OFR WORKING PAPER
OFFICE OF FINANCIAL RESEARCH

17-07 | December 15, 2017
Revised June 2, 2020

Investor Concentration, Flows, and Cash Holdings: Evidence from Hedge Funds

Mathias S. Kruttli
Board of Governors of the Federal Reserve System
Oxford-Man Institute of Quantitative Finance
mathias.s.kruttli@frb.gov

Phillip J. Monin
Board of Governors of the Federal Reserve System
phillip.monin@frb.gov

Sumudu W. Watugala
Cornell University
sumudu@cornell.edu

Investor Concentration, Flows, and Cash Holdings: Evidence from Hedge Funds*

Mathias S. Kruttli, Phillip J. Monin, and Sumudu W. Watugala†

June 2020

Abstract

We show that when only a few investors contribute a substantial portion of a fund's equity, the probability of large liquidity-driven fund outflows increases because investors' idiosyncratic liquidity shocks are not diversified away. Using confidential regulatory filings, we find the five largest investors on average own 50% of a hedge fund. Consistent with our predictions, we confirm that high investor concentration hedge funds are more likely to experience large liquidity-driven outflows. Such funds hold more precautionary cash and implement other portfolio adjustments that help absorb outflows, but result in lower risk-adjusted returns. We find no evidence that hedge funds with a concentrated investor base impose longer share restrictions.

JEL classification: G11, G20, G23.

Keywords: Hedge funds, investor concentration, flows, precautionary cash.

*We thank Vikas Agarwal, Turan Bali, Daniel Barth, Murillo Campello, Grant Farnsworth, Mark Flood, Byoung Hyoun-Hwang, David Johnson, Andrew Karolyi, Dasol Kim, Andrew Lo, Jose Martinez, David Ng, Andrew Patton, Lubomir Petrasek, Tarun Ramadorai, David Thesmar, Stathis Tompaidis, Kelsey Wei, Russ Wermers, Scott Yonker, Peyton Young and seminar participants at the Office of Financial Research, Federal Reserve Board of Governors, SEA Annual Meeting 2017, Federal Reserve Bank of Boston, MFA Annual Meeting 2018, CICF 2018, and SIAM Annual Meeting 2018 for helpful comments. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Board of Governors of the Federal Reserve System, its research staff, or the Office of Financial Research, U.S. Department of the Treasury.

†Kruttli: The Board of Governors of the Federal Reserve System and Oxford-Man Institute of Quantitative Finance. Phone: (202) 530-6214. Email: mathias.s.kruttli@frb.gov. Monin: The Board of Governors of the Federal Reserve System. Phone: (202) 365-1866. Email: phillip.monin@frb.gov. Watugala: Cornell University. Phone: (617) 858-6178. Email: sumudu@cornell.edu. This paper and an earlier version were produced while Monin and Watugala were at the Office of Financial Research.

SCOPE OF RESEARCH

The research and analysis conducted in this paper is in accordance with the mandate set forth in the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 that the Office of Financial Research study and monitor potential threats to financial stability and issues related to systemic risk in U.S. financial markets.

1 Introduction

A large literature analyzes hedge funds' investment decisions and portfolio risk (see, for example, Fung and Hsieh (1997, 2001, 2004); Agarwal and Naik (2004); Patton and Ramadorai (2013)). However, unlike the risks inherent in hedge fund investments, the risks inherent in hedge fund investor compositions have received relatively little attention. This paper helps fill this gap. We use a novel dataset of hedge fund regulatory filings to investigate the risks to hedge funds from high investor concentration (IC), when a handful of large investors own a substantial portion of a fund's net asset value (NAV). In our data, 50% of a hedge fund's NAV is on average held by the largest five investors, suggesting that investor concentration risk is particularly salient for hedge funds relative to other asset managers.

We first illustrate in a simple theoretical framework that when a hedge fund's investor base is highly concentrated, it is fragile because the fund is not diversified against the effect of idiosyncratic liquidity shocks to individual investors. Consequently, the fund is more likely to face large "liquidity" outflows.[1] These liquidity outflows, unlike "fundamental" outflows, are orthogonal to the hedge fund's fundamentals like past performance or portfolio holdings. This mechanism is related to the stock price fragility studied in Greenwood and Thesmar (2011), who define an asset as fragile if it is "susceptible to non-fundamental shifts in demand" from investors and show that a stock can be fragile due to concentrated ownership.[2] Our analysis differs in that the asset in question, the hedge fund, can take into account the risk of non-fundamental shifts in investor demand and mitigate the fragility introduced by a high investor concentration. If and how hedge funds account for a concentrated investor base is important for hedge fund investors and for assessing financial stability risks that hedge funds could pose.[3]

A high-IC hedge fund faces a trade-off between the costs of large liquidity outflows and the costs of the portfolio adjustments that account for the outflow risk. Being forced to sell assets to generate cash and absorb large outflows can lead to substantial losses for a hedge fund in two ways. First, outflows can force a hedge fund to obtain cash by exiting an arbitrage trade prior to convergence, thus realizing losses. The convergence of such trades can fail to materialize for an extended period of time and the trades can even diverge further before converging, making arbitrage inherently risky if early exit could be necessary (see, for example, De Long, Shleifer, Summers, and Waldmann (1990); Shleifer and Vishny (1997); Hombert and Thesmar (2014)).[4] Second, for hedge

[1] Examples of such exogenous liquidity shocks to investors include an institutional investor such as a fund of hedge funds facing outflows or a high net worth investor experiencing sudden large personal expenditures.

[2] In a related paper, Ben-David, Franzoni, Moussawi, and Sedunov (2018) show that concentrated institutional ownership of stocks can lead to an increase in volatility and price inefficiency.

[3] Regulatory reforms that take into consideration investor concentration risk in asset managers include the Investment Company Liquidity Risk Management Programs rule of the Securities and Exchange Commission that requires open-end mutual funds and exchange-traded funds to establish a liquidity risk management program, which, among other factors, must consider the "fund's shareholder ownership concentration," because the fund "could experience considerable cash outflows from redemptions by a single or small number of shareholders."[SEC (2016), p. 81]. The regulation does not cover hedge funds.

[4] Prime examples are the bet by Long-Term Capital Management on the convergence of the stock prices of Royal Dutch and Shell described in Lowenstein (2000) and Tiger Management's failed shorting of technology stocks during

funds that invest in illiquid assets, being forced to sell such assets quickly can lead to losses due to price impact.

Our framework predicts that a hedge fund with a high investor concentration will hold higher levels of precautionary cash to address the risk of outflows if the hedge fund is willing to pay the associated liquidity premium. Such precautionary cash holdings allow the hedge fund to accommodate large outflows and avoid losses from forced asset sales. We test these hypotheses empirically: (i) High-IC funds have a higher likelihood of liquidity outflows (i.e., flows orthogonal to a hedge fund's performance); (ii) High-IC funds hold more precautionary cash and make other portfolio adjustments to account for the higher likelihood of large outflows due to idiosyncratic investor shocks; (iii) These adjustments are costly and result in lower risk-adjusted returns. We find robust evidence for all three predictions. We further examine the frictions that can cause these effects to persist and rule out alternative explanations.

To measure the investor concentration of a hedge fund, we use confidential Form PF filings data reported to the Securities and Exchange Commission (SEC) starting in 2012. Large hedge funds report the proportion of the fund's equity owned by the top five investors. This "five-investor concentration ratio" of a hedge fund's investor base is on average 50% in our sample. Our sample of large hedge funds, which comprises \$1.8 trillion in aggregate NAV at the end of 2017, is well suited for our analysis.[5] The Long-Term Capital Management crisis of 1998 has shown the risks that a large hedge fund can pose to financial stability. Additionally, because Form PF filing is mandatory, we avoid self-selection and reporting issues that are known to exist in commercial hedge fund databases (see, for example, Bollen and Pool (2008, 2009) and Patton, Ramadorai, and Streatfield (2015)).[6]

To test the first prediction of our framework, we separate the liquidity outflows from the total net flows that we estimate from the data and use a logistic regression model to test if lagged investor concentration predicts a higher probability of large liquidity outflows in the next quarter. Our results support this hypothesis. High-IC hedge funds are 1.5-2.0 times more likely to experience large liquidity outflows.

High-IC hedge funds do indeed hold more cash relative to low-IC hedge funds, which supports the second prediction of our framework. The differences are economically significant. A one standard deviation (22 percentage points) increase in investor concentration is associated with a 3 percentage point increase in the hedge fund's cash relative to the hedge fund's NAV. This increase is a substantial fraction of the average and median cash holdings of 15% and 7% of NAV, respectively.

While holding precautionary cash helps hedge funds absorb outflows without selling assets, holding cash is also costly. We expect that the larger portfolio share of precautionary cash for high-IC hedge funds, together with the reduction in long-term arbitrage trades and the larger share

the dot-com boom (Brunnermeier and Nagel (2004)).

[5]The total NAV of all the hedge funds that file Form PF, including smaller hedge funds that file a pared down version of the form at an annual frequency, was \$3.9 trillion at the end of 2017: https://www.sec.gov/divisions/investment/private-funds-statistics.shtml.

[6]A comparison of the data from Form PF and the Thomson Reuters Lipper TASS Database (a commercially available hedge fund database) can be found in the Online Appendix.

of liquid assets holdings, have a negative effect on the funds' risk-adjusted returns. Confirming this hypothesis, we find that the risk-adjusted returns of high-IC hedge funds, estimated based on the Fung and Hsieh (2004) seven factor model, are significantly lower than those of low-IC hedge funds. This effect is economically significant: a one standard deviation increase in investor concentration is associated with a decrease in a hedge fund's annualized unlevered and levered risk-adjusted returns of around 110 and 130 basis points, respectively.

Our results are based on cross-sectional variation across hedge funds, rather than within-fund variation, because the investor concentration measure fluctuates little for a hedge fund through time. However, we also consider two portfolio adjustments in addition to holding more cash and confirm our findings. First, we find that high-IC hedge funds refrain from investing in risky arbitrage trades where the mispricing can worsen further before converging to the fundamental value of the asset. The volatility of long-term hedge fund returns acts as the proxy for risky arbitrages trades as in Hombert and Thesmar (2014). Second, we document that high-IC hedge funds generally invest in more liquid assets in addition to holding more cash, as these assets can be sold with little price impact in the case of large liquidity outflows.

We include strategy fixed effects in all of our regression specifications. Further, our results are robust to the inclusion of the financing constraints, leverage, size, flows, performance, and manager stake of a hedge fund as controls in our regression specifications. Our results hold both for subsamples of hedge funds for which the majority of investors are institutional and for hedge funds for which the majority of investors are individuals. Also, the documented effects of investor concentration are not driven by hedge funds with a very small total number of investors such as certain family offices.

High-IC hedge funds could also use gates and longer share restrictions to help mitigate the risk of liquidity outflows due to a high investor concentration. However, we find that gates and share restrictions are surprisingly not correlated with investor concentration.[7] A potential explanation for this finding is that share restrictions and gates are regulated by the limited partnership agreement that is set at the inception of a hedge fund and needs the investors' approval to be changed (Agarwal, Daniel, and Naik, 2009). When a hedge fund's investor concentration increases (for instance, because a large investor decides to invest in the fund) the hedge fund would likely face resistance from its existing investors to imposing longer share restrictions. The fund's investors would be unwilling to accept that their investments become more illiquid. In contrast, the portfolio adjustments we document do not need investor approval and are more difficult for hedge fund investors to observe than changes in share restrictions and gates.

The lower risk-adjusted returns of high-IC hedge funds raises the question of why an investor would invest in a hedge fund with a concentrated investor base. There are several possible explanations for this empirical observation. First, it is possible that investor awareness regarding this issue is low.[8] Our findings suggest that this inattention is costly. Second, there are factors that could

[7]Share restrictions include lock-ups, redemption and redemption notice periods.

[8]A hedge fund investor does not have a large panel dataset (as used in this paper) available to analyze the impact of investor concentration on risk-adjusted returns.

make it optimal for an investor to invest in a high-IC hedge fund despite the lower risk-adjusted returns. For example, the presence of a large investor could be a positive signal about the quality of the fund when assuming that the large investor conducted proper due diligence before investing.

To our knowledge, this is the first paper that investigates the effect of investor concentration on *liquidity* outflows. An emerging literature has obtained mixed findings on how the presence of large investors, proxied by institutional ownership, affects *fundamental* outflows of asset managers. For example, Chen, Goldstein, and Jiang (2010) and Goldstein, Jiang, and Ng (2017) find that institutional investors of a mutual fund that holds an illiquid portfolio are *less* likely to run on the fund following a poor performance than retail investors (the flow-performance sensitivity is weaker for institutional investors), and conclude this is because institutional investors internalize the price impact of their redemptions. However, in contrast to these findings on mutual funds, Schmidt, Timmermann, and Wermers (2016) report that large institutional investors exhibit greater flow-performance sensitivity and are *more* likely to run on money market funds holding illiquid assets than smaller institutional or retail investors, and conclude this is likely because institutional investors have more resources to monitor their investments. For hedge funds, we find no evidence that investor concentration affects fundamental flows or the flow-performance relationship. A likely reason behind the contrasting findings in these three studies is that the analyzed asset manager types differ along a range of dimensions.[9]

The remainder of the paper has the following structure. The related literature is discussed in Section 2. Section 3 presents the data and summary statistics. Section 4 reports the results from our empirical analysis, and Section 5 concludes.

2 Related literature

Hedge funds are often thought to be arbitrageurs that contribute to market efficiency (see, for example, Hombert and Thesmar (2014)). Limits to arbitrage theories predict that mispricing can persist because arbitrageurs are financially constrained (see, for example, Shleifer and Vishny (1997); Gromb and Vayanos (2002); Brunnermeier and Pedersen (2009)). We show that the composition of the investor base can also impose constraints on the ability of a hedge fund to correct mispricing, as a high-IC hedge fund has to hold more precautionary cash and refrain from risky arbitrage trades to avoid realizing losses should a large liquidity outflow occur that forces an early exit from such a trade. Further, forced asset sales of hedge funds can also pose a risk to financial stability. An individual hedge fund selling assets because of outflows can have a contagious effect on other hedge funds and a widespread impact on the stability of asset markets. Hedge funds often have large overlaps in their portfolios, and the sales of one hedge fund can depress asset prices and lead to losses for other hedge funds with similar portfolios.[10]

[9] In particular, hedge funds, unlike mutual funds or money market funds, adjust for the illiquidity of their holdings through longer share restrictions (see Aragon (2007) and Agarwal, Daniel, and Naik (2009)), have no retail investors but only institutional investors and high net-worth individuals, and do not disclose their portfolio holdings to investors.

