance measures, as well as on their quadratic terms, are highly positive and significant, consistent with the well-documented convexity in the relation between flow and performance. The statistical significance and pattern of coefficients on the Morningstar dummy variables illustrate that controlling for other types of performance and fund characteristics, Morningstar ratings, and especially 5-star ratings, indeed have independent power for explaining the cross-section of fund flows. While this finding establishes that funds rated more highly by Morningstar experience more flow on average, it does not really address whether Morningstar ratings drive fund flows, or, whether fund flows are related to rating *changes*.

In principle, we could maintain the panel approach of Panel B by simply adding a dummy variable for each type of possible rating change (e.g, from a 4- to a 5-star rating, from a 2- to a 3-star rating, etc.). If the assumption that all funds have the same sensitivity to the right hand side variables were valid, this pooled panel approach would successfully estimate the average incremental flow to funds in the various Morningstar rating change categories. If this assumption is invalid, however, this method is not appropriate for the case of an autocorrelated dependent variable. As shown in Table 1 and elsewhere in the flow literature, the coefficient on lagged fund flow is positive and highly significant. When parameter heterogeneity is coupled with a lagged dependent variable, the panel estimator yields downwardly biased coefficient estimates.<sup>6</sup>

---

<sup>6</sup> For a thorough analysis of bias in dynamic heterogeneous panels, see Pesaran, Smith, and Im (1996) and references cited therein. It is the presence of the lagged dependent variable (fund flow) as a right-hand side variable that induces the bias in all estimated coefficients.

Indeed, all funds are not identical in terms of their sensitivity to performance changes. F-tests for homogeneity of the panel slope coefficients across funds reject this null hypothesis at the 1% significance level (not reported). We explicitly model this heterogeneity by using standard event study methods, estimating a flow-performance benchmark regression for each fund. Fund-specific benchmarks allow for both the heterogeneity observable in the form of performance sensitivities and unobserved heterogeneity across funds, such as reputation. Carrying ‘normal’ flow benchmark predictions into an event window, we are able to estimate the abnormal flow response to a Morningstar rating change for each event. The cross-sectional average of these estimates is then an unbiased and consistent estimate of the average incremental flow due to a rating change. In fact, this cross-sectional average is effectively the ‘mean group estimator’ commonly recommended in the econometrics literature as a solution to the bias problem in dynamic heterogeneous panel regressions.

In addition to unbiased estimation of the mean effect across funds, an important advantage of an event study approach is that it preserves the maximum amount of information for analysis of other questions of interest. We can, for example, easily test whether a Morningstar effect is stronger for larger funds, for rating upgrades vs. downgrades, for year-end performance changes, etc. Relative to a panel, we can more easily determine the horizon over which rating changes affect fund flows by extending the event window. In sum, by fitting a model of normal flow for each fund, we most accurately benchmark the expected flow around a Morningstar rating change and hence most accurately estimate the abnormal flow that is reliably purged of the influence of other variables. The following sections describe the details of how we apply event study methods in the context of mutual fund flows.

### *3.1 Frequency of Star Rating Change Events*

Panels C and D of Table 1 provide sample summary statistics on the frequency of star rating change events. We observe 10,735 rating changes, which represent close to 10% (10,735/111,715) of total fund-months. The vast majority of these changes are upgrades or downgrades of one star. We do

not analyze changes greater than one star, which represent less than 1% of all rating changes. Panel D reports the frequency count of funds receiving an initial star rating upon their third anniversary as a fund, which is a special case of a change in star rating. We identify 1,637 initial star rating events, with a breakdown by initial star rating that closely follows that of 10%, 22.5%, 35%, 22.5%, 10% expected under Morningstar's algorithm.

### 3.2 Event study method for calculating abnormal flow

An event-study allows us to isolate the incremental flow due to a rating change as separate from normal flow to that fund. Similar to a traditional event study of stock returns, to compute expected, or normal, flow we estimate a benchmark regression for each individual fund  $i$ :

$$F_t^i = \mathbf{g}^i + \mathbf{b}_1^i SF_t^i + \mathbf{b}_2^i RET_{t-1}^i + \mathbf{b}_3^i \Delta \mathbf{a}_{t-1}^i + \mathbf{b}_4^i (\Delta \mathbf{a}_{t-1}^i)^2 + \mathbf{b}_5^i F_{t-1}^i + \mathbf{e}_t^i$$

where  $F_t^i$  is the net dollar flow to fund  $i$  at month  $t$  and  $SF_t^i$  is the aggregate net flow to all funds in the same style category as fund  $i$  at month  $t$ ,  $RET_{t-1}^i$  is fund  $i$ 's return at month  $t-1$ ,  $\Delta \mathbf{a}_{t-1}^i$  is the change in fund  $i$ 's Jensen's alpha from month  $t-2$  to  $t-1$ , and  $F_{t-1}^i$  is the net flow to fund  $i$  at month  $t-1$ . We define the first month an investor would have information available about a new Morningstar rating as event time 0. To estimate the coefficients for the benchmark flow regressions we use 24 months of data, ending 3 months before time 0 (i.e., event month -26 to -3). We calculate the fund alphas using 36 months of return data and the S&P500 as the market index, but drop funds with less than 24 months of data. We use nineteen Morningstar-defined style categories to identify each fund's style and to compute aggregate style flows. The appendix contains details of our procedure for constructing these style flow benchmarks.

The ultimate goal in fitting a benchmark to obtain estimates of normal flow is to parsimoniously, but sufficiently, capture the influence on fund flow of all other performance and non-performance characteristics besides the change in Morningstar rating. Our chosen regressors are

intended to control for the current popularity of a fund's investment style, a fund's long-term performance record, and recent innovations in a fund's raw and risk-adjusted performance. We include a squared term on the change in Jensen's alpha to account for potential convexities between flow and recent performance. Including lagged flow as a regressor is equivalent to an infinitely distributed lag model that contains the fund's full performance history, and thus is a way to control for the fund's long-term performance record. While there are clearly other candidates for inclusion in the benchmark model, most notably fund size and Sharpe ratio, we later provide supportive evidence that our specification sufficiently captures the important systematic influences on flow.<sup>7</sup>

Regarding the choice of estimation period, because our data occurs monthly, we face the typical tradeoff between parameter stability and the statistical power that comes from a larger sample. Analysis of benchmark residuals and other diagnostics suggest that a 24-month estimation period provides the best balance between parameter stability and precision. If we extend the horizon to 36 months, diagnostics on benchmark residuals display significant serial correlation, especially for the young funds with a short performance history. While we have used a variety of estimation periods ranging from 12 to 36 months, the 24-month period that we report are the most statistically well-behaved. The results are qualitatively similar when we vary the length of the estimation period. In addition, our chosen test statistics described in 3.4 explicitly adjust for benchmark fit by deflating each estimate of abnormal flow by the RMSE from the benchmark estimation. This ensures that imprecisely estimated benchmarks do not exert a large influence on our results. The next section provides supportive evidence of the reliability of our benchmark.

### *3.3 Benchmark Diagnostics*

We have undertaken a battery of diagnostic tests to assess the reliability of the benchmark model. The top half of Table 1 panel E contains the sample distribution of coefficients and the  $R^2$  of

---

<sup>7</sup> For example, fund size might be an obvious candidate given that several cross-sectional studies have documented that larger funds receive more flow on average. In a time-series benchmark regression, this effect is likely absorbed into the intercept, which picks up the average flow not due to other included factors. We provide direct evidence that our benchmark has controlled for fund size and Sharpe ratios later in the paper.

the benchmark regressions, and suggests that our relatively parsimonious model results in an overall good fit. The median  $R^2$  across events is 0.43, with 25% of these events exhibiting an  $R^2$  of over 0.66. We also report in the table that the number of rating change events per month is reasonably dispersed throughout the sample period, indicating that calendar time clustering is not a concern.

If our benchmark is working properly we should observe much higher normal flow for those funds that have performed well in the past. Figure 2, which plots the average time 0 normal flow for all funds according to their Morningstar star rating at time  $-1$ , illustrates that the benchmark is performing as expected. Average normal flow is negative for funds below 3-stars, while the typical 5-star fund has large monthly normal flow of \$34 million, closely mirroring the pattern in the raw total flow picture of Figure 1. Thus, our specification of normal flow appears to be capturing a fund's longer-term performance track record reasonably well, which in turn gives us confidence in properly labeling the flow response to a rating change as 'abnormal' flow.

The bottom half of Table 1 panel E contains cross-sectional correlations, pooled over all funds and over all months in the estimation period, of residual flow with benchmark regressors, as well as of predicted (normal) flow with benchmark regressors. There is a strong positive correlation between predicted flows and fund size implying that large funds have higher normal flow. This is consistent with the positive relation commonly reported in the annual cross-sectional literature, and is evidence that the benchmark captures the relation between flow and fund size. As expected from a well-specified benchmark, the cross-sectional correlation between the estimation residuals and all of the regressors, including fund size, are not significantly different from zero. The insignificant positive correlation between the standardized residuals and fund size implies that our measure of abnormal flow is not related to fund size and also suggests that heteroskedasticity of the residuals is unlikely to be a problem. Overall, we are reasonably confident that the benchmark is well-specified, and we provide additional evidence on benchmark performance when we describe the results.

### 3.4 Statistical Tests

Our measure of abnormal flow around a change in Morningstar star rating is therefore:

$$AF_t^i = F_t^i - \hat{g}^i - \hat{b}_1^i SF_t^i - \hat{b}_2^i RET_{t-1}^i - b_3^i \Delta a_{t-1}^i + b_4^i (\Delta a_{t-1}^i)^2 - \hat{b}_5^i F_{t-1}^i.$$

In words, the abnormal flow to fund  $i$  at time  $t$  equals the actual flow at time  $t$  minus the expected flow due to its aggregate style flow, lagged return, lagged change in alpha, the square of lagged change in alpha, and lagged flow, minus the average abnormal flow to fund  $i$  ( $\hat{g}^i$ ). This average abnormal flow captures fund-specific determinants of flow that do not vary over time.

In our main tests we report abnormal flow statistics over the event window month 0 to +6.<sup>8</sup> Unlike an event study of returns where market efficiency implies an immediate stock price reaction to new information, the impact of a rating change on fund flows may persist over many months, and because there is no mispricing to exploit in an open-end mutual fund, a delayed response would not be anomalous. An immediate flow reaction to a rating change is presumably due to vigilant investors who monitor funds on a monthly basis. A delayed flow response is also plausible either because more casual investors take time to respond to the new rating, or because the majority of these investors primarily respond to fund advertisements touting the new rating, which take time to place in publications.

We apply the event study methods described in Dodd and Warner (1983) to these abnormal flow estimates. Specifically, we focus on standardized abnormal flows, where event period abnormal flows are standardized by the estimated forecast variance of the abnormal flow. An advantage of analyzing standardized abnormal flows is that it prevents funds with large forecast variances (noisy flows) from dominating the statistical tests. Under this method, funds with more precisely measured abnormal flows are implicitly weighted more heavily.

---

<sup>8</sup> Given that Morningstar releases ratings within a few business days of month end, event month 6 is technically less than seven months after a rating change (i.e., six months plus a fractional month). Nonetheless, we refer to this throughout the paper as seven months after a rating change for simplicity.

In addition to these month-by-month measures we also compute cumulative standardized abnormal flows by summing the standardized abnormal flows for each fund from event time 0 to  $t$ , and then dividing by the square root of the number of months used in the cumulation. Following standard practice, at each event date  $t$  we compute average standardized abnormal flows ( $ASTAF_t$ ) and average cumulative standardized abnormal flows ( $ACSTAF_t$ ) by averaging across  $N$  events (funds). By grouping funds according to their pre-change star rating, we define events as a change from 3- to 4-stars, from 4- to 5-stars, etc., and assess the flow response measures for statistical significance.

