

What Do Investors Really Care About?

Itzhak Ben-David, Jiacui Li, Andrea Rossi, Yang Song*

(Very Preliminary, Please Do Not Circulate)

Abstract

Recent studies utilize mutual fund flows to infer which asset pricing model investors use. Among the tested models, the Capital Asset Pricing Model (CAPM) was found to be “closest to the true asset pricing model.” We show that, in fact, fund flow data is most consistent with investors relying on fund rankings (e.g., Morningstar ratings) and chasing past returns. We also show that investors do not adjust for market beta or exposures to other risk factors when allocating capital among mutual funds. Flows are weaker for high-volatility funds because Morningstar penalizes funds for high total volatility.

*We thank Brad Barber, Xing Huang, and Terrance Odean for kindly sharing their data. Ben-David is with The Ohio State University and NBER, Li is with Stanford University, Rossi is with University of Arizona, and Song is with the University of Washington. Email: ben-david.1@osu.edu, jiacui@stanford.edu, rossi2@email.arizona.edu, and songy18@uw.edu.

I. Introduction

How investors allocate capital across mutual funds has been the focus of a heated academic debate in recent years. Financial economists have argued that studying investors' mutual fund choices can provide us with a lens to the way investors perceive risk in financial markets. Two celebrated studies, Barber, Huang, and Odean (2016) (henceforth BHO) and Berk and van Binsbergen (2016) (henceforth BvB), study mutual fund flows and, using different empirical techniques, reach the same conclusion: among the asset pricing models tested, investors appear to use the Capital Asset Pricing Model (CAPM). BvB concluded that the CAPM is the “closest to the asset pricing model investors are actually using” (p.2). While the idea that investors evaluate financial assets using a risk-based rational model is appealing,¹ it is potentially at odds with empirical findings from the mutual fund literature documenting that investors respond to rankings issued by rating agencies (Del Guercio and Tkac (2008), Reuter and Zitzewitz (2015)) or the Wall Street Journal (Kaniel and Parham (2017)), and chase past raw returns (Clifford, Fulkerson, Jordan, and Waldman (2013)).

In this study, we attempt to reconcile the results from the two streams of the literature. Our results show that the main determinant of capital allocation across mutual funds is external ratings, followed by past returns. We find no evidence that investors account for exposure to risk factors.

In the first part of this article, we adopt the diagnostic test proposed by BvB and test how external fund ratings (we use Morningstar ratings) compare to alphas (from the CAPM or other models) in predicting mutual fund flows. BvB's test measures the correlation between the direction of net fund flows (inflows or outflows) and different signals (e.g., the sign of a fund's alpha in different asset pricing models or Morningstar ratings, in our case). We first replicate BvB's main finding. Consistent with their results, the sign of alphas from different asset pricing models commonly used by academics agrees with the sign of fund flows 57.8% to 59.6% of the times, and the CAPM dominates other models by a small margin (60.4%). Morningstar ratings, in contrast, predict the direction of flows much better (up to 68% of the times). Based on BvB's methodology, our test results show that investors use ratings, not alphas from risk-based factor models, to guide their capital allocation.

Next, we look in depth into BHO's methodology and results. BHO decompose fund returns into components associated with a host of commonly-used risk factors and an alpha. They find that while fund flows respond to alpha and to the returns originated by most factor models, they react weakly to past returns originated from exposure to the market

¹For example, Agarwal, Green, and Ren (2018) and Blocher and Molyboga (2018) use the empirical methods of BvB and BHO in the hedge fund space.

factor. BHO conclude that investors care about market risk and therefore discount returns that originate from exposure to the market.

Our analysis shows that BHO’s findings may be interpreted in a different way. Specifically, BHO’s result is based on panel regression with time fixed effects, which is the standard method used in most of the fund flows literature. We show that, in this particular case, these regressions overweight periods with extreme market returns because in those periods the dispersion in the independent variable of interest (i.e., the market-related component of fund returns) is the highest. Also, during the same periods fund flows are significantly less responsive to fund performance, an empirical fact first documented by Franzoni and Schmalz (2017). Put together, a panel regression would convey the impression that flows respond less to the market-related component of fund returns even if investors do not use the CAPM.

In order to address this econometric issue, we examine the distribution of the coefficients from period-by-period cross-sectional regressions of fund flows on the different components of fund returns. We find that, in fact, there is no evidence that investors discount fund returns related to market risk exposure or to other risk factors. For example, if we assign equal weights to all time periods (i.e., a Fama-MacBeth specification (Fama and MacBeth (1973))), mutual fund flows respond similarly to all return components.

To provide additional insight regarding the preferences of mutual fund investors, we independently explore the determinants of fund flows. Consistently with previous studies, we find that investors chase past returns and invest according to external rankings. We find little evidence that flows respond to funds’ market beta, which confirms the results we obtained when reexamining BvB and BHO. Between Morningstar ratings and past returns, Morningstar ratings are by far the stronger determinant of fund flows. Even if we include up to 30 lags of past monthly returns in our regressions, they only explain up to 3.4% of flow variation. In contrast, the most recent Morningstar ratings explain 9.9% of the variation. This finding corroborates our initial results indicating that, using the BvB horse-race test, Morningstar ratings are the most important determinant of fund flows.

The fact that investors rely so much on Morningstar ratings also helps to explain other patterns in fund flows. Morningstar uses a methodology that does not adjust rankings for systematic risk factors, however, it does adjust for a fund’s total return volatility. In fact, we observe that fund flows are weaker for volatile funds.² This result raises the important question of whether investors independently consider volatility as a source of risk, or whether they rely solely on Morningstar rankings. We document that the latter is most likely to be

²Clifford et al. (2013) report that “net flows show aversion to risk”, which they measure as fund volatility. Besides the papers cited until this point, several other mutual fund papers used various proxies for fund-level risk or volatility, usually as control variables, in flow-performance regressions. The results in these papers are mixed.

true. Consistent with the formula that Morningstar uses to rank funds, we find that return volatility is an important determinant of Morningstar ratings, and that volatility is related to fund flows *only* through its effect on these ratings. The part of the dispersion in fund return volatility that is unrelated to ratings, which accounts for 97% of the total variation, is unrelated to flows. This suggests that volatility is negatively related to flows simply because, *ceteris paribus*, Morningstar gives lower ratings to funds with higher volatility, rather than because investor independently calculate volatility and use it to direct their flows.

In summary, our results indicate that investors do not use the CAPM, or any other of the commonly-used factor models, to allocate capital to mutual funds. Rather, they simply chase past winners, especially within the historically best-rated funds.

This paper fits into the literature that examines the relationship between mutual fund performance and investment flows into mutual funds. Early work includes Ippolito (1992), Chevalier and Ellison (1997), Sirri and Tufano (1998), Lynch and Musto (2003), Frazzini and Lamont (2008), Pástor and Stambaugh (2012), Pástor, Stambaugh, and Taylor (2015), Franzoni and Schmalz (2017), and Song (2018), among many others.³ We contribute to this literature by demonstrating that mutual fund investors behave in a less sophisticated way than asset pricing models would predict.

There are two other papers that put forward explanations for the results of BvB and BHO. Chakraborty, Kumar, Muhlhofer, and Sastry (2018) argue that the reason why investors appear to adjust for market returns and not for other risk factors is because market returns are readily available. To support their claim, they show that in the subsample of sector funds, where both market returns and sector-specific average returns are available to investors, flows also respond to the sector-specific returns. Jegadeesh and Mangipudi (2017) contest the tests proposed by BvB, arguing that simple factor models are easier to estimate and thus their alphas are more likely to win a horse race test. For the same reason, they also argue that the tests by BHO are contaminated by measurement error and therefore are tilted towards favoring a simple asset pricing model such as the CAPM.

The rest of the paper is organized as follows. Section II introduces the Morningstar ratings system. Section III describes the dataset and the linear factor models used in this paper. Section IV shows that mutual fund ratings explain fund flows much better than the CAPM model and other commonly-used asset pricing models. Section V explores the econometric framework of Barber, Huang, and Odean (2016) and finds no evidence that investors discount market-related returns more than other components of fund returns. Section VI shows that investors discount volatility only through the Morningstar ratings channel. Section VII provides concluding remarks. Robustness checks are found in appendices.

³See Barber, Huang, and Odean (2016) for a more comprehensive review.

II. Overview of Morningstar Ratings

The popularity of mutual funds as a way to own stocks has been growing for at least the past 35 years (French (2008)). The increasing demand has led to an explosion in the number of funds offered, and currently, the number of existing US equity funds exceeds the number of publicly-traded firms. The large number of available products created the need for investors to classify and rate these funds. The fund rating industry emerged to satisfy this need.

In the United States, Morningstar is the undisputed leader of this industry (Del Guercio and Tkac (2008)). Its most well-known product, the five-star rating system, was introduced in 1985 and is highly regarded and widely employed by financial professionals and advisors and used by asset management companies for the purpose of advertising (Blake and Morey (2000), Morey (2003)). This rating system has been shown to have a strong independent influence on investors' flows (Del Guercio and Tkac (2008)).

Morningstar explains its rating method in a publicly-available manual.⁴ Ratings are assigned using a relative ranking system and updated every month. Mutual funds are benchmarked against their peer funds based on their past risk-adjusted performance. Peer groups are defined as category groups (e.g., Foreign Large Value) within broadly-defined groups (e.g., International Equities). Consistent with the relevant literature, we focus on US equities funds in our study, which are categorized into nine groups⁵ based on their size tilt (Small, Mid-Cap or Large) and style tilt (Value, Blend, or Growth). The top 10% of funds within each category are assigned five stars. The following 22.5%, 35%, 22.5% and 10% of funds are assigned four, three, two and one stars, respectively.

