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

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

December 2017

Abstract

We show that when only a few investors own a substantial portion of a hedge fund's net asset value, flow volatility increases because investors' exogenous, idiosyncratic liquidity shocks are not diversified away. Using confidential regulatory filings, we confirm that high investor concentration hedge funds experience more volatile flows. These hedge funds hold more cash and liquid assets, which help absorb large, unexpected outflows. Such funds have to pay a liquidity premium and generate lower risk-adjusted returns. Small investors of a fund are likely unaware of the fund's investor concentration and its associated costs, while larger investors are compensated through the increased economies of scale in monitoring and manager access. We find no evidence that high investor concentration hedge funds have longer lock-up and redemption periods. Investor concentration does not augment the flow response to a hedge fund's performance. These results are robust to the hedge fund strategy, extent of manager ownership, and other controls.

JEL classification: G11, G20, G23.

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

*We thank Turan Bali, Daniel Barth, Murillo Campello, Mark Flood, Byoung Hyoun-Hwang, David Johnson, Andrew Karolyi, Dasol Kim, Andrew Lo, Jose Martinez, David Ng, Andrew Patton, Lubomir Petrasek, Tarun Ramadorai, Stathis Tompaidis, Russ Wermers, Scott Yonker, Peyton Young and seminar participants at the Office of Financial Research, Federal Reserve Board of Governors, SEA Annual Meeting 2017, and Federal Reserve Bank of Boston 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. Email: mathias.s.kruttli@frb.gov. Monin: Office of Financial Research. Email: phillip.monin@ofr.treasury.gov. Watugala: Cornell University and Office of Financial Research. Email: sumudu@cornell.edu.

1 Introduction

Hedge funds implement investment strategies that often require substantial leverage and the trading of illiquid assets. A large literature analyzes hedge funds' investment decisions and portfolio risk. However, unlike the risks inherent in hedge fund investment decisions, the risks inherent in hedge fund investor compositions have received relatively little attention in the academic literature.[1] This paper helps fill this gap. We investigate the risk of high investor concentration (IC) to a hedge fund, where a few large investors own a substantial portion of the hedge fund's net asset value (NAV). Hedge fund investors can be subject to idiosyncratic liquidity shocks.[2] Consequently, a hedge fund can experience inflows and outflows of assets based on liquidity shocks to its investors. These liquidity shocks are independent of the hedge fund's fundamentals like past performance or portfolio holdings.

Recent regulatory changes take into consideration such IC risk to asset managers. For example, the Investment Company Liquidity Risk Management Programs rule of the Securities and Exchange Commission (SEC) requires open-end mutual funds and exchange-traded funds to establish a liquidity risk management program, which, among other factors, specifically asks funds to consider the "fund's shareholder ownership concentration" in its liquidity management, as the fund "could experience considerable cash outflows from redemptions by a single or small number of shareholders".[3] While hedge funds are not covered by this rule, hedge funds are thought to have a more concentrated investor base and invest in more illiquid assets than mutual funds or exchange-traded funds. On the other hand, hedge funds have lock-up and redemptions periods that allow them to manage the liquidity provided to their investors. Therefore, hedge funds pose a particularly rich set of questions when studying the IC risk of asset managers.

We first show that when a hedge fund's investor base is highly concentrated, the effect of idiosyncratic liquidity shocks to individual investors is not dampened through investor diversification. Consequently, a hedge fund with a high IC is more likely to face substantial unexpected flows, such that the volatility of its flows is higher on average than for an equivalent hedge fund with a low IC. This mechanism is related to the stock price fragility mechanism of Greenwood and Thesmar (2011), who find that concentrated ownership makes a stock's return more volatile, and to Gabaix, Gopikrishnan, Plerou, and Stanley (2006) and

[1]There are papers that investigate one specific aspect of hedge fund investor composition, namely the stake of the hedge fund manager (see, for example, Ackermann, McEnally, and Ravenscraft (1999); Brown, Goetzmann, Liang, and Schwarz (2008); and Agarwal, Daniel, and Naik (2009)).

[2]For example, an institutional investor such as a fund of hedge funds could face outflows or a high net worth investor could experience sudden large personal expenditures.

[3]See page 81 of the Securities and Exchange Commission (2016) Investment Company Liquidity Risk Management Programs Rule 22e-4, adopted by the SEC on October 13, 2016.

Ben-David, Franzoni, Moussawi, and Sedunov (2017), who show that the trading of large institutional investors leads to more volatile stock returns. In the case of a hedge fund, the mechanism is similar as non-fundamental shifts in the demand of large investors lead to volatile flows. However, the main analysis presented in this paper differs in that the asset in question, the hedge fund, can take into account the risk of non-fundamental shifts in investor demand and adjust, for example, by holding more precautionary cash. These cash holdings would allow the hedge fund to accommodate large outflows without being forced to sell assets.

We test these two hypotheses empirically. To measure the IC of a hedge fund, we use the confidential Form PF filings data reported to the SEC. Beginning in 2012, large hedge funds are required to report the proportion of the fund's equity owned by the top five investors on Form PF, which was adopted as part of the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010.[4] We use this five-investor concentration ratio as our measure of IC and estimate whether IC predicts a hedge funds' flow volatility. We find that the data support our hypothesis that hedge funds with a high IC experience significantly more volatile flows.

Consistent with our main prediction, we find that high-IC hedge funds hold more cash relative to low-IC hedge funds to account for the higher likelihood of large outflows. The differences in cash holdings between low-IC and high-IC hedge funds are economically significant. The average IC in our sample is 50%. A one standard deviation (22 percentage points) increase in IC is associated with a 3.3 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 16.1% and 7.5% of NAV, respectively. We find that this adjustment in cash holdings ensures that there is no significant difference in the probability of quarterly outflows exceeding cash for low-IC and high-IC hedge funds. We also examine whether high-IC hedge funds not only hold more cash but also hold more liquid assets in general, and our empirical tests confirm this hypothesis. Further, we show that the effect of IC on precautionary cash exists not only in the cross-section of hedge funds, but also when estimating a predictive model, where changes in IC predict changes in cash.

Investors deciding idiosyncratically whether to invest in or withdraw from a hedge fund clearly affect the concentration of a hedge fund's investor base. One might also consider whether hedge funds try to steer the concentration of their investor base. However, several measures in our data indicate that this effect is limited. First, the share restrictions (which include lock-up, redemption, and redemption notice periods and are chosen at the incep-

[4] Hedge funds report the proportion of the fund's equity held by the top five investors in response to Question 15 of the Form PF: https://www.sec.gov/about/forms/formpf.pdf.

tion of the hedge fund (see Agarwal, Daniel, and Naik (2009))) show no positive correlation with IC. Therefore, it does not appear that hedge funds with short share restrictions, and thus, lower ability to manage redemptions, try to keep their IC low. Controlling for share restrictions in our regression specifications does not affect the significance of IC. Second, there exists substantial variation in terms of cash and portfolio illiquidity across hedge fund strategies, but IC is not affected by the strategy of the hedge fund, indicating that hedge funds do not target IC based on their strategy. We include strategy fixed effects in our regression specifications and find that the effect of IC remains statistically and economically significant. Third, a tool with which hedge funds could control IC to some degree is the minimum investment level of an investor, but the correlation between minimum investment and IC is close to zero. This low correlation indicates that hedge fund managers who set a low minimum investment requirement are generally not attempting to diversify their investor base by allowing small investments instead of large investments, and end up accepting investments that are substantially larger than the required minimum. Management fees are a likely driver behind the reluctance of hedge fund managers to steer IC, because they incentivize hedge funds to accept large investments.

In addition to share restrictions, investment strategy, and minimum investment, the effect of IC on flow volatility and precautionary cash is robust to the inclusion of financing constraints, leverage, size, flows, and performance of the hedge fund, as controls in our regression specifications. Also, the documented effect of IC is not driven by hedge funds with a very small total number of investors such as certain family offices.[5] We control for manager stake to ensure that the larger precautionary cash holdings are not driven by the manager owning a large share of the fund and being more risk averse because of the personal exposure to the fund's performance.[6] Also, the IC effect holds 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. This result suggests that the mechanism is valid for both investor types as both can be subject to idiosyncratic liquidity shocks.

We expect that the larger portfolio share of precautionary cash and the higher aggregate portfolio liquidity of high-IC hedge funds have an effect on their risk-adjusted returns. Because high-IC hedge funds have to pay a liquidity premium to hold cash, they are expected to generate lower 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

[5]For details on which family offices have to register with the SEC, see rule 202(a)(11)(G)-1, which was adopted in 2011. A summary of the rule can be found here: https://www.sec.gov/rules/final/2011/ia-3220-secg.htm.

[6]We obtain information on manager ownership of the hedge funds from their Form ADV filings, which are used by investment advisers to register with the SEC: https://www.sec.gov/files/formadv.pdf.

(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 IC is associated with a decrease in a hedge fund's annualized unlevered and levered risk-adjusted returns of 90 and 130 basis points, respectively.

Our findings have several important implications. First, whether hedge funds account for the diversification of their investor base by holding more precautionary cash is an important consideration for the stability of financial markets. An individual hedge fund selling assets in a fire sale can have a contagious effect on other hedge funds and a widespread impact on asset markets. Hedge funds often have large overlaps in their portfolios, and the fire sales of one hedge fund can depress asset prices and lead to losses for other hedge funds with similar portfolios. 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 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 IC risk into account and hold more liquid assets, which mitigates the risk of fire sales and contagious losses. This result is important for policymakers wary of the potential of hedge funds to spark widespread market instability.

Second, our findings are important for hedge fund investors who want to efficiently allocate their portfolios. If they know that hedge funds adjust their cash holdings and portfolio liquidity based on IC, hedge fund investors may wish to request information on IC before investing in a fund. Obtaining information on IC is particularly important for investors with a relatively small share of the hedge fund's NAV. First, while an investor with a large share can, to some extent, infer IC based on the knowledge of her investment and the hedge fund's NAV, an investor with a small share can infer little about a hedge fund's IC. Second, being one of the largest investors of a hedge fund likely comes with certain advantages, such as better access to the hedge fund manager which reduces monitoring costs. These advantages can compensate the large investors for the lower risk-adjusted returns that the hedge fund generates. However, investors that own only a small share of the hedge fund's NAV lack these advantages, but still receive lower risk-adjusted returns if the hedge fund has a high IC.

This paper contributes to the hedge fund literature by extending 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 (see, for example, Fung and Hsieh (1997, 2001, 2004); Agarwal and

Naik (2004); Bollen and Whaley (2009); Patton and Ramadorai (2013); Bali, Brown, and Caglayan (2012); Buraschi, Kosowski, and Trojani (2014); and Agarwal, Arisoy, and Naik (2016)). In contrast, we analyze the risk posed by unexpected outflows due to liquidity shocks to investors. Our paper also complements existing work on how hedge fund portfolio allocations are driven by investor flows. For example, Teo (2011) analyzes the performance of hedge funds with short share restrictions and finds that hedge funds with high net inflows significantly outperform hedge funds with low net inflows through exposure to liquidity risk. Ben-David, Franzoni, and Moussawi (2012) show that during the financial crisis of 2007-2009, hedge funds sold about 29% of their portfolios due to investor redemptions and portfolio margin calls. We present a novel perspective on the relationship between flows and portfolio allocation, and show that simply an increase in the likelihood of a large outflow—without the outflow necessarily materializing—can affect portfolio allocation.

Further, this paper adds to existing work on hedge funds and their exposure to liquidity risk by documenting that the investor composition can have an effect on the liquidity risk that hedge funds are willing to take. There are several papers that analyze the liquidity risk of hedge funds. For example, Getmansky, Lo, and Makarov (2004) show that a high autocorrelation in reported hedge fund returns proxies for illiquid portfolio holdings, and Kruttli, Patton, and Ramadorai (2015) find that a hedge fund illiquidity index based on the return correlation averaged across the hedge fund universe has strong predictive power for asset returns. Aragon (2007) and Agarwal, Daniel, and Naik (2009) report that longer share restrictions allow hedge funds to generate higher risk-adjusted returns, likely due to investments in more illiquid assets. Sadka (2010) shows that liquidity risk is an important determinant in the cross-section of hedge fund returns, and hedge funds that take on substantial liquidity risk outperform hedge funds with low liquidity risk. Cao, Chen, Liang, and Lo (2013) test if hedge funds can time market liquidity by adjusting their market exposure, and their results confirm that some hedge funds are highly skilled at timing liquidity. Aiken, Clifford, and Ellis (2015) analyze the portfolio liquidity risk after hedge funds impose discretionary liquidity restrictions and find that these hedge funds surprisingly continue to sell illiquid stocks. Agarwal, Aragon, and Shi (2017) show that funds of hedge funds that engage in more liquidity transformation have greater incentives to attract capital and perform worse during crisis periods. Further, researchers have investigated if hedge funds provide liquidity in asset markets (see, for example, Jylha, Rinne, and Suominen (2014) and Franzoni and Plazzi (2015)) and find that hedge funds generally stop providing liquidity when funding conditions tighten.