We use the standardized cross-sectional test of Boehmer, Musumeci, and Poulsen (1991) to assess the statistical significance of event period  $ASTAF_t$  and  $ACSTAF_t$ . Specifically, we divide  $ASTAF_t$  and  $ACSTAF_t$  by its contemporaneous cross-sectional standard error. Because this measure uses event-period abnormal flows to calculate the standard errors, it adjusts for any change in the variance of the standardized abnormal flow from the estimation period to the event-window. This is the most appropriate statistic to use when testing out of sample event-window abnormal flows. Since we observe larger event-window cross-sectional variances, if anything, this statistic will bias against finding a significant Morningstar effect. For example, if funds exhibit wide variation in their response to a change from 5- to 4-stars, a large cross-sectional standard error and lower t-statistic will make it more difficult to reject the null hypothesis of no abnormal flow response.

We also report statistics on the percentage of funds that experience positive standardized abnormal flows for each event date. Specifically, we report the results of a chi-square test under the null hypothesis that 50% of sample funds have positive standardized abnormal flow. This nonparametric test statistic complements the parametric tests by providing information on the distribution of standardized abnormal flow.

#### **4. Results**

Our event study method of analyzing rating changes is designed to isolate and quantify the magnitude of the “Morningstar effect” apart from other influences on fund flow. Toward this end, we

first establish statistical significance using the standardized abnormal flow measures and test statistics described in the previous section, which have the most well-behaved statistical properties for testing the null hypothesis of zero abnormal flow response.<sup>9</sup> The focus on standardized measures, however, does not lend itself to assessing the economic significance of abnormal flow. We address the economic significance of the Morningstar effect separately in section 4.1.2, using non-standardized estimates of abnormal flow.

We report results based on the full sample of funds experiencing a rating change, including those that undergo subsequent rating changes within the seven-month event window. Thus, estimates of average abnormal flow represent the unconditional Morningstar effect, equivalent to the ex-ante expected value of a Morningstar rating change. This is the most appropriate sample to study because investors cannot know in advance whether a rating will change again. For completeness, however, section 4.3 contains analysis of a sub-sample of funds without further event-window rating changes.

#### *4.1 Abnormal Flow Response to Star Rating Changes*

This section presents event study results for over 10,000 monthly star rating changes during the 1996 to 1999 period. Table 2, panels A and B report the average standardized abnormal flow (ASTAF) for all rating upgrades and downgrades for each event month from 0 to 6, while panels C and D contain the corresponding average cumulative standardized abnormal flow (ACSTAF) estimates. Abnormal flow in this context can be interpreted as a fund's actual flow relative to that expected if it had maintained its pre-change star rating. In most event months, we can reject the null hypothesis of

---

<sup>9</sup>To minimize the possibility that extreme observations drive the results, we also delete the top and bottom 1% of standardized abnormal flows at each event date in each rating subsample, and we note that inferences do not change when we include them. We examine the estimation period residuals and find that serial correlation is low (mean of  $-0.06$ ), indicating a reasonably good benchmark fit. In an unreported specification test, we have also checked that average standardized abnormal flows (ASTAF<sub>t</sub>) are indeed zero for 'non-event' months when the rating remains unchanged. Specifically, we identify funds where the star rating has not changed over the 8 months prior to time 0 (i.e., the last 6 months of the estimation period and pre-'non-event' window). For this exercise, we group funds according to their constant-over-time star rating, and compute ASTAF in the same way that we do for star rating change events. We find that in all but one case, the ASTAF are not statistically different from zero. Specifically, the ASTAF for three-star funds at time 0 is  $-0.09$  and statistically significant ( $n=9894$ ). Most notably, for four- and five-star funds where we later report the largest flow response, we find zero flow response to these non-events. This lends additional support to our interpretation of the flow response of a rating change as causal, or as truly 'abnormal' flow. Finally, we have also examined the ASTAF of rating changes in the pre-event period ( $-2, -1$ ), and find no evidence of a flow response at time  $-2$ , but some evidence of a response at time  $-1$ . In all of these cases, the ASTAF at time  $-1$  is the same sign as it is at time 0, indicating that this is

no Morningstar effect, as most ASTAF and ACSTAF are significantly different from zero. In fact, in many cases abnormal flow is significant out to event month six, suggesting that some investors respond to a rating change with a considerable lag.

We generally observe a surprisingly consistent flow response across the rating change categories, positive for rating upgrades and negative for rating downgrades, with confirmatory evidence from the non-parametric  $c^2$  tests. In particular, the attainment of 5-star status stands out as an overwhelmingly positive event for the typical fund. In the 4- to 5 star upgrade category, all test statistics are highly significant in every month in both panels, the magnitude of the mean standardized response is nearly three times that for any other rating change category, and the percentage of positive abnormal flow observations ranges from 53 to 61%.

There are a few exceptions to the general finding that upgrades are followed by positive abnormal flow and downgrades are followed by negative abnormal flow. Although the abnormal flow has the correct sign, upgrades from 1- to 2-stars and downgrades from 2- to 1-star exhibit only weak evidence of a flow response. This might not be surprising because these funds still remain near the bottom of the performance distribution. More surprising is the result that downgrades from 5- to 4-stars are associated with *positive* ASTAF and ACSTAF, although, the  $c^2$  tests indicate that the median fund response is not different from zero.

#### 4.1.1 *The star rating effect is not driven by Jensen's alphas or Sharpe ratios*

As stated earlier, an advantage of our fund-specific benchmark is its ability to control for the influence of all variables included in the benchmark, including a fund's performance history. Table 3 presents evidence that our measure of abnormal flow that we attribute to a star rating change is indeed net of the effect of alternative performance measures, such as the Sharpe ratio and Jensen's alpha.

---

most likely a timing issue. We note that if we include up to  $t = -1$  in the estimation period, we find similar results in the event window. All unreported diagnostic tests are available upon request.

Specifically, panel A of Table 3 presents analysis of variance tests that compare the ability of Morningstar star rating changes to explain variation in abnormal flow against the ability of these alternative contemporaneous performance measures.

Table 3 reports the results of two-way analysis of variance (Type II sum of squares) tests on seven-month cumulative standardized abnormal flows (CSTAF) following a fund's Morningstar rating change. To put alpha and Sharpe measures on equal footing with Morningstar's star rating in terms of its discreteness, we examine various groupings of funds based on rankings by these two measures. We examine three alternative definitions of alpha and Sharpe groups: quartiles, quintiles, and a "Morningstar-type" grouping where the top 10% of alphas (Sharpes) are in quintile 5, the next 22.5% in quintile 4, the next 35% in quintile 3, the next 22.5% in quintile 2, and the bottom 10% in quintile 1. These groupings might be thought of as a simulated rating system, with each representing a different method investors might use to translate the continuous alpha or Sharpe measures onto a coarser grid of fund quality.

The insignificant F-statistics for the Sharpe and alpha groups indicate that conditional on a fund's Morningstar star rating change, none of the alpha or Sharpe categorizations has the power to explain cross-sectional variation in abnormal flows. Under each of the three grouping methods, we cannot reject the null hypothesis that the conditional average abnormal flows are equivalent across alpha and Sharpe measure groups. At the same time, Morningstar's star rating remains highly significant conditional on any of the alpha or Sharpe categorizations; we can reject the null hypothesis that Morningstar star rating change does not matter (i.e., that average abnormal flows are equivalent across all star ratings) at the 1% confidence level. Therefore, the abnormal flow around star rating changes of Table 2 are not driven by fund alpha and Sharpe measures. In particular, the large positive abnormal flow following an upgrade from 4- to 5-stars is not driven by high alpha, or high Sharpe measure funds, suggesting that our benchmark successfully incorporates a fund's long-term performance history.

We also examine whether *changes* in the simulated alpha and Sharpe ratings contemporaneous with Morningstar's rating change can explain cross-sectional variation in abnormal flows, where the simulated ratings are measured the same way as in Panel A. Our benchmark specification includes a monthly change in alpha as a control variable, so we would expect abnormal flow to be net of the effect of changes in Jensen's alpha. Panel B of Table 3 reports analysis of variance tests in which funds are grouped according to the change in a fund's alpha or Sharpe measure occurring at the same time as the Morningstar rating change. Similar to Panel A, the F-statistics indicate that conditional on the star rating change, changes in alpha cannot explain variation in abnormal flow. Although not explicitly controlled for in the benchmark, changes in the Sharpe measure also cannot explain variation in abnormal flow. The star rating change, in contrast, remains significant even after controlling for these contemporaneous changes in alpha and Sharpe. Thus, while all three performance measures are based on the same underlying set of returns data, and are therefore changing at the same time, these tests indicate that our benchmark has successfully purged any flow effect of a change in alpha or Sharpe measure. This gives us more confidence to interpret our measure of abnormal flow as attributable to an update in the packaging of the underlying performance information, the Morningstar rating, and net of the update in the underlying performance information itself.

#### *4.1.2 The dollar value of a Morningstar rating change: A comparison with the cross-sectional flow-performance literature*

While statistical significance indicates whether star ratings reliably influence investors, the economic significance determines whether Morningstar's influence is likely to alter the behavior of investment companies and fund managers. One obvious question of interest is how many millions of dollars attaining a 5-star rating is worth to the average fund. To answer such questions, we analyze non-standardized abnormal dollar flow, rather than the standardized version used in the statistical tests.

Figure 3 graphs the non-standardized average normal and average abnormal flows cumulated through the 6 event months following a rating change in each of the upgrade and downgrade categories. The pattern of normal flow in the figure suggests that the benchmark reliably captures the flow resulting from the *recent* performance that led to an upgrade or downgrade. For example, funds that were downgraded from 4- to 3-stars have a normal flow of \$1.4 million, while funds that were upgraded from 4- to 5-stars have a normal flow of \$131 million. Thus, normal flow is starkly different for these two sets of funds despite the fact that they were both at the same pre-event 4-star rating.

Many of the rating changes are economically significant events for the funds, resulting in flows that are much different than the expected flow. We find that an upgrade from 4- to 5-stars results in an average of \$32 million in abnormal flow, or 25% above average expected flow, over the 7 months following the upgrade. The average fund downgraded from 4- to 3-stars experiences outflow of \$16.7 million, nearly 12 times larger than the expected normal flow of only \$1.4 million. Rating changes also affect funds with negative normal flow, such as those at the low end of the performance spectrum. A downgrade from 3- to 2-stars, for example, increases the outflow in the average fund by 27%.

Prior studies have used cross-sectional panel regressions such as that in Table 1, to draw conclusions about the implicit incentives facing managers as they seek to accumulate more assets under management. However, the major unstated assumption is that the flow response that a typical fund could expect were *its* performance to change can be accurately inferred from the cross-section. In other words, it is assumed that the difference in average flow between a 4-star fund and a 5-star fund in the cross-section represents the expected change in flow for a performance improvement from a 4- to 5-star rating.

Using this logic and the estimated coefficients for 4- and 5-star dummies in Table 1 Panel B, an upgrade from 4- to 5-stars should result in an average of \$140 million of additional flow over the seven months following this rating change (the coefficients imply a \$20 million difference in average flow per month). Comparing this to our abnormal flow estimate for an upgrade to 5-stars of \$32

million indicates that relying on a traditional cross-sectional analysis grossly overestimates the average response to this performance change, and most certainly overestimates an isolated Morningstar effect.<sup>10</sup> In addition, because this specification of the cross-sectional analysis necessarily assumes a symmetric flow response for both upgrades and downgrades, it also makes the analogous prediction that a downgrade from 5- to 4-stars should result in a \$140 million decrease in flow on average. In contrast, our method allows for separate estimation of responses to upgrades and downgrades, revealing that losing a 5-star rating has a much milder effect on fund flows than the cross-section would predict.

