

# Flow Diversification\*

Sunil Wahal<sup>+</sup>  
WP Carey School of Business  
Arizona State University  
Tempe, AZ 85287  
[Sunil.Wahal@asu.edu](mailto:Sunil.Wahal@asu.edu)

Albert Y. Wang  
Harbert College of Business  
Auburn University  
Auburn, AL 36849  
[albertwang@auburn.edu](mailto:albertwang@auburn.edu)

May 2026

## Abstract

We measure flow diversification in mutual funds using the cross-sectional correlation of daily flows across investor clienteles. Greater flow diversification is associated with lower future flow volatility, smaller subsequent outflows, and lower cash holdings. We decompose flow correlation into predictable and unexpected components. Unexpectedly low flow correlation is associated with higher post-outflow fund returns, mitigating outflow-induced externalities on incumbent shareholders. Flow diversification also dampens the higher flow-performance sensitivity of illiquid funds, partially offsetting strategic complementarities that can generate fragility.

---

\* We are grateful to Lili Ge and Huan Li for research assistance. We thank the Center for Responsible Investing at ASU for financial support. Wahal is a consultant to Avantis Investors. Avantis did not provide data or funding for this research. We thank Vincent Bogousslavsky, Sergey Chernenko (discussant), Mitch Firestein, Massimo Massa, Elena Pikulina (discussant), Clemens Sialm, Yuri Tserluckevich, and participants at the American Finance Association (2023) and China International Conference in Finance (2023) meetings for helpful comments and suggestions.

<sup>+</sup> Corresponding author.

## 1. Introduction

A number of studies show that outflows force mutual funds to engage in costly trading, generating an externality for non-redeeming shareholders and contributing to fragility in underlying markets (Edelen (1999), Wermers (2000), Greene and Hodges (2002), Alexander, Cici, and Gibson (2007), Coval and Stafford (2007), Chen, Goldstein, and Jiang (2010), Falato, Goldstein, and Hortaçsu (2021) and others).<sup>1</sup> In this paper, we study whether the imperfectly correlated liquidity needs of investors can mitigate the negative externalities induced by outflows. Imperfect correlation in liquidity demands can arise from a variety of sources: investors may differ in their sensitivities to economic conditions or fundamentals that drive a fund’s NAV; behavioral differences may lead to heterogeneity in trading patterns (Christie and Huang (1995), Chang et al. (2000), and Bessembinder et al. (1996)); or geographic distance between investors and portfolio holdings might shape differences in liquidity demand (Ferreira, Massa, and Matos (2017)). Our central thesis is that for a mutual fund, diversification across investors with variation in such demands confers a “flow diversification” benefit that reduces the magnitude of liquidity externalities.

The channel is not simply the mechanical offsetting of investor flows; rather, it operates through a reduction in unexpected outflows that lowers immediacy demands. Rakowski (2010) argues that flow volatility, which proxies for unexpected outflows, causes funds to hold excess

---

<sup>1</sup> In addition to performance deterioration due to trading costs, redemptions can also result in the realization of capital gains, creating a tax externality (Dickson, Shoven, and Sialm (2000) and Sialm and Zhang (2020)). This also generates incremental monitoring costs since a portfolio manager needs to ensure that securities sold to accommodate redemptions are not bought back within 30 days for wash sale considerations.

cash and trade more frequently, incurring incremental explicit and implicit (price impact) costs, all of which reduce short-horizon returns. Flow volatility is a reduced-form outcome that reflects two underlying forces: the volatility of individual investor flows and the correlation structure across those flows. Even if all investors exhibit identical flow volatilities, lower correlation across investor flows reduces net flow volatility, allowing funds to have smaller cash buffers and incur lower trading costs. In this sense, understanding the correlation in investor liquidity demands provides direct insight into the mechanism through which unexpected outflows generate liquidity externalities.<sup>2</sup>

Empirically identifying the effects of flow diversification is challenging for two reasons. First, investor-level daily demand is not observable, even to a fund manager; mutual fund plumbing is such that fund managers only receive notifications of net flows across all investors in a share class, which means that investor-level flow correlations are not calculable. But because share classes pool investors with similarly correlated liquidity needs, correlation in flows across share classes, which are calculable, serves as a proxy for differences in the correlation structure across investor-specific liquidity needs. The literature is consistent with this view. For example, flows of broker-sold funds are less sensitive to poor performance (Del Guercio and Reuter (2014)) and deliver lower returns relative to directly-sold funds (Bergstresser, Chalmers, and Tufano

---

<sup>2</sup> One can also view the mechanism through a market microstructure lens because a mutual fund is effectively a market maker who is required to provide end-of-day immediacy. Net flows are equivalent to order imbalance, and flow volatility corresponds to order imbalance volatility. Bogousslavsky and Collin-Dufresne (2023) show that order imbalance volatility influences price impact, even when holding the level effect of order imbalance constant, because liquidity providers face risk in offloading inventory in the future. In other words, holding net flows constant, flow volatility affects fund-level transaction costs because it is related to unexpected future outflows.

(2009)), retail clients monitor their funds less extensively than institutional clients (Evans and Fahlenbrach (2012)), investors in retirement accounts exhibit inertia (Agnew, Balduzzi, and Sundén (2003)), and tax clienteles are important (Sialm and Starks (2012)). In the practitioner arena, it is well known that fee-only Registered Investment Advisors that serve as fiduciaries attempt to reduce fire-sale behavior by encouraging investors to stay invested in market downturns whereas retail investors who do not delegate their assets do not receive such treatment. Exploiting these insights, we measure daily fund-level flow correlation (*FlowCorr*) as the scaled value-weighted average covariance of flows across share classes for over 6,500 active mutual funds over the 2007-2024 period.

A second identification challenge is that persistent flow correlation is unlikely to be random across funds: to the extent that fund managers know their clientele's tendencies, they can *ex ante* adjust cash buffers to minimize trading costs, or even adjust their clientele mix to reduce flow correlation. Because holding additional cash generally involves a lower and more predictable opportunity cost than forced short-run trading, it is optimal for fund managers to adjust cash buffers *ex ante* rather than incur trading costs *ex post*.<sup>3</sup> In fact, we observe precisely such behavior: there is a positive relation between end-of-quarter cash buffers and average flow correlation in the quarter, and the addition of share classes is associated with both a sharp decline in flow correlation as well as a reduction in cash buffers. This evidence of endogenous adjustments underscores the

---

<sup>3</sup> This tradeoff is economically meaningful. Wermers (2000) estimates the return difference between equity and non-equity holdings (a measure of cash drag) to be about 0.70% per year while Edelen, Evans, and Kadlec (2013) report trading costs of about 1.44%. Differentials between trading costs and cash drag are even bigger in small cap and value funds, where trading costs in the underlying securities are higher (see also Yan (2008) and Simutin (2014)).

economic importance of flow diversification.<sup>4</sup> But it also makes it difficult to interpret the return consequences of flow correlation from persistent cross-sectional differences alone. To sharpen inference, we decompose *FlowCorr* into an average component that represents its conditional expectation (*ExpFlowCorr*) and a deviation from the average that represents unexpected flow correlation (*UnexpFlowCorr*). Under this decomposition, predictable differences in flow correlation should be absorbed, at least partially, through managerial responses such as cash holdings; on the other hand, unexpected flow correlation captures short-run deviations in the synchronization of investor liquidity demands that are harder to manage *ex ante* and therefore more likely to generate variation in immediate trading needs and post-outflow returns.

Consistent with this mechanism, in daily univariate sorts on *UnexpFlowCorr* for funds experiencing outflows, the average spread in subsequent 3-day benchmark-adjusted returns between low and high *UnexpFlowCorr* quintiles is 1.07 basis points (t-statistic = 3.61). In contrast, and as expected, analogous sorts based on *ExpFlowCorr* show no such variation in returns. In multivariate regressions that control for a variety of fund characteristics as well as fund and day fixed effects, we continue to see a reliable negative relation between fund returns and *UnexpFlowCorr*. These return differences are neither transitory nor reversed: over a 20-day return horizon, the average spread in benchmark-adjusted returns is 4.13 basis points (t-statistic = 3.76), and the coefficient on *UnexpFlowCorr* in multivariate regressions remains negative and

---

<sup>4</sup> The economic underpinnings of flow correlation are also evident in other ways. For instance, flow correlation is related to prior market returns and investor sentiment. But most interestingly, flow correlation spikes in December and towards the last three trading days of the month, lining up with correlated trading arising from the payment cycle (Etula et al. (2020)), retirement contributions, or tax motivated trading (Sialm and Zhang (2020)).

statistically significant. We interpret these results as evidence on the short-run return consequences of unexpectedly synchronized liquidity needs, rather than as estimates from a fully exogenous shock to flow correlation.

These differences in returns are in the same order of magnitude as trading costs and therefore economically meaningful to investors, but there are other ways to think about economic significance. In daily data, the average benchmark-adjusted return difference between small and large outflow quintiles over the 2007-2024 period is 1.22 basis points. Thus, the return spread associated with unexpected flow diversification in sort-based tests (1.07 basis points) is an order of magnitude comparable to that of the liquidity externality itself. A second comparison is to Rakowski (2010) who reports that funds in the lowest quartile of daily flow volatility outperform those in the highest quartile by 0.89 basis points per day. Again, the mitigation due to flow diversification is comparable to this benchmark. It is also instructive to see how practitioners consider liquidity management in similar contexts. Through a “real-life example”, Cheng et al. (2024) describe the liquidity management techniques used by a particular fund to mitigate trading costs when facing a small, unexpected outflow of comparable magnitude – emblematic of the lengths to which portfolio managers go to avoid trading costs.

Given these results, we consider the effects of flow diversification on shareholder runs. Chen, Goldstein, and Jiang (2010) argue that the expectation that some investors will exit a fund and impose a cost on others generates a multiplier effect. Their approach to studying such payoff complementarities is to determine if flow-performance sensitivity is higher in funds with illiquid assets, leveraging the idea that redemptions impose higher costs on non-redeeming shareholders

in illiquid funds (see also Goldstein, Jiang, and David (2017), Falato, Hortaçsu, Li, and Shin (2021), and Jiang et al. (2022)). We reproduce their monthly regressions with daily data and then ask whether flow-performance sensitivities vary with flow diversification, even in the presence of higher flow-performance sensitivities for illiquid funds. This is indeed the case: independent of asset liquidity, lower levels of flow diversification are associated with higher flow-performance sensitivities. We then assess whether flow diversification mitigates higher flow-performance sensitivities in illiquid funds subject to outflows, essentially asking whether flow diversification can partially offset strategic complementarities. Consistent with this idea, we find that flow-performance sensitivity in illiquid funds is higher in funds when flow diversification is low and attenuated when flow diversification is high.

Our tests show important benefits for mutual funds from diversifying their investor bases. In particular, they imply that liquidity in fund management not only involve the often-studied liquidity of portfolio holdings (e.g., Pástor, Stambaugh, and Taylor (2020)), but also the liquidity demands of the investor base – an aspect of delegated asset management that has received limited attention. The diversification-based liquidity mechanism we describe is also visible in other financial markets. In banking, Allen and Gale (2000) argue that risk-sharing takes place when banks hold interregional claims on other banks, and diversification enables insurance against local liquidity shocks. Campello and Gao (2017) study customer concentration (the flip side of client diversification) in loan contracts and conclude that concentration is associated with higher interest rate spreads. In mortgage-backed securities, prepayment risk from one group of mortgages is effectively transferred to all investors in the pool, but variation in holding periods and liquidity

requirements affects pricing.<sup>5</sup> A similar insurance mechanism against temporary liquidity shocks is also evident in funds-of-funds (Bhattacharya, Lee, and Pool (2013)). In money market funds, Kacperczyk and Schnabl (2013) show that funds with a more diversified client base take on less risk, arguing that this is because reputational costs would generate spillover effects (outflows) to other mutual funds managed by the same organization (see also, Schmidt, Timmermann, and Wermers (2016)). At the family level, Massa (2003) argues that investor heterogeneity motivates fund families to expand product offerings, albeit with negative effects on performance.

Some market participants have developed innovative solutions to address the liquidity externalities of mutual funds. Jin, Kacperczyk, Kahraman, and Suntheim (2022) report that swing pricing rules (in which redeeming shareholders bear the cost of a fund's price impact) reduce liquidity externalities and spillover effects in U.K. domiciled bond funds. In our sample, however, none of the U.S. equity funds we study use swing pricing rules between 2019 and 2024. Of course, liquidity and tax externalities are substantially lower in ETFs because investors trade in the secondary market, and in-kind redemption, facilitated by Authorized Participants and market makers, reduce trading activity and allow low-cost-basis securities to be transferred without tax consequences. Given these structural advantages, it is perhaps unsurprising that there has been a steady shift from mutual funds towards ETFs in the last decade (see ICI Factbook, 2024), the conversions of mutual funds to ETFs under Rule 6c-11 (Du, Starks, and Xiaolan (2023)), and dual fund-ETF share class offerings by prominent asset managers.

---

<sup>5</sup> Original insight into this parallel offered by the late Rick Green, as cited in Johnson (2004).

## 2. Institutional Details

Investors who submit a trade prior to market close on day  $t$  receive the NAV struck at market close on the same day. A portfolio manager's ability to observe aggregated client purchases or sales in a share class depends on the trading platform through which investors submit orders. In most standard trading platforms, net flows are reported to portfolio managers by 10:00 PM PST on day  $t$ . In the majority of retirement platforms, however, portfolio managers may learn about flows as late as 10:45 AM PST on day  $t+1$ . Much of the transaction processing between fund companies and distributors takes place via Fund/Serv, an operational subsidiary of the DTCC (for details, see <https://www.dtcc.com/wealth-management-services/mutual-fund-services/fund-serv>).

