

Timing Poorly: A Guide to Generating Poor Returns While Investing in Successful Strategies

JASON HSU, BRETT W. MYERS, AND RYAN WHITBY

JASON HSU is co-founder and vice chair at Research Affiliates, LLC, in Newport Beach, CA, an adjunct professor of finance at the University of California Los Angeles Anderson School of Management in Los Angeles, CA, and CEO of Rayliant Global Advisors in Hong Kong, China.
hsu@rallc.com

BRETT W. MYERS is an assistant professor of finance at Texas Tech University's Rawls College of Business in Lubbock, TX.
brett.myers@ttu.edu

RYAN WHITBY is an assistant professor of economics and finance at Utah State University in Logan, UT.
ryan.whitby@usu.edu

Value investing is a historically successful investment strategy.¹ The literature generally agrees on the strategy's robustness but disagrees on the explanations for the success. Although the empirical research focuses exclusively on the time-series returns—or the buy-and-hold returns—of a value portfolio, the investor experience is of course driven by the internal rate of return (IRR), or the dollar-weighted average return. Although the buy-and-hold average portfolio return may be the proper way to document the anomaly, the dollar-weighted average return can shed light on some interesting questions that cannot be addressed by analyzing the buy-and-hold returns. In particular, examining the dollar-weighted returns lets us ask whether investors have actually generated superior IRR consistent with value strategies' reported buy-and-hold outperformance.

One of the most compelling arguments against the behavioral interpretation of the value anomaly is the value effect's persistence and magnitude. How can investors continue to make the same mistake after the anomaly has been so thoroughly documented? It is often argued that any investor or manager promoting value as a persistent anomaly—a free lunch of sorts—must be able to identify the other side of the trade that is funding the excess profit. We believe this question is equally valuable to those who prefer the risk-

based explanation for the value premium. Who is selling value stocks and why do they find value exposure unattractive, given the premium? To shed light on this question, we first examine the dollar-weighted returns (IRR) that value investors have earned in mutual funds and compare them with traditional buy-and-hold returns. In contrast to much of the research on value mutual funds, our primary interest is in the mutual fund investors' investment outcome, rather than the mutual fund managers' performance.

We find that investors in value mutual funds have produced an average IRR that is meaningfully lower than the average returns reported by the corresponding value mutual funds. On average, investors who invest in value mutual funds do not benefit from the excess returns those funds report, because of their allocations' timing. Thus, we must use caution in interpreting mutual funds' documented value premium, because it does not reflect what the average investor actually received. In fact, over time periods with a documented high-value premium, the average value investor in mutual funds has actually done worse than a buy-and-hold investor in an S&P 500 Index fund! We can conclude from this finding that, on average, value mutual fund investors time poorly, directing money to value mutual funds when value stocks are expensive and offer a lower premium and redeeming money

from value mutual funds when value stocks offer a high premium.

This finding could support the following interpretations. First, it suggests that value mutual fund investors have not been extracting profits from investors on the other side of their trades; thus, we would not expect the anomaly to be arbitrated away. In fact, we could argue that value mutual fund investors' activities might contribute to the persistence of the value anomaly, rather than arbitrage it away. A simple behavioral interpretation would argue that glamor investing creates the value premium and trend-chasing contributes to its persistence. Of course, a seemingly poor timing skill could potentially be rational; it is also possible that the average value investor in mutual funds understands that the value premium is time varying and chooses to sell away from value stocks when they are riskiest and offer the highest premium.

We also find that poor timing is not unique to value investors in mutual funds. As we extend our research to examine other style categories—such as growth, small cap, and large cap—we find the same pattern. Regardless of the fund category or style, we observe that the average mutual fund investor's IRR is always meaningfully lower than the buy-and-hold return and always worse than a strategy that naively buys and holds the S&P 500 Index. This is true even gross of fees. Specifically, we find that certain mutual fund investors—such as growth strategy investors—are more prone to generate large return gaps. This suggests the possibility that specific categories of investors time their trades more poorly than others. These findings could also suggest that mutual fund investors, through their poor timing decisions, provide alpha to other investors.²

