JOIM
www.joim.com

Journal Of Investment Management, Vol. 17, No. 3, (2019), pp. 4–20

DOES TRADING BY ETF AND MUTUAL FUND INVESTORS HURT PERFORMANCE? EVIDENCE FROM TIME- AND DOLLAR-WEIGHTED RETURNS

Ananth Madhavan and Aleksander Sobczyk**

This paper analyzes the "return gap" between internal rate of returns that account for intermediate investor flows ("dollar-weighted returns") and more familiar buy-and-hold returns that funds typically must report. Our sample constitutes all US-domiciled open-end mutual funds and exchange-traded funds (ETFs), and covers both fixed income and equity funds, as well as active and index styles of management. We find that return chasing behavior explains the cross-sectional pattern of the return gap. We conclude that high turnover of liquid ETFs does not lead to sub-par returns for investors in these funds.

1 Introduction

Buy-and-hold returns are the performance measure that investment funds are required to report. These returns also constitute historical performance data. Buy-and-hold returns reflect the performance of an investor who makes no subsequent contributions or withdrawals beyond their initial investment. In reality, investors usually contribute to or withdraw from funds over time potentially leading to a "return gap" between the dollar-weighted return that accounts for intermediate cash flows (defined as the internal rate of return on the investment) and the reported buy-and-hold return. Understanding the magnitudes and determinants of the return gap is important in an era where many studies report that retirees have insufficient savings and return expectations are low.

A growing number of studies have examined the return gaps for investors in mutual funds and individual stocks. We contribute to this literature by analyzing the returns of all US-domiciled open-end *active* and *index* equity and fixed income funds, over 6,700 funds with $13.1 trillion in assets.[1] Of particular interest, our universe comprises not only mutual funds, but also all US-domiciled exchange-traded funds (ETFs). Our study is the first to compare and contrast return gaps across ETFs and mutual funds, vehicles that

*The views expressed here are those of the authors alone and not of BlackRock, Inc.
BlackRock, Inc. 400 Howard Street, San Francisco CA 94105, USA.

are of considerable interest, given differences in turnover and clientele.

Assets in exchange-traded products now exceed $6 trillion and continue to grow rapidly.[2] Indeed, the very benefits of ETFs—broad exposures, intraday tradability on organized exchanges, transparency, and low cost—have raised concerns that their liquidity and ease of use encourages excessive turnover by retail investors, possibly to the detriment of investment performance.[3] Further, ETF transactions incur commissions and require an investor to have a brokerage account. Consequently, mutual funds are often chosen for automated investing or offered in 401(k) plans, suggesting the possibility of major differences between ETFs and mutual funds.

There is another feature of ETFs that is very important for our research. Unlike mutual funds, ETFs flows (i.e., changes in shares outstanding arising from creation or redemption activity) are available daily. Higher-frequency flow data allows us to directly relate observed return gaps over horizons such as weeks or months to the time-series of investor fund sensitivity to past returns, while controlling for fund level attributes such as size, volatility, and fees. Explaining the behavioral drivers of differences between realized and reported returns across funds is important for advisors and policymakers seeking to improve investment outcomes for investors.

Our analysis of US-domiciled ETFs and mutual funds provides several new findings:

1. Empirically, the return gap for larger, less volatile equity, and fixed income ETFs is generally economically small and similar in economic magnitude to those of open-end mutual funds. So, the ability to trade ETFs intraday does not necessarily lead to sub-par returns relative to traditional investment vehicles that offer only end-of-day liquidity.
2. Methodology matters: Some widely-cited industry reports conclude that investor return gaps are economically significant and negative, as much as −3.52% annually for the S&P 500 over the past 20 years,[4] numbers much larger than what we find using an internal rate of return approach. When we use a common industry approach that does not explicitly adjust for the time value of money we do indeed find much larger gaps, indicating that methodology matters.
3. Return gaps are directly related to investor trend-chasing, as measured by the time-series relation between flow and past returns, confirming the importance of behavioral drivers of realized investment performance. Across funds and time, we find no evidence that gaps are related to broad market returns.

This paper proceeds as follows: In Section 2, we provide an overview of the literature on investor returns and offer a conceptual framework to understand return gaps. Section 3 describes our data and methodology. We then turn in Section 4 to an empirical analysis of return gaps, cut by asset class and investment vehicle. Section 5 focuses on ETFs, relating the gap across funds to time-series estimates of flow sensitivity and to analyze the impact of trend-chasing on performance while controlling for other factors. Finally, we conclude with a summary of our results and the implications for practitioners.

2 Realized investor returns

2.1 Previous research

Here, we briefly survey the relevant literature and then present a framework to understand the determinants of the "return gap." As we show below, the return gap can be positive or negative, depending on the timing and magnitude of intermediate cash flows over the investment horizon. If cash flows are stochastic, the internal rate of return

and buy-and-hold returns are equal on average, and expected return gap is zero. This is so even if the mean contribution is non-zero, as long as we correctly account for the time value of money. But if intermediate cash flows depend on past returns, we expect gaps. In particular, we show that if investors act as trend-chasers, with flows positively related to past returns, the expected return gap will be negative. By contrast, contrarian behavior (such as rebalancing activity) will lead to positive expected gaps.

There is a vast literature on investor behavior. First, there is a literature beginning with Constantinides (1979) that deals with dollar cost averaging. This behavior is distinct from contributions to retirement accounts or through automated investing plans. If correctly computed, the return gap for such steady flows should, on average, be zero. However, the great majority of fund flows reflect changes in investor sentiment. Indeed, there is a large academic literature on investor cognitive biases and its possible impact on investment flows.[5] In general, the previous research on behavioral finance tends to support the hypothesis that investors will allocate more to funds with positive returns, but there are also numerous exceptions including investors who systematically rebalance or those who exhibit contrarian ("buy on the dip") behavior.