Investor redemptions must ultimately be met either with cash balances or through the sale of current holdings. Cash balances prevent the fund from becoming levered and allow stock sales to be deferred for a day or two. Morris, Shim, and Shin (2017) show that cash hoarding arises in response to redemption risk in bond funds. Chernenko and Sunderam (2020) further show that funds with stronger incentives to internalize their price impact adjust their cash buffers. Cash buffers reduce the need for immediate trading, but they are not a complete solution because daily redemptions are difficult to predict and may exceed available liquidity when investor liquidity needs become unexpectedly synchronized.<sup>6</sup>

Regardless of timing and mitigation strategies, investors remaining in the fund ultimately bear the implicit and explicit costs associated with redemptions, including market impact,

---

<sup>6</sup> The median autocorrelation in daily net flows is only 0.03 implying very limited predictability. We also compute the ratio of average daily outflows to beginning-of-quarter cash and find that a significant fraction of funds may not have adequate cash buffers to meet daily redemption requirements.

commissions, stamp duties, and ticket charges. They may also include costs associated with temporarily deviating from target portfolio weights while managers sequence trades to limit price impact. These costs reduce the fund's NAV and therefore appear in our tests as lower subsequent fund returns. Such externalities are well recognized by market participants, as evidenced by the fact that many data providers track and report flows. For example, Morningstar's Fund Flow Report provides considerable detail on flows over a variety of horizons and is widely followed by investors. Anticipating these issues, investment committees of financial advisory firms and other fiduciaries routinely monitor outflows from the funds they own (see, for example, Bailey and Richards (2017) and Illan (2018)). Moreover, industry observers frequently comment on the “slow death of mutual funds” at the hands of ETFs that do not suffer from such liquidity or tax externalities (see, for example, Rekenthaler (2021), Balchunas and Weber (2023), Britton (2019); Helmke (2023) provides an academic counterview).

The cash flow timing mechanism described above implies that trading triggered by redemptions need not take place on day  $t+1$ , and can be stretched out over several days. Therefore, in measuring post-flow performance, we primarily consider returns from day  $t+1$  to  $t+3$ .

### **3. Measurement Issues and Sample Construction**

#### **3.1 Measuring Flow Diversification**

In a perfect empirical setting, we would observe the liquidity needs of each investor and aggregate the correlation structure of those liquidity demands up to a fund-level flow diversification measure. Absent such idealized data, we rely on the well-documented idea that mutual funds use share classes to channel investors with similar liquidity requirements and

tendencies (see, for example, Chordia (1996)). Most funds have specific share classes for financial advisors labeled “advisor” or “institutional,” and retirement share classes are often prefixed “R” (see Nanda, Wang, and Zheng (2009) for early history and nomenclature). A number of studies report behavioral and investment differences between retail investors, retirement accounts, institutional channels, and broker-sold funds. Differences across clienteles have a variety of origins and take numerous forms, including inattention, performance-flow sensitivities, tax-loss harvesting, and rebalancing needs.<sup>7</sup>

Although individual investors can invest through multiple share classes, share classes nonetheless determine the channel by which ownership claims are transmitted and through which liquidity is demanded. We exploit the systematic variation in liquidity demands across share classes in measuring flow diversification. We first define flows for share class  $k$  in fund  $i$  at time  $t$  as:

$$Flows_{i,k,t} = TNA_{i,k,t} - TNA_{i,k,t-1} \times (1 + r_{i,k,t}) \quad (1)$$

where  $TNA$  represents total net assets and  $r$  is the net return after fees.<sup>8</sup> Aggregating flows across all  $k$  share classes yields net flows for fund  $i$  at time  $t$ :

$$Flows_{i,t} = \sum_{k=1}^K Flows_{i,k,t} \quad (2)$$

---

<sup>7</sup> The literature is vast and in addition to the papers cited in the introduction, see Ben-Rephael et al. (2018), Barber et al. (2016), Christoffersen and Musto (2002), Chalmers and Reuter (2010), Del Guercio et al. (2018), Del Guercio and Tkac (2002), James and Karceski (2006), Schmidt et al. (2016), and Williams (2022).

<sup>8</sup> We use dollar flows, rather than proportional flows, for two reasons. First, from the perspective a portfolio manager, it is the dollar value of the trading rather than a portfolio weight that matters in generating trading externalities. Second, if one uses proportional flows and value-weights, then the ratio of covariances to variances is degenerate because of symmetry around the central value (in this case, by construction because the sum of share class assets is equal fund assets). We consider an equal-weighted alternative using proportional flows in Section 5.5.

Our measure of flow diversification,  $FlowCorr$ , is based on the covariance of flows across share classes. Intuitively,  $FlowCorr$  captures the extent to which liquidity demands across investor groups move together, conditional on a given level of fund net flows. We compute  $FlowCorr$  as the average pairwise cross-product of value-weighted deviations of class-level flows from fund-level net flows, scaled by the average squared value-weighted deviation across all share classes. The normalization ensures that  $FlowCorr$  is comparable across funds with different numbers of share classes.

$$\begin{aligned}
FlowCorr_{i,t} &= \frac{\sum_{k \neq l} cov(flows_{i,k,t}, flows_{i,l,t})}{var(flows_{i,k,t})} \\
&= \frac{\frac{2}{K(K-1)} \sum_{k=1}^{K-1} \sum_{j=k+1}^K [w_{k,t-1}(Flows_{i,k,t} - Flows_{i,t}) \times w_{j,t-1}(Flows_{i,j,t} - Flows_{i,t})]}{\frac{1}{K} \sum_{k=1}^K w_{k,t-1}^2 (Flows_{i,k,t} - Flows_{i,t})^2} \quad (3)
\end{aligned}$$

where  $w_{k,t-1} = \frac{TNA_{i,k,t-1}}{TNA_{i,t-1}}$  is the weight of share class  $k$  in fund  $i$  at time  $t-1$ .  $FlowCorr$  is bounded between -1 and +1, and positive (negative) values indicate positively (negatively) correlated flows across share classes.

$FlowCorr$  reflects both a clientele effect, the correlation in liquidity needs across types of investors segmented by share classes, and a client effect, which captures idiosyncratic, uncorrelated liquidity demands across individual investors. As described earlier, clientele effects are likely endogenous because fund managers understand the typical liquidity needs of investors in various share classes. For example, retirement share classes experience periodic contribution and distribution cycles, end-of-year rebalancing by 401(k) platforms is well known, and institutional clients often provide advance notice of large upcoming redemptions. These

predictable features generate a slow-moving component in *FlowCorr* that shifts primarily when funds alter their clientele structure by adding or removing share classes.<sup>9</sup> In contrast, the client effect arises from unpredictable variation in investors' liquidity needs that lies outside managerial control.

To separate these forces, we extract the average (permanent) component of *FlowCorr* using a backward-looking 60-day moving average (MA) process, interpreting it as the conditional expectation of *FlowCorr* from clienteles. Deviations from the average, or the temporary component, represent the idiosyncratic unpredictable correlation in liquidity needs across investors. We denote these components *ExpFlowCorr* and *UnexpFlowCorr*, respectively. Importantly, this decomposition is not intended as a statistical forecasting model. Instead, it provides a parsimonious way to separate predictable liquidity patterns that managers can internalize from deviations from the average that cannot be managed *ex ante*. Unexpected flow correlation may still be correlated with unobserved market, investor-demand, or fund-level shocks. Accordingly, we interpret *UnexpFlowCorr* as a measure of synchronized liquidity needs that moderates the effect of unexpected outflows on future fund returns, rather than a fully exogenous shock.

### **3.2 Data and Sample**

Our primary data source is the Morningstar Mutual Fund Database, which provides daily data on flows and returns for 7,489 funds at the share class level. Morningstar receives daily data

---

<sup>9</sup> The slow-moving component also shifts when a share class is added to or removed from so-called model portfolios used by RIAs, Wirehouses, independent broker-dealers, private banks, wealth managers, and insurance companies. Such changes result in substantial flows, which changes the correlation structure across clienteles.

on NAVs, assets, and returns through direct feeds from fund managers, fund administrators, custodians, trusts, advisors, and fund companies. To construct our sample, we start with all U.S.-domiciled open-end equity mutual funds in the database. We exclude index funds because they attract a particular type of clientele and because Vanguard, the largest index fund provider, does not provide daily flow data to Morningstar. The sample period spans January 2007 through December 2024.

We supplement daily flow data with additional monthly or quarterly data on fund characteristics (i.e. expenses, turnover, load, and age) from the CRSP survivorship-bias-free Mutual Fund database. To merge Morningstar with CRSP, we first match on CUSIP. When CUSIP is unavailable, we use TICKER, and if neither identifier is available (representing 13% of the sample), we manually match observations using fund and share-class names. After matching, our sample consists of 6,548 funds and 22,076 share classes. At the end of 2024, our sample represents approximately 60% of total assets in active U.S. equity mutual funds.

For cash holdings, we use Morningstar rather than the N-SAR filings because Morningstar provides data on both cash and cash equivalents. We extract equity style classifications (using the ubiquitous 3x3 size and value/growth grid) from Morningstar along with their benchmark returns. We aggregate reported variables across share classes at the fund level by value weighting them using lagged total net assets (TNA) of each share class. The average fund in our sample has 4.17 share classes, with more than 91% of funds reporting daily observations for at least two share classes.

To determine clienteles, we map share classes into three economically distinct categories

used by Morningstar and other practitioners: (a) retail, (b) retirement, and (c) advisor/institutional, hereafter referred to simply as institutional. We merge institutional and advisor share classes into one group, because, in practice, fund families use these labels interchangeably to accommodate investment from RIAs and Wirehouse platforms, regardless of whether the underlying investors are 13F filers. To do this, we start with the existing Morningstar mappings, but for share classes not categorized by Morningstar, we rely on CRSP classifications.<sup>10</sup> At the end of 2024, the percentage of aggregate assets corresponding to retail, retirement, and institutional groups are 21%, 18%, and 61%, respectively.

### 3.3 Descriptive Statistics

Definitions for key variables of interest appear in Appendix Table 1. With exceptions noted below, most tests, especially those involving return differences, are based on the sample of fund-days with net outflows, measured after aggregating flows across all share classes. This is because (a) trading costs associated with inflows are borne by both entering and incumbent shareholders, (b) the literature described earlier largely focuses on externalities generated by outflows, and (c) it is relatively inexpensive to equitize inflows in the futures market and then minimize trading costs by slowly purchasing target securities (see, for example, Cheng et al. (2024)). Descriptive statistics for the outflow-only sample are reported in Appendix Table 2.<sup>11</sup>

The average *FlowCorr* in the sample is only 0.20, but its standard deviation is more than

---

<sup>10</sup> CRSP classifies share classes into retail or institutional groups but does not flag retirement classes. We classify non-retail and non-institutional classes as retirement if the fund name has a suffix in the following list: J, K, K6, R, R1, R-1, R2, R-2, R2E, R-2E, R3, R-3, R4, R-4, R5, R-5, R5E, R-5E, R6, R-6, Retire, and Retirement.

<sup>11</sup> To alleviate concerns associated with extreme values, we replicate our main tests involving future returns after winsorizing them at the 1<sup>st</sup> and 99<sup>th</sup> percentiles. There are no substantive changes in inferences.

twice as large. Given the moving average filter used in the decomposition, it is not surprising that the standard deviation of *ExpFlowCorr* is close to its mean (0.16 compared to 0.19), but that the standard deviation of *UnexpFlowCorr* is much higher than its mean (0.39 compared to 0.01). Although *UnexpFlowCorr* is transitory by construction, it is precisely these short-lived deviations that generate immediacy demands, making them economically relevant. The cross-sectional average autocorrelation of *FlowCorr* is 0.10 based on the full sample of all fund days.<sup>12</sup>

Also of interest is the distribution of cash buffers, particularly in the context of expected flow correlation. The average cash buffer (*Cash*), defined as the sum of cash and cash equivalents scaled by total assets, is 3.97%.<sup>13</sup> But the standard deviation is more than twice the mean (9.42%) and the median is 1.83%. Consistent with the literature, this large variation suggests that fund managers actively manage cash levels.

## **4. Flow Diversification**

### **4.1 Determinants of Flow Correlation**

We start by examining the determinants of flow correlation. Since our purpose is to broadly understand flow correlation, we use the full panel of fund-day observations, including days with net inflows and outflows. Panel A of Figure 1 shows monthly average *FlowCorr* across all funds over the sample period. The long-run average is 0.20 with elevated values during the financial

---

<sup>12</sup> The autocorrelation of *FlowCorr* cannot be meaningfully calculated for an intermittent time series of fund-days that only have net outflows.

<sup>13</sup> Morningstar cash holdings include cash and cash equivalents and can exceed 100% of total assets for some funds, potentially reflecting securities lending, short positions, or derivatives. To ensure that our results are not driven by these extreme observations, we winsorize *Cash* at the 0.5<sup>th</sup> and 99.5<sup>th</sup> percentiles. We use the 99.5<sup>th</sup> percentile because it lies just below 100%, which provides a conservative upper-tail adjustment while preserving economically meaningful variation in cash buffers.

crisis (2008-2009) and other common liquidity shocks. Interestingly, each of the local peaks in the graph correspond to December. Panel B plots monthly averages over the 17-year period and confirms the notable increase in flow correlation in December, consistent with end-of-year portfolio rebalancing and tax motivated trading by investors. Etula et al. (2020) document that the monthly payment cycle creates systematic patterns in liquidity drawdowns towards the end of the month. Panel C shows daily flow correlations from 10 days before to 10 days after the end of the month, where day 0 is the last trading day of the month. There is a striking increase in flow correlations two days before the last trading day of the month, consistent with a systematic liquidity demand related to the payment cycle termed “dash for cash” by Etula et al. (2020). Collectively, Panels B and C imply that flow correlation responds systematically to recurring calendar-based liquidity demands.

A more formal regression analysis appears in Table 1. The dependent variable is *FlowCorr* on day  $t$  and the independent variables capture the degree to which flow correlation is related to fund-specific attributes, responds to time-varying market conditions, and reflects changes in investor heterogeneity. Column (1) includes fund fixed effects that control for unobservable, time-invariant fund characteristics. Column (2) additionally includes style-quarter fixed effects to absorb any liquidity shocks that may be clustered within each style and quarter. Robust standard errors, clustered at the fund-quarter level, appear in parentheses. We do not include cash holdings as a control in these regressions because cash holdings are a managerial response to predictable liquidity risk (see discussion in Section 5.1).