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 (2015)). Generally, this literature investigates whether a negative performance leads to investor outflows. Our IC mechanism is distinctly different, as the investor liquidity shocks considered in this paper are independent of the performance of the hedge fund. Even if a hedge fund is performing well, investors can experience an idiosyncratic liquidity shock for exogenous reasons and redeem their investments. The IC mechanism is more closely related to papers in the mutual fund literature that show how investor flows can affect performance (see, for example, Edelen (1999)) and portfolio allocations (see, for example, Chordia (1996)).

However, IC could affect the sensitivity of flows to past performance and consequently affect precautionary cash holdings through this channel. On the one hand, large hedge fund investors potentially internalize the impact of their redemptions on the hedge fund's performance and refrain from redeeming investments when the hedge fund performs poorly, making flows less sensitive to past performance and reducing the need for precautionary cash. On the other hand, the resources of large hedge fund investors might allow them to monitor their investments more closely, which makes the flows more sensitive to the hedge fund's performance and increases the need for precautionary cash. The existing literature suggests that both mechanisms could be present. Chen, Goldstein, and Jiang (2010) find a stronger flow-performance sensitivity for mutual funds with more illiquid assets, but this effect disappears for mutual funds held by institutional investors, likely because institutional investors are larger and internalize the impact of their redemptions on the mutual fund. However, Schmidt, Timmermann, and Wermers (2016) find that large institutional investors were more likely to run on money market funds than smaller institutional or retail investors around the Lehman Brothers collapse in September 2008, likely because the largest institutional investors have the resources to closely monitor their investments and react more quickly to the distress of a money market fund.[7]

For hedge funds, investor type is a less informative measure for the size of investors, because the typical hedge fund investor base is composed of institutional investors and high net worth individual investors, and both investor types are generally large. Form PF provides us with a direct estimate of the investor base concentration, IC, so we do not need to rely on investor type as a proxy for investment size. We use these regulatory data to test the effect of

[7] Ben-David, Franzoni, and Moussawi (2012) come to a similar conclusion in their analysis of hedge fund flows during the financial crisis of 2007-2009. They show that hedge funds held predominately by institutional investors had larger outflows than hedge funds held by individual investors, likely because institutional hedge fund investors monitor their investments more closely because of career concerns and greater sophistication. Their analysis and conclusions are on investor type, not on investor size.

a concentrated investor base, or high IC, on a hedge fund's flow-performance sensitivity, and we find that there is no difference between the flow-performance sensitivity of low- and high-IC hedge funds. Therefore, our results suggest that the two mechanisms, internalizing the price impact and better monitoring, likely cancel each other for large hedge fund investors.

The remainder of the paper has the following structure. We describe the theoretical relationship between investor concentration and flows in Section 2. Section 3 discusses the data and summary statistics. Section 4 reports the main results of our empirical analysis, and Section 5 concludes.

2 Hedge fund investor concentration, flows, and precautionary cash

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

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

where $F_{it}^{\$}$ are the flows in dollars. We decompose F_{it} into two orthogonal components, which are flows driven by fundamentals and flows driven by liquidity shocks:

$$F_{it} = F_{it}^{F} + F_{it}^{L}, \tag{2}$$

where F_{it}^{F} are the flows driven by the fundamentals of the hedge fund, such as past performance, and F_{it}^{L} are the 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.[8]

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

$$F_{it}^{L} = W_{it}' L_{it}, \tag{3}$$

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 $L_{it}' = [l_{1t}, ..., l_{Kt}]$, with l_{kt} being the liquidity shock to investor k in quarter t

[8] 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.

as a share of the investor's portfolio. The variance of F_{it}^L is then given by

$$Var(F_{it}^L) = \sigma_{L,it}^2 = W_{it}' Var(L_{it}) W_{it} = W_{it}' \Omega_{L,i} W_{it}, \tag{4}$$

where $\Omega_{L,i}$ is the $K \times K$ covariance matrix of the liquidity shocks of hedge fund i's investors.

When we assume that the variance of l_{kt} is equal to σ^2 for all $k \in K$ and the correlation between the liquidity shocks of two investors is ρ, then equation (4) can be written as

$$\sigma_{L,it}^2 = \sigma^2 \Big(\sum_{k=1}^{K} w_{ikt}^2 + \sum_{k=1}^{K} \sum_{j=1,k\neq j}^{K} w_{ikt} w_{ijt} \rho \Big) = \sigma^2 \Big[(1-\rho) \sum_{k=1}^{K} w_{ikt}^2 + \rho \Big], \tag{5}$$

where the second equality holds because $1 - \sum_{k=1}^{K} w_{ikt}^2 = \sum_{k=1}^{K} \sum_{j=1,k\neq j}^{K} w_{ikt} w_{ijt}$. 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

$$\frac{\partial \sigma_{L,it}^2}{\partial \sum_{k=1}^{K} w_{ikt}^2} = \sigma^2 (1-\rho). \tag{6}$$

Assuming $\rho < 1$, that is, the liquidity shocks to investors are not perfectly correlated, then a higher concentration of the investor base will lead to a higher liquidity flow volatility, $\sigma_{L,it}^2$. If $\rho = 1$, that is, if liquidity shocks to investors are perfectly positively correlated, then the HHI does not affect the liquidity flow volatility because diversifying the investor base does to reduce the volatility of the liquidity flows to the hedge fund. The assumption $\rho < 1$ is arguably more realistic.[9]

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 different concentration measure, namely the five-investor concentration ratio. In Appendix B, 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. These results can also be found in Appendix B.

[9] 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, Da, Larrain, Sialm, and Tessada (2017) and Dahlquist, Martinez, and Soderlind (2017)). However, such correlated trading patterns are likely less pronounced for hedge fund investors, because they are thought to be more sophisticated than retail investors.

The relationship between volatile flows due to high IC and the portfolio liquidity of a hedge fund is straightforward to discern. When a hedge fund manager wants to hold enough cash to cover potential outflows $x\%$ of the time, a higher flow variance σ_{it}^2 will require more precautionary cash holdings. For example, if we assume that F_{it}^L is normally distributed with mean zero and F_{it}^F is set to zero, then the cash required to cover outflows 95% of the time is $\sigma_{L,it} \times NAV \times 1.645$, which is increasing in $\sigma_{L,it}$. Thus, a hedge fund with a higher IC and consequently, a higher $\sigma_{L,it}$, will hold more cash than a low-IC hedge fund, assuming other relevant hedge fund characteristics are equal—for example, share restrictions, size, and financing liquidity.

This link between flow volatility and the portfolio allocation of asset managers is also documented in other papers. For example, Chordia (1996) shows theoretically and empirically that mutual funds with uncertain redemptions hold more cash. Chernenko and Sunderam (2016) find that mutual funds with a higher flow volatility engage in less liquidity transformation. However, to our knowledge, no other paper investigates IC as a driver of flow volatility and its implications for precautionary cash holdings.

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 who manage at least US$150 million in private funds, such as hedge funds and private equity funds.[10] Small private fund advisers file annually, while large advisers file quarterly and report more detailed information. For example, small advisers report in Section 1 of Form PF on an annual basis the gross and net asset values, total borrowings, investor concentration and composition, monthly returns, strategies, and top counterparties of each hedge fund that they advise. However, Large Hedge Fund Advisers, defined as those with at least US$1.5 billion in assets managed under all of their hedge funds, provide this information on a quarterly basis for each hedge fund they advise. In addition, these advisers report more detailed information in Section 2b about each of their so-called Qualifying Hedge Funds, such as cash holdings, portfolio and funding liquidity, asset class exposures, collateral posted, and risk metrics. 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

[10] Form PF data are confidential. The Office of Financial Research has access to the data through an agreement with the SEC. For additional description of the Form PF hedge fund data see Flood, Monin, and Bandyopadhyay (2015) and Flood and Monin (2016).

fiscal quarter.[11] Our paper uses data on Qualifying Hedge Funds since certain variables crucial to our analysis, particularly those relating to cash holdings and portfolio liquidity, are only reported by these funds.[12]

3.1 Summary statistics

Our sample period runs from 2012:Q4 to 2016:Q4, inclusive, where reporting dates are assigned to their calendar quarters. Details on the sample construction are included in Section C.1 of Appendix C. Summary statistics are reported in Table 1.

Panel A of Table 1 shows the number of observations (hedge fund-quarters), average, standard deviation, and 25$^{\text{th}}$, 50$^{\text{th}}$, and 75$^{\text{th}}$ percentiles for several quarterly variables.

The first variable is our variable of interest, IC, from Question 15 of Form PF, which asks hedge funds for the 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 value across all of the hedge funds in our sample is 50%, and the median is 47%. These values show that unlike mutual funds, hedge funds often depend on a few large investors. The 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles are 33% and 64%, which shows the measure varies considerably across the hedge funds in our sample. In Appendix B, we analyze how the five-investor concentration ratio relates to the HHI.

The next three variables are gross returns, net after fee returns, NAV, and flows. The returns are reported in Question 17 of Form PF. The average gross return is 2.5% with a large standard deviation of 6.5%. For the net returns, the average is 1.8% with a standard deviation of 5.7%. The average hedge fund size is US$2.2 billion with the median being US$1.2 billion, which indicates the presence of some very large hedge funds in our sample. We compute the flows for quarter t and hedge fund i as

$$F_{it} = \frac{NAV_{it} - NAV_{it-1} \times (1 + r_{it})}{NAV_{it-1}}, \tag{7}$$

where r_{it} is the return net of fees. The flows are winsorized at the 5% level. On average

[11] 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.

[12] 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.

[13] Form PF Question 15 asks for the "beneficial owners", that is, the investors, and not the advisors or managers of the hedge fund.

Table 1: Summary statistics

Panel A reports summary statistics for a range of hedge fund variables. The data are quarterly from 2012:Q4 to 2016:Q4, and the "Number of observations" column reports the number of hedge fund-quarter observations. Panel B shows the number of hedge funds for which we have data for the given number of quarters. Panel C reports the number of observations and the averages of all the variables for each hedge fund strategy. "ND" denotes that the number of observations is too small and showing the value would violate confidentially restrictions.

Panel A: Hedge fund variables

	Number of observations	Average	Stand. dev.	25%	50%	75%
IC (%)	12,639	50.0	21.9	33.0	47.0	64.0
Gross returns (%)	12,639	2.5	6.5	-0.5	2.2	5.1
Net returns (%)	12,639	1.8	5.7	-0.7	1.7	4.2
Flows (in %)	12,172	-1.0	8.7	-5.1	-0.5	2.6
NAV (million US$)	12,639	2,182.1	3,385.7	635.7	1,169.9	2,284.3
Unenc. cash (% of NAV)	12,639	16.4	21.7	0.6	7.5	23.1
Port. illiquidity (days)	12,502	50.4	87.5	4.1	13.7	48.0
Share restriction (days)	12,559	163.9	120.8	60.5	142.3	269.7
Fin. duration (days)	10,204	46.7	87.9	0.5	0.5	55.3
Leverage (GAV/NAV)	12,639	1.7	1.9	1.1	1.3	1.7
Manager stake (%)	12,639	7.9	10.0	1.0	4.0	11.0
Number of investors	12,639	183.6	364.1	45.0	98.0	213.0
Min. inv. (million US$)	12,639	3.7	5.6	1.0	1.0	5.0

Panel B: Reporting history in quarters

	All hedge funds	1-3 quart.	4-7 quart.	8-11 quart.	12-15 quart.	16-17 quart.
# of hedge funds	1,355	290	295	226	201	343

Panel C: Summary statistics by strategy

	Credit	Equity	Event Driven	Macro	Mgd. Futures	Multi-strat.	Other	Rel. Value
Number of observations	768	4,740	1,232	636	203	2,218	966	1,876
Avg. IC (%)	48.8	50.5	44.8	50.3	50.9	46.6	52.1	58.6
Avg. gross returns (%)	2.0	2.1	3.5	1.0	1.3	2.1	4.5	2.1
Avg. net returns (%)	1.3	1.5	2.5	0.5	0.6	1.5	3.2	1.5
Avg. flows (in %)	-1.9	0.0	-2.4	-0.7	-0.2	-0.7	-2.3	-1.3
Avg. NAV (million US$)	1,131.9	1,844.4	2,045.0	2,696.8	2,102.8	3,375.5	2,470.2	1,227.5
Avg. unenc. cash (% of NAV)	14.4	10.6	10.7	44.0	61.0	21.8	14.9	16.7
Avg. port. illiquidity (days)	85.1	20.3	77.5	11.8	2.9	55.1	111.0	48.7
Avg. share restriction (days)	209.3	133.9	243.1	93.2	18.9	187.5	184.2	160.7
Avg. fin. duration (days)	95.5	21.2	53.5	16.4	ND	44.2	103.1	52.9
Avg. leverage (GAV/NAV)	1.6	1.5	1.4	3.0	1.2	1.8	1.5	3.1
Avg. manager stake (%)	6.4	9.2	10.3	7.4	1.1	7.9	5.3	6.6
Avg. number of investors	165.2	163.2	211.3	205.2	174.5	276.2	122.3	157.1
Avg. min. inv. (million US$)	3.4	3.6	4.4	3.7	1.7	3.9	4.5	2.3

flows are -1.0% over our sample period.

