Flow Hedging and Mutual Fund Performance*

Du Nguyen†

July 2023

Abstract

This paper studies risk-taking behavior of active mutual funds with respect to flow risk and its implications for fund performance. Recent evidence suggests that shocks to the common component of fund flows are a priced risk factor in expected stock returns. I find that nearly half of U.S. active equity funds tilt their portfolios toward stocks with higher exposure to common flows, suggesting that many funds do not hedge against flow risk. A rational model in which informed managers receive more precise private signals about common flows provides an explanation for this behavior. Using a holdings-based measure of flow risk management, I confirm the model's main prediction that skilled funds have higher exposure to common flows: funds in the top decile of the measure outperform those in the bottom decile by 5% annually in the data. Overall, the paper identifies a new aspect of flow risk management among active equity funds.

*I thank Michael O'Doherty for helpful comments and suggestions. All errors are my own.

†Du Nguyen is a Ph.D. candidate in finance in the Trulaske College of Business, University of Missouri, Columbia, MO 65211, USA. E-mail: ddnhw5@umsystem.edu.

1 Introduction

A vast literature in finance research studies risk-taking behavior of mutual funds.[1] Prior studies document a convex flow-performance relation, which attribute to fund managers' tendency to increase risk for more flows (see, e.g., Chevalier and Ellison, 1997, Sirri and Tufano, 1998). Recent studies examine whether funds strategically adjust risk to avoid flow volatility. Fund managers might want to avoid flow risk because volatile flows can impair fund performance (e.g., Rakowski, 2010), or outflows induced by extreme performance of portfolio holdings can reduce fee revenue (e.g., Di Maggio, Franzoni, Kogan, and Xing, 2023). Notably, Dou, Kogan, and Wu (2023) show that equity funds hedge against flow risk by tilting away from stocks that are more likely to have poor performance during systematic outflows. This flow-hedging behavior predicts (i) stocks with high flow risk have higher expected return, and (ii) funds have lower expected return relative to the market.[2] While Dou, Kogan, and Wu (2023) provide extensive empirical evidence to support the first prediction, this paper examines the empirical evidence of the second.

The latter remains puzzling for two reasons. First, it challenges prior evidence that there are active equity funds that outperform the market (Pástor and Vorsatz, 2020, Bessembinder, Cooper, and Zhang, 2023).[3] Second, it remains unclear why all funds do not exploit the risk premium associated with stocks that have higher exposure to common flow risk. In my sample period, the annual average return of high-flow-risk stocks is 4% more than that of low-flow-risk stocks. Foregoing this opportunity does not appear to align with active funds' objective to add value over their clients' alternative investment opportunity set. Thus, this paper seeks to provide insights about the flow-hedging behavior and its performance consequences in the cross-section of active funds.

I first document a significant variation in the flow-hedging behavior among U.S. domestic active equity funds. Figure 1 shows the distribution of the tilting coefficient from the regression of the deviation of a fund' portfolio weights from their market weights on underlying stocks' exposure to

[1]See Brown, Harlow, and Starks, 1996, Chevalier and Ellison, 1997, Sirri and Tufano, 1998, Huang, Wei, and Yan, 2007, Chen and Pennacchi, 2009, Huang, Sialm, and Zhang, 2011, Lee, Trzcinka, and Venkatesan, 2019.

[2]The first prediction arises because a common component of fund flows acts as a state variable that prices the cross-section of stocks. For the second prediction, I use a simple example for illustration. Suppose that the market portfolio contains only two stocks A and B, and stock A hedges against flow risk while stock B does not. Thus, stock A has lower expected return compared to stock B. If a fund overweights stock A relative to its optimal market weight due to flow hedging, we expect the fund to underperform the market.

[3]While on average active mutual funds deliver negative risk-adjusted net returns (see, e.g., Fama and French, 2010), not all funds underperform the market. For example, Bessembinder et al. (2023) show that almost a third of U.S. domestic equity mutual funds outperform the S&P500 market benchmark after fees over their lifetime.

flow risk, or stocks' flow beta. Flow beta capture stocks' exposure to common flows, which are the first principal component of flows across all active funds (Dou et al., 2023). The red line illustrates the estimated coefficient obtained using the aggregate mutual fund portfolio. Consistent with Dou et al. (2023), the aggregate mutual fund hedges against flow risk by tilting toward stocks with low flow beta. However, the distribution shows that there is a wide heterogeneity in the flow-hedging behavior across funds. The almost symmetrical distribution and zero mean suggest that half of the active funds tilt their portfolio toward stocks with high flow beta, and thus do not appear to hedge against flow risk.

I rationalize this empirical finding in an extended model of Kacperczyk and Seru (2007), which features informed and uninformed investors who differ in the precision of private information they receive about future flows. The model's main prediction is that a fund can increase its holdings of high-flow-beta stocks relative to other funds if it has more accurate private information on the common fund flows. Since flows in this model come from common sources (Dou et al., 2023), private signals about flows can be thought of as information on market-wide shocks that drive flows in and out of the equity market.[4] Such information can come from funds' ability to predict the market's demand for equity assets, which stems from changing investment opportunities and macroeconomic fundamentals. This conjecture further implies that funds who deviate away from their benchmark in a positive relation with flow beta are more likely to be skilled funds. Subsequently, the empirical prediction that I seek to verify in the data is that these funds should outperform funds that deviate less with respect to flow beta.

To capture the extent to which active equity funds manage flow betas, I use fund holdings data and construct an empirical measure AFB (Active Flow Beta), defined as the covariance between deviations of a fund's portfolio weights from the market portfolio and its holdings' flow betas. The interaction between deviations in portfolio weights and flow betas is an important feature of the measure because the model's main prediction is that the higher a stock's flow beta is, the larger the tilt is if the manager is skillful. From this perspective, AFB not only captures the response of a portfolio's holdings to flow risk but also measures the extent to which the fund manager actively manages exposure to common flow risk.

[4] Dou et al. (2023) model endogenous common flows, driven by exogenous macroeconomic shocks such as uncertainty. Since this paper's focus is on the flow-hedging behavior, I assume common flows are exogenous.

Using the holdings data for a sample of over 4,000 unique U.S. domestic active equity funds from 1994 to 2021, I estimate AFB for each fund and quarter. I find that AFB strongly predicts fund performance in subsequent quarters. In a univariate portfolio sort, I document that funds in the top decile of the AFB outperform those in the bottom decile by 0.43% monthly (or 5.16% annually) in net return. After adjusting for risk exposure to Carhart's (1997) four factors, the differential is 0.60% monthly (or 7.20% annually) and statistically significant at the 1% level. Controlling for Pástor and Stambaugh's (2003) liquidity risk factor does not affect the outperformance of high AFB funds. More importantly, AFB is a persistent predictor as top decile funds continue to outperform for up to four years after portfolio formation.

The predictive power of AFB is over and above other fund characteristics that have been shown to predict subsequent fund performance in the literature. In a series of double sorts on AFB and other fund performance predictors, including *Return gap* (Kacperczyk, Sialm, & Zheng, 2008), *Reliance on public information* (Kacperczyk & Seru, 2007), *Active share* (Cremers & Petajisto, 2009), *Active fundamental performance* (Jiang & Zheng, 2018), and *Active fund overpricing* (Avramov, Cheng, & Hameed, 2020), the Carhart four-factor alphas that AFB-based strategy deliver are substantial and statistically significant. For example, *Return gap* is the most consistent predictor across all subgroups of AFB portfolios, but funds in the top quartile of *Return gap* only have a four-factor alpha of 0.24% per month more than that of funds in the bottom quartile. The similar strategy based on AFB sorts delivers a four-factor alpha of 0.45% per month.

In multivariate panel regressions, I continue to document the consistent power of AFB in predicting subsequent fund performance. Controlling for standard fund characteristics that can affect future performance, such as fund size, age, expense ratios, turnover, past flow and past performance, the predictive coefficient of four-factor alpha on AFB remains economically large and statistically significant. For example, a 1-standard-deviation increase in AFB is associated with an increase of 2.32 basis points in average monthly four-factor alpha over the subsequent quarter and the effect is statistically significant at the 1% level. Controlling for other fund predictors as in the double sorts, AFB retains its predictive ability. In the strictest specification with all other fund characteristics and predictors, the predictive coefficient on AFB is economically large and statistically significant at 1% level.

In the last empirical tests, I verify an important prediction from the model to gain more insights about funds' hedging behavior. The model predicts that while the skilled funds deviate from the benchmark positively in the direction of flow beta, they do less so for stocks whose public information is more noisy. In other words, conditional on a stock's flow beta, the more imprecise the public information is, the less a skilled fund tilts away from the benchmark. To test this hypothesis, I construct two measures of public information precision based on the prior literature. Particularly, I use the dispersion in analysts' earnings forecasts and stocks' idiosyncratic volatility to proxy for the volatility of public information. Consistent with the model's prediction, the coefficient on the interaction between flow beta and both disagreement measures is significantly smaller among the high-flow-beta funds. The difference is economically large and statistically significant at the 5% level. This result supports the notion that high AFB funds deviate less from the benchmark when the public information on the stock is less precise.

Finally, I find that the ability of funds to actively manage flow betas is more valuable during times of high public disagreement. I construct the aggregate version of the two public disagreement measures following Huang, Li, and Wang (2021), and document that funds with high AFB perform significantly better during periods of high public disagreement. For example, the high AFB funds earn an additional 0.73% net return on average on months of extreme disagreement among analysts. This return difference between the high- and low-flow-beta funds is economically large and statistically significant. On average, high-flow-beta funds earn 1.71% more than low-flow-beta funds on months when analysts have the most diverse opinions, even after adjusting for risk exposure to Carhart's (1997) four factors.

My contribution to the literature on mutual funds is two-fold. First, I show that there exists a significant heterogeneity of flow hedging across U.S. active equity funds. Almost half of active funds appears to tilt toward high-flow-beta stocks, an empirical finding that is not explicit from Dou et al.'s (2023) theoretical model and empirical findings. I provide both theoretical arguments and empirical results to rationalize this finding by showing that skilled funds who might have private information about future flows may not engage in flow hedging. Second, I add to the broad literature on the mutual fund performance by establishing a measure that is informative about fund performance. I show that active management of flow betas strongly predict subsequent fund performance and its predictive ability cannot be subsumed by other persistent fund predictors.

The remainder of the paper is organized as follows. Section 2 describes the model and generates testable predictions. Section 3 describes the data, and the empirical construction of AFB. Section 4 examines the predictive ability of AFB for fund performance, and Section 5 concludes.

2 The Model

In this section, I extend Kacperczyk and Seru's (2007) model to show that the precision of the private signal about asset flows that an informed investor receives can explain her higher holdings of assets with high flow beta relative to an uninformed investor. The key intuition is that an asset' future flows are correlated positively with its future payoffs, making an informed investor who has more precise information about future flows to invest more in assets with high flow beta.

2.1 Simple Model of Flows and Payoffs

The standard model is an information economy with two periods in which investors make asset allocation decisions today, and receive payoffs from these assets tomorrow. Investors also receive an exogenous flow tomorrow.[5] In the model of Dou et al. (2023), flows are endogenously driven by aggregate exogenous shocks (e.g., economic uncertainty). However, the model presented here is agnostic about the sources of flows by assuming that flows are exogenously determined. It is important to note that flows in the model implicitly come from common sources; therefore, all investors receive either inflows or outflows when flows occur, but the magnitude of flows can be different across investors.

The investors' investment opportunity set includes one risk-free asset with a constant price normalized to one, and one risky asset. The future value (u) and future flow (F) of the risky asset have the following bivariate normal distribution

$$u, F \sim N\left(\begin{bmatrix}\bar{u} \\ \bar{F}\end{bmatrix}, \begin{bmatrix}\rho_u & \psi \\ \psi & \rho_F\end{bmatrix}\right) \tag{1}$$

where $\bar{u}$ (ρ_u) and $\bar{F}$ (ρ_F) are the mean and the variance of u and F, respectively. The parameter ψ is the covariance between future payoff and future flow. I assume that ψ is strictly positive. This

[5] In the mutual fund literature, flows are generally determined by funds' past performance (e.g., Berk and Green, 2004).

assumption is motivated from profound empirical evidence that flows are positively correlated with contemporaneous returns (e.g., Warther, 1999, Edelen and Warner, 2001, Ben-Rephael, Kandel, and Wohl, 2012).[6] ψ captures the economic channel that underlies the decision of informed investors to invest more in assets whose payoffs co-move more with flows. Following Kacperczyk and Seru (2007), the per capita stock of the risky asset follows an independent normal distribution with mean $\bar{t}$ and variance η. The price, p, of the risky asset is endogenously determined today under the market clearing condition.

Investors obtain signals today for the future value and future flow of the risky asset. I assume that a public signal s_1 for the future payoff is observed by both informed and uninformed investors, and a private signal s_2 for the future flow is observed only by informed investors. In empirical settings, examples of public signals are analysts' forecasts for firms' future earnings or analysts' stock recommendation as in Kacperczyk and Seru (2007). Generally, public signals contain information about assets' fundamentals. Since flows in this model come from common sources, private signals about flows can be thought of as information on market-wide, non-fundamental shocks that drives flows in and out of the equity market. Such information can come from funds' ability to predict market demand for equity assets in response to variations in investment opportunities. It is worth to emphasize that these private signals are not about firm-specific private information that can motivate informed trading. The public and private signals have the following bivariate normal conditional distribution

$$s_1, s_2 | u, F \sim N\left(\begin{bmatrix} u \\ F \end{bmatrix}, \begin{bmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{bmatrix}\right) \tag{2}$$

where ρ_1 and ρ_2 are the variance of s_1 and s_2, respectively. Conditional on the value and flow of the risky asset, the public and private signals are independent. I assume that an α $(0 < \alpha < 1)$ fraction of investors are informed (I), and $1 - \alpha$ fraction are uninformed (U). There are J investors in the economy $(j = 1, 2, \ldots, J)$ in which each investor has CARA utility and their coefficient of risk aversion γ is strictly positive. The equilibrium price is obtained by imposing the market clearing condition that the investors' demand of the risky asset is equal to the available supply.

[6] The assumption that ψ is positive implies that flow beta is always positive in this economy. However, empirical flow betas can be negative. Relaxing this assumption does not change the main conclusion that an informed investor's demand for the risky asset relative to an uninformed investor is positively correlated with flow beta. That is, if flow beta is negative, an informed investor would underweight the asset more relative to an uninformed.

