Can Mutual Fund Investors Distinguish Good from Bad Managers?*

TEODOR DYAKOV[†] ⓘ AND MARNO VERBEEK[‡] ⓘ

[†]School of Business and Economics, VU Amsterdam and Tinbergen Institute, Amsterdam, The Netherlands and
[‡]Rotterdam School of Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

ABSTRACT

Mutual fund flows respond significantly to the return gap, which captures information about unobserved actions of mutual funds and predicts future performance. The sensitivity of fund flows to the return gap is: (i) strong and positive; (ii) increasing with investor sophistication; (iii) highly nonlinear; and (iv) decreasing with the informativeness of past fund returns. On average, the response of investors to the return gap enhances their performance. Our findings suggest there is a sophisticated mass of investors who can distinguish good from bad managers using information that may not be directly inferred from standard performance indicators.

I. INTRODUCTION

With currently more than $8.5 trillion in assets under management[1], the equity mutual fund industry holds a substantial amount of the total market portfolio in the USA. Understanding how investors move capital across the plenitude of funds available is therefore important for understanding the allocative efficiency of capital markets. The extensive mutual fund literature has studied various determinants of mutual fund flows, with the overall conclusion that investors tend to make naive decisions. Most notably, past studies have shown that investors make decisions largely based on past performance (e.g., Ippolito 1992; Chevalier and Ellison 1997; Sirri and Tufano 1998), even though past

performance appears to be a poor predictor of future performance (e.g., Carhart 1997). Recent results from the behavioral literature further point to the direction that investors often seem to be naive and inexperienced in their decisions.[2]

In this paper, we want to augment our knowledge on the drivers behind mutual fund flows by investigating whether investors direct flows towards managers likely to add value in the future. We argue that investors may possess information about future performance which is not directly captured by observable fund characteristics. Investors may base their inferences on information coming from qualitative sources, an analysis of fund holdings, reading analysts' reports, and so on. As long as the performance signal that investors derive is not captured by observable fund characteristics, regressing fund flows on fund characteristics might miss important insights about some of the drivers behind fund flows.

We use the return gap of Kacperczyk et al. (2008) to *proxy* such information about future performance. Kacperczyk et al. (2008) show that the return gap, calculated as the difference between the reported fund returns and the hypothetical return of the fund's most recently disclosed holdings, is highly persistent and predicts future performance. The return gap is particularly useful for avoiding poorly performing funds in the future. In contrast to the return gap, conventional performance measurements have very limited ability to distinguish good from bad fund mangers. Moreover, the return gap cannot be explained by observable fund characteristics, such as past performance. These results suggest the existence of information about future performance orthogonal to previously studied observable fund characteristics as determinants of mutual fund flows.

Accordingly, we investigate whether mutual fund flows are related to information about future performance reflected in the return gap. A positive correlation between fund flows and past realizations of the return gap would indicate that mutual fund investors are able to differentiate good from bad managers using information beyond readily available performance indicators. Such positive correlation does not require investors to be able to actually calculate the return gap for each fund. Instead, it suggests that investors use information signals correlated with the information content of the return gap when investing in funds.

Using a large panel of nearly 2500 actively managed US equity mutual funds over the period 1990 to 2010, we find strong support for this conjecture. Our results show a strong sensitivity of fund flows to the return gap, over and above other performance indicators. More specifically, a one standard deviation increase in the return gap during the last year is followed by a 0.74% increase in money flows in the following quarter. This finding indicates that mutual fund investors use information about future performance beyond standard

---

2 Examples include Barber et al. (2005), Cooper et al. (2005), Choi et al. (2010), Bailey et al. (2011), and Frazzini and Lamont (2008).

backward-looking performance measures, like returns and alphas, in their allocation decisions.

Separating bad from good managers is a process that requires a certain degree of investor sophistication. Consistent with this notion, we find that the sensitivity of fund flows to the return gap is stronger for institutional investors than for retail investors. Furthermore, we find that almost all of the sensitivity of fund flows is driven by a response to funds in the top return gap quintile. We also find that the sensitivity of fund flows to the return gap is stronger when there is less cross-sectional dispersion in fund performance, implying that the performance information investors obtain becomes more important when there is less information in past net performance.

We further investigate the economic importance from our main finding that fund flows respond to the return gap. Given that the return gap is related to future performance, the positive sensitivity of fund flows to past realizations of the return gap suggests that investors enhance their returns from directing flows towards high return gap funds and avoiding low return gap funds. To assess the economic magnitude of this effect, we first calculate for each fund the difference between the expected fund flows from a flow-performance model including the return gap with those from a flow-performance model excluding the return gap. This difference captures the differential capital allocated to mutual funds that is attributed to differences in their return gaps. Next, we sort funds into 10 decile portfolios based on this difference and investigate their performance over time. The four-factor alphas of the spreads between the top and bottom portfolios amount to 18 to 21 bp per month, depending on the specification. These effects imply a sizable economic benefit that investors realize from directing flows towards high return gap funds and particularly from avoiding low return gap funds.

We next test whether investors are guided towards better fund managers by brokers and financial advisers. Our results do not offer evidence for this conjecture. We do not find significant differences in the sensitivity of fund flows to the return gap across investors using financial advisers and brokers and those who do not. For robustness, we show that very little of the sensitivity of fund flows to the return gap can be attributed to readily available performance indicators and fund characteristics. This evidence supports our conjecture that investors are able to infer information about future performance which may not be directly observable or easily deduced from fund characteristics.

An alternative explanation for our findings is related to momentum. A high return gap may be the result of funds chasing high momentum stocks. Under this conjecture, funds with high return gaps generate high returns because of momentum. This is unlikely to be the case. First, we show that funds with high return gaps outperform funds with low return gaps even after controlling for exposure to the momentum risk factor. Second, past performance, together

with a number of other observable characteristics, explains a mere 4% of the variation of the return gap.

This paper builds on the literature investigating the drivers of mutual fund flows. A well established finding in this literature postulates that fund flows respond strongly to past performance.[3] Other determinants of fund flows examined include fund size (e.g., Sirri and Tufano 1998), fund ratings (e.g., Del Guercio and Tkac 2008), the presence of a star fund within the family (e.g., Nanda et al. 2004), media coverage (e.g., Solomon et al. 2014), advertisements (e.g., Jain and Wu 2000), and fees (e.g., Barber et al. 2005), among others. Sialm et al. (2015) show that plan sponsors' monitoring of defined contribution plans' available options leads to relatively more volatile fund flows that respond stronger to past performance. Berk and van Binsbergen (2016) and Barber et al. (2016) study the response of fund flows to alternative measures of performance derived from competing asset pricing models. We differ from this literature by showing that flows are correlated with information about future performance beyond readily available backward-looking performance indicators. In other words, we show investors are able to extract information about future performance which is not captured by observable fund characteristics.

This paper is also related to the literature investigating managerial skill. A number of studies document that some fund managers are able to consistently beat their benchmarks.[4] We take this analysis one step further and investigate whether investors are able, in the cross-section, to distinguish good from bad managers. What separates us from other papers is that we document a new channel though which investors (particularly institutional) allocate capital: namely, information beyond readily available performance indicators. This is important, as previous studies (e.g., Evans and Fahlenbrach 2012) only document some evidence of investor sophistication driven by a response to observable characteristics.

Our findings also offer support to the theoretical literature that reconciles the stylized facts of mutual fund underperformance, lack of performance persistence, and the performance-flow relationship with the notion that fund investors are sophisticated. Notably, Berk and Green (2004) show theoretically that both lack of performance persistence and the flow-performance relationship are part of a framework where investors learn about managerial skill from past returns. Similarly, Lynch and Musto (2003), Huang et al. (2007, 2012) incorporate investor sophistication in models attempting to explain stylized mutual fund facts. Our paper provides empirical support in favor of investor sophistication by showing that at least some fund investors can separate good from bad fund managers.

---

3 See, for example, the work of Ippolito (1992), Gruber (1996), Chevalier and Ellison (1997), and Sirri and Tufano (1998).
4 See, for example, Hendricks et al. (1993), Elton et al. (1996), Cohen et al. (2005), Kacperczyk et al. (2005), Jiang et al. (2007), Kacperczyk and Seru (2007), Cremers and Petajisto (2007), and Baker et al. (2010).

II. DATA SELECTION

This study combines a number of commonly used databases—Center for Research in Security Prices (CRSP) Mutual Fund Database, Thomson Financial/ CDA equity holdings database, and the CRSP monthly stock files. The CRSP Mutual Fund Database provides monthly fund net investor returns, total net assets and annual data on expenses, fees, proportion of assets invested in common stocks, bonds, cash and other securities, and other fund characteristics. The Thomson Financial/CDA database covers quarterly/semi-annual holdings of mutual funds, as reported to the Securities and Exchange Commission (SEC) or voluntarily reported by the funds, which we link to the monthly and daily CRSP stock files in order to obtain information on holdings' prices and returns (adjusting for stock splits and other share adjustments). Both mutual fund databases are free of survivorship bias and linked *via* the MFLINKS tool provided by Wharton Research Data Services (WRDS). This study focuses on US domestic actively managed equity mutual funds, for which the data is most complete and reliable. Thus, we exclude index, balanced, bond, money market, sector, and international funds, as well as funds that do not invest primarily in common stocks. Since most actively managed US equity funds offer different share classes to investors, we sum the net assets over different share classes and take asset-weighted share class averages of different attributes such as returns and expense ratios. More details on the merging process and sample selection is available in Appendix A.

Following standard procedures in the literature, we define flows for fund *i* during quarter *t* as the return-adjusted difference in total net assets (TNA) between the start and end date of quarter *t*, scaled by the fund's total net assets at the start of the quarter[5]:

$$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}\left(1 + Return_{i,t}\right)}{TNA_{i,t-1}}, \tag{1}$$

where *TNA* stands for total net assets and *Return* for net fund return.

The summary statistics are presented in Table 1. In total, the sample covers 2486 equity mutual funds, ranging from 373 in 1990 to 1691 in 2006. Over time, the median amount of assets has increased from \$137 million to \$309 million. We also observe a tendency for mutual funds to hold larger numbers of stocks in more recent times. Generally, the first half of our sample period (before 2000) is characterized by larger mean flows and higher returns than the second half. We further note that the mean annual expense ratios have remained about the same throughout the sample period.

---

5 Consistent with Coval and Stafford (2007), we exclude funds whose information is too different between CRSP and CDA $\left(1/1.3 < TNA_{i,t}^{CRSP}/TNA_{i,t}^{CDA}\right)$ and funds with too extreme changes in TNA ($-0.5 < \Delta TNA_{i,t}/TNA_{i,t-1} < 2.0$).