Morningstar summarizes a fund's past performance record using the so-called Morningstar Risk-Adjusted Return (*MRAR*):

$$\text{MRAR}(\gamma) = \left[\frac{1}{T} \sum_{t=1}^T (1 + ER_t)^{-\gamma} \right]^{-\frac{12}{\gamma}} - 1, \quad (1)$$

where ER_t is the geometric return in excess of the risk-free rate in month t , γ is the risk aversion coefficient, and T is the number of past monthly returns utilized. Depending on the age of the fund, this risk-adjusted return is calculated using the past three, five, and ten years of monthly excess returns and is then annualized. The chosen value for γ is 2. No other adjustment is carried out, e.g., exposure to risk factors is not taken into account.

⁴The Morningstar manual is available at https://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf

⁵An additional category, called Leveraged Net Long, has been introduced in the US Equities group as of September 30, 2007. We do not include these funds in our sample.

Morningstar motivates this formula using expected utility theory.

In practice, *ceteris paribus*, this risk adjustment penalizes funds with higher return volatility. Based on our calculations, the magnitude of the risk adjustment increases exponentially with the fund’s monthly return volatility while it is not significantly related to the fund’s average return. The risk-adjusted return is further adjusted for sales charges, loads, and redemption fees. Because these costs can vary across different share classes of the same fund, Morningstar ratings are assigned at the share class level rather than at the fund level. We follow BHO by calculating the TNA-weighted overall star rating across share classes for a fund. As there is very little variation in star ratings across share classes, our results are similar if we use the simple average ratings across share classes.

Morningstar rates funds for different time horizons. Mutual fund share classes with history shorter than three years are not rated.⁶ Three-year star ratings are available for every fund with at least three years of past performance. Morningstar also calculates five-year and ten-year ratings for funds with track records of more than 60 and 120 months, respectively. These three ratings are then consolidated into an overall rating, which is the most salient and influential one. In case the fund is less than five years old, the rating is based on the three-year risk-adjusted return. In case the fund is at least five but not more than ten years old, the overall rating is a weighted average of the five-year and the three-year rating, with weights of 60% and 40%, respectively. In case the track record is longer than ten years, the overall rating is a weighted average of the ten-year rating (50% weight), the five-year rating (30% weight), and the three-year rating (20% weight).

III. Data and Methods

In this section, we describe the mutual fund dataset and the linear factor models used in the study, both of which are standard in the academic literature. In order to make our results directly comparable with the BvB and BHO studies, we restrict our sample to the sample of funds used in BHO, which spans from January 1991 to December 2011.⁷ To limit the extent to which our results are driven by variable construction or other methodological choices, we take the fund flows variable and several other variables (fund expense ratios, fund style assignments, fund ratings assignments, etc.) directly from the BHO dataset. Extending the BHO dataset to include observations up to the end of 2017 does not materially alter our conclusions.

⁶See <http://quicktake.morningstar.com/DataDefs/FundRatingsAndRisk.html> for details.

⁷We thank the authors for generously sharing their data. The dataset of BvB is from January 1977 to March 2011. We restrict the sample to mutual funds that start on 1991 because the CRSP database contains monthly total net assets beginning in 1991.

A. Data

We briefly explain how BHO constructed their dataset for the reader’s convenience. The BHO dataset, spanning from 1991 to 2011, is based on the standard CRSP survivorship-bias-free mutual fund database. BHO focus on actively-managed equity mutual funds. They eliminate index funds, balanced funds, and ETFs. Mutual funds are often marketed to different types of clients through different share classes that are in practice invested in the same portfolio. Since the key difference across these share classes is typically the fee structure, all share classes are aggregated into a single fund in the dataset.

Following the mainstream literature on fund flows, the investment flow for fund p in month t is defined as the net flow into the fund divided by the lagged value of its asset under management. Formally, the flow is calculated as

$$\text{Flow}_{p,t} = \frac{\text{TNA}_{p,t}}{\text{TNA}_{p,t-1}} - (1 + R_{p,t}). \quad (2)$$

This formula assumes that all flows in a given month take place at the end of that month. Here, $\text{TNA}_{p,t}$ is fund p ’s total assets under management at the end of month t , and $R_{p,t}$ is the total return of fund p in month t .

The analysis is limited to mutual funds with a minimum of 10 million in assets at the end of each month, and month t flows of between -90% and $1,000\%$. To obtain Morningstar ratings and fund style, the CRSP mutual fund dataset is merged with the fund-style box from Morningstar equity fund universe by matching on fund CUSIPs. The final sample consists of observations with successful merges. The resulting sample includes 3,432 different funds in total.

In Table I, we provide descriptive statistics for our final sample, which contains over 250,000 fund-month observations. The average fund has a modestly negative monthly flow during our sample period (-0.53%), it manages \$408.79 million, and its average age is 14.0 years. Funds with higher ratings tend to be larger and have higher flows over the following month. Consistent with the algorithm that Morningstar uses to assign ratings (Section II), higher rated funds also tend to have higher past returns and lower return volatility. Table I also presents descriptive statistics for the factor loadings on the Fama-French-Carhart (FFC) four factors (Carhart (1997)) from rolling 60-month regressions. Unsurprisingly, higher-rated funds are, on average, contemporaneously associated with higher value and momentum betas.

Table I **Descriptive statistics for the mutual fund sample.**

	Morningstar Rating						
	1 Star	2 Stars	3 Stars	4 Stars	5 Stars	Rating NA	All
Fund characteristics (fund-month obs.)							
# Observations	17,024	60,416	92,131	60,613	18,279	8,590	257,053
Fund size (\$mil)	161.28	249.98	408.35	688.58	1291.16	181.39	408.79
Fund age (years)	13.37	13.96	14.17	14.33	13.68	13.29	14.04
Fund flow	-1.54%	-1.23%	-0.69%	0.17%	1.14%	-0.42%	-0.53%
Weighted past return	-0.08%	0.18%	0.36%	0.55%	0.78%	0.36%	0.37%
Ret volatility (1 year)	5.51%	5.05%	4.85%	4.81%	4.89%	4.67%	4.93%
Ret volatility (5 years)	6.28%	5.55%	5.22%	4.94%	4.93%	4.97%	5.27%
Correlation with market	0.60	0.61	0.64	0.66	0.70	0.58	0.64
Market beta	0.82	0.76	0.75	0.75	0.79	0.65	0.76
Size beta	0.19	0.13	0.12	0.12	0.13	0.16	0.13
Value beta	-0.031	0.013	0.038	0.063	0.078	0.076	0.038
Momentum beta	-0.011	0.011	0.016	0.023	0.043	0.020	0.017

B. Linear factor models

The tests carried out by BvB and BHO are designed to assess the ability of alphas, derived from different factor models, in explaining mutual fund flows. We now describe how we construct these alphas based on historical fund performance data.

As an example, consider the BHO seven-factor model, which augments the FFC four factors with the three industry factors of Pástor and Stambaugh (2002). Following BHO, for each fund p in month t , we estimate the following time-series regression using the 60 months of returns from month $t - 60$ to month $t - 1$:

$$\begin{aligned}
 R_{p,\tau} - R_{f,\tau} = & a_{p,t}^{7F} + b_{p,t}(\text{MKT}_\tau - R_{f,\tau}) + s_{p,t}\text{SMB}_\tau + h_{p,t}\text{HML}_\tau \\
 & + u_{p,t}\text{UMD}_\tau + \sum_{k=1}^3 \gamma_{p,t}^k \text{IND}_k\tau + \nu_{p,\tau}, \quad \tau \in \{t - 60, \dots, t - 1\}. \quad (3)
 \end{aligned}$$

Here, $R_{p,\tau}$ is the mutual fund return net of fees in month τ , $R_{f,\tau}$ is the risk-free interest rate,⁸

⁸The risk-free interest rate here is the one-month Treasury bill rate. We download the interest rate series together with the factor returns from Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

MKT is the return on the value-weighted market portfolio, SMB, HML, and UMD are the returns on the three factor portfolios in Fama and French (1993) and Carhart (1997). IND1, IND2, and IND3 are three industry factors defined in Pástor and Stambaugh (2002), and they represent the first three principal components of the residuals in multiple regressions of the Fama-French industry returns on the MKT, SMB, HML, and UMD factors. The parameter $a_{p,t}^{7F}$ is the average factor-adjusted return, while $b_{p,t}$, $s_{p,t}$, $h_{p,t}$, $u_{p,t}$ and $\gamma_{p,t}^k$ are the fund exposures to the market, size, value, momentum and industry factors, respectively.

Following BHO, we then calculate the seven-factor alpha for fund p in month t as its realized return less the return related to the fund's factor tilts in month t :

$$\begin{aligned} \hat{\alpha}_{p,t}^{7F} = & R_{p,t} - R_{f,t} - \left[\hat{b}_{p,t}(\text{MKT}_t - R_{f,t}) + \hat{s}_{p,t}\text{SMB}_t + \hat{h}_{p,t}\text{HML}_t \right. \\ & \left. + \hat{u}_{p,t}\text{UMD}_t + \sum_{k=1}^3 \hat{\gamma}_{p,t}^k \text{INDk}_t \right], \end{aligned} \quad (4)$$

where $\hat{b}_{p,t}$, $\hat{s}_{p,t}$, $\hat{h}_{p,t}$, $\hat{u}_{p,t}$, and $\hat{\gamma}_{p,t}^k$ are the estimated coefficients in Equation (3).