Indeed, Figure 3 echoes our finding in Table 2 that funds downgraded to 4-stars have positive average abnormal flow in the following 7 months. Consistent with the nonparametric results in that table, we find that the median fund falling to 4-stars exhibits non-standardized abnormal flows approximately equal to zero. The sum of our evidence suggests that a fall from 5- to 4-stars is a non-event in the sense that it does not generate economically significant negative abnormal flow.

Finally, our results offer an alternative definition of investor ‘punishment’ for poor performance (or lack thereof). While the literature has noted that the total outflows of poor performers are quite small, we argue that the abnormal flow response to a rating change might provide a better measure of punishment since it captures the incremental loss of flow due to the performance decline. Figure 3 shows that 5- to 4-star downgrades are the one exception regarding evidence of ‘punishment’. We find that the negative abnormal flow following rating downgrades is economically meaningful, especially if viewed as a percentage of normal flow. Most notable is the large negative \$16.7 million flow response to a downgrade from 4- to 3-stars, which is almost 12 times the expected normal flow of \$1.4 million. Thus, funds that would have experienced net inflows if they remained at 4-stars instead

---

<sup>10</sup> While we argue in section 3 that the coefficients from a panel regression will be downwardly biased in the presence of fund heterogeneity, we were referring to a specification that includes rating *change* dummies. Here we are comparing our methodology to the specifications in the previous literature that typically include only performance measures in levels, and therefore we include only Morningstar rating dummies. A cross-sectional specification with rating change dummies would allow for an asymmetric response to upgrades and downgrades. However, the event study method allows for both a better estimate of the mean effects of rating changes as well as a measure of abnormal flow for each fund.

experience large net outflows due to the drop in performance. Overall, these comparisons illustrate that the traditional cross-sectional methodology yields incorrect or incomplete inferences about the flow likely to result from a performance change in a given fund.

#### *4.2 Abnormal flow response to initial Morningstar star ratings*

An interesting special case of the more frequent occurrence of a rating change is a rating initiation, whereby Morningstar rates a fund for the very first time. This event is perhaps the most straightforward to study because all funds share the common status of being unrated, and in existence for exactly 36 months. However, because of their relatively short performance history, initiations present difficulties in terms of fitting the same normal flow benchmark that we use for rating changes. Specifically, meaningful Jensen's alphas, and changes in alpha, do not exist for these funds in the 2 years prior to their initiation. We therefore exclude the change in alpha terms from the benchmark, but add a squared term on lagged return to capture potential convexities. Thus, the benchmark contains style flow, lagged own flow, lagged return, and lagged return squared as regressors. We use the same 24-month estimation period as we do for changes, but note that the results are qualitatively similar, but stronger statistically when using the 12-months prior to initiation as the estimation period.

Table 4, panel A reports the average standardized abnormal flow (ASTAF) for each event month from 0 to 6, in five initial Morningstar star rating subsamples, while panel B contains the corresponding average cumulative standardized abnormal flow (ACSTAF) estimates. It is clear from Table 4 that a 5-star initiation is an unambiguously positive event. The ASTAF in Panel A are significantly positive at the 5% level in four of the seven months following a 5-star initiation. In addition, the percentage of positive standardized abnormal flow exceeds 50% for most event dates for 5-star funds, significantly so in event month 1 (66%).

It is much less clear that a rating initiation is a significant event in the other rating categories. The ASTAF and ACSTAF are almost never significant for the 1-, 2-, and 3-star initiations. Even

though the 4-star initiations show significant average abnormal flow, the non-parametric ( $\chi^2$ ) tests are insignificant. The cumulative response in panel B appears to be more strongly positive for 4-star initiations than the month-by-month response revealed in Panel A. ACSTAF is positive and significant in all event months 0 to 6, consistent with funds receiving a modest boost in flow upon receiving an initial 4-star rating.

Figure 4 graphs the non-standardized average normal and abnormal flows cumulated over the 7 months following an initial star rating in each of the five star categories. The pattern of abnormal flows largely mirrors the standardized abnormal flow results of Table 4, but now we can interpret our findings in units of millions of dollars. Table 4 revealed a consistently positive and significant abnormal flow response to an initial 5-star rating, and Figure 4 indicates that the magnitude of this response is large in both absolute and relative terms. A 5-star initiation results in an average of \$19.9 million in abnormal flow over the following 7 months, which represents an impressive 34% increase above average normal (expected) flow. Overall, we conclude that there is mild evidence that 4-star initiations are significant events, that 1-, 2-, and 3-star initial ratings have no discernible impact on average fund flows, and that 5-star initiations stand out as experiencing a Morningstar effect.

#### *4.2.1 The initial star rating effect is not driven by initial Jensen's alphas or Sharpe ratios*

Rating initiations are unique in that they represent an opportunity to study the simultaneous release of three completely new pieces of fund quality information to investors. Morningstar provides an initial one-factor Jensen's alpha (versus the S&P500), an initial Sharpe ratio, and an initial star rating upon a fund's 36-month anniversary. Table 5 reports the results of two-way analysis of variance tests of the power of these three performance measures to explain variation in seven-month CSTAF.

As in Table 3, we examine three alternative definitions of alpha and Sharpe groups: quartiles, quintiles, and a "Morningstar-type" grouping. The Type II sum of squares tests indicate that

conditional on the initial Morningstar star rating, none of the alpha or Sharpe categorizations has the power to explain cross-sectional variation in abnormal flows. Thus, at least among these three plausible groupings, neither the initial alpha nor the initial Sharpe measure appears to influence investors. At the same time, Morningstar's star rating remains significant conditional on any of the alpha or Sharpe categorizations. Therefore, the results of Table 4 appear to be robust and not driven by the simultaneous release of alpha and Sharpe measure information.

#### *4.3. The determinants of cumulative standardized abnormal flow*

The results of tables 2 and 4 suggest that a change in Morningstar rating results in a significant flow response, especially for funds able to achieve 5-star status. This section provides confirmatory evidence by showing that the variation in the magnitude of the Morningstar effect across different types of rating changes is consistent with our earlier findings, after explicitly controlling for contemporaneous event-window fund performance. We also test whether the Morningstar effect, as measured by seven-month cumulative standardized abnormal flows (CSTAF6), is systematically related to fund-specific characteristics typically studied in the cross-sectional flow literature, such as fund age or size.

Using the rating change sample of Table 2, Table 6 contains the results of a multivariate regression of CSTAF6 on star rating change dummies, event window performance measures, and fund-specific characteristics. The included performance measures represent new information released during the event window: the change in alpha, the change in Sharpe ratio, and the compounded return, all from time 0 to 5. We report regression results for a screened sample of 2441 observations that include only funds that did not have another rating change during the seven-month event window. We note that our inferences are the same under the unreported full sample. The advantage of the screened sample is that it provides a cleaner measure of the flow response to a rating change, free from the effect of potentially confounding events.

Except for funds downgraded from 2 to 1-star, the coefficients on dummy variables for rating upgrades and downgrades are significantly different from zero in the expected direction, positive for upgrades and negative for downgrades. Thus, controlling for other performance measures, we find consistent responses to rating changes across the sample. The one clear exception to the general pattern is the funds downgraded from 5- to 4-stars. Consistent with the results in Table 2, the positive and significant coefficient indicates that these downgraded funds continue to experience positive abnormal flow, and perhaps are able to ‘ride’ their previous 5-star rating for several months.

The change in alpha and Sharpe measure during the event window are insignificant for explaining the cross-section of abnormal flows, complementing our earlier tests, which showed that performance changes occurring contemporaneously with the Morningstar rating change are not significantly related to abnormal flow. The evidence on the effect of return in the event window is mixed, the coefficient is significant in one regression, insignificant in another, and relatively small in magnitude.

We find two fund-specific measures to be systematically related to abnormal flow, fund size and the number of funds in the same Morningstar style and star-rating category. The coefficient on fund size is negative, indicating that larger funds have lower average abnormal flows. This may be somewhat surprising given the positive relation between total flow and fund size documented in previous papers. Whether it is reputation or another factor that drives the higher total flow for larger funds, these funds do not seem to be able to leverage it into a magnified response following a performance change.

The number of funds in the same Morningstar style and star-rating category is a proxy for the amount of competition a fund faces under the assumption that investors tend to choose a particular fund after narrowing their search to a certain style and star rating category (e.g., 5-star, small-cap growth funds). We find a significant negative coefficient on this variable, implying that when close-

substitute funds are more plentiful, the abnormal flow to the typical fund is lower. This finding also supports the idea that investors use style category information in their allocation decisions.

We do not find a systematic relation between abnormal flow and fund-specific characteristics such as the load status of the fund, the expense ratio, 12b-1 expenses, or fund age. Note that these results are not inconsistent with the previous literature. We find that these characteristics do not explain variation in *abnormal* flow following a performance change, which does not mean that they are unrelated to total net flow. To check for a seasonal effect around the turn of the year when, anecdotally, most investors evaluate their investments, we include a year-end dummy equal to one if the rating event occurred in December or January. The coefficient on this variable is not significant, suggesting that the tournament among mutual fund managers is more likely to be ongoing than confined to year-end.

#### *4.4 Interpretations of abnormal flow patterns and Morningstar's influence on investor decision-making and fund manager behavior*

In our view, Morningstar's influence is not surprising given the complicated information gathering and decision problem individual investors face when choosing among the thousands of available mutual funds. In addition to being easier for the typical mutual fund investor to understand than the more academic risk-adjusted measures, the information packed into a Morningstar rating plausibly reduces search costs substantially. The 1- to 5-star system provides a quick source of information on the full distribution of funds in that a 5-star fund is known to be in the top 10% of the universe of funds. Thus, Morningstar ratings provide a potentially valuable tool to individual investors with either little expertise or limited time to devote to their allocation decision.<sup>11</sup>

---

<sup>11</sup> Survey evidence suggests that mutual fund investors do not possess a great deal of financial expertise. For example, only 14% of surveyed fund investors in 1996 report using quantitative measures such as beta or alpha to identify risks. (*Shareholder Assessment of Risk Disclosure Methods*, Investment Company Institute).

Morningstar also plausibly influences the allocation decisions of investors consulting financial advisors, which a recent survey suggests is the majority of fund investors.<sup>12</sup> In this segment, Morningstar's influence may result from the agency relationship between advisor and client. Financial advisors are charged with recommending funds in their clients' best interests. The external validation a Morningstar rating provides could be useful to advisors should the fund they recommend perform poorly in the future. In this situation, an advisor could easily justify their fund selection advice ex-post by stating that the recommended fund had a top Morningstar rating. A similar situation occurs in the 401(k) and defined contribution plan market. Trustees responsible for choosing the menu of funds for plan participants might favor 5-star funds out of a desire to fulfill their fiduciary duty by offering participants the "best available funds." These effects likely contribute to the disproportionately large average abnormal flows we find for both 5-star initial ratings and upgrades from 4- to 5-stars.