Table 1 Summary statistics of the sample

	No. of funds	No. of stocks Median	Net assets, $mil Median	Flow, % per quarter Mean	Return, % per quarter Mean	Expense ratio, % per year Mean
1990	373	56	137.19	0.73	−1.18	1.26
1991	420	56	130.83	3.61	8.33	1.27
1992	502	58	142.66	5.17	2.53	1.29
1993	536	63	173.26	5.18	3.73	1.26
1994	678	67	197.58	2.85	−0.11	1.26
1995	809	68	169.89	3.37	6.96	1.25
1996	920	71	185.53	3.94	4.48	1.24
1997	1000	74	222.20	3.61	5.58	1.23
1998	1160	73	229.57	1.99	4.27	1.26
1999	1201	70	233.10	0.62	6.95	1.26
2000	1374	72	272.15	2.87	0.15	1.27
2001	1408	75	276.30	2.66	−1.23	1.29
2002	1517	76	227.60	1.03	−5.39	1.33
2003	1595	75	171.00	2.24	8.25	1.36
2004	1691	81	214.00	1.22	3.12	1.33
2005	1679	78	232.30	1.32	1.86	1.30
2006	1691	78	261.65	0.77	3.13	1.27
2007	1621	77	306.10	−0.15	1.84	1.22
2008	1603	76	320.25	−1.23	−11.01	1.20
2009	1504	76	239.40	0.19	7.63	1.20
2010	1274	82	309.20	−0.43	−2.11	1.14

III. THE RETURN GAP: SEPARATING GOOD FROM BAD FUND MANAGERS

Our study investigates whether mutual fund investors are able to identify funds likely to perform well in the future. We construct a *proxy* which is likely to be highly correlated with the information investors use to distinguish good from bad funds. The proxy we use is the return gap of Kacperczyk et al. (2008), which is constructed as the difference between the performance of the fund and the performance of the portfolio based on the fund's most recently reported holdings. We rely on the return gap because it is known to be persistent and a good predictor of future fund performance. While the return gap is not directly observable, it is more easily derived from observable information than, for example, a measure like Active Share (Cremers and Petajisto 2007), which requires detailed information about a fund's benchmark.

A. Construction of the return gap

Following Kacperczyk et al. (2008), for each fund *i* in quarter *t*, the return gap is constructed as

$$ReturnGap_{i,t} = Return_{i,t} - \left(HoldingsReturn_{i,t} - ExpenseRatio_{i,t} \right). \qquad (2)$$

For each fund i, *HoldingsReturn*$_{i,t}$ refers to the quarter t return of the portfolio holdings disclosed at the end of quarter $t - 1$ and *ExpenseRatio*$_{i,t}$ is the most recently available fund expense ratio at the beginning of quarter t. We use the stockholdings information provided by Thomson Financial in order to identify each common stock in a fund's portfolio. These data come from mandatory reports to the SEC as well as voluntary reports by the mutual funds. After 2004, all funds are required to report their holdings quarterly to the SEC. Before then, they were required to file their holdings semiannually, but about two thirds of the funds already reported quarterly. Even though we select funds with average percentage of assets invested in common stocks above 80% and below 105%, funds still have a proportion of their portfolio invested in other assets. We cannot identify the precise portfolio composition in those other assets and we proxy their returns with the returns of suitable indices. In particular, we approximate returns of bonds and preferred stocks with the Barclays Aggregate Bond Index (formerly known as the Lehman Brothers Aggregate Bond Index) and the return of cash and other assets with the Treasury Bill rate.[6] Since a number of funds included in the Thomson Financial/CDA database have long periods of missing data, we require the latest fund holdings used for calculating the return gap in quarter t to be not older than 6 months at the beginning of quarter t. The expense ratio used is the most recently reported as of the end of quarter t and reported no earlier than 2 years before the end of quarter t, and is calculated as one fourth of the yearly expense ratio. Throughout the paper, we aggregate the quarterly calculated return gap to a yearly return gap measure.

B. Interpretation of the return gap

The return gap of Kacperczyk et al. (2008) captures the impact of unobserved actions of mutual fund managers. Even though funds are subject to extensive disclosure requirements, most of their actions remain unobserved to investors. For example, investors do not observe the transaction costs paid by managers, the timing of their trades, or how many units of each stock they hold between the quarterly portfolio reports. However, the impact of these unobserved actions is reflected in the net return of the fund, without affecting the hypothetical return of the fund's most recently disclosed holdings. Consequently, the difference between the fund's return and the return of the hypothetical portfolio, measured by the return gap, captures the value added (or subtracted) by fund managers *via* their unobserved actions. On the one hand, value-adding unobserved trades would increase the return of the fund relative to the return of the previously disclosed holdings. On the other hand, trading costs and

---

6 The bond index data come from Datastream and the return of Treasuries comes from Kenneth French's Data Library http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Table 2 Summary statistics and Pearson correlations of key variables

| | Mean | StdDev | Correlation with | | | | |
			Return Gap	Alpha	ExpRatio	Flow	Fund Return
YearlyReturnGap	−0.02	0.05	1.00				
Alpha	−0.05	0.86	0.22	1.00			
ExpRatio	1.27	0.66	−0.12	−0.04	1.00		
YearlyFlow	10.73	39.55	0.10	0.13	0.04	1.00	
YearlyFundReturn	9.44	22.60	0.17	0.22	−0.02	0.15	1.00

Alpha is estimated using past 12 months of data with the excess return on the market, SMB, HML, and momentum as risk factors and expressed in monthly value. All other variables are expressed in yearly values.

commissions and other value-decreasing unobserved actions affect negatively the return gap.

We provide summary statistics for the return gap and other key variables, together with their Pearson correlations, in Table 2. The mean return gap is negative, which implies that, on average, the gains of the unobserved interquarterly actions of fund managers do not outweigh the trading costs. This return gap is characterized with a substantial cross-sectional dispersion. Our yearly estimate of return gap is lower than that of Kacperczyk et al. (2008): –0.20% per year *versus* 0.13% (equally weighted) and –0.12% (value-weighted) in Kacperczyk et al. (2008). There are two potential reasons for this difference. First, our sample is more recent. Barras et al. (2010), Fama and French (2010), and Lewellen (2011), among others, document a decreasing mutual fund performance over time. Thus, the return gap could also be decreasing over time. Second, Kacperczyk et al. (2008) include a small number of index funds in their sample (4.5% of all funds), while we exclude them. Index funds are likely to have a return gap that is closer to zero than active funds. Hence, our estimate of the return gap should be slightly smaller than that of Kacperczyk et al. (2008). We further note that the return gap is negatively correlated with fund expenses, which implies that fees, on average, are not compensating for value-enhancing unobserved actions. Not surprisingly, the return gap is positively correlated with past returns and alpha because the return gap contributes to both net returns and alpha, but the correlations are relatively low.

An important driver of the return gap is transaction costs. These costs affect fund performance negatively, without affecting the return of the previously disclosed fund holdings. Thus, funds paying high brokerage fees will typically have more negative return gaps than their peers. Grinblatt and Titman (1989) are the first to use the difference between fund return and the return of the most recently disclosed holdings for approximating transaction costs. Later, the same approximation for inferring transaction costs has been used by Wermers (2000) and Bollen and Busse (2006).

However, the return gap captures more than the effect of trading costs. The return gap may reflect informational advantages, or optimal timing of trades

(Kacperczyk et al. 2008). For example, a mutual fund manager may process news faster than the market and trade before her private information is incorporated into prices. Suppose a manager reported her portfolio holdings at the end of December and then again in the end of March in the following year (a realistic quarterly disclosure policy). A manager may sell an overvalued stock in January, before the rest of the market brings the price of the stock closer to fundamentals in February. In this case, the asset sale positively affects the return of the fund without affecting the return of the most recently disclosed holdings, driving upwards the return gap in that quarter. Using daily fund returns Bollen and Busse (2005) demonstrate that stock selection and market-timing are short-lived phenomena whose effect on fund performance disappear within a quarter. Alternatively, a negative return gap might appear due to, for instance, agency problem within the fund or the fund family (e.g., Gaspar et al. 2006; Casavecchia and Tiwari 2016). Thus, information about the future performance of the fund manager is likely to be reflected in the return gap.

To understand how the return gap is different from alpha, consider the case when there is a shock to a stock in a quarter when the stock is not traded by the fund manager. In that case, the shock affects equally the most recently disclosed holdings and net return of the fund, and therefore does not affect the return gap. Yet, this shock is reflected in the overall risk-adjusted performance. In the hypothetical fund setup above, a manager may receive a private signal in February that a stock will experience surprisingly high earnings in April. The information content of the trade does not affect the return gap in quarter 1, because the net return is not affected until April. Furthermore, the trade does not affect the return gap in quarter 2, because it is not traded in quarter 1 and consequently it does not change the net return in quarter 2 relative to the most recent holdings disclosed at the end of quarter 1.[7] This explains why we find the positive but less than perfect correlation between the return gap and alpha in Table 2.

Moreover, the return gap cannot be explained by observable fund characteristics. We regress the quarterly return gap on a number of variables, which might have an economic link to the return gap. Since our results are similar to those of Kacperczyk et al. (2008) and in the interest of brevity, we skip discussion of the individual relationships between each determinant of the return gap and leave it to Appendix B, where we summarize the results in Table B1. Importantly, we find that the R^2 of the regression is only 4%, which implies that very little of information contained in the return gap is captured by observable fund characteristics.

Importantly, Kacperczyk et al. (2008) show that unobserved actions of some funds can help differentiate good from bad managers. To show this, at the end

---

7 Empirical evidence on such trading behavior comes from, for example, Baker et al. (2010) who demonstrate that mutual fund quarterly trades predict next quarter's unexpected stock earnings.

of each quarter we sort funds on their return gaps over the previous 1, 3, and 5 years and then show that these past sorts predict return gaps in the following quarter. Moreover, we show that information contained in the return gap is not captured by other performance measures. Specifically, funds with higher return gaps outperform funds with lower return gaps, even when we control for net returns or alpha. The portfolios with highest return gap do not have statistically positive alphas, while portfolios with the lowest return gap have statistically negative alphas. Hence, the return gap is particularly useful for avoiding poorly performing funds. The results, similar to Kacperczyk et al. (2008), are presented in Appendix B, Tables B2 and B3, respectively. These two tests suggest that if investors are able to infer information that helps them predict future performance, it can be reflected in the return gap.

In sum, the return gap measures the unobserved actions of mutual fund managers and is a persistent indicator of future performance that predicts returns better than traditional measures of past performance. Moreover, the information contained in the return gap cannot be captured by observable fund characteristics. Therefore, even though we do not observe the information process that sophisticated investors potentially use to select funds, any information they possess that is orthogonal to observable fund characteristics and performance indicators, can be reflected in the return gap. Consequently, investigating the sensitivity of fund flows to the return gap provides us with a powerful setup for testing whether investors can identify good and bad funds in the cross-section of fund managers, using information beyond readily available characteristics and performance measures.

IV. THE SENSITIVITY OF FUND FLOWS TO THE RETURN GAP

In this section, we investigate the sensitivity of fund flows to the return gap. We provide a number of empirical findings consistent with the hypothesis that investors respond to information that predicts future performance, proxied by the return gap.

