Returns to Stock-Picking Skill

By C. Thomas Howard, PhD
Returns to Stock-Picking Skill

By C. Thomas Howard, PhD

We are witnessing a dramatic flow of money out of active equity mutual funds and a similarly sized flow into index funds. A large portion of these outflows are from so-called closet indexers, funds that claim to be active equity managers but, upon closer inspection, closely track an index while charging active fees. Investors have wised up to this and are heading for the exits, moving into much lower-fee passive funds that provide the same underlying equity return.

Many mistakenly believe that this is occurring because active investment teams are unskilled stock pickers, and this misunderstanding is the focus of this article. I argue that the returns generated by the skill of these teams can be overwhelmed by poor portfolio decisions and market environments that do not favor successful stock picking. As a consequence, the returns delivered to investors often are disappointing, underperforming the fund’s benchmark even when team members are skilled stock pickers.

Poor portfolio decisions are the result of industry incentives that drive active equity funds to turn themselves into closet indexers by creating an unhealthy environment that I refer to as the “closet indexing factory.” The resulting industry infrastructure encourages funds to capture all of the economic rents flowing from their skill, while delivering inferior returns to investors. It is the fruit of this misguided industry structure against which investors are rebelling.

But even if these perverse incentives suddenly evaporated, poor fund performance may persist because returns delivered to investors by active equity funds are impacted by market conditions. Much evidence supports the existence of “stock-picking markets” that are followed by periods during which stock picking is not rewarded.

I refer to the returns delivered to investors by active equity funds as “returns-to-skill.” However, before exploring the impact of closet indexing and market environment on returns-to-skill, I address the question of whether stock-picking skill even exists.

**STOCK-PICKING SKILL**

Many who argue that stock-picking skill does not exist base their conclusions on the widely accepted empirical result that the average active equity fund underperforms. But fund returns are driven by a combination of skill, portfolio decisions, and market conditions. So, fund returns are a distorted measure of skill. In order to measure skill by means of actual performance, it is necessary to remove the distortions created by portfolio decisions and market conditions.

Another argument frequently put forward is a mistaken extension of Sharpe (1991), which points out that active equity management must be a zero-sum game around the market portfolio, with as many losers as winners. Those opposed to active management extend this argument to the mutual fund universe, concluding there must be as many losers as winners meaning that the average fund must underperform by the amount of the management fees charged.

The problem with this line of attack is that active equity mutual funds hold only 10 percent of all equities traded on U.S. exchanges ($4.5 trillion in funds versus $46 trillion on U.S. exchanges). It is entirely possible that stocks held by funds may outperform the other 90 percent held by others, and Sharpe explicitly acknowledges this possibility. This is particularly the case when focusing on the “best idea stocks,” discussed below.

A preliminary indication of skill is presented in table 1, which is based on the active equity mutual fund universe from 1997 through 2017. In table 1, the equally weighted annual gross alpha is 104 basis points (bps), offsetting all but 30 bps of the average management fee of 134 bps. The assets under management (AUM)
weighted gross alpha is 34 bps, resulting in a –78 bps net alpha. Thus, both equal-weighted and AUM-weighted gross alphas reported in the bottom row of table 1 reveal stock-picking skill among active funds. Furthermore, the equal-weighted alphas imply that smaller funds deliver higher returns-to-skill than do larger funds, as evidenced by the equal-weighted alphas (both gross and net) being two to three times greater than the AUM-weighted alphas.

Cremers et al. (2018) surveyed 103 mutual fund studies published over 20 years and found that many provide evidence of skilled fund managers. Table 2 summarizes the measures used to gauge the extent of manager skill.

The 103 studies surveyed used a range of measures to find evidence of considerable skill among primarily U.S. active equity mutual fund investment teams. Based on this extensive evidence, Cremers et al. (2018) stated, “While the debate between active and passive is not settled and many research challenges remain, we conclude that the current academic literature finds active management more promising for investors than the conventional wisdom claims.”