VALUE MUTUAL FUND INVESTORS AND THE IMPLICATION OF NEGATIVE VALUE PREMIUM

As of the end of 2012, the Investment Company Institute reported that U.S. mutual funds held more than \$13 trillion, of which \$4.3 trillion was in U.S. equity funds alone, accounting for nearly 18% of the U.S. equity market. As comprehensive and detailed data on institutional flows and products are unavailable, we use the mutual fund data to inform us about retail investor behavior and generalize where appropriate. Historically, mutual fund return data have been used to study

professional money managers' skills and behaviors. In our article, we assert that mutual fund flows are largely driven by the behaviors of retail investors and their wealth advisors. Therefore, this dataset can be used to help us gain additional insight into retail asset owners and their consultants. As such, our study complements previous studies that have focused on data from discount brokerages as a way to understand retail investors.³

We compare the dollar-weighted versus the buy-and-hold average returns of value mutual funds using the CRSP mutual fund database. These calculations help us estimate the value premium that has actually accrued to the average mutual fund value investor and shed light on how those investors time their value strategy allocation. We also apply the same methodology to examine other fund categories to help us better understand their timing behavior. We follow Dichev [2007] and use an IRR-based methodology to estimate the average investors' experience. This methodology recognizes that the average mutual fund investor trades in and out of funds and can earn a return quite different from that estimated by the standard buy-and-hold calculation.

We find that, to the average investor, the value premium has largely been an illusion. Although the buy-and-hold average returns for value mutual funds have outperformed the market portfolio, the dollar-weighted average returns of these same funds meaningfully underperformed the market portfolio. In other words, although value managers have largely been successful in exploiting the value premium to outperform the market, value investors—who ultimately allocate capital to value managers—have managed to reverse the sign on the premium earned through their timing decisions. Specifically, they seem to allocate to value mutual funds when value stocks are relatively expensive and offer a low premium, and vice versa.

We find a similar pattern for the average growth investor in mutual funds. This may be less interesting, however, as the average growth mutual fund already underperforms the market. Indeed, in a later section, we show evidence that growth fund investors produce the largest return deficit from timing.

Our evidence is consistent with results reported by Jason Zweig in his 2002 *Money Magazine* investigative report (Zweig [2002])⁴ and by Russel Kinnel, director of mutual fund research at Morningstar, in his regular online column (Kinnel [2013]).⁵ There is a meaningful gap between the excess return reported by the average

mutual fund manager and the returns to capital earned by the average investor; the former captures manager skill, while the latter includes both manager and investor skill.⁶

Our findings appear to be consistent with the hypothesis that investors attempt to time their mutual fund allocations. Our evidence supports the idea that the average investor chases performance and allocates to value funds after periods of strong relative performance, and vice versa. Unfortunately, this—coupled with the evidence on the mean-reversion tendency for value stock outperformance (see Asness et al. [2000] and Cohen, Polk, and Vuolteenaho [2003])—implies that they time poorly. The average value investor plows money into value funds before periods of poor performance and allocates away before periods of strong performance. This observation may help explain why the value premium has remained so persistent. We simply would not expect the value anomaly to be arbitrated away by poorly timed fund flows that have led to substantial excess losses instead of excess profits for value investors. Quite the opposite: the poor timing might be generating excess profits for others. Because equity mutual funds account for nearly 20% of the U.S. equity market, the dollar alpha supplied by mutual fund investors—through their timing—could be quite substantial.

Given that the average investor in value mutual funds has not earned a positive value premium, the question “who is on the other side of the trade?” takes on a very different spin. Instead of wondering about the potential sucker on the other side of the value trade, we ask, “why does the average investor in value mutual funds consistently give away the value premium?” The “poor timing” interpretation is behavioral in nature and argues that investors overextrapolate from recent fund performance and mistakenly forecast high future returns instead of low future returns.⁷

Of course, poor timing is just one possible interpretation of the gap between buy-and-hold and dollar-weighted average fund returns. Perhaps value investors in mutual funds have rational reasons to sell value funds when their expected returns are relatively high and buy when expected returns are relatively low. For example, missing out on technology stocks in 1999 and investing in bank stocks in 2009 exposed investors to the risk of ignoring a very significant paradigm shift in the economy. Under this framework, an implicit assumption is that value stocks are riskier when cheap and safer when

expensive. However, we would also argue that mutual fund investors have sufficiently higher risk aversion than institutional investors, who are likely on the other side of the mutual fund trades.⁸ The higher relative risk aversion is likely needed to generate willingness to trade away from value funds when they have high expected returns and to plow back into value funds when they have low expected returns. When we examine the return gap for other style investors in later sections, we offer evidence that suggests the poor timing is more likely behavioral in nature.