A. Main effect

We regress quarterly fund flows in quarter t on lagged variables, known to influence investors' capital allocation decisions, and the yearly return gap. More specifically,

$$Flow_{i,t+1} = \boldsymbol{\beta}'\boldsymbol{X}_{i,t} + \epsilon_t. \tag{3}$$

The vector of explanatory variables $\boldsymbol{X}_{i,t}$ includes past net returns and fund flows, alpha, the most recently available expense ratio, and past return gap. All variables are calculated using yearly data. The alpha is estimated at the end of quarter t using monthly data over the preceding 12 months from a four factor

model, including the excess return of the market, the size factor (SMB), the value factor (HML), and momentum.[8] We include the most recently available expense ratio because the return gap is calculated using fund's expenses. This way, we rule out a mechanical relation between fund flows and the return gap that may be due to a response to the expense ratio. Morevoer, we include fund style-fixed effects in each specification.[9] We estimate the models using pooled regressions with time-fixed effects and standard errors clustered on the fund level.[10]

The results are summarized in Table 3. We add sequentially the different components of the return gap in specifications (1) to (3). The results indicate that investors respond strongly to both holdings return and the return gap. In specifications (4) to (6), we add additional control variables, including fund net return, alpha, and flows. We find that flows are persistent and investors strongly chase past returns and alpha, consistent with previous studies on the flow-performance relationship, such as Ippolito (1992) and Chevalier and Ellison (1997). Importantly, the return gap has an incremental power in explaining fund flows in each specification. A one standard deviation increase in the return gap in the previous year leads to subsequent 0.74% flows in the following quarter. In comparison, a one standard deviation increase in yearly alpha results in 1.26% in additional quarterly flows. The evidence suggests that the impact of the sensitivity of fund flows to the return gap is therefore economically important.

Previous studies document negative relationship between fund expenses and fees (e.g., Sirri and Tufano 1998). We find a statistically insignificant relationship between fees and flows, and in an unreported Fama–Macbeth test we even find a statistically positive relationship between flows and fees. Yet, the specifications include total expense ratio, which contains management fees, administrative fees, operating costs, 12b-1 fees, and all other costs potentially incurred by the fund. Barber et al. (2005) argue that investors might be unaware of magnitude of the different components and thus respond to the more salient load fees, which are not part of the total expense ratio. Alternatively, investors might respond negatively to operating and management fees, but positively to the marketing and distribution expenses, known as 12b-1 fees (e.g., Jain and Wu 2000). The positive response to 12b-1 fees could be driven by managers masking payments to brokers and advisors in the 12b-1 fees and marketing themselves as no-load funds in order to attract naive investors (Haslem 2009).

Table B1 in Appendix B demonstrates that observable fund characteristics explain only 4% in the variation of the return gap. This makes it unlikely that the effect we document in this section can be attributed to an omitted fund

---

8 The risk factors are obtained from Kenneth French's data library: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

9 A fund's style is determined as at most two of the following styles: large, small, value, and growth. We base our fund style selection on the basis of on the funds' Lipper objective codes.

10 We find similar results using Fama–Macbeth regressions (Fama and Macbeth 1973) with Newey–West standard errors. For brevity, we do not report them.

Table 3 Investors' response to the return gap

	(1) Flow$_t$		(2) Flow$_t$		(3) Flow$_t$		(4) Flow$_t$		(5) Flow$_t$		(6) Flow$_t$	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
Intercept	-1.40***	-6.75	-1.29***	-4.23	-1.39***	-3.00	-1.42***	-2.96	-1.10**	-2.18	-1.08**	-2.33
YearlyHoldings Return$_{t-1}$	0.22***	27.55	0.22***	27.51	0.25***	28.34						
ExpRatio$_{t-1}$			-0.08	-0.46	0.30	0.95	0.53	1.58	0.57	1.60	0.51	1.55
YearlyReturn Gap$_{t-1}$					0.46***	15.85	0.18***	6.83	0.15***	5.89	0.14***	5.82
YearlyFund Return$_{t-1}$							0.26***	28.16	0.22***	26.81	0.21***	24.46
Alpha$_{t-1}$									1.51**	8.72	1.46***	8.70
YearlyFlow$_{t-1}$											0.01***	4.32
Style FE	Yes		Yes		Yes		Yes		Yes		Yes	
Time FE	Yes		Yes		Yes		Yes		Yes		Yes	
R^2	0.05		0.05		0.07		0.07		0.08		0.09	
Observations	85,914		85,914		85,914		85,914		85,914		85,914	
Time period	Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010	

The dependent variable in each regression specification is fund flow in quarter *t*. Depending on the specification, we include a lagged year holdings return, the most recently available expense ratio, lagged yearly return gap, lagged yearly fund net return, alpha (estimated using past 1 year of monthly fund returns and the excess return on the market, SMB, HML, and momentum as risk factors), and lagged yearly flow. All specifications include style fixed effects. We estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

516

characteristic. However, for robustness, we investigate this conjecture in Section B and find that the results we present in Table 3 remain largely the same after controlling for fund characteristics.

B. The sensitivity of fund flows to the return gap, conditional on investor sophistication

We conjecture that investors' information about future performance is likely to be reflected in the return gap. However, such information is likely to be costly to obtain and difficult to process. Thus, we expect our previous results to be driven by the more sophisticated investors. To empirically test this hypothesis, we repeat the analysis in Section A, conditional on investor type, where institutional investors are expected to be more sophisticated than retail investors (for instance, Evans and Fahlenbrach 2012).

Since 1999, the CRSP database reports whether a share class was distributed to institutional or retail investors, which provides the main identification mechanism in this section. The share class distinction allows us to aggregate separately flow and return data for the retail and institutional part of a fund. Consequently, we obtain flow and return data separately for institutional and retail investors. Note that if a fund does not distribute share classes to institutional (retail) investors it drops out of the institutional (retail) subsample. In total, the institutional investors subsample has 25,706 fund-period observations and the retail investor subsample has 49,653 fund-period observations, covering the period 2000 to 2010.

We estimate the restricted and unrestricted flow performance specifications in Section A, separately for the institutional and retail subsamples. The dependent variable, the lagged net return, expense ratio, and alpha are calculated separately for the institutional and retail subsamples, while the lagged flows and return gap variables are calculated the same way as in the previous analysis, using information on the whole fund level (i.e., both retail and institutional). We report results aggregating lagged flow measures on the whole fund level, but results remain qualitatively the same if we aggregate the flows separately for the institutional and retail subsamples.

The estimation results covering the period 2000–2010 are summarized in Table 4. In specification (1) we report results for the institutional subsample, while specification (2) relates to the retail subsample. Comparing the results across the two subsamples, we do not observe a differential response to past performance data. The main difference comes with respect to the expense ratio variable—institutions avoid funds with high expenses, while individual investors prefer them, possibly due to the effect of advertisement fees (Jain and Wu 2000; Barber et al. 2005).

The findings further suggest that the results in Table 3 presented earlier are mainly driven by the more sophisticated clientèle. Institutional flows respond very strongly to the yearly return gap. On the other hand, the statistical significance using the subset of retail investors is much weaker. Furthermore, the

Table 4 Investors' response to the return gap—institutional *versus* retail investors

	Institutional (1) Flow$_t$		Retal (2) Flow$_t$		Difference in RG (1) − (2)	
	Coeff	*t*-stat	Coeff	*t*-stat	Diff	*t*-stat
Intercept	0.03***	4.48	0.00	−0.49		
ExpRatio$_{t-1}$	−0.18*	−1.90	0.00	0.01		
YearlyFlow$_{t-1}$	0.01***	7.82	0.01**	2.16		
YearlyFundReturn$_{t-1}$	0.20***	9.45	0.26***	9.91		
Alpha$_{t-1}$	2.54***	8.06	2.28***	10.24		
YearlyReturnGap$_{t-1}$	0.14***	2.69	0.08*	1.76	0.05	0.80
Style FE	Yes		Yes			
Time FE	Yes		Yes			
R^2	0.04		0.11			
Observations	25,706		49,653			
Time period	Q1.2000–Q3.2010		Q1.2000–Q3.2010			

We use the identification of retail and institutional share classes introduced by CRSP at the end of 1999 and aggregate the flow, expenses, and return data separately for the retail and institutional part of a fund. The dependent variable in specification (1) is institutional flow in quarter *t*, and in specification (2) is retail flow in quarter *t*. In each specification we include an intercept, the most recently available expense ratio, lagged yearly flow, lagged yearly fund net return, specific to institutional (specification (1)) or retail (specification (2)) investors, alpha (estimated using past 1 year of monthly fund returns to institutional (specification (1)) or retail (specification (2)) investors and the excess return on the market, SMB, HML, and momentum as risk factors), and lagged yearly return gap. All specifications include style fixed effects. In the last two columns, we compare the estimated return gap coefficient in specifications (1) and (2) and report the corresponding *t*-stats. We estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

magnitude of the estimated return gap coefficient using the subset of institutional investors is larger that the one using the subset of retail investors. The last two columns of Table 4 compare the estimated return gap coefficients between institutional and retail investors. Overall, the results are consistent with the notion that the more sophisticated investors are more able to separate good from bad managers than the less sophisticated investors.

C. Asymmetric sensitivity of fund flows to the return gap

A large number of empirical papers document that investors reward highly successful funds, but they tend not to withdraw money from poorly performing funds (e.g., Ippolito 1992; Chevalier and Ellison 1997; Sirri and Tufano 1998). These findings raise the possibility that the sensitivity of fund flows to the return gap might be driven by the investors' flows to funds with high return gap. Below we investigate this conjecture.

In order to test for potential nonlinearities in the sensitivity of fund flows to the return gap, we follow Sirri and Tufano (1998) and employ a piece-wise

linear approach. First, we calculate each fund's fractional rank RG_Rank_t which represents the fund's yearly return gap percentile relative to the rest of the funds in that period and ranges from 0 to 1. Then, we spread each fund's RG_Rank_t over five different quintiles in the following way:

$$RG\_Q1 = \min(0.2, RG\_Rank_t)$$
$$RG\_Q2 = \min(0.2, RG\_Rank_t - RG\_Q1)$$
$$RG\_Q3 = \min(0.2, RG\_Rank_t - RG\_Q1 - RG\_Q2)$$
$$RG\_Q4 = \min(0.2, RG\_Rank_t - RG\_Q1 - RG\_Q2 - RG\_Q3)$$
$$RG\_Q5 = RG\_Rank_t - RG\_Q1 - RG\_Q2 - RG\_Q3 - RG\_Q4$$

We also calculate RG_mid, which combines the middle three quintiles:

$$RG\_mid = \min(0.6, RG\_Rank_t - RG\_Q1)$$

Similarly, we split yearly fund return fractional rank into 5 quintiles $FR\_Q1–5$ and combine the middle three quintiles in an additional bucket, FR_mid. Similarly to most of our previous analyses, we include lagged flows, expense ratio, and alpha in the performance flow relationship. For brevity, we do not report their estimated coefficients.

The results are summarized in Table 5. In specifications (1) and (2), we offer results for the whole sample. The sensitivity of flows to past return gap appears to be very strong for funds with high return gaps. An increase in return gap among funds in the top return gap quintile (say from 80th to 90th percentile) is associated with significantly greater inflows in the following quarter (1.60%). To understand the economic importance of the findings, one has to multiply a given change in return gap percentile rank (scaled between 0 and 1) with the estimated return gap coefficient pertaining to funds in that return gap quintile. Importantly, investors appear to not respond to funds with particularly poor and even average return gaps.