Further, Panel A lists the unencumbered cash as a percent of NAV. Hedge funds report unencumbered cash in Question 33 of Form PF, and it 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][14]

The average unencumbered cash as a percent of NAV is 16.4%, with a standard deviation of 21.7% and median of 7.5%, which indicates that the range of unencumbered cash holdings across the hedge funds in our sample is large. In the remainder of the paper, we will refer to unencumbered cash normalized by NAV simply as cash.

The next three variables in Panel A are measured in days. The portfolio illiquidity measure is computed based on information from Question 32. Hedge funds have to report what percentage of the portfolio 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. Cash is excluded from Question 32. We compute a weighted average to obtain a measure of portfolio illiquidity. This measure can be interpreted as the average time it takes to liquidate assets in a hedge fund's portfolio. The average portfolio illiquidity measure in our sample is 50 days and the median is 14 days. The financing duration and share restriction variables are similar weighted averages. Financing duration uses data from Question 46, which requires hedge funds to report what percentage of the total financing is committed for particular time durations (for <1, 2-7, 8-30, 31-90, 91-180, 181-365, and >365 days). Share restriction uses data from Question 50, where hedge funds are required to list what percentage of the NAV is locked for particular time horizons (for <1, 2-7, 8-30, 31-90, 91-180, 181-365, and >365 days). The average financing duration is 47 days. The median financing duration is substantially lower at 1 day, which indicates that financing terms are low for most hedge funds, with some exceptions. The average share restriction is 163 days with a median of 142 days. We also compute a leverage metric: the balance sheet leverage

[14] Cash equivalents are defined in the 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.

GAV/NAV, which is the gross asset value (GAV) reported in Question 8 divided by the NAV reported in Question 9. We refer to GAV/NAV as leverage in the remainder of this paper. For most hedge funds in our sample, the leverage measure is between 1 and 2, which is in line with Ang, Gorovyy, and van Inwegen (2011).

The fact that the portfolio illiquidity measure (average time in days to liquidate asset in portfolio) is substantially lower than the share restriction measure is noteworthy, as it suggests that the risk of fire sales may be limited. However, Question 32 in Form PF asks hedge funds to asses how long it would take to liquidate an asset under current market conditions, that is, the market conditions of the quarter for which the hedge fund files Form PF. Therefore, the portfolio illiquidity measure is time-varying, not only because of changes in the hedge fund's portfolio allocation, but also because of changes in the liquidity of asset markets. Our sample period starts in 2012:Q4 and ends in 2016:Q4, thus, our sample covers a period of high market liquidity. For example, the average aggregate equity market liquidity measure of Pastor and Stambaugh (2003) was 4.5 times higher for the 2012:Q4 to 2016:Q4 period than for the period from 1963:Q1 to 2012:Q3.[15] Unlike the portfolio illiquidity measure, changes in the hedge fund's share restrictions are rare (see Agarwal, Daniel, and Naik (2009)). Share restrictions are regulated by a hedge fund's limited partnership agreement, and changing the agreement is possible only if the majority of the limited partners consent. Therefore, one would expect that during a sample period of low market liquidity, the portfolio illiquidity measure is lower than the share restriction measure. However, a change in market liquidity can quickly lead to a situation where the portfolio illiquidity measure is greater than the share restrictions, which introduces fire sale risks.

The final three variables presented in Panel A are manager stake, number of investors, and minimum investment. These three variables are obtained through the matching of Form PF data with the publicly available Form ADV filings of the hedge funds.[16] 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 to avoid including family offices and predominantly manager-owned funds in our analysis.[17] 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

[15] The average aggregate market liquidity for the 2012:Q4 to 2016:Q4 period is -0.007 compared to -0.031 for the period from Q1 1963 to Q3 2012.

[16] We are able to merge 98.2% (14,292 of 14,561) of the fund-date observations in Form PF to Form ADV. Schedule D, Section 7.B.(1), Question 13 of Form ADV asks advisers to report on the number of beneficial owners (number of investors) of the fund. Question 14 from the same section asks advisers to report the percentage of the fund beneficially owned by the adviser or its related persons (manager stake).

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

between investors and the manager and leads to fewer unanticipated investor flows. If the manager owns the majority of the hedge fund, then the majority of the hedge fund's NAV is not subject to investor liquidity shocks. Therefore, the mechanism described in Section 2 of how IC affects flow volatility is likely not applicable to these hedge funds. For the minimum investment we have an average of US$3.7 million, but the median is lower with US$1 million.

Applying the filters discussed previously results in a sample of 1,355 hedge funds from 577 advisors. Panel B of Table 1 splits our cross-section of hedge funds based on how many quarters of data are in the sample for each hedge fund. Hedge funds can enter and leave the sample based on inception and attrition, so there is no survivorship bias. A hedge fund can also exit our sample if the NAV falls below US$500 million.

In Panel C, we show the average of each variable for seven hedge fund investment strategies (Credit, Equity, Event Driven, Macro, Managed Futures, Relative Value, and Multi-strategy) and an "Other" category. The strategy classification is based on data from Question 20. We establish a single broad strategy category for each hedge fund and reporting date as described in the Data Appendix section C.2. We have the most fund-quarter observations for equity hedge funds with 4,740. The second largest strategy is multi-strategy hedge funds with 2,218 observations.

Panel C shows that there is little variation in IC across strategies. The average IC value for all the strategies cluster around 50%, with Event Driven hedge funds having the lowest IC with 45%, Relative Value hedge funds having the highest IC with 59%. However, there are large differences across strategies for share restrictions. Managed Futures hedge funds have average share restrictions of 19 days, but the share restrictions of Event Driven hedge funds are on average 243 days. Further, for cash holdings and portfolio illiquidity, the differences across strategies are also large. For cash, the range is from 11% (Equity and Event Driven) to 61% (Managed Futures).[18] For portfolio illiquidity, Managed Futures hedge funds have the most liquid portfolio (3 days), and Other hedge funds have the least liquid portfolio (111 days).

In Appendix C.3, 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.

[18]The cash holdings of managed futures hedge funds are large, as under Form PF, derivatives on US treasury and agency securities are cash equivalents.

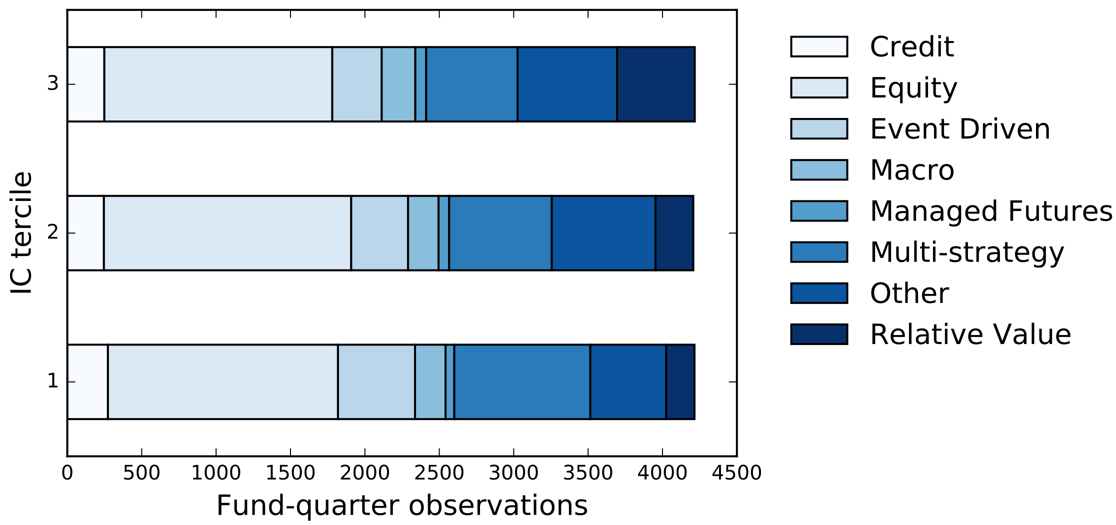


Figure 1: Investor concentration by strategy
This figure shows the number of fund-quarter observations for each strategy and IC tercile. Every quarter, the hedge funds are sorted based on IC. The first tercile contains the hedge funds with the lowest IC values.

3.2 Investor concentration

Compared with other asset managers, a hedge fund's investor base is typically highly concentrated. We show in Panel A of Table 1 that the largest five investors hold on average 50% of the hedge fund's equity. The reason for the concentrated investor base is that hedge funds—unlike, for example, mutual funds or exchange-traded funds—are usually sold not to the general public, but only to "accredited investors," such as institutional investors and high net worth individuals.[19,20] Accredited investors are generally interested in investing large sums of money, which leads to hedge funds having on average a highly concentrated investor base.

To analyze IC in more detail we show the distribution of IC across investment strategies in Figure 1. The number of fund-quarter observations for each strategy and each IC tercile is shown, 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 skewed toward the high-IC tercile.

[19]For a definition of accredited investors see the SEC's Rule 501 of Regulation D (https://www.sec.gov/fast-answers/answers-accredhtm.html).

[20]According to the SEC's Rule 506 of Regulation D, hedge funds can have up to 35 non-accredited investors, but the disclosure documents for these investors must generally be the same as those used in registered offerings. Therefore, having one or more non-accredited investors will increase the hedge fund's disclosure burden substantially. Details on Rule 506 of Regulation D can be found here: https://www.sec.gov/fast-answers/answers-rule506htm.html.

Most hedge funds have minimum investment requirements for their investors despite not being required by regulators. In Panel A of Table 1, we show that the average minimum investment is US$3.7 million and the median is US$1.0 million. If investors tend to simply invest the minimum amount required, we would expect a strong positive correlation between minimum investment and IC. However, the correlation between minimum investment and IC is low. Figure 2 shows the correlation across the entire sample and within subsamples by strategy. Hedge funds pursuing event driven or relative value strategies have the highest correlation between minimum investment and IC, but even for these hedge funds, the correlation is only 0.17. The low correlation suggests that most investors invest substantially more than the required minimum investment, which is likely encouraged by the hedge fund manager, whose compensation depends on a management fee and who thus would have an incentive to accumulate more investor money.

The correlation of the share restriction measure and IC is also small, as depicted in Figure 3, for the entire sample and for individual strategies. The correlations are small in magnitude and often negative. For the total sample, the correlation is -0.16. This result might be unexpected because for a high-IC hedge fund, one way to account for the IC risk would be to have longer share restrictions, in which case, the correlation would be larger in magnitude and positive. However, as discussed in Section 3.1, share restrictions are generally stated in the hedge fund's limited partnership agreement and are difficult to change throughout the hedge fund's life (see Agarwal, Daniel, and Naik (2009)). Therefore, when IC changes because of investors deciding to invest in or withdraw from the hedge fund, a hedge fund cannot simply adjust the share restrictions but would need to adjust the portfolio allocation to account for this risk.

4 Empirical results

In this section, we empirically test two predictions of the simple theoretical framework discussed in Section 2. First, we find that hedge funds with a high IC experience more volatile flows. Second, our main analysis shows that high-IC hedge funds maintain a larger cash share in their portfolios. Further, high-IC hedge funds pay a liquidity premium because they hold more cash and invest more of their portfolio in liquid assets. Our results confirm that risk-adjusted returns are significantly higher for low-IC hedge funds than for high-IC hedge funds. Also, we test whether IC affects the flow-performance sensitivity of a hedge fund and find no significant effect.