[10] An example of overlapping portfolios causing contagious losses in the hedge fund industry is the "quant meltdown" discussed by Khandani and Lo (2011). They show that several equity long-short hedge funds experienced substantial

This paper contributes to the hedge fund literature by adding to our understanding of investor concentration as a novel source of risk for hedge funds. There is a large literature that documents how hedge funds are exposed to systematic risk—proxied by equity, bond, and option factors.[11] In contrast, we analyze the idiosyncratic risk posed by the hedge fund's investor composition. Investor composition has received little attention by the literature with the exception of the hedge fund manager's stake (see, for example, Ackermann, McEnally, and Ravenscraft (1999); Agarwal, Daniel, and Naik (2009); and Gupta and Sachdeva (2017)). Our paper presents a novel mechanism that contributes to a literature on how portfolio allocations of asset managers are affected by investor flows (see, for example, Chordia (1996); Teo (2011); Ben-David, Franzoni, and Moussawi (2012); Agarwal, Aragon, and Shi (2018)) and the impact of such portfolio allocations on performance (see, for example, Edelen (1999); Aragon (2007); Agarwal, Daniel, and Naik (2009)). Further, by showing that investor concentration can have an effect on the cash holdings and portfolio liquidity of hedge funds, our paper contributes to the literature examining hedge funds and their portfolio liquidity (see, for example, Getmansky, Lo, and Makarov (2004); Sadka (2010); Jylha, Rinne, and Suominen (2014); Aiken, Clifford, and Ellis (2015); Kruttli, Patton, and Ramadorai (2015); Aragon, Ergun, Getmansky, and Girardi (2017); Barth and Monin (2018)).

A large literature looks at the flow-performance relationship of mutual funds and hedge funds (see, for example, Chevalier and Ellison (1997); Sirri and Tufano (1998); Lynch and Musto (2003); Chen, Goldstein, and Jiang (2010); Li, Zhang, and Zhao (2011); Christoffersen, Musto, and Wermers (2014); Getmansky, Liang, Schwarz, and Wermers (2015); and Goldstein, Jiang, and Ng (2017)). Generally, this literature investigates whether performance predicts investor flows, that is, *fundamental* flows. Our investor concentration mechanism is distinct because the outflows considered in this paper are driven by *liquidity* flows that are orthogonal to hedge fund performance. Even if a hedge fund is performing well, investors can experience idiosyncratic liquidity shocks for exogenous reasons and redeem their investments.

3 Data and summary statistics

The primary source for our empirical analysis is data from the SEC's Form PF, which was adopted as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010. Form PF is filed by investment advisers that are registered with the SEC and manage at least US$150 million in private funds, such as hedge funds and private equity funds. Small private fund advisers file annually, while large advisers file quarterly and are required to report more detailed information

losses in August 2007, and these correlated losses were likely triggered by one hedge fund or proprietary trading desk unwinding its portfolio. Further, Boyson, Stahel, and Stulz (2010) find evidence of substantial hedge fund contagion in response to liquidity shocks. Our findings suggest hedge funds take investor concentration risk into account and hold more cash and liquid assets, which mitigate the risk of forced asset sales and contagious losses. This result is important for policymakers wary of the potential of hedge funds to spark widespread market instability.

[11] See, for example, Fung and Hsieh (1997, 2001, 2004); Agarwal and Naik (2004); Bollen and Whaley (2009); Bali, Brown, and Caglayan (2012); Patton and Ramadorai (2013); Buraschi, Kosowski, and Trojani (2014).

on their "Qualifying Hedge Funds."[12] A Qualifying Hedge Fund has a NAV of at least US$500 million. We only use data on Qualifying Hedge Funds because their filings are quarterly and certain variables crucial to our analysis, particularly those relating to cash holdings, are only reported by these funds. Details on the sample construction are included in Appendix C.1.

Our sample period runs from 2012:Q4 to 2017:Q4, inclusive. Table C.1 of Appendix C provides a summary of the variable definitions and data sources for quick reference. Table 1 Panel A presents the number of observations (hedge fund-quarters), average, standard deviation, and 10$^{\text{th}}$, 50$^{\text{th}}$, and 90$^{\text{th}}$ percentiles for several quarterly variables capturing key characteristics of the hedge funds in our analysis.

Our main variable of interest, IC_{it}, is based on the hedge funds' reported five-investor concentration ratio, that is, how much of the reporting fund's equity is held by the five investors with the largest investments in the fund.[13] The average five-investor concentration ratio across all of the hedge funds in our sample is 50%, and the median is 47%.[14] These values show that hedge funds often depend on a handful of large investors. The 10$^{\text{th}}$ and 90$^{\text{th}}$ percentiles are 24% and 84%, respectively, which shows the measure varies considerably across the hedge funds in our sample. In the Online Appendix, we analyze how the five-investor concentration ratio relates to the Herfindahl-Hirschman Index.

The other two variables of interest are flows (F_{it}) and cash ($Cash_{it}/NAV_{it}$). We estimate the flows for quarter t and hedge fund i as $F_{it} = (NAV_{it} - NAV_{it-1} \times (1 + r_{it}))/(NAV_{it-1})$, where r_{it} is the return net-of-fees. The flows are winsorized at the 5% level as in Ben-David, Franzoni, and Moussawi (2012). Hedge funds report their "unencumbered cash" in Form PF.[15] The average unencumbered cash as a percent of NAV is 15.3%, with a standard deviation of 20.5% and median of 6.8%. In later discussions, we will refer to unencumbered cash divided by NAV simply as cash.

In Form PF, hedge funds report the percentage of the portfolio, excluding cash, that can be liquidated within particular time horizons (within <1, 2-7, 8-30, 31-90, 91-180, 181-365, and >365 days) using the market liquidity in a given reporting period. We compute the weighted average to obtain a measure of portfolio illiquidity ($PortIlliq_{it}$). The average portfolio illiquidity measure

[12] Large hedge fund advisers are defined as those with at least US$1.5 billion in total regulatory assets under management aggregated across all of their hedge funds.

[13] Form PF Question 15 asks for the "beneficial owners," referring to the investors, not the advisers or managers, of the hedge fund.

[14] Smaller hedge funds that file an abridged version of Form PF at an annual frequency also have a high average investor concentration measure, indicating that investor concentration risk is not simply restricted to large hedge funds. The abridged version reported by annual filers does not include some variables crucial to our analysis (e.g., measures of cash, liquidity and share restrictions), so only hedge funds that are required to file quarterly are included in our baseline analysis.

[15] Unencumbered cash is defined in Form PF: Glossary of Terms as the fund's cash and cash equivalents plus the value of overnight repos used for liquidity management where the assets purchased are U.S. treasury securities or agency securities minus the sum of the following (without duplication): (i) cash and cash equivalents transferred to a collateral taker pursuant to a title transfer arrangement; and (ii) cash and cash equivalents subject to a security interest, lien or other encumbrance (this could include cash and cash equivalents in an account subject to a control agreement).[Pg. 10] Cash equivalents are defined in Form PF as (i) bank deposits, certificates of deposits, bankers acceptances and similar bank instruments held for investment purposes; (ii) the net cash surrender value of insurance policy; (iii) investments in money market funds; (iv) US treasury securities (including derivatives); (v) agency securities (including derivatives); and (vi) any certificate of deposit for any of the foregoing.

in our sample is 52 days and the median is 14 days. Hedge funds are also required to report restrictions on investor withdrawals locked for particular time horizons (for the same horizons as for portfolio illiquidity). We compute the weighted average as a measure of share restrictions for each fund ($ShareRes_{it}$). The average share restriction is 167 days with a median of 147 days.

The portfolio illiquidity measure, with an average of 53 days, is substantially lower than the share restrictions measure, with an average of 167 days. This illiquidity gap (see Aragon, Ergun, Getmansky, and Girardi (2017) and Agarwal, Aragon, and Shi (2018)) can be partially explained by Form PF asking hedge funds to assess how long it would take to liquidate an asset under "current market conditions," that is, the market conditions in the quarter for which the hedge fund is filing the form, and our sample covers a period of relatively high market liquidity. The fact that portfolio illiquidity is substantially lower than share restrictions suggests that the average high-IC hedge fund is relatively unconcerned that forced selling of illiquid assets would generate losses because of price impact. However, as discussed previously, large investor outflows can also force a hedge fund to exit an arbitrage trade early and realize losses, and such non-convergence can occur even in highly liquid assets.[16]

The final three variables presented in Panel A are manager stake ($MgrStake_{it}$), number of investors ($NumInvestors_{it}$), and minimum investment ($MinInv_{it}$). These three variables are obtained through the matching of Form PF data with the publicly available Form ADV filings of the hedge funds.[17] We include in our sample only matched hedge funds with more than five investors and a manager stake of no more than 50%. We apply these filters for two reasons. First, the IC variable is a five-investor concentration measure and fails to capture variation in the concentration of the investor base for hedge funds with five or fewer investors. Second, the filters avoid including family offices and predominantly manager-owned funds in our analysis.[18] In a family office, the investors of the fund know each other and can smooth out liquidity shocks amongst each other. Also, the hedge fund manager likely knows the investors personally, which reduces the asymmetric information about liquidity outflows between investors and the hedge fund manager. In this case, the hedge fund manager will likely learn about the possibility of large outflows long before the redemption request is filed, which mitigates the need for holding precautionary cash. If the manager owns the majority of the hedge fund, then the asymmetric information between the hedge fund manager and the investor base is also clearly limited. Therefore, the mechanism of how IC affects the probability of large liquidity outflows is likely not applicable to these hedge funds. In our sample, minimum investment has a mean of of US$3.8 million and median of US$1 million.

In Table 1 Panel B, we show the average of each variable separately for six hedge fund invest-

[16]Prime examples of convergence trades in liquid assets that failed to converge for an extended period of time are the bet by Long-Term Capital Management on the convergence of the stock prices of Royal Dutch and Shell described in Lowenstein (2000) and Tiger Management's failed shorting of technology stocks during the dot-com boom (Brunnermeier and Nagel (2004)).

[17]We are able to merge 98.5% of the fund-date observations in Form PF to Form ADV. Schedule D, Section 7.B.(1) of Form ADV asks advisers to report on the number of beneficial owners (number of investors) of the fund. In the same section, advisers are asked to report the minimum investment level and the percentage of the fund beneficially owned by the adviser or its related persons (manager stake).

[18]Our results are robust to including these observations in our sample.

ment strategies (Credit, Equity, Event Driven, Macro, Multi-strategy, and Relative Value) and an "Other" category. We establish a single broad strategy category for each hedge fund and reporting date as described in the Appendix C.2. The most fund-quarter observations are for Equity hedge funds with 5,993. The second largest strategy is Multi-strategy hedge funds with 2,702 observations.

Panel B shows that there is little variation across strategies in average IC, which clusters around 50%. Event Driven hedge funds have the lowest average IC with 46% and Relative Value hedge funds have the highest average IC with 59%. In contrast, there are large differences across strategies for share restrictions. Macro hedge funds have average share restrictions of 95 days, but the share restrictions of Event Driven hedge funds are on average 247 days. Further, for cash holdings and portfolio illiquidity, the differences across strategies are also large. For cash, the range is from 10% (Equity) to 43% (Macro). For portfolio illiquidity, Macro hedge funds have the most liquid portfolios (11 days), and "Other" hedge funds have the least liquid portfolios (112 days).

We show the distribution of IC across investment strategies in Figure 1, which depicts the number of fund-quarter observations in each IC tercile for each strategy, where the lowest IC observations are in the first tercile. IC shows little correlation with a specific strategy type, as most strategies are equally distributed across the three terciles. The one exception is the Relative Value strategy, which is slightly skewed toward the high-IC tercile.

In the Online Appendix, we compare the size, net-of-fees returns, and flows of hedge funds from Form PF and from the Thomson Reuters Lipper TASS Database (TASS), as the TASS database and other commercial hedge fund databases have been used extensively in the hedge fund literature. The average measures of the two hedge fund datasets correlate strongly over our sample period.

4 Empirical results

In this section, we empirically test the three predictions of the simple theoretical framework presented in Appendix A. The three predictions are supported by our results. First, our analysis shows that hedge funds with a high IC have a greater probability of large outflows than hedge funds with a low IC. Second, we find that high-IC hedge funds maintain a larger cash share in their portfolios. Third, our results show that high-IC hedge funds generate lower risk-adjusted returns. Further, we examine why high-IC hedge funds do not use other methods to account for a high IC and rule out alternative mechanisms.

4.1 Investor concentration and large outflows

The first prediction that we test is whether high-IC hedge funds have a greater probability of large liquidity outflows than low-IC hedge funds. Our framework in Appendix A distinguishes between flows due to liquidity shocks to investors (F_{it}^L or liquidity flows) and flows due to hedge fund fundamentals (F_{it}^F or fundamental flows), with total flows $F_{it} = F_{it}^F + F_{it}^L$. IC affects flows due to investor liquidity shocks and the likelihood of large liquidity outflows is expected to be higher when IC is high. We first estimate the liquidity flows necessary for our empirical analysis. We use a methodology similar to Coval and Stafford (2007), who regress flows on lagged performance and lagged flows to estimate predicted flows of mutual funds.[19] We use these predicted flows as a measure of flows due to fundamentals.

Unlike other asset managers, hedge funds have share restrictions that affect how flows are predicted by lagged performance and lagged flows (see Getmansky, Liang, Schwarz, and Wermers (2015)). Therefore, we add share restrictions to the regression estimated by Coval and Stafford (2007). We estimate the following panel regression model with the Fama and MacBeth (1973) methodology:

$$\begin{aligned} F_{it} =& a + \sum_{p=1}^{P} (b_p F_{it-p} + c_p r_{it-p} + d_p F_{it-p} * ShareRes_{it-p} \\ &+ e_p r_{it-p} * ShareRes_{it-p} + g_p * ShareRes_{it-p}) + \epsilon_{it}. \end{aligned} \tag{1}$$

The frequency is quarterly, and variations of the model are estimated with values for P of 1 or 2, i.e., including one or two lags, respectively. The estimated liquidity flows of a hedge fund, $\hat{F}^L_{it}$, is the difference between the predicted and actual flows.

In a second stage, we estimate a logit model to test whether IC predicts a higher probability of large liquidity outflows.[20] The logit model is given by

$$P(\hat{F}^L_{it} \leq \tau | IC_{it-1}, W_{it-1}) = L(\psi + \gamma IC_{it-1} + \phi W_{it-1}), \text{ where } L(z) = \frac{\exp(z)}{1+\exp(z)}. \tag{2}$$

For τ, we use two liquidity flow thresholds: -15% and -20% of a hedge fund's NAV. We consider these thresholds because outflows of that magnitude are considerable in our sample. The respective unconditional probabilities for outflows being greater than 15% and 20% are 4% and 2%, respectively. W_{it-1} is a set of control variables that may affect flows. Further, we include quarter and strategy fixed effects. The standard errors are clustered by quarter.