Investors often do not respond to fund performance instantaneously (Coval and Stafford (2007)). To allow for the slow response of flows to returns, we follow BHO and use an exponential decay function to model the response of flows to fund returns in the past 18 months. That is, the seven-factor alpha measure we use in month t is computed as

$$\text{ALPHA}_{p,t}^{7F} = \frac{\sum_{s=1}^{18} e^{-\lambda(s-1)} \hat{\alpha}_{p,t-s}^{7F}}{\sum_{s=1}^{18} e^{-\lambda(s-1)}}, \quad (5)$$

where the decay parameter λ is estimated empirically from the relationship between flows and past returns and $\hat{\alpha}_{p,t}^{7F}$ is from Equation (4). In our implementation, we follow BHO and use $\lambda = 0.20551497$.

Similarly, we calculate the CAPM alpha for fund p in month t as

$$\hat{\alpha}_{p,t}^{\text{CAPM}} = R_{p,t} - R_{f,t} - \hat{\beta}_{p,t}(\text{MKT}_t - R_{f,t}), \quad (6)$$

where $\hat{\beta}_{p,t}$ is estimated in the 60 months prior to t . Thus, the relevant CAPM-alpha measure in month t is a weighted average of the prior eighteen monthly CAPM alphas in Equation (6):

$$\text{ALPHA}_{p,t}^{\text{CAPM}} = \frac{\sum_{s=1}^{18} e^{-\lambda(s-1)} \hat{\alpha}_{p,t-s}^{\text{CAPM}}}{\sum_{s=1}^{18} e^{-\lambda(s-1)}}. \quad (7)$$

We also calculate the weighted averages of the Fama-French three-factor alpha (Fama and French (1993)), $\text{ALPHA}_{p,t}^{\text{FF}}$, and the FFC four-factor alpha, $\text{ALPHA}_{p,t}^{\text{FFC}}$, respectively.

IV. Morningstar Ratings Trump CAPM

In this section, we show that Morningstar ratings explain mutual fund flows significantly better than the CAPM model and other commonly-used models. To this end, we rely on the diagnostic test proposed by Berk and van Binsbergen (2016) (henceforth BvB).

BvB propose that mutual fund flows can be used to infer which asset pricing models investors use. The core idea behind their methodology is that mutual fund investors compete with each other to find positive net present value (NPV) investment opportunities. Funds with positive alphas are positive NPV investment opportunities, and vice versa. As investors observe fund returns and alphas over time, they should respond by directing their money accordingly. Hence, they argue, by investigating how well the signs of alphas match the directions of fund flows, it is possible to deduce which asset pricing model investors are indeed using. Based on this test, BvB find that CAPM alphas match flows the best and therefore conclude that the CAPM is the closest to the “true” asset pricing model that investors use.

For reader’s convenience, as we detail the test we run, we also illustrate BvB’s methodology. For each fund p in each month t , let $F_{p,t}$ denote the fund flow and let $\text{ALPHA}_{p,t}^{\mathcal{M}}$ denote the inferred return alpha under the asset pricing model \mathcal{M} . Notice that $\text{ALPHA}_{p,t}^{\mathcal{M}}$ is calculated using historical returns prior to t , as one can see from, for example, equations (5) and (7). To refrain from making restrictive functional form assumptions on the flow-performance relationship, BvB’s method makes use of the sign of fund flows and the model-implied alphas. Following the method of BvB⁹, for a given asset pricing model \mathcal{M} , we run the following regression:

$$\text{sign}(F_{p,t}) = \beta_0^{\mathcal{M}} + \beta_1^{\mathcal{M}}\text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}}) + \epsilon_{p,t}, \quad (8)$$

where $\text{sign}(F_{p,t})$ and $\text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}})$ take on values in $\{-1, 1\}$. Lemma 2 of BvB shows that a linear transformation of the regression slope, intuitively, equals the frequency in which the alpha and flow signs match each other. In other words,

$$\begin{aligned} & \frac{\beta_1^{\mathcal{M}} + 1}{2} \\ = & \frac{\widehat{\Pr}(\text{sign}(F_{p,t}) = 1 | \text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}}) = 1) + \widehat{\Pr}(\text{sign}(F_{p,t}) = -1 | \text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}}) = -1)}{2}, \end{aligned}$$

where $\widehat{\Pr}(\cdot)$ denotes the occurrence frequency in the sample.

⁹We are a bit different from BvB, as BvB use alphas that are contemporaneous with the flows. We lag the alphas by one month, which avoids a potential look-ahead bias.

Table II **Fraction of fund flows that is positive for each Morningstar rating category.**

Morningstar ratings	Number of observations	Fraction of funds	Fraction of positive flows
1	17,024	6.9%	15.9%
2	60,416	24.3%	19.4%
3	92,131	37.1%	29.7%
4	60,613	24.4%	49.3%
5	18,279	7.4%	67.0%

In their Table 2, BvB find that the signs of CAPM alpha match the flows signs better than the commonly used risk models. CAPM alpha also does better than the “no model” benchmark, defined as simply the fund return in excess of the market. Thus, they conclude that the CAPM is closest to the “true” model used by investors.

However, we find that the simple heuristics of reallocating capital based on Morningstar fund ratings explain the signs of fund flows much better than the CAPM model. To set the stage, Table II shows that Morningstar ratings have a significant explanatory power on fund flows. For instance, only 15.9% of funds with a one-star rating have positive flows in the next month. The fraction of funds with positive flows increases monotonically with ratings, reaching 67.0% for the highest rating category (five-star funds).

We now consider the following simple heuristic model: investors increase allocation to funds with ratings $\geq i$ and decrease allocation to those with ratings $< i$. We consider three possible thresholds, i.e., $i = 3, 4$, and 5 . These rankings comprise of the top 68.9%, 31.8%, and 7.4% mutual fund-months, respectively. We estimate Equation (8) for the asset pricing models and our rating-based heuristic models. Following BvB, standard errors are double-clustered by fund and by time. The results are shown in the first two columns of Table III.

The table shows that we can closely replicate BvB’s finding that the CAPM performs slightly better than the market-adjusted model, the FF three-factor model, and the FFC four-factor model. The CAPM does also better than BHO’s seven-factor model, which augments FFC factors with the three industry factors. We also find that the simple excess return model (the “no-model” benchmark) performs the worst. However, the rating-based heuristics significantly outperforms the CAPM, and the degree of outperformance is larger than the dispersion among the scores of all other models. The best performing heuristics, which has investors reallocating money into five-star funds, gets the flow signs right approximately

Table III **Horse race of different models.** The first two columns are estimates of Equation (8) for each model considered. For ease of interpretation, the table reports $(\beta_1^M + 1)/2$ in percent, and models are ordered in decreasing order of the point estimate of β_1^M . The remaining columns provide statistical significance tests of the pairwise model horse races based on Equation (9). Each cell reports the t-statistic of the hypothesis that $\beta^{\text{row}} > \beta^{\text{column}}$. For both univariate and pairwise tests, standard errors are double clustered by fund and time.

Model	Estimate of $(\beta_1^M + 1)/2$	Univariate t -stat	Rating ≥ 5	Rating ≥ 4	Rating ≥ 3	CAPM	Market- adjusted	FF 3-factor	FFC 4-factor	BHO 7-factor	Excess return
Rating ≥ 5	67.95	29.48	-	5.60	9.57	9.93	10.80	12.70	13.13	13.95	11.20
Rating ≥ 4	64.41	36.32	-	-	10.07	8.27	9.21	12.26	12.82	13.86	8.53
Rating ≥ 3	61.01	32.15	-	-	-	1.43	2.81	5.33	6.22	7.58	5.15
CAPM	60.36	25.62	-	-	-	-	5.59	6.53	7.22	9.28	4.67
Market-adjusted	59.64	22.64	-	-	-	-	-	3.07	4.06	6.16	3.68
FF 3-factor	58.76	26.44	-	-	-	-	-	-	3.37	5.08	2.61
FFC 4-factor	58.36	25.83	-	-	-	-	-	-	-	3.98	2.08
BHO 7-factor	57.76	22.44	-	-	-	-	-	-	-	-	1.27
Excess return	56.90	11.78	-	-	-	-	-	-	-	-	-

68% of the time, while the CAPM gets the flow signs right roughly 60% of the time. The difference is close to 7.6%, which, for comparison, is much larger than the difference between CAPM (60.4%) and the worst performing model (excess returns, 56.9%).

Is this outperformance of rating-based heuristics statistically significant? We follow BvB to conduct pairwise model horse races. For any two models $\mathcal{M}1$ and $\mathcal{M}2$, we run regression

$$\text{sign}(F_{p,t}) = \gamma_0 + \gamma_1 \left(\frac{\text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}1})}{\text{var}(\text{ALPHA}_{p,t}^{\mathcal{M}1})} - \frac{\text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}2})}{\text{var}(\text{ALPHA}_{p,t}^{\mathcal{M}2})} \right) + \xi_{p,t} \quad (9)$$

and we consider $\mathcal{M}1$ to be a better model of investor behavior if $\gamma_1 > 0$ with statistical significance. We double-cluster standard errors by fund and by time. The results are reported in the remaining columns in Table III. The first two rating-based models all outperform the CAPM with strong statistical significance, with t -statistics of 9.93 and 8.27, respectively.