Our main empirical measure of IC is the five-investor concentration ratio and not the HHI shown in Section 2. However, in Appendix B, we show how to compute lower and upper

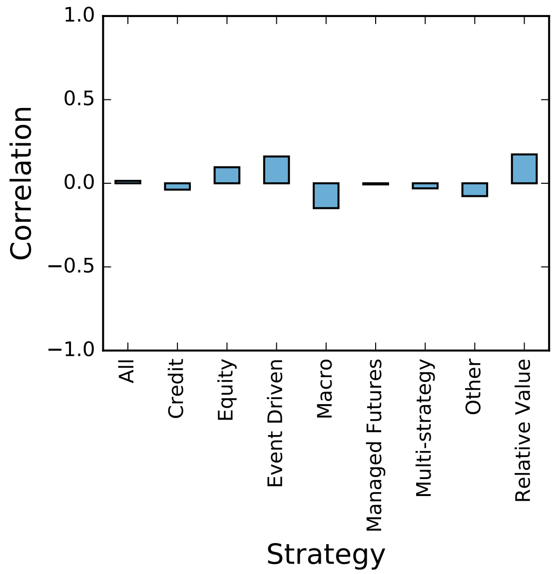


Figure 2: Correlation between investor concentration and minimum investment
This figure shows the correlation between IC and minimum investment at the fund level for the total sample and for the strategy subsamples.

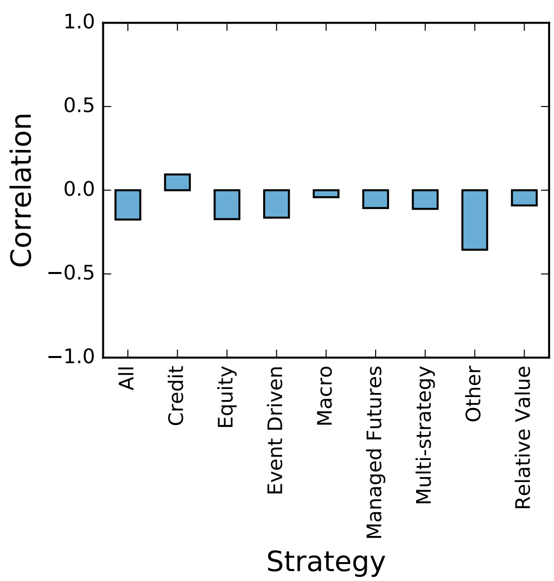


Figure 3: Correlation between investor concentration and share restrictions
This figure shows the correlation between IC and share restrictions at the fund level for the total sample and for the strategy subsamples.

bounds for the HHI based on IC and the total number of investors. To assess the robustness of our empirical results, we re-estimate the regression models discussed in this section with the lower and upper HHI bounds instead of IC and find that our results hold. These results can also be found in Appendix B.

4.1 Investor concentration and flow volatility

To estimate whether high-IC hedge funds experience more volatile flows, we estimate the realized volatility of the flows by computing the standard deviation over a rolling window of quarterly flows. The standard deviation of the flows of hedge fund i over a τ quarter window ranging from t to $t+\tau-1$ is denoted $\hat{\sigma}_{Fi,t:t+\tau-1}$.

Our framework in Section 2 distinguishes between flows due to liquidity shocks to investors, F_{it}^{L}, and flows due to hedge fund fundamentals, F_{it}^{F}. As IC is expected to affect only flows due to liquidity shocks, only the volatility of F_{it}^{L} is expected to be high when IC is high. However, in our empirical analysis, we do not observe F_{it}^{L} alone, but only combined with the total flows F_{it}, where $F_{it} = F_{it}^{L} + F_{it}^{F}$. As F_{it}^{L} and F_{it}^{F} are orthogonal to each other, F_{it} can be used as a noisy proxy for F_{it}^{L}. While we try to account for F_{it}^{F} through control variables, perfectly controlling for it is infeasible. The introduced noise in the dependent variable increases the standard errors of our coefficient estimates and makes observing a significant relationship between flow volatility and IC more difficult.

Figure 4 plots the average hedge fund flow volatility after sorting the hedge funds into terciles based on IC and computing the flow volatility over the subsequent four quarters. Over the whole sample period, the high-IC hedge funds have more volatile flows than the medium- and low-IC hedge funds. Also, the medium-IC hedge funds have more volatile flows than the low-IC hedge funds. This plot suggests that our hypothesis is supported by the data.

To statistically test our hypothesis, we estimate a panel regression model where the dependent variable is $\hat{\sigma}_{Fi,t:t+\tau-1}$. We control for other variables that could affect the flow volatility of a hedge fund. The panel model is given by

$$\hat{\sigma}_{F_{i,t:t+\tau-1}} = \psi + \gamma IC_{it-1} + \phi X_{it-1} + \epsilon_{it}, \tag{8}$$

where X_{it-1} is a column vector that includes lagged control variables for size, return volatility, flow volatility, share restrictions, and manager stake: $log(NAV_{it-1})$, $\hat{\sigma}_{r_{i,t-\tau:t-1}}$, $\hat{\sigma}_{F_{i,t-\tau:t-1}}$, $ShareRes_{it-1}$, and $MgrStake_{it-1}$. The row vector ϕ contains the corresponding regression coefficients. If high IC predicts more volatile flows, the estimate of γ will be significant and positive.

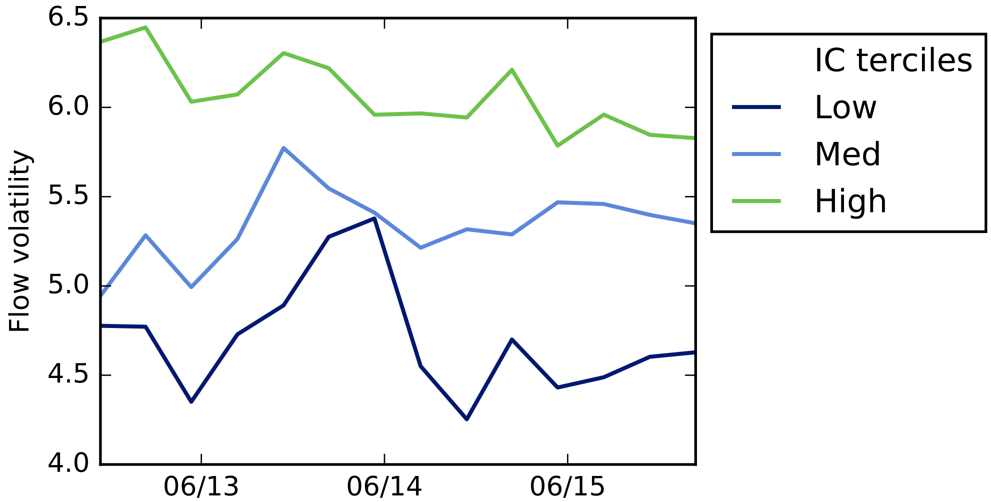


Figure 4: Flow volatility for investor concentration terciles
This figure shows the average hedge fund flow volatility measured over a four-quarter rolling window for three terciles sorted every quarter based on IC.

The estimation results are reported in Table 2 and are strongly supportive of our hypothesis. We use rolling windows of four quarters to compute $\hat{\sigma}_{i,t:t+\tau-1}$. We follow Thompson (2011) and cluster the standard errors by time as we have a large cross-section and a short time series. We include strategy fixed effects for the strategies reported in Panel C of Table 1 and quarter fixed effects.

The coefficient estimates of IC are positive and strongly significant with and without the control variables, which supports our hypothesis that an increase in IC leads to more volatile flows over the subsequent quarters. The results are robust to the inclusion of strategy fixed effects and quarter fixed effects. The results are also economically significant as a one standard deviation (22 percentage points) increase in IC predicts an average increase in flow volatility of 4% to 12% depending on the specification. Of the control variables, size, share restrictions, manager stake, and flow volatility have significant coefficient estimates. Larger hedge funds have less volatile flows, which is potentially related to larger hedge funds being more established and having better investor relations. As expected, longer share restrictions reduce flow volatility because investors cannot withdraw their investments as quickly, and a larger manager stake leads to less volatile flows as managers can smooth out their own withdrawals. Also, flow volatility is positively serially correlated, as shown by the positive and significant coefficient estimate on lagged flow volatility.

These results show that there is a strong relationship between IC and hedge fund flow volatility, as implied by our model in Section 2, despite the added challenge in the empirical analysis of the dependent variable being total flows, which is a noisy proxy for the liquidity-

Table 2: Investor concentration and flow volatility

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (8). The dependent variable is the flow volatility estimated as the standard deviation over a rolling four-quarter window. The independent variables are lagged IC, size, return volatility, flow volatility, share restriction, and manager stake. The data are quarterly from 2012:Q4 to 2016:Q4. Quarter fixed effects and strategy fixed effects are used where indicated. 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: Flow volatility, $\boldsymbol{\hat{\sigma}_{i,t:t+3}}$				
	(1)	(2)	(3)	(4)
IC_{it-1}	0.026***	0.027***	0.008**	0.009**
	(10.35)	(10.92)	(2.79)	(3.08)
$log(NAV_{it-1})$			-0.539***	-0.536***
			(-9.47)	(-9.48)
$\hat{\sigma}_{r_{i,t-4:t-1}}$			-0.000	0.006
			(-0.00)	(0.37)
$\hat{\sigma}_{F_{i,t-4:t-1}}$			0.189***	0.189***
			(18.97)	(19.64)
$ShareRes_{it-1}$			-0.002***	-0.002***
			(-4.39)	(-4.43)
$MgrStake_{it-1}$			-0.021***	-0.021***
			(-4.59)	(-4.45)
Quarter FE	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes
Observations	4681	4681	4681	4681
Adjusted R^2	0.037	0.038	0.110	0.111

driven flows F^L_{it} in our model.

4.2 Investor concentration and cash

The main hypothesis that we test in this paper is whether a hedge fund with a high IC accounts for the increased likelihood of sudden large outflows by holding more cash to be able to absorb these potential outflows. We estimate a panel model that has cash normalized by NAV, $Cash_{it}/NAV_{it}$, as the dependent variable:

$$\frac{Cash_{it}}{NAV_{it}} = \psi + \gamma IC_{it} + \phi Z_{it} + \epsilon_{it}, \tag{9}$$

where cash is defined in Section 3.1. 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 $MgrStake_{it}$. 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 0.15 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 3.3 percentage points in the cash holdings normalized by NAV. This increase is substantial considering that the average cash holdings are 16.1% and the median is 7.5%, as shown in Table 1.

The control variables have coefficient estimates consistent with existing research. For three 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

Table 3: Investor concentration and cash

This table reports the coefficient estimates and t-statistics of the panel regression models given in equations (9) and (10). 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 2016: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, $\boldsymbol{Cash_{it}/NAV_{it}}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IC_{it}	0.087***	0.087***			0.146***	0.146***		
	(2.92)	(2.91)			(4.79)	(4.76)		
IC 2^{nd} tercile$_{it}$			0.139	0.153			1.842*	1.864*
			(0.13)	(0.14)			(1.74)	(1.76)
IC 3^{rd} tercile$_{it}$			3.953***	3.967***			6.687***	6.699***
			(2.74)	(2.76)			(4.71)	(4.72)
$log(NAV_{it})$					3.091***	3.091***	2.820***	2.823***
					(5.45)	(5.43)	(5.22)	(5.22)
F_{it}					-0.056	-0.049	-0.059*	-0.049
					(-1.60)	(-1.32)	(-1.69)	(-1.34)
$ShareRes_{it}$					-0.026***	-0.026***	-0.027***	-0.027***
					(-4.90)	(-4.92)	(-4.95)	(-4.96)
$FinDur_{it}$					0.002	0.002	0.002	0.002
					(0.28)	(0.30)	(0.30)	(0.31)
$Leverage_{it}$					1.390***	1.382***	1.382***	1.374***
					(2.95)	(2.92)	(2.92)	(2.89)
$MgrStake_{it}$					-0.083	-0.082	-0.086	-0.086
					(-1.56)	(-1.54)	(-1.62)	(-1.60)
Quarter FE	No	Yes	No	Yes	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9800	9800	9800	9800	9800	9800	9800	9800
Adjusted R^2	0.170	0.170	0.170	0.170	0.231	0.231	0.228	0.228

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), who shows that longer share restrictions lead to a hedge fund holding a more illiquid portfolio. Interestingly, the economic significance of share restrictions is similar to IC. A one standard deviation decrease in the share restrictions (121 days) leads to an increase of cash holdings normalized by NAV of 3.6 percentage points (compared with an increase of 3.3 percentage points when IC decreases by one standard deviation). Further, leverage has a positive coefficient estimate, suggesting that highly leveraged hedge funds hold more cash. This result can be explained by highly leveraged hedge funds being more exposed to an increase in funding constraints and therefore, holding hold more cash as a precautionary measure. A one standard deviation increase in leverage (1.9) leads to an increase in cash holdings normalize by NAV of 2.6 percentage points, which is again comparable to the effect of IC on cash.