In the rests of the specifications, we investigate the asymmetries in flows' sensitivity to the return gap, separately for institutional and retail investors. We find similar patterns as in the previous specifications. Both retail and institutional investors are characterized with a nonlinear sensitivity of fund flows to the return gap, over the 2000–2010 period. However, institutional investors respond stronger to past poor performance and put lower weight on stellar past performance. Among institutional investors, an increase in return gap among funds in the return gap quintile (say from 80th to 90th percentile) results in additional 1.6% flows in the following quarter. Under the same scenario, the increase in retail flows is expected to be 1.1%.

D. Time-varying sensitivity of fund flows to the return gap

In this paper, we hypothesize that investors can obtain information about future fund performance, which we proxy with the return gap. Of course,

Table 5 Asymmetric response to the return gap

	All funds				Institutional				Retail			
	(1) Flow$_t$		(2) Flow$_t$		(3) Flow$_t$		(4) Flow$_t$		(5) Flow$_t$		(6) Flow$_t$	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
FR_Q1$_{t-1}$	5.71***	2.66	6.67***	3.49	17.52***	4.67	17.92***	5.15	10.42***	4.79	10.37***	5.20
FR_Q2$_{t-1}$	9.36***	8.42			8.58***	2.89			7.00***	5.07		
FR_Q3$_{t-1}$	6.56***	5.99			5.99*	1.91			7.83***	5.48		
FR_Q4$_{t-1}$	8.31***	5.74			9.04***	2.52			6.48***	3.38		
FR_mid$_{t-1}$			7.86***	19.26			7.58***	7.70			7.23***	13.62
FR_Q5$_{t-1}$	38.11***	13.15	38.15***	15.14	25.08***	4.12	26.03***	4.86	50.02***	10.86	49.49***	12.24
RG_Q1$_{t-1}$	-0.20	-0.11	-1.99	-1.15	1.44	0.33	-0.13	-0.03	-1.52	-0.54	-2.62	-1.10
RG_Q2$_{t-1}$	-1.80	-1.37			-1.84	-0.58			-1.35	-0.70		
RG_Q3$_{t-1}$	0.16	0.13			5.63*	1.73			-2.39	-1.49		
RG_Q4$_{t-1}$	2.48*	1.84			-2.25	-0.65			2.88*	1.66		
RG_mid$_{t-1}$			0.27	0.64			1.28	1.41			-0.62	-1.17
RG_Q5$_{t-1}$	16.02***	6.02	17.87***	7.70	15.78***	2.69	13.81***	2.61	11.23***	3.41	13.86***	4.80
Controls	Yes		Yes		Yes		Yes		Yes		Yes	
Style FE	Yes		Yes		Yes		Yes		Yes		Yes	
Time FE	Yes		Yes		Yes		Yes		Yes		Yes	
R^2	0.09		0.09		0.05		0.05		0.11		0.11	
Observations	85,914		85,914		25,706		25,706		49,653		49,653	
Time period	Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.2000–Q3.2010		Q1.2000–Q3.2010		Q1.2000–Q3.2010		Q1.2000–Q3.2010	

The dependent variable is fund flows in quarter t in specifications (1) and (2), institutional flow in quarter t in specifications (3) and (4), and retail flow in quarter t in specifications (5) and (6). We use the identification of retail and institutional share classes introduced by CRSP at the end of 1999 and aggregate the flow, expenses, and return data separately for the retail and institutional part of a fund. We calculate each fund's fractional return gap rank RG_Rank which represents the fund's lagged yearly return gap percentile relative to the rest of the funds in that period and ranges from 0 to 1. We spread each fund's RG_Rank over five different quintiles: $RG\_Q1 = \min(0.2, RG\_Rank_t)$, $RG\_Q2 = \min(0.2, RG\_Rank_t - RG\_Q1)$, $RG\_Q3 = \min(0.2, RG\_Rank_t - RG\_Q1 - RG\_Q2)$, $RG\_Q4 = \min(0.2, RG\_Rank_t - RG\_Q1 - RG\_Q2 - RG\_Q3)$, and $RG\_Q5 = RG\_Rank_t - RG\_Q1 - RG\_Q2 - RG\_Q3 - RG\_Q4 - RG\_Q5$. We also combine the middle three quintiles in $RG\_mid = \min(0.6, RG\_Rank_t - RG\_Q1)$. In specifications (1) and (2) we use total return, in specifications (3) and (4) we use return specific to institutional investors, and in specifications (5) and (6) we use return specific to retail investors. We similarly split FR_Rank in FR_Q1, FR_Q2, FR_Q3, FR_Q4, FR_Q5, and FR_mid. In each specification, we include an intercept, the most recently available expense ratio, lagged yearly flow, alpha (estimated using past 1 year of monthly fund returns to all (specifications (1) and (2)), or institutional (specifications (3) and (4)), or retail (specifications (5) and (6)) investors and the excess return on the market, SMB, HML, and momentum as risk factors), and lagged yearly return gap. All specifications include style fixed effects. We estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

investors also take into account more direct performance indicators, such as past net returns and alpha. Given that the information content embedded in each of those measures may vary with time, we expect to find an increasing fund flow sensitivity to the return gap when the information embedded in other performance measures decreases. In other words, in times when standard performance measures are more informative about future performance, their relative importance ought to increase.

We use the cross-sectional standard deviation of fund returns to proxy for the amount of information embedded in fund returns. When the cross-sectional dispersion of fund returns is relatively low, investors extract less information from fund returns to distinguish good from bad managers than in periods when the dispersion is relatively high. Consequently, when the cross-sectional dispersion of fund returns is low, investors have to rely relatively more on other information.[11] Empirically, we include the interaction of the lagged yearly return gap with the standard deviation of fund returns during that year as an explanatory variable in the flow-performance relationship. We estimate the model using pooled regressions where we include quarter fixed effects and cluster the standard errors on the fund level.

The results from this exercise are summarized in Table 6. Consistent with the hypothesis that the information component captured by the return gap becomes more important when there is less information in total fund returns, we find the impact of the interactions between the return gap and the standard deviation of fund returns to be negative. The results suggest that there is substantial time-variation in the sensitivity of fund flows to the return gap. Moreover, the evidence is in accordance with the hypothesis that the relative importance of the information proxied by the return gap depends on the informativeness of other performance measures.

V. DO INVESTORS BENEFIT FROM DIRECTING CAPITAL TO HIGH RETURN GAP FUNDS?

The results in Section IV provide a number of empirical patterns consistent with the hypothesis that investors direct capital towards high return gap funds. This evidence suggests that investors realize positive risk-adjusted returns from directing capital towards funds likely to outperform and from avoiding funds likely to exhibit a poor performance.

Therefore, in this section we investigate to what extent investors enhance their returns by allocating capital towards funds likely to perform well in the future and withdrawing capital from funds likely to underperform in the future. To test this, we first estimate a flow-performance model using the 1990–2010

---

11 Similarly to Kacperczyk et al. (2016), we argue that information variables with high dispersion contain relatively more information. Important difference between our work and theirs is that whereas they aim to answer when return dispersion increases, we study investors' response when dispersion is low *versus* high.

Table 6 Investors' response to the return gap—Time-dimension

	(1) Flow$_t$	
	Coeff	*t*-stat
Intercept	−0.01**	−1.97
ExpRatio$_{t-1}$	0.64*	1.66
YearlyFlow$_{t-1}$	0.00	−1.05
YearlyFundReturn$_{t-1}$	0.22***	27.08
Alpha$_{t-1}$	1.48***	8.69
YearlyReturnGap$_{t-1}$	0.34***	5.42
YearlyReturnGap$_{t-1} \times \sigma$(Return$_{t-1}$)	−2.66***	−3.19
Style FE	Yes	
Time FE	Yes	
R^2	0.08	
Observations	85,914	
Time period	Q1.1990–Q3.2010	

The dependent variable is fund flow in quarter *t*. We include an intercept, the most recently available expense ratio, lagged yearly flow, lagged yearly fund net return, alpha (estimated using past 1 year of monthly fund returns and the excess return on the market, SMB, HML, and momentum as risk factors), lagged yearly return gap, and the interaction between the lagged yearly return gap and the standard deviation of fund returns during the same period. All specifications include style fixed effects. We estimate the model using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

sample, where the dependent variable is Flow$_t$ and on the right hand side there are lagged alpha, lagged expense ratio, lagged flows, and *FR_Q1*, *FR_mid*, and *FR_Q5*. We call this the restricted model. For each fund in each quarter, we calculate an expected flow using the estimated coefficients from the restricted model and the respective realizations of the independent variables. Next, we estimate a more general model, which expands the restricted specification with three additional explanatory variables—*RG_Q1*, *RG_mid*, and *RG_Q5*, which we orthogonalize with respect to the variables in the restricted model.

Then, for each fund in each quarter, we calculate the difference between the expected flow based on the unrestricted model and the expected flow based on the restricted model. We term this difference "Expected Flow Difference". At the end of each quarter, we sort funds in 10 portfolios based on that quarter's "Expected Flow Difference" and track their performance over the subsequent quarter. The top decile contains funds with the highest "Expected Flow Difference" and the bottom one those with the lowest "Expected Flow Difference". This way we obtain a time-series of portfolio returns and evaluate their performance using a four-factor model, including the return on the market, SMB, HML, and momentum. We report results using both equally and flow-weighted portfolios.

This methodology allows us to evaluate the performance of fund flows that are due to the information component that we proxy with the return gap. If

the "Expected Flow Difference" score for a fund is positive (negative), investors' response to the return gap has increased (decreased) the assets under management for that particular fund. Consequently, the difference in subsequent risk-adjusted performance between funds with positive and negative "Expected Flow Difference" captures the extent to which investor returns are enhanced by allocating capital towards high return gap funds and withdrawing capital from funds with low return gap.

The results, using the whole set of funds over 1990–2010, are summarized in panel A of Table 7. The excess return on each of the spread portfolios is positive and statistically significant at conventional levels, irrespective of the estimation method and the weighting scheme. The four-factor monthly alpha of the spread portfolio is economically important, ranging between 0.18% and 0.21% per month, depending on the specification. The Spearman rank correlation between the portfolio rank and the calculated flows rejects the null of no relationship, indicating that despite the small differences between portfolios, the patterns are monotonic. Overall, the results suggest that investors realize non-negligible gains from directing capital towards funds with high return gaps and more importantly, from avoiding funds with low return gaps.[12]

The results are also consistent with the hypothetical return of a trading strategy, documented by Kacperczyk et al. (2008). They sort funds in 10 deciles based on their average monthly return gap during the past 12 months, and examine their subsequent results. Their results indicate that a strategy long in the top decile and short in the bottom decile generates a subsequent four factor alpha of 0.22% per month, consistent with the 0.21% we find.

In panels B and C of Table 7 we repeat the exercise in panel A, using the subsets of institutional and retail investors (and necessarily restricting the sample to the most recent decade). The only difference with respect to the exercise using all funds is that we estimate separately the restricted and unrestricted models for each subgroup of funds, on the basis of which we construct the expected flow measures. Even though there is a similar pattern of increasing performance from bottom to top deciles, the spread portfolios for both institutional and retail investors are generally not statistically different from zero. We attribute this to the lower statistical power of the test since the analysis of the institutional and retail subsamples is based on 10 years of data only.

