What Do Mutual Fund Investors Really Care About?

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

October 7, 2018

(Very Preliminary, Please Do Not Circulate)

Abstract

Recent studies use 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 (Morningstar ratings) and chasing recent 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 only because Morningstar penalizes funds for high total volatility.

*We thank Brad Barber, Xing Huang, and Terrance Odean for kindly sharing their data. We thank Jonathan Berk, Justin Birru, Sylvester Flood (Morningstar), Francesco Franzoni, John Graham, Cam Harvey, David Hirshleifer, Steve Kaplan, Martin Schmalz, Berk Sensoy, Rick Sias, and seminar participants at The Ohio State University and Boston College for helpful comments. Ben-David is with The Ohio State University and NBER, Li is with Stanford University, Rossi is with the 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 academic debate in recent years. Financial economists have argued that studying investors' mutual fund choices can provide 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 using different empirical techniques.[1] Both 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 fund managers based on risk-adjusted returns is appealing, it is potentially at odds with empirical findings from the mutual fund literature documenting that investors respond to external rankings (Morningstar: Del Guercio and Tkac (2008), Reuter and Zitzewitz (2015), Wall Street Journal: Kaniel and Parham (2017), sustainability: Hartzmark and Sussman (2018)), chase past returns (Chevalier and Ellison (1997), Choi and Robertson (2018)), and display behaviors that may be considered suboptimal or unsophisticated.[2]

In this study, we attempt to reconcile the results from the two streams of the literature. Motivated by the fact that the vast majority of mutual fund assets is held by households,[3] we test whether simple and readily available signals explain investors behavior better than common asset pricing models. Specifically, we test whether Morningstar's star rating system explains mutual fund flows better than risk-adjusted returns. Morningstar ratings are the ideal candidate for our tests for several reasons. First, Morningstar is the leader of the US fund rating industry and its star ratings are often provided to investors by financial advisors, brokers, defined-contribution retirement plan sponsors and by fund companies themselves through their marketing material. The ratings can also be checked for free on Morningstar's website. Second, Morningstar ratings do not adjust for fund exposure to any systematic risk factor (see section II for an in-depth discussion). Third, these ratings are available for most of the mutual funds typically studied in the academic literature.

[1] Agarwal, Green, and Ren (2018a) and Blocher and Molyboga (2018) applied these empirical methods in the hedge fund space.

[2] Academic studies have found that mutual fund investors prefer funds that report holdings of recent winners and lottery stocks (Solomon, Soltes, and Sosyura (2014), Agarwal, Jiang, and Wen (2018b), Chuprining and Ruf (2018)), react to advertisements and media coverage that do not signal skill (Jain and Wu (2000) and Solomon et al. (2014)), generate 'dumb money' flows (Frazzini and Lamont (2008), Akbas, Armstrong, Sorescu, and Subrahmanyam (2015), Friesen and Nguyen (2018)) and make suboptimal retirement planning choices (Xiao, Zhang, and Kalra (2018)).

[3] According to the 2011 ICI Fact Book, at the end of 2010, 93.7% of long-term mutual fund assets in the US, i.e., equity and bond funds, were held by households. These assets were owned by 90.2 million US individuals.

Our results show that ratings are the main determinant of capital allocation across mutual funds, followed by past returns. We find no evidence that investors account for mutual fund exposure to the market and other risk factors. We also show that fund flows are weaker for high-volatility funds only because Morningstar penalizes funds for high total volatility.

In the first part of this article, we adopt the diagnostic test proposed by BvB and compare the performance of Morningstar ratings to that of alphas from asset pricing models in predicting mutual fund flows. BvB's test measures the degree of agreement 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 common asset pricing models agrees with the sign of fund flows 57.8% to 59.6% of the time, 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 time).

To further sharpen the BvB test, we also analyze the spread between flows to top and bottom funds ranked according to various asset pricing models or Morningstar ratings. In all tests, ratings decisively outperform all asset pricing models considered. At the aggregate level, funds rated highest by Morningstar received more money than the funds ranked highest according to any asset pricing model in every single year. Moreover, when using either fractional flows or dollar flows, the CAPM model no longer consistently outperforms other models in explaining flows, including raw return (the 'no-model' benchmark).

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 attributable to exposure to most risk factors, they react only weakly to returns originated from exposure to the market factor. BHO conclude that investors care about market risk and therefore discount returns that originate from exposure to the market risk.

Our analysis indicates that BHO's findings should be interpreted in a different way. Specifically, BHO's conclusion is based on a 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 with time fixed effects would convey the impression that flows respond less to the market-related component of fund returns even if investors do not use the CAPM.

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 the 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 the same to all components of past returns. As a possible interpretation of their results, BHO suggest that mutual fund investors are relatively unsophisticated because they do not distinguish between returns generated by managerial skill from returns due to exposure to factors such as size and momentum. We contribute to this debate by showing that these investors also fail to adjust for exposure to market risk, which supports the interpretation that they are unsophisticated.

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 invest according to external rankings and chase past returns. 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 120 lags of past monthly returns in our regressions, they only explain up to 5.4% of the variation in fund flows. In contrast, the most recent Morningstar rating explains 9.2% of the variation. Moreover, when included in the same regression, the incremental explanatory power (marginal R^2) of Morningstar ratings is more than twice as that of 120 lags of past monthly returns combined. This evidence 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 documented 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.[4] This result raises the important question of whether investors independently consider total volatility as a source of risk, or whether they rely solely on Morningstar rankings. We document that the latter is most likely to be 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 fund flows are related to volatility *only* through Morningstar ratings. Specifically, fund flows

[4]Clifford, Fulkerson, Jordan, and Waldman (2013) report that "net flows show aversion to risk", which they measure as fund volatility. In addition to 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.

are negatively related only with the 3% of the variation that is correlated with Morningstar ratings, and not with the remaining 97%.

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, relying heavily on past rankings to do so.

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), Del Guercio and Reuter (2014), and Song (2018), among many others.[5] 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 investors. To support their claim, they show that in the subsample of sector funds, where both market returns and sector-specific historical returns are presented to investors, flows treat sector-specific returns as a source of risk. Jegadeesh and Mangipudi (2017) contest the validity of the tests proposed by BvB. They assert that estimated alphas of simple factor models are less noisy than estimated alphas of complex models, and therefore are more likely to win a horse race test. For the same reason, they 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.

[5] See 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 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)). Morningstar ratings have been shown to have a strong independent influence on investors' flows (Del Guercio and Tkac (2008), Reuter and Zitzewitz (2015)).