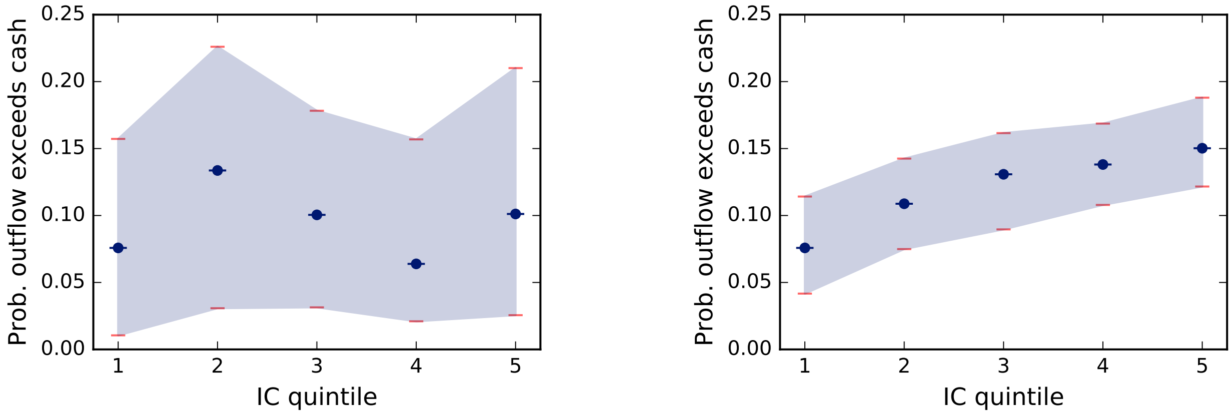
To test whether there exists non-monotonicity in the effect of IC on cash, we estimate a panel regression with hedge funds being sorted into three terciles based on IC in each quarter:

$$\frac{Cash_{it}}{NAV_{it}} = \psi + \sum_{n=2}^{3} I_{i \in n_t} \gamma_n + \phi Z_{it} + \epsilon_{it}. \tag{10}$$

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 quarterly 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 lead to the question whether the increase in cash is "sufficient". While the answer to this question is to some degree subjective, we assume that if hedge funds sufficiently account for IC, then the probability of the 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 with IC, 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 flows, standard deviation of flows, and cash are computed. Assuming that the flows are normally distributed, we compute the probability that the outflows exceed cash within a quarter based on the median



(a) Adjusted cash

(b) Constant cash

Figure 5: Probability of outflows exceeding cash
This figure shows the quarterly probability of outflows exceeding cash for each IC quintile with 95% bootstrapped confidence intervals. The first quintile contains the hedge funds with the lowest IC values. Flows are assumed to be normally distributed. For Plot (a), the probability is computed based on the median quarterly flows, standard deviation of flows, and cash of each quintile. For Plot (b), the probability is again computed based on the median quarterly flows and standard deviation of flows of each quintile, but the median cash level of the first quintile is used to compute the probabilities for all quintiles.

cash and flow values for each IC quintile. We also compute bootstrapped standard errors for these probabilities.

Plot (a) of Figure 5 depicts the probability of 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 outflows exceeding cash is unchanged.

In contrast, Plot (b) of Figure 5 shows the probabilities of outflows exceeding cash when we set the median cash level for each quintile equal to the median cash level of the first quintile. We can see that the probabilities for quintile three, four, and five are significantly higher than for the first quintile. This result suggests that the cash adjustment enables high-IC hedge funds to avoid a significant increase in the likelihood of outflows exceeding cash.

4.2.1 Predictive model

Our main analysis of how IC relates to cash holdings is based on the contemporaneous model shown in equation (9). To analyze if the results are robust to a predictive model specification, we test if changes in a fund's IC predict changes in its cash holdings through the panel model given by

$$\Delta \frac{Cash_{it}}{NAV_{it}} = \psi_i + \gamma \Delta IC_{it-1} + \phi_1 F_{it-1} + \phi_2 r_{it-1} + \phi_3 \Delta Z_{it-1} + \phi_4 \frac{Cash_{it-1}}{NAV_{it-1}} + \epsilon_{it}, \tag{11}$$

where $\Delta IC_{it-1} = IC_{it-1} - IC_{it-2}$. We difference the variables because the dependent variable, $Cash/NAV$, is highly persistent, which would result in a likely unit root for a predictive regression in levels. The control variables included in Z_{it-1} are size, share restrictions, financing duration, leverage, and manager stake. We also include lagged $Cash_{it-1}/NAV_{it-1}$ to account for potential mean reversion.[21] To ensure that our data contain some within fund variation of cash and IC, we estimate the model at a semi-annual frequency. We include quarter and either fund or strategy fixed effects. The standard errors are clustered by time. For the estimates of the predictive model to be in line with the results reported in Table 3, the γ estimate should be positive and significant, such that, an increase in IC predicts an increase in cash.

The results are reported in Table 4. The coefficient estimate of ΔIC_{it-1} is significant and positive with and without controls, indicating that an increase in IC predicts an increase in cash holdings. The result is robust to the inclusion of quarter, strategy, or fund fixed effects. When including fund fixed effects, the statistical significance of ΔIC_{it-1} weakens, but the adjusted R^2 is also reduced indicating that fund fixed effects have little explanatory power. The magnitude of the coefficient estimate is around 0.10-0.15. This estimate is economically significant as it implies that an increase in IC of 10 percentage points predicts an increase in cash holdings of 1.0-1.5 percentage points. This increase in cash holdings is considerable relative to the average and median cash holdings of 16.4% and 7.5%, respectively.

These results show that the effect of IC on cash can also be observed through a predictive model. However, as expected based on the nature of the data sample, these results are statistically weaker than the results for the contemporaneous regression, that is, the statistical significance of the coefficient estimates in Table 3 are higher than in Table 4. The reason is that our sample consists of a large cross-section of hedge funds, but includes at most 16 and generally fewer quarters of time series observations for each hedge fund. Therefore, we observe less time-series variation in IC for an individual hedge fund, and our sample is better suited for utilizing the cross-sectional variation in IC.

4.2.2 Portfolio illiquidity

To further assess the robustness of our main result that high-IC hedge funds account for the risk of IC by holding more cash, we also use portfolio illiquidity as the dependent variable

[21] This specification is similar to the leverage regression model of Ang, Gorovyy, and van Inwegen (2011).

Table 4: Predictive model of investor concentration and cash

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (11). The dependent variable is the change in cash normalized by NAV. The independent variables are lagged flows, returns, and cash, and lagged changes in IC, size, share restriction, financing duration, leverage, and manager stake. The data are semi-annual from 2012:Q4 to 2016:Q4. Fund, time, 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: Changes in cash normalized by NAV, $\boldsymbol{\Delta(Cash_{it}/NAV_{it})}$				
	(1)	(2)	(3)	(4)
ΔIC_{it-1}	0.155**	0.154**	0.118*	0.155**
	(2.65)	(2.71)	(2.43)	(2.81)
$\Delta log(NAV)_{it-1}$		1.762	2.498	1.621
		(1.22)	(1.33)	(1.19)
r_{it-1}		-0.023	0.013	-0.018
		(-0.74)	(0.31)	(-0.57)
F_{it-1}		-0.019	-0.003	-0.020
		(-0.87)	(-0.11)	(-0.90)
$\Delta ShareRes_{it-1}$		0.012**	0.012*	0.012**
		(3.18)	(2.36)	(3.23)
$\Delta FinDur_{it-1}$		-0.001	-0.000	-0.001
		(-0.19)	(-0.07)	(-0.14)
$\Delta Leverage_{it-1}$		-0.529	-0.751	-0.552
		(-1.20)	(-1.97)	(-1.22)
$\Delta MgrStake_{it-1}$		-0.139	-0.113**	-0.142
		(-1.70)	(-3.42)	(-1.73)
$Cash_{it-1}/NAV_{it-1}$			0.050	0.027
			(0.61)	(0.99)
Quarter FE	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	No	Yes
Fund FE	No	No	Yes	No
Observations	2769	2769	2769	2769
Adjusted R^2	0.038	0.041	0.020	0.042

instead of cash and estimate the models given in equations (9) and (10). Using portfolio illiquidity as the dependent variable allows us to measure if high-IC hedge funds also hold more liquid assets in addition to holding more cash. Hedge funds have to report on Form PF what percentage of the portfolio (excluding cash) 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 3.1, and use it as a measure of portfolio illiquidity. If hedge funds increase portfolio illiquidity when IC is low, we expect a significant and negative γ estimate. 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 its actual portfolio liquidity and introduce measurement error.

The results are reported in Table 5 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. When using tercile dummy variables sorted on IC, the results also support the hypothesis that high-IC hedge funds hold more liquid portfolios. The results are again economically significant. The γ estimate is around -0.40 when including control variables. This coefficient estimate implies that a one standard deviation (22 percentage points) increase in the top five investors' holdings of the hedge fund is associated with a decrease in the hedge fund's portfolio illiquidity by 8.8 days. Considering that the average and median portfolio illiquidity measures are 50.4 and 13.7 days, respectively, this decrease is substantial.

The coefficient estimates of the control variables are again as expected. Size has a negative coefficient estimate, which shows that larger hedge funds, similar to larger mutual funds, invest in more liquid assets. The coefficient estimate of share restrictions is positive, which shows that hedge funds that grant investors less favorable (longer) redemption terms invest in more illiquid assets. Financing duration also has a positive coefficient estimate. This estimate indicates that when counterparties grant a hedge fund long financing terms, the hedge fund tends to hold a more illiquid portfolio. On the other hand, the coefficient estimate of leverage is negative, suggesting that highly levered hedge funds tend to hold more liquid assets. The negative and strongly significant coefficient estimate on the manager stake variable indicates that hedge funds in which the manager is more invested engage in less liquidity risk-taking and hold a more liquid portfolio.

Table 5: Investor concentration and portfolio illiquidity

This table reports the coefficient estimates and t-statistics of the panel regression models given in equations (9) and (10), 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 2016: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, $\boldsymbol{PortIlliq_{it}}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
IC_{it}	-0.363***	-0.364***			-0.391***	-0.396***		
	(-3.71)	(-3.71)			(-4.17)	(-4.21)		
IC 2nd tercile$_{it}$			0.939	1.013			-3.373	-3.379
			(0.17)	(0.18)			(-0.77)	(-0.77)
IC 3rd tercile$_{it}$			-14.484***	-14.420***			-14.995***	-14.988***
			(-2.72)	(-2.70)			(-2.98)	(-2.98)
$log(NAV_{it})$					-12.625***	-12.682***	-11.412***	-11.424***
					(-5.70)	(-5.71)	(-5.29)	(-5.28)
F_{it}					0.081	0.110	0.082	0.104
					(0.55)	(0.71)	(0.56)	(0.68)
$ShareRes_{it}$					0.281***	0.282***	0.285***	0.285***
					(12.34)	(12.36)	(12.43)	(12.46)
$FinDur_{it}$					0.263***	0.263***	0.263***	0.263***
					(7.52)	(7.50)	(7.47)	(7.46)
$Leverage_{it}$					-1.607**	-1.624**	-1.587**	-1.603**
					(-2.46)	(-2.47)	(-2.40)	(-2.41)
$MgrStake_{it}$					-0.601***	-0.600***	-0.604***	-0.603***
					(-3.82)	(-3.81)	(-3.77)	(-3.77)
Quarter FE	No	Yes	No	Yes	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9797	9797	9797	9797	9797	9797	9797	9797
Adjusted R^2	0.142	0.142	0.140	0.140	0.437	0.437	0.434	0.434

4.2.3 Investor type

So far, our analysis has focused on differences in IC across hedge funds without differentiating between investor characteristics. One investor characteristic for which Form PF provides data is the investor type. Question 16 of Form PF asks for the percentage of the reporting hedge fund's equity held by individual investors and institutional investors, where individual investors are generally high net worth individuals, as discussed in Section 3.2.[22] The summary stats for these data are reported in Table A.1 of Appendix A. To assess whether the effect of IC on cash is robust to differences in the investor composition type, we estimate the model in equation (9) for hedge funds for which individual investors own more than or equal to 75%, 50%, and 25% of the hedge fund's equity. We also estimate the same model for hedge funds for which institutional investors own more than 75%, 50%, and 25% of the hedge fund's equity.

The results are reported in Table 6. The first three columns show the regression model estimates for hedge funds with an individual investor share of greater than or equal to 75%, 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 across the six subsamples and significant for all except one subsample. The insignificant IC coefficient estimate for the subsample of hedge funds with individual investors owning 75% or more of the equity is likely because of the reduced power due to the small sample size of only 337 fund-quarter observations. The average individual investor share is 20% with a standard deviation of 22%, and the 95$^{\text{th}}$ percentile is 67%. Therefore, there are only a few hedge funds for which individual investors hold 75% or more of the equity.