The results of the first stage regression in equation (1) are shown in Appendix B. The estimation results for the logit model in equation (2) are in Table 2. Panel A shows the results for the dependent variable being an indicator variable that takes the value 1 if outflows are greater in magnitude than 20% of a hedge fund's NAV. The coefficient estimates of IC are greater than 1 and strongly significant for the specifications without and with controls and when one lag ($P = 1$) or two lags ($P = 2$) are used in the first stage regression. Because the reported coefficient estimates are odds ratios, a coefficient estimate greater than 1 implies that the probability of outflows being greater than or equal to 20% increases when IC increases. The coefficient estimates are around 1.01, which means that if IC increases by 1 percentage point, then the odds ratio increases by 1%. When looking at the coefficient estimates of the IC terciles, we can see that the odds of a hedge fund in the third tercile to experience outflows greater than or equal to 20% are 1.5-2 times higher

[19] Jotikasthira, Lundblad, and Ramadorai (2012) also use this methodology to estimate predicted flows based on lagged performance and lagged flows.

[20] Using a probit model obtains results that are qualitatively the same.

than for a hedge fund in the first tercile. Similar results are obtained when setting the threshold for the outflows to greater than or equal to 15%. The coefficient estimates of IC or higher IC terciles remain greater than 1 and significant under the different specifications.

These results confirm our hypothesis that high-IC hedge funds face a significantly higher probability of large liquidity outflows and motivate our next analysis.

4.2 Investor concentration and cash

We showed in the previous section that high-IC hedge funds have a greater probability of large outflows. Next, we test our second hypothesis that a hedge fund with a high IC accounts for the increased probability of sudden large outflows by holding more cash. Holding insufficient cash can lead to hedge funds being forced to sell assets and incur losses because an arbitrage trade has to be exited early and, for hedge funds that invest in illiquid assets, because of price impact when illiquid assets have to be sold quickly. Importantly, our simple framework predicts that a manager of a high-IC hedge fund does not necessarily need to experience large outflows to hold precautionary cash, because the manager can rationally anticipate that a concentrated investor base increases the probability of large outflows.

We estimate a panel model that has cash normalized by NAV, $Cash_{it}/NAV_{it}$, as the dependent variable:

$$\frac{Cash_{it}}{NAV_{it}} = \psi + \gamma IC_{it} + \phi Z_{it} + \epsilon_{it}, \tag{3}$$

where cash is as defined in Section 3. The control variables are included in the column vector Z_{it}. We use the control variables size, flow, share restrictions, financing duration, leverage, and manager stake: $log(NAV_{it})$, F_{it}, $ShareRes_{it}$, $FinDur_{it}$, $Leverage_{it}$, and $MgrStake_{it}$, respectively. If hedge funds take IC into account when making portfolio allocation decisions and hold more cash, we would expect γ to be significant and positive.

The estimates of the panel regression are shown in Table 3. Because of the persistence of our dependent variable, that is, $Cash_{it}/NAV_{it}$, we account for potential serial correlation in the error terms by clustering by hedge fund in addition to clustering by time (see Petersen (2009) and Thompson (2011)). We also include strategy fixed effects and quarter fixed effects. The results strongly support our hypothesis. We find that the coefficient estimate of IC, γ, is positive and strongly significant with and without the control variables included. Consequently, the results are in line with the mechanism that high-IC hedge funds hold more cash than low-IC hedge funds to absorb large outflows that are more likely to occur because of a concentrated investor base.

These results are economically significant. The γ estimate is around 0.14 when including control variables. This coefficient estimate implies that a one standard deviation (22 percentage points) increase in IC is associated with an increase of 3.1 percentage points in the cash holdings normalized by NAV. This increase is substantial considering that the average cash holdings are 15.3% and the median is 6.8%, as shown in Table 1.

The control variables have coefficient estimates consistent with existing research. For three of the control variables, the coefficient estimates are highly significant. Size has a positive coefficient estimate, which is likely a result of larger hedge funds generally investing in more liquid assets. This finding is not surprising because trading strategies in illiquid assets are difficult to scale due to trading costs and price impact (see, for example, Fung, Hsieh, Naik, and Ramadorai (2008)). Also, holding cash is expensive, and larger hedge funds may be better able to afford the costs of holding large cash positions. The coefficient estimate of share restrictions is negative, which shows that hedge funds that grant investors less favorable (longer) redemption terms hold less cash. This result is in line with the finding of Aragon (2007) and Agarwal, Daniel, and Naik (2009), who show that longer share restrictions lead to a hedge fund holding a more illiquid portfolio. Interestingly, the economic significance of share restrictions is similar in magnitude to IC. A one standard deviation decrease in the share restrictions (121 days) leads to an increase of cash holdings normalized by NAV of 3.1 percentage points (compared with an increase of also 3.1 percentage points when IC increases by one standard deviation). Further, leverage has a positive coefficient estimate, suggesting that highly leveraged hedge funds hold more cash. This result is in line with highly leveraged hedge funds being more vulnerable to an increase in funding constraints and holding more cash as a precautionary measure. A one standard deviation increase in leverage (2.1) leads to an increase in cash holdings normalized by NAV of 2.5 percentage points, which is again comparable to the effect of IC on cash.

To test whether there is non-monotonicity in the effect of IC on cash, we estimate a panel regression with hedge funds being sorted into three terciles based on IC in each quarter:

$$\frac{Cash_{it}}{NAV_{it}} = \psi + \sum_{n=2}^{3} I_{i\in n_t}\gamma_n + \phi Z_{it} + \epsilon_{it}. \tag{4}$$

The third tercile corresponds to the hedge funds with the highest IC. The estimates of γ_2 and γ_3 should be positive and significant, with γ_2 smaller than γ_3 if high-IC hedge funds hold more cash. Columns (3), (4), (7), and (8) of Table 3 show that the γ_2 and γ_3 estimates are indeed positive and significant for the specifications where the control variables are included. The estimates are robust to including quarter fixed effects and strategy fixed effects. Further, the γ_2 estimate is significantly lower than γ_3, which indicates that the IC effect on cash is stronger for hedge funds with very high IC.

The results in Table 3 suggest that hedge funds account for a high IC by holding more cash, and the magnitude of the changes in cash are economically significant. These results raise the question of whether the increase in cash is "sufficient" given the increase in the probability of large liquidity outflows shown in Table 2 for high-IC hedge funds. While the answer to this question is to some degree subjective, we assume that if hedge funds adjust cash to fully account for IC risk, then the probability of the liquidity outflows exceeding cash holdings in a given quarter would be the same for low and high-IC hedge funds.

We sort each hedge fund in our sample that have IC, liquidity flow, and cash data for at least four quarters into a quintile based on average IC. The hedge funds in the first quintile have the lowest IC values. Then, for each quintile, the median of quarterly liquidty flows, standard deviation of liquidity flows, and cash are computed. The liquidity flows are computed with the model given in equation (1) and $P = 1$. Assuming that the liquidity flows are normally distributed, we compute the probability that the liquidity outflows exceed cash within a quarter based on the median cash and liquidity flow values for each IC quintile. We also compute bootstrapped standard errors for these probabilities.

Plot (a) of Figure 2 depicts the probability of liquidity outflows exceeding cash in a quarter for each IC quintile with 95% confidence intervals. The figure shows that the probability is around 10%. There is only little variation across the quintiles, and for none of the quintiles is the probability significantly different from any of the other quintiles. This result indicates that hedge funds adjust for a high IC by holding more cash such that the probability of liquidity outflows exceeding cash is unchanged.

In contrast, Plot (b) of Figure 2 presents the counterfactual scenario. Here, the probabilities of liquidity outflows exceeding cash when we set the median cash level for each IC quintile equal to the median cash level of the first quintile. We can see that the probabilities for quintiles three, four, and five are significantly higher than for the first quintile. This result suggests that high-IC hedge funds avoid this significant increase in the likelihood of liquidity outflows exceeding cash by their adjustment in precautionary cash levels.

4.2.1 Investor type

So far, our analysis has focused on differences in investor concentration across hedge funds without differentiating among investor characteristics. In this section, we investigate whether there is variation in the effect of IC on cash based on the predominant investor type invested in a hedge fund. Question 16 of Form PF asks for the percentage of the reporting hedge fund's equity held by individual versus institutional investors, where individual investors are generally high net worth individuals.[21] The summary stats for these data are reported in Table B.1 of Appendix B. To assess whether the effect of IC on cash is robust to differences in the investor composition type, we estimate the model in equation (3) for subsamples of hedge funds where individuals own greater than or equal to 50% and 25% of the hedge fund's equity. We also estimate the same model for the complimentary subsamples of hedge funds where institutions own greater than 50% and 75% of the hedge fund's equity (where individuals own less than 50% and 25% of the fund, respectively).

The results are reported in Table 4. The first two columns show the regression model estimates for hedge funds with an individual investor share of greater than or equal to 50%, and 25%, respectively. The subsequent columns present the model estimates for the complementary subsamples. The effect of IC is robust to the sample split. The coefficient estimates on IC are positive and

[21] Individual investors are split into US persons and non-US persons. Institutional investors are split into: broker-dealers, insurance companies, investment companies registered with the SEC, private funds, non-profits, pension plans (excluding governmental pension plans), banking or thrift institutions (proprietary), state or municipal government entities (excluding governmental pension plans), state or municipal governmental pension plans, sovereign wealth funds and foreign official institutions, unknown non-US persons, and others.

significant across the four subsamples. The lower significance of the IC coefficient estimate for the subsample of hedge funds with individual investors owning 50% or more of the equity is likely because of the reduced power due to the reduced sample size. The average individual investor share is 18% with a standard deviation of 22%, and the 95$^{\text{th}}$ percentile is 66%. Therefore, there are only a few hedge funds for which individuals, as opposed to institutions, hold 50% or more of the equity.

The coefficient estimates in the first two columns are not significantly different from the coefficient estimates of the complementary subsamples in the last two columns. These results indicate that the effect of IC on cash discussed applies to hedge funds that are predominantly held by both individual and institutional investors. A further implication is that from a hedge fund's perspective, the risk that one of its investors suffers a liquidity shock is likely similar for individual and institutional investors.

4.3 Additional portfolio adjustments

The results in the previous section are in line with our hypothesis that high-IC hedge funds hold more cash to absorb potential large outflows. This prediction on the use of precautionary cash to absorb outflows due to idiosyncratic shocks to large investors is motivated by the theoretical framework described in Appendix A. There are other ways through which hedge funds can potentially adjust their portfolios to account for a high IC, in addition to holding more cash. In this section, we test for two additional portfolio adjustments with the caveat that these adjustments can only be measured through proxies, unlike cash, which we observe directly in Form PF.

In our framework, high-IC hedge funds hold more cash because they want to avoid a scenario where they have to exit an arbitrage trade early and realize losses. In addition to holding more cash, high-IC hedge funds may also reduce taking on risky arbitrage trades where the mispricing can worsen further before the trade yields a profit (see Shleifer and Vishny (1997)). To test this hypothesis, we test whether high-IC hedge funds have less volatile long-term returns. This test is based on Hombert and Thesmar (2014), who show that hedge funds with short share restrictions have less volatile returns, in line with these hedge funds refraining from risky arbitrage trades, because the funds are concerned about outflows. We estimate the regression model given by

$$|r^g_{it} - \hat{r}^g_i| = \psi + \gamma IC_{it} + \phi Z_{it} + \epsilon_{it}, \tag{5}$$

where $\hat{r}^g_i$ is the average gross return of hedge fund i over the sample period. The regression is estimated at a semi-annual and annual frequency. As discussed in Hombert and Thesmar (2014), longer return frequencies are required for this analysis, as monthly returns are likely smoothed (Getmansky, Lo, and Makarov (2004)). The control variables are again included in the column vector Z_{it}. The control variables are size, flow, share restrictions, financing duration, leverage, and manager stake. If high-IC hedge funds refrain from risky arbitrage trades, their return volatility would be lower and the γ estimate would be negative and significant.

The results are in line with our prediction and are shown in Table 5. The γ estimate is negative and significant for all the specifications. The estimates are also economically significant, with a one standard deviation (22 percentage point) increase in IC predicting an absolute return deviation from the mean return that is around 95 basis points lower at the annual return frequency. These results are in line with high-IC hedge funds refraining from taking on arbitrage trades with high volatility. Share restrictions, which are included as a control, have a coefficient estimate that is positive and significant. This estimate is consistent with the finding of Hombert and Thesmar (2014) that hedge funds with long share restrictions avoid risky arbitrage trades. Interestingly, the economic significance of share restrictions is close to IC. A one standard deviation decrease in the share restrictions (121 days) leads to decrease in the absolute return deviation from the mean return of around 85 basis points.

A second portfolio adjustment that can help a hedge fund account for the risks due to a high-IC, at least to some extent, is holding liquid assets in addition to holding cash. A low portfolio illiquidity can reduce price impact when assets have to be sold quickly due to outflows caused by an idiosyncratic shock to a large investor. However, unlike holding cash, a low portfolio illiquidity does little to prevent realizing losses on arbitrage trades. Arbitrage trades in liquid assets can also be subject to the risk that prices might diverge further before converging.[22]

Hedge funds report in Form PF the percentage of the portfolio (excluding cash) that can be liquidated within particular time horizons. We compute the average time in days that it takes a hedge fund to liquidate an asset in its portfolio, as described in Section 3, and use it as a measure of portfolio illiquidity. A drawback of this portfolio illiquidity measure compared with the cash measure used in the preceding analysis is that it depends on a hedge fund's subjective assessment of its portfolio liquidity, which might differ from actual portfolio liquidity and introduce measurement error.

We use portfolio illiquidity as the dependent variable in place of of cash and estimate the models given in equations (3) and (4). The results are reported in Table 6 and confirm our previous results. The coefficient estimates of IC are strongly significant and negative with and without controls. Accordingly, these results support our hypothesis that hedge funds with higher IC hold a more liquid portfolio to absorb potential idiosyncratic liquidity shocks to investors. The results with IC sorted into terciles also support the hypothesis that high-IC hedge funds hold more liquid portfolios. The results are again economically significant. The γ estimate is around -0.37 with control variables included. This coefficient estimate implies that a one standard deviation (22 percentage points) increase in the investor concentration of the hedge fund is associated with a decrease in the hedge fund's portfolio illiquidity by 8.1 days. Considering that the average and median portfolio illiquidity measures are 52.4 and 14.2 days, respectively, this decrease is substantial.

[22] For example, the hedge fund Tiger Management bet on falling technology stock prices, which were highly liquid, during the dot-com boom and did not survive this episode as described in Brunnermeier and Nagel (2004).

4.4 Implications for hedge fund investors

In the previous section, we show high IC is associated with higher levels of cash. Also, high-IC hedge funds tend to refrain from arbitrage trades that could lead to losses in the short run and hold a more liquid portfolio. On the one hand, from the perspective of hedge fund investors, this finding shows that hedge funds account for a risk that could potentially lead to substantial losses to the fund and ultimately its investors. On the other hand, our findings raise the question of whether the portfolio adjustments of high-IC hedge funds generate lower risk-adjusted returns. When high-IC hedge funds pay a liquidity premium for holding more cash and liquid assets and also refrain from riskier arbitrage trades that are lucrative in the long run but might lead to losses in the short run, then returns to investors are potentially affected.