VI. ADDITIONAL TESTS

In this part of the paper, we conduct a number additional test which aims at providing a clearer understanding of the drivers of the documented fund flows sensitivity to the return gap. We first test whether investors are guided towards funds with expected positive future performance by financial advisers and

---

12 The finding that investors enhance their returns through their response to the information captured by the return gap does not necessarily imply that their overall allocation is "smart" in the sense of Gruber (1996) and Zheng (1999).

brokers. Next, we show that information contained in the return gap is not captured by observable fund characteristics. Next, we incorporate the accuracy of the return gap in our analysis. If investors respond to information signals for which the return gap is a noisy measure, we expect the responsiveness of money flows to the return gap to be stronger if the return gap is more accurately estimated. Finally, we examine the robustness of our findings to a model specification using quarterly measured control variables.

A. The role of financial advisors and brokers

The empirical results in the previous sections suggest that investors can distinguish between value-adding and value-destroying funds. A potential explanation to this finding is that investors are directed towards good fund managers by financial advisers and brokers (e.g., Bergstresser et al. 2009; Del Guercio et al. 2010). To test for this conjecture, we check if the previously documented sensitivity of fund flows to the return gap is driven by investors who use financial advisers and brokers.

We split the data sample in two subsamples—load and no-load funds. We define a load fund share class as a share class with a front-load or a back-end load or with 12b-1 fees above 25 basis points. Information on load fees is available in the CRSP database since 1999. Similarly to the split of institutional *versus* retail investors in Section B, we aggregate fund information separately for the load and no-load part of a fund and obtain separate flow and return data for investors using the services of brokers and financial advisers and those who do not. This allows us to separately estimate the flow-performance relationship for two subsamples—one for the subset of investors using the services of brokers and financial advisers, and one for the subset of investors who do not use such services.

If the flows' sensitivity to the return gap documented previously is entirely driven by the advise of brokers and financial advisors, we should observe no sensitivity to the return gap in the no-load subsample. The results in Table 8 suggest that this is not the case and are in line with studies finding that there are limits to advice by professional investors (e.g., Bodnaruk and Simonov 2015). Investors in no-load funds respond very strongly to the lagged return gap measures where almost all of the coefficients are larger in magnitude than those in the load sample. This indicates that the sensitivity of fund flows to the return gap cannot be explained by the help investors receive by financial advisors and brokers.

B. The return gap and observable information

Table B1 demonstrates that very little in the variation of the return gap can be explained by observable fund characteristics. However, for robustness, we include the determinants of the return gap in the flow-performance relationship and check if any of these variables drives the main effects. As an additional

Table 7 Economic effect from the sensitivity of fund flows to the return gap

	A: All funds				B: Institutional funds				C: Retail funds			
	Equally weighted		Flow-weighted		Equally weighted		Flow-weighted		Equally weighted		Flow-weighted	
1 (lowest)	-0.16	-1.40	-0.17	-1.32	-0.15	-1.51	-0.13	-1.24	-0.11	-0.76	-0.10	-0.64
2	-0.09**	-2.35	-0.07	-1.47	-0.16**	-2.45	-0.16**	-2.21	-0.11	-1.38	-0.08	-1.23
3	-0.06	-1.38	-0.03	-0.98	-0.06	-0.79	-0.10	-1.48	-0.03	-0.53	-0.04	-0.67
4	-0.07	-1.54	-0.06	-1.36	-0.07	-1.18	-0.06	-1.05	-0.04	-0.54	-0.05	-0.83
5	-0.05	-0.86	-0.04	-0.88	-0.05	-0.71	-0.01	-0.15	-0.03	-0.36	-0.06	-0.71
6	-0.04	-0.91	-0.08	-1.40	-0.05	-0.68	-0.09	-1.22	0.02	0.21	-0.04	-0.46
7	-0.04	-0.80	-0.07	-1.16	-0.04	-0.42	-0.05	-0.17	-0.09	-1.61	-0.05	-0.56
8	-0.05	-0.83	-0.07	-1.06	-0.04	-0.91	-0.02	-0.25	-0.05	-0.93	-0.03	-0.44
9	-0.07	-0.92	-0.04	-0.62	-0.08	-0.95	0.01	0.15	-0.06	-1.05	-0.07	-0.92
10 (highest)	0.03	0.20	0.04	0.36	-0.04	-0.48	0.04	0.46	-0.03	-0.28	-0.01	-0.14
10–1	0.18**	1.98	0.21**	2.06	0.11	1.43	0.17*	1.85	0.08	0.56	0.09	0.64
SpCorr t-stat	3.41		2.04		3.32		4.29		1.60		2.90	
Time period	Q2.1990–Q4.2010		Q2.1990–Q4.2010		Q2.2000–Q4.2010		Q2.2000–Q4.2010		Q2.2000–Q4.2010		Q2.2000–Q4.2010	

We regress fund flows on an intercept, alpha (estimated using past 1 year of monthly fund returns and the excess return on the market, SMB, HML, and momentum as risk factors), the most recently available expense ratio, lagged yearly fund flow, and FR_Q1_{t-1}, FR_mid_{t-1}, and FR_Q5_{t-1} (defined in Table 5). We call this the restricted model. We also regress fund flows on the same set of variables and RG_Q1_{t-1}, RG_mid_{t-1}, and RG_Q5_{t-1} (defined in Table 5), which we orthogonalize with respect to the variables in the restricted model. We call this the unrestricted model. At the end of each quarter t we use the estimated coefficients from the restricted and the unrestricted models and the respective time-specific realizations of the independent variables (i.e., fund flows in quarters t, $t-1$, $t-2$, and $t-3$) to construct two expected flow scores. We calculate "Expected Flow Difference" for each fund in quarter $t+1$ as the difference between the expected flow based on the unrestricted model and the expected flow based on the restricted model. Next, we sort funds in 10 portfolios based on their "Expected Flow Difference" scores and track their performance until the end of quarter $t+1$ when we rebalance the portfolios. This way we obtain a time-series of monthly returns for each portfolio. Next, we evaluate the performance of each time-series of portfolio returns using a four-factor asset pricing model, where we use the excess return on the market, SMB, HML, and momentum as risk factors. We report results estimating the restricted and unrestricted models we report the alpha and the corresponding t-statistic. In panel A, the sample covers the whole data set. In panels B and C, we use the subsamples of institutional and retail investors (defined in Table 4). Note that the explanatory variables used for estimating the restricted and unrestricted are defined in the same way as in Table 4. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively. SpCorr t-stat denotes the t-statistic on the spearman rank correlation coefficient between the estimated economic effects and the ranks of the portfolios.

Table 8 Investors' response to the return gap—load *versus* no-load funds

	Load (1) Flow$_t$		No-load (2) Flow$_t$	
	Coeff	*t*-stat	Coeff	*t*-stat
Intercept	0.00	0.22	0.02***	3.63
ExpRatio$_{t-1}$	0.80***	2.89	−0.01	−0.02
YearlyFlow$_{t-1}$	0.01**	2.13	0.01	1.38
YearlyFundReturn$_{t-1}$	0.06***	6.32	0.05***	8.67
Alpha$_{t-1}$	2.29***	11.17	3.31***	11.74
YearlyReturnGap$_{t-1}$	0.24***	6.92	0.15***	4.76
Style FE	Yes		Yes	
Time FE	Yes		Yes	
R^2	0.07		0.04	
Observations	54,955		42,731	
Time period	Q1.2000–Q3.2010		Q1.2000–Q3.2010	

We define load share classes as share classes having front-end or rear-end load (CRSP reports this information from the end of 1999) or with a 12b-1 fee that is higher than 0.25% per year. Consequently, we aggregate the flow, expenses, and return data separately for the load and no-load part of a fund. The dependent variable in specification (1) is load flow in quarter *t*, and in specifications (2) is no-load flow in quarter *t*. In each specification we include an intercept, alpha (estimated using past 1 year of monthly fund returns to load (specification (1)) or no-load (specification (3)) investors and the excess return on the market, SMB, HML, and momentum as risk factors), the most recently available expense ratio, specific to load (specification (1)) or no-load (specification (3)) investors, lagged yearly fund flow, and lagged yearly fund return, specific to load (specification (1)) or no-load (specifications (2)) investors. Both specifications also include lagged yearly return gap, calculated according to the procedure described in Section II. All specifications include style fixed effects. We estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

control, we add a variable indicating whether a fund is a "star fund". To construct this varibale, we collect data on Morningstar's star ratings. Previous research has documented that funds that experience an increase in their star ratings during the last year receive significantly higher inflows from investors (Del Guercio and Tkac 2008; Nanda et al. 2004). Therefore, in specifications (1) of Table 9, we include a dummy for an increase in a fund's star rating by Morningstar following Del Guercio and Tkac (2008). Results indicate that the return gap significantly predicts flows, even after including the star dummy variable.

Barber et al. (2005) show that marketing expenses are important determinants of fund flows. They propose front-load and 12b-1 fees as proxies for marketing expenses—the former is related to distribution payments to brokers and the latter captures advertising expenditure. In specification (2), we find results consistent with Barber et al. (2005)—fund flows are negatively related to front-load charges and positively to 12b-1 fees. Thus, as Barber et al. (2005) argue, investors respond negatively to the salient front-load charges but marketing

Table 9 Investors' response to the return gap, controlling for correlated performance measures

	(1) Flow$_t$		(2) Flow$_t$		(3) Flow$_t$		(4) Flow$_t$		(5) Flow$_t$		(6) Flow$_t$	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
ReturnGap$_{t-1}$	0.10***	3.79	0.15***	3.47	0.08**	2.23	0.13***	4.48	0.14***	4.64	0.14***	4.67
Star fund$_t$	0.02***	22.33			0.03***	9.68						
Front-load$_t$			-0.12**	-2.28	0.10	0.61						
12b-1$_t$			1.07**	1.99	1.30	0.89						
Trading costs$_t$							0.15***	4.57				
Weight of recent IPOs$_t$									0.13	0.96		
ρ(holding return and net returns)$_{t-1}$											-0.19***	-4.93
Yearly Turnover$_t$												
log(fund TNA)$_{t-1}$												
log(family TNA)$_{t-1}$												
log(age)$_{t-1}$												
σ(Return$_{t-5\ to\ t-1}$)												
Controls	Yes		Yes		Yes		Yes		Yes		Yes	
Style FE	Yes		Yes		Yes		Yes		Yes		Yes	
Time FE	Yes		Yes		Yes		Yes		Yes		Yes	
R^2	0.13		0.11		0.13		0.11		0.11		0.11	
Observations	75,868		49,781		44,997		85,771		85,914		85,914	
Time period	Q1.1990–Q3.2010		Q1.1993–Q3.2010		Q1.1993–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010	

	(7) Flow$_t$		(8) Flow$_t$		(9) Flow$_t$		(10) Flow$_t$		(11) Flow$_t$		(12) Flow$_t$	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
ReturnGap$_{t-1}$	0.13***	4.52	0.14***	4.74	0.14***	4.78	0.13***	4.42	0.14***	4.89	0.07**	2.27
Star fund$_t$											0.03***	18.70
Front-load$_t$											0.03	0.59
12b-1$_t$											-0.47	-0.96
YearlyFundReturn$_t$ × Front-load$_t$												