Based on BvB’s logic, the test results suggest that Morningstar ratings explain investors’ capital reallocation better than the CAPM and all other asset pricing models considered.

V. Investors Do Not Adjust for Market Beta

Similar to BvB, Barber, Huang, and Odean (2016) (BHO) analyze mutual fund flows in order to infer which asset pricing model investors use. BHO, however, take a different approach. They decompose fund returns into factor-related returns and an alpha, and estimate how mutual fund flows respond to these different components. Using a pooled regression with time fixed effects (FEs), BHO find that fund flows are much less responsive to a fund’s market-related returns than to other components of fund returns. Since investors appear to discount returns arising from exposure to market risk, BHO conclude that investors presumably use a model akin to the CAPM.

In this section, we suggest a different explanation for BHO’s result. The difference in interpretation has to do with the fact that, by construction, panel regressions overweight periods in which there is more dispersion in the independent variable. Also, most of the variation in the independent variable of interest, i.e., the market-related component of fund returns, is concentrated in periods with extreme market returns, when the sensitivity of fund flows to cross-sectional differences in fund returns is particularly low. Once we account for this issue, we

find no evidence that investors differentiate market-related returns from returns related to other factors or alphas. In other words, investors don’t account for market beta or fund exposures to other factors when allocating capital across mutual funds.

A. BHO's return decomposition approach

We briefly explain BHO's methodology for the reader's convenience. For each fund, they use rolling time series regressions to decompose monthly fund excess returns into seven factor-related components (market, size, value, momentum, and the three industry factors of Pástor and Stambaugh (2002)) and the residual, which they refer to as the seven-factor alpha. They account for the slow response of flows to past returns by applying an exponential decay function to each of the return components in the past 18 months. For instance, the relevant market-related return in month t is

$$\text{MKTRET}_{p,t} = \frac{\sum_{s=1}^{18} e^{-\lambda(s-1)} \hat{b}_{p,t-s} (\text{MKT}_{t-s} - R_{f,t-s})}{\sum_{s=1}^{18} e^{-\lambda(s-1)}}, \quad (10)$$

where $\hat{b}_{p,\tau}$ is the fund exposure to the market factor under the seven-factor model in (3), as estimated using the past 60-month return prior to month τ . They also calculate returns related to the fund's size, value, momentum, and three industry tilts, which are labeled SIZRET, VALRET, MOMRET, INDRET1, INDRET2, and INDRET3, respectively.

To infer investor response to different return components, BHO estimate the following panel regression with time fixed effects:

$$\begin{aligned} F_{p,t} = & \mu_t + \gamma X_{p,t} + b_0 + b_{\text{ALPHA}} \text{ALPHA}_{p,t}^{7F} + b_{\text{MKT}} \text{MKTRET}_{p,t} + b_{\text{SMB}} \text{SIZRET}_{p,t} \\ & + b_{\text{HML}} \text{VALRET}_{p,t} + b_{\text{MOM}} \text{MOMRET}_{p,t} + \sum_{k=1}^3 b_{\text{IND}_k} \text{INDRET}_{k,p,t} + e_{p,t}, \end{aligned} \quad (11)$$

where $F_{p,t}$ is the monthly fund flow, μ_t is the time fixed effect in month t , and $X_{p,t}$ is a vector of control variables. The controls include the total expense ratio, a dummy variable for no-load, a fund's return standard deviation over the prior one year, the log of fund size in month $t - 1$, the log of fund age, and lagged fund flows from month $t - 19$. The coefficients $b_{\text{ALPHA}}, b_{\text{MKT}}, \dots$, measure how fund flows respond to different return components. Standard errors are two-way clustered by month and fund.

Using the data provided to us by BHO, we are able to exactly reproduce their key result, which we report in Column (1) of Table IV (see Table 5 of BHO). In Column (2) of Table IV, we also report the difference between each reported coefficient and the coefficient on the market-related return component. As noted in BHO and reproduced in Column (1) of Table IV, the response coefficient to market-related returns, ($b_{\text{MKT}} = 0.25$), is significantly lower than the coefficients on all other components of returns. Based on this result, BHO concluded that investors discount market-related returns more than other components when assessing mutual fund performance, implying that investors appear to be using the CAPM

Table IV **Response of fund flows to components of fund returns.** This table presents coefficient estimates from panel regressions of percentage fund flow (dependent variable) on the components of a fund's return in Equation (11). The controls include the total expense ratio, a dummy variable for no-load, a fund's return standard deviation over the prior one year, the log of fund size in month $t - 1$, the log of fund age, and lagged fund flows from month $t - 19$. Columns (1) and (3) are based on pooled regression with time FEs and Fama-Macbeth regression, respectively. Columns (2) and (4) report the difference between the flow-response to other return components and the flow-response to MKTRET. Column (5) shows the change in each of the coefficient estimates by the two different regression methods (Columns (1) and (3)). The t -statistics (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	BHO panel regression with time FEs		Fama-Macbeth regression		Change in coefficients
	(1)	(2)	(3)	(4)	(5)
	Coefficients	Difference	Coefficients	Difference	
ALPHA ^{7F}	0.88*** (32.74)	0.63*** (10.15)	1.04*** (39.70)	0.24* (1.96)	18%
MKTRET	0.25*** (4.52)	-	0.80*** (6.65)	-	216%
SIZERET	0.76*** (14.06)	0.51*** (6.50)	0.54*** (3.24)	-0.26 (-1.27)	-29%
VALRET	0.67*** (10.56)	0.42*** (4.89)	0.93*** (5.63)	0.13 (0.65)	40%
MOMRET	1.06*** (17.65)	0.81*** (9.82)	0.65** (2.28)	-0.15 (-0.47)	-38%
INDRET1	0.92*** (12.43)	0.67*** (7.19)	0.76*** (4.91)	-0.04 (-0.18)	-17%
INDRET2	0.70*** (7.38)	0.45*** (4.06)	0.98*** (3.74)	0.18 (0.62)	40%
INDRET3	0.69*** (7.97)	0.44*** (4.25)	1.14*** (3.40)	0.24 (0.95)	64%
Month FE	Yes	-	Yes	-	-
Controls	Yes	-	Yes	-	-
Observations	257,053	-	257,053	-	-
Adjusted R^2	0.173	-	0.204	-	-

in their capital allocation decisions.

Compared to the methodology of BvB, the econometric specification of BHO has the advantage that it exploits the full variation in fund flows as opposed to just using the sign of the flow. However, BHO’s test has an important drawback. We argue that the results in the first column of Table IV are partially driven by the time-varying nature of the flow-performance sensitivity (FPS). As pointed out by Pástor, Stambaugh, and Taylor (2017), the coefficient estimates in a pooled regression with time FEs (as used by BHO) are weighted averages of period-by-period cross-sectional coefficient estimates, with more weights placed on periods where the independent variable has larger cross-sectional variation. At the same time, Franzoni and Schmalz (2017) show that fund flows are less responsive to past returns right after periods when market returns have been extreme, both positively and negatively so. By construction, most of the cross-sectional variation the market-related component of fund returns happens precisely in these periods, when flows respond weakly to fund returns. As a consequence, when all periods are pooled together in the panel regression, it appears as if flows respond weakly to the market-related component of fund returns.

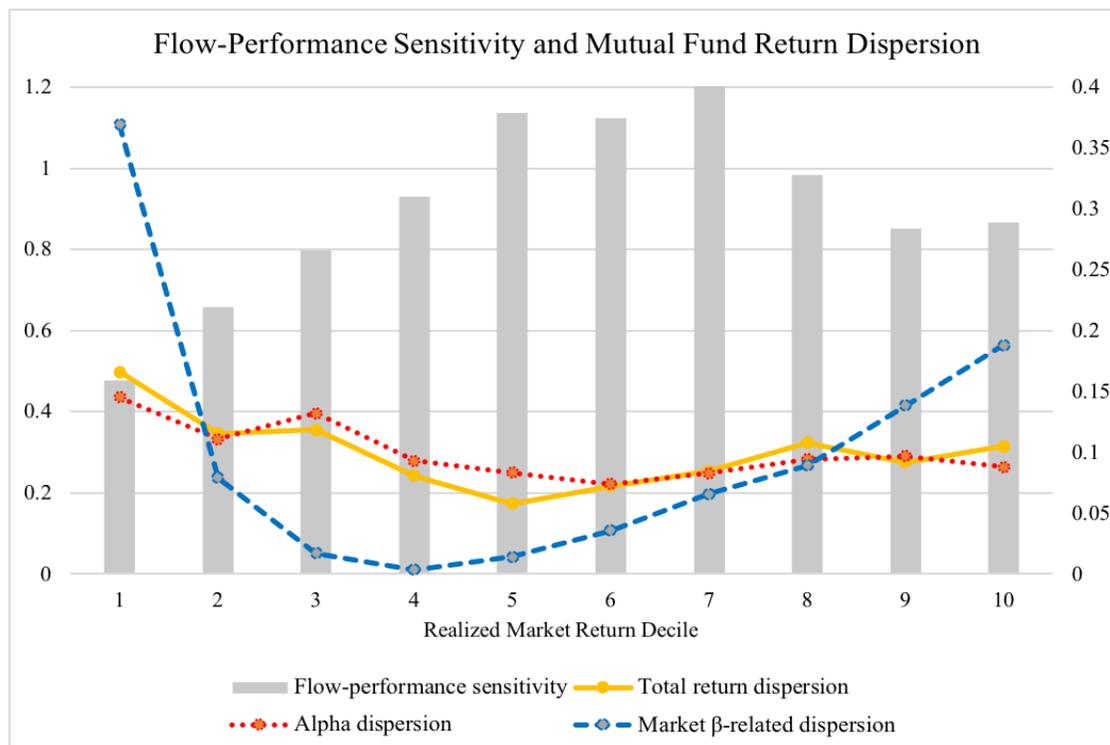
In the next section, we will show that the time-varying nature of the FPS causes the estimated average response of fund flows to marker-related returns to be downward biased. After adjusting for this effect, we no longer find evidence that investors discount market-related returns more than other return components.