Morningstar explains its rating method in a publicly-available manual.[6] 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 groups[7] 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 ($MRAR$):

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

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. Morningstar motivates this formula using expected utility theory.

[6] The Morningstar manual is available at `https://corporate.morningstar.com/US/documents/MethodologyDocuments/FactSheets/MorningstarRatingForFunds_FactSheet.pdf`

[7] An 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.

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.[8] Three-year star ratings are available for funds 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.[9] 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.

[8] See `http://quicktake.morningstar.com/DataDefs/FundRatingsAndRisk.html` for details.

[9] We thank the authors for generously sharing their data. The dataset of BvB ranges 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 fund flow literature, 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

$$\mathrm{F}_{p,t} = \frac{\mathrm{TNA}_{p,t}}{\mathrm{TNA}_{p,t-1}} - (1 + R_{p,t}). \tag{2}$$

This formula assumes that all flows in a given month take place at the end of that month. Here, $\mathrm{TNA}_{p,t}$ 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)	500.70	751.88	1293.52	2136.05	3460.10	606.97	1443.50
Fund age (years)	16.22	16.67	16.95	17.37	16.50	15.99	16.87
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%
Market beta	0.99	0.96	0.94	0.91	0.90	0.82	0.93
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
Fraction of positive flows	15.9%	19.4%	29.7%	49.3%	67.0%	36.7%	33.9%

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, to explain 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 $t-1$:

$$\begin{aligned} R_{p,\tau} - R_{f,\tau} &= a^{7\mathrm{F}}_{p,t} + b_{p,t}(\mathrm{MKT}_\tau - R_{f,\tau}) + s_{p,t}\mathrm{SMB}_\tau + h_{p,t}\mathrm{HML}_\tau \\ &\quad + u_{p,t}\mathrm{UMD}_\tau + \sum_{k=1}^{3}\gamma^k_{p,t}\mathrm{INDk}_\tau + \epsilon p,\tau, \quad \tau \in \{t-60,\ldots,t-1\}. \end{aligned} \tag{3}$$

Here, $R_{p,\tau}$ is the mutual fund return net of fees in month τ, $R_{f,\tau}$ is the risk-free interest rate,[10]

[10]The 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 $a_{p,t}^{7\mathrm{F}}$ 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:

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

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

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

where the decay parameter λ is estimated empirically from the relationship between flows and past returns and $\hat{\alpha}_{p,t}^{7\mathrm{F}}$ 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

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

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

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

We also calculate the weighted averages of the Fama-French three-factor alpha (Fama and French (1993)), $\mathrm{ALPHA}_{p,t}^{\mathrm{FF}}$, and the FFC four-factor alpha, $\mathrm{ALPHA}_{p,t}^{\mathrm{FFC}}$, 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) and then perform additional tests to study the relation between fund flows and past performance.

A. BvB's test

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 verse. 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 the 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 $\mathrm{ALPHA}^{\mathcal{M}}_{\mathrm{p,t}}$ denote the inferred return alpha under the asset pricing model $\mathcal{M}$. Notice that $\mathrm{ALPHA}^{\mathcal{M}}_{\mathrm{p,t}}$ 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 of the model-implied alphas. Following the method of BvB,[11] for a given asset pricing model $\mathcal{M}$, we run the following regression:

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

where $\mathrm{sign}(F_{p,t})$ and $\mathrm{sign}(\mathrm{ALPHA}^{\mathcal{M}}_{\mathrm{p,t}})$ 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

[11] Our tests differ slightly from those of 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 and is more consistent with the flow-performance literature.

the alpha and flow signs match each other. Specifically,

$$\frac{\beta_1^{\mathcal{M}}+1}{2} = \frac{\widehat{\Pr}(\text{sign}(F_{p,t}) = 1|\text{sign}(\text{ALPHA}_{p,t}^{\mathcal{M}}) = 1) + \widehat{\Pr}(\text{sign}(F_{p,t}) = -1|\text{sign}(\text{ALPHA}_{\text{p,t}}^{\mathcal{M}}) = -1)}{2}, \tag{9}$$

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

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 market-adjusted benchmark. Thus, they conclude that the CAPM is closest to the "true" model used by investors.

In our analysis, 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, the last row of Table I 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. Funds with ratings greater than 3, greater than 4, and equal to 5 comprise, respectively, 68.9%, 31.8%, and 7.4% of fund-month observations. 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 II.

Consistent with BvB's findings, our replication shows that the CAPM performs better than the market-adjusted model, the FF three-factor model, and the FFC four-factor model.[12] We also find that the excess return model (the return of the fund in excess of the risk-free rate) performs the worst. However, the rating-based heuristics significantly outperforms the CAPM and the other models, and the degree of outperformance is larger than the entire dispersion among the scores of all other models. The best-performing heuristics, which has investors reallocating money into five-star funds, gets the sign of the flows right approximately 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 the CAPM (60.4%) and the worst performing model (excess returns, 56.9%).

Is this outperformance of rating-based heuristics statistically significant? We follow BvB

[12] BvB also include some dynamic equilibrium models in their tests. In the original study, these models are generally dominated by the CAPM and by multifactor models, therefore we do not include them in our tests.

Table II **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 $(\beta_1^{\mathcal{M}}+1)/2$ in percent, and models are ordered in decreasing order of the point estimate of $\beta_1^{\mathcal{M}}$. The remaining columns provide statistical significance tests of the pairwise model horse races based on Equation (10). 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 $(\beta_1^{\mathcal{M}}+1)/2$	Univariate t-stat	Rating ≥ 4	Rating ≥ 3	CAPM	Market-adjusted	FF 3-factor	FFC 4-factor	Excess return
Rating ≥ 5	67.95	29.48	5.60	9.57	9.93	10.80	12.70	13.13	11.20
Rating ≥ 4	64.41	36.32	-	10.07	8.27	9.21	12.26	12.82	8.53
Rating ≥ 3	61.01	32.15	-	-	1.43	2.81	5.33	6.22	5.15
CAPM	60.36	25.62	-	-	-	5.59	6.53	7.22	4.67
Market-adjusted	59.64	22.64	-	-	-	-	3.07	4.06	3.68
FF 3-factor	58.76	26.44	-	-	-	-	-	3.37	2.61
FFC 4-factor	58.36	25.83	-	-	-	-	-	-	2.08
Excess return	56.90	11.78	-	-	-	-	-	-	-

to conduct pairwise model horse races. For any two models $\mathcal{M}1$ and $\mathcal{M}2$, we run regression

$$\text{sign}(F_{p,t}) = \gamma_0 + \gamma_1 \left(\frac{\text{sign}(\text{ALPHA}_{\text{p,t}}^{\mathcal{M}1})}{\text{var}(\text{ALPHA}_{\text{p,t}}^{\mathcal{M}1})} - \frac{\text{sign}(\text{ALPHA}_{\text{p,t}}^{\mathcal{M}2})}{\text{var}(\text{ALPHA}_{\text{p,t}}^{\mathcal{M}2})} \right) + \xi_{p,t} \tag{10}$$

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 II. 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 diagnostic, the test results suggest that Morningstar ratings explain investors' capital reallocation better than the CAPM and all other asset pricing models considered.