In terms of prior empirical research, several industry groups publish widely distributed reports about how realized returns may differ from the buy-and-hold index returns aimed at financial advisors and investors. Kinnel (2017), using global Morningstar investor returns across markets and asset classes, finds that the five-year investor return gap ranged widely from −1.40% to −0.53% for the year ended 2016. DALBAR publishes an annual report "Quantitative Analysis of Investor Behavior" (QAIB) that measures the effects of investor decisions to trade mutual funds. Their results—which date to 1994—consistently suggest that the average investor earns significantly less than reported mutual fund performance. For the period ending 12/31/2015, QAIB reports the 20-year annualized return of the average equity mutual fund investor was 4.67% versus 8.19% for the S&P 500, a gap of −3.52%. This finding is not limited to equity investors. QAIB reports that the corresponding return average for fixed income mutual fund investors over the 20-year period ending 12/31/2015 was 0.51% versus 5.34% for the broad Bloomberg Barclays Aggregate Bond Index.[6] The relatively large return gaps reported in industry studies may reflect methodology, as we discuss below.[7]

Academic studies (see, e.g., Fischer and Wermers, 2012) have largely focused on mutual fund or individual stock investors. Dichev (2007) focuses on individual stocks and reports average return gaps between dollar- and buy-and-hold returns of around −1.3% to −5.3% per annum over a long horizon. Bessembinder (2017) finds that the entire gain in the US stock market since 1926 is attributable to the best-performing 4% of stocks, so that most investors in individual common stocks would have underperformed broad indexes. Friesen and Sapp (2007) find that investor timing decisions in equity mutual fund reduce fund investor average returns by 1.56% annually. Hsu *et al.* (2016) examine the difference between buy-and-hold returns and the average dollar-weighted returns by mutual fund investors from January 1991 to June 2013, focusing on value funds. They find that value fund investors underperformed the benchmark by 0.92% in terms of realized returns. The average growth fund investor, however, experienced a performance gap of −3.16% a year compared to the fund's return. Dichev and Yu (2011) examine dollar-weighted returns of hedge funds and find that the alpha of investors in these funds is close to zero.

2.2 *Conceptual framework*

Why and under what conditions do return gaps exist? Consider the simplest case with two intervals, where intermediate cash flows occur after the first period's returns are observed. The argument is easily generalized to longer intervals. We use continuously compounded returns for simplicity because they can be added over intervals and assume that returns are independently distributed with mean μ. In the first period the time- and asset-weighted returns are the same by definition, denoted by r_1. In period 2, conditional upon the first period return outcome, the expected return is $E[r_2|r_1] = \mu$. The expected average two-period buy-and-hold return at the beginning of period 2, conditional on the return in period 1 is defined as:

$$E[r^{BH}|r_1] = (r_1 + \mu)/2. \tag{1}$$

Note that Equation (1) holds irrespective of any intermediate cash flows after the first period, and is also independent of the dispersion in returns. The internal rate of return or dollar-weighted return—denoted r^{DW}—can be written (see, e.g., Fischer and Wermers, 2012) as a weighted average of the sub-interval returns, and we write:

$$r^{DW} = \sum_{i=1}^{2} w_i r_i, \tag{2}$$

where w_i is the weight in interval i based on that period's asset base, with the sum of weights equal to 1. We consider three possibilities for intermediate period cash flows.

Case 1: Trend-chasing behavior
In the first case, we assume that in period 2, trend-chasing investors adjust their assets up or down in the direction of returns in excess of μ. The relation between flows and returns need not be symmetric.[8] So, conditional upon the first period's return r_1, we have:

- If $r_1 > \mu$, cash flows are positive, and $w_2 > \frac{1}{2} > w_1$.
- If $r_1 < \mu$, cash flows are negative, and $w_2 < \frac{1}{2} < w_1$.

The asset-weighted return will place more weight on period 2 than period 1 if $r_1 > \mu$ because the asset base has risen. Then, the expected return gap is:

$$E[r^{DW} - r^{BH}|r_1] = E[w_2|r_1] - \frac{1}{2} < 0.$$

If flows are very sensitive to past returns, the intermediate cash flows are large relative to assets in absolute value. The same is true for volatile funds or assets. Greater return dispersion for a given sensitivity of flows to returns means greater (absolute) cash flows, and hence a more negative return gap.

Case 2: Contrarian behavior/portfolio rebalancing
When investors are contrarians or rebalance to a target, we have the exact opposite of the first case where, if $r_1 < \mu$, assets go up as cash flows are positive leading to a positive gap.

Case 3: Stochastic cash flows
If there are no intermediate cash flows or if expected intermediate cash flows are zero, then it follows that *ex ante* we expect $E[w_i] = 1/2$ and again, $E[r^{DW} - r^{BH}|r_1] = 0$.

Implications:

1. In the absence of intra-period cash flows, time- and dollar-weighted returns are equal. If there are cash inflows, but these are unrelated to returns, the expected gap is zero.
2. Trend chasing by fund investors can explain why dollar-weighted returns are lower than buy-and-hold returns. Conversely, if investors are contrarians (or rebalance to a target) the opposite result holds. Importantly, this result does not require security prices to exhibit mean

reversion, as our model assumes independent returns.

3. The expected return gap is related negatively to the covariance between flows and past returns. We can write the covariance as $\sigma^2(r)\gamma$, where $\sigma^2(r)$ is the variance of returns and γ is the regression coefficient of flow on past returns, a measure of flow sensitivity. We can independently test this idea by estimating γ through time-series regressions of flow on returns and relating this estimate in the cross-section to estimated return gaps.
4. Greater return variability $\sigma^2(r)$ thus means a larger return gap for any given flow sensitivity. So, we predict that narrower, more focused funds will experience larger return gaps than funds with a broad set of constituents. The initial asset base of the fund does not affect buy-and-hold returns. However, as the initial asset base of the fund relative to the size of the cash flows affects the flow sensitivity parameter γ, we expect greater return gaps for smaller funds.

3 Data and methodology

In this section, we operationalize the simplified model framework above. We then discuss our data sources and procedures.

3.1 Definitions

The industry-standard for reporting fund performance is to use buy-and-hold (i.e., compounded) returns. As noted earlier, this convention captures the total return of an investor who remains fully invested, but does not contribute or withdraw funds over the interval of analysis. In empirical work we will use geometric (as opposed to continuous) returns[9]

$$(1 + r^{BH})^T = \prod_{t=1}^{T}(1 + r_t)^t \tag{3}$$

where r_t is the fund's return for period t.