Even though we generally observe a negative response to rating downgrades, we distinctly find no such response to a downgrade from 5- to 4-stars, and suspect that these same institutional features underlie this result. Perhaps the 5-star funds added to 401(k) plans or advisor "recommended lists" are only removed if performance lags either more dramatically (e.g., a fall to 3-stars), or more permanently. This conjecture is also supported by the large negative flow response after a downgrade from 4- to 3-stars. It is also possible that fund advertising aggressively promotes the perception that 4-star funds are also high quality. Of the 86 distinct funds advertising in June 1999, 36 (42%) are 4-star funds and 32 (37%) are 5-star funds, implying that a larger percentage of the advertised funds are 4-stars than are 5-stars.<sup>13</sup> Jain and Wu (2000) find that mutual funds advertised in *Barron's* or *Money* magazine have significantly higher flows than a control sample after controlling for fund performance. Alternatively, the tax lock-in effect discussed in Bergstresser and Poterba (2002) might explain the

---

<sup>12</sup> According to a 1997 survey, 59% of mutual fund investors consult a financial advisor when choosing a fund. In addition, 74% of these investors choose funds from a short list supplied by the advisor, and another 13% give the advisor total control over the choice of funds. (*Understanding Shareholders use of Information and Advisors*, Investment Company Institute)

<sup>13</sup> To investigate this we sampled all mutual fund ads featuring domestic equity funds appearing in *Money*, *Barron's* and the *Wall Street Journal's* Money & Investing Section during the month of June 1999.

lack of response exclusive to 5- to 4-downgrades, as investors in these funds might be reluctant to realize the capital gains of these recent good performers.

Our time-series event-study also provides new evidence on the timeliness of investor response to fund quality information. In most cases the flow response to a rating change is significant beginning in event month 0, indicating that investor response is detectable at the same monthly frequency that Morningstar releases new ratings. This implies that some investors vigilantly monitor rating information and, given that our method accounts for other influences on fund flow, view the rating change as “new” information on fund quality. While somewhat surprising, the quick response by some investors is consistent with anecdotal evidence. We have learned from an official at Morningstar that traffic on their website spikes on the days surrounding the release of new ratings. Consistent with this apparent investor interest, MSN’s Moneycentral website allows investors to screen on funds that Morningstar upgraded since last month, and Morningstar rating changes are listed as ‘news alert’ items under each fund’s page on their website. Our results show that this vigilant investor group is indeed large enough to measurably affect fund flow.

Our results also imply the existence of a substantial group of investors who respond to rating changes with a considerable lag. We report that abnormal flow is significant out to month 6, and confirm in unreported month-by-month results that this holds for the screened sample of funds that do not experience any further rating changes. In other words, even in the subsample of funds without confounding events, we generally observe a significant flow response out to month 6. This slow response might be driven in part by investors that rely on the somewhat stale information in fund advertising or magazine summary lists. For example, of 37 mutual fund advertisements touting star ratings in the month of June 1999, 12 featured star ratings as of April 1999, 20 as of March, one as of February, and only 4 reflected ratings as of May 1999.<sup>14</sup> In addition, one of Morningstar’s popular

---

<sup>14</sup>The June issues of Barron’s, Money Magazine or the of the Wall Street Journal contained 187 mutual fund advertisements, 43 of which featured Morningstar star ratings. Of these, 37 are in our sample of domestic equity funds.

products is a quarterly subscription to its CD-ROM database. Subscribers relying on this source would necessarily get the rating change with a lag. Thus news of a rating change, or rating status, may take several months to reach investors.

Taken together, our results can inform the design of more powerful tests of managerial incentives and risk-shifting behavior. Most tests of risk-shifting behavior are based on the empirical flow-performance relation found in a traditional cross-sectional analysis of annual flow. Given the differences in results we find from analyzing monthly time-series of flow and performance, the mixed results of the risk-shifting literature are perhaps not surprising.<sup>15</sup> We suspect that the substantial heterogeneity in the time-series relation between flow and performance across individual funds is partly responsible for the mixed evidence. We find that within the cross-section of funds, there are a meaningful number serving clientele who are relatively insensitive to changes in performance. Including these funds in tests of risk-shifting clearly adds noise, as such managers have little incentive to risk-shift. By estimating a flow-performance sensitivity for each fund, our time-series approach suggests a simple, but more powerful method for identifying subsamples of managers with stronger implicit incentives.

The features of the Morningstar rating system that allow us to measure its influence apart from other performance metrics are the same features that make it a powerful incentive for fund managers.<sup>16</sup> Our findings suggest specific points in the performance distribution where risk-shifting incentives are strongest. Rather than separating funds by performance quartile, or relative to the median fund, as in Brown, Harlow, and Starks (1996), our findings indicate that risk-shifting should be more prevalent along the boundaries between star ratings. In particular, the finding that upgrades to 5-stars are

---

<sup>15</sup> Brown, Harlow and Starks (1996) and Chevalier and Ellison (1997) find evidence that mutual fund managers respond to the implicit incentives in the convex relation between flow and performance and systematically alter the riskiness of their portfolios during the last part of the year. Busse (2001) and Koski and Pontiff (1999) argue that the empirical relation between performance and risk are driven by methodology or mechanically by flows, and is not the result of incentives. Chen and Pennachi (1999) provide an alternative test based on tracking error.

<sup>16</sup> The more dominant a rating system is among the other information inputs decision-makers use, and the easier the rating's effect is to measure, the greater the incentive for suppliers (e.g., fund managers) to focus their energies on improving their rating. Undesirable supply-side responses on the part of university administrators deriving from the influence of *U.S. News and World Report's* rankings in the college application market are discussed in Ehrenberg (2002), Ehrenberg and Monks (1999), and the *Wall Street Journal* 4/5/95 page A1).

handsomely rewarded with positive abnormal flow, but downgrades to 4-stars often do not elicit an outflow response suggests that the relation between flow and performance is especially convex at the boundary between 4- and 5-stars. Managers of 4-star funds that are close to the 5-star cutoff might have the greatest incentive to increase the risk of their portfolio, while managers attaining 5-star status have little incentive to reduce their risk. The large penalty that we find for dropping to 3-stars also suggests that managers of 4-star funds close to this precarious cutoff have a strong incentive to guard their position and reduce risk.

Our evidence suggests that some investors respond immediately to *monthly* rating changes, and we detect no differences in this response at year-end. Thus, the tournament for investor flow more likely operates on an ongoing basis, rather than at the annual frequencies assumed in previous studies. Thus, more work in the theoretical modeling of managerial behavior that better captures the dynamic and repeated game aspects of the tournament may be fruitful. A theoretical model of managerial incentives in such an environment could point the way toward new empirical tests of strategic managerial behavior.

## **5. Conclusion**

In the market for colleges, cars, and mutual funds, the choice set is large and the list of relevant attributes that translate into quality are difficult for the typical decision-maker to assess. Information intermediaries, such as *U.S. News and World Report* and Morningstar, can add value in this decision-making environment by providing low-cost summary product quality information from a reputable and unbiased source. Using discrete informational events occurring at regular monthly intervals, we provide evidence that Morningstar's re-packaging of fund quality information in the form of a simple one- to five-star rating has a large independent influence on mutual fund flow, and therefore on investor allocation decisions.

While it may at first seem surprising that Morningstar's re-packaging of publicly available information can rationally influence investors, this result is consistent with what we know about fund investor decisions from other studies and contexts. For example, Elton, Gruber, and Busse (2003) find a surprising amount of dispersion in prices (expense ratios) in the market for homogeneous S&P500 index funds, and document that seemingly dominated products appear to prosper in this market. In the 401(k) context, Madrian and Shea (2001) find that employees exhibit much higher plan enrollment rates when the default option is enrollment (opt-out plans), than when the default is non-enrollment (opt-in plans). In both cases, investor behavior is attributed to irrational decision-making. Our interpretation of this evidence, however, is that making a financial decision is an especially costly undertaking for many mutual fund and 401(k) investors. For example, in the Madrian and Shea case, accepting the default option is equivalent to choosing not to incur the high cost of making a financial decision. Similarly, investors may choose not to incur the search cost of finding a lower-priced index fund, but instead purchase the one their advisor recommends.

Current research along these lines is Hortascu and Syverson (2003), who formally model the distribution of search costs across investors in the S&P500 index fund market. They find that reasonable magnitudes of search costs can explain the observed price dispersion, and that their empirical estimates of search costs increase over the sample period, a result that they attribute to the recent entry of mutual fund investors with less expertise and higher decision costs. Unlike a study of security prices where the arbitrage behavior of the marginal, lowest decision-cost investor matters, studies of fund flow or fund investor allocation decisions reflect the decision and search costs of the average investor. We believe it plausible that the typical mutual fund investor's search and decision costs are indeed high enough to induce these empirical regularities, but further evidence on this issue awaits future research.

## References

- Bergstresser, Daniel and James Poterba, 2002, Do After-Tax Returns Affect Mutual Fund Inflows?, *Journal of Financial Economics* 63, 381-414.
- Blume, Marshall, 1998, An Anatomy of Morningstar Ratings, *Financial Analysts Journal* vol. 54, 1-9.
- Blake, Christopher R., and Matthew R. Morey, 2000, Morningstar Ratings and Mutual Fund Performance, *Journal of Financial and Quantitative Analysis* 35, 451-483.
- Boehmer, Ekkehart, Jim Musumeci, and Annette Poulsen, 1991, Event-study Methodology under Conditions of Event-induced Variance, *Journal of Financial Economics* 30, 253-272.
- Brown, Keith, W.V. Harlow, and Laura Starks, 1996, Of Tournaments and Temptations: An Analysis of Managerial Incentives in the Mutual Fund Industry, *Journal of Finance* 51, 85-110.
- Brown, Stephen, and William N. Goetzmann, 1995, Performance Persistence, *Journal of Finance* 51, 85-110.
- Busse, Jeffrey, 2001, Another look at mutual fund tournaments, *Journal of Financial and Quantitative Analysis*, 36, 53-73.
- Carhart, Mark, 1997, On Persistence in Mutual Fund Performance, *Journal of Finance* 52, 57-82.
- Chevalier, Judith and Glenn Ellison, 1997, Risk Taking by Mutual Funds as a Response to Incentives, *Journal of Political Economy* 105, 1167-1200.
- Chen, Hsiu-Lang, and George Pennacchi, 1999, Does prior performance affect a mutual fund's choice of risk? Theory and Further Empirical Evidence, working paper University of Illinois at Urbana-Champaign.
- Del Guercio, Diane and Paula A. Tkac, 2002, The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds versus Pension Funds, *Journal of Financial and Quantitative Analysis* 37, 523-557.
- Dodd, Peter and Jerold B. Warner, 1983, On Corporate Governance: A Study of Proxy Contests, *Journal of Financial Economics* 11, 401-438.
- Ehrenberg, Ronald G., and James Monks, 1999, The Impact of *US News and World Report* College Rankings on Admission Outcomes and Pricing Decisions at Selective Private Institutions, NBER working paper 7227.
- Ehrenberg, Ronald G., 2002, Reaching for the Brass Ring: The *US News and World Report* Rankings and Competition, *The Review of Higher Education* 26, 145-162.
- Elton, Edwin, Martin Gruber, and Jeffrey Busse, 2003, Are Investors Rational? Choices Among Index Funds, *Journal of Finance* forthcoming.