B. *Best- and worst-performing funds and other measures of flows*

The test proposed by BvB, i.e., analyzing the degree of agreement between the sign of a fund's alpha and the sign of the flow, is a theoretically-grounded application of the NPV rule. This particular test, however, focuses only on the signs of alphas and flows, and therefore likely disregards valuable information and might be susceptible to noise. In light of the results reported in Table II, this issue appears particularly important because many of the asset pricing models considered have similar scores. In this section, we carry out tests that exploit additional variation in fund performance and fund flows.[13]

We groups funds into best and worst performing using cutoffs based on the number of funds that are top rated and bottom rated by Morningstar. First, we focus on funds with extremely high and low rankings, i.e., 5-star and 1-star rated funds, respectively. Each month, we rank funds based on different measures of past performance, i.e., raw returns, CAPM alphas, etc. We then define top and bottom ranked funds for each of these measures based on the number of funds that have 5 stars and 1 star, respectively. For instance, if in a month there are 150 5-star rated funds, the 150 funds with the highest CAPM alpha are defined as being top-ranked according to the CAPM. On average, the fraction of fund-month observations that are defined as being top and bottom ranked is 7.4% and 6.9%, respectively. For each of these groups, we calculate the fraction of funds with positive flows, as well as the average fractional flows and the average dollar flows. The results are reported in Panel A of

[13] In a similar spirit, in one of their robustness tests, BvB restricted their sample to funds with extreme returns (see Table 9 of BvB). They did not, however, consider variation in fund flows other than the sign. In these robustness tests, they find that the CAPM performs better than the other models considered by a small margin, e.g., the CAPM score is around or less than 1 percentage point higher than the score for the return in excess of the market model

Table III **Flows to best and worst performing funds**

Panel A — **High ranked: five-star funds and the best 7.4% of funds for each model**
Low ranked: one-star funds and the worst 6.9% of funds for each model

	Fraction positive flow			**Fund flow (%)**			**Fund flow ($ Mn)**		
	High	**Low**	**Diff**	**High**	**Low**	**Diff**	**High**	**Low**	**Diff**
Morningstar	67.0%	15.9%	51.1%	1.15%	−1.53%	2.68%	37.3	−8.0	45.4
Market-adjusted	48.8%	25.9%	22.8%	0.29%	−1.19%	1.48%	11.6	−9.5	21.1
CAPM	43.9%	23.2%	20.6%	0.04%	−1.38%	1.41%	8.2	−10.6	18.7
FF 3-factor	41.0%	23.8%	17.2%	−0.11%	−1.31%	1.20%	5.3	−9.5	14.8
FFC 4-factor	40.3%	24.6%	15.6%	−0.16%	−1.26%	1.11%	4.1	−8.2	12.2

Panel B — **High ranked: 4- & 5-star funds and the best 31.8% of funds for each model**
Low ranked: 1- & 2-star funds and the worst 31.2% of funds for each model

	Fraction positive flow			**Fund flow (%)**			**Fund flow ($ Mn)**		
	High	**Low**	**Diff**	**High**	**Low**	**Diff**	**High**	**Low**	**Diff**
Morningstar	54.2%	19.3%	34.9%	0.41%	−1.29%	1.71%	12.7	−9.6	22.3
Market-adjusted	46.5%	24.6%	21.8%	0.10%	−1.09%	1.19%	8.5	−11.0	19.6
CAPM	45.9%	23.7%	22.2%	0.07%	−1.14%	1.21%	8.1	−11.7	19.8
FF 3-factor	44.3%	25.2%	19.0%	0.00%	−1.07%	1.08%	6.4	−9.9	16.3
FFC 4-factor	43.9%	25.6%	18.2%	−0.02%	−1.06%	1.04%	5.9	−9.3	15.2

Table III. In Panel B, we report the results of a similar test by classifying funds with 4 or 5 stars to be top ranked, and funds with 1 or 2 stars to be bottom ranked. In this case, the fraction of fund-month observations that is classified as being top-ranked and bottom-ranked under each model is 31.8% and 31.2%, respectively.[14]

The tests performed in Table III confirm the results in the BvB test that Morningstar ratings are the best predictor of fund flows. In particular, the spread between flows to the top-rated and the bottom-rated funds under Morningstar ratings is significantly high than that generated by all other asset pricing models, regardless of whether we use the sign of the flow, the fractional flow, or the dollar flow. On the contrary, the relative performance of the asset pricing models does vary across the different tests, suggesting that the ability of different asset pricing models to predict flows is not consistent across different definitions of flows.