Table 9 (*continued*)

	(7) Flow$_t$		(8) Flow$_t$		(9) Flow$_t$		(10) Flow$_t$		(11) Flow$_t$		(12) Flow$_t$	
	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat	Coeff	t-stat
YearlyFundReturn$_t$ × 12B1$_t$												
Star fund$_t$ × Front-load$_t$												
Star fund$_t$ × 12B1$_t$												
Trading costs$_t$											0.07	1.13
Weight of recent IPOs$_t$											0.10	0.61
ρ(holding return and net returns)$_{t-1}$											−0.15***	−2.80
Yearly Turnover$_t$	0.01**	2.52									0.00	−0.85
log(fund TNA)$_{t-1}$			−0.01***	−11.59							−0.01***	−7.90
log(family TNA)$_{t-1}$					0.00	−0.04					0.00***	6.30
log(age)$_{t-1}$									−0.09	−1.59	−0.01***	−4.21
σ(Return$_{t-5}$ to $_{t-1}$)							−0.02***	−20.23			0.15*	1.76
Controls	Yes		Yes		Yes		Yes		Yes		Yes	
Style FE	Yes		Yes		Yes		Yes		Yes		Yes	
Time FE	Yes		Yes		Yes		Yes		Yes		Yes	
R^2	0.11		0.11		0.10		0.13		0.11		0.15	
Observations	85,771		85,914		79,561		85,914		85,914		41,309	
Time period	Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1993–Q3.2010		Q1.1990–Q3.2010		Q1.1990–Q3.2010		Q1.1993–Q3.2010	

The dependent variable in each regression specification is fund flow in quarter *t*. In each specification we include an intercept, alpha (estimated using past 1 year of monthly fund returns and the excess return on the market, SMB, HML, and momentum as risk factors), the most recently available expense ratio, lagged yearly fund flow, lagged yearly fund net return, and lagged yearly return gap, calculated according to the procedure described in Section II. Each of the variables is defined in Appendix B. All specifications include style fixed effects. We estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

expenses bring more money under management. To shed more light on the advertising channel, we further interact the two marketing expense variables with past performance and the star dummy variable. Importantly, even after including the new variables, the return gap remains a statistically significant predictor of fund flows. In specifications (3) to (11), we examine the separate effect of each of the determinants of the return gap in conjunction with the return gap. We document very small changes in the return gap coefficient, indicating that none of the controls single-handedly subsumes the effect of the return gap. In specification (12), we include all of the flow drivers on the right hand side. Again, the return gap coefficient remains significant. In sum, the results in Table 9 suggest that observable fund characteristics cannot explain the sensitivity of fund flows to the return gap.

C. Precision of the return gap

Our main results in Table 3 show that the sensitivity of fund flows to the return gap is positive. However, some of the calculated return gaps might be noisy indicators of future performance. For example, managers may manipulate their reported holdings in order to present themselves as more able. The managers may window dress their portfolios, which refers to buying (selling) stocks with past positive (negative) performance shortly before reporting the holdings to the public in order to convey stock-picking skills. Portfolio pumping, referring to buying shares in the stocks the fund already owns on the last day of the reporting period, is another practice used by some managers to inflate their performance.[13] Both practices would add noise to the return gap as an indicator for future performance.

Another reason why there might be noise in the return gaps we estimate comes from the data limitations of our sample. Although small, the share of non-equity holdings in the portfolios of the mutual funds in our sample is non-zero. The quarterly snapshots of the funds' portfolios do not include their non-equity holdings. Consequently, to calculate the quarterly return of the portfolio of fund holdings we assume that the fund's yearly asset class allocation provided by CRSP is constant over time. However, funds may decide to actively manage their asset class allocations and have, for example, lower cash holdings in some quarters, while having higher cash holdings in other quarters. This, in turn, would add noise to the return gap measures we calculate.

Consequently, if investors base their capital allocation decisions on information about future performance which is correlated with the return gap, one would expect that the sensitivity of fund flows is weaker if the precision of the return gap is lower. To investigate this, we first calculate monthly return gaps in the 12 previous months. Next, we construct two additional return gap variables: *t*-statistic (*RG_t*) and standard deviation (*RG_stdev*). In specification (1) of

---

13 Window dressing and portfolio pumping, for example, see Lakonishok et al. (1991), Musto (1999), Carhart et al. (2002), and Agarwal et al. (2014).

Table 10, we find that fund flows respond positively to *RG_t.* Thus, the more precise the return gap, the higher the inflows. In specification (2), we include the return gap, *RG_stdev*, and the interaction between the two. We find negative coefficients on the interactions of the return gap with its standard deviation, implying that investors allocate more capital towards funds with more precise return gaps. Overall, the results suggest that a more precisely estimated return gap results in higher fund flows.

D. Robustness tests

Our main test is based on results using yearly estimated return gap, alpha, fund return, alpha, and flows as control variables. In Table 11, we investigate the sensitivity of fund flows to the return gap, where the key independent variables are measured on quarterly frequency. Results remain: fund flows respond strongly to the return gap, even when variables are measured on quarterly level. In

Table 10 The precision of the return gap

	(1) Flow$_t$		(2) Flow$_t$	
	Coeff	*t*-stat	Coeff	*t*-stat
Intercept	−0.01***	−2.60	−0.01*	−1.76
ExpRatio$_{t-1}$	0.25	1.09	0.37	1.35
YearlyFundReturn$_{t-1}$	0.23***	23.98	0.22***	23.64
Alpha$_{t-1}$	2.61***	13.75	2.50***	16.12
YearlyFlow$_{t-1}$	0.01***	4.22	0.01***	4.23
YearlyReturnGap$_{t-1}$			0.17***	6.15
RG_t	0.06**	2.10		
RG_stdev			0.59	1.16
ReturnGap$_{t-1}$ × RG_stdev			−3.41***	−2.78
Style FE	Yes		Yes	
R^2	0.11		0.11	
Observations	85,914		85,914	
Time period	Q1.1990–Q3.2010		Q1.1990–Q3.2010	

The dependent variable in each regression specification is fund flow in quarter *t.* In each specification, we include four lagged return gap scores, calculated according to the procedure described in Section II. In specifications (1) and (3), we include interactions of the four return gaps with the *t*-statistic of the return gap, calculated from monthly return gap scores in the past 12 months. In specifications (2) and (4), we include interactions of the four return gaps with the standard deviation of the monthly return gaps during the past 12 months. In each specification, we include an intercept, alpha (estimated using past 1 year of monthly fund returns and the excess return on the market, SMB, HML, and momentum as risk factors), the most recently available expense ratio, four lagged quarterly fund flow measures, and four lagged fund net return measures. All specifications include style fixed effects. In specifications (1) and (2), we estimate the models using Fama–Macbeth regressions where we report *t*-statistics based on Newey–West standard errors with 3 lags. In specifications (3) and (4), we estimate the models using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

Table 11 Investors' response to the return gap using quarterly measured variables

	(1) $Flow_t$	
	Coeff	*t*-stat
Intercept	−0.01[*]	−1.80
$ExpRatio_{t-1}$	0.35	1.31
$ReturnGap_{t-1}$	0.12[***]	3.37
$ReturnGap_{t-2}$	0.17[***]	5.64
$ReturnGap_{t-3}$	0.15[***]	4.40
$ReturnGap_{t-4}$	0.13[***]	3.55
$FundReturn_{t-1}$	0.32[***]	15.68
$FundReturn_{t-2}$	0.14[***]	7.29
$FundReturn_{t-3}$	0.16[***]	10.01
$FundReturn_{t-4}$	0.10[***]	6.17
$Alpha_{t-1}$	0.09[***]	5.08
$Alpha_{t-2}$	0.07[***]	4.60
$Alpha_{t-3}$	0.06[***]	4.38
$Alpha_{t-4}$	0.05[***]	3.03
$Flow_{t-1}$	0.09[***]	3.40
$Flow_{t-2}$	0.06[***]	5.86
$Flow_{t-3}$	0.02[***]	2.91
$Flow_{t-4}$	0.01[***]	2.64
Style FE	Yes	
Time FE	Yes	
R^2	0.13	
Observations	84,087	
Time period	Q1.1990–Q3.2010	

The dependent variable specification is fund flow in quarter *t*. Control variables include four lagged quarterly alphas. To calculate an alpha in a given quarter, we first estimate factor loadings on the excess return on the market, SMB, HML, and momentum using the 12 months preceding the quarter. Then, we use realized fund and risk factor returns together with the estimated alphas to compute quarterly alpha. We further include four lagged return gaps, the most recently available expense ratio, four lagged expense ratios and for lagged fund net returns. All specifications include style fixed effects. We estimate the model using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

unreported tests, we further establish that our key results (i) hold for the subset of fund funds which hold more than 95% of their assets in equities; and (ii) are robust to the inclusion of fund fixed effects.

VII. CONCLUSION

We conjecture that mutual fund investors possess information about future performance which is likely to not be reflected in fund characteristics and past performance. We proxy this information with the return gap of Kacperczyk et al. (2008), which is constructed as the difference between the net return of the fund and the return of the most recently disclosed portfolio

holdings. Apparently, the return gap captures information that has high predictive value for future fund performance. Consequently, if fund investors possess information about future performance, it is likely to be reflected in the return gap.

Our main findings are consistent with the hypothesis that investors are able to distinguish good from bad funds. We find a strong sensitivity of fund flows to the return gap, which increases with investors sophistication. We find that the sensitivity of fund flows of the return gap is highly nonlinear, potentially due to the costs associated with identifying information about future performance, and increases in times when the information content embedded in other performance measures decreases. In order to assess the economic importance of investors' response to the return gap, we analyze the flow component associated with the sensitivity to the return gap. We document that investors enhance their returns with about 2% per year, particularly by avoiding funds likely to destroy value in the future.

This paper contributes to our understanding of investor sophistication and the drivers of fund flows. Much of the empirical work has focused on showing that investors are naive and inexperienced in their investment decisions. However, we show that investors posses the ability to separate funds likely to perform well in the future from those likely to perform poorly, using information which may not be readily available. Thus, this paper provides empirical evidence for the empirically contested assumption in most of the theoretical literature that there is a significant degree of investor sophistication.

Marno Verbeek
Rotterdam School of Management
Erasmus University Rotterdam
Burgemeester Oudlaan 50, P.O. Box 1738
3000DR Rotterdam
The Netherlands
mverbeek@rsm.nl

REFERENCES

Agarwal, V., G. D. Gay, and L. Ling (2014), 'Window Dresing in Mutual Funds', *The Review of Financial Studies*, 27, 3133–70.

Bailey, W., A. Kumar, and D. Ng (2011), 'Behavioral Biases of Mutual Fund Investors', *Journal of Financial Economics*, 102, 1–27.

Baker, M. et al. (2010), 'Can Mutual Fund Managers Pick Stocks? Evidence from Their Trades Prior to Earnings Announcements', *Journal of Financial and Quantitative Analysis*, 45, 1111–31.

Barber, B. M., T. Odean, and L. Zheng (2005), 'Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows', *The Journal of Business*, 78, 2095–120.