Using the specification in Column (2), larger and older funds have significantly lower flow

correlation, consistent with the view that more established funds attract a broader and more heterogeneous investor base (Pástor, Stambaugh, and Taylor (2020)). A one-standard-deviation increase in fund size and age reduces flow correlation by 4.53% and 14.89%, respectively. In contrast, funds with higher expense ratios or loads display higher flow correlation, reinforcing the notion that fee structures segment investors into more homogeneous clienteles. A one-standard-deviation increase in expense ratios (loads) is associated with a 9.83% (2.80%) increase in flow correlation. Higher-turnover funds, which trade more frequently and tend to have shorter holding periods, may attract performance-sensitive investors, increasing synchronous flows.

Market-level factors also play an important role. Both contemporaneous and lagged market returns are negatively related to *FlowCorr*, consistent with the notion that declining markets generate redemptions across investors. The negative coefficient on sentiment aligns with the view that mutual fund flows reflect investor sentiment (Frazzini and Lamont (2008)) and that shifts in sentiment can generate synchronous trading behavior, producing more correlated liquidity demands.<sup>14</sup> Turning to time-series patterns, the coefficients on the December and month-end dummies are both positive and highly significant, matching the spikes in flow correlation observed in Figure 1. Quantitatively, flow correlation in December is almost 20% higher than in other months, and flow correlation in the final trading days of the month is 4% higher than the rest of the month. These findings reinforce the interpretation that broad market conditions and recurring calendar effects contribute meaningfully to the synchronization of investor liquidity demands.

---

<sup>14</sup> Sentiment is measured as the monthly change in the University of Michigan Index of Consumer Sentiment. Because the index is available monthly, we assign the same monthly change to all daily observations within that month.

In Columns (3) and (4), we augment the prior regressions with an indicator variable equal to one for days when a new share class is added. The coefficient on the indicator variable is  $-0.02$  with a standard error of  $0.005$ . Given that the average flow correlation in the sample is  $0.20$ , this implies that share class additions reduce flow correlation by  $10\%$  on average. Columns (5) and (6) contain separate indicator variables for retail, institutional, or retirement class additions. The addition of retail or institutional share classes substantially reduces flow correlations, implying that new distribution channels introduce investors with liquidity needs that are less synchronized with those of existing clientele. Additions of retirement share classes, which typically attract investors with slow-moving, contribution-driven flows, do not significantly affect flow correlation. The lack of significance may reflect limited statistical power, as only 258 retirement-class additions occur in our sample.

#### **4.2 Background Externalities and Prerequisites**

In this section, we quantify the flow-related externalities central to our analysis and establish the prerequisites for our main tests.

A number of studies show that large outflows force mutual funds to engage in costly trading, generating an externality for non-redeeming shareholders and contributing to fragility in underlying markets (Edelen (1999), Wermers (2000), Greene and Hodges (2002), Alexander, Cici, and Gibson (2007), Coval and Stafford (2007), Falato, Goldstein, and Hortaçsu (2021) and others). Most of these studies rely on quarterly holdings or monthly fund flows and returns. However, as discussed in Section 2, redemption-driven trading occurs over the next few days (intra-month or intra-quarter). We therefore replicate these findings in a daily setting to provide a benchmark for

interpreting the economic significance of flow diversification. We form separate quintiles for outflows and inflows on day  $t$ , and examine returns on  $t+1$ ,  $t+2$ ,  $t+3$ , and the cumulative return from  $t+1$  to  $t+3$ . Panel A of Appendix Table 3 shows that the difference in  $t+1$  to  $t+3$  cumulative returns between small versus large outflow quintiles is 1.22 basis point, with a standard error that is roughly one-third of that. In other words, funds experiencing larger outflows have lower subsequent returns. Panels B and C show that these return differences exist in both liquid and illiquid funds (as defined by Chen, Goldstein, and Jiang (2010)). As expected, return differences only exist for fund-days with outflows, and the remainder of our analyses focus on the outflows-only subsample.

Since our measure of flow correlation is constructed from class-, not investor-level flows, it is important to verify that higher flow correlation is associated with greater flow volatility. To do so, we estimate regressions of subsequent flow volatility on *FlowCorr*. We measure flow volatility using the square root of the sum of squared daily flows over a 3-day window (using a  $t+1$  to  $t+3$  timing convention) and control for market volatility (VIX), fund size and past flows.<sup>15</sup> Robust standard errors are clustered at the fund-quarter level.

Panel A of Table 2 reports these flow-volatility regressions for fund-days with net outflows. Column (1) presents a univariate regression with fund fixed effects (to absorb time-invariant fund heterogeneity) and Column (4) additionally includes day fixed effects (to absorb common volatility spikes). In both specifications, future flow volatility is reliably positively

---

<sup>15</sup> A short window isolates flow volatility in close proximity to flows and flow diversification on day  $t$ . The cost is limited degrees of freedom and, therefore, imprecision in the measurement of flow volatility. However, somewhat larger windows (e.g., five days) do not affect our conclusions.

related to *FlowCorr*. Since flow volatility is persistent, the regressions in Columns (2) and (5) add lagged flow volatility as a predictor. As expected, the coefficient on lagged volatility is positive but the coefficient on *FlowCorr* remains significantly positive. Lastly, Columns (3) and (6) add measures of market volatility (VIX), fund size and lagged flows. Each successive model is more stringent and throughout the specifications, the coefficient on *FlowCorr* remains positive.<sup>16</sup>

Flow volatility is a reduced-form proxy for unexpected flows, but we can also assess whether flow correlation is related to unexpected flows directly. One could measure unexpected flows using a time series model but the median autocorrelation in daily flows is only 0.03, implying that at the daily level, flows are largely unpredictable. To avoid overfitting a weakly persistent process, we use outflows in  $t+1$  as the dependent variable. Panel B of Table 2 reports the results for fund-days with net outflows. In the univariate specifications (with either fund, or fund and day fixed effects), the coefficients on *FlowCorr* are -0.229 and -0.163, showing that higher flow correlation, or lower flow diversification, is associated with larger next-day outflows. Adding lagged flow volatility in Columns (2) and (5) attenuates the coefficients modestly, confirming the results in Rakowski (2010). In the full specifications that include controls for past flows (which further accounts for the modest serial correlation in daily flows described above), fund size, expense ratios, turnover, cash levels and loads, the coefficients on *FlowCorr* remain statistically significant at -0.249 and -0.177. These estimates imply that a one-standard-deviation increase in *FlowCorr* is associated with larger next-day redemptions: the implied effects are 3.78% and 4.10%

---

<sup>16</sup> We also estimate models in which we first regress *FlowCorr* measured on day  $t$  on lagged flow volatility measured on days  $t-2$  to  $t$  and use the residuals (representing orthogonalized flow correlation, *OrthFlowCorr*) in second-stage regressions. The results are quantitatively similar and *OrthFlowCorr* continues to predict future flow volatility.

of the mean next-day outflow in the outflow-only sample, respectively.

Overall, the results in Table 2 establish that flow correlation has meaningful implications for future fund flows. Higher *FlowCorr* is related to both elevated future flow volatility and larger unexpected outflows, the mechanisms through which liquidity externalities arise. These findings validate the central premise of our identification strategy: flow correlation, even when measured through share-class flows rather than investor-level flows, captures economically significant differences in the synchronization of liquidity demands that are not captured by flow volatility alone.

## **5. The Effects of Flow Diversification**

### **5.1 Flow Diversification and Cash Buffers**

If fund managers recognize that greater flow diversification lowers outflows then, *ceteris paribus*, they should optimally hold less precautionary cash. The economic tradeoff underlying this prediction is between holding costly precautionary cash *ex ante* and incurring trading costs *ex post* when redemptions arrive. Holding cash creates cash drag, but this cost is relatively predictable and can be managed through the size of the buffer. In contrast, redemption-induced trading may require funds to sell securities quickly, generating explicit trading costs, bid-ask spreads, and price impact that reduce fund NAVs. Thus, predictable flow correlation can be incorporated into cash-buffer decisions, whereas higher unexpected flow correlation is more likely to generate short-run trading pressure and lower post-outflow returns.

We test whether managers incorporate predictable flow correlation into cash-buffer decisions in two complementary ways. The first approach relates end-of-quarter cash buffers to

average flow correlation over the quarter (*ExpFlowCorr*). Column (1) of Panel A in Table 3 reports a univariate regression with fund-year fixed effects, and Column (3) augments this specification with fund and style-year fixed effects. The coefficients on *ExpFlowCorr* are 0.744 and 0.898, both of which are statistically significant. A one-standard-deviation increase in *ExpFlowCorr* predicts a 2.78% and 3.35% increase in end-of-quarter cash buffers, respectively. In Columns (2) and (4), we add controls for fund size, fund flows over the quarter, fund returns, daily flow volatility, expense ratios, loads, and turnover. The inclusion of these controls attenuates the coefficients on *ExpFlowCorr* but they remain statistically significant.

The second approach exploits changes in clientele composition by examining share-class additions, which should expand the investor base, reduce average flow correlation, and reduce cash buffers. Over our sample period, there are 6,552 share class additions from a different clientele. Panel A of Figure 2 plots the cross-sectional average *FlowCorr* from 60 days before to 60 days after the addition of a share class. The graph shows a 6.39% decrease in flow correlation sharply centered around day 0. Panels B through D repeat the analysis separately for retail, institutional and retirement share classes. For retail and institutional additions, we continue to observe large declines in flow correlation around day 0. There is no significant decline around the addition of retirement share classes, likely reflecting the limited number of such events in our sample (only 258 events, compared to 1,834 retail class additions and 4,460 institutional share class additions).

Panel B of Table 3 examines whether these declines in flow correlation translate into reductions in cash buffers. Cash holdings are reported at month- or quarter-end, but share-class additions are not subject to such discreteness. To manage these differing frequencies, we assume

that daily cash equals the most recently reported value, updating it when new cash data becomes available. The dependent variable in our regressions is the change in cash buffers from day  $t-20$  to day  $t+20$ .<sup>17</sup> Because this daily sample includes both inflows and outflows, we first ask whether changes in cash buffers are symmetric for inflows and outflows. Columns (5) and (6) show regressions of change in cash on *ExpFlowCorr*, an indicator variable equal to one for outflows, and the interaction between the two variables. Column (5) uses fund fixed effects with standard errors based on Driscoll and Kraay (1998), while Column (6) uses fund and day fixed effects with standard errors clustered by fund and day.<sup>18</sup> The interaction term is positive, implying that the relation between predictable flow correlation and changes in cash buffers is stronger for outflows than inflows.

Next, we examine the effect of changes in the number of share classes on changes in cash. Columns (7) and (8) regress this change in cash on an indicator variable equal to one if the number of share classes increases on day  $t$  with the same set of control variables as in Panel A. In both specifications (that use fund or fund and day fixed effects), share-class additions are associated with a reduction of roughly 3% in cash holdings.

The impact of adding a share class is likely nonlinear and depends on the pre-increase number of share classes. For example, an increase from two to three share classes is quite different

---

<sup>17</sup> Using this convention maximizes the precision in measurement because some cash holdings in the Morningstar database are reported on a monthly and others on a quarterly basis.

<sup>18</sup> Because cash holdings are reported at month- or quarter-end and assigned to daily observations until updated, adjacent fund-days may share the same underlying cash observation. In addition, the  $t-20$  to  $t+20$  cash-change window creates dependence across nearby observations, and cash holdings may be correlated across funds on the same day. We therefore use standard-error adjustments that allow for serial and cross-sectional dependence. Section 5.2.2 discusses related inference issues in the return regressions.

from seven to eight. In Columns (9) and (10), we replace the share-class increase indicator variable with separate indicators for large and small share-class indicators, in which “large” is defined as cases where the post-increase number of share classes is four or less. The coefficients on large additions remain negative and statistically significant, implying a 5.3% decline in cash holdings. In contrast, the coefficients on small additions are indistinguishable from zero, suggesting that only economically meaningful expansions in clientele drive reductions in precautionary cash buffers.<sup>19</sup>

## **5.2 The Effects of Flow Diversification on Fund Performance**

### **5.2.1 Univariate Sorts and Multivariate Evidence**

We start our analysis of flow diversification and fund performance with sorts because they provide model-free and easy-to-interpret summary estimates. In Table 4, we sort all funds with outflows on day  $t$  into quintiles based on  $FlowCorr$ ,  $UnexpFlowCorr$ , and  $ExpFlowCorr$ , and then compute benchmark-adjusted returns over several short horizons. We focus on fund-days with outflows because redemption-driven trading costs are borne by remaining shareholders and should therefore be most visible following outflows. Risk effects are unlikely to be important at short horizons, so factor model alphas are unnecessary. However, since liquidity differences associated with investment styles can be meaningful (e.g., small- versus large-cap funds), we use benchmark-adjusted returns based on each fund’s equity style classification. Because flow diversification is

---

<sup>19</sup> To assess whether the addition of retail, institutional, or retirement classes has differential effects on cash holdings, we also estimate regressions with separate indicator variables. In these unreported regressions the coefficients on large additions of retail and institutional share classes are associated with declines in cash holdings. For example, an addition of institutional share classes when pre-increase number of share classes is four and less is associated with a 5% decrease in cash holdings from the month before to the month after the addition.

measured on day  $t$  and trading externalities arise through short-horizon trading costs, we focus on benchmark-adjusted returns over days  $t+1$ ,  $t+2$ ,  $t+3$ , and the cumulative return from  $t+1$  to  $t+3$ . These horizons accommodate the cutoff times and reporting delays described in Section 2. Standard errors are computed from the time series of quintile returns with Newey-West adjustments to account for overlapping return windows.

Panel A contains results for quintiles based on *FlowCorr*. Funds in the lowest quintile earn higher subsequent returns than funds in the highest quintile, but the standard errors are large. For example, the 3-day spread in benchmark-adjusted returns between low and high *FlowCorr* quintiles is 0.55 basis points with a standard error of 0.35. The muted relation is unsurprising because the permanent component of *FlowCorr* is anticipated by fund managers and baked into cash buffers, reducing the need for immediate trading. Panel B shows quintiles based on *UnexpFlowCorr* and reveals a starkly different picture. Here we observe a strong, monotonic relation between unexpected flow correlation and subsequent returns; the 3-day benchmark-adjusted return spread between the lowest and highest *UnexpFlowCorr* quintiles is 1.07 basis points with a standard error of 0.30. By contrast, Panel C shows no meaningful difference in benchmark-adjusted returns across quintiles formed on *ExpFlowCorr*, consistent with the interpretation that fund managers understand predictable flow patterns from their share-class based clientele and adjust cash buffers accordingly.