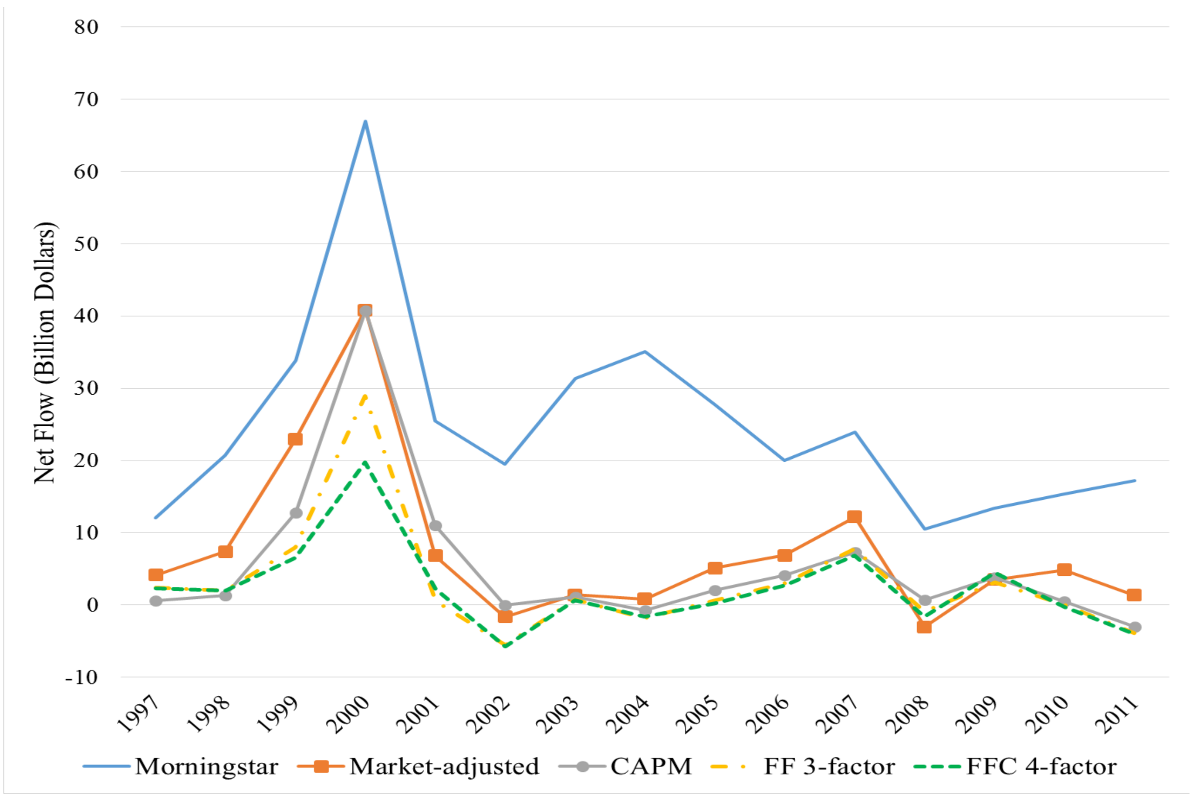
To provide additional insight, we also plot the annual aggregate net flows to the top-rated funds (based on 5-star rating) in Figure 1. There are two main takeaways. First, in each year, funds with top Morningstar ratings receive more inflows than funds that are deemed best-performing according to any of the asset pricing models considered, and the difference is economically large, i.e., on average, 20.3 billion dollars per year. Second, none of the asset pricing models considered appears to clearly outperform the others. For example, flows to funds that are ranked highest by the CAPM model and by the market-adjusted return model, which are the two best-performing asset pricing models in the BvB test of Table II, appear to move together and it does not seem that one model decisively dominates the other. Notice that, by construction, differences in rankings between these two models are driven by differences in fund's market betas, hence, this result is consistent with the idea that investors do not adjust for market beta - which is the main result of the analysis we present in the next section.

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

[14]Notice that, by construction, this test is cross-section in nature, and therefore rankings based on raw fund returns and on fund returns in excess of the risk-free rate and of the market return are equivalent. Therefore, we report the results for these ranking rules only once using the label 'market-adjusted.'

Figure 1: **Flows to best-performing funds.** This figure presents annual aggregate new flows to best-performing funds ranked according to five different measures of performance and according to the Morningstar rating system. Funds are ranked within each month, therefore, rankings based on raw fund returns and on fund returns in excess of the risk-free rate and of the market return are equivalent. For ease of exposition, we report the results for these ranking rules only once using the label 'market-adjusted'.



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 do not 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 a 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

$$\mathrm{MKTRET}_{p,t} = \frac{\sum_{s=1}^{18} e^{-\lambda(s-1)} \hat{b}_{p,t-s}(\mathrm{MKT}_{t-s} - R_{f,t-s})}{\sum_{s=1}^{18} e^{-\lambda(s-1)}}, \tag{11}$$

where $\hat{b}_{p,\tau}$ is the fund exposure to the market factor under the seven-factor model in Equation (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:

$$\begin{aligned} F_{p,t} = b_0 + & \quad \mu_t + \gamma X_{p,t} + b_{\mathrm{ALPHA}}\mathrm{ALPHA}^{\mathrm{7F}}_{p,t} + b_{\mathrm{MKT}}\mathrm{MKTRET}_{p,t} + b_{\mathrm{SMB}}\mathrm{SIZRET}_{p,t} \\ + & \quad b_{\mathrm{HML}}\mathrm{VALRET}_{p,t} + b_{\mathrm{MOM}}\mathrm{MOMRET}_{p,t} + \sum_{k=1}^{3} b_{\mathrm{IND}_k}\mathrm{INDRET}k_{p,t} + e_{p,t}, \end{aligned} \tag{12}$$

where $F_{p,t}$ is the monthly fund flow, μ_t is the time fixed effects 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 $b_{\text{ALPHA}}, b_{\text{MKT}}, \ldots$, 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 of returns when assessing mutual fund performance, implying that investors appear to be using the CAPM 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 simply 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 following extreme market return periods, both positive and negative. By construction, most of the cross-sectional variation of the market-related component of fund returns happens precisely in these periods,[15] when flows respond weakly to fund returns. As a consequence, when all periods are pooled together in the panel regression, it appears as if fund 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.

[15] The market-related component of returns is computed as the product of beta (which does not vary much over time) and the market return (see Equation ((11))). During periods of extreme market returns, this component has a large variation, since the Equation ((11)) multiplies beta by a larger number. Whatever cross-sectional variation there is beta, it is magnified once we multiply it by a large number.

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 (12). 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 MKTRET and the flow-response to other return components. 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 by 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***	0.63***	1.04***	0.24*	18%
	(32.74)	(10.15)	(39.70)	(1.96)	
MKTRET	0.25***	-	0.80***	-	216%
	(4.52)		(6.65)		
SIZERET	0.76***	0.51***	0.54***	−0.26	−29%
	(14.06)	(6.50)	(3.24)	(−1.27)	
VALRET	0.67***	0.42***	0.93***	0.13	40%
	(10.56)	(4.89)	(5.63)	(0.65)	
MOMRET	1.06***	0.81***	0.65**	−0.15	−38%
	(17.65)	(9.82)	(2.28)	(−0.47)	
INDRET1	0.92***	0.67***	0.76***	−0.04	−17%
	(12.43)	(7.19)	(4.91)	(−0.18)	
INDRET2	0.70***	0.45***	0.98***	0.18	40%
	(7.38)	(4.06)	(3.74)	(0.62)	
INDRET3	0.69***	0.44***	1.14***	0.34	64%
	(7.97)	(4.25)	(3.40)	(0.95)	
Month FE	Yes	-	-	-	-
Controls	Yes	-	Yes	-	-
Observations	257,053	-	257,053	-	-
Adjusted R^2	0.173	-	0.204	-	-

B. *Flow-performance sensitivity across 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 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, and report the average FPS per buckets in Figure 2.