Each investor in the economy faces the following budget constraint

$$c^j + px^j = e^j, \tag{3}$$

where x^j is the amount of risky assets the investor j purchases in the first period, c^j is the amount of cash she holds, and e^j is the initial wealth. The terminal wealth, ω^j, in the second period is

$$\omega^j = c^j + ux^j + F^j. \tag{4}$$

Using Equation 3 to rewrite Equation 4 in terms the investor's initial wealth, her subsequent capital gains and additional flow

$$\omega^j = e^j + (u - p)x^j + F^j. \tag{5}$$

The investor chooses her demand for the risky asset that maximizes her expected utility. She uses the signals to update her beliefs about the payoff and flow $(u^j, F^j|\mathbf{s} = \{s_1, s_2\})$, which follows a conditional bivariate normal distribution.[7] The CARA utility implies that the investor's asset allocation decision is to choose x^j that maximizes

$$E_{\mathbf{s}}[\omega^j] - \frac{\gamma}{2}\text{Var}_{\mathbf{s}}[\omega^j]. \tag{6}$$

Follwing Kacperczyk and Seru (2007), I solve for the equilibrium price p by conjecturing its form as a linear combination of the variables in the model. The partially revealing price for the risky asset in this economy has the following solution

$$p = a_1\bar{u} - a_2\bar{F} + bs_1 + cs_2 - dt + e\bar{t} + g, \tag{7}$$

where $a_1 = \frac{(\rho_\theta+\rho_F)\kappa_1+\alpha(\rho_\theta-\rho_2)\kappa_2}{\kappa}\rho_1$, $a_2 = \frac{\rho_2(\rho_\theta-\rho_2)\kappa_1}{\kappa}\rho_1$, $b = \frac{\kappa_2}{\kappa}$, $c = \frac{\psi\kappa_1+\alpha\rho_u(\rho_\theta-\rho_2)}{\kappa}\rho_1$, $d = \frac{\kappa_1\kappa_2\rho_1\gamma}{\kappa(\psi\kappa_2+\alpha\psi\rho_u(\rho_\theta-\rho_2))}$, $e = \frac{(1-\alpha)\psi\gamma\rho_1\kappa_1^2\kappa_2}{(\psi\kappa_2+\alpha\psi\rho_u(\rho_\theta-\rho_2)+(1-\alpha)\kappa_1)\kappa}$, $g = \frac{\rho_1\psi[\alpha(\rho_u\rho_F-\psi^2)(\rho_2-\rho_\theta)+\rho_\theta\kappa_1]}{\kappa}$, where $\kappa_1 = \rho_u\rho_2 + \rho_u\rho_F - \psi^2$, $\kappa_2 = \rho_u\rho_\theta + \rho_u\rho_F - \psi^2$, and $\kappa = \alpha(\rho_\theta - \rho_2)\rho_1\psi^2 + (\rho_F\rho_1 + \rho_1\rho_\theta + \kappa_2)\kappa_1$.

[7] Uninformed investors do not observe private signals s_2 directly. Instead they learn only noisy estimates θ; therefore, the set of signals for uninformed investors is $\mathbf{s} = \{s_1, \theta\}$.

The optimal allocation for investor j is determined as

$$
\begin{aligned}
x^{j*} &= \frac{1}{\gamma}\frac{E_{\mathbf{s}}(u^j - p)}{\mathrm{Var}_{\mathbf{s}}(u^j)} - \frac{\mathrm{Cov}_{\mathbf{s}}(u^j, F^j)}{\mathrm{Var}_{\mathbf{s}}(u^j)} \\
&= \frac{1}{\gamma}\frac{E_{\mathbf{s}}(u^j - p)}{\mathrm{Var}_{\mathbf{s}}(u^j)} - \frac{\mathrm{Cov}_{\mathbf{s}}(u^j, F^j)}{\mathrm{Var}_{\mathbf{s}}(F^j)}\frac{\mathrm{Var}_{\mathbf{s}}(F^j)}{\mathrm{Var}_{\mathbf{s}}(u^j)} \\
&= \underbrace{\frac{1}{\gamma}\frac{E_{\mathbf{s}}(u^j - p)}{\mathrm{Var}_{\mathbf{s}}(u^j)}}_{\text{mean-variance tradeoff}} - \underbrace{\beta^j_{\text{flow}}\frac{\mathrm{Var}_{\mathbf{s}}(F^j)}{\mathrm{Var}_{\mathbf{s}}(u^j)}}_{\text{hedging component}}.
\end{aligned} \tag{8}
$$

Additional details of the optimal demands and equilibrium price can be found in Section B.1 of the Appendices. Since the main interest is in the relative holdings of an informed investor to an uninformed investor in terms of the private signal s_2 and the flow risk β_{flow}, I analyze the difference in the holdings between the two groups of investors ignoring irrelevant terms in the standard mean-variance tradeoff. Because $\beta^I_{\text{flow}} = \beta^U_{\text{flow}} = \frac{\rho_1\psi}{\rho_F\rho_1 + (\rho_u\rho_F - \psi^2)}$, the difference in the holdings can be rewritten in terms of the model's parameters

$$
\Delta \propto \underbrace{\left[\frac{\psi}{\gamma}\frac{(\rho_\theta - \rho_2)(\rho_u + \rho_1)}{\kappa}\right]}_{\text{private signal coefficient}} s_2 + \underbrace{\left[\frac{(\rho_F\rho_1 + \rho_u\rho_F - \psi^2)(\rho_u\rho_F - \psi^2)(\rho_\theta - \rho_2)}{\rho_1\kappa_1\kappa_2}\right]}_{\text{flow risk coefficient}} \beta_{\text{flow}}. \tag{9}
$$

The terms on the numerators are non-negative because $\rho_u\rho_F - \psi^2 \geq 0$, $\rho_\theta - \rho_2 > 0$, and all the terms under the denominators are strictly positive. For assets that the flow risk coefficient is strictly positive (i.e., $\rho_u\rho_F - \psi^2 > 0$), Equation 9 implies that informed investors have higher holdings of the risky asset relative to those of the uninformed investors given the asset's flow risk. More importantly, this difference increases in the flow beta: informed investors boost their holdings of the risky asset the higher its flow risk is.

The intuition behind this relation is that an informed investor can bet on high-flow-beta assets to take advantage of the positive correlation between future flows and payoffs. If flow risk commands a premium as shown in Dou et al.'s (2023) model (or implied from the hedging component in Equation 8), it is possible that informed investors make investment decisions to capture this risk premium. It is clear that this is an optimal allocation decision if future flows are positive (i.e., inflows). In case of outflows, an informed investor who receives private signals about potential outflows (i.e., s_2 is negative) lowers her holdings by a magnitude of the private signal coefficient, offsetting the demand for high-flow-risk assets captured by the flow risk coefficient. Moreover, Equation 9 implies that the more precise the private signal is (i.e., lower ρ_2), the more weight an informed investor put to the private signal and flow beta.

Based on the analyses so far, I conjecture that more skilled investors receive more accurate private information about potential flows (s_2 with lower σ_2), and their portfolios have higher exposure to flow risk. To gauge the relation between a portfolio's holdings and underlying stock flow betas in the cross section, I rewrite the flow risk coefficient from Equation 9 in terms of the covariance between the difference in risky holdings and β_{flow}

$$\begin{aligned}\text{Flow risk coefficient} &\propto \text{Cov}(x^I - x^U, \beta_{\text{flow}}) \\ &= \text{Cov}(x^I, \beta_{\text{flow}}) - \text{Cov}(x^U, \beta_{\text{flow}}) \geq 0. \end{aligned} \tag{10}$$

I use this covariance representation to later motivate an empirical measure that captures the heterogeneous skill in managing the flow betas among US active mutual fund managers. It is important to discuss what β_{flow} captures. To be precise, β_{flow} captures the co-movement between an asset's payoff and its future flows. If we extend an asset to a portfolio with N stocks in the context of an equity fund, $\beta_{\text{flow,i}}$ captures the co-movement between stock i's payoff and the future flows into the fund's portfolio. Since Dou et al. (2023) show that flows in and out of mutual funds share a common structure, I use the common flows to proxy for all funds' flows. That is, the empirical measure $\beta_{\text{flow,i}}$ captures the co-movement between stock i's payoff and the common fund flows.

2.2 Empirical Predictions

Following the analyses from Section 2.1, in this section I formally state the testable predictions.

Since I conjecture that more skilled investors have higher exposure to flow risk, the underlying hypothesis is that funds who deviate more from the benchmark in a positive direction with flow beta are skilled funds. This hypothesis directly leads to the first empirical prediction: a fund whose covariance between its holdings' deviation from a benchmark and underlying stock flow beta has higher subsequent performance.

A second testable prediction is related to how the fund deviates from the benchmark with respect to the precision of public information. The model predicts that while the skilled funds deviate from the benchmark positively in the direction of flow beta, they do less so when the public information on the stock is more volatile. In other words, the more imprecise the public information on the stocks, the less the funds deviate from the benchmark. The formal prediction is that funds deviate less from the benchmark for stocks that has volatile public information, conditional on stocks' flow beta.

3 Data

Section 3.1 provides construction details of the mutual fund sample. Section 3.2 describes construction details of common flow shocks and stock flow beta, and Section 3.3 provides construction details of the AFB measure.

3.1 Mutual Fund Sample

I obtain data on monthly fund returns and total net assets (TNA) from Center for Research in Security Prices Survivor-Bias-Free U.S. Mutual Fund (CRSP MF) from January 1991 to December 2021. The restriction on the start date of the data is due to poor data coverage on monthly TNA before 1991 (Dou et al., 2023). The returns are net of fees, expenses, and brokerage commissions, but before any loads. I convert the net returns to excess returns by subtracting the risk-free rate.[8]. I obtain quarterly fund equity holdings data from the Thomson Reuters Mutual Fund Holdings Database (S12) for the sample period from 1991 to the third quarter of 2008, and the CRSP mutual fund holdings data for the rest of the sample. The use of CRSP data on portfolio holdings is to minimize concerns related to data quality of Thomson Reuters holdings data before 2008 (e.g., Zhu, 2020).

I use the CRSP MF database to collect information on fund characteristics such as expenses, fund portfolio turnovers, and percentage of portfolio invested in common stocks and other asset classes. Since a mutual fund can have multiple share classes, I use the MFLINKS database to identify such funds and combine different share classes into fund-level portfolios. For each period, I use the most recent TNA to construct fund-level TNA, returns, and characteristics. In particular, I take the sum of TNA across all share classes of a fund to construct the fund's TNA. The fund's returns

[8] I obtain data on the monthly factor returns (i.e., the market, size, value, momentum, and liquidity factors) and the risk-free rate from Kenneth French's and Robert Stambaugh's website at `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/` and `https://finance.wharton.upenn.edu/~stambaug/`. I thank Kenneth French and Robert Stambaugh for making these data available.

and other characteristics are TNA-weighted averages. Similar to prior studies (e.g., Kacperczyk et al., 2008, Jiang and Zheng, 2018), I estimate monthly gross returns by dividing the annual expense ratio by 12 and adding that to the monthly net returns. I also use the MFLINKS database to merge the holdings data with the CRSP MF data.

I follow Dou et al. (2023) and restrict the sample to domestic actively managed U.S. equity funds. In particular, I eliminate index funds, balanced funds, sector funds, international funds, bond funds, money market funds, and exchange-traded funds.[9] I also remove funds for which fund names are missing. To address concerns related to omission bias (Elton, Gruber, & Blake, 2001) and incubation bias (Evans, 2010), I perform additional screens on the sample. In particular, I delete any fund-month observations prior to the first offer dates of funds, and exclude observations if the fund's TNA in the previous month is below \$15 million. Finally, I include only funds that have more than 80% of their holdings on average in common stocks. The final mutual fund sample contains 4,258 unique funds from 1991 to 2021.

Panel A of Table 1 shows the summary statistics for mutual funds in my sample. The average fund manages \$1.57 billion of assets. On average, a fund exists for over 13 years during the sample period. The quarterly mean return is 2.72% and its distribution appears symmetric since the median is close to the mean. Consistent with prior studies (e.g., Jiang and Zheng, 2018), fund flow is positively skewed as the mean of quarterly flow (1.03%) is significantly higher than the median (-1.10%). The average annual expense ratio is 1.04% in my sample and the turnover ratio is 80.67% annually. In general, the distribution of these characteristics is consistent with previous literature.

3.2 Construction of Common Flows and Stock Flow Beta

Common Flow Shocks. I follow Dou et al. (2023) to estimate the time-series common fund flows. Since I use these estimates later to evaluate future mutual fund performance, it is important to avoid look-ahead biases. As a result, my estimation procedure adopts an expanding-window design in which I re-estimate all the parameters to construct the common fund flows at month t using

[9] To exclude index and exchange-traded funds, I use both CRSP index fund flag and check for funds' name with the following key words: 'index', 'inde', 'indx', 'inx', 'idx', 'dow jones', 'ishare', 's&p', 's &p', 's& p', 's & p', '500', 'wilshire', 'russell', 'msci', 'etf', 'exchange-traded', 'exchange traded'. I identify balanced, sector, international, bond, and money market funds by using the following CRSP policy code: 'C & I', 'Bal', 'Bonds', 'Pfd', 'B & P', 'GS', 'MM', 'TFM'. U.S. equity funds are further selected by using the following policy code: Lipper classes and objective codes 'EIEI', 'G', 'LCCE', 'LCGE', 'LCVE', 'MCCE', 'MCGE', 'MCVE', 'MLCE', 'MLGE', 'MLVE', 'SCCE', 'SCGE', 'SCVE', 'CA', 'EI', 'GI', 'MC', 'MR', 'SG'; Strategic Insight objective codes 'AGG', 'GMC', 'GRI', 'GRO', 'ING', 'SCG'; Wiesenberger objective codes 'G', 'GCI', 'IEQ', 'LTG', 'MCG', 'SCG'.

data only up to month t. This is different from Dou et al.'s (2023) main procedure in which they use the full sample for estimation. The detailed process is as follows.

Starting from December 1991, I first run a pooled panel regression of fund flows on funds' current and prior performance and prior flows using data from January to December 1991

$$F_{j,t} = \beta_0 + \beta_1 R^e_{j,t} + \beta_2 R^e_{j,t-1} + \beta_3 F_{j,t-1} + \gamma_t + \varepsilon_{j,t}, \tag{11}$$

where $F_{j,t}$ is fund j's flow at month t, $R^e_{j,t}$ is fund j's excess return relative to the market return over month t, and γ_t is the month t fixed effects.[10] The fund-flow shock for fund j at month t is estimated as

$$flow_{j,t} = \gamma_t + \varepsilon_{j,t}. \tag{12}$$

Second, I sort all funds in each month into five groups based on their TNA in the previous month, and use fund-flow shock $flow_{j,t}$ to calculate the TNA-weighted average flow shock for each group. The process produces five time-series flow shocks. Similar to Dou et al. (2023), I detrend the series of each quintile to account for the time trend in asset size of the mutual funds. Finally, I obtain the common fund flows ($flow_t$) by extracting the first principal component of the fund flow shocks across the quintiles using principal component analysis. I repeat the procedure for each month until December 2021 to obtain the monthly time-series of common fund flows from 1991 to 2021. Panel A of Figure 2 plots the time-series of the common fund flow shocks during my sample.

Stock Flow Beta. I estimate the exposure of stock i to the common fund flows in month t using 36-month rolling regressions, controlling for the market exposure

$$r_{i,t-\tau} = \alpha_{i,t} + \beta_{mkt,i,t} MKT_{t-\tau} + \beta_{flow,i,t} flow_{t-\tau} + \varepsilon_{i,t-\tau}, \quad \tau = 0, 1, \ldots, 35, \tag{13}$$

where $r_{i,t-\tau}$ is stock i's monthly excess returns, $MKT_{t-\tau}$ is the market excess returns, and $flow_{t-\tau}$ is the common fund flow shocks.[11] I require at least 12 months of observations for each regression

[10] Flow $F_{j,t}$ is defined as $[A_{j,t} - A_{j,t-1}(1 + R_{j,t})]/[A_{j,t-1}(1 + R_{j,t})]$, where $A_{j,t}$ is fund j's TNA at month t. The return adjustment in the denominator is to minimize large distortions in flows due to intermediate contemporaneous flows and returns within month t (e.g., Berk, Van Binsbergen, and Liu, 2017, Sialm and Zhang, 2020). I show in the robustness section that the main results do not change significantly when I use the traditional flow measure defined as $[A_{j,t} - A_{j,t-1}(1 + R_{j,t})]/A_{j,t-1}$.

[11] Dou et al. (2023) do not control for the market exposure. I include the market factor because the asset pricing model in Dou et al. (2023) represents an ICAPM model in which the market risk is priced. I show in the robustness section that the main results do not change significantly when Equation 13 does not include the market factor.

to ensure reliable estimation for $\beta_{flow,i,t}$. Because I use an expanding-window design to estimate the common fund flows, the shocks in the first few years are subject to estimation risk. Therefore, I discard the first two years of the estimates to address this concern.[12] As a result, the sample used for performance evaluation starts from January 1994.