It is uncommon to measure trading-cost effects over longer horizons because cumulating

returns over many days introduces substantial noise.<sup>20</sup> Nonetheless, we examine returns over a 20-trading-day horizon for two reasons. First, a longer horizon allows us to verify that the return differences associated with flow diversification are not reversed, ensuring that the benefits accrue to longer-term shareholders. Second, the 20-day horizon provides a useful bridge to the monthly return frequencies commonly used in the mutual fund literature, enabling a more direct comparison of economic magnitudes. The last column in each panel in Table 4 shows results using 20-day benchmark-adjusted returns with 19 lags used in the Newey-West standard errors. In Panel A, the difference in returns between the low and high *FlowCorr* quintile rises to 1.66 basis points but with a standard error of 1.43. In contrast, the equivalent spread in returns for *UnexpFlowCorr* quintiles quadruples relative to the 3-day window, reaching 4.13 basis points with a standard error of 1.10. This indicates that the return differences associated with unexpected flow correlation are not transitory.

Table 5 complements the above evidence with multivariate regressions of 3-day cumulative benchmark-adjusted returns on prior flow correlation measures. The regressions allow us to explicitly control for outflows on day  $t$ , as well as other covariates that may affect short-horizon fund returns. In addition to the usual controls (TNA, turnover, expense ratios, and loads), two variables are of particular importance. First, including cash buffers ensures that flow diversification is relevant beyond precautionary cash holdings. Second, lagged flow volatility isolates the influence of flow diversification from the broader flow volatility channel. Even with

---

<sup>20</sup> An exception is if a fund has sustained daily outflows over the cumulating horizon. We do not condition our tests on recurring outflows, so our longer-horizon tests are conservative.

these control variables, a remaining concern is that unobservable fund-manager skill could be correlated with flow diversification, creating an omitted-variable bias. In Column (1) we use fund fixed effects to absorb the variation in returns due to skill. Given our return measures are constructed over overlapping short-horizon windows, fund fixed effects also allow us to exploit within-fund variation over time without restricting identification to within-quarter changes. Column (2) further adds day fixed effects to capture common liquidity or market-wide shocks on certain days.

Unlike the sorts, the multivariate regressions reveal a statistically significant negative relation between *FlowCorr* and subsequent returns. In Columns (1) and (2), the coefficients on *FlowCorr* are -0.44 and -0.52 with standard errors of 0.13 and 0.12; a one-standard-deviation decrease in *FlowCorr* is associated with a 3.88% and a 4.53% increase in the average 3-day benchmark-adjusted return, respectively. Columns (3) and (4) replace *FlowCorr* with its decomposed counterparts. Consistent with the portfolio sorting results in Table 4, the coefficients on *UnexpFlowCorr* are negative and statistically significant while the coefficient on *ExpFlowCorr* is indistinguishable from zero. A one-standard-deviation decrease in *UnexpFlowCorr* implies a 3% improvement in the 3-day benchmark-adjusted return (0.22 basis points).

Table 6 extends the analysis to 20-day benchmark-adjusted returns. Although longer-horizon return regressions naturally introduce more noise, an issue we noted in the context of Table 4, our earlier discussion regarding inferences from longer horizons also applies in this regression setting: the longer window allows us to assess whether the trading-cost effects associated with flow diversification persist rather than reverse. Consistent with the univariate results, the coefficients

on *FlowCorr* remain negative and statistically significant. Moreover, decomposing *FlowCorr* again reveals that the predictive power resides in *UnexpFlowCorr*: the coefficient on *ExpFlowCorr* remains statistically indistinguishable from zero, while the coefficient on *UnexpFlowCorr* continues to be negative and statistically significant. Quantitatively, a one-standard-deviation decrease in *UnexpFlowCorr* implies a 1.5% increase in the 20-day benchmark-adjusted return (0.47 basis points).

Taken together, the return evidence in this section and the cash-holding results in Section 5.1 support the interpretation that predictable differences in flow correlation are partly absorbed through managerial liquidity management, while unexpected increases in flow correlation are associated with lower post-outflow returns.

### **5.2.2 Statistical Inference and Economic Magnitudes**

In Tables 5 and 6, inference is based on robust standard errors clustered at the fund-quarter level. Given the nature of panel data, cross-sectional dependence, serial correlation, and overlapping in returns when the horizon extends to 20 days, we evaluate several alternative methods of computing standard errors. In Appendix Table 4 we report three additional specifications: (a) fund fixed effects with Driscoll and Kraay (1998) standard errors, which directly address cross-sectional dependence and overlapping return windows; (b) fund and day fixed effects with standard errors clustered by fund and day, which accommodates cross-sectional dependence and within-fund serial correlation; and (c) fund and day fixed effects with standard errors clustered by fund, ensuring that inference is not overly conservative when two-way clustering is unnecessary. Across all specifications, the statistical significance of our estimates

remains unchanged, indicating that our results are robust to the challenges posed by the panel structure and overlapping returns.

Moving from statistical inference to economic quantities, several comparisons help contextualize the size of the effects. The baseline impact of large outflows on returns is about 1.22 basis points (Appendix Table 3) and the sort-based performance spread associated with unexpected flow correlation is 1.07 basis points (Table 4). Thus, the mitigation of liquidity externalities due to flow diversification is of the same order of magnitude as the externality it offsets. A second benchmark is the underperformance attributable to flow volatility documented by Rakowski (2010). His estimates come from regressions using daily data and show that the difference in 3-factor alphas between low and high flow volatility quartiles is 0.89 basis points per day (as reported in his Table 3), again comparable to the performance spread in Table 4.

One concern is that because daily *FlowCorr* is volatile, it may have limited influence on managerial decisions or longer-term performance.<sup>21</sup> Several considerations indicate otherwise. First, our decomposition shows that daily *FlowCorr* contains an average component that is relevant for managerial decisions. As shown in Table 3, higher *ExpFlowCorr* is associated with higher cash buffers, consistent with managers incorporating predictable flow correlation into liquidity management. Second, the performance differences over longer horizons are also economically large. Again, a comparison is helpful. Chen, Goldstein, and Jiang (2010) report that for the least liquid funds in their sample, large outflows in the current month are associated with 13 to 19 basis

---

<sup>21</sup> Of course, a variable need not be persistent to be economically important on an ex ante basis. As a case in point, the daily market risk premium is not predictable and yet portfolio managers routinely equitize cash inflows because of the *expected* underperformance of being in cash.

point decline in next month's returns. Our estimate of the 20-day (approximately monthly) return spread associated with unexpected flow correlation from Table 4 is 4.13 basis points, representing 34% to 24% of their estimates.<sup>22</sup> Lastly, in interpreting these results, it is important to emphasize the distinction between clientele and client diversification in the context of the decomposition of *FlowCorr* into its expected and unexpected components. The former reflects predictable features of the investor base that can, to some degree, be influenced by fund managers through share class additions. The latter is outside fund managers' control but the less correlated the client mix, the larger the benefits of unexpected flow correlation.

There are economic and managerial implications of these results. Predictable differences in flow correlation appear to be incorporated into liquidity management, but managers' ability to influence flow diversification is limited. In practice, the main levers available to managers are the creation of new share classes and the opening of additional distribution channels, both of which can broaden the clientele base and alter expected flow patterns. Unexpected synchronization in individual investors' liquidity needs, however, is harder to manage directly. This limitation motivates the exploration of alternative architectural mechanisms through which asset managers or regulators might mitigate outflow externalities. We turn to these next.

### **5.3 Alternative Mechanisms**

One mechanism to manage trading induced by redemptions is to impose swing pricing rules in which redeeming shareholders bear the cost of a fund's trading rather than passing those

---

<sup>22</sup> The performance degradation in Chen, Goldstein, and Jiang (2010) is for the most illiquid funds whereas ours are based on the full sample of all funds, implying that this comparison is, if anything, conservative

costs to remaining investors. Jin, Kacperczyk, Kahraman, and Suntheim (2022) use investor-level transaction data in U.K. corporate bond funds to show that swing pricing rules reduce outflows that occur during times of market stress. To determine if swing pricing is prevalent in our sample, we extract information from N-CEN for all US-domiciled mutual funds on whether the fund is engaged in swing pricing between 2019 and 2024.<sup>23</sup> Unlike their U.K. peers, none of the U.S. equity funds in our sample employ swing pricing during this period. Thus, swing pricing does not seem to operate as a mitigating mechanism for the funds we study.

Another mechanism is redemption-in-kind, which allows funds to meet redemptions by transferring securities rather than selling them. Agarwal, Ren, Shen, and Zhao (2023) document that while a sizeable and growing fraction of mutual funds reserve the option to redeem in kind, actual usage is rare: they observe only 2,985 redemptions-in-kind across 58,443 fund-years. A more recent structural response by asset managers is the conversion of mutual funds into ETFs, which exploits in-kind creation and redemption processes to minimize trading and tax externalities. Prominent asset management companies including Fidelity and JP Morgan have undertaken such conversions under Rule 6c-11 (see Du, Starks, and Xiaolan (2023)). Relatedly, the expiration of Vanguard’s patent on ETFs as a share class in 2023 has prompted some asset managers to seek regulatory approval to include ETFs as a share class within a mutual fund.<sup>24</sup> Both innovations

---

<sup>23</sup> The specific information we extract from N-CEN is based on question C.21, “Did the fund (if not a Money Market Fund, Exchange-Traded Fund, or Exchange-Traded Managed Fund) engage in swing pricing?” N-CEN is only available after 2019.

<sup>24</sup> Funds that are granted dual class permission effectively give mutual funds investors an exchange privilege which permits them to trade their mutual fund holdings for the equivalent ETF share class without triggering capital gains distributions. As pointed out by numerous practitioners, the “cost” of the subsequent shared tax exposure is that ETF shareholders could pay for taxes on gains that they would otherwise not be exposed to (see, for example, <https://www.morningstar.com/funds/etf-share-classes-are-go-dimensional-heres-what-investors-need-know>).

share a common objective: shifting trading activity away from the fund and thereby sidestepping liquidity externalities imposed by redemptions.

Notwithstanding these alternative mechanisms, our results indicate that flow diversification itself operates as a natural buffer. While redemptions create immediacy demands that can depress fund returns, greater heterogeneity in investor liquidity needs reduces the likelihood that redemptions arrive simultaneously. This mitigating effect is particularly important in settings where shareholders anticipate the redemptions of others, since such expectations can generate multiplier effects or shareholder runs, contributing to fragility in financial markets. We turn to this multiplier effect next.

#### **5.4 Multiplier Effects**

Chen, Goldstein, and Jiang (2010) argue that investors in illiquid funds face a higher degree of strategic complementarity because redemptions from illiquid assets impose larger trading costs on remaining shareholders. Exploiting this insight, their test of strategic complementarity consists of regressions of flows on prior performance in which the key explanatory variable is the interaction between an illiquidity indicator and prior performance. Using monthly flow and return data at the share-class (not fund) level, they find a positive interaction coefficient, concluding that outflows are more sensitive to poor performance in illiquid funds.

If a fund has diversified flows, then shareholders may be less concerned with the exit of other investors, weakening strategic complementarities. This intuition translates into two tests. First, we aim to determine whether funds with higher flow diversification have lower flow-performance sensitivity, notwithstanding the higher flow-performance sensitivity of illiquid funds.

Second, we ask whether flow-performance sensitivities in illiquid funds vary with flow diversification; in other words, can flow diversification at least partially offset strategic complementarities, even when they exist? To implement these tests, we make two modifications to the Chen, Goldstein, and Jiang (2010) approach. First, because we use higher-frequency daily flow data to measure liquidity effects, we measure performance over the prior three days rather than using alphas from a factor model.<sup>25</sup> Second, we estimate regressions at the fund rather than the share-class level to align our fund-level flow correlation measure built from share-class flows. Aside from these changes, the regressions closely mirror their Table 2 and contain the same set of control variables. Our key additions are: (i) an indicator variable *FlowCorr(I)*, equal to one for funds with higher flow correlation (above median), or lower flow diversification, (ii) interaction terms between *FlowCorr(I)* and prior performance, and (iii) triple interaction terms between *FlowCorr(I)*, prior performance, and the illiquidity indicator. We estimate these regressions separately for the fund-days with inflows versus outflows because strategic complementarities should be most relevant for outflows.

Panels A and B in Table 7 report results for outflows and inflows, respectively. The dependent variable is winsorized proportional fund flows (at the 1<sup>st</sup> and 99<sup>th</sup> percentiles). The first two columns in each panel include fund fixed effects, while the last two columns add day fixed effects to absorb common shocks in the time series. Across all specifications, the interaction term

---

<sup>25</sup> Daily data are more precise to measure liquidity effects but likely poorer in measuring flow-performance sensitivities. Chen, Goldstein, and Jiang (2010) use factor models and style-adjusted excess returns measured over the past six months. However, there is little theoretical guidance on the appropriate lookback period for flow-performance sensitivities, so all choices are somewhat arbitrary. We use a 3-day period to roughly correspond to tests in earlier tables. The results are similar using a 10-day window.

between the illiquidity indicator and prior performance is consistently positive and statistically significant, replicating the results in Chen, Goldstein, and Jiang (2010). Using monthly data, their estimates imply that flow-performance sensitivity is 52% higher in illiquid funds; by comparison, in Column (3) with fund and day fixed effects, the equivalent increase in flow-performance sensitivity is 49.4% ( $0.166 / (0.170+0.166)$ ), remarkably similar in magnitude.<sup>26</sup> This correspondence suggests that our higher-frequency empirical setting captures the same strategic complementarities mechanism.

Columns (1) and (3) show that the interactions between *FlowCorr(I)* and prior performance are positive and statistically significant, indicating that flow-performance sensitivity is higher in funds with larger flow correlation (or equivalently, lower flow diversification). In Column (3), which includes fund and day fixed effects, the magnitude of this difference is economically meaningful: dividing the coefficient on the interaction between *FlowCorr(I)* and *Perf.* by the coefficient on *Perf.* alone implies that flow-performance sensitivity is 49% ( $0.084/0.170$ ) higher for funds with lower flow diversification. Panel B shows a similar positive interaction for inflows, suggesting that less diversified clienteles may be more performance sensitive. These results show that even in the presence of strategic complementarities, flow diversification continues to play an important mitigating role.