Barber, B. M., X. Huang, and T. Odean (2016), 'Which Factors Matter to Investors? Evidence from Mutual Fund Flows', *Review of Financial Studies*, 29, 2600–42.

Barras, L., O. Scaillet, and R. Wermers (2010), 'False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas', *The Journal of Finance*, 65, 179–216.

Bergstresser, D., J. M. R. Chalmers, and P. Tufano (2009), 'Assessing the Costs and Benefits of Brokers in the Mutual Fund Indsutry', *The Review of Financial Studies*, 22, 4129–56.

Berk, J., and R. Green (2004), 'Mutual Fund Flows and Performance in Rational Markets', *Jounal of Political Economy*, 112, 1269–95.

Berk, J. B., and J. H. van Binsbergen (2016), 'Assessing Asset Pricing Models Using Revealed Preferences', *Journal of Financial Economics*, 119, 1–23.

Bodnaruk, A., and A. Simonov (2015), 'Do Financial Experts Make Better Financial Decisions?', *Journal of Financial Intermediation*, 24, 514–36.

Bollen, N. P. B., and J. A. Busse (2005), 'Short-Term Persistence in Mutual Fund Performance', *The Review of Financial Studies*, 18, 569–97.

Bollen, N. P. B., and J. A. Busse (2006), 'Tick Size and Institutional Trading Costs: Evidence from Mutual Funds', *Journal of Financial and Quantitative Analysis*, 41, 915–37.

Carhart, M. M. (1997), 'On Persistence in Mutual Fund Performance', *The Journal of Finance*, 52, 57–82.

Carhart, M. M. et al. (2002), 'Learning for the Tape: Evidence of Gaming Behavior in Equity Mutual Funds', *The Journal of Finance*, 57, 661–93.

Casavecchia, L., and A. Tiwari (2016), 'Cross Trading by Investment Advisers: Implications for Mutual Fund Performance', *Journal of Financial Intermediation*, 25, 99–130.

Chen, J. et al. (2004), 'Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization', *The American Economic Review*, 94, 1276–302.

Chevalier, J., and G. Ellison (1997), 'Risk Taking by Mutual Funds as a Response to Incentives', *Journal of Political Economy*, 105, 1167–200.

Choi, J. J., D. Laibson, and B. C. Madrian (2010), 'Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds', *The Review of Financial Studies*, 23, 1405–32.

Cohen, R. B., J. D. Coval, and L. Pastor (2005), 'Judging Fund Managers by the Company They Keep', *The Journal of Finance*, 60, 1057–96.

Cooper, M. J., H. Gulen, and P. R. Rau (2005), 'Changing Names with Style: Mutual Fund Name Changes and Their Effects on Fund Flows', *The Journal of Finance*, 60, 2825–58.

Coval, J., and E. Stafford (2007), 'Asset Fire Sales (and Purchases) in Equity Markets', *Journal of Financial Economics*, 86, 479–512.

Cremers, K. J. M., and A. Petajisto (2007), 'How Active Is Your Fund Manager? A New Measure That Predicts Performance', *The Review of Financial Studies*, 22, 3329–65.

Del Guercio, D., and P. A. Tkac (2008), 'Star Power: The Effect of Morningstar Ratings on Mutual Fund Flows', *Journal of Financial and Quantitative Analysis*, 43, 907–36.

Edelen, R., R. Evans, and G. Kadlec (2013), 'Shedding Light on "Invisible" Costs: Trading Costs and Mutual Fund Performance', *Financial Analyst Journal*, 69, 33–44.

Elton, E., M. J. Gruber, and C. R. Blake (1996), 'The Persistence of Risk-Adjusted Mutual Fund Performance', *The Journal of Business*, 69, 133–57.

Evans, R. B., and R. Fahlenbrach (2012), 'Institutional Investors and Mutual Fund Governance: Evidence from Retail-Institutional Fund Twins', *The Review of Financial Studies*, 25, 3530–71.

Fama, E. F., and K. R. French (2010), 'Luck Versus Skill in the Cross-Section of Mutual Fund Returns', *The Journal of Finance*, 65, 1915–47.

Fama, E., and J. D. Macbeth (1973), 'Risk, Return, and Equilibrium: Empirical Tests', *Journal of Political Economy*, 81, 607–36.

Frazzini, A., and O. A. Lamont (2008), 'Dumb Money: Mutual Fund Flows and the Cross-Section of Stock Returns', *Journal of Financial Economics*, 88, 299–322.

Gaspar, J.-M., M. Massa, and P. Matos (2006), 'Favoritism in Mutual Fund Families? Evidence on Strategic Cross-Fund Subsidization', *The Journal of Finance*, 61, 73–104.

Grinblatt, M., and S. Titman (1989), 'Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings', *The Journal of Business*, 62, 393–416.

Gruber, M. J. (1996), 'Another Puzzle: The Growth in Actively Managed Mutual Funds', *The Journal of Finance*, 51, 783–810.

Del Guercio, D., J. Reuter, and P. A. Tkac (2010). 'Demand for Financial Advice, Broker Incentives, and Mutual Fund Market Segmentation', Working Paper https://ssrn.com/abstract=1361710.

Haslem, J. A. (2009), 'Investor Learning and Mutual Fund Advertising and Distribution Fees', *Journal of Investing*, 18, 53–6.

Hendricks, D., J. Patel, and R. Zeckhauser (1993), 'Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance', *The Journal of Finance*, 48, 93–130.

Huang, J., K. D. Wei, and H. Yan (2007), 'Participation Costs and the Sensitivity of Fund Flows to Past Performance', *The Journal of Finance*, 62, 1273–311.

Huang, J., K. D. Wei, and H. Yan (2012). 'Investor Learning and Mutual Fund Flows', Working Paper https://ssrn.com/abstract=972780.

Ippolito, R. A. (1992), 'Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry', *Journal of Law and Economics*, 35, 45–70.

Jain, P. C., and J. S. Wu (2000), 'Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows', *The Journal of Finance*, 55, 937–58.

Jiang, G. J., T. Yao, and T. Yu (2007), 'Do Mutual Funds Time the Market? Evidence from Portfolio Holdings', *Journal of Financial Economics*, 86, 724–58.

Kacperczyk, M., and A. Seru (2007), 'Fund Manager Use of Public Information: New Evidence on Managerial Skill', *The Journal of Finance*, 62, 485–528.

Kacperczyk, M., C. Sialm, and L. Zheng (2005), 'On the Industry Concentration of Actively Managed Mutual Funds', *The Journal of Finance*, 60, 1983–2011.

Kacperczyk, M., C. Sialm, and L. Zheng (2008), 'Unobserved Actions of Mutual Funds', *The Review of Financial Studies*, 21, 2379–416.

Kacperczyk, M., S. van Nieuwerburgh, and L. Veldkamp (2016), 'A Rational Theory of Mutual Funds' Attention Allocation', *Econometrica*, 84, 571–626.

Lakonishok, J. et al. (1991), 'Window Dressing by Pension Fund Managers', *The American Economic Review*, 81, 227–31.

Lewellen, J. (2011), 'Institutional Investors and the Limits of Arbitrage', *Journal of Financial Economics*, 102, 62–80.

Lynch, A., and D. K. Musto (2003), 'How Investors Interpret Past Fund Returns', *The Journal of Finance*, 58, 2033–58.

Musto, D. K. (1999), 'Investment Decisions Depend on Portfolio Disclosure', *The Journal of Finance*, 54, 935–52.

Nanda, V., Z. J. Wang, and L. Zheng (2004), 'Family Values and the Star Phenomenon: Strategies of Mutual Fund Families', *The Review of Financial Studies*, 17, 667–98.

Nimalendran, M., J. R. Ritter, and D. Zhang (2007), 'Do Today's Trades Affect Tomorrow's IPO Allocations', *Journal of Financial Economics*, 84, 87–109.

Reuter, J. (2006), 'Are IPO Allocations for Sale? Evidence from Mutual Funds', *The Journal of Finance*, 61, 2289–324.

Sialm, C., L. T. Starks, and H. ZHang (2015), 'Defined Contribution Pension Plans: Sticky or Discerning Money?', *The Journal of Finance*, 70, 805–38.

Sirri, E., and P. Tufano (1998), 'Costly Search and Mutual Fund Flows', *The Journal of Finance*, 53, 1589–622.

Solomon, D. H., E. Soltes, and D. Sosyura (2014), 'Winners in the Spotlight: Media Coverage of Fund Holdings as a Driver of Flows', *Journal of Financial Economics*, 113, 53–72.

Wermers, R. (2000), 'Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transaction Costs, and Expenses', *The Journal of Finance*, 55, 1655–95.

Zheng, L. (1999), 'Is Money Smart? A Study of Mutual Fund Investors' Fund Selection Ability', *The Journal of Finance*, 54, 901–33.1.

APPENDIX A

A. Database construction and sample selection

We start by selecting all US open-ended mutual funds from the CRSP mutual fund database and Thomson Financial/CDA database from January 1990 to June 2010. To ensure that we cover the universe of domestic diversified equity funds, for which the holdings data are most reliable, we select in our sample only funds with one of the following objective codes, provided by Lipper, Wiesenberger, and Strategic Insight and available in the CRSP Mutual Fund Database:

- Lipper: "EI", "EIEI", "EMN", "FLX", "G", "GI", "I", "LCCE", "LCGE", "LCVE", "LSE", "MC", "MCCE", "MCGE", "MCVE", "MLCE", "MLGE", "MLVE", "SCCE", "SCGE", "SCVE", "SESE", "SG"
- Wiesenberger:"SCG", "AGG", "G", "G-S", "S-G", "GRO", "LTG", "I", "I-S", "IEQ", "ING", "GCI", "G-I", "G-I-S", "G-S-I", "I-G", "I-G-S", "I-S-G", "S-G-I", "S-I-G", "GRI", "MCG"
- Strategic insight: "SCG", "GRO", "AGG", "ING", "GRI", "GMC"

Furthermore, we include funds only if they have one of the following investment objective codes in the Thomson Financial database: aggressive growth, growth, growth and income, or unclassified, thus excluding any international, bond, asset allocation, precious metal, and sector funds. Then, we drop funds that hold less than 80% or more than 105% in common stocks, as reported by CRSP. We also drop index funds by removing funds that contain in their CRSP-reported fund name the strings "INDEX", "INDE", "INDX", "S&P", or "MSCI". From Thomson Financial database, we remove overlapping report dates and file dates caused by fund mergers and name changes. We also delete funds that hold less than 10 stocks or manage less than $5 million.

If a fund offers multiple share classes to investors, we aggregate across different share classes. For TNA under management, we sum the TNAs of individual shares. For funds' age, we select the age of the oldest share class. For the other fund attributes (expenses, turnovers, *etc.*), we take the weighted average of the attributes of the individual share classes, where the weights are the lagged TNAs of the individual share classes.

We link the two mutual fund databases, using the MFLINKS database provided by WRDS. More information on how MFLINKS assigns a unique fund identifier to each fund in the two databases can be found in Wermers (2000). We check the MFLINKS databases for assigning reports from different Thomson Financial/CDA funds to the same fund in MFLINKS, and resolve such problems manually.