- Fama, Eugene and Kenneth French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics* 33, 3-56.
- Gruber, Martin, 1996, Another Puzzle: The Growth in Actively Managed Mutual Funds, *Journal of Finance* 51, 783-810.
- Goetzmann, William and Massimo Massa, 2002, Daily Momentum and Contrarian Behavior of Index Fund Investors, *Journal of Financial and Quantitative Analysis* 37, 375-389.
- Hortacsu, Ali, and Chad Syverson, 2003, Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds, working paper University of Chicago.
- Ippolito, Richard, 1992, Consumer reaction to measures of poor quality, *Journal of Law and Economics* 35, 45-70.
- Jain, Prem C. and Joanna Shuang Wu, 2000, Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows, *Journal of Finance* 55, 937-958.
- Koski, Jennifer and Jeffrey Pontiff, 1999, How are derivatives used? Evidence from the mutual fund industry, *Journal of Finance* 54, 791-816.
- Madrian, Brigitte and Dennis Shea, 2001, The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior, *Quarterly Journal of Economics* 116, 1149-1187.
- Morey, Matthew R., 2002, Mutual Fund Age and Morningstar Ratings, *Financial Analysts Journal* 58,
- Patel, J., R. Zeckhauser, and D. Hendricks, 1994, Investment flows and performance: Evidence from mutual funds, cross-border investments and new issues, in R. Satl, R. Levitch, and R. Ramachandran, Eds.: *Japan, Europe and the International Financial Markets: Analytical and Empirical Perspectives* (Cambridge University Press, New York).
- Pesaran, J., R., Ron P. Smith, and K.S. Im, 1996, Dynamic Linear Models for Heterogeneous Panels, in L. Matyas and P. Sevestre, Eds.: *The Econometrics of Panel Data* (Kluwer Academic Publishers, Dordrecht, The Netherlands).
- Pozen, Robert C., 1998, *The Mutual Fund Business*, MIT Press, Cambridge, MA.
- Sharpe, William, 1997, Morningstar's Performance Measures, unpublished manuscript (Stanford University).
- Sharpe, William, 1998, Morningstar's Risk-Adjusted Ratings, *Financial Analysts Journal* vol. 54, 21-33.
- Sirri, Erik R. and Peter Tufano, 1998, Costly Search and Mutual Fund Flows, *Journal of Finance* 53, 1589-1622.

Walter, Ingo, 1999, The Asset Management Industry in Europe: Competitive Structure and Performance, BSA Gamma Foundation Conference on the Future of Asset Management, Lugano, Switzerland.

## Appendix

### Morningstar Style Categories and the Construction of Style Benchmarks

We use Morningstar's style category as a style identifier for sample funds. Morningstar divides the domestic equity star rating category peer group into nineteen mutually exclusive categories. We list the nineteen categories in the table below. For each fund in our sample, we assign it to the category Morningstar places it in as of October 1999. (For dead funds we use the Morningstar category reported in the month before its disappearance). This procedure should result in an accurate categorization since according to Morningstar's definition, the style category assignment is based on the fund's average investment style over the previous three years. At each point in time, Morningstar assigns a fund to a style category based on the style and size of the stocks the fund owns.

The style flow we use in our benchmarking procedure is an in-sample measure, using all funds in the sample within the same Morningstar category. Thus, large-cap growth funds are benchmarked using their sensitivity to the overall flow to large-cap growth funds. Because our sample is constructed to include all domestic equity funds that were rated at any time during the sample period, the only funds missing from the aggregate style flow measure are funds coming into existence after November 1996 (and thus not rated by Morningstar throughout our sample period). Below we report the number of funds in each category that we use to compute aggregate style flows.

Morningstar category	Number of funds
Large-cap growth	376
Large-cap blend	552
Large-cap value	406
Medium-cap growth	279
Medium-cap blend	168
Medium-cap value	147
Small-cap growth	246
Small-cap blend	132
Small-cap value	123
Specialty-Health	30
Specialty-Technology	51
Specialty-Utilities	83
Specialty-Communications	16
Specialty-Financials	25
Specialty-Real Estate	59
Specialty-Natural Resources	43
Specialty-Precious Metals	40
Convertible Bond	47
Domestic Hybrid	564

**Table 1. Descriptive statistics****Panel A. Summary Statistics on Mutual Fund Flow by Morningstar Star Rating Category**

This panel contains the total dollar flow to mutual funds in six Morningstar star rating categories. The sample includes all funds in Morningstar's domestic equity star rating category over the period November 1996 through October 1999. In addition to the diversified equity funds that Morningstar assigns to one of nine equity styles (e.g. large-cap growth), this category also includes sector funds, convertible bond funds, and domestic hybrid funds (e.g., asset allocation funds). The sample contains flow data for a total of 111,715 fund-months and includes data for 3,388 unique funds. For these aggregate flow calculations, and in our empirical tests, we delete 220 fund-months that coincide with a fund merger. Fund-months appear in the not rated category if they are less than three years old during the month we measure flows, and therefore not yet rated by Morningstar. We only include unrated fund-months if the fund receives a rating by October 1999, the end of our sample period.

Morningstar star rating	Number of Fund-months	Total Flow (\$ Millions)	Percent of Total Flow to Category
1	8,564	-20,497	-7.2
2	19,403	-57,106	-20.3
3	29,849	-55,714	-19.8
4	19,233	111,126	39.6
5	7,285	238,141	84.8
Not rated	27,381	64,998	23.1
Total	111,715	280,949	100

**Table 1 Panel B. Pooled Panel Regression of Monthly Dollar Fund Flow on Performance and Fund Characteristics**

This panel reports a panel regression pooled over all fund-months from November 1996 through October 1999. The intercept has been omitted in this specification. The Jensen's alpha is calculated using returns over the previous 36 months, with the S&P500 as a proxy for the market. Fund age is measured in years. Funds are classified as having a value style if their Morningstar category as listed in the Appendix is of the value style (e.g., Large-cap value). Funds are classified as having a growth style if their Morningstar category as listed in the Appendix is of the growth style (e.g., Large-cap growth). T-statistics are in parentheses. The symbols \*, \*\*, \*\*\* indicate statistical significance at the 10, 5, and 1% level.

One star dummy	0.25 (0.25)
Two star dummy	-2.68 <sup>***</sup> (-3.98)
Three star dummy	-4.12 <sup>***</sup> (-7.99)
Four star dummy	-1.01 <sup>*</sup> (-1.74)
Five star dummy	19.31 <sup>***</sup> (22.32)
Jensen's alpha (t = -1)	912.15 <sup>***</sup> (12.73)
Jensen's alpha squared	19897 <sup>***</sup> (8.06)
Lagged return (t = -1)	15.36 <sup>***</sup> (3.92)
Lagged return squared	106.27 <sup>***</sup> (3.60)
Lagged flow (t = -1)	0.09 <sup>***</sup> (25.87)
Log of total net assets (t = -1)	0.003 <sup>**</sup> (40.32)
Fund age	-0.002 (-1.40)
No-load fund dummy	1.58 <sup>***</sup> (3.51)
Growth style dummy	4.13 <sup>***</sup> (7.29)
Value style dummy	2.86 <sup>***</sup> (5.26)
Adjusted R-squared	0.061
N	80879

**Table 1. Descriptive statistics (continued)****Panel C. Frequency of Fund-Months with Changes in Morningstar Star Rating**

This panel reports the frequency count of fund-months where the fund's Morningstar star rating is different from its previous month's star rating over the period November 1996 to October 1999. These star rating changes are divided into one-star rating upgrades (i.e., a change from one to two stars), and one-star rating downgrades. We also report the frequency of rating changes greater than one-star (e.g., a change from one to three stars), but we do not analyze these.

Morningstar star rating after change	One-star upgrade (N)	One-star downgrade (N)
1	N/A	739
2	653	1,678
3	1,588	2,065
4	2,053	931
5	944	N/A
Subtotal	5,238	5,413
Total number of one-star changes	10,651	
Upgrades greater than one-star	40	
Downgrades greater than one-star	44	
Total star rating changes	10,735	

**Panel D. Frequency of Funds Receiving an Initial Star Rating from Morningstar**

This panel reports the frequency count of funds receiving a Morningstar star rating for the very first time over the period November 1996 to October 1999. Receiving an initial rating coincides with the fund turning three-years old, and therefore having a 36-month return history.

Initial Morningstar star rating	Number of Fund-months	Percent of Total Initial Ratings
1	161	9.8
2	394	24.1
3	588	35.9
4	317	19.4
5	177	10.8
Total	1,637	100

**Table 1. Descriptive statistics (continued)**

**Panel E. Estimation period Statistics of the Event-study Methodology**

This panel reports diagnostic statistics on the event study benchmark model used to calculate abnormal flow for star rating change events. The benchmark model, analogous to a market model regression, regresses a fund's monthly flow on aggregate flow at time t to funds in its same style group, its own time t-1 flow, and its own time t-1 return, its change in Jensen's alpha from t-2 to t-1, and its change in Jensen's alpha from t-2 to t-1 squared. We separately fit this benchmark for each of the 10,651 rating change events, and define the estimation period as months (-26,-3). The pairwise cross-sectional correlations reported in the table are computed by pooling observations over all funds and months in the estimation period. The symbols \*, \*\*, and \*\*\* indicate statistical significance at the 10, 5, and 1% levels.

Estimated Benchmark Coefficients			25 <sup>th</sup> percentile	Median	75 <sup>th</sup> percentile			
R <sup>2</sup>			0.25	0.43	0.66			
Intercept			-1.89	-0.06	0.095			
Style flow			-0.00002	0.0004	0.003			
Lagged return			-12.47	0.00	12.19			
Lagged flow			-0.18	0.03	0.34			
Change in alpha			-5.73	0.21	10.56			
Change in alpha squared			-51.32	0.36	54.24			
# events per month			36	43	58			
Pairwise Correlations			Style flow	Lagged return	Lagged flow	Change in alpha	Change in alpha squared	Fund size
Standardized residuals			-0.000	0.0002	0.0004	0.0001	-0.0002	0.003
Normal flow			0.096***	0.020***	0.030***	0.033***	0.004***	0.042***

**Table 2. Morningstar Star Rating Changes of Domestic Equity Funds**

**Panel A. Average Standardized Abnormal Flow for Rating Upgrades**

Average standardized abnormal flow (ASTAF<sub>t</sub>) reported below is averaged across domestic equity funds within the same star-rating change group. We define standardized abnormal flow in month t as the actual dollar flow in month t minus the normal, or expected, flow standardized by the estimated forecast variance of the normal flow. Normal flow is based on a market-model regression whereby a fund's monthly flow is regressed on aggregate flow at time t to funds in its same style group, its time t-1 flow, its time t-1 return, its change in Jensen's alpha from t-2 to t-1, and its change in Jensen's alpha from t-2 to t-1 squared. We use nineteen Morningstar-defined style categories to compute aggregate style flows. See the appendix for details on the style categories we use. The estimation period for computing the market-model parameters is months (-26,-3). The standard error is equal to the cross-sectional standard deviation over N sample events of the standardized abnormal flows, divided by the square root of N. Average standardized abnormal flows significantly different from zero at the 5% level or higher in a two-tailed test are indicated in bold. For each event date, we also report the percentage of sample events with positive abnormal flows. We indicate in bold the percentage positive that differs from 50% at the 5% significance level or better, using a chi-square test with one degree of freedom. The number of observations differs from Panel D of Table 1 because we delete fund-months that coincide with mergers, observations that do not have the data necessary to compute flows, and the top and bottom 1% of standardized abnormal flows at each event date.