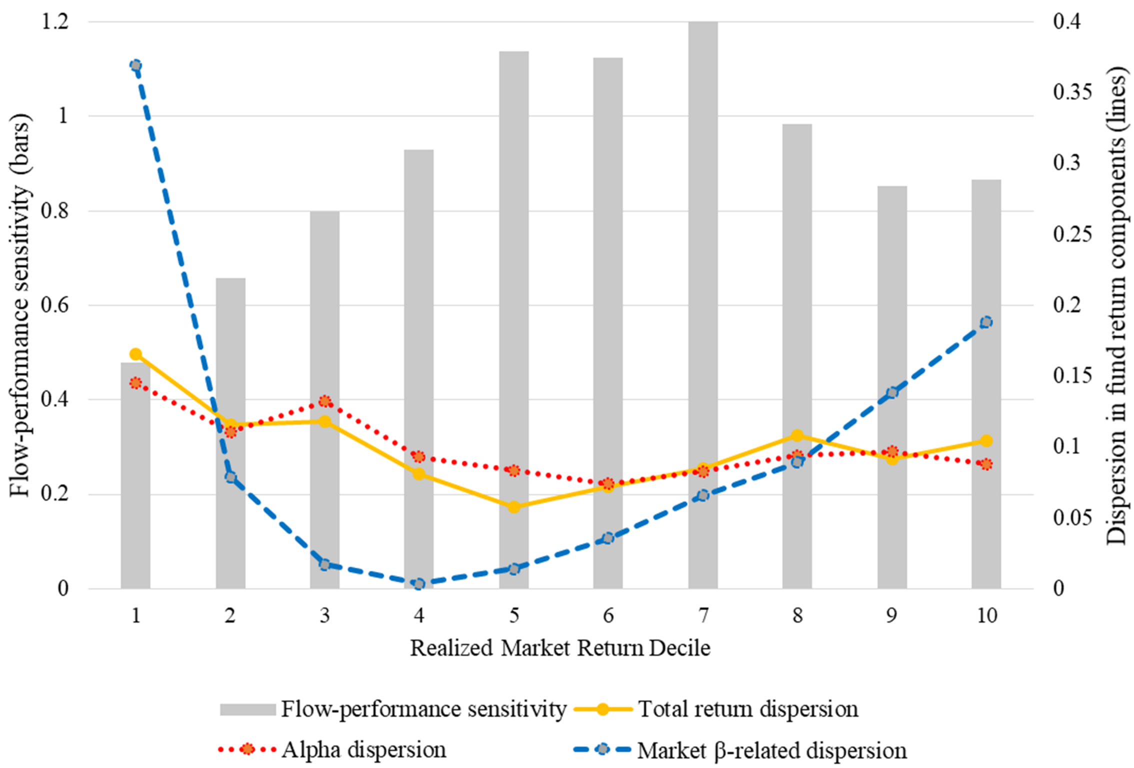
The figure shows that the FPS is a hump-shaped function of aggregate market realizations (left axis). This is consistent with the finding of Franzoni and Schmalz (2017). The FPS is more than twice as large in moderate states as in the states when the aggregate market has extremely negative returns. While 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 different market states.[16] 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 $b_{\mathrm{MKT}} = 0.25$ is likely downward-biased relative to other coefficient estimates in Equation (12).

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 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 (12), 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

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

Figure 2: **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.

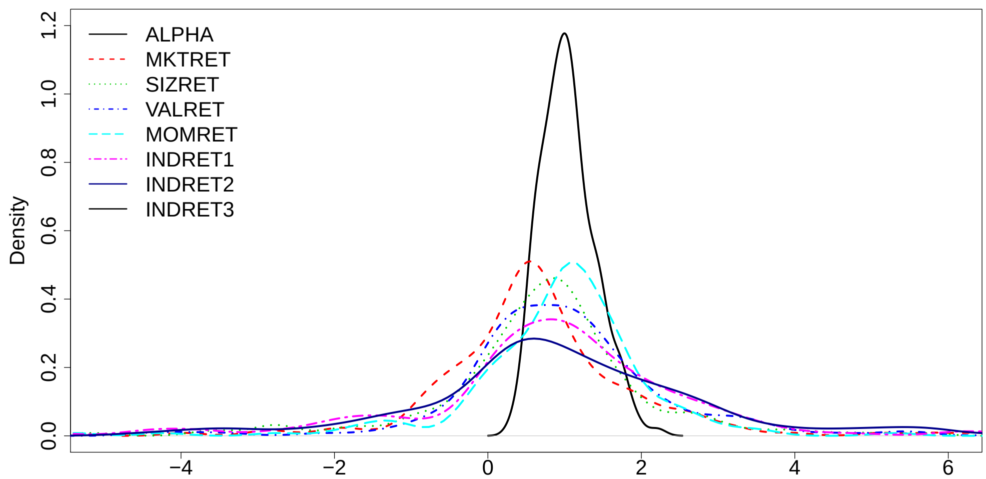


of the fund response to the market-related returns, which becomes more than three times as large (from 0.25 to 0.80). As shown in Column (5) of Table IV, the changes in the other coefficients are much smaller and have no clear pattern, with 3 out of 7 decreasing and the other 4 increasing. While the 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 3, 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, $\text{ALPHA}^{7\text{F}}$, 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 II 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 $\text{ALPHA}^{7\text{F}}$ 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 always highly correlated with cross-sectional differences in alphas, but this is not the case for the 7 factor-related returns. In Table A.II of Appendix A, we report summary statistics for the cross-sectional Spearman's rank correlation between total fund return and its eight components, i.e., $\text{ALPHA}^{7\text{F}}$, MKTRET, etc. Only alpha has a positive correlation with the fund return in all the months in the sample, ranging from a minimum of 0.34 to maximum of 0.94 and averaging 0.71. On the contrary, the factor-related return components are not always highly correlated with the total fund return. The average cross-sectional correlation between the factor-related components and the total fund return is between -0.03 and 0.26 and is negative in at least 10% of the months in the sample for all factor-related 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,

Figure 3: **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.



the cross-sectional dispersion in alphas is always relatively large and stable.[17]. 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 cannot 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 that of the other coefficients.

BHO also offer several robustness checks with different subsamples.[18] We repeat their 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 A.I 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.

[17]These patterns can also be observed in Figure 3

[18]BHO 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.