The dollar-weighted return (also known as "internal rate of return" or IRR) captures the effect of the timing and magnitude of intermediate capital flows. The IRR is more reflective of the returns realized by fund investors in the aggregate than buy-and-hold returns. IRR equates the present value of all cash flows to the present value of the final period assets under management (AUM) of the fund:

$$AUM_0 + \sum_{t=1}^{T} \frac{Flow_t}{(1 + r^{DW})^t} = \frac{AUM_T}{(1 + r^{DW})^T} \tag{4}$$

where:

$$AUM_t = (1 + r_t) \cdot AUM_{t-1} + Flow_t. \tag{5}$$

From Equations (3) and (4), we see that in the absence of intermediate cash flows, both returns are identical. From a mathematical point of view, this definition of dollar-weighted return is unique and well defined, as AUM is always positive.[10]

3.2 An illustrative example

In the hypothetical example of Exhibit 1, consider three Investors and two periods, with only one chance for intermediate cash flows at the very end of period 1. Investor A is a buy-and-hold investor who invests \$1,000 in a fund and experiences returns over two consecutive periods of 10% and 10%, respectively. The investor's overall gross return is thus $(1 + 0.10) \times (1 - 0.10) - 1 = -0.01$, or -0.50% annualized over the two-periods for a final period investment value of \$990. From Equation (4), the internal rate of return is the solution r^{DW} to the present value equation analogue $\$1{,}000 = \frac{\$990}{(1+r^{DW})^2}$ which is a -0.50%, also the buy-and-hold return. Note that in a more realistic example with more than just one window for possible intermediate cash flows, an investor in an automated program that

Exhibit 1 Hypothetical illustration of dollar-weighted returns and buy-and-hold returns.

Initial assets AUM_0	First period return r_1	Cash flows $Flow_1$	Second period return r_2	Final assets AUM_2	Annual buy-and-hold return r^{BH}	Annual dollar-weighted return r^{DW}	Return gap $r^{DW} - r^{BH}$
Investor A							
\$1,000	10%	0	−10%	\$990	−0.50%	−0.50%	0.00%
Investor B							
\$500	10%	\$500	−10%	\$945	−0.50%	−3.71%	−3.21%
Investor C							
\$500	10%	−\$200	−10%	\$315	−0.50%	1.85%	2.35%
Fund (A+B+C)							
\$2,000	10%	\$300	−10%	\$2,250	−0.50%	−1.17%	−0.67%

The table illustrates the computations of buy-and-hold returns and dollar-weighted (internal rate of return) for three hypothetical investors and fund with all three investors.

mechanically invests a constant amount would also experience a return gap of zero too.

Investor B, a trend-chaser, experiences a lower dollar-weighted return over the two periods than the buy-and-hold return, because his cash inflows coincide with future lower returns in the second period. The investor's initial investment is \$500, which is worth \$550 in the first period. Following the addition of \$500 to the account ($Flow_1$) the investor experiences a −10% second period return, so the final investment is worth $\$945 = (\$550 + \$500) \times (1 - 10\%)$. It can be verified that $r^{DW} = -3.71\%$ is the solution to the IRR Equation (4) that equates the present value of all cash flows with the present value of the assets:

$$\$500 + \frac{\$500}{(1 + r^{DW})} = \frac{\$945}{(1 + r^{DW})^2}. \quad (6)$$

Investor C is a contrarian so that faced with a first period positive return, she withdraws \$200. Her final investment is worth $\$315 = (\$550 - \$200) \times (1 - 10\%)$. She experiences a higher dollar-weighted return than the buy-and-hold return because her cash outflows happen to coincide with lower returns in the second period.

What about a fund that comprises the three investors? The last line of the table shows that an individual investor's return may be different from the fund's dollar-weighted return. The fund's dollar-weighted return in this illustrative example is −1.17%. Note that actual account level data (history of purchases and sales) is needed to calculate individual investor returns, but the fund return provides an overall view of the collective (or average) experience of the three investors.

3.3 An aside on alternative approximations

In industry, the Dietz Method[11] is commonly used to approximate the realized returns of investors based on the total gain relative to the cost basis:

$$DM = \frac{AUM_T - AUM_0 - \sum_{t=1}^{T} Flow_t}{AUM_0 + \sum_{t=1}^{T} Flow_t}. \quad (7)$$

Here *Flow* includes net contributions plus dividends and interest, less fees, and costs. The approximation of Equation (7) will tend to understate the returns of investors who invest steadily over time in rising markets because these investors will—on average—have less terminal wealth over long horizons than investors who make a single, lump-sum investment at the start

of the period. But this return gap is not the result of poor timing but rather reflects simply the time value of money. Even in rising markets, the expected return gap, using the IRR method of Equation (4) is zero. Methodology and data may help explain why some of the return gaps reported in industry studies can differ so widely, especially in rising markets.[12]

Data sources

We gathered daily data from January 2010 to December 2016 on all US-domiciled ETF returns, assets, and flows from a variety of primary sources including Bloomberg, IDC, and Thomson Reuters. To facilitate comparisons with mutual funds, we focus on physically backed[13] equity and fixed income ETFs, and exclude other types of exchange-traded products such as commodity funds, inverse and leveraged funds, and exchange-traded notes (ETNs), that use leverage, swaps, or offer synthetic exposures. The sample universe is large, comprising as of 12/31/2016 some 1,077 equity ETFs and 257 fixed income ETFs, with assets of approximately $2.0 trillion and $0.42 trillion, respectively.

We complement the ETF data with monthly return, flow, and asset data from Morningstar on all US-domiciled open-end mutual funds, both active and index, for the same period, January 2010–December 2016. The sample comprises, as of 12/31/2016, of 4,044 equity mutual funds with

Exhibit 2 Sample summary statistics by year, fund type, and asset class.