Ultimately, identifying the precise explanation for the return gap between fund return and investor IRR is beyond the scope of this article. We are interested in documenting that the average value investor in mutual funds does not earn the value premium. This evidence allows us to argue that value investors in mutual funds might, in fact, be subsidizing the value premium by actively trading value mutual funds. This last point warrants additional exposition. In the aggregate, all investors earn the market return. For an investor to generate \$1 worth of value premium, someone must give up \$1 of return. It is often argued that value investors in mutual funds have been on the positive side of that trade, and there has been much discussion about who is funding this excess return. Our findings suggest that the reality is quite the opposite.

MUTUAL FUND DATA AND IRR CALCULATION METHODOLOGY DESCRIPTION

In order to analyze investor behavior around mutual fund allocations, we use the CRSP Survivorship-Bias-Free U.S. Mutual Fund Database as our source for monthly return and quarterly fund characteristic data. We examine only equity mutual funds and classify a fund as an equity fund if its portfolio allocates an average of more than 80% to equities. If we do not have any information on a fund’s equity allocation, then we classify it as an equity fund if its Wiesenberger, Strategic Insights, Lipper, or Morningstar classification indicates that it is an equity fund. We also exclude funds with total net assets (TNA) of less than \$10 million. The sample period extends from 1991 through 2013 and is limited by the availability of the data necessary for us to perform our analysis (e.g., monthly data on fund TNA, data on fund expense ratios, etc.).

In order to classify funds by size or style (e.g., value or growth), we use the fund's prospectus-stated benchmark, which we obtain from Morningstar Direct and map to the CRSP data using the fund's CUSIP identifier. When a stated benchmark is not available, we assign a benchmark index using the active share methodology of Cremers and Petajisto [2009]. To do this, we first calculate the active share of each fund relative to 24 of the benchmarks that equity mutual funds most commonly use, a list that includes the S&P 500, S&P 400, S&P 600, Russell 1000, Russell 2000, Russell 2500, Russell 3000, and Russell Midcap indices, along with their value and growth derivatives. We then assign the index corresponding to a fund's lowest active share as that fund's benchmark.

Much of our analysis involves creating TNA-weighted portfolios of equity mutual funds and comparing their buy-and-hold return with their dollar-weighted return.⁹ The dollar-weighted return (or the realized IRR) adjusts the average return according to the timing and magnitude of capital flows and is a more accurate measure of the actual returns realized by equity fund investors.

We calculate the IRR for each portfolio of mutual funds, and later for each individual mutual fund, using the methodology of Dichev [2007]. First, we estimate the capital flows relative to our portfolio of mutual funds by:

$$Distribution_t = TNA_{t-1}(1 + r_t) - TNA_t \quad (1)$$

where TNA_{t-1} is the portfolio's total net assets at the beginning of month t , TNA_t is the portfolio's total net

assets at the end of month t , and r_t is the return of the portfolio for month t . Normalizing the starting TNA value at 1 and using the distribution and return data to arrive at the end-period TNA value, we are able to calculate the realized IRR, or dollar-weighted return, for both portfolios of mutual funds, as well as each individual mutual fund. We report the summary statistics in Exhibit 1.

THE GAP BETWEEN BUY-AND-HOLD RETURN AND ACTUAL INVESTOR IRR

We begin our analysis of the effect of timing on investor results by comparing the buy-and-hold return with the dollar-weighted average return, or the IRR. The dollar-weighted average return accounts for the flows to and from the mutual fund and is representative of the return an average investor would have received over the period after adjusting for capital flow.