B. Flow-performance sensitivity in different market states

To illustrate the relationship between market returns and the sensitivity of fund flows to returns, we reproduce the observation of Franzoni and Schmalz (2017) in Figure 1. In particular, we split the entire sample period into ten buckets depending on the past-18-month-weighted excess returns of the aggregate market factor. We measure FPS as the slope from monthly cross-sectional regressions of fund flows on prior 18-month weighted fund returns. We report the average FPS for all the ten buckets in Figure 1.

The figure shows that the FPS is a hump-shaped function of aggregate market realizations (left axis). This is consistent with the finding of Franzoni and Schmalz (2017). The FPS is more than twice as large in moderate states as in the states when the aggregate market has extremely negative returns. While the FPS is a hump-shaped function of past realized market returns, the cross-sectional dispersion in the market-related component of fund returns is an inverse hump-shaped function of it, by construction. In contrast, the cross-sectional dispersion in seven-factor alpha or in other factor-related returns is essentially flat across

Figure 1: **Flow-performance sensitivity in different market states.** We split the entire sample period into ten market-state buckets depending on the past-18-month-weighted excess returns of the aggregate market. We then measure the flow-performance sensitivity (FPS) each month as the estimated coefficient from the monthly cross-section regressions of percentage flows on the past-18-month-weighted fund returns. We also calculate the monthly cross-sectional standard deviation of the fund market-related returns, the BHO 7F-alphas, and the total fund returns, respectively. The grey bars (the left axis) present the time-series averages of the FPS for each of the ten market-state buckets. The blue, red, and yellow lines (the right axis) show the time-series averages of the cross-sectional variation in the market-related returns, the BHO 7F-alphas, and the total returns for each market-state buckets, respectively.



different market states.¹⁰ Therefore, based on the mathematical relationship between cross-sectional and pooled regression estimates derived by Pástor et al. (2017) and mentioned in the previous subsection, the estimate of the flow response to market-related returns in a pooled regression with time FEs is likely to overweight periods with smaller flow-performance sensitivities, and this does not apply to the other coefficients. In other words, the pooled regression estimate of $b_{\text{MKT}} = 0.25$ is likely downward biased relative to other coefficient estimates in Equation (11).

Can the relation between market returns and the sensitivity of flows to returns explain BHO’s finding that flows are less sensitive to the part of a fund’s return that is attributable

¹⁰ We also find that, after controlling for the market factor, the flow-performance sensitivity does not meaningfully depend on the volatility of other factors.

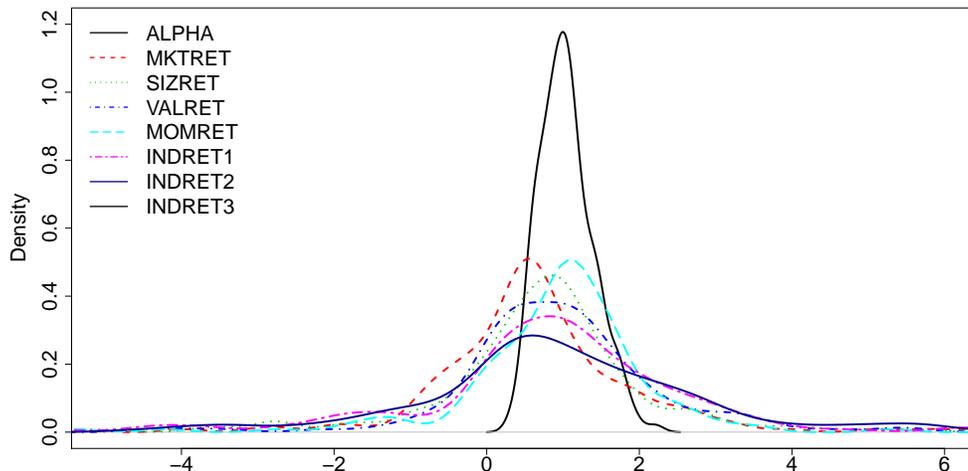
to its exposure to the market factor? To answer this question, we run a Fama-Macbeth (FM) regression (Fama and MacBeth (1973)) of fund flows on different return components. That is, for each month, we run cross-sectional regressions of fund flows on the eight components of fund returns (and controls) in Equation (11), and then we calculate the time-series averages of the estimated cross-sectional coefficients. In contrast with the pooled regression with time FEs, this amounts to equally weighting the period-by-period cross-sectional coefficient estimates. We report the results in Columns (3) and (4) of Table IV. We also report the changes in the estimated coefficients between the FM regression and the pooled regression with time FEs in Column (5).

As expected, when comparing the results from the FM regression (Column (4)) with those of the panel regression (Column (2)), the most significant change is in the point estimate of the fund response to the market-related returns, which becomes more than three times as large (from 0.25 to 0.80). As shown in Column (5) of Table IV, the changes in the other coefficients are much smaller and have no clear pattern, with 3 out of 7 decreasing and the other 4 increasing. While the market-related coefficient is significantly smaller than all other coefficients in the panel regression specification, in the FM specification it is no longer statistically significantly different from all the other factor-related components at the conventional 5% confidence level, as shown in Column (4) of Table IV. Moreover, in the latter specification, the coefficient on the market-related return has a higher point estimates than the size-, momentum-, and the first industry factor-related returns.

Here, we are not trying to argue that one of the two econometric approaches is superior to the other. After all, the FM regression and the pooled regression with time FEs offer different schemes for weighting the period-by-period cross-sectional coefficient estimates (Pástor et al. (2017)). In fact, a more direct way to gain insight on this issue is to look at the distributions of the coefficient estimates across all cross-sectional regressions. In Figure 2, we plot the kernel density of period-by-period cross-sectional regression coefficients for different return components. While the distribution of the coefficient on the factor-adjusted return, ALPHA^{7F} , is more concentrated, the coefficients on all factor-related components are all highly dispersed and not clearly different from each other. We have shown in Column (4) of Table III that one cannot reject the null that the coefficient on the market-related return is different from the coefficient on the other factor-related returns. The coefficient of market-related returns is only different from that of the alpha measure at 10% confidence level.

One may wonder why the distribution of the ALPHA^{7F} coefficients from cross-sectional regressions is much more concentrated than that of the coefficients for the other factors. This happens because, within each time period, cross-sectional differences in fund returns are

Figure 2: **Density of the period-by-period cross-sectional coefficient estimates.** This figure presents the density of the coefficient estimates from cross-sectional regression of percentage fund flow on the components of a fund’s return: a fund’s BHO 7F alpha and seven factor-related returns.



always highly correlated with cross-sectional differences in alphas, but this is not the case for the 7 factor-related returns. In Appendix Table X, we report average and median Pearson and Spearman’s rank correlations between total fund return and its eight components, ALPHA^{7F} , MKTRET , etc. Only alpha has always a positive correlation with the fund return, for example, the Spearman correlation across each month is 0.71. The same figure is between -0.03 and 0.26 for the other seven fund returns components. This is not surprising because, during periods in which a given factor has a return close to zero, the dispersion in returns associated with that factor is also close to zero. In contrast, the cross-sectional dispersion in alphas is always relatively large and stable. Hence, as investors move money from low-return funds into high-return funds, in most cases they also happen to be moving money from low-alpha funds into high-alpha funds, hence, the flow-performance coefficient on the alpha component is always relatively close to its mean. The same can not be said for the other components on fund returns. This observation also explains why, in Columns (1) and (3) of Table IV, the t -statistic of the ALPHA^{7F} coefficient is significantly larger than that of the other coefficients.

BHO also have several robustness checks with different subsamples.¹¹ We repeat their

¹¹BHO also conduct a nonlinear pairwise test of asset pricing models in their Table 4. They find that the CAPM and the market-adjusted model clearly win against all other models with more factors, and that the CAPM slightly beats the market-adjusted model. We redo their exercise using the Fama-MacBeth regression. We find that CAPM does not outperform the market-adjusted model if one uses this econometric approach. Detailed results for this analysis are presented in Appendix A.

Table V **Response of fund flows to market beta.** This table presents coefficient estimates from panel regressions of percentage fund flow on past returns and market beta in Equation (12). The t -statistics (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Pooled	Pooled	Pooled	FM	FM	FM
	(1)	(2)	(3)	(4)	(5)	(6)
Past Ret	0.22*** (12.28)	0.18*** (12.24)	0.17*** (11.53)	0.31** (28.19)	0.32*** (11.33)	0.29*** (12.72)
Mkt-beta	-0.00010 (-0.24)	-0.000033 (-0.086)	-0.00015 (-0.42)	-0.00014 (-0.35)	0.00053 (0.64)	-0.00068 (-1.19)
Month FE	Yes	No	No	Yes	No	No
Month-style FE	No	Yes	No	No	Yes	No
Month-rating FE	No	No	Yes	No	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs	257,053	257,053	257,053	257,053	257,053	257,053
Adjusted R^2	0.090	0.11	0.16	0.095	0.10	0.074

exercises but using the FM regression approach instead. In all of these additional exercises, we again find that one cannot reliably conclude that investors discount market-related returns more than other factor-related returns. The results are presented in Table V of Appendix A. Therefore, based on these tests, we argue that there is no evidence that investors discount market-related returns differently than they do for other components of fund returns.