Event-month	From 1 to 2-stars (n=534)			From 2 to 3-stars (n=1382)			From 3 to 4-stars (n=1763)			From 4 to 5-stars (n=805)		
	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0
0	0.01	0.15	51.7	<b>0.22</b>	3.46	<b>54.5</b>	<b>0.21</b>	3.24	49.9	<b>0.56</b>	5.40	<b>55.8</b>
1	0.02	0.25	52.4	<b>0.22</b>	2.94	51.4	<b>0.19</b>	3.12	51.8	<b>0.76</b>	6.56	<b>57.8</b>
2	0.01	0.08	49.4	<b>0.24</b>	3.58	<b>55.5</b>	<b>0.35</b>	4.80	52.3	<b>0.99</b>	6.59	<b>57.3</b>
3	0.06	0.62	52.6	<b>0.35</b>	4.41	<b>55.7</b>	<b>0.54</b>	6.78	<b>52.6</b>	<b>0.90</b>	5.33	<b>55.1</b>
4	<b>0.25</b>	2.44	<b>55.5</b>	<b>0.24</b>	3.01	<b>54.1</b>	<b>0.55</b>	6.39	<b>52.8</b>	<b>1.04</b>	5.84	<b>55.3</b>
5	0.19	1.33	52.4	<b>0.52</b>	4.98	<b>57.8</b>	<b>0.41</b>	4.46	<b>53.8</b>	<b>1.00</b>	5.16	<b>58.1</b>
6	<b>0.26</b>	2.25	54.0	<b>0.29</b>	2.70	<b>54.4</b>	<b>0.45</b>	4.43	50.5	<b>1.44</b>	5.57	52.7

**Panel B. Average Standardized Abnormal Flow for Rating Downgrades**

Event-month	From 2 to 1-star (n=589)			From 3 to 2-stars (n=1407)			From 4 to 3-stars (n=1718)			From 5 to 4-stars (n=770)		
	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0
0	-0.08	-1.16	47.5	<b>-0.18</b>	-2.98	48.5	<b>-0.16</b>	-2.66	47.7	<b>0.46</b>	4.26	<b>53.9</b>
1	-0.15	-1.81	48.9	-0.08	-1.43	49.5	<b>-0.12</b>	-2.05	49.1	<b>0.60</b>	3.95	53.0
2	<b>-0.23</b>	-3.09	46.5	<b>-0.16</b>	-2.51	47.9	-0.09	-1.49	<b>46.1</b>	0.04	0.36	47.6
3	-0.10	-1.22	47.6	<b>-0.20</b>	-3.14	<b>46.7</b>	<b>-0.21</b>	-3.02	<b>47.0</b>	<b>0.26</b>	2.03	50.7
4	<b>-0.20</b>	-1.98	46.9	<b>-0.15</b>	-1.98	<b>47.1</b>	-0.03	-0.46	<b>46.9</b>	<b>0.82</b>	4.76	52.4
5	-0.17	-1.81	<b>44.8</b>	<b>-0.27</b>	-3.66	<b>44.8</b>	<b>-0.21</b>	-2.91	<b>45.7</b>	<b>0.67</b>	3.44	49.4
6	<b>-0.32</b>	-3.26	<b>44.4</b>	-0.18	-1.91	48.1	<b>-0.15</b>	-1.79	<b>45.7</b>	<b>0.85</b>	4.09	49.9

**Table 2. Morningstar Star Rating Changes of Domestic Equity Funds (continued)****Panel C. Average Cumulative Standardized Abnormal Flow for Rating Upgrades**

This panel reports the average cumulative standardized abnormal flow (ACSTAF<sub>t</sub>) for each date t in the event window. We first compute the cumulative standardized abnormal flow for each fund by summing the standardized abnormal flow from 0 to t, and then dividing by the square root of the number of months used in the cumulation. We then average these standardized cumulated abnormal flows over the N sample events in each star-rating change group to obtain ACSTAF at each event date t. ACSTAF significantly different from zero at the 5% level or higher in a two-tailed test are indicated in bold. For each event date, we also report the percentage of sample events with positive cumulative standardized abnormal flows. We indicate in bold the percentage positive that differs from 50% at the 5% significance level or better, using a chi-square test with one degree of freedom. The number of observations differs from Panel D of Table 1 because we delete fund-months that coincide with mergers, observations that do not have the data necessary to compute flows, and the top and bottom 1% of standardized abnormal flows at each event date.

Event-month	From 1 to 2-stars (n=534)			From 2 to 3-stars (n=1382)			From 3 to 4-stars (n=1763)			From 4 to 5-stars (n=805)		
	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0
0	0.01	0.15	51.7	<b>0.22</b>	3.46	<b>54.5</b>	<b>0.21</b>	3.24	49.9	<b>0.56</b>	5.40	<b>55.8</b>
1	0.02	0.19	48.7	0.25	1.39	<b>54.2</b>	<b>0.31</b>	2.54	51.7	<b>0.97</b>	7.08	<b>58.1</b>
2	-0.08	-0.59	49.4	<b>0.49</b>	3.97	<b>54.9</b>	<b>0.37</b>	2.86	52.0	<b>1.81</b>	3.90	<b>59.5</b>
3	0.08	0.41	48.7	<b>0.58</b>	4.49	<b>55.7</b>	<b>0.64</b>	3.12	<b>52.8</b>	<b>2.17</b>	4.65	<b>60.5</b>
4	0.14	0.88	51.1	<b>0.77</b>	4.34	<b>55.1</b>	<b>0.77</b>	4.22	<b>52.6</b>	<b>2.81</b>	5.01	<b>61.6</b>
5	0.16	0.87	50.4	<b>0.96</b>	4.92	<b>56.5</b>	<b>0.88</b>	4.47	<b>53.0</b>	<b>2.89</b>	5.16	<b>61.0</b>
6	0.39	1.93	51.5	<b>1.02</b>	4.84	<b>57.1</b>	<b>0.86</b>	4.00	<b>53.3</b>	<b>3.59</b>	5.78	<b>60.4</b>

**Panel D. Average Cumulative Standardized Abnormal Flow for Rating Downgrades**

Event-month	From 2 to 1-star (n=589)			From 3 to 2-stars (n=1407)			From 4 to 3-stars (n=1718)			From 5 to 4-stars (n=770)		
	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0
0	-0.08	-1.16	47.5	<b>-0.18</b>	-2.98	48.5	<b>-0.16</b>	-2.66	47.7	<b>0.46</b>	4.26	<b>53.9</b>
1	-0.14	-1.24	46.6	<b>-0.19</b>	-1.99	47.5	<b>-0.27</b>	-2.80	48.2	<b>0.79</b>	4.70	53.1
2	-0.09	-0.51	47.2	<b>-0.25</b>	-2.63	<b>46.7</b>	<b>-0.29</b>	-1.98	<b>46.4</b>	<b>0.56</b>	3.35	52.0
3	<b>-0.33</b>	-2.81	<b>44.2</b>	<b>-0.28</b>	-2.66	<b>46.0</b>	<b>-0.35</b>	-2.20	<b>46.7</b>	<b>2.00</b>	2.05	52.5
4	<b>-0.38</b>	-2.62	<b>45.3</b>	<b>-0.35</b>	-2.74	<b>46.0</b>	<b>-0.39</b>	-2.20	<b>46.8</b>	<b>1.84</b>	3.36	52.0
5	-0.29	-1.78	<b>42.6</b>	<b>-0.52</b>	-3.96	<b>44.3</b>	<b>-0.52</b>	-2.82	<b>45.2</b>	<b>1.78</b>	3.32	52.2
6	-2.12	-1.26	<b>43.1</b>	<b>-0.53</b>	-3.62	<b>44.6</b>	-0.34	-1.68	<b>44.8</b>	<b>2.99</b>	2.62	50.9

**Table 3. Determinants of the Six-Month Cumulative Standardized Abnormal Flow Response to a Change in Star Rating**  
**Panel A. Jensen’s Alpha and Sharpe ratio versus Morningstar Star Rating**

This table reports the results of a two-way analysis of variance of cumulative standardized abnormal flows from month 0 to 6 (CSTAF<sub>6</sub>) for the full sample of funds experiencing a Morningstar star rating change (N=7060). We compare the explanatory power of a fund’s Morningstar rating change at t=0 to either a fund’s Jensen’s alpha group or a fund’s Sharpe ratio group, measured one month prior to the Morningstar rating change. We present three ways of defining Jensen’s alpha or Sharpe ratio groups: equal-sized quartiles, equal-sized quintiles, and “Morningstar-type” quintiles where the top 10% of alphas (Sharpe ratios) are in quintile 5, the next 22.5% in quintile 4, the next 35% in quintile 3, the next 22.5% in quintile 2, and the bottom 10% in quintile 1. For each comparison below labeled (1) through (6), we report the F-statistic and the corresponding p-value for the test of the null hypothesis that the sample means are identical across groups in the one rating system, conditional on the other rating system. The F test has (n<sub>1</sub> – 1, n-p) degrees of freedom, where n<sub>1</sub> equals the number of groups in the rating system being tested for significance, n equals the number of observations, and p equals the number of groups in both rating systems. For example, the F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across Jensen’s alpha quartiles, holding constant the effect of Morningstar star rating changes, is 1.20 and the degrees of freedom for this test are (3, 7048). The F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across Morningstar star ratings changes, holding constant the effect of Jensen’s alpha quartiles, is 3.44, with (7,7048) degrees of freedom. The symbols \*, \*\*, \*\*\* indicate statistical significance at the 10, 5, and 1% level.

Categorizations		Type II Sum of Squares F-statistic	Prob > F
(1)	Jensen’s alpha quartiles	1.20	0.308
	Morningstar star rating change groups	3.44***	0.000
(2)	Jensen’s alpha quintiles	1.02	0.395
	Morningstar star rating change groups	10.38***	0.000
(3)	Jensen’s alpha “Morningstar-type” quintiles	0.97	0.202
	Morningstar star rating change groups	10.63***	0.000

Categorizations		Type II Sum of Squares F-statistic	Prob > F
(4)	Sharpe ratio quartiles	1.40	0.240
	Morningstar star rating change groups	10.27***	0.003
(5)	Sharpe ratio quintiles	1.20	0.307
	Morningstar star rating change groups	10.04***	0.000
(6)	Sharpe ratio “Morningstar-type” quintiles	1.73	0.141
	Morningstar star rating change groups	10.54***	0.000

**Table 3. Determinants of the Six-Month Cumulative Standardized Abnormal Flow Response to a Change in Star Rating**  
**Panel B. Contemporaneous Change in Jensen’s Alpha and Sharpe ratio versus Morningstar Star Rating**

This table reports the results of a two-way analysis of variance of cumulative standardized abnormal flows from month 0 to 6 (CSTAF<sub>6</sub>) for the full sample of funds experiencing a Morningstar star rating change (N=7060). We separately compare the explanatory power of a fund’s Morningstar rating change at t=0 to either a change in a fund’s Jensen’s alpha group or a change in a fund’s Sharpe ratio group, measured at the same time as the Morningstar rating change. We present three ways of defining Jensen’s alpha or Sharpe ratio groups: equal-sized quartiles, equal-sized quintiles, and “Morningstar quintiles” where the top 10% of alphas (Sharpe ratios) are in quintile 5, the next 22.5% in quintile 4, the next 35% in quintile 3, the next 22.5% in quintile 2, and the bottom 10% in quintile 1. For each comparison below labeled (1) through (6), we report the F-statistic and the corresponding p-value for the test of the null hypothesis that the sample means are identical across groups in the one rating system, conditional on the other rating system. The F test has (n<sub>1</sub> – 1, n-p) degrees of freedom, where n<sub>1</sub> equals the number of change groups in the rating system being tested for significance, n equals the number of observations, and p equals the number of change groups in the both rating systems. For example, the F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across groups representing changes in Jensen’s alpha quartiles, holding constant the effect of Morningstar star rating change, is 0.55 and the degrees of freedom for this test are (10, 7042). The F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across Morningstar star ratings, holding constant the effect of Jensen’s alpha quartiles, is 9.21, with (7,7042) degrees of freedom. The symbols \*, \*\*, \*\*\* indicate statistical significance at the 10, 5, and 1% level.