We next turn to the specifications most relevant for our mechanism, which include the triple interaction terms between *FlowCorr(I)*, prior performance, and the illiquidity indicator. The

---

<sup>26</sup> As expected, the increase in the flow-performance sensitivity for inflows in Panel B is much smaller at only 19% ( $0.092 / (0.092 + 0.385)$ ).

goal is to ask whether flow-performance sensitivities in illiquid funds vary with flow diversification. Across all columns in Panel A, the coefficients on the triple interaction are significantly positive, indicating that illiquid funds with low flow diversification exhibit the strongest strategic complementarities. In Column (4), that includes fund and day fixed effects, the flow-performance sensitivity of illiquid funds with high flow diversification is  $0.200 + 0.113 = 0.313$ . For low flow-diversification funds, the coefficient on the triple interaction is 0.100, implying that the increase in flow-performance sensitivity is almost 31% higher ( $0.100/0.313$ ). In contrast, in Panel B, the triple interaction effect is indistinguishable from zero. Taken together, these results indicate that while low flow diversification is associated with greater performance sensitivity more generally, the amplification of strategic complementarities in illiquid funds is concentrated in outflows, where redemption externalities are most relevant.

### **5.5 Alternative Measures of Flow Diversification**

Our main measure of flow diversification relies on the covariance structure of flows across the K share classes within a fund, computed as the value-weighted average of pairwise cross-products of deviations of class-level flows from the fund's net flow, where weights are lagged share-class total net assets. As robustness checks, we consider two alternative approaches that vary either by the weighting scheme or the underlying construction of flow diversification.

A first alternative replaces value weights with equal weights while using proportional (instead of dollar) flows, assigning weight  $1/K$  to each share-class deviation from the fund-level flow. This equal-weighting scheme removes the influence of share-class size and places greater emphasis on dispersion across smaller share classes. We re-estimate the performance regressions

in Table 5 using this equal-weighted flow correlation measure and report the results in Panel A of Appendix Table A5. The first column includes fund fixed effects and the second adds day fixed effects. In both specifications, the coefficients on *FlowCorr* remain negative and similar in magnitude to those obtained using value-weighted correlations, indicating that our results are not sensitive to the weighting scheme. Columns (3) and (4) further decompose the equal-weighted construction of *FlowCorr* into its expected and unexpected components. As before, *UnexpFlowCorr* is negatively associated with future returns but the coefficient on *ExpFlowCorr* is insignificant.

Our second alternative measure focuses on flow imbalance, summing signed dollar flows across the K share classes and scaling by the sum of their absolute values.

$$FlowImb_{i,t} = \frac{\sum_{k=1}^K Flow_{i,k,t}}{\sum_{k=1}^K |Flow_{i,k,t}|} \quad (4)$$

The measure is analogous to order imbalance in the market microstructure literature and resides in the interval [-1,+1]. Since the numerator is signed, the measure is negative when a fund has outflows, and as a result, higher flow imbalance in absolute value (distance from zero) indicates less offsetting flow across share classes and hence lower flow diversification. Since our focus is on outflows, we use the absolute value of flow imbalance in our tests ( $|FlowImb|$ ). The regression results with  $|FlowImb|$  are reported in Panel B of Appendix Table A5. The coefficient on  $|FlowImb|$  is -1.29 with a standard error of 0.24 in Column (1) (for fund fixed effects), and -0.86 with a standard error of 0.23 in Column (2) (for fund and day fixed effects). In Columns (3) and (4), we continue to observe a negative coefficient on the unexpected component of  $|FlowImb|$

but the coefficient on the expected component is indistinguishable from zero.

Thus, these alternative measures, whether varying the weighting in the cross share-class correlation calculation or using a conceptually distinct flow imbalance metric, yield qualitatively identical inferences. In both cases, the return evidence remains concentrated in the unexpected component of flow correlation.

## **6. Conclusion**

The pooling of claims in mutual funds, combined with redemption-on-demand, generates liquidity externalities: transaction costs generated by existing shareholders are passed on to non-redeeming shareholders. A large literature documents that redemptions transmit trading costs, amplify illiquidity, and in some settings induce strategic complementarities that can generate fragility. In this paper, we study a mechanism that mitigates these externalities. When liquidity needs across investors and investor clienteles are imperfectly correlated, funds benefit from flow diversification: outflows from one client or clientele can be offset by inflows from another, thereby reducing daily flow volatility and lowering unexpected outflows. This leads to lower immediacy requirements, mitigating the liquidity externality inherent in the mutual fund structure.

We develop a novel measure of flow diversification using share-class level variation in daily flows. Funds with more diversified flows experience lower future flow volatility, smaller subsequent outflows, and hold smaller cash buffers. While predictable differences are absorbed through managerial liquidity management, unexpected increases in flow correlation are associated with lower post-outflow returns. The magnitude of the return spread associated with unexpected flow diversification is comparable to the liquidity externality itself, underscoring its relevance for

both investors and asset managers.

Our findings connect to several strands of research. First, they reinforce the view that liquidity provision in delegated asset management involves managing not only portfolio liquidity (e.g., Pástor, Stambaugh, and Taylor (2020)) but also the liquidity characteristics of the investor base. Second, our results complement evidence from the financial intermediation literature showing that diversification of funding sources insulates institutions from run dynamics and liquidity shocks (Allen and Gale (2000); Kacperczyk and Schnabl (2013)). Third, by highlighting the role of investor heterogeneity, our analysis links to work on flow-performance sensitivities and strategic complementarities: even in settings where complementarities are strong, diversified flows meaningfully attenuate fragility. Finally, our results help explain emerging practices, such as share-class engineering and fund-to-ETF conversions, as attempts to reshape the correlation structure of investor flows.

## References

- Agarwal, Vikas, Honglin Ren, Ke Shen, and Haibei Zhao, 2023, Redemption in Kind and Mutual Fund Liquidity Management, *The Review of Financial Studies* 36(6), 2274–2318.
- Agnew, Julie, Pierluigi Balduzzi, Annika Sundén, 2003. Portfolio choice and trading in 401(k) Plan. *American Economic Review* 93(1): 193-215.
- Alexander, Gordon, Gjergji Cici, and Scott Gibson, 2007. Does motivation matter when assessing trade performance? An analysis of mutual funds. *Review of Financial Studies* 20(1): 125-150.
- Allen, Franklin, and Douglas Gale, Financial Contagion, 2000. *Journal of Political Economy* 108(1): 1-33.
- Bailey, Jeffrey, and Thomas Richards, 2017. A primer for investment trustees: Understanding investment committee responsibilities. CFA Institute Research Foundation.
- Balchunas, Eric and Joel Weber, 2023, The mutual fund is very much alive and well. Trillions, Bloomberg podcast, August 31, 2023.
- Barber, B. M., Huang, X., and Odean, T., 2016. Which factors matter to investors? Evidence from mutual fund flows. *Review of Financial Studies*, 29(10), 2600-2642.
- Ben-Rephael, A., and R. Israelsen, 2018, Are some clients more equal than others? An analysis of asset management companies' execution costs, *Review of Finance* 22, 1705-1736.
- Bergstresser, Daniel, John Chalmers, and Peter Tufano, 2009. Assessing the costs and benefits of brokers in the mutual fund industry. *Review of Financial Studies* 22(10), 4129-4156.
- Bessembinder, Hendrik, Kalok Chan, and Paul Seguin, 1996. An empirical examination of information, differences of opinion, and trading activities. *Journal of Financial Economics* 40: 105-134.
- Bhattacharya, Utpal, Jung Lee, and Veronika Pool, 2013. Conflicting family values in mutual fund families. *Journal of Finance* 68(1): 173-200.
- Bogouslavsky, Vincent and Pierre Collin-Dufresne, 2023. Liquidity, volume, and order imbalance volatility. *Journal of Finance* 78(4), 2189-2232.
- Britton, Diana, 2019. The slow demise of mutual funds. *WealthManagement.com*, December 10, 2019.
- Campello, Murillo, and Janet Gao, 2017. Customer concentration and loan contract terms. *Journal of Financial Economics* 123(1): 108-136.
- Chalmers, J., and Reuter, J., 2010. What is the Impact of Financial Advisors on Retirement Portfolio Choices & Outcomes? (No. orrc10-05). National Bureau of Economic Research.
- Chang, Eric, Joseph Cheng, and Ajay Khorana, 2000. An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance* 24: 1651-1679.
- Chen, Qi, Itay Goldstein, and Wei Jiang, 2010. Payoff complementarities and financial fragility: Evidence from mutual fund outflows. *Journal of Financial Economics* 97: 239-262.
- Cheng, E., Frank Fabozzi, Robert Harlow, Wai Lee, and Shaojun Zhang, 2024. Applications of equity derivatives to portfolio management. *Journal of Asset Management* 25: 589-599.
- Chernenko, Sergey and Adi Sunderam, 2020. Do fire sales create externalities? *Journal of Financial Economics* 135(3): 602-628.
- Chordia, Tarun, 1996. The structure of mutual fund charges. *Journal of Financial Economics* 41(1), 3-39.

- Christie, William and Roger Huang, 1995. Following the pied piper: Do Individual returns herd around the market? *Financial Analysts Journal* 4: 31-37.
- Christoffersen, Susan E. K., and David K. Musto, 2002. Demand curves and the pricing of money management. *Review of Financial Studies* 15 (5): 1499–1524.
- Coval, Joshua and Erik Stafford, 2007. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics* 86(2): 479–512.
- Del Guercio, D., Genc, E., and Tran, H., 2018. Playing favorites: Conflicts of interest in mutual fund management. *Journal of Financial Economics* 128, 535–557
- Del Guercio, D., and Jonathan, Reuter, 2014. Mutual fund performance and the incentive to generate alpha. *Journal of Finance* 69(4), 1673-1704.
- Del Guercio, D., and Tkac, P. A., 2002. The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds. *Journal of Financial and Quantitative Analysis*, 37(04), 523-557.
- Dickson, Joel, John B. Shoven, and Clemens Sialm, 2000. Tax Externalities of Equity Mutual Funds, NBER Working Papers 7669, National Bureau of Economic Research, Inc.
- Driscoll, J. C., & Kraay, A. C., (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *The Review of Economics and Statistics*, 80(4), 549-560.
- Du, Linda, Laura Starks, and Mindy Z. Xiaolan, 2023. Active ETFs cloned from mutual funds: competing for investor flows. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4653702](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4653702)
- Edelen, Roger, 1999. Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics* 53(3): 439–466
- Edelen, Roger, Richard Evans, and Gregory Kadlec, 2013. Shedding light on “invisible” costs: Trading costs and mutual fund performance. *Financial Analysts Journal* 69(1), 33–44.
- Etula, Erkko, Kalle Rinne, Matti Suominen, and Lauri Vaittinen, 2020. Dash for cash: Monthly market impact of institutional liquidity needs. *Review of Financial Studies* 33(1), 75–111.
- Evans, Richard and Rudiger Fahlenbrach, 2012. Institutional investors and mutual fund governance: evidence from retail-institutional fund twins. *Review of Financial Studies* 25(12), 3530-3571.
- Falato, Antonio, Itay Goldstein, and Ali Hortaçsu, 2021. Financial fragility in the COVID-19 crisis: The case of investment funds in corporate bond markets. *Journal of Monetary Economics* 123: 35-52.
- Falato, Antonio, Ali Hortaçsu, Dan Li, and Chaehee Shin, 2021. Fire-Sale Spillovers in Debt Markets. *Journal of Finance* 76(6): 3055-3102.
- Ferreira, Miguel, Massimo Massa, and Pedro Matos, 2017, Investor-stock decoupling in mutual funds. *Management Science* 64(5), 2144-2163.
- Goldstein, Itay, Hao Jiang, and David Ng, 2017, Investor flows and fragility in corporate bond funds. *Journal of Financial Economics* 126(3): 592-613.
- Greene, Jason and Charles Hodges, 2002. The dilution impact of daily fund flows on open-end mutual funds. *Journal of Financial Economics* 65(1): 131-158.
- Helmke, Anna, 2023. Will ETFs drive mutual funds extinct? [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4614021](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4614021)

- Illan, Ivan, 2018. How monitoring mutual fund flows gives unique stock market insight. Forbes Financial Council. <https://www.forbes.com/sites/forbesfinancecouncil/2018/03/05/how-monitoring-mutual-fund-flows-gives-unique-stock-market-insight/?sh=9847fae35f65>
- Investment Company Institute Factbook, 2024. Washington, DC: Investment Company Institute.
- James, C., and Karceski, J., 2006. Investor monitoring and differences in mutual fund performance. *Journal of Banking and Finance* 30(10), 2787-2808
- Jiang, Hao, Yi Li, Zheng Sun, and Ashley Wang, 2022. Does mutual fund illiquidity introduce fragility into asset prices? Evidence from the corporate bond market. *Journal of Financial Economics* 143(1), 277-302.
- Jin, Dunhong, Marcin Kacperczyk, Bige Kahraman, and Felix Suntheim, 2022. Swing pricing and fragility in open-end mutual funds. *Review of Financial Studies* 35(1), 1-50.
- Johnson, Woodrow, 2004. Predictable investment horizons and wealth transfers among mutual fund shareholders. *Journal of Finance* 59(5): 1979-2012.
- Kacperczyk, Marcin and Philipp Schnabl, 2013. How safe are money market funds? *Quarterly Journal of Economics*: 1073-1122.
- Massa, Massimo, 2003. How do family strategies affect fund performance? When performance-maximization is not the only game in town. *Journal of Financial Economics* 67: 249-304.
- Morris, Stephen, Ilhyock Shim, and Hyun Song Shin, 2017. Redemption risk and cash hoarding by asset managers. *Journal of Monetary Economics* 89: 71-87.
- Nanda, Vikram, Jay Wang, and Lu Zheng, 2009. The ABCs of mutual funds: On the introduction of multiple share classes. *Journal of Financial Intermediation* 18(3): 329-361.
- Pástor, Lubos, Robert F. Stambaugh, and Lucian A. Taylor, 2020. Discretionary Liquidity Management, *Journal of Financial Economics*, 137(1), 135–160.
- Rakowski, David, 2010. Fund Flow Volatility and Performance. *Journal of Financial and Quantitative Analysis* 45: 223-237.
- Rekenthaler, John, 2021. Farewell, mutual funds. Morningstar, September 2, 2021.
- Schmidt, Lawrence, Allan Timmermann, and Russ Wermers, 2016. Runs on money market mutual funds. *American Economic Review* 106(9): 2625-2657.
- Sialm, Clemens, and Laura, Starks, 2012. Mutual fund tax clienteles. *Journal of Finance* 67(4), 1397-1422.
- Sialm, Clemens, and Hanjiang Zhang, 2020. Tax-efficient asset management: Evidence from equity mutual funds. *Journal of Finance* 75(2): 735-777.
- Simutin, Mikhail, 2014. Cash Holdings and Mutual Fund Performance. *Review of Finance* 18(4), 1425–1464.
- Wermers, Russ, 2000. Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses. *Journal of Finance* 55(4):1655–1695.
- Williams, Beggs, 2022. The company you keep: Investment adviser clientele and mutual fund performance, *Journal of Financial Intermediation* 50.
- Yan, Xuemin (Sterling), 2008. Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance. *Journal of Financial and Quantitative Analysis* 43(3), 741–767.