Exhibit 2 reports the differences in dollar-weighted and buy-and-hold returns for several standard mutual fund classifications, all net of fees. The first column reports the dollar-weighted returns and the second reports the buy-and-hold returns. As discussed, this allows us to examine the average mutual fund investor's ability to time flows in and out of mutual funds. The timing of fund allocation plays a crucial role in the final return investors receive. The third column reports the differences between the dollar-weighted returns and the buy-and-hold returns, with the corresponding p -values reported in the fourth column. The returns in columns 1 and 2 can also be compared with the buy-and-hold

EXHIBIT 1

Descriptive Statistics—Monthly Return Data (January 1991–June 2013)

This table contains summary statistics for the monthly equity mutual fund data in our sample. Summary statistics are separated by fund type and include the total number of monthly fund observations, the number of unique funds, and information on the total net assets, expense ratio, and percentage of funds that charge a front-end load.

	Total Observations	Unique Funds	Total Net Assets (\$ MM)			Expense Ratio			Percent Load
			Mean	Median	Std Dev	Mean	Median	Std Dev	
All funds	1,076,143	18,665	623.6	96.3	2,873.5	1.35%	1.26%	0.58%	23.2%
Growth funds	281,733	3,280	438.6	86.6	1,593.5	1.46%	1.35%	0.56%	24.4%
Value funds	235,865	2,836	558.1	93.0	2,123.4	1.36%	1.26%	0.54%	23.0%
Small-cap funds	169,751	1,868	308.4	80.3	1,047.0	1.44%	1.36%	0.56%	19.6%
Large-cap funds	656,165	7,882	752.1	101.9	3,406.6	1.30%	1.23%	0.59%	24.4%

Source: Research Affiliates, using data from the CRSP Mutual Funds Database.

EXHIBIT 2

Comparison of Dollar- and Time-Weighted Returns for Mutual Fund Portfolios Based on Fund Type (January 1991–June 2013)

This table reports buy-and-hold versus dollar-weighted return portfolios of different categories of mutual funds weighted by total net assets. The buy-and-hold return is the geometric average return over the sample period. The dollar-weighted return is the category IRR (inclusive of all category mutual funds weighted by their total assets), calculated as in Dichev [2007]. We report the difference between the dollar-weighted and buy-and-hold return, as well as a derived *p*-value estimated from bootstrap tests under the null hypothesis that the difference between the realized dollar-weighted and time-weighted returns is not statistically different from the difference between a random dollar-weighted return and the realized time-weighted return.

Fund Classification	Dollar-Weighted Return	Buy & Hold Return	Difference (DW Ret – TW Ret)	<i>p</i> -Value on Difference
S&P 500 Total Return Index		8.97%		
All Funds	6.87%	8.81%	–1.94%	0.000
Growth Funds	5.22%	8.38%	–3.16%	0.000
Value Funds	8.05%	9.36%	–1.31%	0.003
Small-Cap Funds	8.23%	9.78%	–1.55%	0.000
Large-Cap Funds	6.76%	8.66%	–1.90%	0.000
Index Funds	6.95%	9.66%	–2.72%	0.003
Active Funds	6.85%	9.73%	–1.88%	0.000

Source: Research Affiliates, using data from the CRSP Mutual Funds Database and Morningstar Direct.

return for the S&P 500 Index, which dominates the dollar-weighted average returns reported for all fund categories.

The difference between the dollar-weighted and buy-and-hold returns is negative and economically significant in every mutual fund style category that we examined.¹⁰ As previously discussed, the IRR for value mutual funds is meaningfully lower than the buy-and-hold return, with a difference of 1.31%. Consistent with the literature on value outperformance, the average value fund manager has outperformed the buy-and-hold market benchmark by 39 basis points per year, net of fees (unadjusted for risk or style). In contrast, the average value investor in mutual funds has actually earned a negative 92 basis point value premium per year, relative to the market index. The observation that value investors in mutual funds may actually be supplying premium to others instead of extracting premium has interesting implications, as we discussed in previous sections. Our

focus now expands to examine the timing-induced return gap across other classifications.

Over the last 20 years, mutual fund investors appear to have given up almost 2% per year because of their trade timing. There is ample evidence that excessive securities trading is detrimental to returns (see Odean [1999], Barber and Odean [2000, 2002], Barber et al. [2009], and Bailey, Kumar, and Ng [2011]). We examine this from the perspective of the mutual fund investor and find that the average mutual fund investor also suffers because of the timing of trades.