I construct the stock sample using the universe of firms covered by the Center for Research in Security Prices (CRSP) and the Compustat Fundamentals Annual (Compustat). Similar to Dou et al. (2023), I include only U.S. common stocks that are listed on NYSE, NASDAQ, and Amex. To ensure sufficient data for estimation, I require a stock to have at least 2 years of data on Compustat and 36 months of return observation on CRSP. Panel B of Table 1 shows the summary statistics of the stock sample. The statistics are in line with Dou et al. (2023). The mean of flow beta is 0.17; however, the distribution is skewed to the left as the median is close to zero. The average firm in my sample has a market capitalization of $338 million. On average, the book-to-market ratio is 0.52. Both liquidity and uncertainty betas are positively skewed with mean of 0.016 and 0.024, respectively. The average firm has an Amihud's illiquidity measure of 4.68.

Table A1 in the Appendix provides the summary statistics for decile portfolios of stocks sorted on their flow beta. Consistent with main findings in Dou et al. (2023), high-flow-beta stocks has higher monthly returns on average because of the flow risk premium associated with the common fund flows. The average monthly return for the decile portfolio from 1994 to 2021 is 1.08% compared to 0.74% of the bottom decile portfolio.[13] High-flow-beta stocks are smaller and less liquid. They are more likely to be value stocks. Stocks with high flow risk also have higher exposure to aggregate liquidity risk and uncertainty risk.

3.3 Construction of Active Flow Beta

In this section I describe the construction of an empirical measure that captures a fund manager's portfolio exposure to common fund flows. Equation 10 suggests that one can capture the portfolio exposure to flow risk by estimating the covariance between the portfolio weights and underlying stock flow betas. This type of performance measure has been analyzed and adopted in the literature on fund manager skills (e.g., Grinblatt and Titman, 1993, Jiang and Zheng, 2018). Grinblatt and

[12] In untabulated results, I find that this data choice does not affect the main results materially.

[13] In untabulated results, I confirm that CAPM risk-adjusted return of the top decile portfolio is higher and the difference is statistically significant.

Titman (1993) show that the covariance between a fund's portfolio weights and its underlying asset returns is a reasonable proxy for active management. Jiang and Zheng (2018) improve the measure by changing the assets' returns to their abnormal returns around earnings announcements to better capture fundamental values of the assets. Moreover, because mutual fund managers evaluate their performance relative a benchmark (Cremers & Petajisto, 2009), Jiang and Zheng (2018) use the relative portfolio weights instead of absolute holdings.

I adopt the covariance logic and empirically measure the active management of fund j's flow betas across its holdings as follows

$$AFB_{j,q} = \sum_{i=1}^{N_j} \mathrm{Cov}(\omega_{j,i,q} - \omega_{bm,i,q}, \beta_{flow,i,q}) \approx \sum_{i=1}^{N_j} (\omega_{j,i,q} - \omega_{bm,i,q})\beta_{flow,i,q}, \tag{14}$$

where $AFB_{j,q}$ is the active flow beta of fund j in quarter q, $\omega_{j,i,q}$ and $\omega_{bm,i,q}$ are the portfolio weights of asset i in fund j's portfolio and its benchmark portfolio, respectively. Equation 14 requires identifying the benchmark portfolio for each fund j. I discuss two potential concerns relating to the empirical benchmark identification.

First, the theoretical relation from Equation 10 is silent on benchmark identification. Nevertheless, the model implies that the investment opportunity set, or the benchmark, of all fund managers is the same. Motivated by this observation, I use all available stocks in the stock market as the benchmark for main analyses. This benchmark choice is also consistent with the empirical choice in Dou et al. (2023). However, mutual funds differ in their benchmark empirically (e.g., Cremers and Petajisto, 2009), motivating the use of fund-specific benchmarks for performance evaluation. I show in the robustness section that the main results do not change significantly when fund-specific benchmarks, constructed following Cremers and Petajisto's (2009), are used to measure AFB.

Second, a same market benchmark for all fund managers requires assigning zero weights to stocks that are in the benchmark but not in the funds' portfolio. In other words, since funds only report the stocks with non-negative holdings, the absence of other stocks in their portfolios implies that funds completely deviate from the benchmark. According to the model, this is an ideal empirical assumption. However, the assumption requires the stacking of large holdings data, and thus limits deeper empirical analyses. Therefore, I restrict the estimation of AFB to only those stocks that funds report. In untabulated results, I find that the stacking design does not affect the main results significantly.

Table 2 shows the mean of fund characteristics for portfolios sorted by AFB. I provide construction details of the fund predictors in Section B.2 in the Appendices. For each quarter from 1994 to 2021, I calculate the AFB for each fund and sort the funds into ten decile portfolios based on their lagged AFB. The high (low) decile portfolio contains funds with the highest (lowest) AFB. In each quarter, I calculate the cross-sectional mean for each characteristic and portfolio, and report the time-series average of these cross-sectional means. The last column reports the mean difference of the characteristics between the top and bottom deciles. Funds with high AFB appear to be smaller and exist longer. However, their differences are not statistically significant. These funds have significantly higher flows. They also have significantly lower expense and turnover ratios. The differences are 0.09% and 16.4%, respectively.

It is perhaps not surprising that funds in the top decile of AFB generally have higher value of the fund characteristics that positively predict their future performance (e.g., higher active share, or lower R^2) as these funds may share a common source of skill. However, most of differences are not statistically significant. There are two exceptions. Kacperczyk et al.'s (2008) *Return gap* appears to be strongly correlated with activeness in flow beta management. Particularly, the proxy for unobserved actions of mutual funds is 0.15% higher in the top decile AFB portfolio than that in the bottom decile, and the difference is statistically significant at the 5% level. The top decile funds also have significantly lower Avramov et al.'s (2020) AFO, suggesting that these funds may possess the skill to identify mispriced stocks. These results suggest the need to control for these fund characteristics when evaluating the predictive power of AFB for fund performance.

4 Active Flow Beta and Mutual Fund Performance

4.1 Portfolio Sorts

In this section, I evaluate the performance of a strategy that invests in mutual funds based on their active management of flow beta. In particular, at the beginning of each calendar quarter from 1994Q1 to 2021Q4, I sort all funds into decile portfolios according to AFB. The high (low) decile portfolio includes funds with the highest (lowest) activeness with regard to flow beta. I compute the equal-weighted returns for the decile portfolios in the first month of each quarter and track their excess returns over the next two months. I rebalance the portfolios quarterly. To account for the portfolios' exposures to risk factors, I compute the risk-adjusted returns on the portfolios as the intercept from the time-series regressions based on Carhart's (1997) four-factor model

$$r_{p,t} - r_f = \alpha_p + \beta_{MKTRF} MKTRF_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{UMD} UMD_t + \varepsilon_{p,t}, \quad (15)$$

where $r_{p,t} - r_f$ is portfolio p's monthly excess returns, net of expenses. $MKTRF_t$, SMB_t, HML_t, UMD_t are the excess returns on the market, size, value, and momentum factors, respectively. I also add the Pástor and Stambaugh's (2003) liquidity factor to the four-factor model to account for the portfolios' exposure to aggregate liquidity risk.

Table 3 summarizes the results from these portfolio tests. Panel A reports the results on the portfolios' excess returns, and Panels B and C reports the results on the risk-adjusted returns according to the two factor models. Panel A shows that the difference in net-of-expenses performance of low- and high-AFB funds is about 0.43% per month, or 5.16% per year.[14] This difference is economically large and statistically significant at the 10% level. After adjusting for exposure to Carhart's (1997) four risk factors in Panel B, the return difference is 0.60% per month, or 7.20% per year. Again, the difference is statistically significant at the 1% level. Panel B also reports the exposure of each decile portfolio to the risk factors. The high AFB portfolio has significantly lower exposure to the market and size factors, but the loadings are positive. The top decile portfolio also loads positively on the value factor, and has significantly higher exposure to this factor compared to the low decile portfolio. The flow-beta strategy does not appear to be significantly related to the momentum factor. Panel C shows that adding the liquidity factor does not affect the performance differential significantly. The return differential is still economically large (about $0.53 \times 12 = 6.36\%$ annually), and statistically significant ($t-$stat $= 2.50$). Finally, the high-flow-beta funds have significantly positive exposure to the liquidity factor.

A common issue with identifying mutual fund skill is that the documented performance might not be persistent (Carhart, 1997). If the measure AFB captures the skill of fund managers to manage their exposure to common flow risks, it is important to examine whether this skill is persistent.

[14] Since mutual funds cannot be shorted, this strategy is not implementable in real life. Therefore, the difference should be interpreted as the net-of-expenses return that an investor would earn on average by buying the high-flow-beta funds than the low-flow-beta funds.

Similar to prior studies (e.g., Kacperczyk et al., 2008), I establish the persistence of the mutual fund skill with regard to active flow beta by tracking the funds over time based on their portfolio rank and active flow beta. Specifically, at the beginning of each calendar quarter from 1994Q1 to 2021Q4, I sort all funds into decile portfolios according to AFB as the investment strategy above. Then I track the portfolio rank and AFB of each fund for the subsequent 20 quarters, or five years. For each quarter, I compute the TNA-weighted average of the portfolio rank and AFB. Finally, I compute the time-series average of these values for each quarter series in the full sample to obtain the trajectory of the portfolio ranks and active flow betas.

Figure 3 plots the trajectory of the portfolio rank (Panel A) and the active flow beta (Panel B). Both panels suggest that there exists a persistence in the performance of funds in both top and bottom deciles. For example, Panel A shows that funds in the high-flow-beta portfolio continue to stay on top in the ranking for up to more than 12 quarters, or three years. On the other hand, funds with low-flow-beta port continue to stay at the bottom for up to 16 quarters, or four years. Panel B shows that the persistence in the portfolio ranking comes from the persistence in active flow betas. The top decile funds maintain high flow beta throughout and their AFB is positive for up to five years. Overall, the evidence suggests that the mutual funds' ability to manage flow beta is persistent.

4.2 Double Sorts

The evidence so far suggests that AFB is a strong predictor of mutual fund performance. However, the literature has documented a number of fund characteristics that also have strong predictive power for funds' future returns. As Table 2 shows, some of these fund characteristics (e.g., *Return gap* or *AFO*) share significant correlation with AFB. Therefore, it is important to examine whether the AFB can provide incremental information for fund performance above and beyond what other characteristics have provided.

The five characteristics I consider include Kacperczyk et al.'s (2008) *Return gap*, Cremers and Petajisto's (2009) *Active share*, Kacperczyk and Seru's (2007) *Reliance on public information* (RPI), Jiang and Zheng's (2018) *Active fundamental performance* (AFP), and Avramov et al.'s (2020) *Active fund overpricing* (AFO). I adopt portfolio double sorts to examine the interaction between AFB and these fund characteristics in predicting fund performance. Specifically, at the beginning of each calendar quarter from 1994Q1 to 2021Q4, I independently sort all funds into quartile portfolios according to AFB. The high (low) decile portfolio includes funds with the highest (lowest) activeness with regard to flow beta. Simultaneously, I independently sort all funds into quartile portfolios according to one of the five characteristics. The high (low) decile portfolio includes funds with the highest (lowest) values of the characteristic. I compute the equal-weighted returns for the portfolios in the first month of each quarter and track their excess returns over the next two months. The rebalancing frequency is quarterly. I report the risk-adjusted returns according to Carhart's (1997) four-factor model for the portfolios.

Table 4 summarizes the results from these double sorts. Panel A reports the results for Kacperczyk et al.'s (2008) *Return gap*. Consistent with the findings from Kacperczyk et al. (2008), *Return gap* positively predict fund performance, The four-factor alpha of the high-low portfolio is 0.24% monthly and statistically significant at the 1% level. The predictive power is consistent throughout all groups of AFB. Nevertheless, AFB remains a strong predictor of future fund returns controlling for different *Return gap* levels. The alpha is consistently large and statistically significant in all quartiles of *Return gap*. Panel B shows that AFB still predicts future fund returns after controlling for *Active share*.[15] Panel C shows that AFB remains a strong predictor after controlling for *Reliance on public information*. Interestingly, RPI has the strongest predictive power among funds with the extreme AFB. Since my model in Section 2 retains the prediction of funds' response to public information from Kacperczyk and Seru (2007), these results suggest that the skill of some low RPI funds may come from having private information about future flows.

Recent studies on skills of mutual funds argue that funds differ in their ability to predict future earnings (Jiang & Zheng, 2018) and identify mispriced stocks (Avramov et al., 2020). Panels D and E show that both Jiang and Zheng's (2018) AFP and Avramov et al.'s (2020) AFO predict fund future performance as found in the original studies. However, they cannot subsume the predictive power of AFB. The alpha is economically large and statistically significant across all subgroups of AFP and AFO. Interestingly, funds' ability to predict future earnings and profit from such ability is only profound among funds in the low-flow-beta portfolio. The alpha of the long-short strategy based on AFP is 0.14% monthly and statistically significant only among the bottom quartile portfolios

[15] The insignificant alpha of high-flow portfolio sorted on *Active share* does not contradict Cremers and Petajisto's (2009) results. They show that the alpha is only significant when the funds' specific benchmark is used. The alpha in this table (0.03% monthly) is close to what the authors find when using the four-factor model for risk adjustment.

sorted by AFB. For other quartile portfolios, the alpha is not statistically significant. I conjecture that since these funds do not have private information or their private signals are noisy, the ability to use public information to predict future earnings is more rewarding. On the other hand, funds' differential ability to identify mispriced stocks is only pronounced among funds in the high-flow-beta portfolio. Among funds in the top quartile of AFB, funds that incorrectly price stocks earn 0.08% less per month and this differential is statistically significant. This differential is small and not statistically significant for other subgroups of AFB.

4.3 Predictive Panel Regressions

In this section, I use panel regressions to assess the incremental power of AFB to predict fund performance. The panel is at fund-quarter level and includes fund performance, AFB, and the five fund characteristics that I use in Section 4.2. Following prior studies, I additionally control for fund size, age, expense ratio, turnover, and flows and performance in the previous quarter. I use realized alpha to proxy for fund performance. The monthly realized alpha is estimated as the difference between fund's realized net return and expected return estimated from Carhart's (1997) four-factor model. The factor loadings are estimated from rolling-window regressions of fund excess returns on the factors in the previous 36 months. I take the average of monthly realized alphas within a quarter to measure the fund performance in the quarter.

Table 5 summarizes the results from the panel regressions. Following Jiang and Zheng (2018), all measures of skill are winsorized at the top and the bottom 5%. I standardize all measures of skill to have mean of 0 and standard deviation of 1 for ease of comparison, and report the coefficient estimates in basis points. All specifications include time fixed effects and the standard errors are clustered at the fund level. Column 1 shows the univariate result in which I examine how AFB predicts fund performance independently. Subsequent columns report multivariate results when I include other fund characteristics.

The results in Table 5 show that AFB predicts fund performance consistently. In Column 1, the predictive coefficient is 3.38 and statistically significant at the 1% level. This number implies that a 1-standard-deviation increase in AFB is associated with an increase of 10.14 basis points in risk-adjusted return over the subsequent quarter. In Column 2 where I control for other standard fund characteristics, the predictive magnitude decreases but remains statistically significant at the 1% level. The predictive sign of other fund characteristics is also consistent with established findings in the literature: larger funds have lower risk-adjusted performance (Chen, Hong, Huang, & Kubik, 2004), funds with higher expense ratio underperform relative to funds with lower expense ratio (Elton, Gruber, Das, & Hlavka, 1993), and past flows predict positively future fund returns (Zheng, 1999). The results from Columns 3 to 9 show that AFB remains a strong predictor of fund performance after controlling for other fund characteristics individually. All the measures of fund skill predict future fund performance in the correct direction as shown in the original studies, and their predictive coefficients are statistically significant, except for R^2.[16] When I include all predictors and report the regression result in Column 10, *Return gap* and *RPI* are no longer significant. However, AFB continue to predict future performance significantly.