Categorizations		Type II Sum of Squares F-statistic	Prob > F
(1)	Jensen’s alpha change (quartiles)	0.55	0.840
	Morningstar star rating groups	9.21***	0.000
(2)	Jensen’s alpha change (quintiles)	0.70	0.754
	Morningstar star rating change groups	8.50***	0.000
(3)	Jensen’s alpha change (“Morningstar quintiles”)	0.83	0.617
	Morningstar star rating change groups	9.26***	0.000
Categorizations		Type II Sum of Squares F-statistic	Prob > F
(4)	Sharpe ratio change (quartiles)	0.71	0.696
	Morningstar star rating change groups	9.53***	0.000
(5)	Sharpe ratio change (quintiles)	0.70	0.752
	Morningstar star rating change groups	9.48***	0.000
(6)	Sharpe ratio change (“Morningstar quintiles”)	0.97	0.470
	Morningstar star rating change groups	10.22***	0.000

**Table 4. Initial Morningstar Star Ratings of Domestic Equity Funds****Panel A. Average Standardized Abnormal Flow**

Average standardized abnormal flow (ASTAF<sub>t</sub>) reported below is averaged across domestic equity funds within the same initial star-rating group. Morningstar assigns an initial star rating in the 36th month of a fund's existence. We define standardized abnormal flow in month t as the actual dollar flow in month t minus the normal, or expected, flow standardized by the estimated forecast variance of the normal flow. Normal flow is based on a market-model regression whereby a fund's monthly flow is regressed on aggregate flow at time t to funds in its same style group, its time t-1 flow, its time t-1 return, and its time t-1 return squared. The estimation period for computing the market-model parameters is months (-26,-3). The standard error is equal to the cross-sectional standard deviation over N sample events of the standardized abnormal flows, divided by the square root of N. ASTAF<sub>t</sub> significantly different from zero at the 5% level or higher in a two-tailed test are indicated in bold. For each event date, we also report the percentage of sample events with positive standardized abnormal flows. We indicate in bold the percentage positive that differs from 50% at the 5% significance level or better, using a chi-square test with one degree of freedom. The number of observations differs from Panel C of Table 1 because we delete fund-months that coincide with mergers, observations that do not have the data necessary to compute flows, and the top and bottom 1% of standardized abnormal flows at each event date.

Event-month	One-star (n=141)			Two-stars (n=344)			Three-stars (n=527)			Four-stars (n=277)			Five-stars (n=154)		
	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0	ASTAF <sub>t</sub>	t-stat	%>0
0	0.12	0.84	49.6	-0.15	-1.40	<b>41.6</b>	0.01	0.08	<b>45.2</b>	<b>0.42</b>	2.65	49.8	<b>0.54</b>	2.13	55.8
1	-0.16	-0.83	42.9	-0.16	-1.49	<b>44.0</b>	0.17	1.84	47.2	<b>0.38</b>	2.50	51.6	<b>1.38</b>	3.81	<b>66.2</b>
2	-0.06	-0.37	48.5	-0.09	-0.78	<b>42.2</b>	0.12	1.20	46.7	0.03	0.21	48.9	<b>0.79</b>	2.48	56.6
3	0.18	0.90	44.4	-0.08	-0.85	<b>41.5</b>	0.12	1.24	<b>44.8</b>	<b>0.37</b>	2.18	51.9	0.29	0.99	51.2
4	0.01	0.07	47.3	0.05	0.47	46.5	-0.07	-0.72	<b>44.7</b>	0.23	1.14	48.2	0.64	1.65	54.0
5	0.27	0.96	50.8	0.04	0.37	47.4	-0.03	-0.30	<b>41.4</b>	0.31	1.49	51.0	0.26	1.15	57.6
6	0.35	1.38	55.3	0.29	1.68	46.1	-0.16	-1.35	<b>44.7</b>	<b>0.51</b>	2.46	49.8	<b>0.68</b>	1.94	52.2

**Panel B. Average Cumulative Standardized Abnormal Flow**

This panel reports the average cumulative standardized abnormal flow (ACSTAF<sub>t</sub>) for each date t in the event window. We first compute the cumulative standardized abnormal flow for each fund by summing the standardized abnormal flow from 0 to t, and then dividing by the square root of the number of months used in the cumulation. We then average these over the N sample events to obtain ACSTAF<sub>t</sub> at each event date t. ACSTAF<sub>t</sub> significantly different from zero in a two-tailed test, and the percentage positive different from 50% using a chi-square test with one degree of freedom, both at the 5% significance level or better, are indicated in bold.

Event-month	One-star (n=141)			Two-stars (n=344)			Three-stars (n=527)			Four-stars (n=277)			Five-stars (n=154)		
	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0	ACSTAF <sub>t</sub>	t-stat	%>0
0	0.12	0.84	49.6	-0.15	-1.40	<b>41.6</b>	0.01	0.08	<b>45.2</b>	<b>0.42</b>	2.65	49.8	<b>0.54</b>	2.13	55.8
1	0.02	0.07	45.7	<b>-0.34</b>	-2.07	<b>39.5</b>	0.24	1.22	47.2	<b>0.56</b>	2.74	52.3	<b>1.28</b>	3.74	<b>61.6</b>
2	-0.21	-1.41	43.4	-0.22	-1.61	<b>37.5</b>	0.10	0.82	45.5	<b>0.66</b>	2.72	53.8	<b>1.30</b>	3.44	<b>62.5</b>
3	-0.04	-0.18	43.6	-0.19	-1.36	<b>38.9</b>	0.15	1.00	46.9	<b>0.86</b>	3.28	53.5	<b>1.51</b>	2.90	<b>64.3</b>
4	-0.03	-0.13	48.1	-0.14	-0.86	<b>42.1</b>	0.15	0.85	46.5	<b>0.80</b>	2.94	53.0	<b>1.44</b>	2.96	<b>60.3</b>
5	0.25	0.80	49.2	-0.22	-0.74	<b>43.6</b>	-0.06	-0.37	<b>45.3</b>	<b>1.08</b>	2.98	50.6	<b>1.45</b>	2.89	55.9
6	0.10	0.38	48.0	-0.01	-0.03	45.8	-0.08	-0.43	<b>44.4</b>	<b>1.45</b>	3.18	50.7	<b>1.59</b>	2.81	53.9

**Table 5. Determinants of the Six-Month Cumulative Standardized Abnormal Flow Response to an Initial Star Rating: Initial Jensen’s Alpha and Initial Sharpe ratio versus Initial Morningstar Star Rating**

This table reports the results of a two-way analysis of variance of cumulative standardized abnormal flows from month 0 to 6 (CSTAF<sub>6</sub>) for the full sample of funds receiving an initial Morningstar star rating (N=1267). We separately compare the explanatory power of a fund’s initial Morningstar rating (one to five stars) to either a fund’s Jensen’s alpha group or a fund’s Sharpe ratio group. We present three ways of defining Jensen’s alpha or Sharpe ratio groups: equal-sized quartiles, equal-sized quintiles, and “Morningstar-type” quintiles where the top 10% of alphas (Sharpe ratios) are in quintile 5, the next 22.5% in quintile 4, the next 35% in quintile 3, the next 22.5% in quintile 2, and the bottom 10% in quintile 1. For each comparison below labeled (1) through (6), we report the F-statistic and the corresponding p-value for the test of the null hypothesis that the sample means are identical across groups in the one rating system, conditional on the other rating system. The F test has (n<sub>1</sub> – 1, n-p) degrees of freedom, where n<sub>1</sub> equals the number of groups in the rating system being tested for significance, n equals the number of observations, and p equals the number of groups in the both rating systems. For example, the F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across Jensen’s alpha quartiles, holding constant the effect of Morningstar star rating, is 0.06 and the degrees of freedom for this test are (3, 1258). The F-statistic for the test that the mean values of CSTAF<sub>6</sub> are identical across Morningstar star ratings, holding constant the effect of Jensen’s alpha quartiles, is 6.12, with (4,1258) degrees of freedom. The symbols \*, \*\*, \*\*\* indicate statistical significance at the 10, 5, and 1% level.

Categorizations		Type II Sum of Squares F-statistic	Prob > F
(1)	Jensen’s alpha quartiles	0.06	0.982
	Morningstar star rating groups	6.12**	0.014
(2)	Jensen’s alpha quintiles	0.20	0.937
	Morningstar star rating groups	6.40**	0.011
(3)	Jensen’s alpha “Morningstar-type” quintiles	1.49	0.202
	Morningstar star rating groups	3.71***	0.005

Categorizations		Type II Sum of Squares F-statistic	Prob > F
(4)	Sharpe ratio quartiles	0.53	0.661
	Morningstar star rating groups	10.29***	0.001
(5)	Sharpe ratio quintiles	0.35	0.845
	Morningstar star rating groups	9.06***	0.003
(6)	Sharpe ratio “Morningstar-type” quintiles	0.80	0.522
	Morningstar star rating groups	10.13***	0.002

**Table 6. Cross-sectional Regressions of Cumulative Standardized Abnormal Flow on Event-Window Performance and Fund Characteristics for Star Rating Change Events**

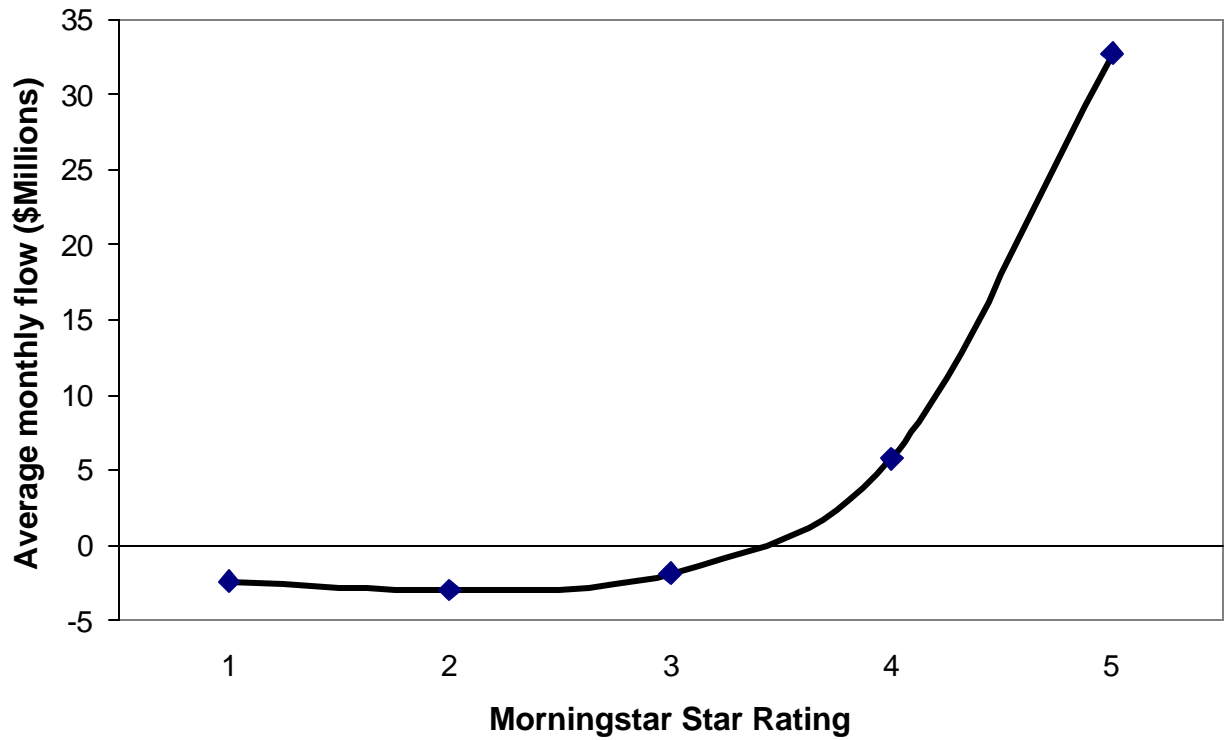
This table reports cross-sectional regressions for funds experiencing a change in star rating of cumulative standardized abnormal flows from month 0 to 6 (CSTAF<sub>6</sub>) on contemporaneous performance measures and other fund characteristics. We indicate the event timing of the measurement of the contemporaneous performance measures in the table. For example, the change in alpha (0, 5) equals the alpha of the fund at month 5 (which uses returns over the previous 36 months in its estimation) minus the alpha of the fund at month 0. Return (0,5) is the fund's compounded return from month 0 to 5. The year-end event date dummy equals one if month 0 occurred in December or January and zero otherwise. T-statistics are in parentheses and N represents the number of observations. The screened sample is the subsample of rating changes where the fund did not have any further rating changes during the (0,6) event window. The symbols \*, \*\*, \*\*\* indicate statistical significance at the 10, 5, and 1% level.