In the next section, we will provide direct evidence that mutual fund investors do not behave as if they account for market beta when they allocate capital among mutual funds.

C. *Additional Evidence that Investors Do Not Attend to Market Beta*

In the previous section, we showed that the econometric test proposed by BHO might deliver spurious evidence in support of the CAPM because the dispersion in the market-related component of fund returns varies systematically over time with the FPS. Note, however, that market beta itself is less likely to be affected by this problem, because the dispersion in market betas across funds is relatively stable over time. In this section, we show that, in a simple flow-performance regression setting, market beta is not a significant predictor of fund future flows. We reach this conclusion by using both pooled regression with time FEs and Fama-Macbeth regressions.

To this end, we first estimate the following regression with time FEs:

$$F_{p,t} = \nu_t + \psi \text{RET}_{p,t} + \phi \hat{\beta}_{p,t} + \gamma X_{p,t} + \epsilon_{p,t}, \quad (12)$$

where ν_t is the time fixed effect, $\text{RET}_{p,t}$ is the weighted average of the 18-month returns prior to month t , $\hat{\beta}_{p,t}$ is the estimated market beta in the CAPM model from time $t - 60$ to $t - 1$, and $X_{p,t}$ is a vector of controls as in Equation (11). Standard errors are double-clustered by time and by fund. We also estimate the same equation by means of a FM regression. The results are displayed in Table V.

The results suggest that controlling for past fund returns, market beta does not influence fund flows in a significant manner. This result holds with both pooled regressions with time FEs and with the FM procedure, and are robust to controlling for month-style and month-rating fixed effects. In other words, investors do not seem to adjust for market beta when they allocate flows among mutual funds. This implies that CAPM is unlikely to be the model that investors use. The results presented in Table V, although based on a simple model specification, provide a meaningful sanity check for the other results we present through the present study.

VI. Morningstar Ratings are Primary Drivers of Flows

In this section, we provide additional evidence that Morningstar ratings are important in explaining mutual fund flows. This evidence is complementary to the findings presented in Sections IV and V. We first show that the explanatory power of Morningstar ratings is much higher than that of past fund returns. We then consider one of the common empirical findings in the mutual fund flows literature, that is, that flows are negatively associated with fund return volatility. We show that this empirical relation is driven by the fact that Morningstar ratings penalize volatility, rather than because investors are averse to funds with high volatility.

A. Explanatory power of Morningstar ratings

So far we have argued that mutual fund investors do not seem to be using asset pricing models when allocating money across funds, but that they rather simply chase past returns and rely on popular fund ranking systems, the most prominent of which is the Morningstar five-star rating. In this section, we ask how important fund ratings are in explaining fund

flows. To put the importance of Morningstar ratings in perspective, we compare the explanatory power of Morningstar rating with that of fund returns. Past returns have been the most cited and studied determinant of fund flows (Christoffersen, Musto, and Wermers (2014)).

A simple comparison of the adjusted R^2 in Columns (3) and (4) in Table VI suggests that ratings are even more important than past fund returns in predicting future flows. The only difference between those two regression specifications is that the first includes the weighted past fund returns, while the second includes ratings. The second regression achieves an adjusted R^2 that is 5.1 percentage points higher. A potential confounding factor when comparing the R^2 across these regressions is that some of the explanatory power of past return might be absorbed by the other controls. However, a univariate regression of fund flows on past return alone only achieves an R^2 of 2.90%, while ratings alone explain 9.86% of the variation in fund flows.

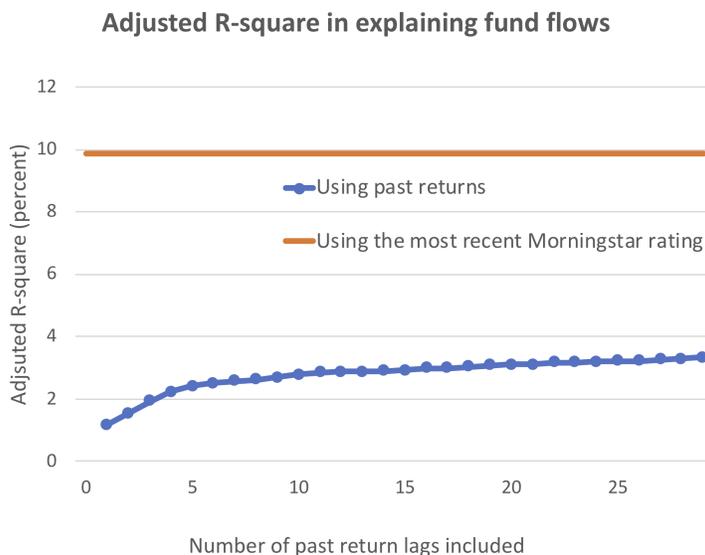


Figure 3: The adjusted R-square when regressing fund flows on up to 30 lags of past fund returns (blue) or when regressing on the most recent Morningstar fund rating (orange).

One may be concerned that the response of flows to past returns varies for different horizons. To address this concern, we regress fund flows on 30 lags of past returns,

$$F_{p,t} = a + \beta_1 R_{p,t-1} + \beta_2 R_{p,t-2} + \dots + \beta_{30} R_{p,t-30} + \epsilon_{p,t},$$

and plot the adjusted R^2 achieved when including $h = 1, \dots, 30$ lags in Figure 3. It is apparent that the incremental benefit of including additional lags quickly plateaus at around 20 months, and the highest adjusted R^2 achieved is only 3.38%, which is roughly a third of the

explanatory power of ratings. Put together, these results indicate that Morningstar ratings are much more important than past returns in determining fund flows. This is remarkable, especially as ratings are only on a coarse scale with five possible values.

B. Morningstar ratings account for fund return volatility, investors do not

As we have discussed in Section II, Morningstar ratings do not adjust for exposure to the market factor or other risk factors, however, they do account for a fund’s return volatility. Given that, as we have seen above, Morningstar ratings are paramount in determining flows across funds, one may wonder how ratings influence the relationship between flows and return volatility, which is known to be negative in the data. We carry out several tests to study this relationship.

We start by estimating the following regression model:

$$F_{p,t} = \xi \text{Ratings}_{p,t} + \phi \text{Vol}_{p,t}^5 + \pi \text{Vol}_{p,t}^1 + \psi \text{RET}_{p,t} + \gamma Y_{p,t} + \nu_{t,\text{style}} + \epsilon_{p,t}, \quad (13)$$

where $\text{Ratings}_{p,t}$ is the Morningstar rating, $\text{Vol}_{p,t}^5$ and $\text{Vol}_{p,t}^1$ are the monthly standard deviations of fund returns, as estimated over the prior 5 years and prior 1 year, respectively, $\text{RET}_{p,t}$ is the prior-18-month return weighted according to the exponential decay model in Equation (12), and $Y_{p,t}$ is a vector of controls that include the total expense ratio, a dummy variable for no-load, the log of fund size, the log of fund age prior to month t , market beta over the prior 5 years, and lagged fund flows from month $t - 19$. We also include time-style fixed effects, and the standard errors are double-clustered by time and by fund. The results are reported in Table VI.

In the first three specifications, where we do not control for Morningstar ratings, return volatility has a negative and statistically significant coefficient. This implies that controlling for the level of past returns, an increase in return volatility is, on average, associated with a decrease in fund flows in the next month. When included separately, both the one-year and the five-year volatility measures are significant. Interestingly, in Column (3), we find that the five-year volatility is statistically more important than the one-year volatility in predicting negative flows. This result is consistent with what we would expect to see if investors used Morningstar ratings to direct their flows, because Morningstar uses up to 10 years of past returns to assign the rating.

We conjecture that the negative effect of a fund’s volatility on future flows is not due to the fact that investors actually calculate a fund’s volatility and use that information to direct flows, but rather it is due to the fact that Morningstar takes volatility into account when assigning ratings. In Column (4) we confirm that controlling for a fund’s return, volatility is

Table VI **Response of fund flows to return volatility through Morningstar ratings.** This table presents coefficient estimates from panel regression of percentage fund flow on a fund's return volatility over the prior 5 years or 1 year in Equation (13). The t -statistics (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Flow	Flow	Flow	Ratings	Flow	Flow	Vol ⁵	Flow
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vol ⁵	-0.054*** (-4.85)		-0.041*** (-3.15)	-7.84*** (-8.81)	0.0044 (0.38)	-0.0059 (-0.50)		
Vol ⁵ _{predicted}								-1.88*** (-36.25)
Vol ⁵ _{residual}								0.0091 (0.85)
Vol ¹		-0.065*** (-3.49)	-0.035 (-1.61)	-7.10*** (-5.00)	0.012 (0.67)	-0.00069 (-0.041)		
Ratings					0.0060*** (34.85)		-0.0032*** (-11.77)	
Past Ret	0.18*** (12.28)	0.18*** (12.24)	0.18*** (12.34)	6.63*** (13.32)	0.14*** (11.15)	0.17*** (11.56)		0.14 (11.31)
Month-style FE	Yes	Yes	Yes	Yes	Yes	No	No	Yes
Month-rating FE	No	No	No	No	No	Yes	No	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Number of obs	257,053	257,053	257,053	257,053	257,053	257,053	257,053	257,053
Adjusted R ²	0.113	0.118	0.113	0.231	0.173	0.160	0.030	0.173

a significant negative predictor of fund ratings. This is, of course, consistent with how the Morningstar Risk-Adjusted Return is calculated (see Equation (1)). Based on this evidence, controlling for a fund’s rating seems imperative. Once we add the Morningstar ratings in the regression (Column (5)), return volatility loses its ability to predict future flows for both return horizons considered. Moreover, the coefficient estimates are so small that, even if the effects were statistically significant, the economic meaning would be negligible. On the other hand, the coefficient on the ratings is highly significant. A one-notch increase in ratings is on average associated with a 0.60% increase in fund flows in the next month. Alternatively, if we include month-rating fixed effects (Column (6)), the effect of fund return volatility on fund flows becomes again insignificant.