To further understand the contribution of AFB in predicting fund performance, I follow Avramov et al. (2020) to decompose the measure into three components

$$
\begin{aligned}
AFB_{j,t} \approx & N_{j,t} \times \operatorname{Cov}(x_{j,i,t} - x_{mkt,i,t}, \beta_{flow,i,t}) \\
= & \underbrace{\rho(x_{j,i,t} - x_{mkt,i,t}, \beta_{flow,i,t})}_{\text{correlation between active share and flow beta}} \times \overbrace{N_{j,t}\sigma(x_{j,i,t} - x_{mkt,i,t})}^{\text{standard deviation of active share}} \times \sigma(\beta_{flow,i,t}).
\end{aligned} \tag{16}
$$

Similar to Avramov et al.'s (2020) AFO, the correlation term from Equation 16 captures the direction of funds to deviate from the benchmark with respect to the flow risk, while the standard deviation term captures the activeness of the fund. I estimate the two components for each fund-quarter observation and repeat the panel regressions.

Table 6 shows the regression results from the predictive regressions using the components of AFB. Column 1 shows that both the correlation and the activeness terms strongly and positively predict fund performance. Particularly, a 1-standard-deviation increase in the correlation between funds' deviation from the market benchmark and flow beta is associated with an increase of 3.42 basis points in monthly alpha over the next quarter. The effect of a 1-standard-deviation increase in *STDAS* is 3.56 basis points. Interestingly, the inclusion of the two terms in the regressions make a number of fund characteristics to lose their predictive power. *RPI* and *Active share* are no longer significant. If the correlation term is a better proxy for funds' use of public versus private

[16] The decrease in predictive magnitude and absence of statistical significance for some fund performance predictors can come from the out-of-sample effect (Jones & Mo, 2021).

information, it is possible that the predictive power of RPI is subsumed by this measure. The absence of significance for *Active share* is not surprising because $STDAS$ is also an alternative measure of active share. When I use all the predictors and fund characteristic in Column 10, only the correlation term remains a strong predictor of fund performance. Although $STDAS$ still positively predicts fund performance, its significance reduce by a large magnitude. These results suggest that funds' increasing holdings of stocks in the same direction as flow beta is more informative about their performance rather than just their activeness.

4.4 AFB and Precision of Public Information

The analyses so far provide strong evidence to support the main hypothesis from the model in Section 2 that funds with higher flow beta exposure have better performance. In this section, I test the model's predictions relating to how the fund deviates from the benchmark with respect to the precision of public information. The model predicts that while the skilled funds deviate from the benchmark positively in the direction of flow beta, they do less so when public information on underlying stocks is more volatile. In other words, the more imprecise the public information on the stocks, the less the funds deviate from the benchmark. Intuitively, since the funds' expectation of stock payoff and flow is positively related to the precision of public information, skilled funds should bet less on the high-flow-beta stock if the public information is more noisy.

To test this hypothesis, I construct two measures of public information precision based on the prior literature. Particularly, I use the dispersion in analysts' earnings forecasts and stocks' idiosyncratic volatility to proxy for the volatility of public information. These two variables have been used extensively in the literature to examine the relation between disagreement and asset prices (e.g., Yu, 2011, Huang et al., 2021). I follow Huang et al. (2021) to construct the stock-level analysts' forecast dispersion and idiosyncratic volatility. To construct the aggregate holdings for each decile portfolio sorted on AFB, I aggregate fund-level holdings and compute the portfolio-level holdings for each stock. I perform the following Fama-MacBeth regressions for each portfolio

$$\omega^j_{i,q+1} - \omega^m_{i,q+1} = \gamma_{0,q} + \gamma_{1,q}\beta_{\text{flow},i,q} + \gamma_{2,q}\beta_{\text{market},i,q} + \gamma_{3,q}\sigma_{i,q} + \gamma_{4,q}\beta_{\text{flow},i,q}\sigma_{i,q} + \varepsilon_{i,q+1}, \tag{17}$$

where $\omega_i^j - \omega_i^m$ is the deviation of stock i in portfolio's j from the market allocation and σ_i is a measure of precision of public information for stock i. The prediction is that the coefficient estimate γ_4 is significantly lower for high-flow-beta funds.

Table 7 reports the coefficient estimates and their statistical significance from Equation 17 for the high- and low-flow-beta portfolios. Panels A and B use analysts' disagreement and stocks' idiosyncratic volatility as the proxy for the imprecision of public information, respectively. The last row reports the difference in the coefficients between the two portfolios. First, the difference in the coefficient β_{flow} between the high- and low-flow-beta funds is positive and statistically significant. This is consistent with the main prediction that skilled funds who have more accurate private signals should deviate more with respect to flow beta. Second, the coefficient $\beta_{\text{flow}} \times \sigma$ is smaller in the high-flow-beta portfolio by a large magnitude compared to that of the low-flow-beta portfolio. Their difference is -0.035 and statistically significant at the 5% level. Consistent with the prediction from the model, high AFB funds deviate less from the market benchmark when the public information on the stocks is less precise.

Finally, I test if the ability of funds to actively manage flow beta is more valuable during times of high public disagreement. I construct the aggregate version of the two public disagreement measures following Huang et al. (2021) using the stocks' market capitalization as the weights. Over the sample period from 1994 to 2021, I construct an indicator variable equal to 1 if a month belongs to the top decile of the aggregate measure, and 0 otherwise. For each decile portfolio sorted on AFB, I perform the time-series regression that adjusts for risk exposure to Carhart's (1997) four factors

$$\begin{aligned} r_{p,t} - r_f = &\alpha_p + \beta_{vol}\text{Volatility Indicator} + \beta_{MKTRF}MKTRF_t \\ &+ \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{UMD}UMD_t + \varepsilon_{p,t}, \end{aligned} \tag{18}$$

I conjecture that β_{vol} is positive and statistically significant if the funds' ability to manage flow beta is more profound during periods of high public disagreement.

Table 8 summarizes the results from Equation 18 for the high- and low-flow-beta fund portfolios. Panels A and B use analysts' disagreement and stocks' idiosyncratic volatility, respectively. The last row reports the results for the high-minus-low portfolio. Consistent with the conjecture, funds with high AFB perform significantly better during periods of high public disagreement. For example, the high AFB portfolio earns an additional 0.73% on months of extreme disagreement among analysts.

The difference between the high- and low-flow-beta portfolios is economically large and statistically significant. On average, high AFB funds earn 1.71% more than low AFB funds on months when analysts has the most diverse opinions. The number is 1.70% when idiosyncratic volatility is used as the proxy for the imprecision of public information.

4.5 Robustness Tests

The empirical measure AFB is constructed from the difference between funds' holdings and the market benchmark, and the underlying stock flow betas. In this section, I examine the robustness of the main results to the choice of flow beta construction and funds' benchmark.

Alternatives of flow beta. Dou et al. (2023) use three different measures of flow beta in their study. The main measure in this paper is constructed as the first principal component from flow shocks obtained from size-sorted quintile portfolios. I use the flow definition that adjusts for contemporaneous return to avoid large distortion in flow measures. Similar to the first measure, I construct a second measure for common flow shocks but fund flows are measured as $[A_{j,t} - A_{j,t-1}(1 + R_{j,t})]/A_{j,t-1}$. Moreover, I also construct common flow shocks using the TNA-weighted average of the fund-level flow shocks at the cross section of all funds, instead of using the first principal component for quintile portfolios. For these two alternative measures, I construct AFB as in Equation 14. Finally, instead of using Equation 14 to proxy for the covariance, I use the cross-section of a fund's holdings to capture AFB and measure AFB directly as the covariance between holdings' deviation from the market benchmark and the main measure of stock flow beta.

Table A2 in the Appendix reports the results from these tests. Panels A, B, and C report the results using the first, second, and third definition of AFB as described above. Across the panels, the difference in performance between funds in the high AFB portfolios versus those in the low AFB portfolios is positive. The differences are statistically significant for the risk-adjusted returns according to Carhart's (1997) four-factor model and the five-factor model that adds the liquidity factor. The weaker results for the third measure using the direct covariance measure is potentially because the measure cannot fully capture the activeness of funds with respect to flow betas across all stocks. That is, the covariance is only meaningful to capture the sign of the two variables, but is not indicative of how strong the direction is.

Alternatives of fund benchmark. I construct the AFB under the assumption that all funds use the market benchmark. Since funds differ in the benchmark, I adopt the benchmark selection criteria as in Cremers and Petajisto (2009) to examine the robustness of the main results. Particularly, I use the active share data obtained from Martijn Cremers' website to identify the benchmark for each fund.[17] Similar to Jiang and Zheng (2018), I obtain the holdings of the benchmarks using the holdings of index funds that closely resemble the underlying indices. I then use these holdings to compute the benchmark weights and subsequently the deviation of funds' holdings from their benchmark. Apart from the fund-specific benchmark, I also adopt the trade-based benchmark, which is the change in the portfolio holdings over time. Particularly, I measure the percentage changes in each fund's portfolio holdings from the previous quarter, taking into account stock price changes over the quarter.

Table A3 in the Appendix reports the results from these tests. Panels A and B report the results using the fund-specific and trade-based benchmark as described above. Panel A shows that while the economic magnitude in the return differential between the high- and low-flow-beta portfolio is smaller, it is still statistically significant. For example, the high-minus-low portfolio earns a Carhart's (1997) alpha of 0.42% monthly ($t-$stat=2.65). Panel C shows that the economic magnitude reduces by almost half when the trade-based benchmark is used, but the strategy still earns significant Carhart's (1997) alpha (0.22% with $t-$stat of 2.64).

5 Conclusion

Dou et al. (2023) show that active equity funds tilt their portfolios toward low-flow-beta stocks to hedge against common flows. This flow-hedging behavior rationally explains a flow risk premium in the cross section of expected stock returns, but also predicts lower expected return for active funds. However, it is still not clear why all active mutual funds forgo the premium associated with high-flow-beta stocks.

Using a sample of over 4,000 U.S. domestic active equity mutual funds from 1994 to 2021, I first document that there is a significant heterogeneity in the flow-hedging behavior: almost half of active equity funds do not appear to hedge against flow shocks but rather tilt toward high-flow-beta stocks. I rationalize this finding in an extended model of Kacperczyk and Seru (2007) that

[17] The benchmark is defined as the one among a set of 21 indices that funds have the lowest active share.

features informed and uninformed investors who differ in the precision of private information they receive about future flows. The main intuition is that an informed investor that receives private signals about future flows allocate more to the high-flow-beta asset because such assets' future payoff are strongly and positively correlated with the future flow. In the mutual fund context, the main empirical prediction from the model is that funds who deviate more from the benchmark in a positive direction with flow beta are skilled funds and should have higher subsequent performance.

To test the model's prediction, I construct an empirical measure that captures the active management of mutual funds with respect to flow beta (AFB). I find that funds in the top decile of the measure outperform those in the bottom decile. The performance differential is economically large and statistically significant, even after adjusting for exposure to risk factors. The predictive ability of AFB is above and beyond other established fund predictors in the literature. Thus, I show that the cross-sectional difference in flow hedging among active funds can be informative about managerial skill, and provide a new insight into the management of flow risk in the mutual fund industry.

References

Amihud, Y., & Goyenko, R. (2013). Mutual fund's R^2 as predictor of performance. *Review of Financial Studies*, *26*(3), 667–694.

Avramov, D., Cheng, S., & Hameed, A. (2020). Mutual funds and mispriced stocks. *Management Science*, *66*(6), 2372–2395.

Ben-Rephael, A., Kandel, S., & Wohl, A. (2012). Measuring investor sentiment with mutual fund flows. *Journal of Financial Economics*, *104*(2), 363–382.

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, *112*(6), 1269–1295.

Berk, J. B., Van Binsbergen, J. H., & Liu, B. (2017). Matching capital and labor. *Journal of Finance*, *72*(6), 2467–2504.

Bessembinder, H., Cooper, M. J., & Zhang, F. (2023). Mutual fund performance at long horizons. *Journal of Financial Economics*, *147*(1), 132–158.

Brown, K. C., Harlow, W. V., & Starks, L. T. (1996). Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance*, *51*(1), 85–110.

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57–82.

Chen, H., & Pennacchi, G. G. (2009). Does prior performance affect a mutual fund's choice of risk? Theory and further empirical evidence. *Journal of Financial and Quantitative Analysis*, *44*(4), 745–775.

Chen, J., Hong, H., Huang, M., & Kubik, J. D. (2004). Does fund size erode mutual fund performance? The role of liquidity and organization. *American Economic Review*, *94*(5), 1276–1302.

Chevalier, J., & Ellison, G. (1997). Risk taking by mutual funds as a response to incentives. *Journal of Political Economy*, *105*(6), 1167–1200.

Cremers, K. M., & Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, *22*(9), 3329–3365.

Di Maggio, M., Franzoni, F., Kogan, S., & Xing, R. (2023). *Avoiding idiosyncratic volatility: Flow sensitivity to individual stock returns* (Tech. Rep.). National Bureau of Economic Research.

Dou, W. W., Kogan, L., & Wu, W. (2023). Common fund flows: Flow hedging and factor pricing. *Journal of Finance*, Forthcoming.

Edelen, R. M., & Warner, J. B. (2001). Aggregate price effects of institutional trading: A study of mutual fund flow and market returns. *Journal of Financial Economics*, *59*(2), 195–220.

Elton, E. J., Gruber, M. J., & Blake, C. R. (2001). A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases. *Journal of Finance*, *56*(6), 2415–2430.

Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*, *6*(1), 1–22.

Evans, R. B. (2010). Mutual fund incubation. *The Journal of Finance*, *65*(4), 1581–1611.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance*, *65*(5), 1915–1947.

Grinblatt, M., & Titman, S. (1993). Performance measurement without benchmarks: An examination of mutual fund returns. *Journal of Business*, 47–68.

Huang, D., Li, J., & Wang, L. (2021). Are disagreements agreeable? evidence from information aggregation. *Journal of Financial Economics*, *141* (1), 83–101.

Huang, J., Sialm, C., & Zhang, H. (2011). Risk shifting and mutual fund performance. *Review of Financial Studies*, *24* (8), 2575–2616.

Huang, J., Wei, K. D., & Yan, H. (2007). Participation costs and the sensitivity of fund flows to past performance. *Journal of Finance*, *62* (3), 1273–1311.

Jiang, H., & Zheng, L. (2018). Active fundamental performance. *Review of Financial Studies*, *31* (12), 4688–4719.

Jones, C. S., & Mo, H. (2021). Out-of-sample performance of mutual fund predictors. *The Review of Financial Studies*, *34* (1), 149–193.

Kacperczyk, M., & Seru, A. (2007). Fund manager use of public information: New evidence on managerial skills. *Journal of Finance*, *62* (2), 485–528.

Kacperczyk, M., Sialm, C., & Zheng, L. (2008). Unobserved actions of mutual funds. *Review of Financial Studies*, *21* (6), 2379–2416.

Lee, J. H., Trzcinka, C., & Venkatesan, S. (2019). Do portfolio manager contracts contract portfolio management? *Journal of Finance*, *74* (5), 2543–2577.

Lou, D. (2012). A flow-based explanation for return predictability. *Review of Financial Studies*, *25* (12), 3457–3489.

Pástor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, *111* (3), 642–685.

Pástor, L., & Vorsatz, M. B. (2020). Mutual fund performance and flows during the COVID-19 crisis. *Review of Asset Pricing Studies*, *10* (4), 791–833.

Rakowski, D. (2010). Fund flow volatility and performance. *Journal of Financial and Quantitative Analysis*, *45* (1), 223–237.

Sialm, C., & Zhang, H. (2020). Tax-efficient asset management: Evidence from equity mutual funds. *Journal of Finance*, *75* (2), 735–777.

Sirri, E. R., & Tufano, P. (1998). Costly search and mutual fund flows. *Journal of Finance*, *53* (5), 1589–1622.