To test this prediction, we follow a procedure proposed for mutual funds by Carhart (1997) and used for hedge funds by Teo (2011). First, we regress the monthly gross excess returns of each hedge fund i on the seven factors of the Fung-Hsieh model (see Fung and Hsieh (2004)). We use the gross excess return, because it allows us to measure whether a hedge fund can profit from lower cash holdings and a higher portfolio illiquidity without the noise introduced by performance and management fees. The Fung-Hsieh seven factor model is widely used to estimate hedge fund alphas (see, for example, Fung, Hsieh, Naik, and Ramadorai (2008); Teo (2009, 2011); and Patton and Ramadorai (2013)). The seven factors are: the excess return on the S&P 500 index (market factor); a small minus big factor (S-B factor) constructed as the difference between the return on the Russell 2000 index and the S&P 500; the change in the constant maturity yield of the 10-year Treasury bond (bond factor); the change in the Moody's Baa yield minus the change in the 10-year Treasury bond constant maturity yield (credit factor); and the returns on portfolios of lookback straddle options on currencies (currency trend factor), commodities (commodities trend factor), and bonds (bond trend factor) from Fung and Hsieh (2001). To ensure that we have enough data points to estimate the model, we select only hedge funds with 24 or more monthly return observations as in Patton and Ramadorai (2013). The return regression is given by

$$r^e_{im} = \alpha_i + \beta_i G_{im} + \epsilon_{im}, \text{ where } i = 1, 2, ..., N. \tag{6}$$

The gross excess return of hedge fund i in month m is given by r^e_{im}. The regressor G_{im} is a column vector of the seven Fung-Hsieh factors. The row vector of coefficient estimates $\hat{\beta}_i$ is then used to compute a monthly α_{im}:

$$\alpha_{im} = r^e_{im} - \hat{\beta}_i G_{im}. \tag{7}$$

We compute an average α_{it} for each quarter t based on the monthly α_{im} and estimate Fama and MacBeth (1973) cross-sectional regressions on the quarterly α_{it}:

$$\alpha_{it} = \psi + \gamma IC_{it-1} + \phi Y_{it-1} + \epsilon_{it}. \tag{8}$$

The control variables are included in the column vector Y_{it}. We use the control variables size, flows, share restrictions, financing duration, and manager stake: $log(NAV_{it-1})$, F_{it-1}, $ShareRes_{it-1}$, $FinDur_{it-1}$, and $MgrStake_{it-1}$. We also include strategy fixed effects.

The results are given in Table 7. We show the results for hedge funds' levered and delevered quarterly excess returns. When delevering, we divide the excess returns by the leverage measure, GAV/NAV. The coefficient on IC is negative and strongly significant for both the levered and delevered returns. As expected, the coefficient estimates for the delevered returns are slightly lower due to the reduction in the volatility of the dependent variable caused by deleveraging. The effect of IC on the risk-adjusted returns is economically significant. A one standard deviation (22 percentage points) increase in IC is associated with a reduction in the levered (delevered) annualized risk-adjusted return of 133 (106) basis points.

The estimated relationships between the control variables and risk-adjusted returns are as established in other papers of the asset management literature. Size and flows have a negative effect on performance in line with the hypothesis of negative returns to scale. The coefficient estimate of share restrictions is positive, indicating that hedge funds with long lock-up and redemption periods generate higher risk-adjusted returns. We are not aware of any paper investigating the effect of financing duration on risk-adjusted returns, but it is sensible to believe that this relation is positive: a longer financing duration allows hedge funds to pursue more illiquid strategies and generate an illiquidity premium. The ownership stake of the hedge fund manager does not appear to significantly affect the risk-adjusted returns.

The lower risk-adjusted returns of high-IC hedge funds raises the question why an investor would invest in a hedge fund with a concentrated investor base. In other words, why do we observe so many high-IC hedge funds in our sample? There are different possible explanations for why investors invest in high-IC hedge funds. First, the mechanism presented in this paper has to our knowledge not been analyzed in any other paper. Therefore, it is possible that investor awareness regarding this issue is low. Unlike in this paper, investors do not have a large regulatory panel dataset available to estimate the effect of IC on risk-adjusted returns. Our analysis suggests that such inattention is costly and it would beneficial for investors to monitor the investor concentration of hedge funds. Second, there are factors that could make it optimal for an investor to invest in a high-IC hedge fund despite the lower risk-adjusted returns. Being a large investor of a hedge fund could lead to better access to the hedge fund manager and facilitate monitoring (see, for example, Schmidt, Timmermann, and Wermers (2016)). On the other hand, small investors might be willing to invest in a high-IC hedge fund because a large investor could internalize "runs" on a hedge fund after a poor performance (see, for example, Chen, Goldstein, and Jiang (2010)). If large investors monitored hedge funds more closely or internalized runs, we would expect that high-IC hedge funds have a stronger or weaker flow-performance relationship, respectively. However, our analysis in Section 4.5.2 shows that IC has *no* effect on the flow-performance relationship of hedge funds. These results are consistent with investors investing in high-IC hedge funds because of inattention.

4.5 Alternative mechanisms

There are alternative ways through which hedge funds might deal with a high IC risk other than adjusting their portfolios such that potential large outflows can be absorbed. In this section, we examine these alternative ways of accounting for high-IC risk and analyze other mechanisms that could potentially explain some of our results. While we do not find evidence for these alternatives, we include factors that control for these alternatives in our baseline regressions.

4.5.1 Methods to prevent outflows

In this section, we discuss methods that a hedge fund manager could potentially use to prevent large outflows driven by a high investor concentration from occurring in the first place, instead of adjusting the portfolio such that large outflows can be absorbed. We examine the likelihood that these methods are employed. Ways to account for a high IC without adjusting the portfolio could be through share restrictions or redemption gates.

One could expect that hedge funds with a high IC have longer share restrictions. These longer share restrictions would give the hedge fund more time to sell assets and generate cash when facing large outflows. However, the correlation between share restrictions and IC is small, as shown in Figure 3, both for the entire sample and within individual strategies. The correlations are small in magnitude and often negative. For the total sample, the correlation is -0.16. If high-IC hedge funds used share restrictions to account for IC risk, the correlation would be larger in magnitude and positive. Further, we control for share restrictions in all our regression specifications and IC remains significant.

An explanation for the low correlation between IC and share restrictions is that share restrictions are generally set at the inception of a hedge fund and stated in the hedge fund's limited partnership agreement, so these restrictions are difficult to change throughout the hedge fund's life (see Aragon (2007) and Agarwal, Daniel, and Naik (2009)) as the investors' approval is required. This means that if IC changes because of investors deciding to invest in or withdraw from the hedge fund, a hedge fund cannot simply adjust its share restrictions to corresponds to its new IC. A hedge fund that tries to impose longer share restrictions because its IC increased would likely face resistance from its investors, because the investors would be unwilling to accept that their investment becomes more illiquid without any changes in the investment strategy of the hedge fund that would justify longer share restrictions. Therefore, we expect that hedge funds account for a high IC through portfolio adjustments that do not need investor approval and are more difficult for investors to observe, rather than through changes in share restrictions. Our results support this prediction.

In addition to the granular share restriction horizons captured in $ShareRes$, we examine whether the extent to which hedge funds can restrict (impose gates on) or suspend withdrawals and redemptions is related to the investor concentration of a fund. Figure 4 illustrates the fraction of funds that are able to impose restrictions on investor withdrawals and redemptions.[23] The depth of color,

[23] As captured in Question 49 (a), (b), and (c) of Form PF.

from darkest to lightest, correspond to the strength of the fund's restrictions, with each observation assigned to a set depending on the strongest restriction the fund is able to impose. The four sets represent the fraction of funds that (1) do not allow investors to withdraw their investment in the ordinary course; (2) are able to suspend any investor withdrawal of 100% the fund's NAV; (3) are able to materially restrict (via the imposition of "gates") any investor withdrawal of 100% the fund's NAV; and (4) have restrictions on 0% of their fund's NAV. Only 14% of observations fall in category (4). 86% of the observations are funds that are able to restrict or suspend 100% of investor withdrawals. We define the indicator $HasGates_{it}$ as 0 if a fund is in category (4), i.e., if the fund has no restrictions on investor withdrawals and redemptions, and 1 otherwise. $HasGates_{it}$ is uncorrelated with a fund's investor concentration. Inclusion of $HasGates$ as a control does not affect our results. This is unsurprising given that imposing discretionary gates on redemptions can lead to high reputational costs and subsequent difficulty in raising capital (Aiken, Clifford, and Ellis, 2015). It is a tool that is more likely to be used during periods of market stress when peer hedge funds are also imposing discretionary restrictions. As such, it would be a less effective tool for high-IC hedge funds to manage idiosyncratic investor liquidity shocks.[24]

4.5.2 Investor concentration and flow-performance sensitivity

The IC risk that we focus on in this paper is concerned with outflows due to idiosyncratic liquidity shocks to a hedge fund's investors who own a large share of the fund's NAV. These liquidity shocks are independent of the performance or other fundamentals of the hedge fund. Even if a hedge fund is performing well, a large investor can experience an idiosyncratic liquidity shock and redeem his/her investment in the fund. Having a diversified investor base reduces this risk of large outflows from idiosyncratic liquidity shocks and reduces the need to hold precautionary cash. However, separate from this mechanism, a concentrated investor base could potentially also affect the sensitivity of a hedge fund's flows to past performance, that is, its fundamental flows. On the one hand, large hedge fund investors potentially internalize the impact of their redemptions on the hedge fund and refrain from redeeming investments when the hedge fund performs poorly, in which case flows would be less sensitive to the hedge fund's performance, reducing the need for precautionary cash holdings. On the other hand, large hedge fund investors might have the resources to monitor their investments more closely, in which case flows would be more sensitive to the hedge fund's performance, increasing the need for precautionary cash holdings.

We can infer from existing research that evidence for either mechanism, internalizing the impact of redemptions or better monitoring, could be present in our data. Chen, Goldstein, and Jiang (2010) show for equity mutual funds that the flow-performance sensitivity is stronger for funds that hold more illiquid assets, but this effect disappears for mutual funds held by large in-

[24]We do not show $HasGates$ in our main regression results as there is little variation in the ability of the funds in our sample to restrict or suspend investor withdrawals (86% of fund observations have $HasGates$=1, i.e., can impose restrictions on withdrawals of 100% of the fund's NAV) and the variable is collinear to a combination of other regressors. More granular information on investor share restrictions is captured in $ShareRes$ and shown in all regressions.

stitutional investors as opposed to retail investors, because unlike the latter, large institutional investors internalize the price impact of their redemptions and are *less* likely to run on a mutual fund that is in distress. In contrast, Schmidt, Timmermann, and Wermers (2016) find that large institutional investors were *more* likely to run on money market funds with illiquid assets than smaller institutional or retail investors around the collapse of Lehman Brothers in September 2008. The authors posit that the largest institutional investors have more resources to monitor their investments, and thus, react more quickly when a money market fund is in distress. Schmidt, Timmermann, and Wermers (2016) discuss how differences between mutual funds and money market funds along dimensions such as NAV structure and payoffs can explain why their findings differ from Chen, Goldstein, and Jiang (2010). Importantly, hedge funds have several unique features not present in mutual funds or money market funds. For example, hedge funds have share restrictions which allows them to ameliorate the impacts of asset illiquidity. Further, hedge funds have no retail investors, do not disclose their portfolio holdings to investors, and on average have investor bases that are much more concentrated than mutual funds and money market funds. Therefore, it is a priori difficult to determine what effect investor concentration will have on the flow-performance relationship for hedge funds.

To test whether IC affects the flow-performance sensitivity of hedge funds, we estimate the panel model given by

$$F_{it} = \psi + \gamma_1 IC_{it-1} + \gamma_2 Performance_{it-1} \times IC_{it-1} + \phi_1 Performance_{it-1} + \phi_2 \delta Z_{it-1} + \epsilon_{it}, \quad (9)$$

where $Performance_{it-1}$ is a measure of the hedge funds' lagged performance. We try four measures of quarterly performance: net returns, negative net returns, net returns terciles, and net returns quintiles. The control variables in vector Z_{it-1} are lagged size, flows, share restrictions, and manager stake: $log(NAV_{it-1})$, F_{it-1}, $ShareRes_{it-1}$, and $MgrStake_{it-1}$. We include quarter and strategy or fund fixed effects. The standard errors are clustered by quarter. If a high IC is associated with flows that are less sensitive to performance, we would expect the estimate of γ_2 to be negative and significant. If a high IC is associated with flows that are more sensitive to performance, then the estimate of γ_2 would be positive and significant.

The results with quarter and fund fixed effects are given in Table 8, and the results with quarter and strategy fixed effects are reported in Table B.3 in Appendix B. In line with the existing literature on hedge fund flows, we find evidence that higher returns lead to higher subsequent flows. However, the coefficient estimates of IC_{it-1} and of the interaction term $Performance_{it-1} \times IC_{it-1}$ are insignificant for all specifications. These results suggest that the concentration of the investor base does not affect the flow-performance sensitivity of a hedge fund, and the documented relationship of IC and precautionary cash holdings is not affected by any differences in the flow-performance sensitivity between low- and high-IC hedge funds.

4.5.3 Do high-IC funds focus on providing portfolio protection?

Could it be that investors of high-IC funds are attracted to these hedge funds because the funds offer portfolio protection by outperforming peer funds during market downturns? This alternative hypothesis could explain the underperformance of high-IC hedge funds given that the regulatory data used for our analysis is available from 2012Q4 to 2017Q4, a relatively quiet period in financial markets. However, we can test if high-IC hedge funds are focused on such "crisis alpha" strategies. In Form PF, managers are required to disclose the projected performance of their investment portfolios under several market stress scenarios.[25] These values should directly reflect a manager's expectations on how the portfolio will fare during a market downturn. The manager's expectations would indicate how the fund's investment strategy and portfolio risks are represented to investors.

If high-IC funds are expected to outperform in crisis scenarios, it should be reflected in the expectations disclosed by the manager. The crisis scenarios available in the data are (i) "Equity prices decrease 20%", (ii) "Credit spreads increase 250bps", (iii) "Implied volatilities increase 10 percentage points", (iv) "Default rates increase 5 percentage points (corporate bonds and CDS)" (v) "Default rates increase 5 percentage points (ABS)". If a high-IC hedge fund is meant to provide downside protection to its investors, this should be directly reflected in the manager's expectations of portfolio performance being relatively higher under these scenarios.