Growth fund investors see the most dramatic difference in their timing-induced return gap. The average growth fund investor experienced an annual gap of –3.16% between the dollar-weighted and buy-and-hold returns; this is more than twice the gap the average value fund investor experienced. We can only speculate on the reason behind this larger return gap. Using the behavioral argument, we could imagine that investors in growth mutual funds might be more prone to chase hot managers just as they would hot stocks, to their own detriment. Given the documented cyclicity in value/growth performance, this pro-cyclical timing strategy would have worked poorly.

We note that the same comments that apply to value investors in mutual funds also apply to small-cap mutual fund investors. The average small-cap mutual fund manager has outperformed the S&P 500 by 81 basis points per year, but the average small-cap mutual fund investor has underperformed the S&P 500 by 74 basis points per year. Similarly, we again observe that the gap experienced by large-cap mutual fund investors is larger than the gap experienced by small-cap mutual fund investors, although the differences are less stark than in the comparison of value and growth. Again, it is interesting to wonder why retail investors in small-cap strategies might be less prone to commit timing mistakes in their fund allocation.

In addition to analyzing the return gap with respect to the traditional fund-style categories, we also wish to examine other fund characteristics that might provide additional insight into retail fund-trading behavior. Specifically, we are interested in further exploring the story that more sophisticated fund investors are less likely to actively trade their mutual funds in an unproductive manner. A mutual fund's expense ratio has been shown to influence fund returns and flows.¹¹

In Exhibit 3, we separate funds into quintiles based on their expense ratios. Although all the expense ratio quintiles have time-weighted returns that are higher than their dollar-weighted returns, the differences between the low and high expense ratios is striking. Column 3 reports the differences, which monotonically increase as the expense ratio gets larger. The differences range from -1.34% per year for the lowest expense-ratio funds to -4.01% per year for the high expense-ratio funds.

Again, identifying the exact reason for this pattern is beyond the scope of this article. However, this pattern seems consistent with our story about financial knowledge level and mutual fund trading skill. We argue that investors who invest in high expense-ratio funds might, on average, be less financially knowledgeable or less able to access to professional wealth advisors. Thus they are more likely to overextrapolate and aggressively trade their mutual funds in order to chase past fund performance. On the other hand, more sophisticated investors are more likely to adopt either a buy-and-hold strategy

EXHIBIT 3

Comparison of Dollar- and Time-Weighted Returns for Mutual Fund Portfolios Based on Expense Ratio (January 1991–June 2013)

This table reports time-weighted versus dollar-weighted returns for mutual funds that are weighted by total net assets and ranked by expense ratio. We sorted mutual funds into quintiles based on monthly expense ratio. The time-weighted return is the geometric average return over the sample period. The dollar-weighted return is the IRR of each mutual fund category, calculated as in Dichev [2007]. We report the difference between the dollar- and time-weighted returns, as well as a derived p -value estimated from bootstrap tests under the null hypothesis that the difference between the realized dollar-weighted and time-weighted returns is not statistically different from the difference between a random dollar-weighted return and the realized time-weighted return.

Expense Ratio	Dollar-Weighted Return	Time-Weighted Return	Difference (DW Ret – TW Ret)	p -Value on Difference
Low	7.88%	9.22%	-1.34%	0.042
2	6.93%	8.85%	-1.92%	0.000
3	6.07%	8.35%	-2.28%	0.000
4	4.80%	7.84%	-3.03%	0.000
High	2.87%	6.88%	-4.01%	0.000

Source: Research Affiliates, using data from the CRSP Mutual Funds Database.

EXHIBIT 4

Comparison of Dollar- and Time-Weighted Returns for Institutional & Retail Funds (January 1991–June 2013)

This table reports time-weighted versus dollar-weighted returns for institutional and retail funds. The time-weighted return is the geometric average return over the sample period. The dollar-weighted return is the IRR of each mutual fund category, calculated as in Dichev [2007]. We report the difference between the dollar- and time-weighted returns, as well as a derived p -value estimated from bootstrap tests under the null hypothesis that the difference between the realized dollar-weighted and time-weighted returns is not statistically different from the difference between a random dollar-weighted return and the realized time-weighted return.