We estimate two additional regressions to help us interpret these results. The fact that, controlling for Morningstar ratings, volatility is no longer a significant predictor of fund flows might stem from two mutually-exclusive reasons. The first is that investors might want to account for fund return volatility when allocating across funds, but delegate the calculation of fund volatility to Morningstar. Alternatively, investors do not actually intend to account for fund return volatility, and the negative correlation between flows and volatility is only due to the fact that Morningstar’s formula takes volatility into account.

To evaluate which of the two potential explanations describes better investors’ behavior, we decompose return volatility into a component that is correlated with Morningstar ratings, and a component that is orthogonal to the ratings but may reflect investors’ preferences beyond the ratings. We focus on five-year volatility because, as seen in Column (3), it is more strongly related to fund flows than the one-year volatility; we find similar results if we use the one-year volatility instead. We execute this test in two stages. In the first stage, we run a regression of volatility on fund ratings and report the results in Column (6). Consistent with Column (4) and with the Morningstar Risk-Adjusted Return formula, the correlation is negative and highly significant. However, R^2 is only 3%, indicating that most of the cross-sectional dispersion in return volatility is unrelated to Morningstar ratings. In the second stage, reported in Column (7), we use the predicted value of volatility and the residual to explain fund flows. The estimates show very clearly that volatility has an effect on flows only through the ratings. The residual part, which represents 97% of the dispersion in volatility across funds, does not predict lower flows.

We carry out an additional test to further study the relation between flows, volatility, and ratings. In Table VII we split our sample into five subgroups based on Morningstar rating assignments. We estimate Equation (13) for each of the five rating groups. Consistent with the results presented in VI, within each rating group, return volatility is no longer negatively correlated with future flows at the conventional 5% confidence level.

Table VII **Response of fund flows to return volatility within Morningstar ratings groups.** This table presents coefficient estimates from panel regression of percentage fund flow on a fund’s return volatility over the prior 5 years or 1 year for each of the five Morningstar ratings groups. The t -statistics (double-clustered by fund and month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Rating 1	Rating 2	Rating 3	Rating 4	Rating 5
	(1)	(2)	(3)	(4)	(5)
Vol ⁵	-0.034* (-1.71)	0.010 (0.65)	0.025 (1.56)	-0.0068 (-0.29)	-0.068 (-1.49)
Vol ¹	0.024 (0.76)	0.069*** (2.74)	0.020 (0.86)	-0.037 (-1.23)	-0.049 (-0.99)
Past Ret	0.058*** (3.39)	0.12*** (8.34)	0.14*** (10.16)	0.18*** (8.60)	0.31*** (7.46)
Month-style FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Number of obs	17,024	60,416	92,131	60,613	18,279
Adjusted R^2	0.11	0.086	0.084	0.092	0.14

The correlation between fund flows and fund return volatility has little economic meaning for two reasons. First, the correlation between flows and return volatility is driven only by 3% of fund return volatility (which impact Morningstar ratings); the remaining that 97% of return volatility does not impact flows. Second, even before controlling for ratings, the impact of return volatility is economically insignificant, e.g., a one standard deviation increase in 1-year volatility is associated with a 0.07 percentage point decrease in monthly flows. For comparison, a one standard deviation increase in star ratings is associated with a 0.65 percentage point increase in monthly flows.

In summary, the results in Tables VI and VII suggest that investor flows are sensitive to fund return volatility only through the Morningstar ratings channel, and are ignoring 97% of return volatility (which is unrelated to the ratings). The impression from the combined evidence presented above is that it is unlikely that investors really intend to account for a fund’s volatility when allocating capital across funds. Rather, the negative correlation between fund flows and return volatility is most likely simply a byproduct of the fact that Morningstar accounts for a fund’s return volatility when assigning ratings.

VII. Conclusion

The key to understanding investor behavior and market prices is to understand how investors form beliefs. Two recent studies, Berk and van Binsbergen (2016) and Barber et al. (2016), took on the task to reconstruct investor beliefs by studying the drivers of mutual fund flows. The idea is that by allocating funds across active mutual funds, investors reveal their preferences and dislikes. Berk and van Binsbergen (2016) and Barber et al. (2016) used completely different empirical approaches to address the problem but reached the same conclusion: among the commonly-used asset pricing models in academia, mutual fund investors use the “simplest” model — the CAPM.

In this paper, we contrast the results of these studies with another line of research from the mutual fund literature that finds that mutual fund flows respond strongly to external rankings (e.g., Del Guercio and Tkac (2008), Kaniel and Parham (2017)). Our results show that mutual fund investors primarily follow external (Morningstar) ratings and then past returns. They pay no attention to whether past returns were generated by the systematic component of any of the commonly-used asset pricing models. We use the test proposed by Berk and van Binsbergen (2016) to show that Morningstar ratings dominate alphas from any other commonly-used asset pricing models. We also show that the tests run by Barber et al. (2016) are not robust to specification and thus are not conclusive. Finally, it is not plausible that Morningstar ratings serve as a proxy for alpha (of the CAPM or of another asset pricing model) since these ratings do not account for systematic exposure to any risk factor.

Where do our results leave the study of investor behavior and asset pricing? It is clear that mutual fund investors do not use any of the commonly-used asset pricing models for their investment decisions. Mutual fund flows indicate that investors’ revealed preferences are to pursue easy-to-follow signals (Morningstar ratings and past returns), which are ultimately not informative. If anything, the study of revealed preferences of mutual fund investors suggests that their behavior is irrational and uninformative about asset pricing.

REFERENCES

- Agarwal, Vikas, Clifton Green, and Honglin Ren, 2018, Alpha or beta in the eye of the beholder: What drives hedge fund flows, *Journal of Financial Economics* 127, 417–434.
- Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which factors matter to investors? Evidence from mutual fund flows, *Review of Financial Studies* 29, 2600–2642.

- Berk, Jonathan B., and Jules H. van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.
- Blake, Christopher R., and Matthew R Morey, 2000, Morningstar ratings and mutual fund performance, *Journal of financial and Quantitative Analysis* 35, 451–483.
- Blocher, Jesse, and Marat Molyboga, 2018, The revealed preference of sophisticated investors, *European Financial Management* forthcoming.
- Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.
- Chakraborty, Indraneel, Alok Kumar, Tobias Muhlhofer, and Ravi Sastry, 2018, Does limited investor attention explain mutual fund flows? evidence from sector funds, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2869426.
- Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.
- Christoffersen, Susan EK, David K Musto, and Russ Wermers, 2014, Investor flows to asset managers: Causes and consequences, *Annu. Rev. Financ. Econ.* 6, 289–310.
- Clifford, Christopher P., Jon A. Fulkerson, Bradford D. Jordan, and Steve Waldman, 2013, Risk and fund flows, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1752362.
- Coval, Joshua, and Eric Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.
- Del Guercio, Diane, and Paula A Tkac, 2008, Star power: The effect of morningstar ratings on mutual fund flow, *Journal of Financial and Quantitative Analysis* 43, 907–936.
- Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.
- Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 607–636.
- Franzoni, Francesco, and Martin C. Schmalz, 2017, Fund flows and market states, *Review of Financial Studies*, 30.
- Frazzini, Andrea, and Owen A. Lamont, 2008, Dumb money: Mutual fund flows and the cross-section of stock returns, *Journal of Financial Economics* 88, 299–322.

- French, Kenneth R., 2008, Presidential address: The cost of active investing, *Journal of Finance* 63, 1537–1573.
- Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from mutual fund industry, *Journal of Law and Economics* 35, 45–70.
- Jegadeesh, Narasimhan, and Chandra Sekhar Mangipudi, 2017, What do fund flows reveal about asset pricing models, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3090541.
- Kaniel, Ron, and Robert Parham, 2017, WSJ category kings - the impact of media attention on consumer and mutual fund investment decisions, *Journal of Financial Economics* 123, 337–356.
- Lynch, Anthony W., and David K. Musto, 2003, How investors interpret past fund returns, *Journal of Finance* 58, 2033–2058.
- Morey, Matthew R., 2003, The kiss of death: A 5-Star morningstar mutual fund rating?, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=455240.
- Pástor, Ľuboš, and Robert F. Stambaugh, 2002, Mutual fund performance and seemingly unrelated assets, *Journal of Financial Economics* 63, 315–349.
- Pástor, Ľuboš, and Robert F. Stambaugh, 2012, On the size of the active management industry, *Journal of Political Economy* 120, 740–781.
- Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.
- Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A Taylor, 2017, Do funds make more when they trade more?, *Journal of Finance* 72, 1483–1528.
- Reuter, Jonathan, and Eric Zitzewitz, 2015, How much does size erode mutual fund performance? a regression discontinuity approach, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1661447.
- Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Song, Yang, 2018, The mismatch between mutual fund scale and skill, Working paper, University of Washington, Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3065700.