	Equity		Fixed income	
Year	Funds	AUM ($MM)	Funds	AUM ($MM)
Panel A: Exchange-traded funds				
2010	497	$731,113	81	$127,793
2011	596	$727,165	113	$176,744
2012	742	$951,703	147	$231,218
2013	772	$1,344,734	179	$234,828
2014	820	$1,599,605	212	$289,890
2015	913	$1,691,767	240	$333,800
2016	1,077	$1,992,965	257	$421,167
Panel B: Open-end mutual funds				
2010	3,003	$4,717,361	915	$1,841,062
2011	3,151	$4,456,203	977	$2,065,657
2012	3,363	$5,097,943	1,076	$2,520,065
2013	3,564	$6,789,960	1,164	$2,499,531
2014	3,773	$7,533,061	1,263	$2,669,482
2015	3,972	$7,425,701	1,369	$2,595,402
2016	4,020	$7,873,147	1,412	$2,841,421

The table provides summary statistics of on the number and assets under management of the funds in our sample by year, fund type and asset class. Panels A (equity and fixed income ETFs) and B (equity and fixed income mutual funds) cover the period January 2010–December 2016. The number of funds and assets under management (AUM) reported at the end of each calendar year. Data are drawn from Morningstar, Bloomberg, BlackRock, IDC, and Thomson Reuters.

a total AUM of $7.9 trillion and data on 1,415 fixed income mutual funds with approximately $2.8 trillion in assets.

Exhibit 2 presents summary statistics across US-listed ETFs (Panel A) and mutual funds (Panel B), by year and by asset class, from January 2010 to December 2016. The number of funds and assets under management (AUM) reported is the total at the end of each calendar year. The number of ETFs increases over time, as shown in Exhibit 2, Panel A. The growth in fixed income ETFs, both in number and assets, over this period is especially notable.

We estimate the return gap (defined as dollar-weighted return minus buy-and-hold return) for each ETF and for each calendar year, beginning 1/1/2010 through 12/31/2016 using monthly flows and returns (aggregated from daily data). ETF flows are primary market flows, defined as the product of the change in shares outstanding and the fund's net asset value. Note that an ETF does not need to trade for the entire period to be included in the sample. Specifically, we include ETFs that were launched or delisted during this period. For each ETF, we estimate buy-and-hold returns each calendar year, using official daily close prices in our sample, adjusted for corporate actions and dividends, sourced from IDC and Thomson Reuters. The daily flow/return data for ETFs was aggregated into monthly periods, to make them directly comparable with mutual fund flow data (which is monthly). We similarly estimate dollar-weighted returns each calendar year to check that our results are not specific to a given time period.

4 Empirical results

We first examine the results for ETFs across asset class and by time interval, and then compare and contrast these results with active and index mutual funds.

4.1 Results for equity and fixed income ETFs

In each calendar year, we ranked ETFs and mutual funds into quintiles based on their return volatility in that calendar year, with quintile 1 being the lowest and quintile 5 the highest.[14] In Exhibit 3 Panel A (equity ETFs) and Panel B (fixed income ETFs) show the estimated mean and median return gaps by quintiles of return volatility averaged across the panel of all funds in the period 2010–2016, the number of funds, and assets as of 12/31/2016. Note that the number of observations may differ in other years, as funds are incepted or delisted.

Beginning with Panel A of Exhibit 3, in each volatility quintile there are approximately 215 equity ETFs as of 12/31/2016. Over the entire sample period we have 5,396 ETF-yearly observations. By quintiles of fund return volatility, we see that the mean return gap is negative in all cases. This is not to say that some funds do not have positive gaps, but rather that the gaps are negative on average. The median is smaller than the mean in every quintile, illustrating the skew in the measure. As we hypothesized, the average return gap decreases monotonically with volatility, which is consistent with the hypothesis that trend-chasing is more prevalent in more volatile funds. Across all equity ETFs in Exhibit 3, Panel A, the (unweighted) mean return gap is -0.93% and the median is -0.36%.

We also computed, but do not report, the gap using Equation (7) as an approximation for the internal rate of return. Overall, the approach of looking at total gains suggests much larger return gaps for equity ETFs, a mean of -2.28% and a median gap of -1.26% across all funds. These figures are consistent with our earlier comments about the methodology affecting the estimated return gap.

The corresponding statistics for fixed income ETFs are shown in Exhibit 3, Panel B. There

Exhibit 3 Return gap by volatility quintile for equity and fixed income ETFs.

Fund volatility quintile	Mean fund volatility	# of Funds as of 12/31/2016	AUM ($MM) as of 12/31/2016	Median return gap	Mean return gap	Std. Dev. return gap
Panel A: Equity ETFs						
1	9.2%	216	$851,214	−0.12%	−0.40%	1.86%
2	12.5%	215	$459,824	−0.25%	−0.59%	2.48%
3	15.2%	215	$232,933	−0.49%	−0.82%	4.24%
4	19.3%	215	$332,120	−0.61%	−1.17%	4.91%
5	28.6%	216	$116,874	−0.58%	−1.68%	7.98%
All	**17.0%**	**1,077**	**$1,992,965**	**−0.36%**	**−0.93%**	**4.82%**
Panel B: Fixed income ETFs						
1	0.8%	52	$62,478	−0.02%	−0.08%	0.24%
2	2.4%	51	$86,936	−0.10%	−0.24%	0.57%
3	4.1%	51	$131,193	−0.42%	−0.63%	0.85%
4	6.1%	51	$107,334	−0.32%	−0.67%	1.32%
5	11.2%	52	$33,226	−0.50%	−0.98%	3.63%
All	**4.9%**	**257**	**$421,167**	**−0.15%**	**−0.52%**	**1.82%**

The table shows summary statistics on return gaps based on quintiles of return volatility. Panel A represents equity ETFs and Panel B shows fixed income ETFs. Data is sourced from Bloomberg, BlackRock, IDC, and Thomson Reuters. Return gap is estimated annually for each fund, for period January 2010–December 2016 using monthly flow and return data.

are 257 fixed income ETFs in our sample with a total AUM of $421 billion as of 12/31/2016. Again, the same patterns noted above for equity ETFs are evident: Overall, the mean return gap is −0.52%, smaller than the equity average, and the median is only −0.15%. The median return gap is smaller than the mean in every quintile. The return gap is negative in all quintiles, with the smallest quintiles in fund risk exhibiting the smallest return gaps. Indeed, in the lowest risk quintile, the mean return gap is only −0.08%. Given the large cross-sectional dispersion (1.82%), the average return gap is not significantly different from zero. This is consistent with our intuition: Fixed income funds exhibit much less return volatility and hence we expect smaller return gaps. How do fixed income funds do using the Dietz Method? Consistent with the results for equity funds, the average gap using the Dietz Method widens significantly from −0.52% to −1.55%, again showing that methodology matters.