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 propose a simple test to whether investors use market beta to guide their investments. The logic of the test is if investors would care about market beta, then when the market has positive returns they would discount returns of high beta funds. This relation predicts that at times with positive market returns, the correlation between flows and beta is negative, controlling for observed returns. In contrast, when the market has negative returns investors who care about market beta would understand that funds with high beta have low returns because of their market exposure, and therefore would not penalize them with low flows. In other words, at times of negative market return, the relation between flows and beta should be positive, given the observed returns of funds.

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

$$F_{p,t} = \nu_t + \psi \mathrm{RET}_{p,t} + \phi \hat{\beta}_{p,t} + \xi \mathrm{Rating}_{p,t} + \gamma X_{p,t} + \epsilon_{p,t}, \tag{13}$$

where ν_t is the time fixed effect, $\mathrm{RET}_{p,t}$ 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 $t-1$, $\mathrm{Ratings}_{p,t}$ is the Morningstar rating, and $X_{p,t}$ is a vector of controls as in Equation (12). Standard errors are double-clustered by time and by fund. We also estimate the same equation by means of a Fama-Macbeth regression. Columns (1) and (4) show results for the entire sample whereas Columns (2)-(3) and (5)-(6) present results for subsamples based on the signs of the weighted average of the past 18-month returns of the aggregate market. The results are displayed in Table V.

The table suggests that controlling for past fund returns, market beta does not influence fund flows in a significant manner for either the entire sample or any of the subsamples. This result holds with both pooled regressions with time FEs and with the FM procedure. In other words, investors do not seem to adjust for market beta when they allocate flows among mutual funds. In Table A.IV of Appendix A, we show that the results are robust to controlling for month-style or month-style-rating fixed effects. 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. In 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 (13). The t-statistics (double-clustered by fund and by month) are in parentheses. *, **, and *** indicate significance at the 10% ,5%, and 1% level, respectively.

	Panel regression			**Fama-Macbeth regression**		
	All	**+MKT**	**−MKT**	**All**	**+MKT**	**−MKT**
	(1)	(2)	(3)	(4)	(5)	(6)
Weighted past return	0.68***	0.85***	0.53***	0.81**	0.93***	0.59***
	(23.39)	(24.32)	(16.48)	(31.11)	(29.07)	(20.87)
Market beta	−0.000080	−0.00039	0.00015	0.000085	0	0.00023
	(−0.30)	(−1.27)	(0.46)	(0.71)	(0.060)	(1.37)
Ratings	0.0049***	0.0051***	0.0043***	0.0047***	0.0050***	0.0042***
	(29.03)	(28.00)	(18.67)	(47.22)	(45.48)	(22.70)
Month FE	Yes	Yes	Yes	-	-	-
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs	257,053	167,936	89,117	257,053	167,936	89,117
Adjusted R^2	0.206	0.222	0.190	0.242	0.250	0.220

similar exercise in the Appendix A, we also find that investors don't discount size beta, value beta, and momentum beta.

VI. Morningstar Ratings are Primary Drivers of Mutual Fund 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 even 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 capital 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 category. 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)). In order to allow for flexible dependence of flows on past returns, we regress fund flows on the most recent Morningstar rating and 120 lags of past monthly returns:

$$F_{p,t} = b_0 + b_1 R_{p,t-1} + b_2 R_{p,t-2} + ... + b_{120} R_{p,t-120} + c \text{ Rating}_{p,t-1} + \epsilon_{p,t}. \tag{14}$$

We choose 120 lags because Morningstar ratings are computed using up to 10 years of historical data. Requiring all these lags to exist reduces the sample size to 136,707 which is 53% of the original dataset. We also run the same regression with only returns or ratings and reported the adjusted R^2 numbers in Table VI.

The first column of Table VI shows that even when including 120 lags flexibly, past returns only achieve 5.4% R^2, but one lagged Morningstar rating alone achieves 9.2%. A similar picture shows up when we look at marginal R^2 reported in the last two rows. When including ratings, returns add an additional 3.9% explanatory power; when including returns, ratings add 7.6% more.

Table VI **Explanatory power of ratings and past fund returns on fund flows**. The first three rows report the adjusted R^2 when regressing fund flows on different sets of regressors: 120 lagged monthly returns and one lagged ratings, just returns, or just ratings. The last two rows report the marginal adjusted R^2 of returns and ratings regressors. The first column is a pooled regression, and the next three columns add various fixed effects.

Specification	**Adjusted R^2**			
	No FE	**Month FE**	**Style FE**	**Month-Style FE**
Return and Ratings	0.130	0.155	0.132	0.168
Only past 120 returns	0.054	0.082	0.055	0.096
Only Rating	0.092	0.127	0.093	0.145
	Marginal Adjusted R^2			
Past 120 returns	0.039	0.028	0.039	0.024
Rating	0.076	0.073	0.077	0.072

The next three columns of Table VI show that ratings consistently beat past returns when we add month fixed effects, Morningstar 3×3 fund style (e.g. Small-value) fixed effects, or month-style fixed effects. 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 discuss 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 (e.g., Clifford et al. (2013)). We carry out several tests to study this relationship.

We start by estimating the following regression model:

$$F_{p,t} = b_0 + \xi \text{Rating}_{p,t} + \phi \text{Vol}^5_{p,t} + \pi \text{Vol}^1_{p,t} + \gamma Y_{p,t} + \nu_t + \epsilon_{p,t}, \tag{15}$$

where $\text{Ratings}_{p,t}$ is the Morningstar rating, $\text{Vol}^5_{p,t}$ and $\text{Vol}^1_{p,t}$ are the monthly standard deviations of fund returns, as estimated over the prior 5 years and prior 1 year, respectively, 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 fixed effects, and the standard errors are double-clustered by time and by fund. The results are reported in Table VII. We also estimate the regression model (15) with the Fama-Macbeth procedure or controlling for time-style fixed effects, and we get similar results in Tables A.V and A.VI of Appendix A.