Fund Type	Dollar-Weighted Return	Time-Weighted Return	Difference (DW Ret – TW Ret)	p -Value on Difference
Institutional Funds	7.07%	8.90%	-1.83%	0.132
Retail Funds	6.82%	8.82%	-2.00%	0.000

Source: Research Affiliates, using data from the CRSP Mutual Funds Database.

or a contrarian strategy, which would have been more successful historically.

Exhibit 4 reports the differences between dollar-weighted and time-weighted returns for institutional versus retail funds. Investors who qualify for and purchase institutional share class funds experience a smaller return gap, whereas investors in retail funds experience a larger return gap. Although the institutional definition also captures large retirement funds that are dominated by small investors, the idea that these retirement funds or institutions do a better job timing the market or are more focused on buy-and-hold strategies is not surprising.

MULTIVARIATE ANALYSIS

We recognize that there are a variety of fund characteristics that could influence the gap between dollar-weighted and time-weighted returns. In this section, we extend our analysis to include multivariate tests, to more rigorously examine the potential determinants of the return gap. We estimate the parameters of the following cross-sectional mutual fund model in our sample:

$$R_i^{DW} - R_i^{TW} = \delta + \beta x_i + \varepsilon_i \quad (2)$$

In this model, R_i^{DW} and R_i^{TW} are the dollar- and time-weighted returns for fund i as measured over the life of the fund, and x_i is a vector of fund characteristics. These include the log of the fund's average total net asset ratio (defined as the ratio of a fund's monthly TNA to the monthly cross-sectional average total assets for funds in the sample), the average expense ratio for each fund, and a dummy for whether or not the fund charges a front-end load fee. Also included are dummy variables indicating whether a fund is a growth, value, small-cap, large-cap, index, or institutional fund. To be included in the sample, funds must have a minimum of 48 months of return data. Exhibit 5 reports the results from this multivariate analysis. Independent variables include the natural log of the total net assets of the fund $Ln(TNA)$, the fund's expense ratio, and the fund's age. We use robust standard errors to compute the t -statistics reported in parentheses.

Column 5 of Exhibit 5 reports the results from our most complete specification, so for brevity, we will focus our analysis on that model. Consistent with our univariate results, we find that our main variables of interest—the growth and value indicators—are significant with the expected signs. Investors in value funds have experienced a smaller gap, on average, between the dollar- and time-weighted returns; on the other hand, investors in growth funds have experienced a larger gap between the dollar- and time-weighted returns, all else being equal. Similarly, index fund investors have experienced a smaller return gap on average, while non-index fund investors have experienced a larger return gap. Again, investors who qualify for and purchase institutional share class funds experience a smaller return gap, whereas investors in retail funds experience a larger return gap. These findings are consistent with our findings in the previous section. Although the institutional indicator also captures large

EXHIBIT 5

Return Gap Regressions (January 1991–June 2013)

This table reports the result of regressions where the dependent variable is the difference between a fund's dollar- and time-weighted returns:

$$R_i^{DW} - R_i^{TW} = \delta + \beta x_i + \varepsilon_i$$

where R_i^{DW} and R_i^{TW} are the dollar- and time-weighted returns for fund i as measured over the life of the fund, and x_i is a vector of fund characteristics. These include the log of the fund's average total net asset ratio (defined as the ratio of a fund's monthly total net assets to the monthly cross-sectional average total assets for funds in the sample), the average expense ratio for each fund, and a dummy for whether or not the fund charges a front-end load fee. Also included are dummy variables indicating whether a fund is a growth, value, small-cap, large-cap, index, or institutional fund. To be included in the sample, funds must have a minimum of 48 months of return data. Standard errors are adjusted for clustering at the portfolio level, and ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
ln(TNA ratio)	-0.0019*** (-4.00)				-0.0015*** (-3.45)
Expense ratio	-0.0049*** (-8.45)				-0.0027*** (-3.70)
Front load fund	-0.0038*** (-4.34)				-0.0010 (-1.05)
Growth fund		-0.0019 (-1.31)		-0.0018 (-1.21)	-0.0015 (-1.02)
Value fund		0.0074*** (5.14)		0.0075*** (5.15)	0.0076*** (5.23)
Small-cap fund			0.0004 (0.24)	0.0008 (0.43)	0.0009 (0.50)
Large-cap fund			0.0003 (0.20)	0.0008 (0.57)	0.0007 (0.46)
Index fund		0.0103*** (10.11)	0.0106*** (10.51)	0.0103*** (10.14)	0.0071*** (5.75)
Institutional fund		0.0074*** (3.73)	0.0067*** (3.55)	0.0074*** (3.73)	0.0050** (2.34)
Constant	-0.0077*** (-12.43)	-0.0141*** (-15.41)	-0.0132*** (-9.91)	-0.0148*** (-9.79)	-0.0131*** (-8.63)
Observations	8401	8401	8401	8401	8401
Adjusted R2	0.015	0.023	0.016	0.023	0.026