Since ETF flows are available on a daily basis, unlike mutual funds where we typically observe only monthly or perhaps quarterly flows, we can examine the return gaps over different intervals such as days, weeks, months, and quarters. Such an analysis can help shed light on the intervals over which trend-chasing induces a return gap or difference. In Exhibit 4, we estimate the return gap using daily, weekly and monthly horizons for each calendar year from 1/1/2010 through 12/31/2016.

The mean return gap is most negative for both equity and fixed income ETFs at the daily measurement level. As before, the return gaps for fixed income ETFs are smaller in absolute magnitude than in equity ETFs, and there is less variation over the observation interval for bond

Exhibit 4 Return gaps for ETFs for daily, weekly and monthly estimation intervals.

	Panel A: Equity ETFs		Panel B: Fixed income ETFs	
Interval	Median return gap	Mean return gap	Median return gap	Mean return gap
Daily	−0.66%	−1.43%	−0.25%	−0.60%
Weekly	−0.51%	−1.24%	−0.18%	−0.55%
Monthly	−0.36%	−0.93%	−0.15%	−0.52%

The table provides estimates of the median and mean return gap for equity ETFs (Panel A) and fixed income ETFs (Panel B) for three intervals of estimation. Data are sourced from Bloomberg, BlackRock, IDC, and Thomson Reuters. The return gap is estimated annually for each fund, for period January 2010–December 2016 using daily, weekly or monthly flow and return data.

funds. The conclusion then is that trend-chasing is most evident at the daily level, especially for equity funds, but is less of a factor for fixed-income funds (−0.60% at the daily level versus −0.52% at the monthly level) than for equity funds (−1.43% at the daily level versus −0.93% at the monthly level), as we expected.

Is there a relation between market direction and the return gap, given the significant market increase over the sample period? To drill into this issue, we computed the correlation of the gap by fund-year with the S&P 500 returns over monthly and quarterly (using daily flow and return data to estimate IRRs) non-overlapping intervals. At the monthly level, the correlation between S&P 500 returns and return gap is negligible, only −2.1%. This correlation increases, as the estimation interval increases for quarterly epochs the correlation is around 12.5%. The higher quarterly figures is consistent with our conceptual model where fund volatility—which increases with the length of estimation interval—explains the most cross-sectional variation in the return gap. Also worth noting, our study covers both fixed income and equity funds, and we see the same aggregate patterns over both asset classes, despite a very different environment for the two asset classes in the sample period. As shown in the model, the gap reflects the timing of intermediate cash flows and not security returns per se, even with positive flows over a period when returns are positive. Further, the statistics we report are averages across many funds. Within any given year, we see wide dispersion in the gap with some funds positive and others negative.

4.2 Results for equity and fixed-income mutual funds

In Exhibit 5, Panels A and B report estimated return gap by quintiles of fund return volatility for equity and fixed income active mutual funds, respectively. These funds offer an interesting contrast to ETFs, the vast majority of which (see Madhavan, 2016) seek to track market capitalization-weighted indexes. There is a small but growing fraction of ETFs pursuing non-market capitalization strategies that may be termed "active" or "Smart Beta" (these are typically rules-based) but vast majority of AUM is in products that seek to track an index, versus active funds that try to beat the benchmark.[15]

The mean return gaps for active equity mutual funds are −0.15% versus −0.09% for active fixed-income mutual funds. The differences between index and active equity and fixed-income funds are minimal. ETFs and mutual funds can have different clienteles (ETFs are increasingly used by institutions and hedge funds[16] while mutual funds, both index and active, are largely

Exhibit 5 Return gaps by volatility quintile for active and index equity and fixed income mutual funds.

Fund volatility quintile	Mean fund volatility	# of Funds as of 12/31/2016	AUM ($MM) as of 12/31/2016	Median return gap	Mean return gap	Std. Dev. return gap
Panel A: Active equity mutual funds						
1	8.9%	735	$1,618,503	−0.02%	−0.11%	0.97%
2	11.4%	735	$1,330,350	−0.03%	−0.10%	1.13%
3	13.6%	734	$1,157,346	−0.02%	−0.15%	1.79%
4	16.7%	735	$964,392	−0.04%	−0.15%	1.91%
5	22.7%	735	$709,944	−0.15%	−0.24%	2.84%
All	**14.7%**	**3,674**	**$5,780,535**	**−0.04%**	**−0.15%**	**1.85%**
Panel B: Index equity mutual funds						
1	8.9%	70	$705,673	−0.04%	−0.17%	0.97%
2	11.0%	69	$638,712	−0.03%	−0.09%	1.56%
3	13.0%	69	$389,603	−0.04%	−0.09%	2.03%
4	15.9%	69	$209,194	−0.04%	−0.22%	2.10%
5	21.6%	69	$149,430	−0.07%	−0.05%	3.04%
All	**14.1%**	**346**	**$2,092,612**	**−0.04%**	**−0.12%**	**2.06%**
Panel C: Active fixed income mutual funds						
1	0.9%	272	$427,185	0.00%	−0.02%	0.12%
2	2.2%	272	$474,434	−0.02%	−0.06%	0.26%
3	3.2%	271	$676,800	−0.02%	−0.09%	0.54%
4	4.5%	272	$469,073	−0.05%	−0.08%	0.60%
5	8.0%	272	$323,653	−0.07%	−0.19%	1.34%
All	**3.7%**	**1,359**	**$2,371,144**	**−0.02%**	**−0.09%**	**0.71%**
Panel D: Index fixed income mutual funds						
1	1.5%	11	$57,501	−0.01%	−0.06%	0.35%
2	2.4%	10	$76,135	0.00%	0.02%	0.18%
3	3.1%	11	$161,857	−0.02%	−0.19%	1.16%
4	4.0%	10	$163,619	−0.20%	−0.36%	0.67%
5	10.4%	11	$11,164	−0.12%	−0.33%	1.64%
All	**4.3%**	**53**	**$470,277**	**−0.03%**	**−0.18%**	**0.97%**