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 all else equal, 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 research or calculate fund return volatility and use that

Table VII **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 (15). The t-statistics (double-clustered by fund and by 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	−0.055***		−0.050***	−7.26***	−0.0080	−0.011		
	(−4.55)		(−3.53)	(−9.65)	(−0.63)	−(0.86)		
$\text{Vol}^5_{\text{predicted}}$								−2.02***
								(−35.47)
$\text{Vol}^5_{\text{residual}}$								0.0037
								(0.31)
Vol^1		−0.049***	−0.012	−5.31***	0.027	0.0166		
		(−2.74)	(−0.54)	(−4.18)	(1.56)	(0.85)		
Ratings					0.0064***		−0.0032***	
					(33.81)		(−11.77)	
Month 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	0.067	0.066	0.068	0.22	0.14	0.15	0.030	0.14

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 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 0.64% 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, the 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 VIII we split our sample into five subgroups based on Morningstar rating assignments. We estimate Equation (15) for each of the five rating groups. Consistent with

Table VIII **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 by 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	−0.026	−0.016	0.015	−0.025	−0.043
	(−1.35)	(−1.01)	(0.88)	(−0.99)	(−0.99)
Vol^1	−0.0098	0.058***	0.027	0.0073	−0.030
	(−0.35)	(2.61)	(1.26)	(0.26)	(−0.66)
Month 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	0.71	0.050	0.049	0.054	0.086

the results presented in VII, within each rating group, return volatility is no longer negatively correlated with future flows at the conventional 5% confidence level.

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 is 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 VII and VIII and in Tables A.V and A.VI of Appendix A 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 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 recent 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 about systematic risk or managerial skill. Using the same logic that guided Berk and van Binsbergen (2016) and Barber et al. (2016), we can conclude that neither the CAPM, nor any of the asset pricing models that are commonly-used in academia, are close to the asset pricing model investors are actually using.

REFERENCES

Agarwal, Vikas, Clifton Green, and Honglin Ren, 2018a, Alpha or beta in the eye of the beholder: What drives hedge fund flows, *Journal of Financial Economics* 127, 417–434.

Agarwal, Vikas, Lei Jiang, and Quan Wen, 2018b, Why do mutual funds hold lottery stocks?, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3164692.

Akbas, Ferhat, Will J. Armstrong, Sorin Sorescu, and Avanidhar Subrahmanyam, 2015, Smart money, dumb money, and capital market anomalies, *Journal of Financial Economics* 118, 355–382.

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* 23, 839–872.

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.

Choi, James J., and Adriana Z. Robertson, 2018, What matters to individual investors? Evidence from the horse's mouth, Working paper, Yale University.

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.

Chuprining, Oleg, and Thomas Ruf, 2018, Rent seeking and mutual fund inceptions, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3189986.

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 Jonathan Reuter, 2014, Mutual fund performance and the incentive to generate alpha, *Journal of Finance* 69, 1673–1704.

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.

Friesen, Geoffrey C., and Viet Nguyen, 2018, The economic impact of mutual fund investor behaviors, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3160271.

Hartzmark, Samuel M., and Abigail Sussman, 2018, Do investors value sustainability? A natural experiment examining ranking and fund flows, Working paper, University of Chicago.

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

Jain, Prem C., and Joanna S. Wu, 2000, Truth to mutual fund advertising: Evidence on future performance and fund flows, *Journal of Finance* 55, 937–958.

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, Mattew 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.

Solomon, David H., Eugene Soltes, and Denis Sosyura, 2014, Winners in the spotlight: Media coverage of fund holdings as a driver of flows, *Journal of Financial Economics* 113, 53–72.

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.

Xiao, Liu, Wei Zhang, and Ajay Kalra, 2018, The costly zero bias in target retirement fund choice, Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3214811.

Appendix A Additional Results

A Robustness checks of Section V

We verify that our findings in Section V are robust to using different fixed effects specifications and sub-samples. In Table A.I, 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 the 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.

These tests confirm our main finding that the sensitivity of flows to the market-related return component is much higher when using the Fama-MacBeth procedure. The coefficients on the market-related component increase by over 200% for all specifications except for 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:

$$F_{p,t} = a + \sum_i \sum_j b_{i,j} D_{i,j,p,t} + cX_{p,t} + \mu_t + \epsilon_{p,t}, \tag{16}$$

where $D_{i,j,p,t}$ is a dummy variable indicating that fund p is ranked ith decile by model 1 and jth 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.

In Table A.III, 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 do not behave as if they adjust for

Table A.I **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***	0.74***	0.84***	0.90***	0.92***	0.87***	0.70***	0.89***
	(26.30)	(27.33)	(25.55)	(29.97)	(25.50)	(31.00)	(19.47)	(24.75)
MKTRET	0.21***	0.19***	0.24***	0.26***	0.25***	0.25***	0.16***	0.24***
	(4.14)	(4.73)	(4.21)	(4.64)	(4.42)	(4.60)	(2.81)	(4.36)
SIZERET	0.69***	0.64***	0.53***	0.89***	0.73***	0.78***	0.71***	0.59***
	(12.92)	(13.31)	(8.67)	(14.27)	(11.69)	(12.50)	(11.33)	(7.88)
VALRET	0.59***	0.57***	0.70***	0.65***	0.70***	0.65***	0.52***	0.68***
	(10.35)	(10.72)	(10.58)	(9.80)	(9.96)	(10.20)	(7.16)	(9.63)
MOMRET	0.94***	0.85***	0.93***	1.11***	1.20***	1.00***	0.91***	1.00***
	(15.16)	(17.02)	(13.93)	(15.58)	(15.03)	(16.60)	(11.92)	(13.32)
INDRET1	0.82***	0.84***	0.91***	0.92***	0.94***	0.92***	0.71***	0.95***
	(11.23)	(11.17)	(11.01)	(10.89)	(9.28)	(11.92)	(7.58)	(9.64)
INDRET2	0.59***	0.63***	0.68***	0.71***	0.69***	0.70***	0.54***	0.74***
	(7.06)	(7.08)	(5.92)	(6.61)	(5.57)	(6.81)	(5.23)	(5.74)
INDRET3	0.64***	0.48***	0.74***	0.66***	0.73***	0.68***	0.43***	0.78***
	(7.83)	(5.91)	(6.60)	(7.30)	(6.60)	(7.13)	(4.14)	(7.61)
Month FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-style FE	Yes	No	No	No	No	No	No	No
Month-style-rat FEs	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	
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***	0.89***	0.92***	1.10***	1.14***	1.02***	0.96***	1.19***
	(38.58)	(33.13)	(31.74)	(39.31)	(36.02)	(37.40)	(27.72)	(32.64)
MKTRET	0.79***	0.75***	0.41***	1.00***	0.92***	0.81***	0.59***	1.08***
	(6.57)	(6.82)	(2.76)	(7.67)	(5.12)	(6.82)	(3.87)	(6.41)
SIZERET	0.70***	0.56***	0.18	0.71***	−0.09	0.79***	0.41***	0.43**
	(3.98)	(3.80)	(1.04)	(4.00)	-(0.34)	(4.19)	(2.74)	(1.99)
VALRET	1.01***	0.68***	0.76***	1.04***	0.63**	1.00***	0.60***	1.15***
	(5.43)	(4.94)	(3.68)	(6.38)	(2.45)	(5.92)	(4.36)	(4.82)
MOMRET	0.75***	0.66**	0.19	0.88***	0.46	0.77***	0.83**	0.51
	(2.78)	(2.52)	(0.58)	(3.13)	(1.03)	(3.26)	(2.48)	(1.54)
INDRET1	0.75***	0.64***	0.79***	0.76***	0.88***	0.91***	0.66***	1.12***
	(4.57)	(4.09)	(5.20)	(3.79)	(4.16)	(4.84)	(3.89)	(5.16)
INDRET2	1.11***	0.90***	0.69**	1.20***	0.57*	1.17***	0.67*	1.42***
	(4.55)	(4.35)	(2.52)	(3.80)	(1.95)	(4.20)	(1.85)	(5.11)
INDRET3	1.15***	0.83***	0.87***	1.22***	0.99***	1.11***	0.81***	1.56***
	(3.66)	(2.67)	(2.63)	(3.23)	(2.69)	(3.69)	(2.98)	(4.00)
Month FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Month-style FE	Yes	No	No	No	No	No	No	No
Month-style-rat 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	
Adj. R-squared	0.179	0.125	0.196	0.234	0.231	0.217	0.152	0.166