**Table 1: Determinants of Flow Diversification**

The dependent variable is *FlowCorr* on day *t*, defined as the value-weighted average covariance of flows across share classes, scaled by the average of squared deviations of share-class level flows from fund-level net flows. Independent variables are defined in Appendix Table 1 and are contemporaneous unless specified otherwise. The sample consists of all fund-days over the 2007-2024 period. Robust standard errors, clustered at the fund-quarter level are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Log (Fund Size)	-0.004*** (0.001)	-0.003*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)
Expense Ratio	0.027*** (0.004)	0.045*** (0.005)	0.027*** (0.004)	0.045*** (0.005)	0.027*** (0.004)	0.045*** (0.005)
Turnover	-0.006* (0.003)	-0.005* (0.003)	-0.006* (0.003)	-0.005* (0.003)	-0.006* (0.003)	-0.005* (0.003)
StdDev (Flow <sub>t-2,t</sub> )	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)	0.011*** (0.002)
Log (Age)	-0.011*** (0.001)	-0.027*** (0.002)	-0.011*** (0.001)	-0.027*** (0.002)	-0.011*** (0.001)	-0.027*** (0.002)
Load	-0.002 (0.144)	0.577*** (0.166)	-0.002 (0.144)	0.577*** (0.166)	-0.002 (0.144)	0.577*** (0.166)
Flow-Perf. Sensitivity	0.000 (0.003)	0.003 (0.003)	0.000 (0.003)	0.003 (0.003)	0.000 (0.003)	0.003 (0.003)
Rm <sub>t</sub>	-0.025** (0.011)	-0.023** (0.011)	-0.025** (0.011)	-0.023** (0.011)	-0.025** (0.011)	-0.023** (0.011)
Rm <sub>t-3,t-1</sub>	-0.086*** (0.009)	-0.084*** (0.008)	-0.086*** (0.009)	-0.084*** (0.008)	-0.086*** (0.009)	-0.084*** (0.008)
Sentiment	-0.016*** (0.005)	-0.023*** (0.005)	-0.016*** (0.005)	-0.023*** (0.005)	-0.016*** (0.005)	-0.023*** (0.005)
December Indicator	0.036*** (0.001)	0.034*** (0.001)	0.036*** (0.001)	0.034*** (0.001)	0.036*** (0.001)	0.034*** (0.001)
Month-End Indicator	0.008*** (0.001)	0.007*** (0.001)	0.008*** (0.001)	0.007*** (0.001)	0.008*** (0.001)	0.007*** (0.001)
Share Class Increase Indicator			-0.021*** (0.005)	-0.019*** (0.005)		
Retail Class Increase Indicator					-0.035*** (0.009)	-0.032*** (0.009)
Institutional Class Increase Indicator					-0.024*** (0.006)	-0.022*** (0.006)
Retirement Class Increase Indicator					-0.019 (0.016)	-0.019 (0.016)
Constant	0.255*** (0.020)	0.240*** (0.020)	0.255*** (0.020)	0.240*** (0.020)	0.255*** (0.020)	0.240*** (0.020)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Style-Quarter FE		Yes		Yes		Yes
N	8,172,510	8,172,510	8,169,259	8,169,259	8,172,510	8,172,510
Adj-R <sup>2</sup>	0.060	0.062	0.060	0.062	0.060	0.062

**Table 2: Flow Diversification, Future Flow Volatility, and Subsequent Flows**

This table examines whether *FlowCorr* predicts future flow volatility and subsequent outflows. In Panel A, the dependent variable is daily flow volatility ( $\text{StdDev}(\text{Flow}_{t+1,t+3})$ ), measured as the square root of the sum of squared daily fund flows from  $t+1$  to  $t+3$ . In Panel B, the dependent variable is Fund Flow on day  $t+1$ . Independent variables are defined in Appendix Table 1 and are measured at day  $t$  unless specified otherwise. *FlowCorr* is the value-weighted average covariance of flows across share classes, scaled by the average of squared deviations of share-class level flows from fund-level net flows. The sample consists of fund-days with net outflows over the 2007-2024 period. Robust standard errors, clustered at the fund-quarter level are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $\text{StdDev}(\text{Flow}_{t+1,t+3})$ Regressions						
<i>FlowCorr</i>	0.680*** (0.030)	0.470*** (0.030)	0.114*** (0.029)	0.433*** (0.029)	0.273*** (0.028)	0.054* (0.030)
StdDev ( $\text{Flow}_{t-2,t}$ )		0.299*** (0.024)	0.298*** (0.024)		0.290*** (0.024)	0.291*** (0.024)
VIX $_{t-2,t}$			0.001 (0.002)			0.523 (0.320)
Log (Fund Size)			0.634*** (0.043)			0.740*** (0.038)
Log (Fund Flow)			0.475*** (0.045)			0.315*** (0.037)
Constant	2.607*** (0.030)	1.942*** (0.062)	-9.365*** (0.921)	2.656*** (0.031)	2.001*** (0.062)	-22.188*** (6.514)
N	5,149,587	5,144,264	5,071,573	5,149,457	5,144,135	5,071,510
Adj-R <sup>2</sup>	0.038	0.087	0.087	0.051	0.096	0.096
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE				Yes	Yes	Yes
Panel B: $\text{Flow}_{t+1}$ Regressions						
<i>FlowCorr</i>	-0.229*** (0.020)	-0.209*** (0.022)	-0.249*** (0.025)	-0.163*** (0.019)	-0.148*** (0.021)	-0.177*** (0.023)
StdDev ( $\text{Flow}_{t-2,t}$ )		-0.030** (0.014)	-0.030* (0.016)		-0.026* (0.014)	-0.026* (0.016)
Fund Flow			0.008 (0.013)			0.009 (0.013)
Log (Fund Size)			-1.155*** (0.036)			-1.141*** (0.037)
Expense Ratio			0.392*** (0.083)			0.225*** (0.078)
Turnover			0.019 (0.082)			0.064 (0.088)
Cash			-0.005*** (0.002)			-0.006*** (0.002)
Load			-0.025 (0.041)			0.150*** (0.042)
Constant	-1.781*** (0.018)	-1.714*** (0.036)	20.359*** (0.707)	-1.793*** (0.018)	-1.735*** (0.036)	20.129*** (0.703)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes
Day FE				Yes	Yes	Yes
N	5,121,114	5,116,074	4,741,277	5,120,999	5,115,960	4,741,127
Adj-R <sup>2</sup>	0.021	0.022	0.023	0.027	0.028	0.029

**Table 3: Flow Diversification and Cash Holdings**

Panel A contains regressions in which the dependent variable is end-of-quarter cash and cash equivalents scaled by total assets, winsorized at the 0.5<sup>th</sup> and 99.5<sup>th</sup> percentiles. The unit of observation is a fund-quarter and the sample consists of all fund-quarters. *ExpFlowCorr* is the average flow correlation over the quarter. Regressions (2) and (4) include the following control variables: Log (Fund Size), Log (Fund Flow), Fund Returns, StdDev (Flow), Expense Ratio, Turnover and Load. In Panel A, robust standard errors clustered at the fund level appear in parentheses. Panel B contains regressions in which the dependent variable is cash on day t+20 minus cash on day t-20. The unit of observation is a fund-day and the sample consists of all fund-days over the 2007-2024 period. All regressions include the same control variables as in Panel A measured on day t, with the exception of StdDev (Flow<sub>t-2,t</sub>) which uses a 3-day window. The Share Class Increase Indicator is equal to one if the number of share classes increases on day t. The Large Share Class Increase Indicator is equal to one if the number of post-increase share classes is less than or equal to four (i.e., if the number of share classes increases from one, two, or three). The Small Share Class Increase Indicator is equal to one if the number of post-increase share classes is greater than four. Outflow Indicator is equal to one if the fund receives net outflows on day t. In Columns 6, 8, and 10, standard errors are clustered at the fund and day level. In Columns 5, 7, and 9, standard errors are based on Driscoll and Kraay (1998). Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Regressions of End-of Quarter Cash				Panel B: Regressions of Daily Changes in Cash					
<i>ExpFlowCorr</i>	0.744*** (0.201)	0.384** (0.195)	0.898*** (0.272)	0.476* (0.260)	-0.083 (0.072)	-0.099 (0.078)				
Share Class Increase Indicator							-0.123** (0.050)	-0.119** (0.053)		
Large Share Class Increase Indicator									-0.216** (0.087)	-0.217** (0.090)
Small Share Class Increase Indicator									-0.038 (0.055)	-0.029 (0.058)
Outflow Indicator					0.024** (0.011)	0.021* (0.011)				
<i>ExpFlowCorr</i> x Outflow					0.170*** (0.063)	0.158** (0.065)				
Controls		Yes		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund-Year FE	Yes	Yes								
Style-Year FE			Yes	Yes						
Fund FE			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FE						Yes		Yes		Yes
N	140,345	130,736	141,240	131,425	8,021,220	8,021,140	8,017,388	8,017,194	8,021,069	8,021,261
Adj-R <sup>2</sup>	0.834	0.846	0.691	0.690	0.021	0.021	0.019	0.021	0.019	0.021

**Table 4: Flow Diversification and Short-Horizon Fund Returns**

This table reports portfolio sorts that examine the relation between flow diversification and subsequent fund returns. We sort all fund-days with net outflows on day  $t$  into quintiles based on *FlowCorr* (Panel A), *UnexpFlowCorr* (Panel B), and *ExpFlowCorr* (Panel C). *FlowCorr* is defined as the value-weighted average covariance of flows across share classes, scaled by the average of squared deviations of share-class level flows from fund-level net flows. *ExpFlowCorr* is the predictable component of *FlowCorr* based on a backward-looking 60-day moving average, and *UnexpFlowCorr* is the deviation of *FlowCorr* from this predictable component. The table reports benchmark-adjusted returns on  $t+1$ ,  $t+2$ ,  $t+3$ , the cumulative return from  $t+1$  to  $t+3$ , and the cumulative return from  $t+1$  to  $t+20$ , all in basis points. The last row in each panel shows the difference in returns between low and high quintiles with Newey-West standard errors in parentheses. The last column in each panel reports the time series average of the number of funds in each quintile. The sample consists of all fund-days with net outflows over the 2007-2024 period. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	XR <sub>t+1</sub>	XR <sub>t+2</sub>	XR <sub>t+3</sub>	XR <sub>t+1, t+3</sub>	XR <sub>t+1, t+20</sub>	# of Funds
Panel A: Benchmark-Adjusted Returns on Sorts of <i>FlowCorr</i>						
1 (Low)	-1.959	-1.390	-1.400	-4.670	-29.215	230
2	-1.641	-1.254	-1.623	-4.622	-28.742	229
3	-1.566	-1.522	-1.445	-4.514	-28.345	229
4	-1.937	-1.415	-1.564	-4.883	-28.425	229
5 (High)	-2.173	-1.678	-1.491	-5.224	-30.872	229
Low-High	0.214 (0.184)	0.288 (0.176)	0.091 (0.184)	0.554 (0.354)	1.662 (1.426)	
Panel B: Benchmark-Adjusted Returns on Sorts of <i>UnexpFlowCorr</i>						
1 (Low)	-1.796	-1.405	-1.281	-4.460	-28.345	230
2	-1.729	-1.301	-1.378	-4.385	-28.049	229
3	-1.781	-1.310	-1.447	-4.616	-27.210	229
4	-1.859	-1.413	-1.790	-4.828	-28.242	229
5 (High)	-2.247	-1.796	-1.543	-5.534	-32.654	229
Low-High	0.451** (0.182)	0.391** (0.168)	0.262 (0.163)	1.071*** (0.297)	4.128*** (1.098)	
Panel C: Benchmark-Adjusted Returns on Sorts of <i>ExpFlowCorr</i>						
1 (Low)	-2.077	-1.599	-1.643	-5.458	-31.400	230
2	-1.975	-1.493	-1.628	-4.895	-29.382	229
3	-1.738	-1.135	-1.306	-4.138	-27.581	229
4	-1.652	-1.338	-1.341	-4.284	-27.458	229
5 (High)	-1.996	-1.604	-1.532	-4.981	-29.389	229
Low-High	-0.080 (0.220)	0.005 (0.191)	-0.110 (0.196)	-0.467 (0.426)	-1.461 (2.634)	