Stambaugh, R. F., Yu, J., & Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104* (2), 288–302.

Warther, V. A. (1999). Aggregate mutual fund flows and security returns. *Journal of Financial Economics*, *53* (3), 439–466.

Yu, J. (2011). Disagreement and return predictability of stock portfolios. *Journal of Financial Economics*, *99* (1), 162–183.

Zheng, L. (1999). Is money smart? A study of mutual fund investors' fund selection ability. *Journal of Finance*, *54* (3), 901–933.

Zhu, Q. (2020). The missing new funds. *Management Science*, *66* (3), 1193–1204.

Figure 1. Distribution of funds tilting away from stocks with high flow beta
This figure shows the distribution of the tilt away from stocks with high flow beta among U.S. domestic active equity mutual funds. Fund-level tilt for fund j is the time-series average of the tilting coefficients $\gamma_{1,j,q}$ estimated for each fund j and quarter q from the following the Fama-MacBeth regression

$$\omega_{j,i,q+1} - \omega_{m,i,q+1} = \gamma_{0,j,q} + \gamma_{1,j,q}\beta_{\text{flow},i,q} + \gamma_{2,j,q}\beta_{\text{market},i,q} + \varepsilon_{i,q+1},$$

where $\omega_{j,i,q+1} - \omega_{j,i,q+1}$ is the deviation of stock i in fund j from the market allocations in quarter $q+1$. β_{flow} is estimated following Dou et al. (2023) and described in details in Section 3.2, and β_{market} is estimated using a 60-month rolling regression of stock i's monthly excess returns on the market excess returns. Each variable is standardized to have a mean of 0 and standard deviation of 1. The red line illustrates the estimated coefficient obtained from the aggregate mutual fund portfolio as in Dou et al. (2023). The sample period is from 1994Q4 to 2021Q4.

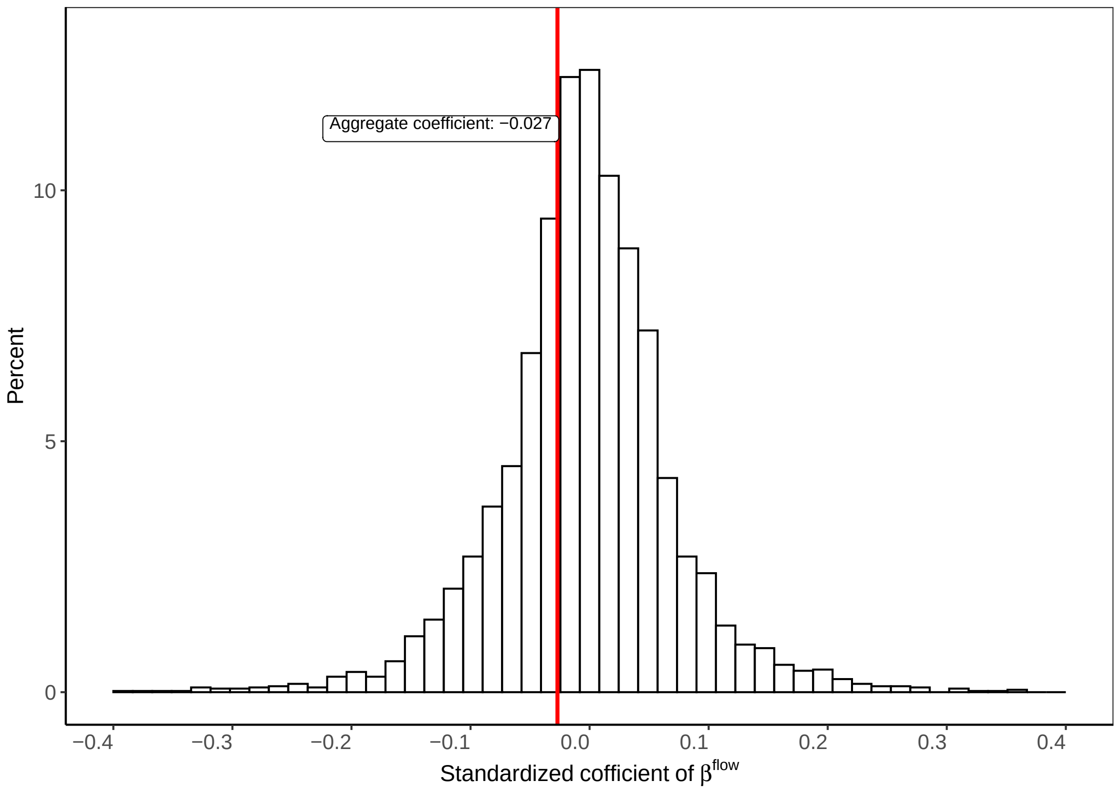


Figure 2. Common fund flows and net returns of the active flow beta strategy
This figure shows the time-series of the common fund flows and the net returns of the active flow beta (AFB) strategy for U.S. domestic active equity funds from 1994Q1 to 2021Q4. Panel A shows the time-series of the common fund flows, which are constructed following Dou et al. (2023) and described in details in Section 3.2. To construct the AFB-based strategy, at the beginning of each calendar quarter, funds are sorted into deciles according to their AFB and their performance is tracked for the subsequent quarter. The rebalancing frequency is quarterly. A fund's active flow beta is estimated as the sum of the product between the difference of the fund's portfolio weights from the market portfolio's weights and the underlying stock flow betas. I compute monthly equally-weighted average net returns on the portfolio that is long on the top decile portfolio and short on the bottom decile portfolio. Panel B shows the time-series net return of this portfolio.

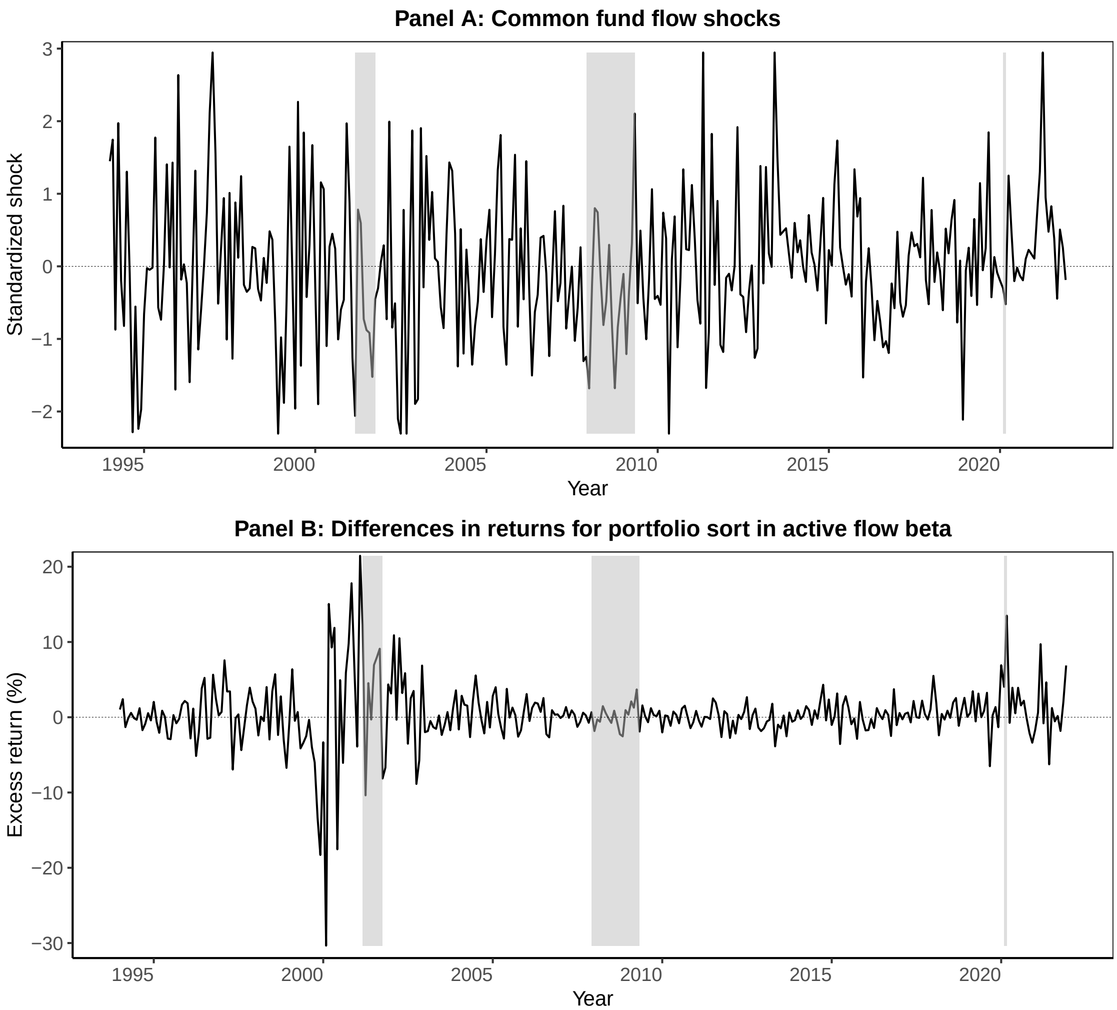
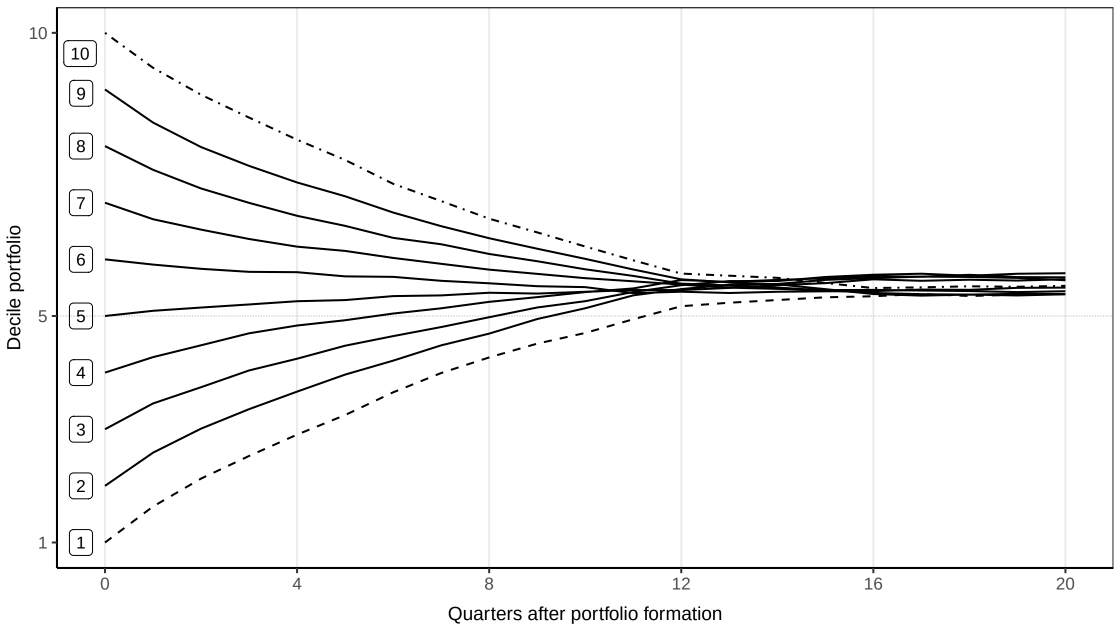


Figure 3. Persistence of flow risk management

This figure illustrates the average portfolio rank and active flow beta of mutual fund portfolios sorted on funds' active flow beta (AFB) over a 5-year (20 quarters) period between 1994 and 2021. A fund's active flow beta is estimated as the sum of the product between the difference of the fund's portfolio weights from the market portfolio's weights and the underlying stock flow betas. At the beginning of each calendar quarter from 1994Q1 to 2017Q1, funds are sorted into deciles according to their AFB. Each portfolio is subsequently tracked over the next 5-year period. Panel A reports the value-weighted average decile rank, and Panel B reports the value-weighted average active flow beta.

Panel A: Persistence in portfolio ranking



Panel B: Persistence in active flow beta

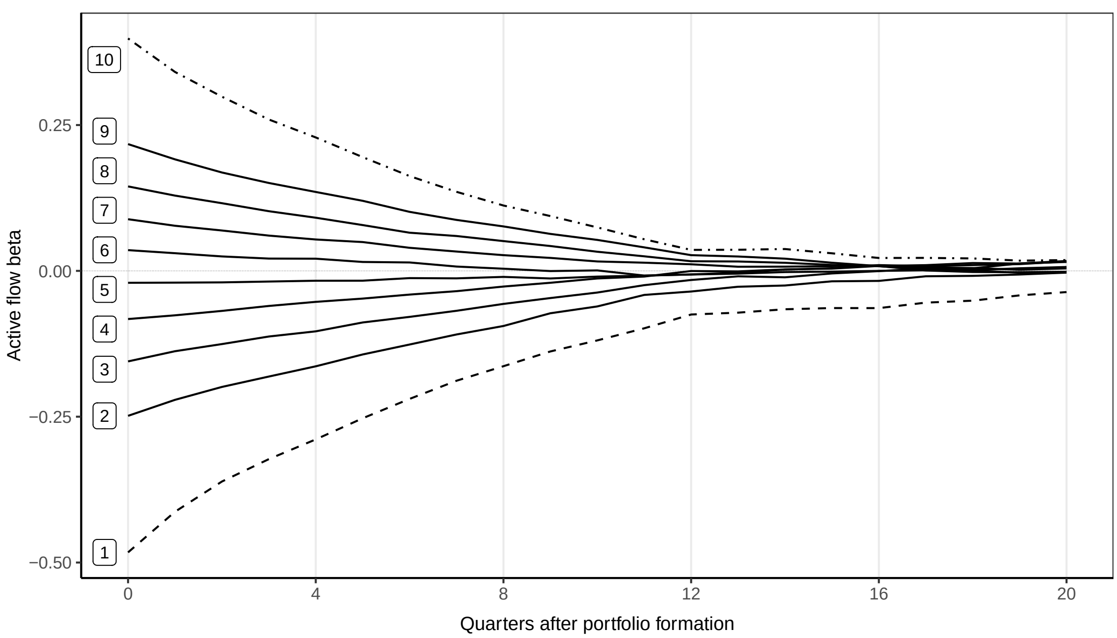


Table 1. Summary statistics

This table reports the summary statistics for the stock and the fund samples from 1994:01 to 2021:12. Panel A reports the statistics for the stock sample that includes U.S. common stocks listed on the NYSE, NASDAQ, and Amex and has at least 36 months of return observations on CRSP and two years of data on Compustat. β_{flow} is estimated monthly from a 36-month rolling regression of stocks' excess returns on the common flow shocks, controlling for the market factor. $\beta_{\text{liquidity}}$ is estimated monthly from a 36-month rolling regression of stocks' excess returns on the market liquidity factor, controlling for the market return, the size and value factors, the momentum factor. $\beta_{\text{uncertainty}}$ is estimated monthly from a 36-month rolling regression of stocks' excess returns on macro economic uncertainty shocks, controlling for the market return, the size and value factors, the momentum factor, the market liquidity factor, and the investment and profitability factors. *AIM* is the Amihud's illiquidity measure. Panel B reports the statistics for the fund sample that includes US active equity funds. *TNA* is the monthly total net fund assets. *Age* is the fund age in years. *Quarterly return* is the quarterly net fund return. *Quarterly flow* is the quarterly growth rate of assets under management. *Expense ratio* is the fund expense ratio. *Turnover ratio* is the turnover ratio of the fund.