In Table 9 column 1, we show the unconditional correlation between a hedge fund's IC and expected portfolio changes under different stress scenarios. Across the five scenarios, these correlations are small or negative, indicating that it is unlikely that high-IC funds are targeting portfolio protection during market stress. Further, we control for other fund characteristics to estimate the conditional relationship between the IC level of a fund and its portfolio performance under different scenarios using the regression specification, similar to equation (3),

$$y_{it} = \psi + \gamma IC_{it} + \phi Z_{it} + \epsilon_{it}, \tag{10}$$

where y_{it} is the percent portfolio change under a given stress scenario. The control variables are included in the vector Z_{it}. We use the control variables size, flow, share restrictions, financing duration, leverage, and manager stake: $log(NAV_{it})$, F_{it}, $ShareRes_{it}$, $FinDur_{it}$, $Leverage_{it}$, and $MgrStake_{it}$, respectively. We also include quarter and strategy fixed effects and cluster standard errors by quarter and fund. If hedge funds with higher IC provide downside protection during crisis scenarios, we would expect γ to be significant and positive. Table 9 column 2 shows the coefficient estimates for γ. It is not significantly positive any of the stress scenarios. In fact, the relationship is marginally negative for scenario (i).

We do not find that high-IC funds are expected to provide significantly more portfolio protections under stress scenarios than low-IC funds. This indicates that investors of high-IC funds are not paying (by accepting lower fund performance during quiet market periods) for enhanced portfolio performance during crisis periods.

[25] See Form PF, Section 2b, Question 42.

5 Conclusion

We investigate a novel source of hedge fund fragility stemming from how diversified hedge funds are with respect to their investors. Using a simple theoretical framework, we show that a hedge fund with a high investor concentration (IC) is more exposed to the risk of idiosyncratic liquidity shocks to its investors. Negative liquidity shocks to an investor can lead to outflows that are unexpected and independent of the hedge fund's fundamentals, and such outflows are more likely for a hedge fund with ownership structure that is concentrated. We predict that to address the risk of large unexpected outflows, a high-IC hedge fund holds a larger share of precautionary cash in its portfolio and pays a liquidity premium. We test these hypotheses using a novel regulatory dataset on hedge funds.

The SEC's Form PF requires hedge funds to report the percentage of NAV held by the five investors who are the largest investors in the fund. We use this five-investor concentration ratio as our empirical measure of IC. First, we find that high-IC hedge funds have a greater probability of experiencing large liquidity driven outflows. Second, in line with our prediction, high-IC hedge funds make portfolio adjustments, including holding more precautionary cash, which help absorb sudden large outflows. These results are not driven by differences in share restrictions, investment strategy, or manager ownership.

Our paper complements the existing hedge fund literature that focuses on how hedge funds are exposed to risk factors through the assets they hold. We show that the a fragile ownership structure can also pose a substantial risk, and we analyze how hedge funds account for this risk. Our main finding that high-IC hedge funds hold more precautionary cash is important for policymakers who monitor the impact of hedge funds on market efficiency and stability. Further, awareness of the mechanism documented in this paper is also important for hedge fund investors when requesting risk disclosures and allocating their portfolios efficiently.

References

Ackermann, Carl, Richard McEnally, and David Ravenscraft, 1999, The performance of hedge funds: Risk, return, and incentives, *Journal of Finance* 54, 833–874.

Agarwal, Vikas, George O. Aragon, and Zhen Shi, 2018, Liquidity transformation and financial fragility: Evidence from funds of hedge funds, *Journal of Financial and Quantitative Analysis* pp. 1–58.

Agarwal, Vikas, Naveen D. Daniel, and Narayan Y. Naik, 2009, Role of managerial incentives and discretion in hedge fund performance, *Journal of Finance* 64, 2221–2256.

Agarwal, Vikas, and Narayan Y. Naik, 2004, Risks and portfolio decisions involving hedge funds, *Review of Financial Studies* 17, 63–98.

Aiken, Adam L., Christopher P. Clifford, and Jesse A. Ellis, 2015, Hedge funds and discretionary liquidity restrictions, *Journal of Financial Economics* 116, 197–218.

Aragon, George O., 2007, Share restrictions and asset pricing: Evidence from the hedge fund industry, *Journal of Financial Economics* 83, 33–58.

———, A. Tolga Ergun, Mila Getmansky, and Giulio Girardi, 2017, Hedge fund liquidity management, *Working Paper*.

Bali, Turan G., Stephen J. Brown, and Mustafa O. Caglayan, 2012, Systematic risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 106, 114–131.

Barth, Daniel, and Phillip J. Monin, 2018, Illiquidity in intermediary portfolios: Evidence from large hedge funds, *Working Paper*.

Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi, 2012, Hedge fund stock trading in the financial crisis of 2007-2009, *Review of Financial Studies* 1, 1–54.

———, and John Sedunov, 2018, The granular nature of large institutional investors, *Working Paper*.

Bollen, Nicolas P.B., and Veronika K. Pool, 2008, Conditional return smoothing in the hedge fund industry, *Journal of Financial and Quantitative Analysis* 43, 267–298.

———, 2009, Do hedge fund managers misreport returns? Evidence from the pooled distribution, *Journal of Finance* 64, 2257–2288.

Bollen, Nicolas P.B., and Robert E. Whaley, 2009, Hedge fund risk dynamics: Implications for performance appraisal, *Journal of Finance* 64, 987–1037.

Boyson, Nicole M., Christof W. Stahel, and Rene M. Stulz, 2010, Hedge fund contagion and liquidity shocks, *Journal of Finance* 65, 1789–1816.

Brunnermeier, Markus K., and Stefan Nagel, 2004, Hedge funds and the technology bubble, *Journal of Finance* 59, 2013–2040.

Brunnermeier, Markus K., and Lasse H. Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Buraschi, Andrea, Robert Kosowski, and Fabio Trojani, 2014, When there is no place to hide: Correlation risk and the cross-section of hedge fund returns, *Review of Financial Studies* 27, 581–616.

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

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.

Chevalier, Judith, and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Chordia, Tarun, 1996, The structure of mutual fund charges, *Journal of Financial Economics* 41, 3–39.

Christoffersen, Susan E.K., David K. Musto, and Russ Wermers, 2014, Investor flows to asset managers: Causes and consequences, *Annual Review of Financial Economics* 6, 289–310.

Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.

Da, Zhi, Borja Larrain, Clemens Sialm, and José Tessada, 2018, Destabilizing financial advice: Evidence from pension fund reallocations, *Review of Financial Studies* 31, 3720–3755.

Dahlquist, Magnus, José V. Martinez, and Paul Soderlind, 2017, Individual investor activity and performance, *Review of Financial Studies* 30, 866–899.

De Long, Bradford J., Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Edelen, Roger M., 1999, Investor flows and the assessed performance of open-end mutual funds, *Journal of Financial Economics* 53, 439–466.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilbrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Flood, Mark D., and Phillip Monin, 2016, Form PF and the systemic risk of hedge funds: Risk-measurement precision for option portfolios, *Journal of Alternative Investments* 18, 125–147.

———, and Lina Bandyopadhyay, 2015, Gauging Form PF: Data tolerances in regulatory reporting on hedge fund risk exposures, *Office of Financial Research Working Paper.*

Fung, William, and David A. Hsieh, 1997, Empirical characteristics of dynamic trading strategies: The case of hedge funds, *Review of Financial Studies* 10, 275–302.

——— , 2001, The risk in hedge fund strategies: Theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

——— , 2004, Hedge fund benchmarks: A risk-based approach, *Financial Analysts Journal* 60, 65–80.

——— , Narayan Y. Naik, and Tarun Ramadorai, 2008, Hedge funds: Performance, risk, and capital formation, *Journal of Finance* 63, 1777–1803.

Getmansky, Mila, Bing Liang, Christopher Schwarz, and Russ Wermers, 2015, Share restrictions and investor flows in the hedge fund industry, *Working Paper*.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov, 2004, An econometric model of serial correlation and illiquidity in hedge fund returns, *Journal of Financial Economics* 74, 529–609.

Goldstein, Itay, Hao Jiang, and David Ng, 2017, Investor flows and fragility in corporate bond funds, *Journal of Financial Economics* 126, 592–613.

Greenwood, Robin, and David Thesmar, 2011, Stock price fragility, *Journal of Financial Economics* 102, 471–490.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Gupta, Arpita, and Kunal Sachdeva, 2017, Skin or skim? Inside investment and hedge fund performance, *Working Paper*.

Hombert, Johan, and David Thesmar, 2014, Overcoming limits of arbitrage: Theory and evidence, *Journal of Financial Economics* 111, 26–44.

Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *Journal of Finance* 67, 2015–2050.

Jylha, Petri, Kalle Rinne, and Matti Suominen, 2014, Do hedge funds supply or demand liquidity, *Review of Finance* 18, 1259–1298.

Khandani, Amir E., and Andrew W. Lo, 2011, What happened to the quants in August 2007? Evidence from factors and transactions data, *Journal of Financial Markets* 14, 1–46.

Kruttli, Mathias S., Andrew J. Patton, and Tarun Ramadorai, 2015, The impact of hedge funds on asset markets, *Review of Asset Pricing Studies* 5, 185–226.

Li, Haitao, Xiaoyan Zhang, and Rui Zhao, 2011, Investing in talents: Manager characteristics and hedge fund performances, *Journal of Financial and Quantitative Analysis* 46, 59–82.

Lowenstein, Roger, 2000, When genius failed: The rise and fall of Long-Term Capital Management, Random House New York.

Lynch, Anthony W., and David K. Musto, 2003, How investors interpret past fund returns, *Journal of Finance* 58, 2033–2058.

Patton, Andrew J., and Tarun Ramadorai, 2013, On the high-frequency dynamics of hedge fund risk exposures, *Journal of Finance* 68, 597–635.

——— , and Michael Streatfield, 2015, Change you can believe in? Hedge fund data revisions, *Journal of Finance* 70, 963–999.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

Sadka, Ronnie, 2010, Liquidity risk and the cross-section of hedge fund returns, *Journal of Financial Economics* 98, 54–71.

Schmidt, Lawrence, Allan Timmermann, and Russ Wermers, 2016, Runs on money market mutual funds, *American Economic Review* 106, 2625–2657.

SEC, 2016, Investment company liquidity risk management programs, *Final Rule*.

Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Teo, Melvyn, 2009, The geography of hedge funds, *Review of Financial Studies* 22, 3531–3561.

——— , 2011, The liquidity risk of liquid hedge funds, *Journal of Financial Economics* 100, 24–44.

Thompson, Samuel B., 2011, Simple formulas for standard errors that cluster by both firm and time, *Journal of Financial Economics* 99, 1–10.

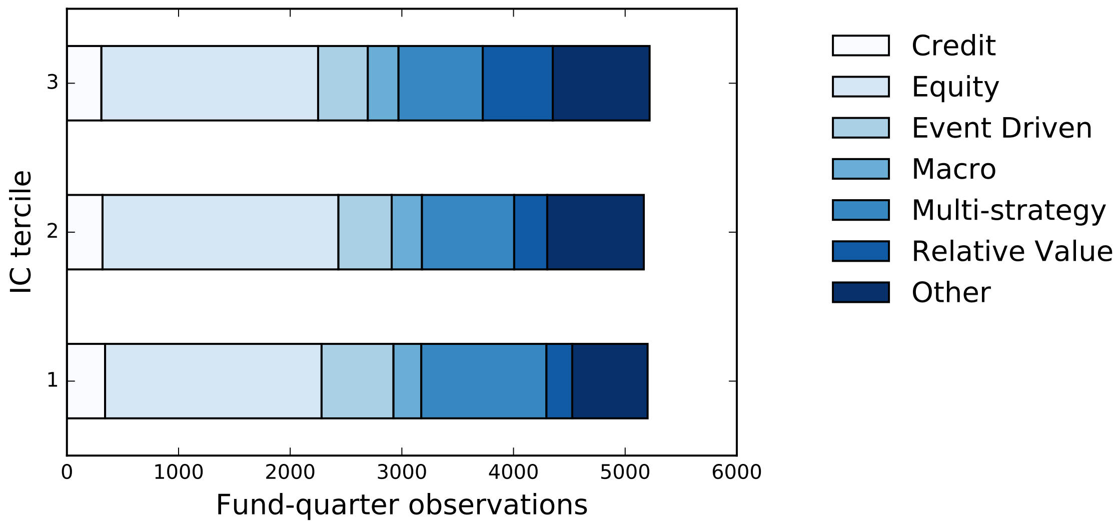
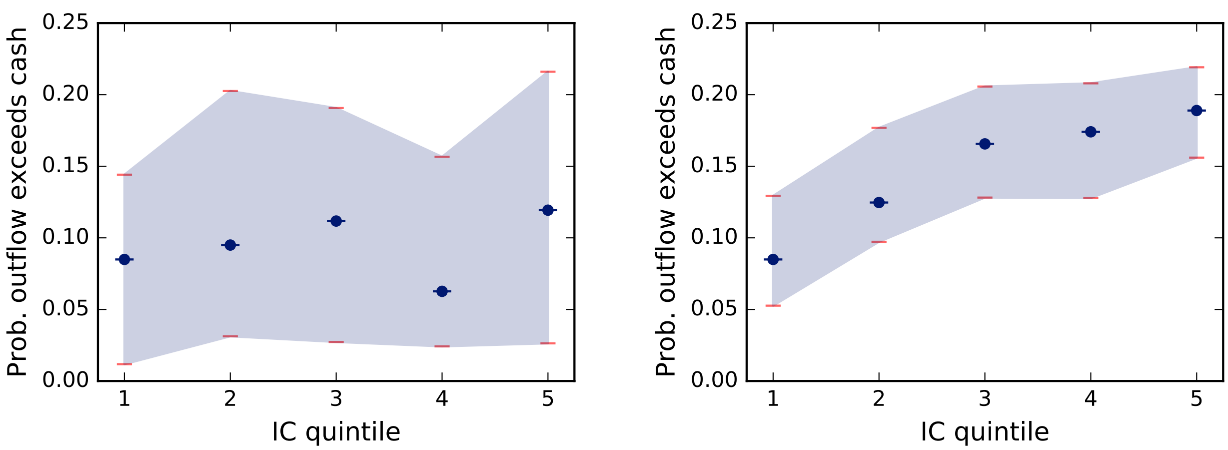


Figure 1: Investor concentration by strategy

This figure shows the number of fund-quarter observations for each strategy and IC tercile. Every quarter, the hedge funds are sorted based on IC. The first tercile contains the hedge funds with the lowest IC values.



(a) Adjusted cash levels (actual)

(b) Constant cash levels (hypothetical case)

Figure 2: Probability of outflows exceeding cash

This figure shows the quarterly probability of liquidity outflows exceeding cash for each IC quintile with 95% bootstrapped confidence intervals. The liquidity flows are estimated with the regression model in 1 when setting $P = 1$. The first quintile contains the hedge funds with the lowest IC values. Liquidity flows are assumed to be normally distributed. For Plot (a), the probability is computed based on the median quarterly liquidity flows, standard deviation of liquidity flows, and cash of each quintile. For Plot (b), the probability is again computed based on the median quarterly liquidity flows and standard deviation of liquidity flows of each quintile, but the median cash level of the first quintile is used to compute the probabilities for all quintiles.