Source: Research Affiliates, using data from the CRSP Mutual Funds Database.

retirement funds that are dominated by small investors, the idea that these retirement funds or institutions do a better job timing the market or are more focused on buy-and-hold strategies is not surprising. The finding that index funds also have a smaller return gap is also consistent with index funds as the focus of buy-and-hold investors. The multivariate analysis confirms the robustness of the results from our univariate analysis and supports the hypothesis that less sophisticated investors are more likely to time poorly.

CONCLUSION

When examining the source, persistency, and the benefit of a return premium, it is important to consider how the average investor might implement an investment strategy to take advantage of that premium. The average return from a buy-and-hold implementation ignores the fact that many investors trade on a regular basis and that the actual return the average investor receives might be very different from the buy-and-hold return. Dollar-weighted returns account for this trading by augmenting the average with information about fund flows. If investors are able to time their movements in and out of mutual funds well (for example, getting into value before value does well and getting out of value before it does poorly), then they could receive returns that are substantially higher than the simple buy-and-hold average. On the other hand, if investors do not time their movements into and out of mutual funds well, then they could in theory miss all the return premium associated with a specific strategy. Indeed, when poor timing leads to an outcome where mutual fund investors do not capture any value or small-cap premium in reality, then we might be more willing to believe that the anomalies might persist.

We examine the difference between dollar-weighted returns, or the IRR of a mutual fund that considers the flows into and out of each fund, and the time-weighted or buy-and-hold returns that are usually reported for each fund. We find that average investors do not time their allocations well and actually underperform the buy-and-hold benchmark by almost 2% per year. This has significant implications. This loss in return through poor timing must be captured by someone else and addresses the question of how the value premium persists and who is on the losing end of the trade. Indeed, for fund categories where the average manager has outperformed the broad market benchmark net of fees, as measured by buy-and-hold returns such as value and small-cap style category, the investor's negative timing skill actually results in an underperformance, as measured by IRR.

Although the average fund investor times poorly, we also find that the return gap is quite different for different investor cohorts. The return gap is much larger for investors in growth funds compared with value funds, and is much larger for investors in large-cap funds compared with small-cap funds. We also find that the return gap is much larger for investors who invest in funds

with higher expense ratios, in non-index funds, and in non-institutional share class funds. We hypothesize that investors who time poorly tend to be less sophisticated investors, as evidenced by their investments in growth-oriented, large-cap, high-fee, active, and/or retail-oriented mutual funds. Our evidence is consistent with this hypothesis and has meaningful implications. As more pension savings occur in defined contribution plans, where individual investors can actively allocate to mutual funds, financial education will be critical in helping individual investors refrain from their natural instinct to actively trade their funds.

ENDNOTES

¹See Chan and Lakonishok [2004] for a review of the literature on the value effect.

²This observation is true whether one believes that the poor timing is rational or behaviorally based.

³See articles by Brad Barber and Terrance Odean for a number of the most foundational articles on individual investor behavioral biases, based on analyzing detailed trading data from investors' brokerage account.

⁴Zweig looked at the gap between time- and dollar-weighted average returns for mutual funds covered by Lippers from 1998 to 2001. See [http://www.winlab.rutgers.edu/~narayan/Course/Wireless_Revolution/real inv gains-fds.pdf](http://www.winlab.rutgers.edu/~narayan/Course/Wireless_Revolution/real%20inv%20gains-fds.pdf).

⁵Kinnel [2013] looked at the gap between time-weighted and dollar-weighted average returns for mutual funds tracked by Morningstar over the most recent one-year, three-year, five-year, and ten-year horizons. See <http://www.morningstar.com/advisor/t/70540275/mind-the-gap-why-investors-lag-funds.htm>.