The coefficient estimates in columns (1) to (3) are not significantly different from the coefficient estimates of the complementary subsamples in columns (4) to (6). These results indicate that the effect of IC on portfolio allocation discussed in this paper 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.

[22]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.

Table 6: 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 (9) estimated for hedge fund subsamples. Columns (1) to (3) show the model estimates for the subsample of hedge funds for which individual investors hold 75%, 50%, and 25% or more of the hedge fund's equity. The complementary subsamples are given in columns (4) to (6). 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 2016: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, $\boldsymbol{Cash_{it}/NAV_{it}}$						
	Individual investor share greater or equal to			Institutional investor share greater than		
	75%	50%	25%	25%	50%	75%
IC_{it}	0.046	0.121*	0.155***	0.151***	0.143***	0.116***
	(0.43)	(1.77)	(3.66)	(4.76)	(4.25)	(2.92)
$log(NAV_{it})$	3.799**	1.934*	2.498***	3.099***	3.100***	2.836***
	(2.52)	(1.76)	(3.17)	(5.15)	(4.89)	(3.79)
F_{it}	-0.428*	-0.264**	-0.097	-0.046	-0.038	-0.049
	(-1.71)	(-2.11)	(-1.43)	(-1.24)	(-1.01)	(-1.14)
$ShareRes_{it}$	0.018	0.001	-0.007	-0.027***	-0.029***	-0.034***
	(1.16)	(0.06)	(-0.88)	(-4.92)	(-5.15)	(-5.35)
$FinDur_{it}$	-0.002	0.026	0.023**	0.001	0.001	-0.002
	(-0.11)	(1.57)	(2.06)	(0.24)	(0.13)	(-0.34)
$Leverage_{it}$	6.490**	-1.143	0.102	1.378***	1.527***	1.513***
	(2.02)	(-1.13)	(0.15)	(2.91)	(3.34)	(3.14)
$MgrStake_{it}$	-0.309**	-0.058	0.046	-0.071	-0.065	-0.094
	(-2.24)	(-0.60)	(0.69)	(-1.27)	(-1.05)	(-1.18)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	337	1126	2714	9463	8674	7086
Adjusted R^2	0.315	0.240	0.247	0.212	0.213	0.209

4.3 Investor concentration and risk-adjusted returns

In the previous section, we discussed how high IC is associated with higher levels of cash and lower portfolio illiquidity. A large literature shows that illiquid assets carry a premium (see, for example, Amihud (2002) and Pastor and Stambaugh (2003)). Consequently, our findings raise the question of whether high-IC hedge funds generate lower risk-adjusted returns because their portfolios are more liquid.

To answer this question, 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

$$r^e_{im} = \alpha_i + \beta_i FH_{im} + \epsilon_{im}, \text{ where } i = 1, 2, ..., N. \tag{12}$$

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

$$\alpha_{im} = r^e_{im} - \hat{\beta}_i FH_{im}. \tag{13}$$

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

$$\alpha_{it} = \psi + \gamma IC_{it-1} + \phi Y_{it-1} + \epsilon_{it}. \tag{14}$$

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}, $ShareRes_{it}$, $FinDur_{it}$, 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 excess returns. When delevering, we divide the excess returns by our leverage measure, GAV/NAV. The coefficient on IC is negative and strongly significant for both the levered and delevered returns when including control variables. 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 (93) basis points. Relative to the average risk-adjusted returns of 55 basis points and 39 basis points for levered and delevered returns across all hedge funds, respectively, these effects are substantial.

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 of the hedge fund manager does not appear to affect the risk-adjusted returns.

The decreasing returns to scale of asset managers at the fund level have been theoretically investigated by Berk and Green (2004). Empirical examinations of decreasing returns to scale have resulted in mixed results. For example, Chen, Hong, Huang, and Kubik (2004) find evidence of decreasing returns to scale for mutual funds, and Fung, Hsieh, Naik, and Ramadorai (2008) report that capital inflows attenuate the ability of fund of hedge funds to deliver positive risk-adjusted returns. However, Pastor, Stambaugh, and Taylor (2015) find significant decreasing returns to scale at the mutual fund industry level, but at the fund level they find only some, but mostly insignificant, evidence of decreasing returns to scale. Our paper analyzes an additional dimension that has not yet been considered in the discussion about asset manager returns to scale. We show that new investor money can have a positive effect on the performance of a hedge fund if it diversifies the investor base and consequently reduces IC. However, if new investor money leads to a more concentrated investor base, the investor money can have a negative effect on performance not only because of negative

Table 7: Investor concentration and risk-adjusted returns

This table reports the coefficient estimates and t-statistics when estimating the model given in equation (14) 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 (13). 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 2016: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, $\boldsymbol{\alpha_{it}}$				
	(1)	(2)	(3)	(4)
IC_{it-1}	-0.004**	-0.006***	-0.003**	-0.004***
	(-2.42)	(-3.56)	(-2.84)	(-3.93)
$log(NAV_{it-1})$	-0.186***	-0.231***	-0.149***	-0.183***
	(-5.08)	(-6.14)	(-6.38)	(-7.70)
F_{it-1}	-0.006**	-0.004*	-0.007***	-0.006**
	(-2.48)	(-1.80)	(-3.10)	(-2.66)
$ShareRes_{it-1}$	0.002***	0.002***	0.002***	0.002***
	(6.97)	(7.99)	(7.40)	(8.09)
$FinDur_{it-1}$	0.006***	0.006***	0.005***	0.005***
	(7.34)	(6.73)	(6.73)	(6.43)
$MgrStake_{it-1}$	-0.003	-0.002	-0.000	0.000
	(-1.28)	(-0.71)	(-0.19)	(0.11)
Strategy FE	No	Yes	No	Yes
Deleveraged	No	No	Yes	Yes
Observations	7488	7488	7488	7488

returns to scale but also because of the high IC.

4.4 Investor concentration and flow-performance sensitivity

The IC risk that we focus on in this paper is concerned with an increased flow volatility due to idiosyncratic liquidity shocks to 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 the investment. 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 also affect the sensitivity of a hedge fund's flows to past performance. 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, which would make flows less sensitive to the hedge fund's performance and reduce the need for precautionary cash holdings. On the other hand, large hedge fund investors might have the resources to monitor their investments more closely, which would make the flows more sensitive to the hedge fund's performance and increase the need for precautionary cash holdings.

A large literature looks at the flow-performance sensitivity of mutual funds (see, for example, Chevalier and Ellison (1997) and Sirri and Tufano (1998)) and hedge funds (see, for example, Li, Zhang, and Zhao (2011) and Getmansky, Liang, Schwarz, and Wermers (2015)). 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 institutional 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. However, Schmidt, Timmermann, and Wermers (2016) find that large institutional investors were more likely to run on money market funds 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.

Ben-David, Franzoni, and Moussawi (2012) analyze if the investor type, institutional versus individual investor, had an effect on the redemptions from hedge funds during the financial crisis of 2007-2009 and find that hedge funds predominantly held by institutional investors experienced larger outflows than hedge funds held by individual investors. The authors explain this finding by institutional investors facing periodic performance evaluations and being more sophisticated, which make them more reactive to market events. The investor size is likely less of a factor behind their results. As discussed in Section 3.2, a hedge fund's investor base, unlike a mutual or money market fund's, generally does not include any retail investors. Therefore, distinguishing between institutional and individual investors in the case of a hedge fund is less informative about the size of the investor, because a hedge fund's individual investors are high net worth individuals whose investment size can be comparable to the investment size of institutional investors.

The IC variable in Form PF provides us with a measure of a hedge fund's investor base concentration without relying on investor type as a proxy for investment size. To test whether IC affects the flow-performance sensitivity of hedge funds, we estimate the panel model given by

$$F_{it} = \psi + \gamma_1 IC_{it-1} + \gamma_2 Performance_{it-1} \times IC_{it-1} + \phi_1 Performance_{it-1} + \phi_2 \delta Z_{it-1} + \epsilon_{it}, \quad (15)$$

where $Performance_{it-1}$ is a measure of the hedge funds' lagged performance. We try four measures of 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}, $ShareRes_{it}$, 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 significant and negative. If a high IC is associated with flows that are more sensitive to performance, then the estimate of γ_2 would be significant and positive.

The results with fund fixed effects are given in Table A.2, and the results with strategy fixed effects are reported in Table 8 in Section A. In line with the existing literature on hedge fund flows, we find evidence that higher returns lead to higher subsequent flows. The results also indicate that there is some persistence in flows, with the coefficient on lagged flows being positive and significant. 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. The two mechanisms, internalizing the price impact and better monitoring, might be canceling each other for large hedge fund investors.

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 (15). 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 2016: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, $\boldsymbol{F_{it}}$				
	(1)	(2)	(3)	(4)
IC_{it-1}	-0.017	-0.018	-0.015	-0.018
	(-0.94)	(-1.04)	(-0.85)	(-0.94)
$r_{it-1} \times IC_{it-1}$	-0.000			
	(-0.67)			
$r_{it-1} \times I_{r_{it-1}<0} \times IC_{it-1}$		-0.001		
		(-0.77)		
r 2$^{\text{nd}}$ tercile$_{it-1} \times IC_{it-1}$			-0.003	
			(-0.48)	
r 3$^{\text{rd}}$ tercile$_{it-1} \times IC_{it-1}$			-0.001	
			(-0.20)	
r 2$^{\text{nd}}$ quintile$_{it-1} \times IC_{it-1}$				-0.003
				(-0.27)
r 3$^{\text{rd}}$ quintile$_{it-1} \times IC_{it-1}$				0.000
				(0.00)
r 4$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.010
				(1.20)
r 5$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				-0.006
				(-0.68)
r_{it-1}	0.109***	0.056*		
	(3.07)	(1.82)		
$r_{it-1} \times I_{r_{it-1}<0}$		0.118		
		(1.72)		
r 2$^{\text{nd}}$ tercile$_{it-1}$			0.689*	
			(2.02)	
r 3$^{\text{rd}}$ tercile$_{it-1}$			1.086***	
			(3.69)	
r 2$^{\text{nd}}$ quintile$_{it-1}$				0.681
				(1.42)
r 3$^{\text{rd}}$ quintile$_{it-1}$				0.727
				(1.74)
r 4$^{\text{th}}$ quintile$_{it-1}$				0.627
				(1.41)
r 5$^{\text{th}}$ quintile$_{it-1}$				1.534***
				(3.58)
Quarter FE	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Observations	10444	10444	10444	10444
Adjusted R^2	0.153	0.153	0.153	0.153

5 Conclusion

We investigate a novel source of hedge fund risk, namely, how diversified hedge funds are with respect to their investors. Using a simple theoretical framework, we show that a hedge fund with only a few large investors, that is, 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 an investor base that is not diversified. 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. We confirm these hypotheses through our empirical analysis using a novel regulatory dataset on hedge funds.

The SEC's Form PF requires hedge funds to report the percentage of their NAV that is held by the five largest investors. We use this five-investor concentration ratio as our empirical measure of IC. First, in line with our prediction, we find that high-IC hedge funds have more volatile flows. Second, high-IC hedge funds hold more precautionary cash, which helps absorb sudden large outflows. Third, high-IC hedge funds generate lower levered and unlevered risk-adjusted returns which is consistent with such funds having to pay a liquidity premium to hold more cash and liquid assets. Further, we show that IC does not affect the flow-performance sensitivity of hedge funds. These results are robust to a variety of controls, including, share restrictions, investment strategy, and 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 investor composition of a hedge fund can also pose a substantial risk, and we analyze how this risk affects the hedge fund's investments. Our main finding that high-IC hedge funds hold more precautionary cash is important for policymakers who assess the financial stability impact of hedge funds. Further, knowing that high-IC hedge funds tend to reduce their exposure to liquidity risk helps hedge fund investors to allocate their portfolios more efficiently.