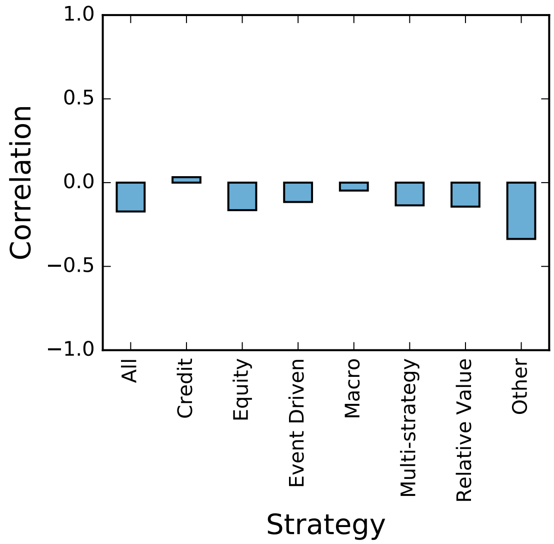


Figure 3: Correlation between investor concentration and share restrictions

This figure shows the correlation between IC and share restrictions at the fund level for the total sample and for the strategy subsamples.

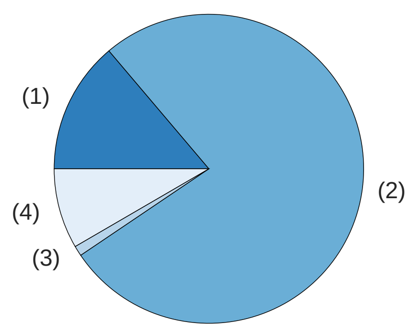


Figure 4: Restrictions on withdrawals and redemptions

This figure illustrates the fraction of funds that are able to impose restrictions on investor withdrawals and redemptions. The depth of color, from darkest to lightest, correspond to the strength of fund's restrictions, with each observation assigned to a set depending on the strongest restriction the fund is able to impose. The four sets represent the fraction of funds that (1) do not allow investors to withdraw their investment in the ordinary course; (2) are able to suspend any investor withdrawal for 100% the fund's NAV; (3) are able to materially restrict (via the imposition of "gates") any investor withdrawal for 100% the fund's NAV; and (4) have restrictions on 0% of their fund's NAV.

Table 1: Summary statistics

Panel A reports summary statistics for a range of hedge fund variables. A summary of the variable definitions and their data sources are included in Table C.1 in the Data appendix. The data are quarterly from 2012:Q4 to 2017:Q4, and the "Number of observations" column reports the number of hedge fund-quarter observations. Panel B reports the number of observations and the averages of all the variables for each hedge fund strategy.

Panel A: Hedge fund variables

		Number of observations	Average	Stand. dev.	10%	50%	90%
Investor Concentration (%)	IC_{it}	15,587	50.5	21.8	24.0	47.0	84.0
Gross returns (%)	g_{it}	15,587	2.8	6.4	-3.6	2.4	9.1
Net returns (%)	r_{it}	15,587	2.0	5.5	-3.6	1.9	7.7
Flows (%)	F_{it}	15,069	-1.2	8.7	-13.2	-0.6	9.6
Net asset value (million US$)	NAV_{it}	15,587	2,184.2	3,431.7	342.4	1,177.5	4,715.8
Unencumbered cash (% of NAV)	$Cash_{it}/NAV_{it}$	15,587	15.3	20.5	0.0	6.8	46.0
Portfolio illiquidity (days)	$PortIlliq_{it}$	15,450	52.4	89.4	1.4	14.2	164.2
Share restriction (days)	$ShareRes_{it}$	15,507	166.6	121.4	19.0	147.4	366.0
Financing duration (days)	$FinDur_{it}$	12,681	47.1	88.7	0.5	0.5	135.5
Leverage (GAV/NAV)	$Leverage_{it}$	15,587	1.8	2.1	1.0	1.3	2.6
Manager stake (%)	$MgrStake_{it}$	15,587	8.1	10.2	0.0	4.0	22.0
Number of investors	$NumInvestors_{it}$	15,587	182.8	353.3	19.0	98.0	416.0
Minimum investment (million US$)	$MinInv_{it}$	15,587	3.8	5.7	0.1	1.0	10.0

Panel B: Number of observations and average values by strategy

	Credit	Equity	Event Driven	Macro	Multi-strat.	Rel. Value	Other
Number of observations	970	5,993	1,566	795	2,702	1,154	2,407
IC_{it}	49.2	51.0	45.9	51.2	47.0	58.9	52.6
g_{it}	2.3	2.5	3.7	1.0	2.2	2.2	4.7
r_{it}	1.6	1.9	2.6	0.5	1.5	1.6	3.3
F_{it}	-2.0	-0.4	-2.7	-1.1	-0.8	-1.7	-2.4
NAV_{it}	1,144.9	1,874.0	1,974.3	2,612.3	3,407.5	1,248.1	2,446.2
$Cash_{it}$	13.7	10.4	10.6	43.3	21.7	17.2	14.0
$PortIlliq_{it}$	87.9	20.1	81.5	11.0	56.5	49.5	112.4
$ShareRes_{it}$	212.6	132.6	246.7	94.7	187.3	160.1	185.5
$FinDur_{it}$	102.9	21.5	53.0	17.5	44.6	51.5	103.3
$Leverage_{it}$	1.6	1.5	1.4	3.3	1.9	3.3	1.5
$MgrStake_{it}$	6.5	9.3	10.4	6.7	7.9	6.8	5.4
$NumInvestors_{it}$	164.8	163.3	207.3	202.0	279.4	155.3	121.0
$MinInv_{it}$	3.3	3.6	4.6	3.9	3.9	2.4	4.5

Table 2: Probability of large liquidity outflows

This table reports the coefficient estimates (odds ratios) and t-statistics of the logit regression model given in equation (2). The dependent variable is an indicator variable that takes the value 1 if liquidity flows are less than -20% of NAV in Panel A and -15% of NAV in Panel B. The independent variables are lagged IC, IC tercile dummies, size, and manager stake. The liquidity flows are estimated as the residuals of the first stage regression given in equation (1). The first stage regression is estimated with one and two lags of the independent variables. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter fixed effects and strategy fixed effects are used. The standard errors are clustered by quarter. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Probability of liquidity flows $\leq$ - 20%

Dependent variable: $P(\hat{F}_{it}^{L} \leq -20\% \mid IC_{t-1}, W_{t-1})$	First stage regression: one lag				First stage regression: two lags			
IC_{it-1}	1.014***	1.008***			1.017***	1.013***		
	(8.179)	(3.311)			(5.961)	(3.477)		
IC 2$^{\text{nd}}$ tercile$_{it-1}$			1.529**	1.302			1.184	1.060
			(2.212)	(1.348)			(0.797)	(0.258)
IC 3$^{\text{rd}}$ tercile$_{it-1}$			2.236***	1.645***			2.319***	1.848***
			(6.714)	(3.107)			(5.438)	(3.042)
$log(NAV_{it-1})$		0.744***		0.743***		0.815**		0.799**
		(-3.719)		(-3.698)		(-2.119)		(-2.350)
$MgrStake_{it-1}$		0.991		0.991		0.987		0.986
		(-1.416)		(-1.580)		(-1.560)		(-1.601)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,107	13,107	13,107	13,107	12,284	12,284	12,284	12,284
Pseudo R^2	0.034	0.043	0.034	0.044	0.042	0.046	0.041	0.047

Table 2: Probability of large liquidity outflows (continued)

Panel B: Probability of liquidity flows ≤ -15%

Dependent variable: $P(\hat{F}_{it}^{L} \leq -15\% | IC_{t-1}, W_{t-1})$

	First stage regression: one lag				First stage regression: two lags			
IC_{it-1}	1.010***	1.004*			1.013***	1.007***		
	(6.335)	(1.850)			(8.371)	(3.890)		
IC 2$^{\text{nd}}$ tercile$_{it-1}$			1.633***	1.378**			1.678***	1.437**
			(3.670)	(2.376)			(3.207)	(2.224)
IC 3$^{\text{rd}}$ tercile$_{it-1}$			1.777***	1.291**			2.035***	1.508***
			(5.265)	(2.029)			(6.205)	(3.106)
$log(NAV_{it-1})$		0.734***		0.734***		0.754***		0.747***
		(-8.010)		(-8.461)		(-7.401)		(-7.176)
$MgrStake_{it-1}$		0.996		0.995		0.991*		0.990*
		(-0.943)		(-1.106)		(-1.652)		(-1.866)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,107	13,107	13,107	13,107	12,284	12,284	12,284	12,284
Pseudo R^2	0.018	0.028	0.019	0.029	0.019	0.028	0.019	0.028

Table 3: Investor concentration and cash

This table reports the coefficient estimates and t-statistics of the panel regression models given in equations (3) and (4). The dependent variable is cash normalized by NAV. The independent variables are IC, IC tercile dummies, size, flows, share restriction, financing duration, leverage, and manager stake. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter fixed effects and strategy fixed effects are used where indicated. The standard errors are clustered by quarter and hedge fund. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Cash normalized by NAV, $Cash_{it}/NAV_{it}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IC_{it}	0.089***	0.091***			0.136***	0.140***		
	(3.095)	(3.150)			(4.606)	(4.706)		
IC 2^{nd} tercile$_{it}$			0.166	0.162			1.716*	1.712*
			(0.165)	(0.161)			(1.776)	(1.768)
IC 3^{rd} tercile$_{it}$			4.105***	4.127***			6.285***	6.303***
			(2.986)	(3.003)			(4.607)	(4.621)
$log(NAV_{it})$					2.749***	2.770***	2.511***	2.507***
					(4.965)	(5.006)	(4.731)	(4.716)
F_{it}					-0.055*	-0.055	-0.058*	-0.055
					(-1.720)	(-1.618)	(-1.814)	(-1.619)
$ShareRes_{it}$					-0.026***	-0.026***	-0.027***	-0.027***
					(-5.053)	(-5.057)	(-5.092)	(-5.101)
$FinDur_{it}$					0.001	0.001	0.001	0.001
					(0.116)	(0.167)	(0.125)	(0.161)
$Leverage_{it}$					1.187***	1.195***	1.184***	1.189***
					(2.589)	(2.610)	(2.595)	(2.608)
$MgrStake_{it}$					-0.082*	-0.082*	-0.086*	-0.086*
					(-1.696)	(-1.693)	(-1.769)	(-1.758)
Quarter FE	No	Yes	No	Yes	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,242	12,242	12,242	12,242	12,242	12,242	12,242	12,242
Adjusted R^2	0.164	0.165	0.164	0.165	0.220	0.222	0.218	0.219

Table 4: Cash regressions with individual and institutional investor split

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (3) estimated for hedge fund subsamples. Shown are the model estimates for the subsample of hedge funds for which individual investors hold 50% and 25% or more of the hedge fund's equity. The complementary subsamples are given in the last two columns. The dependent variable is cash normalized by NAV. The independent variables are IC, size, flows, share restriction, financing duration, leverage, and manager stake. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter fixed effects and strategy fixed effects are used. The standard errors are clustered by quarter and hedge fund. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Cash normalized by NAV, $Cash_{it}/NAV_{it}$				
	Individual investor share greater or equal to		Institutional investor share greater than	
	50%	25%	50%	75%
IC_{it}	0.113*	0.153***	0.135***	0.108***
	(1.812)	(3.782)	(4.144)	(2.826)
$log(NAV_{it})$	1.530	2.216***	2.796***	2.534***
	(1.575)	(3.077)	(4.571)	(3.495)
F_{it}	-0.179*	-0.119**	-0.051	-0.053
	(-1.675)	(-2.081)	(-1.439)	(-1.360)
$Leverage_{it}$	-0.799	0.407	1.302***	1.258***
	(-0.781)	(0.490)	(2.865)	(2.649)
$ShareRes_{it}$	0.005	-0.003	-0.030***	-0.035***
	(0.403)	(-0.427)	(-5.426)	(-5.661)
$FinDur_{it}$	0.021	0.021**	0.000	-0.002
	(1.369)	(2.045)	(0.023)	(-0.460)
$MgrStake_{it}$	-0.034	0.017	-0.065	-0.075
	(-0.369)	(0.272)	(-1.140)	(-1.039)
Quarter FE	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes
Observations	1,366	3,336	10,876	8,906
Adjusted R^2	0.235	0.226	0.226	0.226

Table 5: Investor concentration and return volatility

This table reports the coefficient estimates and t-statistics of the panel regression models given in equation (5). The dependent variable is the absolute deviation from the average for semi-annual and annual returns, as used by Hombert and Thesmar (2014). The independent variables are IC, size, flows, share restrictions, financing duration, leverage, and manager stake. The frequency of the data is semi-annual (SA) for the first four columns and annual (A) for the last four columns. The time period is from 2012:Q4 to 2017:Q4. Time fixed effects and strategy fixed effects are used where indicated. The standard errors are clustered by time. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Absolute deviation to average return, $\lvert r_{it} - \hat{r}_i \rvert$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IC_{it}	-0.004	-0.004	-0.022***	-0.020***	-0.034**	-0.034**	-0.043**	-0.041**
	(-0.863)	(-0.846)	(-3.584)	(-3.698)	(-2.712)	(-3.346)	(-3.559)	(-3.925)
$log(NAV_{it})$			-0.866***	-0.828***			-0.756***	-0.682***
			(-6.757)	(-6.715)			(-6.986)	(-6.794)
F_{it}			-0.097**	-0.096**			-0.126	-0.156*
			(-2.345)	(-2.304)			(-1.424)	(-2.289)
$ShareRes_{it}$			0.004**	0.004***			0.007***	0.007**
			(2.997)	(3.214)			(4.069)	(4.554)
$FinDur_{it}$			0.015***	0.015***			0.016**	0.017**
			(7.151)	(7.140)			(2.840)	(2.733)
$Leverage_{it}$			-0.055	-0.052			-0.038	-0.023
			(-1.562)	(-1.437)			(-0.755)	(-0.487)
$MgrStake_{it}$			-0.004	-0.003			0.000	-0.000
			(-0.393)	(-0.338)			(0.031)	(-0.031)
Frequency	SA	SA	SA	SA	A	A	A	A
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,623	4,623	4,623	4,623	3,207	3,207	3,207	3,207
Adjusted R^2	0.026	0.049	0.067	0.090	0.030	0.052	0.052	0.075