**Table 5: Flow Diversification and Short-Horizon Fund Returns: Multivariate Regressions**

The table contains panel regressions in which the dependent variable is the 3-day cumulative benchmark-adjusted return from  $t+1$  to  $t+3$ . Columns (1) and (2) use *FlowCorr* measured on day  $t$ . Columns (3) and (4) decompose *FlowCorr* into *ExpFlowCorr* and *UnexpFlowCorr* using a backward-looking 60-day moving-average filter, measured from  $t-61$  to  $t-1$ , corresponding to one quarter. Other independent variables are defined in Appendix Table 1 and are measured at day  $t$  unless specified otherwise. The sample consists of fund-days with net outflows over the 2007-2024 period. Robust standard errors, clustered at the fund-quarter level are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
<i>FlowCorr</i>	-0.443*** (0.128)	-0.517*** (0.121)		
<i>ExpFlowCorr</i>			0.004 (0.600)	-0.024 (0.577)
<i>UnexpFlowCorr</i>			-0.480*** (0.129)	-0.557*** (0.122)
Log (Fund Size)	-1.893*** (0.122)	-2.005*** (0.119)	-1.889*** (0.122)	-2.000*** (0.119)
Log (Fund Flow)	0.056 (0.041)	-0.031 (0.038)	0.055 (0.041)	-0.032 (0.038)
Expense Ratio	2.075*** (0.681)	1.074 (0.744)	2.059*** (0.680)	1.050 (0.743)
Turnover	-0.008 (0.039)	-0.010 (0.040)	-0.008 (0.039)	-0.010 (0.040)
StdDev ( $Flow_{t-2,t}$ )	0.002 (0.001)	-0.001 (0.001)	0.002 (0.001)	-0.001 (0.001)
Cash	-0.022 (0.024)	-0.027 (0.023)	-0.022 (0.024)	-0.027 (0.023)
Load	1.292*** (0.263)	-0.175 (0.276)	1.292*** (0.263)	-0.176 (0.276)
Constant	29.456*** (2.559)	33.713*** (2.488)	29.296*** (2.558)	33.558*** (2.489)
Fund FE	Yes	Yes	Yes	Yes
Day FE		Yes		Yes
N	4,648,066	4,647,985	4,648,066	4,647,985
Adj-R <sup>2</sup>	0.003	0.104	0.003	0.104

**Table 6: Flow Diversification and 20-Day Benchmark-Adjusted Fund Returns**

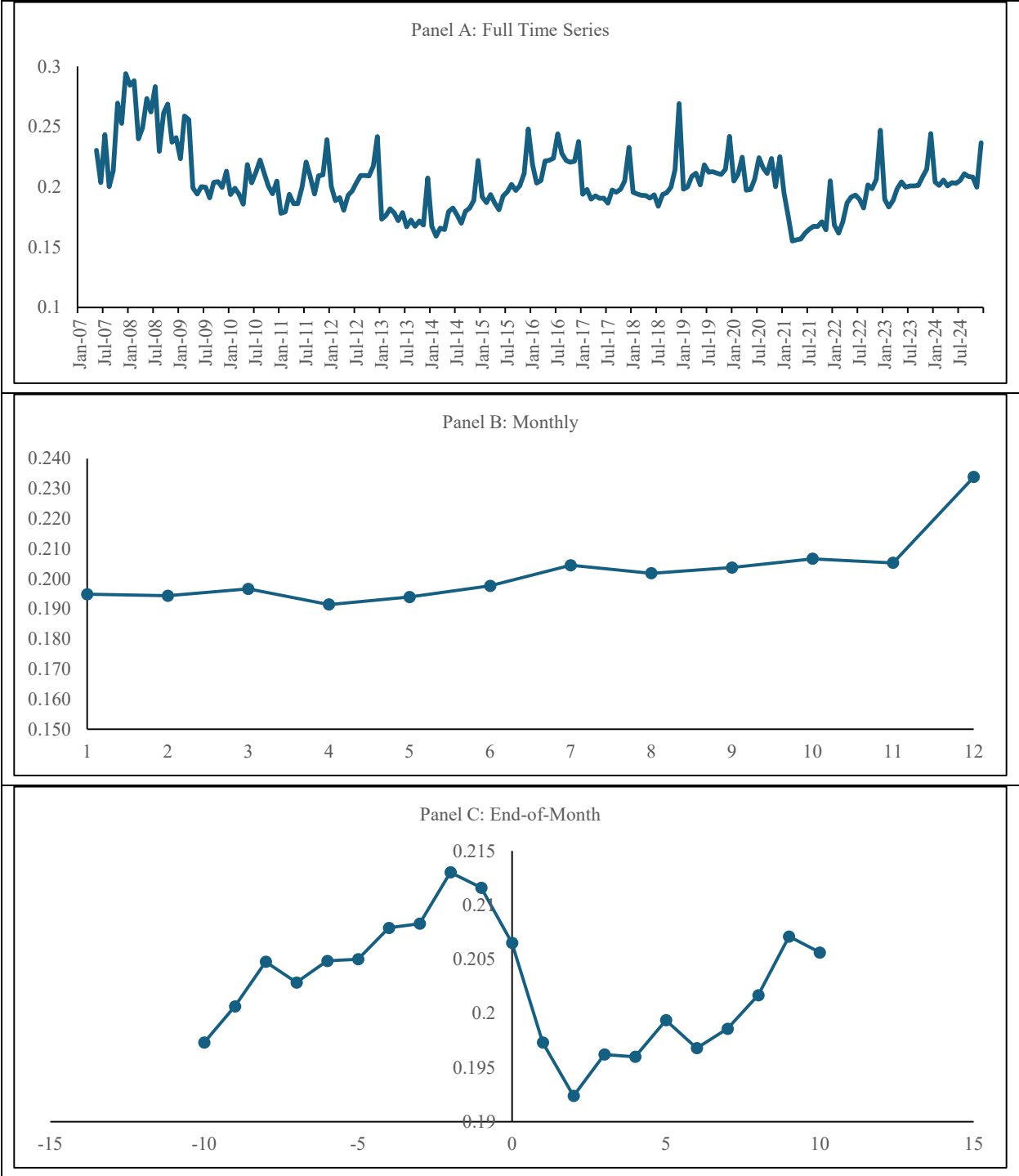
This table reports panel regressions in which the dependent variable is the 20-day cumulative benchmark-adjusted return from  $t+1$  to  $t+20$ . Columns (1) and (2) use *FlowCorr* measured on day  $t$ . Columns (3) and (4) decompose *FlowCorr* into *ExpFlowCorr* and *UnexpFlowCorr* using a backward-looking 60-day moving-average filter, measured from  $t-61$  to  $t-1$ , corresponding to one quarter. Other independent variables are defined in Appendix Table 1 and are measured at day  $t$  unless specified otherwise. The sample consists of fund-days with net outflows over the 2007-2024 period. Robust standard errors, clustered at the fund-quarter level are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
<i>FlowCorr</i>	-0.010** (0.004)	-0.010** (0.004)		
<i>ExpFlowCorr</i>			0.010 (0.036)	0.017 (0.035)
<i>UnexpFlowCorr</i>			-0.012*** (0.004)	-0.012*** (0.003)
Log (Fund Size)	-0.151*** (0.008)	-0.151*** (0.008)	-0.151*** (0.008)	-0.151*** (0.008)
Log (Fund Flow)	0.010*** (0.001)	0.003*** (0.001)	0.010*** (0.001)	0.003** (0.001)
Expense Ratio	0.082* (0.046)	0.042 (0.051)	0.082* (0.046)	0.041 (0.050)
Turnover	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)	0.002 (0.002)
StdDev ( $Flow_{t-2,t}$ )	0.000*** (0.000)	-0.000*** (0.000)	0.000*** (0.000)	-0.000*** (0.000)
Cash	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)	-0.002 (0.002)
Load	0.086*** (0.019)	-0.019 (0.020)	0.086*** (0.019)	-0.019 (0.020)
Constant	2.534*** (0.162)	2.656*** (0.159)	2.527*** (0.161)	2.647*** (0.158)
Fund FE	Yes	Yes	Yes	Yes
Day FE		Yes		YES
N	4,329,763	4,329,747	4,329,763	4,329,747
Adj-R <sup>2</sup>	0.022	0.115	0.022	0.115

**Table 7: Daily Flow-Performance Regressions**

This table reports regressions of daily fund flows on prior fund performance, flow correlation, asset illiquidity, and their interactions. The dependent variable is proportional flow on day  $t$ , winsorized at the 1<sup>st</sup> and 99<sup>th</sup> percentiles. *Perf.* is the cumulative fund return over the prior three days. Panel A reports results for fund-days with cumulative outflows over the prior three days, and Panel B reports results for fund-days with cumulative inflows over the prior three days. *FlowCorr(I)* is an indicator equal to one for funds with higher flow correlation (above median) or lower flow diversification. *Illiq. Indicator* is defined as Chen, Goldstein, and Jiang (2010). All regressions include the following control variables:  $\log(\text{Fund Flow})$ ,  $\log(\text{Fund Size})$ , Expense Ratio, Turnover,  $\text{StdDev}(\text{Flow})$ ,  $\log(\text{Age})$ , and Load. Control variables are defined in Appendix Table 1 and measured on day  $t-1$ , with the exception of  $\text{StdDev}(\text{Flow})$ , which is measured over a 3-day window. Standard errors, clustered at the fund-quarter level, are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Prior 3-Day Outflows				Panel B: Prior 3-Day Inflows			
Perf.	0.358*** (0.017)	0.392*** (0.018)	0.170*** (0.030)	0.200*** (0.030)	0.394*** (0.027)	0.379*** (0.027)	0.385*** (0.044)	0.367*** (0.044)
Illiq. Indicator	0.006 (0.004)	0.007* (0.004)	0.003 (0.004)	0.004 (0.004)	0.001 (0.007)	-0.000 (0.007)	-0.000 (0.007)	-0.001 (0.007)
Perf. * Illiq. Indicator	0.116*** (0.022)	0.054** (0.025)	0.166*** (0.022)	0.113*** (0.025)	0.059* (0.034)	0.085** (0.036)	0.092*** (0.035)	0.121*** (0.037)
<i>FlowCorr(I)</i>	0.009*** (0.000)	0.009*** (0.001)	0.008*** (0.000)	0.009*** (0.001)	0.014*** (0.001)	0.013*** (0.001)	0.012*** (0.001)	0.011*** (0.001)
Perf. * <i>FlowCorr(I)</i>	0.098*** (0.016)	0.035 (0.022)	0.084*** (0.016)	0.029 (0.021)	0.265*** (0.024)	0.299*** (0.038)	0.270*** (0.024)	0.307*** (0.037)
Illiq. Indicator * <i>FlowCorr(I)</i>		-0.001 (0.001)		-0.001 (0.001)		0.002 (0.001)		0.001 (0.001)
Illiq. Indicator * <i>FlowCorr(I)</i> * Perf.		0.116*** (0.032)		0.100*** (0.031)		-0.056 (0.049)		-0.064 (0.049)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes			Yes	Yes		
Fund FE and Day FE			Yes	Yes			Yes	Yes
N	5,062,421	5,062,421	5,062,014	5,062,014	3,213,593	3,213,593	3,213,296	3,213,296
Adj-R <sup>2</sup>	0.025	0.025	0.036	0.036	0.072	0.072	0.080	0.080



**Figure 1 Time-Series Patterns in Flow Correlation:** Panel A plots the monthly average daily flow correlation from January 2007 to December 2024. Panel B reports the average daily flow correlation in each calendar month. Panel C contains the average daily flow correlation from 10 trading days before to 10 trading days after the end of the month (day 0 denotes the last trading day of the month).



**Figure 2 Change of Flow Correlation around Share-Class Additions:** This figure presents an event-study analysis of flow correlation around share-class additions. Each panel shows average flow correlations from 60 trading days before to 60 trading days after the addition of a share class from a different clientele on day 0. Panel A presents results for all share class additions. Panels B through D present results separately for retail, institutional/advisor, and retirement share classes additions.

**Appendix Table 1: Variable Definitions**

<i>FlowCorr</i>	Value-weighted average pairwise cross-product of deviations of share-class-level flows from fund-level net flows, scaled by the average squared value-weighted deviation across share classes.
<i>ExpFlowCorr</i>	The expected (permanent) component of <i>FlowCorr</i> , based on a backward-looking moving average (MA) process of order $q$ , with $q=60$ .
<i>UnexpFlowCorr</i>	The unexpected (transitory) component of <i>FlowCorr</i> , defined as the difference between <i>FlowCorr</i> and <i>ExpFlowCorr</i> .
Log (Fund Size)	The natural logarithm of the fund's daily total net assets (in \$ million).
Fund Flow	Daily dollar fund flow, aggregated across share classes (in \$ million).
Log (Fund Flow)	The natural logarithm of the absolute value of fund's daily dollar flows.
Proportional Flow (%)	Daily fund flow scaled by lagged fund's total net assets.
# Share Classes	The number of share classes associated with a fund on day $t$ .
Expense Ratio (%)	Expense ratio in percent.
Turnover	Turnover ratio.
Age	Fund age in months.
StdDev (Flow)	The square root of the sum of squared daily dollar flows over three consecutive days.
Load (%)	The sum of a fund's front-end load and back-end load in percent.
Cash (%)	Cash and cash equivalents scaled by total assets. Cash equivalents include short-term instruments such as time deposits, commercial paper, money market funds, repurchase agreements, Treasury bills and bonds with nominal maturity below 92 days, currencies, and currency derivatives.
Flow-Perf. Sensitivity	The slope of a regression of daily fund flows on 3-day past returns using 60-day rolling window, requiring at least 20 observations.
Rm	S&P500 daily index return.
$R_{t+1,t+n}$	Cumulative fund returns from $t+1$ to $t+n$ .
$XR_{t+1,t+n}$	Cumulative benchmark-adjusted fund returns from $t+1$ to $t+n$ , where the benchmark is the Morningstar assigned benchmark for the fund.
$VIX_{t-2,t}$	Average daily VIX index from $t-2$ to $t$ .
Sentiment	The monthly change in the University of Michigan Index of Consumer Sentiment.
Perf.	Cumulative fund return over the prior three days
Indicator Variables	
December	A dummy variable equal to one if the month is December, and zero otherwise.
Month-End	A dummy variable equal to one if day $t$ is in the last three trading days of the month, and zero otherwise.
Share Class Increase	A dummy variable equal to one if there is an increase in the number of share classes on day $t$ , and zero otherwise.
Retail Class Increase	A dummy variable equal to one if a retail share class is added on day $t$ and no such class existed prior, and zero otherwise.
Institutional Class Increase	A dummy variable equal to one if an institutional/advisor share class is added on day $t$ and no such class existed prior, and zero otherwise.
Retirement Class Increase	A dummy variable equal to one if a retirement share class is added on day $t$ and no such class existed prior, and zero otherwise.
Large Share Class Increase	A dummy variable equal to one if the number of post-increase share classes is less than or equal to four.
Small Share Class Increase	A dummy variable equal to one if the number of post-increase share classes is more than four.
<i>FlowCorr(I)</i>	A dummy variable equal to one if <i>FlowCorr</i> is above the sample median.
Illiq.	A dummy variable equal to one if the fund is either a small or mid-cap fund, or a single country non-US equity fund, and zero otherwise.
Outflow	A dummy variable equal to one if the fund has an outflow on $t$ , and zero otherwise.