Appendix A Other Exercises of BHO

A Robustness checks of Section V

We verify that our findings in Section V are robust to using different sub-samples. In Table VIII, we first reproduce Columns (2) to (9) in BHO’s Table 5 using panel regressions in Panel A, and then run the same regressions using Fama-MacBeth procedure in Panel B. The change of coefficients are then reported in Panel C. Specifications in Columns (1) and (2) use all funds but include different fixed effects from the main specification in Table IV. Columns (3) and (4) split the sample by median fund sizes; Columns (5) and (6) split the sample by median fund age, and Columns (7) and (8) split the sample by median fund return.

We conclude our main finding that fund flow sensitivity to the market-related return component is much higher under the Fama-MacBeth procedure. The increase is over 200% for all specifications except the small fund sample (Column (3)). The coefficient changes for other return components are smaller and similar in magnitude to our findings in Table IV.

B Horse-race results of BHO

BHO also conduct a nonlinear pairwise test of asset pricing models in their Table 4. We again summarize their methodology for the reader’s convenience. To compare two asset pricing models, in each period, funds are sorted into deciles using both models. Then BHO runs a panel regression with time fixed-effects on fund flows:

$$F_{p,t} = a + \sum_i \sum_j b_{i,j} D_{i,j,p,t} + cX_{p,t} + \mu_t + \epsilon_{p,t}, \quad (14)$$

where $D_{i,j,p,t}$ is a dummy variable indicating that fund p is ranked i th decile by model 1 and j th decile in model 2 (10th decile means the highest alpha), $X_{p,t}$ are a vector of controls, and μ_t are time fixed effects. The authors then compute test statistic $\hat{\theta} = \sum_{i < j} \hat{b}_{i,j} - \sum_{i > j} \hat{b}_{i,j}$ using the cases where the two models rank funds differently. If $\hat{\theta}$ is statistically larger than zero, then this indicates that flows are more responsible to the ranking by model 2 than model 1, and vice versa.

Table VIII **Robustness checks of the findings in Table IV.** Panel A reproduces the BHO panel regressions in Columns (2) to (9) in BHO's Table 5; Panel B estimates the same regressions using Fama-MacBeth procedure, and Panel C presents the change in coefficients.

Panel A: BHO Panel regressions								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All funds	All funds	Small funds	Big funds	Young funds	Old funds	Below-median ret	Above-median ret
ALPHA ^{7F}	0.79*** (26.30)	0.74*** (27.33)	0.84*** (25.55)	0.90*** (29.97)	0.92*** (25.50)	0.87*** (31.00)	0.70*** (19.47)	0.89*** (24.75)
MKTRET	0.21*** (4.14)	0.19*** (4.73)	0.24*** (4.21)	0.26*** (4.64)	0.25*** (4.42)	0.25*** (4.60)	0.16*** (2.81)	0.24*** (4.36)
SIZERET	0.69*** (12.92)	0.64*** (13.31)	0.53*** (8.67)	0.89*** (14.27)	0.73*** (11.69)	0.78*** (12.50)	0.71*** (11.33)	0.59*** (7.88)
VALRET	0.59*** (10.35)	0.57*** (10.72)	0.70*** (10.58)	0.65*** (9.80)	0.70*** (9.96)	0.65*** (10.20)	0.52*** (7.16)	0.68*** (9.63)
MOMRET	0.94*** (15.16)	0.85*** (17.02)	0.93*** (13.93)	1.11*** (15.58)	1.20*** (15.03)	1.00*** (16.60)	0.91*** (11.92)	1.00*** (13.32)
INDRET1	0.82*** (11.23)	0.84*** (11.17)	0.91*** (11.01)	0.92*** (10.89)	0.94*** (9.28)	0.92*** (11.92)	0.71*** (7.58)	0.95*** (9.64)
INDRET2	0.59*** (7.06)	0.63*** (7.08)	0.68*** (5.92)	0.71*** (6.61)	0.69*** (5.57)	0.70*** (6.81)	0.54*** (5.23)	0.74*** (5.74)
INDRET3	0.64*** (7.83)	0.48*** (5.91)	0.74*** (6.60)	0.66*** (7.30)	0.73*** (6.60)	0.68*** (7.13)	0.43*** (4.14)	0.78*** (7.61)
Month FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-cat FE	Yes	No						
Month-cat-rating FEs	No	Yes	No	No	No	No	No	No
Controls	Yes							
Observations	257,053	248,463	257,053		257,053		257,053	
Adj. R-squared	0.190	0.216	0.175		0.173		0.175	

Panel B: Fama-MacBeth regressions								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All funds	All funds	Small funds	Big funds	Young funds	Old funds	Below-median ret	Above-median ret
ALPHA ^{7F}	1.07*** (38.58)	0.89*** (33.13)	0.92*** (31.74)	1.10*** (39.31)	1.14*** (36.02)	1.02*** (37.40)	0.96*** (27.72)	1.19*** (32.64)
MKTRET	0.79*** (6.57)	0.75*** (6.82)	0.41*** (2.76)	1.00*** (7.67)	0.92*** (5.12)	0.81*** (6.82)	0.59*** (3.87)	1.08*** (6.41)
SIZERET	0.70*** (3.98)	0.56*** (3.80)	0.18 (1.04)	0.71*** (4.00)	-0.09 (-0.34)	0.79*** (4.19)	0.41*** (2.74)	0.43** (1.99)
VALRET	1.01*** (5.43)	0.68*** (4.94)	0.76*** (3.68)	1.04*** (6.38)	0.63** (2.45)	1.00*** (5.92)	0.60*** (4.36)	1.15*** (4.82)
MOMRET	0.75*** (2.78)	0.66** (2.52)	0.19 (0.58)	0.88*** (3.13)	0.46 (1.03)	0.77*** (3.26)	0.83** (2.48)	0.51 (1.54)
INDRET1	0.75*** (4.57)	0.64*** (4.09)	0.79*** (5.20)	0.76*** (3.79)	0.88*** (4.16)	0.91*** (4.84)	0.66*** (3.89)	1.12*** (5.16)
INDRET2	1.11*** (4.55)	0.90*** (4.35)	0.69** (2.52)	1.20*** (3.80)	0.57* (1.95)	1.17*** (4.20)	0.67* (1.85)	1.42*** (5.11)
INDRET3	1.15*** (3.66)	0.83*** (2.67)	0.87*** (2.63)	1.22*** (3.23)	0.99*** (2.69)	1.11*** (3.69)	0.81*** (2.98)	1.56*** (4.00)
Month FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-cat FE	Yes	No						
Month-cat-rating FE	No	Yes	No	No	No	No	No	No
Controls	Yes							
Observations	257,053	248,463	257,053		257,053		257,053	
Adj. R-squared	0.179	0.125	0.196	0.234	0.231	0.217	0.152	0.166

In Table IX, we reproduce the horse race between CAPM and the market-adjusted model. While BHO find that CAPM beats the market-adjusted model, we find that the outperformance of CAPM disappears once we use the Fama-Macbeth regression. This is consistent with our analysis in Section V that investors don't behave as if they adjust for market beta.

Panel C: Change in Coefficients								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	All funds	All funds	Small funds	Big funds	Young funds	Old funds	Below-median ret	Above-median ret
ALPHA ^{7F}	+35%	+21%	+9%	+22%	+24%	+18%	+37%	+33%
MKTRET	+283%	+286%	+75%	+291%	+266%	+219%	+263%	+344%
SIZERET	+2%	-13%	-65%	-20%	-112%	+1%	-43%	-28%
VALRET	+71%	+20%	+9%	+60%	-9%	+53%	+14%	+69%
MOMRET	-20%	-23%	-79%	-20%	-62%	-23%	-8%	-49%
INDRET1	-8%	-24%	-13%	-17%	-6%	-1%	-6%	+19%
INDRET2	+86%	+42%	+1%	+67%	-18%	+67%	+25%	+92%
INDRET3	+79%	+73%	+17%	+84%	+36%	+63%	+89%	+101%
Month FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-cat FE	Yes	No	No	No	No	No	No	No
Month-cat-rating FE	No	Yes	No	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	257,053	248,463	257,053		257,053		257,053	

Table IX Results of horse race between CAPM and Market-adjusted return.

	BHO panel regression with time FEs	Fama-Macbeth regression
	CAPM – MAR	CAPM – MAR
Sum of coefficient differences	7.41***	0.62
<i>t</i> -stat	(3.46)	(0.032)
% of coefficient difference > 0	77.78%	46.7%
Binomial <i>p</i> -value	< 1%	> 10%

Table X **Correlation between total fund return and its components.** In this table, we report the correlation between total fund return and the 8 components into which it is decomposed using in Equation (11). Every month, we calculate the cross-sectional Pearson correlation and the Spearman's rank correlation between the total fund return and the 8 return components. We report mean and median values for these correlation measures across the 175 months in the sample.

Correlation with total fund return				
	Pearson		Spearman	
	Mean	Median	Mean	Median
	(1)	(2)	(3)	(4)
ALPHA ^{7F}	0.70	0.76	0.71	0.72
MKTRET	0.07	0.06	0.04	0.05
SIZERET	0.21	0.23	0.22	0.22
VALRET	0.25	0.24	0.26	0.26
MOMRET	0.09	0.13	0.09	0.14
INDRET1	0.17	0.19	0.15	0.17
INDRET2	-0.04	-0.03	-0.03	-0.03
INDRET3	0.00	0.01	-0.01	0.02