The table shows statistic on the return gap for active equity mutual funds (Panel A), Index equity mutual funds (Panel B) and correspondingly for fixed income funds (Panels C and D) based on quintiles of fund return volatility. Data is from BlackRock, Morningstar, and Thomson Reuters. The return gap is estimated annually for each fund, for period January 2010–December 2016 using monthly flow and return data.

held by retail investors), implying differences in the pattern of cash flows over investment horizons. So, it is not surprising that these return gaps are smaller, and in fact are insignificantly different from zero. The same patterns noted above for ETFs are present for mutual funds: the return gap is smaller for fixed-income funds than equity funds. Further, the mean return gap is generally smaller for less volatile mutual funds. Although the medians are smaller than the means, as before, the skew is less noticeable. We computed, but do not report, the Dietz Method for both equity and

fixed-income funds. These differences are not as sharp as with ETFs, presumably because turnover in mutual funds is lower.

5 Cross-sectional analysis

5.1 Multivariate regression models

In this section, we analyze multivariate models of the return gap across funds. Given the relatively small number of fixed-income funds, we focus here on equity ETFs and mutual funds. In Exhibit 6, we report a pooled panel (fund × year) OLS regression where the dependent variable is the return gap, expressed in percent, estimated separately for ETFs (Panel A) and mutual funds (Panel B). The independent variables include proxies for funds likely to experience trend-chasing or have a retail focus (see Hsu *et al.*, 2016) based on factors such as size, fees, volatility, concentration or breadth, and popular retail styles:

- Beginning period assets (log);
- Fund volatility (annualized);
- Fund total expense ratio (TER)
- Dummy variable for sector funds;
- Dummy variable for Large-Cap Growth funds;
- Dummy variable for Small-Cap Growth funds;
- Dummy variable for Large-Cap Value funds;
- Dummy variable for Small-Cap Value funds;
- For ETFs: Dummy variable for Non-market Cap-Weighted funds (examples include "Smart Beta" or "Active" ETFs); and for Mutual Funds: Dummy variable for Active funds.

For both ETFs and mutual funds, higher volatility, after controlling for other factors such as fund size, is strongly negatively significant. This multivariate result is consistent with Exhibits 2 and 3. Fund size matters only for mutual funds and enters with a positive coefficient. That is consistent with investors in larger equity mutual funds

Exhibit 6 Cross-sectional determinants of return gap.

Panel A: equity ETFs			Panel B: Equity mutual funds		
	Beta	*t*-Stat		Beta	*t*-Stat
Constant	0.583	2.50	Constant	−0.074	−1.44
Log AUM ($MM)	−0.025	−1.01	Log AUM ($MM)	0.029	6.21
Fund volatility	−4.468	−6.66	Fund Volatility	−1.417	−7.90
Total expense ratio	−0.008	−3.49	Total Expense Ratio	−0.036	−1.58
Sector Fund Dummy	−0.674	−5.83	Sector Fund Dummy	0.214	0.46
Large-Cap Growth fund dummy	0.529	1.48	Large-Cap Growth Fund Dummy	0.047	1.70
Small-Cap Growth Fund Dummy	−1.278	−2.83	Small-Cap Growth Fund Dummy	0.069	1.79
Large-Cap value fund dummy	0.316	0.88	Large-Cap Value Fund Dummy	0.090	2.78
Small-Cap value fund dummy	0.228	0.48	Small-Cap Value Fund Dummy	0.047	0.90
Non-market Cap Fund Dummy	−0.071	−0.62	Active Fund Dummy	−0.019	−0.56
Adj. R^2	2.22%		Adj. R^2	0.56%	
Observations	5,396		Observations	24,772	

The tables represent a pooled panel (fund × year) OLS regression where the dependent variable is the return gap, expressed in percent for equity ETFs (Panel A) and equity mutual funds (Panel B). Source: Bloomberg, BlackRock, IDC, Morningstar, Thomson Reuters; Results for pooled panel regressions are based on annual estimates for period January 2010–December 2016 using monthly return and flow data.

not chasing past returns or perhaps even acting as contrarians. Of particular interest, for equity ETFs, we find that sector funds are associated with more negative return gap, again indicative of traders using sector ETFs to attempt timing. We do not find an equivalent result for mutual funds, likely a clientele effect. Higher expense ratio funds also have more negative return gaps. These results are consistent with those of Hsu *et al.* (2016) who provide a multivariate analysis of individual mutual funds.[17] Their results show that the return gap is larger for investors in growth-oriented mutual funds that are high-fee, non-index, and retail-oriented, all characteristics they attribute to lack of sophistication.

5.2 *Evidence of timing*

Our hypothesis is that negative return gaps observed across the cross-section of ETFs and mutual funds reflect investor buying/selling decision based on past performance. As shown above, we would observe a negative return gap even if returns did not exhibit any time-dependency as long as investor flows were positively related to past returns.

To test this hypothesis empirically, we estimated the flow-return coefficient γ for each ETF, using time-series OLS in the form:

$$\frac{Flow_t}{AUM_0} = c + \gamma \cdot r_{t-1} + \epsilon_t. \tag{8}$$

We restrict attention to ETFs because there are insufficient time-series observations for mutual funds. We require an ETF to have full trading history within the calendar year to estimate γ parameter, using weekly flow and return data, but do not require that the ETF is present in the entire sample period. We estimate the regression

Exhibit 7 Return gap and flow sensitivity parameter for ETFs.