Table 6: Investor concentration and portfolio illiquidity

This table reports the coefficient estimates and t-statistics of the panel regression models given in equations (3) and (4), but with the dependent variable being portfolio illiquidity. The independent variables are IC, IC tercile dummies, size, flows, share restriction, financing duration, leverage, and manager stake. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter fixed effects and strategy fixed effects are used where indicated. The standard errors are clustered by quarter and hedge fund. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Portfolio illiquidity, $PortIlliq_{it}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IC_{it}	-0.374***	-0.386***			-0.362***	-0.377***		
	(-3.854)	(-3.952)			(-4.017)	(-4.165)		
IC 2^{nd} tercile$_{it}$			-0.221	-0.081			-3.762	-3.703
			(-0.042)	(-0.015)			(-0.887)	(-0.873)
IC 3^{rd} tercile$_{it}$			-15.748***	-15.761***			-14.437***	-14.486***
			(-3.038)	(-3.036)			(-3.019)	(-3.028)
$log(NAV_{it})$					-12.237***	-12.365***	-11.234***	-11.240***
					(-5.717)	(-5.763)	(-5.406)	(-5.401)
F_{it}					0.069	0.111	0.072	0.106
					(0.501)	(0.778)	(0.528)	(0.745)
$Leverage_{it}$					-1.305**	-1.360**	-1.299**	-1.347**
					(-2.372)	(-2.467)	(-2.337)	(-2.411)
$ShareRes_{it}$					0.283***	0.283***	0.286***	0.287***
					(13.175)	(13.173)	(13.282)	(13.289)
$FinDur_{it}$					0.276***	0.275***	0.276***	0.276***
					(8.378)	(8.356)	(8.360)	(8.338)
$MgrStake_{it}$					-0.524***	-0.522***	-0.523***	-0.522***
					(-3.312)	(-3.300)	(-3.231)	(-3.232)
Time FE	No	Yes	No	Yes	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,239	12,239	12,239	12,239	12,239	12,239	12,239	12,239
Adjusted R^2	0.151	0.152	0.149	0.150	0.446	0.447	0.444	0.444

Table 7: Investor concentration and risk-adjusted returns

This table reports the coefficient estimates and t-statistics when estimating the model given in equation (8) with the estimation method of Fama and MacBeth (1973). The dependent variable is the quarterly average of the monthly Fung-Hsieh seven factor risk-adjusted returns given in equation (7). The returns are deleveraged where indicated. The independent variables are lagged IC, size, flows, share restriction, financing duration, and manager stake. The data are quarterly from 2012:Q4 to 2017:Q4. Strategy fixed effects are used where indicated. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Risk-adjusted returns, α_{it}				
IC_{it-1}	-0.004**	-0.005***	-0.003**	-0.004***
	(-2.720)	(-3.713)	(-2.760)	(-3.628)
$log(NAV_{it-1})$	-0.164***	-0.195***	-0.132***	-0.154***
	(-4.973)	(-5.350)	(-6.209)	(-6.609)
$ShareRes_{it-1}$	0.002***	0.002***	0.002***	0.001***
	(7.154)	(8.195)	(7.817)	(8.762)
$FinDur_{it-1}$	0.006***	0.006***	0.004***	0.004***
	(8.224)	(7.755)	(7.715)	(7.514)
F_{it-1}	-0.008***	-0.008***	-0.007***	-0.007***
	(-2.958)	(-2.923)	(-3.452)	(-3.436)
$MgrStake_{it-1}$	-0.000	0.000	0.001	0.001
	(-0.071)	(0.012)	(0.972)	(0.506)
Strategy FE	No	Yes	No	Yes
Deleveraged	No	No	Yes	Yes
Observations	9,450	9,450	9,450	9,450

Table 8: Investor concentration and flow-performance sensitivity

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (9). The dependent variable are quarterly flows. The independent variables are lagged IC, flows, returns, return terciles, return quintiles, size, share restriction, and manager stake. The coefficient estimates of the variables lagged flows, size, share restriction, and manager stake are not shown. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter and fund fixed effects are used. The standard errors are clustered by quarter. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Flows, F_{it}				
	(1)	(2)	(3)	(4)
IC_{it-1}	-0.021	-0.022	-0.021	-0.024*
	(-1.502)	(-1.585)	(-1.604)	(-1.827)
$r_{it-1} \times IC_{it-1}$	-0.000			
	(-0.257)			
$r_{it-1} \times I_{r_{it-1}<0} \times IC_{it-1}$		-0.001		
		(-0.726)		
r 2$^{\text{nd}}$ tercile$_{it-1} \times IC_{it-1}$			0.002	
			(0.270)	
r 3$^{\text{rd}}$ tercile$_{it-1} \times IC_{it-1}$			-0.003	
			(-0.721)	
r 2$^{\text{nd}}$ quintile$_{it-1} \times IC_{it-1}$				0.003
				(0.246)
r 3$^{\text{rd}}$ quintile$_{it-1} \times IC_{it-1}$				0.004
				(0.404)
r 4$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.008
				(1.078)
r 5$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				-0.003
				(-0.329)
r_{it-1}	0.071*	0.036		
	(2.088)	(1.083)		
$r_{it-1} \times I_{r_{it-1}<0}$		0.104		
		(1.610)		
r 2$^{\text{nd}}$ tercile$_{it-1}$			0.454	
			(1.360)	
r 3$^{\text{rd}}$ tercile$_{it-1}$			0.995***	
			(3.366)	
r 2$^{\text{nd}}$ quintile$_{it-1}$				0.561
				(1.189)
r 3$^{\text{rd}}$ quintile$_{it-1}$				0.633
				(1.332)
r 4$^{\text{th}}$ quintile$_{it-1}$				0.702*
				(1.972)
r 5$^{\text{th}}$ quintile$_{it-1}$				1.201***
				(2.884)
Quarter FE	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Observations	13,113	13,113	13,113	13,113
Adjusted R^2	0.147	0.147	0.147	0.147

Table 9: Investor concentration and portfolio changes under crisis scenarios

Column 1 shows the unconditional correlation between a hedge fund's IC and expected portfolio changes under different stress scenarios. Column 2 shows the coefficient estimates for γ in regression equation (10). The significance of the γ estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

		ρ(portfolio change, IC)	γ
(i)	Equity prices decrease 20%	-0.076	-0.021*
(ii)	Credit spreads increase 250bps	-0.098	0.003
(iii)	Implied volatilities increase 10 percentage points	0.069	0.010
(iv)	Default rates increase 5 percentage points (corporate)	-0.002	0.003
(v)	Default rates increase 5 percentage points (ABS)	-0.019	0.021

Appendix A Simple theoretical framework

In this section, we formally explain the relationship between IC, flows, and precautionary cash. We consider flows standardized by the NAV of the hedge fund, such that the flows of hedge fund i in quarter t are given by

$$F_{it} = \frac{F_{it}^{\$}}{NAV_{it-1}}, \tag{11}$$

where $F_{it}^{\$}$ are the flows in dollars. We decompose F_{it} into two orthogonal components,

$$F_{it} = F_{it}^{F} + F_{it}^{L}, \tag{12}$$

where F_{it}^{F} are the "fundamental flows" driven by the fundamentals of the hedge fund, such as past performance, and F_{it}^{L} are the "liquidity flows" caused by liquidity shocks to the investors of hedge fund i. The distinction between the two types of flows is related to the Greenwood and Thesmar (2011) model of concentrated ownership of stocks.[26]

We write the liquidity flows F_{it}^{L} of a hedge fund with K investors as

$$F_{it+1}^{L} = W_{it}' L_{it+1}, \tag{13}$$

where $W_{it}' = [w_{i1t}, ..., w_{iKt}]$, with w_{ikt} being the share of hedge fund i held by investor k in quarter t, and $L_{it+1}' = [l_{1t+1}, ..., l_{Kt+1}]$, with l_{kt+1} being the liquidity shock to investor k in quarter $t+1$ as a share of the investor's portfolio. Assuming L_{it+1} is normally distributed with mean $\mathbf{0_{K\times 1}} = [0, ..., 0]'$ and $K \times K$ covariance matrix $\Omega_{L,it+1}$, we can write

$$F_{it+1}^{L} \sim N(\mathbf{0_{K\times 1}}, W_{it}' \Omega_{L,it+1} W_{it}). \tag{14}$$

Taking the variance of l_{kt+1} as σ^2 for all $k \in K$ and the correlation between the liquidity shocks of any two investors as ρ, the equation (14) can be written as

$$\sigma_{L,it+1}^2 = \sigma^2 \Big(\sum_{k=1}^{K} w_{ikt}^2 + \sum_{k=1}^{K} \sum_{j=1,k\neq j}^{K} w_{ikt} w_{ijt} \rho \Big) = \sigma^2 \Big[(1-\rho) \sum_{k=1}^{K} w_{ikt}^2 + \rho \Big]. \tag{15}$$

The summation $\sum_{k=1}^{K} w_{ikt}^2$ is the Herfindahl-Hirschman index (HHI), which is used to measure concentration in a range of applications, for example, industry or wealth concentration. The larger the concentration, the higher the HHI. The derivative of the liquidity flow variance with respect to the HHI is given by

$$\frac{\partial \sigma_{L,it+1}^2}{\partial \sum_{k=1}^{K} w_{ikt}^2} = \sigma^2 (1-\rho). \tag{16}$$

[26] Greenwood and Thesmar (2011) distinguish between liquidity-driven trading, which is defined as trading that occurs because of liquidity shocks to investors who hold the asset, and active rebalancing, which is trading that corresponds to changes in the weight of an asset in the investor's portfolio that are driven by a change in the stock's fundamentals.

If the liquidity shocks to investors are not perfectly correlated, then $\rho < 1$ and a higher concentration of the investor base will lead to a higher liquidity flow variance, $\sigma^2_{L,it}$. If liquidity shocks to investors are *perfectly* positively correlated, then $\rho = 1$ and the HHI does not affect the liquidity flow variance because diversifying the investor base does not reduce the variance of the liquidity flows to the hedge fund. The assumption $\rho < 1$ is arguably more realistic.[27]

In our empirical analysis, we do not observe the weight of every individual investor of a hedge fund, which prevents us from computing the HHI exactly. However, we observe a comparable concentration measure, the five-investor concentration ratio, which we refer to as the investor concentration (IC) of a fund throughout the paper. In the Online Appendix, we compute lower bounds and upper bounds on the HHI based on the five-investor concentration and the total number of investors of each hedge fund. As a robustness check, we use the upper and lower bound of the HHI instead of IC in the main regression specifications of our empirical analysis and find that our results hold.

A high HHI increases the variance of the liquidity flows and therefore the probability of large outflows. The higher variance also increases the probability of large inflows, but large inflows are clearly less of a concern for a hedge fund manager and do not warrant any precautionary measures. The probability of large outflows is the risk the manager cares about.

A hedge fund manager wants to avoid being forced to sell assets to cover redemptions. First, having to exit an arbitrage trade early can force the hedge fund to realize losses. Second, having to sell illiquid assets quickly can lead to steep price discounts due to price impact. Therefore, the manager has an interest in holding precautionary cash. However, holding more cash comes with an opportunity cost or a liquidity premium. We model the decision of the hedge fund manager regarding how much cash to hold given this trade-off. The timing of the model is the following: in period t the HF manager decides how much cash to hold based on the expected outflows in period $t+1$. The manager of hedge fund i maximizes the following utility function

$$\max_{C_{it}} U = E_t[-\lambda I_{(F_{it+1} < -C_{it})} - C_{it}\psi], \tag{17}$$

where $C_t = C^{\$}_t / NAV_t$, with $C^{\$}$ being the cash in dollars. $\lambda \geq 0$ is the cost when the outflows exceed cash, for example, exiting an arbitrage trade early and realizing losses or selling illiquid assets at steep discounts. $I_{(F_{it+1} < -C_{it})}$ is an indicator variable that takes the value 1 if $F_{it+1} < -C_{it}$ and 0 otherwise. $\psi \geq 0$ is the opportunity cost of holding cash or liquidity premium.

In this simple framework, HHI is not a choice variable of the hedge fund manager for the following reasons. First, changing the HHI results in wide-ranging costs. For example, rejecting the money of a large investor leads to forfeited management fees. Forcing a large investor out of the fund can lead to relationship and reputational costs. Trying to attract small investors involves

[27]There is evidence that retail investors of mutual funds and pension funds often exhibit correlated trading patterns because of financial advisors' recommendations (see, for example, Dahlquist, Martinez, and Soderlind (2017) and Da, Larrain, Sialm, and Tessada (2018)). However, such correlated trading patterns are likely less pronounced for hedge fund investors, because they are thought to be more sophisticated than retail investors.

search costs and potential regulatory costs, through higher disclosure requirements and forfeited performance fees, due to rules such as those in SEC Regulation D governing investor requirements for hedge funds. Second, changing the HHI is arguably substantially slower than adjusting cash; buying or selling securities is likely faster than attracting new investors to a fund. The fact that we empirically observe a large variation in IC values in our dataset and that IC values for a hedge fund are quite persistent supports this assumption, as otherwise hedge funds would actively strive to diversify their investor base and the IC values across different hedge funds would converge. Further, we use an instrumental variable approach in our empirical analysis and find evidence in line with hedge funds choosing cash based on the concentration of the investor base.

To simplify the analysis, we set the expected total flow variance equal to the liquidity flow variance:

$$\sigma^2_{TF,it+1} = \rho\sigma^2 + \sigma^2(1-\rho)HHI_{it}. \tag{18}$$

This implies that the variance of flows driven by fundamentals is zero. However, in our empirical analysis, we control for a range factors other than the concentration of the investor base that can affect the variance of fundamental flows. When using the expected flow variance in equation (18) and the normal distribution of the liquidity flows given in equation (14), the maximization problem of the hedge fund manager becomes

$$\max_{C_{it}} U = -\lambda\Phi\Big(\frac{-C_{it}}{\sqrt{\rho\sigma^2 + \sigma^2(1-\rho)HHI_{it}}}\Big) - C_t\psi, \tag{19}$$

where $\Phi(.)$ is the standard normal cumulative density function (CDF).

The first order condition with respect to C_{it} is

$$\frac{\partial U}{\partial C_{it}} = \lambda\Big(\frac{1}{\sqrt{\rho\sigma^2 + \sigma^2(1-\rho)HHI_{it}}}\Big)\phi\Big(\frac{-C_{it}}{\sqrt{\rho\sigma^2 + \sigma^2(1-\rho)HHI_{it}}}\Big) - \psi = 0. \tag{20}$$

where $\phi(.)$ is the standard normal probability density function (PDF). Solving for C_{it} yields the solution[28]

$$C_{it} = \sqrt{-ln\Big(\frac{\sqrt{2\pi}\psi(\rho\sigma^2 + \sigma^2(1-\rho)HHI_{it})^{1/2}}{\lambda}\Big)2(\rho\sigma^2 + \sigma^2(1-\rho)HHI_{it})}. \tag{21}$$

We illustrate the solution in two figures. Figure A.1 shows the sensitivity of the optimal C_{it} with respect to HHI_{it} for different values of λ. The other parameters are set such that optimal C_{it} are comparable to our empirical values. We set $\rho = 0.06$, $\sigma = 0.03$, and $\psi = 1.5$. The higher the concentration of the investor base, HHI_{it}, the higher the optimal level of cash, C_{it}. Cash reacts more strongly to changes in HHI_{it} for larger values of λ, the cost of outflows exceeding cash (the cost of forced asset sales), as seen by the steeper slopes of the curves.