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-style FE	Yes	No	No	No	No	No	No	No
Month-style-rat 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 A.II **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 (12). Every month, we calculate the cross-sectional Spearman's rank correlation between the total fund return and the 8 return components. We report the mean, minimum, 10^{th} percentile, median, 90^{th} percentile, and maximum values of the correlation measure across the 175 months in the sample.

	Spearman's rank correlation with total fund return						
	Obs (months)	Mean	Min	P10	Median	P90	Max
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ALPHA^{7F}	175	0.71	0.34	0.51	0.72	0.85	0.94
MKTRET	175	0.04	-0.50	-0.29	0.05	0.34	0.55
SIZERET	175	0.22	-0.54	-0.06	0.22	0.58	0.81
VALRET	175	0.26	-0.47	-0.16	0.26	0.69	0.81
MOMRET	175	0.09	-0.59	-0.32	0.14	0.41	0.59
INDRET1	175	0.15	-0.51	-0.19	0.17	0.46	0.63
INDRET2	175	-0.03	-0.46	-0.35	-0.03	0.27	0.53
INDRET3	175	-0.01	-0.49	-0.26	0.02	0.20	0.53

Table A.III **Results of horse race between CAPM and Market-adjusted return (MAR).**

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

market beta.

C Robustness checks of Tables V and VII

We first show that the result in Table V of Section V.C that investors do not discount for market beta is robust to controlling for month-style fixed effects or controlling for month-style-rating fixed effects. That is, we estimate Equation (13) with month-style FEs or month-style-rating FEs. The results are presented in Table A.IV. As one can see, market beta is not a significant determinant of fund flows under these alternative specifications.

We also estimate Equation (15) by the Fama-Macbeth procedure or controlling for month-style fixed effects, and we report the results in Tables A.V and A.VI, respectively. We confirm the results in Table VII that the negative correlation between fund flows and return volatility is a byproduct of the fact that Morningstar accounts for a fund's return volatility when assigning ratings.

Table A.IV **Response of fund flows to market beta: controlling for month-style FEs and month-style-rating FEs.** This table presents coefficient estimates from panel regressions of percentage fund flow on past returns and market beta in Equation (13). The t-statistics (double-clustered by fund and by month) are in parentheses. *, **, and *** indicate significance at the 10% ,5%, and 1% level, respectively.

	All	**+MKT**	**−MKT**	**All**	**+MKT**	**−MKT**
	(1)	(2)	(3)	(4)	(5)	(6)
Weighted past return	0.71***	0.88***	0.55***	0.74***	0.91***	0.56***
	(23.99)	(23.11)	(17.61)	(24.37)	(23.33)	(17.78)
Market beta	0.000013	−0.00014	0.000051	−0.000047	−0.00025	0
	(0.048)	(−0.45)	(0.16)	(−0.18)	(−0.84)	(−0.028)
Ratings	0.0049***	0.0052***	0.0042***			
	(28.42)	(28.00)	(18.45)			
Month-style FE	Yes	Yes	Yes	No	No	No
Month-style-rat FE	No	No	No	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Number of obs	257,053	167,936	89,117	257,053	167,936	89,117
Adjusted R^2	0.22	0.23	0.20	0.23	0.24	0.21

Table A.V **Response of fund flows to return volatility through Morningstar ratings: Fama-Macbeth procedure** This table presents coefficient estimates from the Fama-Macbeth regression of percentage fund flow on a fund's return volatility over the prior 5 years or 1 year in Equation (15). The *t*-statistics (double-clustered by fund and by month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Flow	Flow	Flow	Ratings	Flow	Flow
	(1)	(2)	(3)	(4)	(5)	(6)
Vol^5	−0.046***		−0.030***	−8.20***	0.020**	
	(−4.47)		(−2.97)	(−16.89)	(2.06)	
$\text{Vol}^5_{\text{predicted}}$						−2.02***
						(−49.72)
$\text{Vol}^5_{\text{residual}}$						0.036***
						(3.41)
Vol^1		−0.054***	−0.024	−5.74***	0.022	
		(−3.36)	(−1.29)	(−4.83)	(1.44)	
Ratings					0.0065***	
					(43.28)	
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	0.091	0.094	0.098	0.21	0.17	0.17

Table A.VI **Response of fund flows to return volatility through Morningstar ratings: controlling for time-style FEs.** 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 (15). The t-statistics (double-clustered by fund and by month) are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level, respectively.

	Flow	Flow	Flow	Ratings	Flow	Flow
	(1)	(2)	(3)	(4)	(5)	(6)
Vol^5	−0.057***		−0.052***	−8.24***	−0.0022	
	(−4.90)		(−3.84)	(−9.34)	(−0.19)	
$\text{Vol}^5_{\text{predicted}}$						−1.96***
						(−36.86)
$\text{Vol}^5_{\text{residual}}$						0.0094
						(0.85)
Vol^1		−0.052***	−0.014	−6.33***	0.031	
		(−2.52)	(−0.60)	(−4.26)	(1.53)	
Ratings					0.0063***	
					(35.54)	
Month-style FE	Yes	Yes	Yes	Yes	Yes	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	0.098	0.097	0.098	0.22	0.16	0.16