⁶Here, we use "skill" very loosely to refer to both excess return from taking added risk or exploiting informational advantage.

⁷Our evidence can support the behavioral hypothesis that mutual fund investors chase glamor mutual funds. However, it does appear that mutual fund investors, despite their poor timing, still achieve better results than those of the individual investors studied by Barber and Odean [2000, 2001].

⁸An investigative report that examined the investments and performance of the Norway Sovereign Wealth Fund, by Ang, Goetzmann, and Schaefer [2009] for the Norwegian Central Bank, suggests that even large and professional institutional investors can exhibit similar trend-chasing behaviors. Various institutional asset consultants, who advise large pension funds in their manager selection and asset allocation, confirm this behavior.

⁹Note that we use the TNA of the mutual funds in a category as weight to compute the category return for a particular style. This means that, in our final comparison, we do not capture the allocation effect from the fund investors' timing between funds in the same category.

¹⁰To test for the return difference's significance, we follow Dichev [2007] and simulate the null distribution of dollar-weighted mutual fund returns by holding the vector of realized returns constant but randomizing the distributions as a percentage of beginning-of-month TNA. After each randomization, we use the resulting capital flows to calculate the portfolio's IRR. We repeat this routine 1,000 times. We then compare the difference between the buy-and-hold and each randomized dollar-weighted return with the actual difference between the buy-and-hold return and dollar-weighted return, in order to produce a statistically significant test of our results. Exhibit 2 shows that the differences between dollar-weighted and buy-and-hold returns are highly significant, with *p*-values of less than 2%.

¹¹Numerous studies have found that funds with high expense ratios tend to have poorer performance over time, versus their cheaper category peers. Fee-only advisors, as well as mutual fund consultants such as Morningstar, are also known to focus very heavily on expense ratios when making mutual fund recommendations.

REFERENCES

- Ang, A., W. Goetzmann, and S. Schaefer. *Evaluation of Active Management of the Norwegian GPF*. Norway: Ministry of Finance, 2009.
- Asness, C., J. Friedman, R. Krail, and J. Liew. "Style Timing: Value versus Growth." *The Journal of Portfolio Management*, Vol. 26, No. 3 (Spring 2000), pp. 50-60.
- Bailey, W., A. Kumar, and D. Ng. "Behavioral Biases of Mutual Fund Investors." *Journal of Financial Economics*, Vol. 102, No. 1 (October 2011), pp. 1-27.
- Barber, B. and T. Odean. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *Journal of Finance*, Vol. 55, No. 2 (April 2000), pp. 773-806.
- . "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *Quarterly Journal of Economics*, Vol. 116, No. 1 (February 2001), pp. 261-292.
- . "Online Investors: Do the Slow Die First?" *Review of Financial Studies*, Vol. 15, No. 2 (March 2002), pp. 455-487.
- Barber, B., Y. Lee, Y. Liu, and T. Odean. "Just How Much Do Investors Lose by Trading?" *Review of Financial Studies*, Vol. 22, No. 2 (February 2009), pp. 609-632.
- Chan, L., and J. Lakonishok. "Value and Growth Investing: Review and Update." *Financial Analysts Journal*, Vol. 60, No. 1 (January/February 2004), pp. 71-86.
- Cohen, R., C. Polk, and T. Vuolteenaho. "The Value Spread." *Journal of Finance*, Vol. 58, No. 2 (April 2003), pp. 609-641.
- Cremers, M., and A. Petajisto. "How Active Is Your Fund Manager? A New Measure That Predicts Performance." *Review of Financial Studies*, Vol. 22, No. 9 (September 2009), pp. 3329-3365.
- Dichev, I. "What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns." *American Economic Review*, Vol. 97, No. 1 (March 2007), pp. 386-401.
- Kinnel, R. "Mind the Gap: Why Investors Lag Funds." *Morningstar Advisor* (February 4, 2013).
- Odean, T. "Do Investors Trade Too Much?" *American Economic Review*, Vol. 89, No. 5 (December 1999), pp. 1279-1298.
- Zweig, J. "What Fund Investors Really Need to Know." *Money*, Vol. 31, No. 6 (June 2002), pp. 110-115.

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@ijournals.com or 212-224-3675.