[28]The solution only exists if $0 < \frac{\sqrt{2\pi}\psi(\rho\sigma^2+\sigma^2(1-\rho)HHI_{it})^{1/2}}{\lambda} < 1$, which holds if ψ (the liquidity premium) is not substantially larger than λ (the cost of forced asset sales). The solution satisfies the second order condition of the maximization problem.

Figure A.2 shows the sensitivity of the probability of flows being less than or equal to -20% with respect to HHI_{it} for different values of ρ, the correlation of investor liquidity shocks. An increase in the concentration of the investor base leads to a higher probability of flows being less than or equal to -20%. For higher values of ρ, the probability of large outflows is greater because diversifying the investor base is less effective.

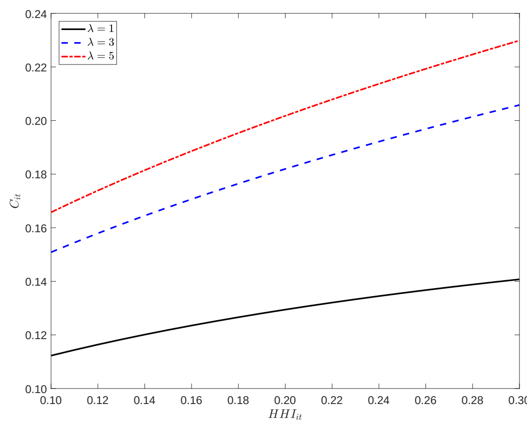


Figure A.1: Sensitivity of cash to HHI

This figure shows the optimal level of C_{it} for a range of HHI_{it} values and for different values of λ. The remaining parameter values are $\rho = 0.06$, $\sigma = 0.03$, and $\psi = 1.5$.

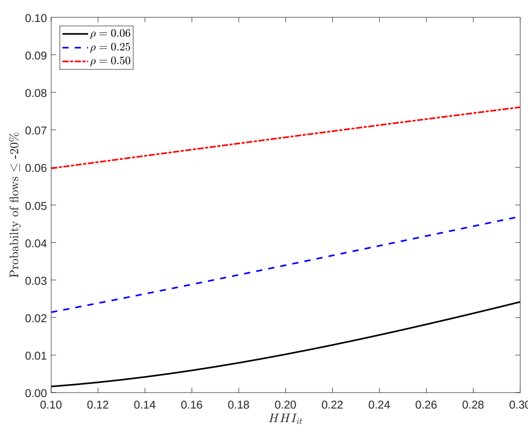


Figure A.2: Sensitivity of probability of large outflows to HHI

This figure shows the probability of flows being less than -20% for a range of HHI_{it} values and for different values of ρ. The remaining parameter values are $\rho = 0.06$, $\sigma = 0.03$, and $\psi = 1.5$.

Appendix B Additional tables

Table B.1: Summary statistics on investor type

This table reports the average and standard deviation of the investor type share of hedge funds. The data are quarterly from 2012:Q4 to 2017:Q4. For each investor type we have 15,586 fund-quarter observations. “US Individuals” and “Non-US Individuals” include trusts owned by the individuals. “Pension plans” and “State or municipal govt. entities” exclude governmental pension plans.

Investor types	Average	Standard deviation
US individuals	15.7	20.4
Non-US individuals	2.9	8.0
Broker-dealers	0.1	1.5
Insurance companies	2.6	5.5
Registered investment companies	1.2	4.5
Private funds	22.0	21.0
Non-profits	14.5	18.8
Pension plans	13.4	18.4
Banking or thrift institutions	1.0	5.4
State or municipal govt. entities	1.3	5.0
State or municipal govt. pension plans	9.0	15.7
Sovereign wealth funds and foreign official inst.	3.3	7.3
Unknown non-US	2.6	11.3
Other	10.5	14.9

Table B.2: Liquidity flows first stage regression

This table reports the coefficient estimates and t-statistics when estimating the model given in equation (1) with the estimation method of Fama and MacBeth (1973). The dependent variable is quarterly hedge fund flows. The independent variables are lagged returns, flows, and share restrictions. The data are quarterly from 2012:Q4 to 2017:Q4. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Flows, F_{it}		
	(1)	(2)
Constant	-0.811***	-1.344***
	(-3.140)	(-5.580)
F_{it-1}	0.342***	0.261***
	(20.418)	(18.007)
r_{it-1}	0.295***	0.240***
	(7.305)	(5.700)
F_{it-2}		0.180***
		(8.935)
r_{it-2}		0.272***
		(8.286)
$ShareRes_{it-1}$	-0.002**	0.008**
	(-2.044)	(2.541)
$ShareRes_{it-2}$		-0.008**
		(-2.662)
$ShareRes_{it-1} \times F_{it-1}$	0.000***	0.000***
	(5.105)	(2.844)
$ShareRes_{it-1} \times r_{it-1}$	-0.001***	-0.001***
	(-10.536)	(-6.862)
$ShareRes_{it-2} \times F_{it-2}$		0.000**
		(2.502)
$ShareRes_{it-2} \times r_{it-2}$		-0.000***
		(-5.479)
Avg adjusted R^2	0.222	0.283

Table B.3: Investor concentration and flow-performance sensitivity (with strategy fixed effects)

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (9). The dependent variable are quarterly flows. The independent variables are lagged IC, flows, returns, return terciles, return quintiles, size, share restriction, and manager stake. The coefficient estimates of the variables lagged flows, size, share restriction, and manager stake are not shown. The data are quarterly from 2012:Q4 to 2017:Q4. Quarter and fund strategy effects are used. The standard errors are clustered by quarter. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Dependent variable: Flows, F_{it}				
	(1)	(2)	(3)	(4)
IC_{it-1}	-0.000	0.000	0.000	-0.004
	(-0.063)	(0.106)	(0.027)	(-0.589)
$r_{it-1} \times IC_{it-1}$	0.001			
	(1.249)			
$r_{it-1} \times I_{r_{it-1}<0} \times IC_{it-1}$		-0.000		
		(-0.466)		
r 2$^{\text{nd}}$ tercile$_{it-1} \times IC_{it-1}$			0.002	
			(0.316)	
r 3$^{\text{rd}}$ tercile$_{it-1} \times IC_{it-1}$			0.000	
			(0.065)	
r 2$^{\text{nd}}$ quintile$_{it-1} \times IC_{it-1}$				0.008
				(0.693)
r 3$^{\text{rd}}$ quintile$_{it-1} \times IC_{it-1}$				0.005
				(0.533)
r 4$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.007
				(0.875)
r 5$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.005
				(0.563)
r_{it-1}	-0.021	-0.068**		
	(-0.619)	(-2.128)		
$r_{it-1} \times I_{r_{it-1}<0}$		0.251***		
		(2.930)		
r 2$^{\text{nd}}$ tercile$_{it-1}$			0.868**	
			(2.716)	
r 3$^{\text{rd}}$ tercile$_{it-1}$			0.924***	
			(3.487)	
r 2$^{\text{nd}}$ quintile$_{it-1}$				0.407
				(0.838)
r 3$^{\text{rd}}$ quintile$_{it-1}$				1.101**
				(2.427)
r 4$^{\text{th}}$ quintile$_{it-1}$				1.239***
				(3.278)
r 5$^{\text{th}}$ quintile$_{it-1}$				0.586
				(1.277)
Time FE	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Observations	13,113	13,113	13,113	13,113
Adjusted R^2	0.225	0.227	0.227	0.228

Appendix C Data appendix

C.1 Hedge fund sample construction

The first Form PF filings for Large Hedge Fund Advisers occurred in 2012:Q2. However, we exclude the 2012:Q2 and 2012:Q3 filings because of data quality concerns. We construct a quarterly hedge fund data sample from 2012:Q4 to 2017:Q4, where reporting dates are assigned to their calendar quarters.

Our analysis focuses on qualifying hedge funds that have to file Form PF at a quarterly frequency. A Qualifying Hedge Fund has a NAV of at least US\$500 million as of the last day in any month in the fiscal quarter immediately preceding the adviser's most recently completed fiscal quarter.[29] The US\$500 million threshold might raise concerns that a lot of hedge funds drop in and out of our sample. However, in our total sample, relatively few observations are from hedge funds that have reporting gaps. Out of a total sample of 15,587 fund-quarter observations, 1,060 observations are from hedge funds with reporting gaps.

We impose several filters to clean the raw Form PF data. As described in Section 3, hedge fund advisers are allowed to file feeder hedge funds separately. Therefore, the raw Form PF data include a few small hedge funds for which several questions in Form PF are unanswered. To avoid including such hedge funds in our sample, we require a hedge funds' NAV to be larger than US\$25 million. Second, we also require the GAV and the gross notional exposure, which is the summation of the long and short values from Form PF's Question 30, to be larger than or equal to the NAV. Third, we delete hedge funds that do not answer Form PF's Question 20, which asks for the investment strategy of the hedge fund, or hedge funds that state that they invest in other funds, as such funds generally file Form PF inconsistently. Also, hedge funds with obvious return outliers, for example, 8888.88, are deleted from our sample. Further, we require that a hedge fund's ratio of unencumbered cash over NAV is between 0 and 1. Lastly, we require that the matching between Form PF and ADV is successful for each hedge fund in the sample. 275 fund-quarter observations could not be matched.

We require that the number of investors in the fund be greater than five and the manager stake be less than or equal to 50%. While our empirical results are robust to removing these data filters, as described in Section 3, we apply these filters for two reasons that improve the precision of our estimates. First, because the IC variable is a five-investor concentration measure, it fails to capture variation in the concentration of the investor base for hedge funds with five or fewer investors, which will all have an IC equal to 100%. Second, theses filters help to exclude family offices and

[29] While the threshold for determining a Qualifying Hedge Fund is in terms of net assets, the thresholds for filing Form PF and for the Large Hedge Fund Adviser classification are on a gross basis. When determining whether a reporting threshold is met, advisers must aggregate private funds, parallel funds, dependent parallel managed accounts, and master-feeder funds. They must also include these items for their related persons that are not separately operated. While the determination of whether a set of funds in a parallel fund structure or master-feeder arrangement constitutes a Qualifying Hedge Fund is on an aggregated basis, advisers are permitted to report fund data either separately or on an aggregated basis. Thus, some funds in our sample have a NAV of less than the Qualifying Hedge Fund threshold of US\$500 million. For additional description of the Form PF hedge fund data, see Flood, Monin, and Bandyopadhyay (2015) and Flood and Monin (2016).

predominantly manager-owned funds in our analysis. In a family office, the investors of the fund know each other and can smooth out liquidity shocks amongst each other. Also, the hedge fund manager likely knows the investors personally, which reduces the asymmetric information about liquidity outflows between investors and the hedge fund manager. The hedge fund manager can learn about the large outflows long before the redemption request is filed, which mitigates the need for holding precautionary cash. If the manager owns the majority of the hedge fund, then the asymmetric information between hedge fund manager and investor base is also clearly limited. Therefore, the mechanism of how IC affects the probability of large liquidity outflows is likely not applicable to these hedge funds. There are 1,131 fund-quarter observations with five or fewer investors. IC is equal to 100 for 4,140 fund-quarter observations. We filter out bad data with IC equal to 0 for 146 fund-quarter observations. The manager stake is greater than 50% for 1,128 fund-quarter observations. A large share of fund-quarter observations that are excluded from our sample violate multiple of these sample restrictions.

C.2 Hedge fund investment strategy classification

The methodology used for classifying a hedge fund's broad strategy is as follows. First, we check the Question 20 description field for the "Other" category to determine if the description can be directly mapped to one of the other broad categories. For example, a description of "Relative Value Fixed Income" is reclassified from "Other" to "Relative Value". Next, the data are normalized so that the sum of each hedge fund's allocation across the 22 sub-categories listed in the form equals 100% of their NAV. These normalized values are then aggregated to the broad strategy categories (credit, equity, event driven, fund of funds, macro, managed futures, multi-strategy, and relative value) and an "other" category. A hedge fund is considered to use a given strategy if 75% or more of its normalized assets are allocated to that strategy. If there is not a strategy to which at least 75% of the normalized assets are allocated, then the fund is classified as a multi-strategy fund. We discard observations from hedge funds identified as "fund of funds" or "managed futures" as these are too few to include given confidentiality restrictions.

Table C.1: Variable Definitions

This table presents definitions of the main variables used in this paper. The first column gives the variable name. The second column includes a short description. The last column gives the reference to the raw data source in Form PF (https://www.sec.gov/about/forms/formpf.pdf) or Form ADV (https://www.sec.gov/about/forms/formadv.pdf). Detailed descriptions and summary statistics of these variables are given in section 3.

Variable Name	Description	Source
IC_{it}	The percentage of fund equity held by the five investors with the largest investments in the fund (the five-investor concentration ratio).	PF Q15
NAV_{it}	Net asset value or the amount of investor equity of the fund.	PF Q9
$Leverage_{it}$	Balance sheet leverage, GAV/NAV.	PF Q8, Q9
g_{it}	Gross hedge fund returns.	PF Q17
r_{it}	Net hedge fund returns (net-of-fees).	PF Q17
F_{it}	Investor flows to a hedge fund, $F_{it} = \frac{NAV_{it} - NAV_{it-1} \times (1+r_{it})}{NAV_{it-1}}$	PF Q9, Q17
$Strategy_{it}$	Investment strategy (Credit, Equity, Event Driven, Macro, Managed Futures, Relative Value, Multi-strategy, or Other). See Appendix C.2.	PF Q20
$Cash_{it}$	Unencumbered cash.	PF Q33
$PortIlliq_{it}$	The average time it would take to liquidate assets in a hedge fund's portfolio (in days).	PF Q32
$FinDur_{it}$	The weighted average maturity of a hedge fund's borrowing (in days).	PF Q46
$ShareRes_{it}$	The weighted average time it would take for investors to withdraw all the fund's NAV (in days).	PF Q50
$HasGates_{it}$	Indicator for whether there are restrictions on investor withdrawals of the fund's NAV. See section 4.5.1.	PF Q49
$MgrStake_{it}$	The percent of NAV owned by the fund's managers or related persons.	ADV Schedule D, Section 7.B.(1), Q14
$NumInvestors_{it}$	The number of investors in the fund.	ADV Schedule D, Section 7.B.(1), Q13
$MinInv_{it}$	The minimum required investment from an investor in the fund.	ADV Schedule D, Section 7.B.(1), Q12