	Screened sample CSTAF <sub>6</sub>	Screened sample CSTAF <sub>6</sub>
Intercept	1.68** (2.48)	1.75*** (2.58)
Upgrade (2 to 3 star) dummy	0.64* (1.65)	0.66* (1.69)
Upgrade (3 to 4 star) dummy	1.35*** (3.46)	1.39*** (3.53)
Upgrade (4 to 5 star) dummy	2.60*** (5.76)	2.68*** (5.86)
Downgrade (2 to 1 star) dummy	-0.52 (-1.16)	-0.57 (-1.28)
Downgrade (3 to 2 star) dummy	-0.88** (-2.31)	-0.89** (-2.36)
Downgrade (4 to 3 star) dummy	-0.69* (-1.77)	-0.67* (-1.74)
Downgrade (5 to 4 star) dummy	0.75* (1.66)	0.80* (1.77)
Change in alpha (0,5) (%)	0.06 (0.15)	
Change in Sharpe ratio (0,5)		1.13 (1.04)
Return (0,5) (%)	0.02* (1.67)	0.01 (0.78)
Number of funds in style and star category	-0.003** (-2.00)	-0.003* (-1.93)
Log of total net assets (t = -1)	-0.31*** (-5.34)	-0.31*** (-5.32)
Log of fund age	0.06 (0.43)	0.06 (0.43)
Expense ratio (%)	-0.26 (-0.90)	-0.27 (-0.91)
12b-1 ratio (%)	0.39 (1.00)	0.39 (0.99)

	Screened sample	Screened sample
No-load fund dummy	-0.21 (-0.98)	-0.21 (-0.98)
Year-end event date dummy	0.23 (1.04)	0.23 (1.05)
Adjusted R-squared	0.069	0.070
N	2441	2441

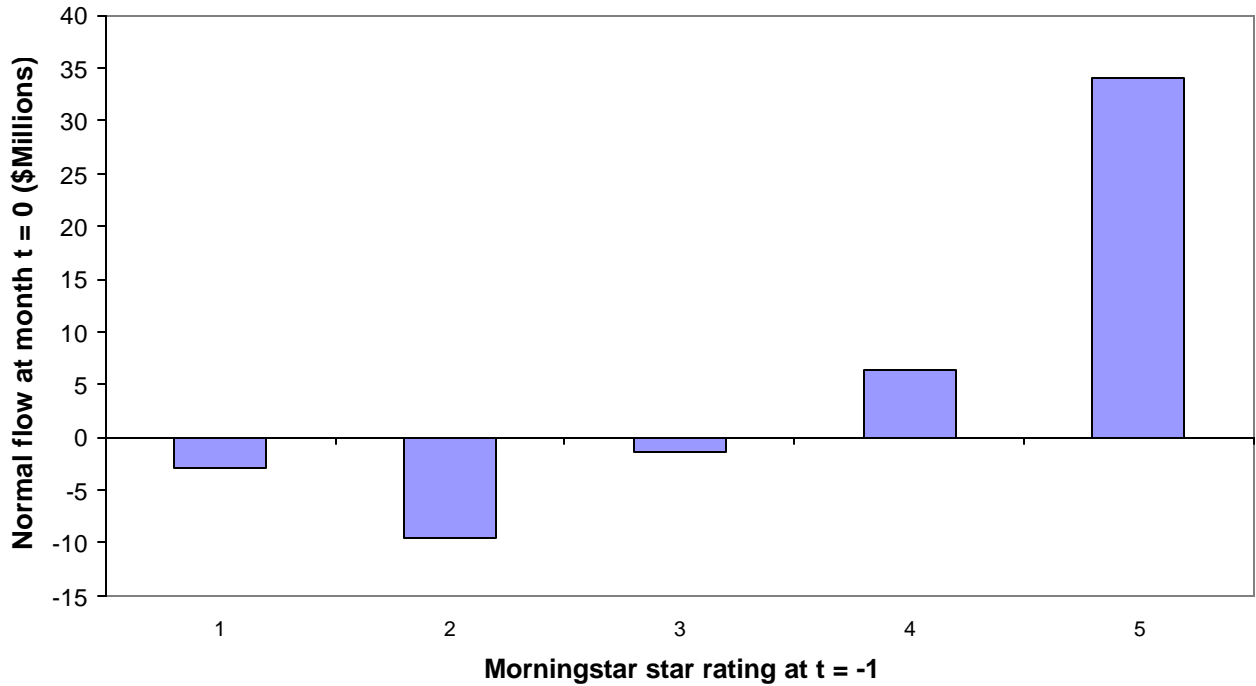
### Figure 1. Average Monthly Dollar Flow in Five Morningstar Star Rating Categories

To create this plot, we average monthly dollar fund flow for funds with a one through five star rating from November 1996 through October 1999. We use all funds in Morningstar's domestic equity star rating category described in Table 1 Panel A (111,715 fund-months).



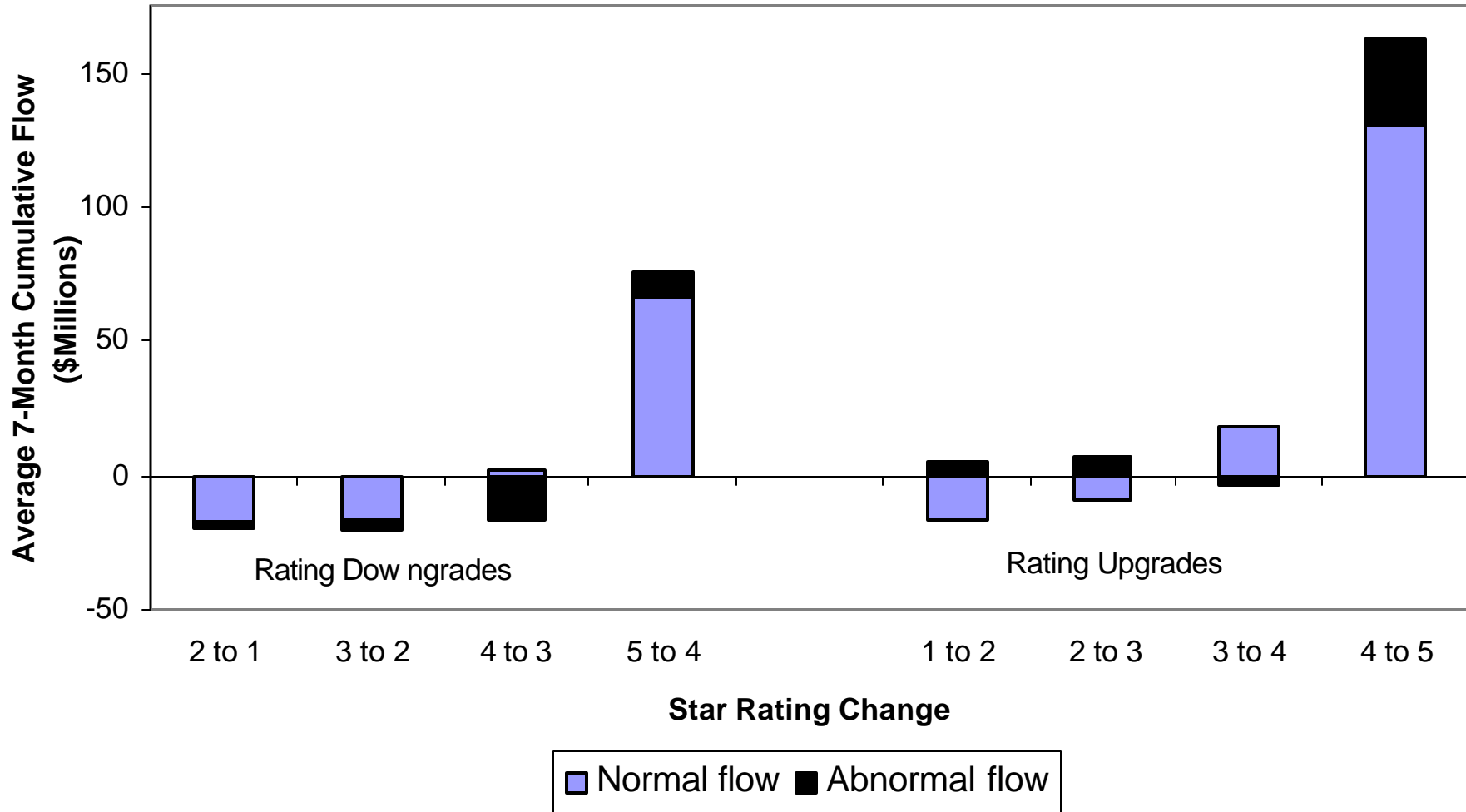
## Figure 2. Average Month Zero Normal Flow in Five Morningstar Star Rating Categories

To create this plot, we average month 0 normal flow across all funds by Morningstar star rating in month -1. Normal flow is based on a benchmark regression, analogous to a market model regression, whereby a fund's monthly flow is regressed on aggregate flow at time  $t$  to funds in its same style group, its own time  $t-1$  flow, and its own time  $t-1$  return, its change in Jensen's alpha from  $t-2$  to  $t-1$ , and its change in Jensen's alpha from  $t-2$  to  $t-1$  squared. We use all funds in Morningstar's domestic equity star rating category from November 1996 through October 1999. The number of funds used to compute each average normal flow in the order of one- to five-stars are: 8016, 18305, 28180, 18176, and 6815.



### Figure 3. Average Seven-Month Cumulative Abnormal Flow for Morningstar Star Rating Changes

To create this plot, we averaged seven-month (0,6) cumulative abnormal flow and seven-month cumulative normal flow across domestic equity funds within the same star-rating change group. Normal flow is based on a benchmark regression, analogous to a market model regression, whereby a fund's monthly flow is regressed on aggregate flow at time  $t$  to funds in its same style group, its own time  $t-1$  flow, and its own time  $t-1$  return, its change in Jensen's alpha from  $t-2$  to  $t-1$ , and its change in Jensen's alpha from  $t-2$  to  $t-1$  squared. Abnormal flow in month  $t$  is the actual dollar flow in month  $t$  minus normal flow. Note that unlike the numbers reported in Table 2, this figure contains measures of abnormal flow that are not standardized by the forecast variance.



### Figure 4. Average Seven-Month Cumulative Abnormal Flow for Initial Morningstar Star Ratings

To create this plot, we averaged seven-month (0,6) cumulative abnormal flow and seven-month cumulative normal flow across domestic equity funds within the same initial star-rating group. Morningstar assigns an initial star rating in the 36th month of a fund's existence. Normal flow is based on a market-model regression whereby a fund's monthly flow is regressed on aggregate flow at time  $t$  to funds in its same style group, its time  $t-1$  flow, its time  $t-1$  return, and its time  $t-1$  return squared. Abnormal flow in month  $t$  is the actual dollar flow in month  $t$  minus normal flow. Note that unlike the numbers reported in Table 4, this figure contains measures of abnormal flow that are not standardized by the forecast variance.