## Appendix Table 2: Descriptive Statistics

The table reports descriptive statistics for the main variables used in the paper. Variable definitions appear in Appendix Table 1. *FlowCorr* measures daily flow correlation across share classes and is decomposed into an expected component (*ExpFlowCorr*), defined as a backward-looking 60-trading-day moving average, and an unexpected component (*UnexpFlowCorr*), defined as the residual. Proportional Flow is winsorized at the 1<sup>st</sup> and 99<sup>th</sup> percentiles, and Cash is winsorized at the 0.5<sup>th</sup> and 99.5<sup>th</sup> percentiles. The sample consists of all fund-days with net outflows over the 2007-2024 period.

	Mean	Median	STD	Min	p1	p10	p25	p75	p90	p99	Max
<i>FlowCorr</i>	0.203	0.226	0.426	-1	-0.977	-0.324	-0.041	0.481	0.749	0.999	1
<i>ExpFlowCorr</i>	0.192	0.177	0.165	-1	-0.167	0.004	0.084	0.288	0.401	0.657	1
<i>UnexpFlowCorr</i>	0.011	0.035	0.394	-1.947	-1.068	-0.486	-0.198	0.234	0.473	0.945	1.598
Fund Flow (\$million)	-1.83	-0.17	33.56	-22687.93	-23.66	-2.23	-0.67	-0.03	-0.01	0.00	0.00
Proportional Flow (%)	-0.142	-0.053	0.304	-1.884	-1.884	-0.280	-0.114	-0.023	-0.009	-0.001	0.000
Fund Size (\$ million)	1479.65	309.94	6820.15	<0.001	1.77	19.61	78.29	1046.07	2797.27	17444.08	307889.20
StdDev (Flow <sub>t-2,t</sub> )	2.372	0.279	31.746	0	0.001	0.017	0.074	0.969	3.179	32.685	17388.290
Expense Ratio (%)	1.126	1.087	0.393	<0.001	0.181	0.708	0.901	1.324	1.588	2.256	6.477
Turnover	0.786	0.520	5.318	<0.001	0.030	0.160	0.290	0.894	1.410	4.380	843.870
Load (%)	0.732	0.333	0.913	0	0	0	0	1.222	2.281	3.257	5.211
Age (Months)	17.834	16.011	12.744	0.003	0.611	4.175	9.096	23.688	31.197	70.293	100.532
Cash (%)	3.970	1.828	9.423	0	0	0.065	0.709	3.629	6.871	56.047	90.669
# Share Classes	4.17	4	2.19	2	2	2	2	5	7	11	21
R <sub>t+1, t+3</sub> (bp)	10.840	20.903	236.71	-6453.16	-708.90	-239.68	-91.92	125.69	247.06	617.90	12423.41
XR <sub>t+1, t+3</sub> (bp)	-4.864	-3.447	105.64	-6422.99	-309.02	-104.83	-44.34	35.69	92.57	289.74	12657.44
R <sub>t+1, t+20</sub> (%)	0.723	1.189	5.81	-90.76	-18.83	-5.55	-1.85	3.79	6.55	14.64	142.73
XR <sub>t+1, t+20</sub> (%)	-0.301	-0.207	2.54	-92.61	-7.76	-2.90	-1.31	0.79	2.14	6.52	145.61

### Appendix Table 3: Post-Flow Average Benchmark-Adjusted Returns

This table reports benchmark-adjusted fund returns following days with net inflows and net outflows. On each day  $t$ , funds are first sorted into inflow and outflow groups, and then into quintiles based on the magnitude of their proportional flows. For outflows, Q1 refers to the quintile with the largest redemptions, while Q5 refers to the quintile with the smallest redemptions. For inflows, Q1 refers to the quintile with the smallest inflows, while Q5 refers to the quintile with the largest inflows. The last row in each panel shows the return spread (5-1). The table reports average benchmark-adjusted returns for  $t+1$ ,  $t+2$ ,  $t+3$ , and the cumulative benchmark-adjusted return from  $t+1$  to  $t+3$ , all in basis points. Panel B reports results for illiquid funds, defined as small-cap, mid-cap, and single country ex-US funds. Panel C contains results for all remaining (i.e. liquid) funds. The sample consists of all fund-days over the 2007-2024 period. Newey-West standard errors appear in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	Outflows				Inflows			
	Ret <sub>t+1</sub>	Ret <sub>t+2</sub>	Ret <sub>t+3</sub>	Ret <sub>t+1,t+3</sub>	Ret <sub>t+1</sub>	Ret <sub>t+2</sub>	Ret <sub>t+3</sub>	Ret <sub>t+1,t+3</sub>
Panel A: All Funds								
1 (Large Redemptions / Small Inflows)	-2.151	-1.764	-1.705	-5.556	-1.309	-1.554	-1.539	-4.308
2	-2.168	-1.464	-1.403	-4.925	-0.699	-1.437	-1.476	-3.499
3	-2.043	-1.304	-1.479	-4.705	-0.658	-1.436	-1.478	-3.432
4	-1.626	-1.506	-1.419	-4.473	-0.944	-1.359	-1.491	-3.664
5 (Small Redemptions / Large Inflows)	-1.576	-1.404	-1.554	-4.339	-0.965	-1.626	-1.743	-4.155
5 - 1	0.575**	0.359*	0.151	1.217***	0.344	-0.072	-0.204	0.153
	(0.231)	(0.201)	(0.224)	(0.401)	(0.220)	(0.232)	(0.204)	(0.441)
Panel B: Illiquid Funds								
1 (Large Redemptions / Small Inflows)	-2.856	-2.477	-2.298	-7.482	-1.994	-2.201	-2.109	-6.109
2	-2.830	-2.048	-2.194	-6.840	-1.581	-2.243	-2.092	-5.700
3	-2.848	-2.047	-2.403	-7.041	-1.265	-1.889	-2.101	-5.016
4	-2.594	-2.049	-2.005	-6.411	-1.347	-1.970	-2.244	-5.367
5 (Small Redemptions / Large Inflows)	-2.005	-2.125	-2.166	-6.021	-1.309	-1.813	-2.193	-5.053
5 - 1	0.851***	0.352	0.132	1.461***	0.686**	0.388	-0.084	1.056*
	(0.272)	(0.274)	(0.268)	(0.533)	(0.292)	(0.286)	(0.283)	(0.556)
Panel C: Liquid Funds								
1 (Large Redemptions / Small Inflows)	-1.249	-1.032	-0.918	-3.160	-0.494	-0.576	-0.671	-1.758
2	-1.449	-0.641	-0.657	-2.736	0.061	-0.659	-0.735	-1.356
3	-1.108	-0.584	-0.705	-2.421	0.025	-0.517	-0.746	-1.238
4	-0.748	-0.718	-0.641	-2.086	-0.355	-0.647	-0.638	-1.611
5 (Small Redemptions / Large Inflows)	-0.948	-0.729	-0.662	-2.262	-0.225	-0.737	-0.836	-1.840
5 - 1	0.301	0.303*	0.257	0.898**	0.269	-0.161	-0.165	-0.081
	(0.184)	(0.172)	(0.178)	(0.361)	(0.210)	(0.214)	(0.207)	(0.409)

**Appendix Table 4: Regressions of Benchmark-Adjusted Returns on Flow Correlation with Alternative Standard Errors**

The table shows return regressions analogous to those in Tables 5 and 6 using alternative methods for computing standard errors. The dependent variable is the benchmark-adjusted fund return measured over either a 3-day (t+1 to t+3) or 20-day (t+1 to t+20) horizon following day t. Columns (1) and (2) report regressions with fund fixed effects and Driscoll-Kraay standard errors. Columns (3) and (4) report regressions with fund and day fixed effects and standard errors clustered by fund. Columns (5) and (6) report regressions with fund and day fixed effects and standard errors clustered by fund and day. All regressions include the same control variables as in Tables 5 and 6. The sample consists of fund-days with net outflows over the 2007–2024 period. Standard errors are reported in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	XR <sub>t+1, t+3</sub> (1)	XR <sub>t+1, t+20</sub> (2)	XR <sub>t+1, t+3</sub> (3)	XR <sub>t+1, t+20</sub> (4)	XR <sub>t+1, t+3</sub> (5)	XR <sub>t+1, t+20</sub> (6)
ExpFlowCorr	0.004 (1.146)	0.010 (0.057)	-0.024 (0.731)	0.017 (0.039)	-0.024 (0.597)	0.017 (0.038)
UnexpFlowCorr	-0.480*** (0.173)	-0.012** (0.005)	-0.557*** (0.133)	-0.012*** (0.003)	-0.557*** (0.125)	-0.012*** (0.003)
Log (Fund Size)	-1.889*** (0.366)	-0.151*** (0.018)	-2.000*** (0.195)	-0.151*** (0.011)	-2.000*** (0.136)	-0.151*** (0.010)
Log (Fund Flow)	0.055 (0.150)	0.010* (0.006)	-0.032 (0.055)	0.003** (0.002)	-0.032 (0.040)	0.003** (0.001)
Expense Ratio	2.059 (2.035)	0.082 (0.117)	1.050 (0.871)	0.041 (0.055)	1.050 (0.830)	0.041 (0.054)
Turnover	-0.008 (0.074)	0.002 (0.003)	-0.010 (0.007)	0.002*** (0.001)	-0.010 (0.008)	0.002*** (0.001)
StdDev(Flow <sub>t-2,t</sub> )	0.002 (0.002)	0.000** (0.000)	-0.001 (0.001)	-0.000* (0.000)	-0.001 (0.001)	-0.000** (0.000)
Cash	-0.022 (0.037)	-0.002 (0.002)	-0.027 (0.033)	-0.002 (0.002)	-0.027 (0.028)	-0.002 (0.002)
Load	1.292 (0.806)	0.086* (0.044)	-0.176 (0.424)	-0.019 (0.026)	-0.176 (0.359)	-0.019 (0.026)
Constant	29.296*** (8.263)	2.527*** (0.422)	33.558*** (3.972)	2.647*** (0.207)	33.558*** (2.677)	2.647*** (0.188)
Fund FE and Cluster DK	YES	YES				
Fund and Day FE and Cluster Fund and Day			YES	YES		
Fund and Day FE and Cluster Fund					YES	YES
N	4,648,069	4,329,765	4,647,985	4,329,747	4,647,985	4,329,747
Adj-R <sup>2</sup>	0.000	0.002	0.104	0.115	0.104	0.115

**Table A5: Alternative Measures of Flow Diversification**

This table reports robustness tests using alternative measures of flow diversification. The dependent variable is the cumulative 3-day (t+1, t+3) benchmark-adjusted returns. Panel A uses a flow correlation measure constructed with equal weights across share classes and based on proportional (rather than dollar) flows. Panel B uses the absolute value of flow imbalance ( $|FlowImb|$ ).  $FlowImb$  is defined as the sum of signed flows across the K share classes, scaled by the sum of their absolute values. In both panels, the decomposition of  $FlowCorr$  and  $|FlowImb|$  uses a backward-looking moving average filter of order  $q=60$  (t-61, t-1), corresponding to one quarter. The sample consists of fund-days with net outflows over the 2007-2024 period. Robust standard errors, clustered at the fund-quarter level are in parentheses. Statistical significance is denoted by \*\*\*, \*\*, and \* at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: $FlowCorr$ constructed from equal weights				Panel B: $ FlowImb $			
$FlowCorr$ or $ FlowImb $	-0.204*	-0.351***			-1.291***	-0.860***		
	(0.121)	(0.115)			(0.236)	(0.225)		
$ExpFlowCorr$ or $Exp FlowImb $			0.310	0.350			1.058	1.631**
			(0.661)	(0.637)			(0.793)	(0.759)
$UnexpFlowCorr$ or $Unexp FlowImb $			-0.237*	-0.395***			-1.524***	-1.110***
			(0.123)	(0.116)			(0.242)	(0.231)
Log (Fund Size)	-1.872***	-1.985***	-1.867***	-1.980***	-2.007***	-2.066***	-1.979***	-2.043***
	(0.122)	(0.119)	(0.122)	(0.119)	(0.125)	(0.122)	(0.126)	(0.122)
Log (Fund Flow)	0.036	-0.049	0.033	-0.052	0.161***	0.025	0.161***	0.026
	(0.040)	(0.038)	(0.041)	(0.038)	(0.049)	(0.046)	(0.049)	(0.046)
Expense Ratio	2.069***	1.051	2.062***	1.047	2.040***	1.038	2.023***	1.058
	(0.681)	(0.745)	(0.680)	(0.744)	(0.682)	(0.745)	(0.682)	(0.745)
Turnover	-0.008	-0.010	-0.008	-0.011	-0.008	-0.010	-0.008	-0.011
	(0.039)	(0.040)	(0.039)	(0.040)	(0.039)	(0.040)	(0.039)	(0.040)
StdDev ( $Flow_{t-2,t}$ )	0.002	-0.001	0.002	-0.001	0.002	-0.001	0.002	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Cash	-0.022	-0.027	-0.022	-0.027	-0.023	-0.027	-0.022	-0.027
	(0.024)	(0.023)	(0.024)	(0.023)	(0.024)	(0.023)	(0.024)	(0.023)
Load	1.294***	-0.177	1.296***	-0.174	1.278***	-0.179	1.289***	-0.169
	(0.263)	(0.276)	(0.262)	(0.276)	(0.263)	(0.276)	(0.263)	(0.276)
Constant	28.937***	33.245***	28.800***	33.081***	32.891***	35.663***	30.473***	33.178***
	(2.554)	(2.483)	(2.551)	(2.481)	(2.689)	(2.601)	(2.795)	(2.699)
Fund FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Day FE		Yes		Yes		Yes		Yes
N	4,648,066	4,647,985	4,648,066	4,647,985	4,648,122	4,648,041	4,648,122	4,648,041
Adj-R <sup>2</sup>	0.003	0.104	0.003	0.104	0.003	0.104	0.003	0.104