	Mean	Standard deviation	p10	p25	Median	p75	p90
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Panel A: Mutual fund sample				
TNA ($ million)	1575.589	6307.106	38.385	94.225	300.696	982.349	2945.965
Age (Years)	13.252	8.042	3	7	12	19	27
Quarterly return (%)	2.719	5.628	-2.317	0.088	2.566	5.200	7.846
Quarterly flow (%)	1.031	12.477	-8.031	-4.094	-1.103	3.119	11.674
Expense ratio (%)	1.043	0.519	0.370	0.627	1.082	1.369	1.655
Turnover ratio (%)	80.670	90.169	16.881	32.725	59.975	101.226	157.476
			Panel B: Stock sample				
β_{flow}	0.017	0.410	-0.344	-0.148	0.004	0.162	0.368
β_{market}	1.104	0.597	0.353	0.677	1.048	1.467	1.932
Ln(Size)	5.812	2.043	3.190	4.300	5.729	7.205	8.564
$\text{Ln(Size)}_{\text{median}}$	0.004	0.749	-0.873	-0.353	0.046	0.396	0.841
Ln(BEME)	-0.653	0.891	-1.796	-1.160	-0.568	-0.092	0.331
$\beta_{\text{liquidity}}$	0.016	0.514	-0.557	-0.244	0.009	0.274	0.604
$\beta_{\text{uncertainty}}$	0.024	0.828	-0.869	-0.374	0.005	0.396	0.930
AIM	4.680	18.284	0.001	0.006	0.082	1.039	7.345

Table 2. Fund characteristics by portfolio sort

This table reports the summary statistics for fund characteristics by the sort of AFB from 1994Q1 to 2021Q4. *TNA* is the monthly total net fund assets. *Age* is the fund age in years. *Quarterly return* is the quarterly net fund return. *Quarterly flow* is the quarterly growth rate of assets under management. *Expense ratio* is the fund expense ratio. *Turnover ratio* is the turnover ratio of the fund. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Low	P2	P3	P4	P5	P6	P7	P8	P9	High	High-Low
TNA ($ million)	1059.165	1332.189	1686.287	1923.153	1967.306	1865.048	1856.949	1638.104	1397.423	957.410	-101.755
Age (Years)	18.523	18.488	18.729	18.632	18.724	18.936	18.874	18.986	19.063	19.249	0.725
Expense ratio (%)	1.174	1.062	1.028	1.006	0.991	0.988	0.988	0.997	1.021	1.082	-0.091***
Turnover ratio (%)	85.773	80.869	77.701	74.518	71.617	70.587	68.896	68.200	68.191	69.368	-16.405***
Quarterly flow (%)	0.442	0.387	0.171	0.301	0.322	0.520	0.623	0.783	0.975	1.228	0.787*
β_{flow}	-0.410	-0.260	-0.163	-0.088	-0.027	0.029	0.081	0.138	0.214	0.327	0.737***
Return gap (%)	-0.251	-0.239	-0.214	-0.184	-0.178	-0.165	-0.144	-0.131	-0.119	-0.101	0.150**
Active share (%)	87.440	82.073	78.785	77.723	77.221	77.093	78.209	80.604	83.371	88.890	1.451
R^2 (%)	0.862	0.892	0.899	0.900	0.898	0.900	0.901	0.897	0.887	0.848	-0.013
RPI (%)	12.432	11.596	10.939	10.841	10.974	10.755	10.742	10.783	10.987	12.020	-0.412
Expected FIT	3.536	3.528	3.525	3.521	3.519	3.518	3.518	3.518	3.530	3.537	0.001
AFP	0.100	0.126	0.119	0.126	0.123	0.132	0.135	0.142	0.155	0.161	0.061
AFO	-0.036	-0.171	-0.168	-0.216	-0.203	-0.236	-0.221	-0.213	-0.152	-0.129	-0.093***

Table 3. Active flow risk and mutual fund performance: Decile portfolios

This table reports the performance of decile fund portfolios sorted on their active flow beta (AFB) from 1994Q1 to 2021Q4. A fund's active flow beta is estimated as the sum of the product between the difference of the fund's portfolio weights from the market portfolio's weights and the underlying stock flow betas. Following Dou et al. (2023), each stock's flow beta is estimated at June of year t as its average flow betas from January to June of year t and remains from July of year t to June of year $t+1$. At the beginning of each calendar quarter, funds are sorted into deciles according to their active flow beta and their performance is tracked for the subsequent three months. The rebalancing frequency is quarterly. *High* (*Low*) is the top (bottom) decile portfolio. I compute monthly equally-weighted average net returns on the portfolios, and report the average excess returns (Panel A), and the risk-adjusted returns based on the Carhart (1997) four-factor model (Panel B), and the Pástor and Stambaugh (2003) five-factor model (Panel C). Panels B and C also report the loadings on corresponding risk factors. The alphas are reported in monthly percentage. Newey-West adjusted $t-$statistics are shown in square brackets. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Low	2	3	4	5	6	7	8	9	High	High-Low
	Panel A: Excess return										
α	0.48	0.57	0.58	0.62	0.67	0.72	0.77	0.79	0.85	0.90	0.43*
	[1.34]	[1.83]	[2.07]	[2.34]	[2.64]	[2.87]	[3.19]	[3.21]	[3.43]	[3.53]	[1.71]
	Panel B: Carhart's (1997) four-factor model										
α	-0.43	-0.31	-0.25	-0.19	-0.11	-0.05	0.04	0.05	0.12	0.17	0.60***
	[-3.26]	[-3.69]	[-4.08]	[-3.66]	[-2.55]	[-0.91]	[0.71]	[0.67]	[1.38]	[1.51]	[2.78]
β_{MKTRF}	1.14	1.07	1.03	1.00	0.98	0.96	0.93	0.94	0.92	0.92	-0.23***
	[27.86]	[38.51]	[54.80]	[62.47]	[76.20]	[61.82]	[56.54]	[40.75]	[33.31]	[25.69]	[-3.27]
β_{SMB}	0.53	0.44	0.31	0.21	0.16	0.12	0.09	0.07	0.08	0.11	-0.42***
	[7.30]	[7.17]	[8.89]	[7.26]	[11.44]	[5.56]	[3.00]	[1.99]	[1.46]	[1.64]	[-3.18]
β_{HML}	-0.41	-0.21	-0.12	-0.04	0.01	0.07	0.13	0.17	0.21	0.23	0.64***
	[-5.22]	[-3.97]	[-3.63]	[-1.54]	[0.54]	[3.12]	[4.46]	[4.88]	[4.39]	[3.63]	[5.05]
β_{UMD}	-0.04	0.04	0.02	0.02	0.01	0.00	-0.01	-0.01	-0.02	-0.01	0.02
	[-0.75]	[1.24]	[1.18]	[1.11]	[0.47]	[-0.10]	[-0.81]	[-0.44]	[-0.56]	[-0.33]	[0.31]
	Panel C: Carhart's (1997) four-factor model + Pástor and Stambaugh's (2003) liquidity factor										
α	-0.41	-0.30	-0.24	-0.18	-0.11	-0.06	0.02	0.02	0.08	0.12	0.53**
	[-3.09]	[-3.44]	[-3.92]	[-3.56]	[-2.46]	[-1.16]	[0.38]	[0.33]	[1.03]	[1.16]	[2.50]
β_{MKTRF}	1.16	1.08	1.04	1.00	0.98	0.95	0.92	0.92	0.90	0.89	-0.27***
	[29.19]	[40.36]	[55.74]	[64.39]	[80.94]	[63.73]	[65.68]	[46.79]	[37.35]	[27.86]	[-4.33]
β_{SMB}	0.54	0.44	0.31	0.21	0.15	0.12	0.08	0.06	0.07	0.10	-0.45***
	[7.58]	[7.45]	[9.19]	[7.24]	[11.05]	[5.59]	[2.90]	[1.84]	[1.33]	[1.51]	[-3.48]
β_{HML}	-0.42	-0.21	-0.13	-0.04	0.01	0.08	0.13	0.18	0.22	0.25	0.67***
	[-5.16]	[-3.95]	[-3.60]	[-1.52]	[0.56]	[3.35]	[4.71]	[5.22]	[4.73]	[3.93]	[5.13]
β_{UMD}	-0.03	0.04	0.03	0.02	0.01	0.00	-0.02	-0.02	-0.03	-0.03	0.00
	[-0.61]	[1.48]	[1.38]	[1.14]	[0.45]	[-0.30]	[-1.12]	[-0.80]	[-0.91]	[-0.66]	[0.04]
β_{LIQ}	-0.08	-0.06	-0.03	-0.01	0.00	0.04	0.06	0.08	0.11	0.15	0.22***
	[-1.86]	[-2.39]	[-1.86]	[-0.44]	[0.19]	[2.31]	[3.61]	[3.96]	[4.47]	[4.82]	[3.57]

Table 4. Active flow risk and mutual fund performance: Double portfolio sorts

This table reports the performance of fund portfolios sorted on their active flow beta (AFB) and skill-related fund characteristics from 1994Q1 to 2021Q4. At the beginning of each calendar quarter, I independently sort funds into four groups based on AFB and into four groups based on the following fund characteristics: Kacperczyk et al.'s (2008) *Return gap* (Panel A), Cremers and Petajisto's (2009) *Active share* (Panel B), Kacperczyk and Seru's (2007) *Reliance on public information* (Panel C), Jiang and Zheng's (2018) *Active fundamental performance* (Panel D), and Avramov et al.'s (2020) *Active fund overpricing* (Panel E). Construction details of these fund predictors can be found in Section B.2 in the Appendices. The portfolio performance is tracked for the subsequent three months. The rebalancing frequency is quarterly. I compute monthly equally-weighted average net returns on the portfolios, and report the risk-adjusted returns based on the Carhart (1997) four-factor model. The alphas are reported in monthly percentage. Newey-West adjusted $t-$statistics are shown in square brackets. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Active flow risk	All	Low	2	3	High	High-Low
			Panel A: Return gap			
All		-0.35***	-0.15***	0.00	0.13	0.48**
		[-3.04]	[-2.72]	[0.08]	[1.39]	[2.59]
Low	-0.25***	-0.45	-0.24	-0.08	0.01	0.46**
	[-3.73]	[-3.42]	[-3.55]	[-1.21]	[0.12]	[2.36]
2	-0.12***	-0.38	-0.14	-0.01	0.12	0.51***
	[-2.78]	[-3.21]	[-2.20]	[-0.30]	[1.36]	[2.69]
3	-0.04	-0.32	-0.12	0.05	0.16	0.48***
	[-1.05]	[-2.88]	[-2.11]	[0.94]	[1.74]	[2.68]
High	-0.01	-0.26	-0.07	0.06	0.19	0.45**
	[-0.24]	[-2.41]	[-1.37]	[0.99]	[1.89]	[2.45]
High-Low	0.24***	0.19***	0.16***	0.14**	0.18***	
	[4.26]	[3.15]	[4.12]	[2.28]	[3.17]	

(Continued on next page)

Table 4 (*continued*)

Active flow risk	All	Low	2	3	High	High-Low
			Panel B: Active share			
All		-0.35***	-0.15***	0.00	0.13	0.48**
		[-3.04]	[-2.72]	[0.08]	[1.39]	[2.59]
Low	-0.07***	-0.32	-0.16	-0.01	0.18	0.50***
	[-2.97]	[-3.11]	[-3.90]	[-0.20]	[2.05]	[2.93]
2	-0.06	-0.24	-0.10	-0.01	0.10	0.34**
	[-1.34]	[-2.62]	[-1.44]	[-0.17]	[1.07]	[2.19]
3	-0.05	-0.30	-0.11	0.00	0.17	0.47***
	[-0.84]	[-2.61]	[-1.88]	[-0.05]	[1.64]	[2.60]
High	-0.04	-0.28	-0.18	0.09	0.12	0.41**
	[-0.64]	[-2.50]	[-3.24]	[1.17]	[1.17]	[2.34]
High-Low	0.03	0.04	-0.01	0.10	-0.05	
	[0.63]	[0.47]	[-0.22]	[1.49]	[-0.86]	
			Panel C: Reliance on public information			
All		-0.35***	-0.15***	0.00	0.13	0.48**
		[-3.04]	[-2.72]	[0.08]	[1.39]	[2.59]
Low	-0.08*	-0.33	-0.12	-0.01	0.18	0.51***
	[-1.75]	[-3.00]	[-1.81]	[-0.24]	[2.01]	[2.87]
2	-0.11**	-0.36	-0.20	0.03	0.11	0.47**
	[-2.37]	[-2.97]	[-3.83]	[0.51]	[1.18]	[2.46]
3	-0.10**	-0.34	-0.15	0.01	0.15	0.49***
	[-2.17]	[-3.01]	[-2.44]	[0.16]	[1.53]	[2.65]
High	-0.11**	-0.40	-0.13	0.00	0.08	0.47**
	[-2.39]	[-3.29]	[-2.12]	[0.03]	[0.75]	[2.43]
High-Low	-0.04	-0.07**	-0.01	0.01	-0.11***	
	[-1.52]	[-2.20]	[-0.17]	[0.29]	[-2.59]	

(Continued on next page)

Table 4 (*continued*)

Active flow risk	All	Low	2	3	High	High-Low
		Panel D: Active fundamental performance				
All		-0.35***	-0.15***	0.00	0.13	0.48**
		[-3.04]	[-2.72]	[0.08]	[1.39]	[2.59]
Low	-0.15**	-0.41	-0.17	0.00	0.07	0.47**
	[-2.13]	[-3.16]	[-2.36]	[-0.01]	[0.72]	[2.57]
2	-0.12***	-0.42	-0.16	-0.04	0.13	0.54***
	[-2.86]	[-3.80]	[-2.78]	[-0.91]	[1.41]	[3.12]
3	-0.10***	-0.34	-0.17	0.01	0.12	0.46**
	[-2.64]	[-3.33]	[-3.30]	[0.20]	[1.22]	[2.58]
High	-0.03	-0.27	-0.07	0.04	0.15	0.41**
	[-0.52]	[-2.37]	[-1.09]	[0.56]	[1.37]	[2.14]
High-Low	0.12*	0.14**	0.10	0.04	0.08	
	[1.79]	[2.43]	[1.44]	[0.48]	[1.29]	
		Panel E: Active fund overpricing				
All		-0.35***	-0.15***	0.00	0.13	0.48**
		[-3.04]	[-2.72]	[0.08]	[1.39]	[2.59]
Low	-0.08*	-0.32	-0.17	0.02	0.15	0.47**
	[-1.83]	[-2.94]	[-3.05]	[0.32]	[1.49]	[2.56]
2	-0.07	-0.37	-0.09	0.01	0.16	0.53***
	[-1.44]	[-3.05]	[-1.22]	[0.19]	[1.77]	[2.80]
3	-0.10**	-0.39	-0.14	0.01	0.14	0.53***
	[-2.39]	[-3.56]	[-2.48]	[0.12]	[1.45]	[2.88]
High	-0.14***	-0.33	-0.19	-0.02	0.06	0.40**
	[-2.72]	[-2.72]	[-3.72]	[-0.47]	[0.68]	[2.13]
High-Low	-0.05*	-0.01	-0.02	-0.04	-0.08***	
	[-1.72]	[-0.29]	[-0.69]	[-1.47]	[-3.10]	