Return gap quintile	Mean return gap	Median return gap	Mean γ	t-Stat $\gamma \neq 0$
Panel A: Equity ETFs				
1	−6.55%	−4.61%	1.11	2.33
2	−1.62%	−1.54%	0.17	5.81
3	−0.50%	−0.50%	0.07	4.95
4	0.07%	0.00%	0.05	3.12
5	2.51%	1.42%	0.02	0.12
Panel B: Fixed income ETFs				
1	−2.59%	−1.86%	1.03	3.71
2	−0.64%	−0.61%	1.87	1.85
3	−0.19%	−0.18%	5.83	1.40
4	−0.02%	−0.01%	4.58	2.44
5	0.72%	0.18%	−2.20	−0.40

The table is based on data sourced from Bloomberg, BlackRock, IDC, and Thomson Reuters. Panels A and B show the mean and median for quintiles of return gap for equity and fixed income ETFs, respectively. We also report the mean flow sensitivity parameter γ which is estimated annually for each fund individually from a time-series OLS regression based on Equation (8), and the average t-statistic for test this is zero. In both panels, the period is January 2010–December 2016 using weekly flow and return data. There are 1,081 equity ETFs in each quintile of Panel A and 246 fixed income ETFs for each quintile of Panel B.

in Equation (8) annually, using data for each calendar year, and for each fund.

Panel A in Exhibit 7 presents the data across quintiles of the return gap for equity ETFs. There are 1,081 equity ETFs in each quintile of Panel A and 246 fixed-income ETFs for each quintile of Panel B. Of course, it should be noted that each fund has different types of investors and equity funds, for example, are not all alike. There could also be period specific effects. Nevertheless, even at this aggregate summary level, the near monotonic relationship between fund in each quintile of the return gap and the estimated average γ parameter is clearly evident, with funds with most negative return gap showing the highest γ.

The statistics presented in the table test the null hypothesis that γ is different from zero, based on a two-tail t-test at the 5% significance level. A similar relationship holds for fixed-income ETFs (Panel B), although the γ estimates are noisier, perhaps because of the lower number of observations. In summary, the time-series analysis provides direct support for the hypothesis that the observed return gaps are attributable to trend-chasing behavior.

6 Conclusions

This paper contributes to the growing literature on the divergence between the return the average fund investor experiences and reported buy-and-hold returns. This is a topic of considerable importance to investment advisors, plan sponsors, and policymakers seeking to encourage investors to save more for retirement. Our sample constitutes all US-domiciled open-end mutual funds (active and index) and exchange-traded funds (ETFs), and covers 6,766 fixed income and equity funds with assets in excess of $13 trillion. To our knowledge, our study is the first to analyze return performance for ETFs as well as mutual funds.

We demonstrate directly that return chasing behavior—based on the time-series sensitivity of flows to past returns—explains the cross-sectional pattern of return gaps across funds. Indeed, we find empirically that ETFs where flows are positively correlated to past returns (return chasing) exhibit a greater return gap between what the average investor experiences and what is reported. Using daily flow data, we find the source of the gap is shorter-term trend chasing at the level of a month. We show that liquidity- and flow-related characteristics explain the cross-sectional variation in return gaps for ETFs and mutual funds. Smaller and more narrowly focused funds have the most negative gaps.

Some commentators (e.g., Bogle, 2016) have noted that the ability to trade ETFs intraday can lead to high turnover and sub-par performance for investors. We find little empirical support for this notion. The return gaps across ETFs are small on average, -0.93% for equity funds and -0.52% for fixed-income funds. This finding is very relevant for administrators of defined contribution plans who must decide what funds are offered to participants, and who are often reluctant to offer ETFs despite their low cost. The results also indicate several areas where public policymakers or advisors can meaningfully enhance investor education. For example, advisors may want to complement standard buy-and-hold returns with flow-weighted returns, as part of their ongoing investor education efforts. Our evidence on the cross-section of return gaps suggests that educational efforts be focused on the trading of narrower, more niche products.

Acknowledgments

We thank Ronald Kahn and Andrew Ang for their many helpful comments, as well as Seth Masters for many fruitful discussions. Victoria Li provided valuable research and data assistance.

Notes

[1] Our data covers 1,334 ETFs with assets of $2.4 trillion as of December 31, 2016. We cover both fixed income and equity mutual funds (a total of 5,432 index and active mutual funds with assets of $10.7 trillion as of December 31, 2016).

[2] Source: ETFGI and Bloomberg, as of August 30, 2017. Exchange-traded products (ETPs) are a broader term that includes exchange traded-funds as well as other products such as exchange-traded notes and commodity funds.

[3] For example, Bogle (2016) voices concern that the ability to trade ETFs intraday induces higher turnover relative to mutual funds that detract from investor returns.

[4] Source: DALBAR's 22nd Annual QAIB 2016 which covers the period ending December 31, 2015. Morningstar's "Mind the Gap" is another well-known annual industry report.

[5] See Barber and Odean (2000, 2001) and Barber *et al.* (2007), among others for analyses of trading by individual investors. Calvet *et al.* (2007, 2009) provide analyses of the costs from over-concentration in investors' home region. Berk and Green (2004) present a rational model where investors use the past returns as a guide to infer manager skill.

[6] Source: DALBAR's 22nd Annual QAIB 2016. For the calendar year 2015, the return "gap" for equity mutual fund investors they reported was 3.66%. Coincidently, this is the same amount by which the average fixed income mutual fund investor underperformed the broad Bloomberg Barclays Aggregate Bond Index.

[7] By contrast, Kinniry and Zilbering (2012) examined 21 Vanguard funds, comparing the returns with investors (who own the same portfolios) in various share classes of index mutual funds and ETFs. They find no statistical differences between dollar- and buy-and-hold returns for these 21 Vanguard ETFs and the same funds' conventional share classes, as would be expected if the ability to trade intraday were detrimental to performance.

[8] Our framework could handle a convex flow–performance relation (see, e.g., Lynch and Musto, 2003, among others) although newer research suggests a more linear relationship using gross flows (see, e.g., Rohleder, 2015).

[9] This is consistent with industry standards. It is straightforward to translate the continuous returns of the previous section to discrete period returns.