Appendix A Additional tables

Table A.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 2016:Q4. For each investor type we have 12,638 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	1.6
Insurance companies	2.6	5.4
Registered investment companies	1.3	4.6
Private funds	22.7	21.2
Non-profits	14.1	18.5
Pension plans	13.2	18.2
Banking or thrift institutions	1.0	5.2
State or municipal govt. entities	1.2	4.8
State or municipal govt. pension plans	9.0	15.8
Sovereign wealth funds and foreign official inst.	3.3	7.3
Unknown non-US	2.7	11.4
Other	10.1	14.7

Table A.2: Investor concentration and flow-performance sensitivity (strategy fixed effects)

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (15). 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 2016: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, $\boldsymbol{F_{it}}$				
	(1)	(2)	(3)	(4)
IC_{it-1}	0.002	0.003	0.003	-0.003
	(0.27)	(0.48)	(0.49)	(-0.47)
$r_{it-1} \times IC_{it-1}$	0.000			
	(0.63)			
$r_{it-1} \times I_{r_{it-1}<0} \times IC_{it-1}$		0.000		
		(0.34)		
r 2$^{\text{nd}}$ tercile$_{it-1} \times IC_{it-1}$			-0.001	
			(-0.16)	
r 3$^{\text{rd}}$ tercile$_{it-1} \times IC_{it-1}$			-0.000	
			(-0.08)	
r 2$^{\text{nd}}$ quintile$_{it-1} \times IC_{it-1}$				0.010
				(0.97)
r 3$^{\text{rd}}$ quintile$_{it-1} \times IC_{it-1}$				0.003
				(0.35)
r 4$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.010
				(1.24)
r 5$^{\text{th}}$ quintile$_{it-1} \times IC_{it-1}$				0.004
				(0.36)
r_{it-1}	0.015	-0.046		
	(0.41)	(-1.26)		
$r_{it-1} \times I_{r_{it-1}<0}$		0.199*		
		(2.08)		
r 2$^{\text{nd}}$ tercile$_{it-1}$			1.071**	
			(2.88)	
r 3$^{\text{rd}}$ tercile$_{it-1}$			1.234***	
			(3.66)	
r 2$^{\text{nd}}$ quintile$_{it-1}$				0.285
				(0.57)
r 3$^{\text{rd}}$ quintile$_{it-1}$				1.169**
				(2.48)
r 4$^{\text{th}}$ quintile$_{it-1}$				1.232**
				(2.55)
r 5$^{\text{th}}$ quintile$_{it-1}$				0.932
				(1.72)
Quarter FE	Yes	Yes	Yes	Yes
Strategy FE	Yes	Yes	Yes	Yes
Additional controls	Yes	Yes	Yes	Yes
Observations	10444	10444	10444	10444
Adjusted R^2	0.224	0.226	0.227	0.228

Appendix B Herfindahl-Hirschman Index bounds: estimation methodology and results

Throughout the paper we use the five-investor concentration ratio (IC) as our measure of the investor concentration of a hedge fund. This value is reported directly for each hedge fund on Form PF. In addition to IC, we know the total number of investors (N) in the fund from the Form ADV. Given IC and N, we first estimate the possible range of the Herfindahl-Hirschman Index (HHI). We then examine the robustness of our results using these bounds.

Given a fund's IC and N, the lower and upper bounds for HHI can be computed using quadratic programming techniques. Proofs are available upon request. The lower bound of HHI is given by

$$HHI_{Min} = 5\left(\frac{IC}{5}\right)^2 + (N-5)\left(\frac{100-IC}{N-5}\right)^2.$$

This is an intuitive result, corresponding to the case in which the fund has the most diversified investor base possible: the top five investors each have an equal share of IC and the rest of the $N-5$ investors each hold an equal share in the remaining $100-IC$.

Next, we compute the upper bound for HHI. Let $i_1, i_2, \ldots, i_N$ be the ordered shares of the investors of the fund, so that i_1 is the largest investor share and i_N is the smallest investor share. It is easily seen that at least one of the top 5 investors must hold at least $\frac{IC}{5}$ of the fund, and that it is possible for the top 5 investors to hold equal amounts $\frac{IC}{5}$ in the fund. Thus the maximal possible value for the share of the sixth largest investor is given by $i_6^{Max} = \min\left(\frac{IC}{5}, 100-IC\right)$. Similarly, at least one of the bottom $N-5$ investors must hold a share of at least $\frac{100-IC}{N-5}$ of the fund. Because i_6 holds the largest share of the bottom $N-5$ investors, it follows that the minimal possible value for i_6 is $i_6^{Min} = \frac{100-IC}{N-5}$. For a given IC, N and value of i_6, one can show that the maximum possible HHI is when i_1 has the largest possible share of IC, i.e., when $i_1 = IC - 4i_6$, and $i_7, i_8, \ldots, i_N$ have the largest possible share that is less than or equal to i_6. We can calculate this largest possible HHI for a given IC, N, and i_6. Formally, this is given by

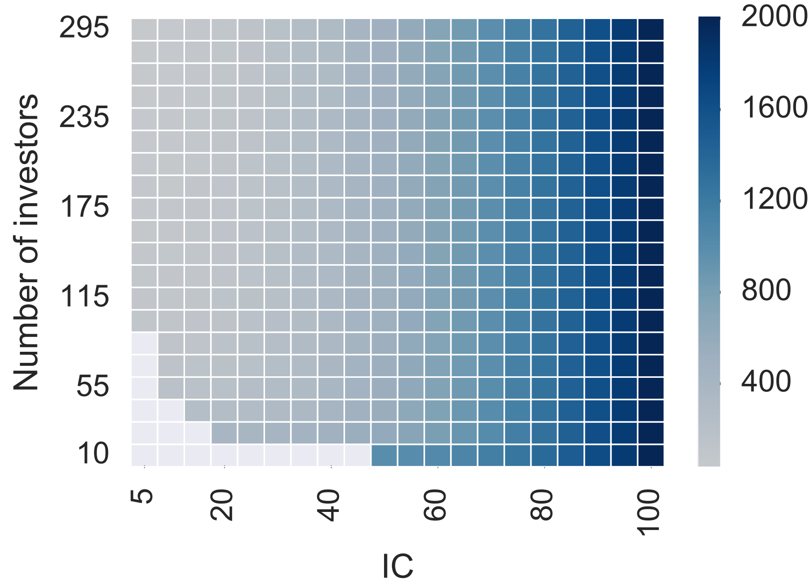
$$h(i_6) = (IC - 4i_6)^2 + 5i_6^2 + \sum_{k=7}^{N} i_k, \quad i_k = \min\left(i_{k-1}, 100 - \sum_{j=1}^{k-1} i_j\right), \; k = 7, \ldots, N.$$

Finally, one can prove that the highest HHI is found at one of the extreme points of i_6. That is,

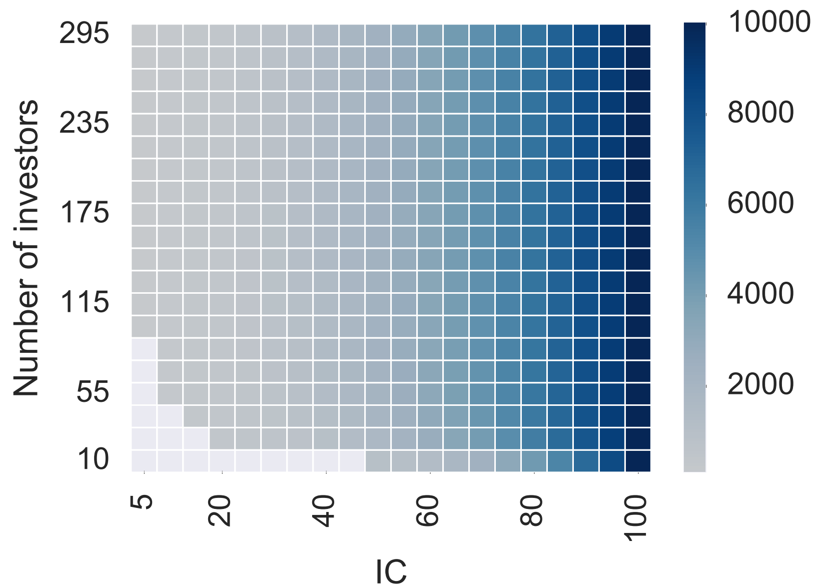
$$HHI_{Max} = \max\left(h(i_6^{Min}), h(i_6^{Max})\right).$$

Figure B.1 show how HHI_{Min} and HHI_{Max} vary with IC and the total number of investors in the hedge fund, N. The gray shaded area in the lower left hand corner correspond to infeasible combinations of IC and N. For lower values of IC, the total number of investors lead to variation in the upper and lower bounds of the HHI. However, when IC is close to a 100%, the total number of investors provides little additional information regarding the concentration of the investor base.

We re-estimate the regressions that analyze that effect of IC on flow volatility, cash, and risk-adjusted returns, but use either HHI_{Min} or HHI_{Max} normalized by 100 instead of IC. The results are shown in Tables B.1, B.2, and B.3 that include the same regression specifications as Tables 2, 3, and 7. Our results are robust to replacing IC with the HHI lower and upper bounds. The coefficient estimates and their significance are comparable to the coefficient estimates of IC.



(a) Lower bound Herfindahl-Hirschman Index



(b) Upper bound Herfindahl-Hirschman Index

Figure B.1: Lower and upper bounds of the Herfindahl-Hirschman Index
This figure shows lower and upper bound of the HHI for a given IC and number of investors.

Table B.1: Herfindahl-Hirschmann Index bounds and flow volatility

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (8). The dependent variable is the flow volatility estimated as the standard deviation over a rolling four-quarter window. The lagged IC variable is replaced by HHI lower or upper bounds normalized by 100 and computed as described in Section B. The remaining independent variables are lagged size, return volatility, flow volatility, share restriction, and manager stake. The data are quarterly from 2012:Q4 to 2016:Q4. Quarter fixed effects and strategy fixed effects are used where indicated. 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: Flow volatility, $\hat{\boldsymbol{\sigma}}_{\boldsymbol{i,t:t+3}}$				
	(1)	(2)	(3)	(4)
$MinHHI_{it-1}$	0.038**	0.040**		
	(2.95)	(3.20)		
$MaxHHI_{it-1}$			0.006*	0.007**
			(2.21)	(2.39)
$log(NAV_{it-1})$	-0.539***	-0.537***	-0.561***	-0.560***
	(-9.53)	(-9.52)	(-10.16)	(-10.11)
$\hat{\sigma}_{r_{i,t-4:t-1}}$	-0.002	0.004	-0.000	0.006
	(-0.12)	(0.25)	(-0.01)	(0.37)
$\hat{\sigma}_{F_{i,t-4:t-1}}$	0.188***	0.189***	0.189***	0.190***
	(18.63)	(19.28)	(18.84)	(19.52)
$ShareRes_{it-1}$	-0.001***	-0.001***	-0.002***	-0.002***
	(-4.27)	(-4.32)	(-4.31)	(-4.35)
$MgrStake_{it-1}$	-0.020***	-0.020***	-0.021***	-0.021***
	(-4.56)	(-4.41)	(-4.55)	(-4.41)
Quarter FE	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes
Observations	4676	4676	4676	4676
Adjusted R^2	0.109	0.110	0.109	0.110

Table B.2: Herfindahl-Hirschman Index bounds and cash

This table reports the coefficient estimates and t-statistics of the panel regression model given in equation (9). The dependent variable is cash normalized by NAV. The IC variable is replaced by HHI lower or upper bounds normalized by 100 and computed as described in Section B. The remaining independent variables are size, flows, share restriction, financing duration, leverage, and manager stake. The data are quarterly from 2012:Q4 to 2016: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, $\boldsymbol{Cash_{it}/NAV_{it}}$				
	(1)	(2)	(3)	(4)
$MinHHI_{it}$	0.664***	0.664***		
	(4.67)	(4.65)		
$MaxHHI_{it}$			0.129***	0.129***
			(4.53)	(4.51)
$log(NAV_{it})$	3.064***	3.061***	2.811***	2.808***
	(5.40)	(5.39)	(5.16)	(5.14)
F_{it}	-0.055	-0.048	-0.043	-0.036
	(-1.55)	(-1.28)	(-1.24)	(-0.97)
$ShareRes_{it}$	-0.026***	-0.026***	-0.026***	-0.026***
	(-4.85)	(-4.87)	(-4.89)	(-4.91)
$FinDur_{it}$	0.001	0.002	0.002	0.002
	(0.27)	(0.29)	(0.37)	(0.39)
$Leverage_{it}$	1.399***	1.391***	1.397***	1.390***
	(2.98)	(2.96)	(2.96)	(2.94)
$MgrStake_{it}$	-0.073	-0.072	-0.080	-0.079
	(-1.38)	(-1.37)	(-1.50)	(-1.49)
Quarter FE	No	Yes	No	Yes
Strategy FE	Yes	Yes	Yes	Yes
Observations	9773	9773	9773	9773
Adjusted R^2	0.232	0.232	0.230	0.230

Table B.3: Herfindahl-Hirschmann Index bounds and risk-adjusted returns

This table reports the coefficient estimates and t-statistics when estimating the model given in equation (14) 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 (13). The returns are deleveraged where indicated. The lagged IC variable is replaced by HHI lower or upper bounds normalized by 100 and computed as described in Section B. The remaining independent variables are lagged size, flows, share restriction, financing duration, and manager stake. The data are quarterly from 2012:Q4 to 2016: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, $\boldsymbol{\alpha_{it}}$				
	(1)	(2)	(3)	(4)
$MinHHI_{it-1}$	-0.024***	-0.018***		
	(-3.66)	(-4.23)		
$MaxHHI_{it-1}$			-0.004***	-0.003***
			(-3.57)	(-3.92)
$log(NAV_{it-1})$	-0.222***	-0.177***	-0.211***	-0.169***
	(-6.10)	(-7.66)	(-6.10)	(-7.47)
F_{it-1}	-0.004*	-0.006**	-0.004*	-0.006**
	(-1.81)	(-2.66)	(-1.83)	(-2.67)
$ShareRes_{it-1}$	0.002***	0.002***	0.002***	0.002***
	(8.13)	(8.14)	(8.23)	(8.22)
$FinDur_{it-1}$	0.006***	0.005***	0.006***	0.005***
	(6.71)	(6.42)	(6.70)	(6.41)
$MgrStake_{it-1}$	-0.002	-0.000	-0.002	0.000
	(-0.93)	(-0.09)	(-0.84)	(0.01)
Strategy FE	Yes	Yes	Yes	Yes
Deleveraged	No	Yes	No	Yes
Observations	7477	7477	7477	7477

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 2016:Q4.