Table 5. Active flow risk and mutual fund performance: Predictive panel regressions
This table reports results from predictive panel regressions of fund performance on active flow beta (AFB) and other fund characteristics from 1994Q1 to 2021Q4. A fund's active flow beta is estimated as the sum of the product between the difference of the fund's portfolio weights from the market portfolio's weights and the underlying stock flow betas. I estimate the fund performance each quarter as the monthly average of the risk-adjusted returns, where the monthly risk-adjusted returns are obtained from 36-month rolling regressions of funds' excess returns on the Carhart (1997) four-factor model. The measures of skill include Kacperczyk et al.'s (2008) *Return gap*, Cremers and Petajisto's (2009) *Active share*, Kacperczyk and Seru's (2007) *Reliance on public information* (RPI), Amihud and Goyenko's (2013) R^2, Lou's (2012) *Expected flow-induced trading* (Expected FIT), Jiang and Zheng's (2018) *Active fundamental performance* (AFP), and Avramov et al.'s (2020) *Active fund overpricing* (AFO). Construction details of these fund predictors can be found in Section B.2 in the Appendices. All measures of skill and fund characteristics are winsorized at the top and the bottom 5% and standardized to have mean of 0 and standard deviation of 1. Coefficient estimates are reported in basis points. The regressions include time fixed effects. Standard errors are clustered by fund and are shown in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Active flow beta	3.382***	2.315***	2.170***	2.223***	2.337***	2.190***	2.307***	2.466***	2.280***	2.168***
	(0.450)	(0.523)	(0.565)	(0.546)	(0.585)	(0.539)	(0.544)	(0.557)	(0.544)	(0.619)
Return gap			2.266**							1.768
			(1.007)							(1.144)
RPI				-1.208*						-1.215
				(0.692)						(0.746)
Active share					3.232*					5.370*
					(1.881)					(2.826)
R^2						-1.783				-3.602*
						(1.092)				(1.919)
Expected FIT							3.075***			3.024***
							(0.608)			(0.723)
AFP								5.357***		5.372***
								(0.988)		(1.072)
AFO									-2.300***	-1.839**
									(0.704)	(0.758)
Ln(TNA)		-3.051***	-3.442***	-3.417***	-4.139***	-3.304***	-3.467***	-3.585***	-3.381***	-4.687***
		(1.093)	(1.140)	(1.185)	(1.376)	(1.141)	(1.160)	(1.211)	(1.160)	(1.454)
Ln(Age)		-2.711*	-3.237**	-3.149*	-4.018**	-3.112**	-3.139*	-3.324**	-3.204**	-3.863**
		(1.418)	(1.594)	(1.613)	(1.886)	(1.572)	(1.606)	(1.682)	(1.610)	(1.859)
Expense ratio		-1.919*	-1.964*	-1.905*	-1.241	-2.326**	-2.021**	-1.886*	-1.982*	-1.315
		(1.033)	(1.033)	(1.055)	(1.139)	(1.095)	(1.031)	(1.082)	(1.037)	(1.162)
Turnover ratio		-3.357***	-2.858***	-3.501***	-3.644***	-3.486***	-3.430***	-3.507***	-3.439***	-3.704***
		(0.743)	(0.775)	(0.768)	(0.795)	(0.746)	(0.746)	(0.775)	(0.754)	(0.905)
Past flow		0.785*	0.600	0.917**	1.235***	0.886*	0.624	0.700	0.685	0.090
		(0.446)	(0.461)	(0.456)	(0.470)	(0.468)	(0.454)	(0.471)	(0.477)	(0.514)
Past performance		6.534***	6.093***	6.401***	6.389***	6.409***	6.358***	6.082***	6.391***	5.689***
		(1.027)	(1.137)	(1.083)	(1.136)	(1.073)	(1.071)	(1.126)	(1.072)	(1.226)
Adj. R^2	0.0243	0.0214	0.0208	0.0207	0.0199	0.0207	0.0208	0.0200	0.0208	0.0196
# of obs.	186,712	147,360	139,706	138,841	129,758	139,557	139,706	133,423	139,706	124,394

Table 6. Active flow risk and mutual fund performance: Decomposition of active flow beta

This table reports results from predictive panel regressions of fund performance on decomposed components of active flow beta (AFB) and other fund characteristics from 1994Q1 to 2021Q4. Specifically, I decompose AFB into three components according to Equation 16 and use the correlation component (ρ_{AFB}) and activeness component ($STDAS$). I estimate the fund performance each quarter as the monthly average of the risk-adjusted returns, where the monthly risk-adjusted returns are obtained from 36-month rolling regressions of funds' excess returns on the Carhart (1997) four-factor model. The measures of skill include Kacperczyk et al.'s (2008) *Return gap*, Cremers and Petajisto's (2009) *Active share*, Kacperczyk and Seru's (2007) *Reliance on public information* (RPI), Amihud and Goyenko's (2013) R^2, Lou's (2012) *Expected flow-induced trading* (Expected FIT), Jiang and Zheng's (2018) *Active fundamental performance* (AFP), and Avramov et al.'s (2020) *Active fund overpricing* (AFO). Construction details of these fund predictors can be found in Section B.2 in the Appendices. Control variables include fund size, age, expense ratio, turnover ratio, past flow and past performance. All measures of skill and fund characteristics are winsorized at the top and the bottom 5% and standardized to have mean of 0 and standard deviation of 1. Coefficient estimates are reported in basis points. The regressions include time fixed effects. Standard errors are clustered by fund and are shown in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\boldsymbol{\rho}_{\text{AFB}}$	3.419***	2.139***	2.075***	2.121***	2.224***	2.098***	2.162***	2.302***	2.179***	1.972***
	(0.526)	(0.636)	(0.679)	(0.666)	(0.712)	(0.649)	(0.664)	(0.681)	(0.665)	(0.721)
STDAS	3.564***	3.382***	3.161***	3.257***	3.324***	3.210***	3.337***	3.423***	3.290***	2.299*
	(0.931)	(1.092)	(1.161)	(1.155)	(1.202)	(1.138)	(1.157)	(1.203)	(1.157)	(1.174)
Return gap			1.886*							1.482
			(1.004)							(1.111)
RPI				-0.786						-0.912
				(0.664)						(0.716)
Active share					2.590					4.687*
					(1.796)					(2.786)
R^2						-1.452				-3.105
						(1.055)				(1.895)
Expected FIT							2.997***			2.991***
							(0.606)			(0.721)
AFP								5.314***		5.354***
								(0.982)		(1.066)
AFO									-2.176***	-1.770**
									(0.703)	(0.764)
Control variables		✓	✓	✓	✓	✓	✓	✓	✓	✓
Adj. R^2	0.0246	0.0215	0.0209	0.0208	0.0200	0.0209	0.0209	0.0201	0.0209	0.0196
# of obs.	186,712	147,360	139,706	138,841	129,758	139,557	139,706	133,423	139,706	124,394

Table 7. Deviation of holdings and precision of public information
This table reports results from Fama-MacBeth regressions of fund portfolios' deviations from benchmark allocations on AFB and its interaction with measures of precision of public information from 1994Q1 to 2021Q4. Specifically, I estimate the following regression

$$\omega_{p,i,q+1} - \omega_{m,i,q+1} = \gamma_{0,p,q} + \gamma_{1,p,q}\beta_{\text{flow},i,q} + \gamma_{2,p,q}\beta_{\text{market},i,q} + \gamma_{3,p,q}\sigma_{i,q} + \gamma_{4,p,q}\beta_{\text{flow},i,q} \times \sigma_{i,q} + \varepsilon_{p,q+1},$$

where $\omega_{p,i,q+1} - \omega_{m,i,q+1}$ is the deviation of stock i in portfolio's p from the market allocations and $\sigma_{i,q}$ is a measure of precision of public information for stock i. At the beginning of each calendar quarter, funds are sorted into deciles according to their active flow beta, where *High* (*Low*) is the top (bottom) decile portfolio. Panel A (B) uses the analysts' forecast dispersion (stocks' idiosyncratic volatility) to proxy for the imprecision of public information. All variables are standardized to have mean of 0 and standard deviation of 1. Coefficient estimates are reported in monthly percentage. Standard errors are Newey-West adjusted and shown in parentheses. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Portfolio	Panel A: Analysts' disagreement				Panel B: Idiosyncratic volatility			
	β_{flow}	β_{market}	σ	$\beta_{\text{flow}} \times\sigma$	β_{flow}	β_{market}	σ	$\beta_{\text{flow}} \times\sigma$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Low (P1)	-0.074***	-0.022**	-0.158***	0.018**	-0.094***	0.029***	-0.304***	0.058***
	(0.006)	(0.009)	(0.020)	(0.009)	(0.007)	(0.008)	(0.018)	(0.007)
High (P10)	0.081***	-0.053***	-0.103***	-0.017	0.089***	0.012	-0.300***	-0.055***
	(0.005)	(0.011)	(0.020)	(0.013)	(0.005)	(0.011)	(0.021)	(0.008)
High-Low	0.155***	-0.032*	0.056	-0.035**	0.183***	-0.017	0.004	-0.113***
	(0.006)	(0.017)	(0.040)	(0.017)	(0.009)	(0.022)	(0.044)	(0.010)

Table 8. Precision of public information and mutual fund performance
This table reports the performance of decile fund portfolios sorted on their AFB conditional on the periods of high variance in public information from 1994:01 to 2021:12. A fund's active flow beta is estimated as the sum of the product between the difference of the fund's portfolio weights from the market portfolio's weights and the underlying stock flow betas. At the beginning of each calendar quarter, funds are sorted into deciles according to their active flow beta, where *High* (*Low*) is the top (bottom) decile portfolio. Panel A (B) use the analysts' forecast dispersion (stocks' idiosyncratic volatility) to proxy for the imprecision of public information. *Volatility indicator* is an indicator equal to 1 if the month belongs to the top decile of the imprecision of public information. I compute monthly equally-weighted average net returns on the portfolios, and report the risk-adjusted returns based on the Carhart (1997) four-factor model. The alphas are reported in monthly percentage. Newey-West adjusted $t-$statistics are shown in square brackets. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Panel A: Analysts' disagreement		Panel B: Idiosyncratic volatility	
Portfolio	Carhart α	Volatility indicator	Carhart α	Volatility indicator
	(1)	(2)	(3)	(4)
Low (P1)	-0.33***	-0.98**	-0.36***	-0.59
	[-2.66]	[-2.31]	[-3.11]	[-1.02]
High (P10)	0.10	0.73**	0.05	1.11**
	[0.88]	[2.52]	[0.47]	[2.48]
High-Low	0.42**	1.71***	0.41**	1.70**
	[2.12]	[2.69]	[2.09]	[2.05]

Appendices

A Additional Tables and Figures

Table A1. Summary statistics: Stock characteristics across portfolio sort
This table reports the summary statistics for decile stock portfolios sorted on stock flow beta from 1994:01 to 2021:12. The sample includes U.S. common stocks listed on the NYSE, NASDAQ, and Amex and has at least 36 months of return observations. Following Dou et al. (2023), in June of each year t I sort all stocks into ten portfolios based on the average flow betas from January to June and track the portfolios from July of year t to June of year $t+1$. β_{flow} is estimated monthly from a 36-month rolling regression of stocks' excess returns on the common flow shocks, controlling for the market excess return. $\beta_{\text{liquidity}}$ is estimated monthly from a 36-month rolling regression of stocks' excess returns on the market liquidity factor, controlling for the market return, the size and value factors, the momentum factor. $\beta_{\text{uncertainty}}$ is estimated monthly from a 36-month rolling regression of stocks' excess returns on macro economic uncertainty shocks, controlling for the market return, the size and value factors, the momentum factor, the market liquidity factor, and the investment and profitability factors. *AIM* is the Amihud's illiquidity measure.

	Low	2	3	4	5	6	7	8	9	High
β_{flow}	-2.985	-1.462	-0.859	-0.458	-0.141	0.159	0.465	0.839	1.381	2.723
Return (%)	0.747	0.770	0.670	0.875	0.773	0.843	0.778	0.964	1.000	1.085
β_{market}	1.202	1.035	0.894	0.850	0.807	0.789	0.770	0.852	0.913	1.075
Ln(Size)	8.423	9.067	9.579	9.713	9.889	9.864	9.853	9.668	9.220	8.255
Ln(Size)$_{\text{median}}$	0.272	0.222	0.183	0.173	0.154	0.184	0.176	0.181	0.167	0.248
Ln(BEME)	-1.264	-1.201	-1.157	-1.092	-1.105	-1.076	-1.125	-1.137	-1.086	-1.162
$\beta_{\text{liquidity}}$	-0.040	-0.020	-0.039	-0.012	-0.008	0.005	-0.006	0.027	0.027	0.024
$\beta_{\text{uncertainty}}$	-0.025	0.008	0.004	-0.001	0.016	0.009	0.024	-0.007	-0.013	0.084
AIM	0.223	0.081	0.046	0.034	0.027	0.030	0.030	0.045	0.088	0.435

Table A2. Active flow risk and mutual fund performance: Alternatives of flow beta

This table examines the performance of decile fund portfolios sorted on their active flow beta from 1994:01 to 2021:12 as in Table 3 but uses different measures of flow beta. At the beginning of each calendar quarter, funds are sorted into deciles according to the active flow beta and their performance is tracked for the subsequent three months. *High* (*Low*) is the top (bottom) decile portfolio. I compute monthly equally-weighted average net returns on the portfolios, and report the average excess returns, the risk-adjusted returns based on CAPM, the Carhart (1997) four-factor model and the Pástor and Stambaugh (2003) five-factor model. Panel A reports the results in which the common flows are estimated based on traditional flow definition. Panel B reports the results in which the common flows are estimated based on the asset-weighted average of fund-specific shocks. Panel C reports the results in which the fund-level AFB is estimated as the direct covariance measure. The values are reported in monthly percentage. Newey-West adjusted $t-$statistics are shown in square brackets. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Low	2	3	4	5	6	7	8	9	High	High-Low
						Panel A: $\beta_{\text{flow},1}$					
Average	0.50	0.61	0.58	0.63	0.65	0.69	0.74	0.78	0.82	0.93	0.43**
	[1.50]	[2.19]	[2.21]	[2.45]	[2.60]	[2.78]	[2.97]	[3.10]	[3.15]	[3.35]	[2.21]
CAPM	-0.44	-0.22	-0.22	-0.15	-0.12	-0.07	-0.02	0.02	0.05	0.15	0.59***
	[-3.03]	[-2.69]	[-3.72]	[-3.35]	[-2.95]	[-1.46]	[-0.33]	[0.34]	[0.64]	[1.25]	[2.86]
Carhart	-0.34	-0.21	-0.21	-0.15	-0.12	-0.08	-0.04	0.00	0.02	0.10	0.44**
	[-2.99]	[-3.04]	[-3.67]	[-3.31]	[-3.30]	[-1.90]	[-0.94]	[0.01]	[0.34]	[1.11]	[2.57]
PS	-0.32	-0.20	-0.21	-0.15	-0.13	-0.09	-0.05	-0.02	0.00	0.07	0.40**
	[-2.92]	[-2.99]	[-3.68]	[-3.35]	[-3.35]	[-2.17]	[-1.30]	[-0.32]	[0.01]	[0.83]	[2.39]
						Panel B: $\beta_{\text{flow},2}$					
Average	0.59	0.64	0.64	0.66	0.67	0.69	0.71	0.76	0.78	0.81	0.22
	[1.65]	[2.03]	[2.21]	[2.46]	[2.64]	[2.81]	[2.94]	[3.13]	[3.21]	[3.25]	[0.91]
CAPM	-0.39	-0.25	-0.20	-0.14	-0.10	-0.06	-0.03	0.03	0.06	0.09	0.48**
	[-2.27]	[-2.03]	[-2.21]	[-2.24]	[-2.12]	[-1.45]	[-0.59]	[0.39]	[0.67]	[0.85]	[1.97]
Carhart	-0.33	-0.24	-0.20	-0.14	-0.11	-0.07	-0.04	0.01	0.03	0.07	0.40**
	[-2.56]	[-2.78]	[-3.13]	[-3.26]	[-2.94]	[-1.70]	[-0.80]	[0.11]	[0.36]	[0.69]	[2.02]
PS	-0.31	-0.22	-0.19	-0.15	-0.11	-0.08	-0.05	-0.02	0.00	0.04	0.34*
	[-2.37]	[-2.62]	[-3.01]	[-3.31]	[-3.11]	[-1.99]	[-1.11]	[-0.25]	[0.00]	[0.36]	[1.73]
						Panel C: $\text{Cov}(w_i^f - w_{\text{i}}^{\text{mkt}}, \beta_{\text{flow}})$					
Average	0.66	0.67	0.68	0.69	0.69	0.72	0.69	0.72	0.70	0.71	0.05
	[2.22]	[2.44]	[2.57]	[2.66]	[2.70]	[2.79]	[2.72]	[2.82]	[2.72]	[2.66]	[0.59]
CAPM	-0.20	-0.15	-0.12	-0.10	-0.09	-0.06	-0.08	-0.05	-0.08	-0.09	0.12
	[-2.08]	[-2.31]	[-2.25]	[-1.96]	[-1.74]	[-1.17]	[-1.57]	[-0.99]	[-1.55]	[-1.21]	[1.45]
Carhart	-0.19	-0.16	-0.13	-0.12	-0.10	-0.07	-0.08	-0.05	-0.07	-0.06	0.12*
	[-3.01]	[-3.43]	[-3.34]	[-3.64]	[-2.85]	[-1.83]	[-1.96]	[-1.08]	[-1.54]	[-0.93]	[1.92]
PS	-0.18	-0.16	-0.14	-0.13	-0.10	-0.08	-0.10	-0.07	-0.08	-0.08	0.11*
	[-2.93]	[-3.33]	[-3.47]	[-3.82]	[-3.06]	[-2.04]	[-2.30]	[-1.41]	[-1.87]	[-1.16]	[1.68]