[10] Internal rates of return are generally not appropriate in comparing real investment projects but are widely used (Fischer and Wermers, 2012) in investment management. The initial asset base of the fund does not enter Equation (3) for buy-and-hold returns. However, the initial asset base of the fund relative to the size of the cash flows does affect the computation in Equation (4), so we normalize cash flows by the asset base for computational purposes. The gap is zero in expectation when (normalized) cash flows and returns are drawn from independent and identical distributions (with constant mean) so that period weights are uncorrelated with returns.

[11] See Dietz (1966) for definition and applications to pension fund performance measurement. The modified Dietz formula assumes that flows occur exactly at halfway point in time within the period of analysis. The Dietz formula can be interpreted (see Fischer and Wermers, 2012, pp. 341–343) as a first-order Taylor series approximation to the internal rate of return within a given interval.

[12] For purely illustrative purposes, we performed 500 simulations where we assumed that fund returns are drawn from an independent and identically random normal distribution with a mean of 5% per annum and annualized volatility of 20% per annum. We further assume that fund flows (scaled by initial AUM) are also random normal with constant volatility but with different means (indicating long-term inflows/outflows to the fund). The fund returns and fund flows are assumed uncorrelated. When the mean fund return and mean fund flows are positive (15% and 20%, respectively), the average return gap between the IRR of Equation (3) and buy-and-hold returns over a year (252 days) is essentially zero (0.01%), as we expect. However, using instead the methodology of Equation (7) to figure dollar-weighted returns yields a materially lower average return gap (−1.14%) for the year.

[13] These ETFs are portfolios of stocks and/or bonds that do not contain derivatives or use leverage.

[14] We also calculated—but do not report—cuts by total dollar AUM, Turnover (the ratio of ADV/AUM) and Cumulative Flow (total net dollar cumulative flows for the calendar year, divided by AUM the beginning of the calendar year).

[15] ETFGI, a consultancy firm, estimates that as of February 2017 there were over 1,200 Smart Beta equity exchange-traded products with assets of $560 billion, including fund targeting low volatility and dividend yield. According to ETFGI, the compound annual growth rate in

Smart Beta assets is over 30% over the previous five years. Source: ETFGI, February 2017.

[16] Over 1,200 US open-end mutual funds and 370 US hedge funds use ETFs as part of their investment strategy. (Source: Cynthia Murphy, "How Investors Are Using Smart Beta ETFs" ETF.com. May 24, 2017.) Crane and Crotty (2018) find that index funds appear to have skill relative to common multi-factor models.

[17] They require that a fund have a minimum of 48 months of data.

References

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

Barber, B. and T. Odean. (2000). "Trading is Hazardous to your Wealth: The Common Stock Investment Performance of Individual Investors," *Journal of Finance* **55**(2), 773–806.

Barber, B. and Odean, T. (2001). "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment," *Quarterly Journal of Economics* **116**(1), 261–292.

Barber, B., Lee, Y., Liu, Y., and Odean, T. (2009). "Just How Much Do Investors Lose by Trading?," *Review of Financial Studies* **22**(2), 609–632.

Berk, J. and R. Green. (2004). "Mutual Fund Flows and Performance in Rational Markets," *Journal of Political Economy* **112**(6), 1269–1295.

Bessembinder, H. (2017). "Do Stocks Outperform Treasury bills?" Working Paper, Arizona State University.

Bogle, J. (2016). "The Lessons We Must Take from ETFs," (December 11) *Financial Times*.

Calvet, L. E., Campbell, J. Y. and Sodini, P. (2007). "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115, pp. 707–747.

Calvet, L. E., Campbell, J. Y., and Sodini, P. (2009)."Measuring the Financial Sophistication of Households," *American Economic Review* **99**, 393–398.

Constantinides, G. (1979). "A Note on the Suboptimality of Dollar-Cost Averaging as an Investment Policy," *Journal of Financial and Quantitative Analysis* **14**(2), 443–450.

Crane, A. and Crotty, K. (2018). "Passive versus Active Fund Performance: Do Index Funds Have Skill?" *Journal of Financial and Quantitative Analysis* **53**(1), 33–64.

Dichev, I. (2007). "What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns," *American Economic Review* **97**(1), 386–401.

Dichev, I. and Yu, G. (2011). "Higher Risk, Lower Returns: What Hedge Fund Investors Really Earn," *Journal of Financial Economics* **100**, 248–263.

Dietz, P. (1966). *Pension Funds: Measuring Investment Performance*. New York: The Free Press, pp. 1–166.

Fischer, B. and Wermers, R. (2012). "Performance Evaluation and Attribution of Security Portfolios," in *Handbook in Economics*. Academic Press, 1–715.

Friesen, G. and Sapp, T. (2007). "Mutual Fund Flows and Investor Returns: An Empirical Examination of Fund Investor Timing Ability," *Journal of Banking and Finance* **31**(9), 2796–2816.

Hsu, J., Myers, B., and Whitby, R. (2016). "Timing Poorly: A Guide to Generating Poor Returns While Investing in Successful Strategies," *Journal of Portfolio Management* **42**(2), 90–98.

Lynch, A. and Musto, D. (2003). "How Investors Interpret Past Fund Returns," *Journal of Finance* **58**, 2033–2058.

Kinnel, R. (30 May, 2017). "Mind the Gap: Global Investor Returns Show the Costs of Bad Timing around the World," *Morningstar Advisor*. Kinniry F., and Y. Zilbering. (2012) *Evaluating Dollar-Weighted Returns of ETFs versus Traditional Fund Returns*. Vanguard Research.

Rohleder, M. (2015). "The Relation between Past Flows and Future Performance: Simple Investment Strategies in the Mutual Fund Sector," *International Journal of Financial Studies* **3**, 3–30.

Zweig, J. (June 2002). "What Fund Investors Really Need to Know," *Money* **31**(6), 110–115.

Important Disclaimer

contained in this material are derived from proprietary and nonproprietary sources deemed by BlackRock to be reliable, are not necessarily all-inclusive and are not guaranteed as to accuracy. Past performance is no guarantee of future results. There is no guarantee that any of these views will come to pass. Reliance upon information in this material is at the sole discretion of the reader.

Keywords: ETFs; mutual funds; performance measurement; investor returns; flows