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 on 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. Lastly, we require that a hedge fund's ratio of unencumbered cash over NAV is between 0 and 1.

These filters are imposed in addition to the filters based on IC, number of investors, manager stake, and Form ADV matching described in Section 3. More specifically, we require that $0<\text{IC}<100$. For 3,721 fund-quarter observations IC is equal to 100, and for 129 fund-quarter observations IC is equal to 0. We require that the number of investors in the fund is greater than 5. There are 896 fund-quarter observations with 5 or fewer investors. The manager stake has to be smaller than or equal to 50%. There are 903 fund-quarter observations with manager stake greater than 50%. We require that the matching between Form PF and ADV is successful for each hedge fund in the sample. 269 fund-quarter observations could not be matched. 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

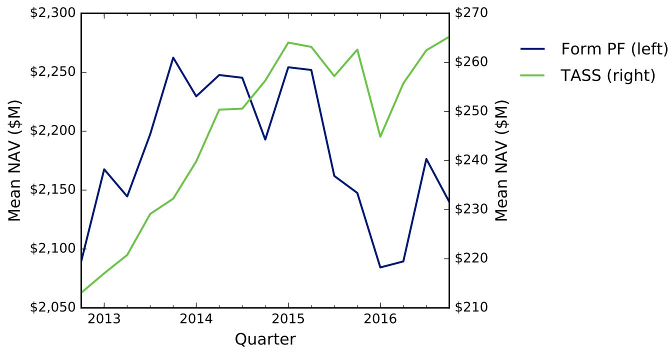
The methodology used for classifying a hedge fund's broad strategy is as follows. First, we check the Question 20 description field for the "Other" category to determine if the description can be directly mapped to one of the other broad categories. For example, a description of "Relative Value Fixed 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, macro, managed futures, relative value, fund of funds, and multi-strategy) 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" as these are too few to include given confidentiality restrictions.

C.3 Form PF and TASS comparison

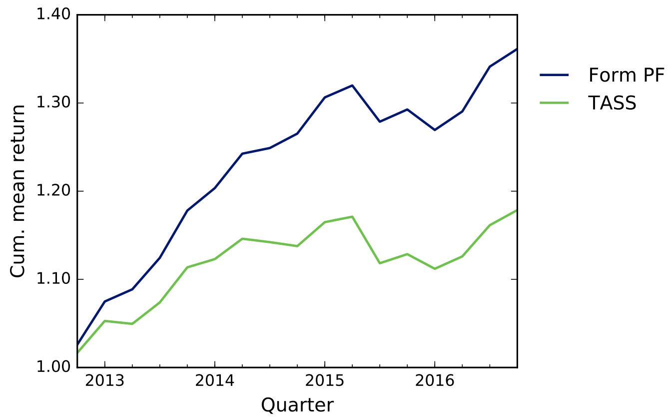
The TASS commercial hedge fund database contains voluntarily reported monthly NAV and net of fees returns. In this section, we compare the Form PF and TASS datasets. We use TASS hedge funds that report in US$. In this comparison, there are 1,355 unique funds in the Form PF dataset and 2,368 funds in the TASS dataset.

In Figure C.1, we plot the size, net of fees returns, and flows of hedge funds from the Form PF and TASS datasets. The Form PF hedge funds are on average an order of magnitude larger than the hedge funds that report to TASS. This difference is caused by the fact that only hedge funds of a certain size are required to file Form PF, as discussed in Section 3. Moreover, hedge funds that voluntarily report to TASS likely do so for the purpose of advertising themselves to potential investors and attracting new investment. Consequently, the TASS hedge funds tend to be smaller on average. The time-variation of the average NAV is similar for Form PF and TASS hedge funds. Both series increase until 2015 and decrease after that, with the Form PF series being more volatile. For the returns, the cumulative return series correlate strongly, with the returns being higher for Form PF than for TASS hedge funds. For the flows, the cumulative flow series also correlate strongly, with the TASS hedge funds—which are on average an order of magnitude smaller—experiencing more outflows during this period.

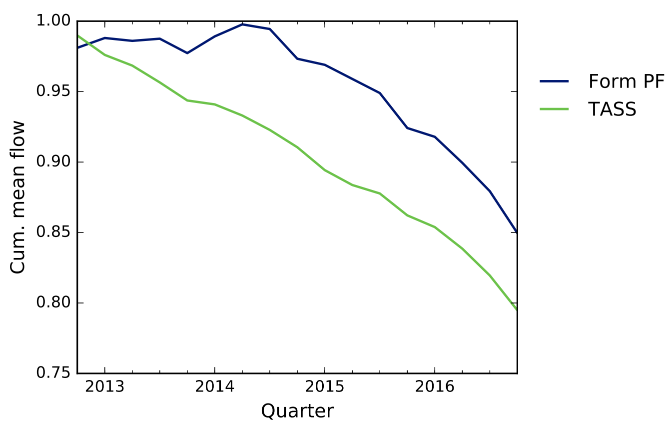
In Figure C.2, we again compare the size, net of fees returns, and flows, of Form PF and TASS hedge funds, but with the TASS hedge funds filtered based on size. Here, we only include TASS hedge funds with NAV equal to or greater than US$500 million. There are 300 such funds in the TASS dataset. The Form PF hedge funds are on average still larger than the hedge funds that report to TASS, which indicates that the larger hedge funds in Form PF do not report to TASS. The cumulative net of fee return series and the cumulative flow series of the Form PF hedge funds and the size-filtered TASS hedge funds still correlated strongly. Here, the Form PF hedge funds experience larger outflows than the size-filtered TASS hedge funds. This difference in flows could be due to TASS hedge funds actively trying to attract new investor money by reporting to TASS and marketing themselves to new investors.



(a) Net asset values

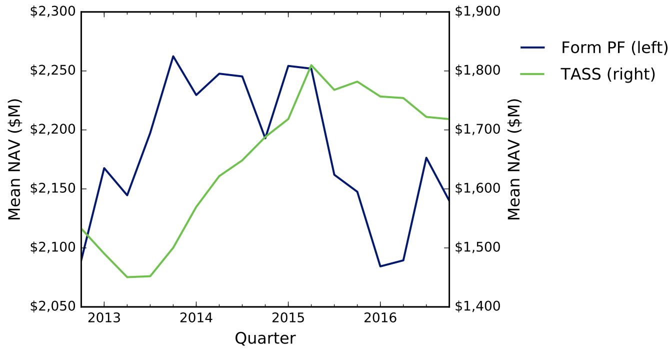


(b) Cumulative net of fees returns

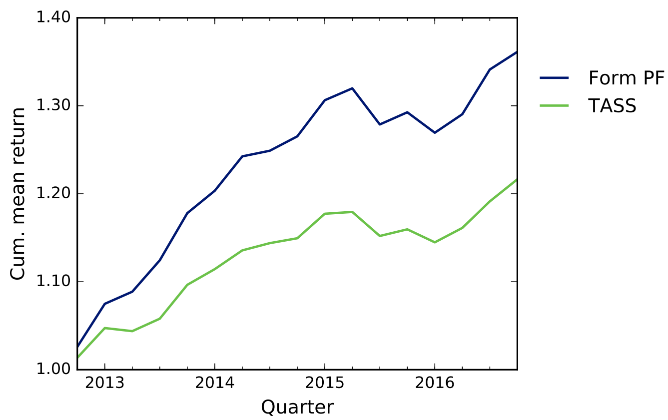


(c) Cumulative flows

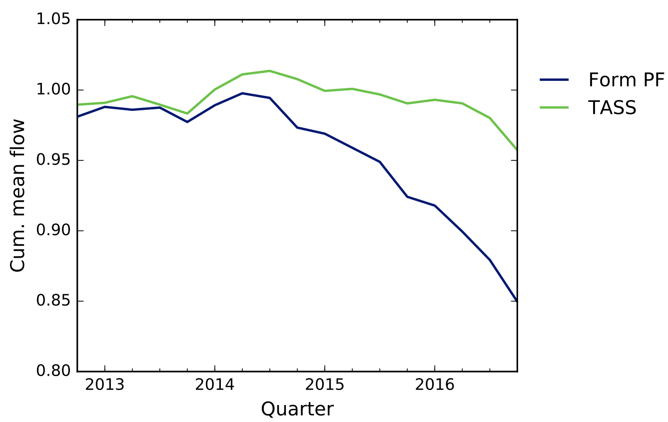
Figure C.1: Form PF and TASS comparison
This figure shows the average NAV, the average cumulative net of fees returns, and the average cumulative flows of Form PF and TASS hedge funds.



(a) Net asset values



(b) Cumulative net of fees returns



(c) Cumulative flows

Figure C.2: Form PF and TASS filtered by size comparison
This figure shows the average NAV, the average cumulative net of fees returns, and the average cumulative flows of Form PF and TASS hedge funds. Only TASS hedge funds with a NAV equal to or greater than US$500 million are included.

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, 2017, Liquidity transformation and financial fragility: Evidence from funds of hedge funds, *Unpublished working paper*.

Agarwal, Vikas, Yakup E. Arisoy, and Narayan Y. Naik, 2016, Volatility of aggregate volatility and hedge fund returns, *Journal of Financial Economics forthcoming*.

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.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effect, *Journal of Financial Markets* 5, 31–56.

Ang, Andrew, Sergiy Gorovyy, and Gregory B. van Inwegen, 2011, Hedge fund leverage, *Journal of Financial Economics* 102, 102–126.

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

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.

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, 2017, The granular nature of lareg institutional investors, *NBER Working Paper*.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

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.

Brown, Stephen, William Goetzmann, Bing Liang, and Christopher Schwarz, 2008, Mandatory disclosure and operational risk: Evidence from hedge fund registration, *Journal of Finance* 63, 2785–2815.

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.

Cao, Charles, Yong Chen, Bing Liang, and Andrew W. Lo, 2013, Can hedge funds time market liquidity?, *Journal of Financial Economics* 109, 493–516.

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

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276–1302.

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.

Chernenko, Sergey, and Adi Sunderam, 2016, Liquidity transformation in asset management: Evidence from the cash holdings of mutual funds, *Harvard Business School Working Paper*.

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.

Da, Zhi, Borja Larrain, Clemens Sialm, and Jose Tessada, 2017, Destabilizing financial advice: Evidence from pension fund reallocations, *Review of Financial Studies forthcoming*.

Dahlquist, Magnus, Jose V. Martinez, and Paul Soderlind, 2017, Individual investor activity and performance, *Review of Financial Studies forthcoming*.

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, *OFR Working Paper.*

Franzoni, Francesco A., and Alberto Plazzi, 2015, What constraints liquidity provision? Evidence from hedge fund trades, *Swiss Finance Institute Research 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.

Gabaix, Xavier, Parameswaran Gopikrishnan, Vasiliki Plerou, and H. Eugene Stanley, 2006, Institutional investors and stock market volatility, *Quarterly Journal of Economics* 121, 461–504.

Getmansky, Mila, Bing Liang, Christopher Schwarz, and Russ Wermers, 2015, Share restrictions and investor flows in the hedge fund industry, *Unpublished 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, 2015, Investor flows and fragility in corporate bond funds, *Journal of Financial Economics forthcoming.*

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

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.

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

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

——— , and Lucian A. Taylor, 2015, Scale and skill in active management, *Journal of Financial Economics* 116, 23–45.

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

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.

Securities and Exchange Commission, 2016, Investment company liquidity risk management programs, *Final Rule.*

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.