Table A3. Active flow risk and mutual fund performance: Alternatives of benchmark

This table examines the performance of decile fund portfolios sorted on their active flow beta from 1994:01 to 2021:12 as in Table 3 but uses different benchmark definition for the funds to estimate its portfolio weights' deviation. At the beginning of each calendar quarter, funds are sorted into deciles according to the active flow beta and their performance is tracked for the subsequent three months. *High* (*Low*) is the top (bottom) decile portfolio. I compute monthly equally-weighted average net returns on the portfolios, and report the average excess returns, the risk-adjusted returns based on CAPM, the Carhart (1997) four-factor model and the Pástor and Stambaugh (2003) five-factor model. Panel A reports the results in which the benchmark is defined as in Cremers and Petajisto (2009). Particularly, a fund's specific benchmark is one of the 21 indices that minimizes its active share. Panel B reports the results in which the benchmar is the fund's holdings in previous quarter. The values are reported in monthly percentage. Newey-West adjusted $t-$statistics are shown in square brackets. $***$, $**$, and $*$ represent statistical significance at the 1%, 5%, and 10% levels, respectively.

	Low	2	3	4	5	6	7	8	9	High	High-Low
	Panel A: Cremers and Petajisto's (2009) benchmark										
Average	0.57	0.60	0.57	0.62	0.64	0.70	0.74	0.79	0.83	0.89	0.32*
	[1.68]	[2.04]	[2.08]	[2.44]	[2.55]	[2.87]	[3.04]	[3.21]	[3.31]	[3.34]	[1.87]
CAPM	-0.38	-0.27	-0.26	-0.16	-0.13	-0.05	0.00	0.05	0.10	0.13	0.52***
	[-2.63]	[-2.90]	[-3.79]	[-3.59]	[-4.34]	[-1.26]	[0.01]	[0.75]	[1.09]	[1.20]	[2.84]
Carhart	-0.32	-0.25	-0.25	-0.17	-0.14	-0.05	-0.01	0.04	0.09	0.10	0.42***
	[-2.98]	[-3.37]	[-4.41]	[-4.31]	[-4.58]	[-1.41]	[-0.27]	[0.55]	[1.03]	[1.11]	[2.65]
PS	-0.31	-0.24	-0.24	-0.17	-0.14	-0.06	-0.03	0.02	0.06	0.07	0.38**
	[-2.86]	[-3.20]	[-4.26]	[-4.11]	[-4.63]	[-1.66]	[-0.63]	[0.24]	[0.69]	[0.80]	[2.40]
	Panel B: Trade-based measure										
Average	0.66	0.64	0.65	0.67	0.70	0.67	0.65	0.68	0.77	0.81	0.15
	[2.16]	[2.19]	[2.38]	[2.55]	[2.79]	[2.74]	[2.66]	[2.74]	[2.98]	[3.10]	[1.27]
CAPM	-0.23	-0.22	-0.17	-0.13	-0.06	-0.07	-0.10	-0.07	-0.01	0.03	0.26**
	[-2.26]	[-2.42]	[-2.61]	[-2.56]	[-1.42]	[-1.42]	[-1.80]	[-1.11]	[-0.09]	[0.40]	[2.17]
Carhart	-0.22	-0.21	-0.17	-0.12	-0.07	-0.07	-0.10	-0.07	-0.02	0.01	0.22***
	[-3.63]	[-3.78]	[-3.65]	[-2.97]	[-1.54]	[-1.47]	[-2.04]	[-1.33]	[-0.36]	[0.11]	[2.64]
PS	-0.21	-0.20	-0.16	-0.13	-0.07	-0.08	-0.11	-0.08	-0.04	-0.02	0.19**
	[-3.51]	[-3.65]	[-3.58]	[-2.96]	[-1.70]	[-1.82]	[-2.36]	[-1.68]	[-0.79]	[-0.28]	[2.35]

B Supplemental Materials

B.1 Model Details

In this section, I provide additional details on the solution of the theoretical model in Section 2.1. Equation 8 shows that an investor's demand for the risky asset depends on her posterior about the asset's risk, return, flow and the covariance between its return and flow

$$x^{I*} = \frac{1}{\gamma}\frac{E_{\mathbf{s}}(u^I - p)}{\mathrm{Var}_{\mathbf{s}}(u^I)} - \beta^I_{\text{flow}}\frac{\mathrm{Var}_{\mathbf{s}}(F^I)}{\mathrm{Var}_{\mathbf{s}}(u^I)}. \tag{B1}$$

Using Bayes' rule, we can obtain the posterior distribution of the asset value and flow. For informed investor I, the distribution is bivariate normal with the conditional mean and variance-covariance matrix given as

$$u^I, F^I | s^I \sim N\left(\begin{bmatrix} \frac{\rho_1(\rho_2+\rho_F)\bar{u}-\rho_1\psi\bar{F}+\kappa_1 s_1+\rho_1\psi s_2}{\rho_1(\rho_2+\rho_F)+\kappa_1} \\ \frac{-\rho_2\psi\bar{u}+\rho_2(\rho_1+\rho_u)\bar{F}+\rho_2\psi s_1+(\rho_F\rho_1+\rho_u\rho_F-\psi^2)s_2}{\rho_1(\rho_2+\rho_F)+\kappa_1} \end{bmatrix}, \begin{bmatrix} \frac{\rho_1\kappa_1}{\rho_1(\rho_2+\rho_F)+\kappa_1} & \frac{\rho_1\rho_2}{\rho_1(\rho_2+\rho_F)+\kappa_1} \\ \frac{\rho_1\rho_2}{\rho_1(\rho_2+\rho_F)+\kappa_1} & \frac{\rho_2(\rho_F\rho_1+\rho_u\rho_F-\psi^2)}{\rho_1(\rho_2+\rho_F)+\kappa_1} \end{bmatrix}\right) \tag{B2}$$

The informed investors' optimal allocation then follows

$$x^{I*} = \frac{\rho_1(\rho_2+\rho_F)\bar{u} - \rho_1\psi\bar{F} + \kappa_1 s_1 + \rho_1\psi s_2 - p[\rho_1(\rho_2+\rho_F)+\kappa_1]}{\gamma\rho_1\kappa_1} - \beta^I_{\text{flow}}\frac{\rho_2(\rho_F\rho_1+\rho_u\rho_F-\psi^2)}{\rho_1\kappa_1}. \tag{B3}$$

Uninformed investors do not directly observe private signals s_2. Instead they infer the signals from the price induced by the informed investors' demand. They conjecture the price in a form as a linear combination of the variables in the model.

$$p = a_1\bar{u} - a_2\bar{F} + bs_1 + cs_2 - dt + e\bar{t} + g. \tag{B4}$$

This is also the equilibrium price. The uninformed investors obtain noisy signals θ that is a random variable defined as

$$\theta = \frac{p - a_1\bar{u} - a_2\bar{F} - bs_1 + \bar{t}(d-e) - f}{c} = s_2 - (t-\bar{t})\frac{d}{c}. \tag{B5}$$

It is straightforward to verify that θ has the following normal distribution

$$\theta \sim N\left(F, \left(\frac{d}{c}\right)^2 \eta + \rho_2\right). \tag{B6}$$

The posterior distribution of the asset value and flow for uninformed investor U then follows

$$u^U, F^U | s^U \sim N\left(\begin{bmatrix} \frac{\rho_1(\rho_\theta+\rho_F)\bar{u}-\rho_1\psi\bar{F}+\kappa_2 s_1+\rho_1\psi s_2}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} \\ \frac{-\rho_\theta\psi\bar{u}+\rho_\theta(\rho_1+\rho_u)\bar{F}+\rho_\theta\psi s_1+(\rho_F\rho_1+\rho_u\rho_F-\psi^2)\theta}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} \end{bmatrix}, \begin{bmatrix} \frac{\rho_1\kappa_2}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} & \frac{\rho_1\rho_\theta}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} \\ \frac{\rho_1\rho_\theta}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} & \frac{\rho_\theta(\rho_F\rho_1+\rho_u\rho_F-\psi^2)}{\rho_1(\rho_\theta+\rho_F)+\kappa_2} \end{bmatrix}\right) \tag{B7}$$

The uninformed investors' optimal holdings of the risky asset can be obtained as

$$x^{U*} = \frac{\rho_1(\rho_\theta+\rho_F)\bar{u} - \rho_1\psi\bar{F} + \kappa_2 s_1 + \rho_1\psi s_2 - p[\rho_1(\rho_\theta+\rho_F)+\kappa_2]}{\gamma\rho_1\kappa_2} - \beta^U_{\text{flow}}\frac{\rho_\theta(\rho_F\rho_1+\rho_u\rho_F-\psi^2)}{\rho_1\kappa_2}. \tag{B8}$$

To solve for the equilibrium price, we impose the market clearing condition that the total supply must equal the total demand for the risky asset

$$\alpha x^{I*} + (1-\alpha)x^{U*} = t. \tag{B9}$$

The solution for the equilibrium price is presented in Equation 7 in the main text. Plugging into Equations B3 and B8, we obtain the solution for the optimal demand of informed and uninformed investors in terms of the model's variables and parameters, respectively. Subtracting the equations leads to the difference in holdings between two types of investors as in Equation 9.

B.2 Variable Construction

In this section, I provide additional details on the construction of mutual fund predictors.

Return gap. I follow Kacperczyk et al. (2008) to measure a fund's return gap in each quarter as the difference in the fund's net return and the return of the fund's holdings using the most recently disclosed holding positions, net of expense ratios. Particularly, the return gap RG of fund j in quarter q is defined as

$$RG_{j,q} = R_{j,q} - (HR_{j,q} - EXP_{j,q}),$$

where $R_{j,q}$ is fund j's net return in quarter q, $HR_{j,q}$ is fund j's holdings return, and $EXP_{j,q}$ is the expense ratio. For portfolio sorts, I follow Kacperczyk et al. (2008) and use average return gaps during the 12 months prior to the portfolio formation.

Reliance on public information. I follow Kacperczyk and Seru (2007) to measure a fund's RPI in each quarter as the R^2 from the regression of the fund's changes in holdings from previous quarter on the changes in analysts' recommendation for the holdings in the last five quarters. Particularly, I estimate the following regression for each fund and quarter

$$\begin{aligned}\%\Delta Hold_{j,i,q} = \beta_{0,j,q} + \beta_{1,j,q}\Delta REC_{i,q-1} + \beta_{2,j,q}\Delta REC_{i,q-2} + \beta_{3,j,q}\Delta REC_{i,q-3} \\ + \beta_{4,j,q}\Delta REC_{i,q-4} + \varepsilon_{j,q}, \quad i = 1, \ldots, N,\end{aligned}$$

where $\%\Delta Hold_{j,i,q}$ is the percentage change in stock split-adjusted holdings of stock i in fund j from quarter $q-1$ to quarter q, and $\Delta REC_{i,q-p}$ is the change in the recommendation of the consensus forecast of stock i from quarter $q-p-1$ to $q-p$.

Active share. Cremers and Petajisto (2009) construct the measure for mutual fund j, holding N stocks in the portfolio, at the end of each quarter q as

$$\mathrm{AS}_{j,q} = \frac{1}{2}\sum_{i=1}^{N} |\omega_{j,i,q} - \omega_{b,i,q}|,$$

where $\omega_{j,i,q}$ is the portfolio weight of stock i in fund j in quarter q, and $\omega_{b,i,t}$ is the weight of stock i in a benchmark portfolio. The authors use 21 benchmark indices and define the active share with respect to the benchmark that minimizes its value (*Active share (Min)*). I obtain the quarterly active share data from Martijn Cremers' website (`https://activeshare.nd.edu/data/`).[18]

***R^2*.** Similar to Amihud and Goyenko (2013), I obtain the R^2 for each fund and month from 24-month rolling regressions

$$R_{j,t} - RF_t = \alpha_p + \beta_{MKTRF} \times MKTRF_t + \beta_{SMB} \times SMB_t + \beta_{HML} \times HML_t + \beta_{UMD} \times UMD_t + \varepsilon_{j,t},$$

where $R_{j,t} - RF_t$ is a fund j's excess return over the risk free rate, $MKTRF_t$, SMB_t, HML_t, and UMD_t are the market, size, value, and momentum factor in the Carhart's (1997) model.

[18] I thank Martijn Cremers for making these data available.

Expected flow-induced trading. I follow Lou (2012) to estimate the expected flow-induced trading for each fund as the portfolio-weighted average expected flow-induced trading across its holdings

$$E[FIT_{j,q}] = \sum_{i=1}^{N_j} E[FIT_{i,q}]\omega_{j,i,q},$$

where $\omega_{j,i,q}$ is the portfolio weight of stock i in fund j in quarter q, and $E[FIT_{i,q}]$ is the expected flow-induced trading for stock i in quarter t. $E[FIT_{i,q}]$ is the adjusted-shares-weighted average of expected capital flows into all mutual funds.

Active fundamental performance. Similar to Jiang and Zheng (2018), I measure a fund's AFB as the covariance between deviations of its portfolio weights from the market portfolio and the underlying stock cumulative abnormal return three days around the stocks' earnings announcement. Particularly, at the end of each quarter, AFB is measured as

$$\mathrm{AFB}_{j,q} = \sum_{i=1}^{N_j} (\omega_{j,i,q} - \omega_{m,i,q})\mathrm{CAR}_{i,q}.$$

where $\omega_{j,i,q}$ is the portfolio weight of stock i in fund j in quarter q, and $\omega_{m,i,q}$ is the weight of stock i in the market portfolio. $\mathrm{CAR}_{i,q}$. is the stock i's 3-day cumulative abnormal return surrounding its quarterly earnings announcement two months following the quarter end.

Active fund overpricing. Similar to Avramov et al. (2020), I measure a fund's AFO as the covariance between deviations of its portfolio weights from the market portfolio and the underlying stock mispricing score. Particularly, at the end of each quarter, AFO is measured as

$$\mathrm{AFO}_{j,q} = \sum_{i=1}^{N_j} (\omega_{j,i,q} - \omega_{m,i,q})\mathrm{O}_{i,q}.$$

where $\omega_{j,i,q}$ is the portfolio weight of stock i in fund j in quarter q, and $\omega_{m,i,q}$ is the weight of stock i in the market portfolio. $\mathrm{O}_{i,q}$. is the stock i's mispricing score. The composite score is calculated based on the stocks' rank among 11 firm characteristics as in Stambaugh, Yu, and Yuan (2012). The higher a stock's score is, the more overpriced the stock is.