APPENDIX B

B. Determinants, persistence, and return predictability of the return gap

We study the determinants of the return gap by regressing the quarterly return gap on a number of fund characteristics. Each variable is defined below and the results are summarized in Table B1.

Table B1 The determinants of the return gap

	(1) ReturnGap$_t$	
	Coeff	*t*-stat
Intercept	−0.01*	−1.81
Trading costs$_t$	0.01**	1.99
Weight of recent IPOs$_t$	0.21***	10.51
ρ(holding return and net returns)$_{t-1}$	0.04	0.83
ExpRatio$_t$	−0.21**	−2.18
Yearly Turnover$_t$	0.00	−1.38
log(fund TNA)$_{t-1}$	−0.01	−0.95
log(family TNA)$_{t-1}$	0.04***	7.90
log(age)$_{t-1}$	−0.02*	−1.66
Flow$_t$	0.11**	2.21
σ(Return$_{t-5 \text{ to } t-1}$)	0.04***	4.60
R^2	0.04	
Observations	78,888	
Time period	Q1.1993–Q3.2010	

The dependent variable is return gap in quarter *t*. The independent variables are described in Appendix B. Observations with missing data are dropped. We estimate the model using a panel regression approach where we include time-fixed effects and cluster standard errors on the fund level. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

Trading costs$_t$: Trading costs in quarter t are calculated as the product of funds' turnover ratio with the relative rank of the average fund dollar holdings position within the same fund size tertile. The approach follows Edelen et al. (2013).

Weight of recent IPOs$_t$: IPOs stocks in quarter t are defined as stocks whose initial public offering (IPO) is conducted after the end of quarter $t - 1$ and before the end of quarter t. The weight of recent IPOs in quarter t is defined as the percentage of the total fund portfolio held in IPO stocks in quarter t, that is the weight of the total portfolio allocated to stock with an IPO in the last 3 months.

Correlation between holdings return and fund net return$_t$: At the end of quarter t, we construct a time series of the performance of the most recently disclosed fund holdings over the last 12 months and calculate its correlation with the fund net return.

Expense ratio$_t$: The most recently available fund expense ratio as of quarter t.

Yearly turnover$_t$: The most recently available fund turnover ratio as of quarter t.

Log(fund TNA)$_t$: The natural logarithm of fund total net assets, expressed in millions of dollars as of quarter t.

Log(family TNA)$_t$: The natural logarithm of the total net assets of the whole family, expressed in millions of dollars as of quarter t.

Log(age)$_t$: The natural logarithm of fund age, expressed in number of months since exception as of quarter t.

Standard deviation of fund returns between $t - 5$ and $t - 1$: The standard deviation of monthly fund return in the 12 months ending in quarter $t - 1$.

As expected, we find that transaction costs, estimated following Edelen et al. (2013), have a negative impact on the return gap. This implies that the interquarterly benefits from trading, on average, cannot offset the trading costs of mutual funds. The next determinant we investigate is the weight of recent IPOs. The previous literature has shown that mutual fund families tend to assign IPOs strategically across the fund family, allocating high weight of recent IPOs to certain funds in the family (Gaspar et al. 2006; Nanda et al. 2004; Nimalendran et al. 2007; Reuter 2006). Given that those IPOs tend to be significantly underpriced, it is not surprising that we find very strong positive relation between the weight of recent IPOs and the return gap.

Next, we look at the transparency of the funds' investment strategy, proxied by the correlation between the funds' reported holdings with the funds' net return. A low correlation might be due to agency problems, such as window dressing or high turnover. A low correlation might also be the result of realizing interquarterly informational advantages. The small positive (aleit statistically insignificant) coefficient on the correlation variable that we present in Table B1 suggests that the opaqueness of the trading strategy is negatively related to the return gap. The negative coefficient on the expense ratio suggests that fund expenses are not a compensation for the value-added of fund managers *via* their interquarterly actions. We further find no significant effect of fund turnover on

the return gap, which suggests that either the effect of fund turnover is captured in the trading cost proxy, or that the benefits of frequent trading just offset the costs of trading frequently.

Similarly to Chen et al. (2004), we find that performance is negatively related to size, but positively related to fund family size. The strong positive relationship to fund family size is probably due to economies of scale associated with transaction costs and lending fees (Chen et al. 2004). We further document a positive relationship between contemporaneous flows and the return gap which is consistent with the "smart money" effect documented by Gruber (1996) and Zheng (1999). Somewhat surprisingly, we find that funds with volatile returns have higher return gaps, though the effect is economically small.

To show that the information captured by return gap is highly persistent, at the end of each quarter we sort funds on their return gaps over the previous

Table B2 Persistence of the return gap

	1 Year		3 Year		5 Year	
	Mean	*t*-stat	Mean	*t*-stat	*t*-stat	*t*-stat
A: Equally weighted						
1 (lowest)	−0.48***	−5.96	−0.53***	−8.41	−0.55***	−6.83
2	−0.44***	−10.39	−0.44***	−8.57	−0.43***	−8.80
3	−0.37***	−8.89	−0.38***	−9.21	−0.40***	−8.70
4	−0.28***	−8.16	−0.29***	−6.26	−0.32***	−6.23
5	−0.26***	−6.90	−0.22***	−6.66	−0.23***	−5.78
6	−0.13**	−1.99	−0.16***	−4.34	−0.20***	−4.67
7	−0.13***	−3.58	−0.13***	−3.10	−0.18***	−3.18
8	−0.08**	−2.08	−0.07	−1.41	−0.12**	−2.00
9	−0.02	−0.42	−0.07	−0.81	−0.09	−1.42
10 (highest)	0.20**	2.37	0.14*	1.95	0.04	0.47
High–low	0.68***	7.91	0.67***	10.44	0.59***	6.86
B: Value-weighted						
1 (lowest)	−0.53***	−4.92	−0.63***	−4.24	−0.65***	−5.86
2	−0.39***	−5.02	−0.40***	−4.92	−0.46***	−4.23
3	−0.32***	−4.39	−0.27***	−4.02	−0.36***	−5.12
4	−0.20***	−3.46	−0.16**	−2.26	−0.29***	−3.98
5	−0.16***	−3.66	−0.22***	−3.33	−0.15**	−2.03
6	−0.10	−1.61	−0.11**	−2.38	−0.14**	−2.00
7	−0.06	−1.31	−0.11*	−1.92	−0.14**	−2.01
8	−0.07	−1.32	−0.07	−1.11	−0.03	−0.54
9	0.04	0.86	0.00	0.07	−0.06	−0.92
10 (highest)	0.08	0.95	0.08	1.01	0.01	0.13
High–Low	0.61***	5.39	0.70***	4.70	0.66***	5.63

At the end of each quarter we sort funds in 10 portfolios based on past 1, 3, or 5 year cumulative return gap. Next, we track the return gap of the portfolios over the next one quarter and repeat the procedure. This way we obtain a time-series of quarterly return gap scores for each portfolio. In panel A, we use equal weights to aggregate returns and in panel B we use fund net assets. We report portfolio means in percentages per quarter with *t*-statistics based on corresponding standard error of the mean. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

Table B3 Fund return predictability based on the return gap

		Yearly return gap					
		Q1	Q2	Q3	Q4	Q5	Q5–Q1
A: Two-way sorts on net return and the return gap							
Net return							
Q1	Alpha	−0.15	−0.19*	−0.02	−0.01	0.00	0.15*
	t-stat	−1.15	−1.90	−0.20	−0.10	0.00	1.67
Q2	Alpha	−0.06	−0.07	−0.08	−0.02	0.05	0.11
	t-stat	−0.86	−1.17	−1.33	−0.33	0.56	1.57
Q3	Alpha	−0.08	−0.07*	−0.06	−0.03	0.02	0.10*
	t-stat	−1.60	−1.75	−1.50	−0.75	0.33	1.67
Q4	Alpha	−0.14*	−0.06	−0.10**	−0.08	0.02	0.16***
	t-stat	−1.75	−1.00	−2.00	−1.60	0.33	2.67
Q5	Alpha	−0.14	−0.14	−0.05	−0.05	0.00	0.15*
	t-stat	−1.08	−1.17	−0.38	−0.42	0.00	1.88
B: Two-way sorts on alpha and the return gap							
Alpha							
Q1	Alpha	−0.41***	−0.22***	−0.18**	−0.17**	−0.10	0.31***
	t-stat	−3.15	−2.75	−2.57	−2.13	−0.91	2.82
Q2	Alpha	−0.19***	−0.13***	−0.10**	−0.05	−0.11	0.09
	t-stat	−3.17	−2.60	−2.00	−1.00	−1.57	1.50
Q3	Alpha	−0.12**	−0.07	−0.05	−0.03	0.00	0.11**
	t-stat	−2.00	−1.40	−1.00	−0.75	0.00	2.20
Q4	Alpha	−0.08	−0.06	0.00	−0.02	−0.01	0.06
	t-stat	−1.14	−1.20	0.00	−0.40	−0.14	1.00
Q5	Alpha	−0.01	0.10	0.13	0.19*	0.17	0.17**
	t-stat	−0.08	1.25	1.44	1.90	1.13	2.13

At the end of each quarter we sort funds in 5 quintiles based on past 1 year net return (panel A) or past alpha (panel B). Alpha is estimated using past 12 months of data and the excess return on the market, SMB, HML, and momentum as risk factors. Next, we sort each of the quintiles in 5 quintiles based on past 1 year return gap. We collect the returns of the portfolios over the next 3 months and repeat the procedure. This way we obtain a time-series of quarterly return gap scores for each portfolio. Next, we evaluate the performance of each time-series of portfolio returns using a four-factor asset pricing model, where we use the excess return on the market, SMB, HML, and momentum as risk factors. For each time-series of portfolio returns, we report the alpha and the corresponding *t*-statistic. *, **, and *** denotes 10%, 5%, and 1% levels of statistical significance, respectively.

1, 3, and 5 years. We then show that the past sorts predict return gaps in the following quarter. The results, summarized in Table B2, show that the spread between the average subsequent return gaps of funds with the highest return gap and those with the lowest return gaps remains economically substantial even after 5 years. Using both equal and value-weighting schemes, the return gaps of the spread portfolios range between 60 and 70 bp per quarter, depending on the time-frame used for sorting.[14] In other words, the information captured by the return gap tends to be highly persistent.

14 The methodology used and the results obtained are very similar to those of Kacperczyk et al. (2008).

We further show that the information contained in the return gap is not already contained in other performance indicators. At the end of each quarter we double-sort funds on past returns or alpha and the return gap. Next, we collect the returns of the portfolios and rebalance. This way we obtain a time-series of double-sorted portfolios and examine their performance. Results are summarized in Table B3. In panel A (B), we first sort on the fund net return (alpha) in the previous year, and then on the return gap in the previous year. Going from left to right in both panels, there is an increasing pattern in abnormal post-ranking performance. The spreads between the top and bottom quintile of funds sorted on return gap ranges between 6 and 31 bp per month. This indicates that the return gap contains predictive power about fund performance which complements the information in past net returns and alpha, rather than substituting it. The return gap is particularly informative about the future performance of the funds with the best/worst past performance, where the spread between funds with the highest and lowest realizations of the return gap ranges between 15 and 31 bp per month and is significantly different from zero at conventional levels.