Considerable evidence shows that mutual funds are able to exploit market mispricings and, in turn, deliver alpha to their investors. Even without adjusting for the performance distortions of subsequent portfolio decisions, the benefits of stock-picking skill show through in fund performance studies.

To measure directly the benefits of stock-picking skill, Wermers (2000) examined individual stocks held by active funds. He found that the average holding outperformed its benchmark by 130 bps, almost identical to the equal-weighted management fees reported in table 1. These results reveal that investment teams have enough stock-picking skill to cover their expenses. If funds held only stocks, the average fund alpha would be zero. But funds also hold cash, with the resulting drag helping to explain the negative alpha for the average fund.

Beyond uncovering overall stock-picking skill, it is possible to identify the “best idea stocks” held by managers. A best idea stock is a stock that the investment team, by means of careful research, has determined has the best chance of generating a superior return. These also are referred to as high-conviction stocks. One of the more interesting papers to accomplish this is Cohen et al. (2010), who find, “...the U.S. stock market does not appear to be efficiently priced, since even the typical active mutual fund manager is able to identify stocks that outperform by economically and statistically large amounts.” These results, presented in figure 1, are based on the performance of the typical manager-ranked best ideas. Figure 1 shows that

---

**Table 1**

<table>
<thead>
<tr>
<th></th>
<th>Equal Weighted</th>
<th>AUM Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net Alpha</td>
<td>–30 bps</td>
<td>–78 bps</td>
</tr>
<tr>
<td>Management Fees</td>
<td>134 bps</td>
<td>102 bps</td>
</tr>
<tr>
<td>Gross Alpha</td>
<td>104 bps</td>
<td>34 bps</td>
</tr>
</tbody>
</table>

*U.S. active equity mutual funds from March 1997 to March 2017, resulting in more than 340,000 fund-month observations. Funds with AUM less than $20 million are eliminated. Alpha is monthly fund return net of best-fit benchmark. Sources: Morningstar and AthenaInvest.

**Table 2**

<table>
<thead>
<tr>
<th>Skill Measure</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data and Benchmark Issues</td>
<td>5</td>
</tr>
<tr>
<td>Fund Insider Information</td>
<td>10</td>
</tr>
<tr>
<td>Fund Public Information</td>
<td>26</td>
</tr>
<tr>
<td>Selection, Timing, Strategy Consistency</td>
<td>43</td>
</tr>
<tr>
<td>Fund Structure and Environment</td>
<td>19</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>103</strong></td>
</tr>
</tbody>
</table>

*Source: Cremers et al. (2018)

**Figure 1**

**STOCK SUBSEQUENT MONTHLY (BPS) ALPHA BY PORTFOLIO RANK**

Based on Cohen et al. (2010), figure 3. The top line represents the mean bps monthly subsequent six-factor alphas for the stock with the largest relative weight in each portfolio, the next highest, and so forth. The bottom line represents two standard deviations below the mean.
This challenges the conventional wisdom that such skill is a rare commodity.

**CLOSEST INDEXING FACTORY**

If stock-picking skill is so widespread, then why do so many funds underperform? The typical manager is able to identify a handful of superior stock ideas and stocks. But rather than limiting the portfolio to these good ideas, managers purchase additional stocks that, by all accounts, hurt portfolio performance. Managers may speak passionately about their best ideas, but they often speak meekly, even apologetically, about the additional stocks purchased to round out the portfolio. This overdiversification is the result of powerful industry forces that even the best managers have a hard time resisting.

As represented in figure 1, the marginal after-the-fact alpha generated by a manager’s best ideas declines as each additional stock is added to a portfolio. Combining this with the diminishing marginal volatility reduction shown in figure 3, one can argue that, if volatility and return optimization is the goal, a properly diversified portfolio comprises 10 to 20 stocks. This means the typical manager should be holding a highly concentrated best ideas portfolio, which has the opportunity to generate the highest long-term compound return for the investor (the annual compound return can be approximated by subtracting one-half of the portfolio variance from the average annual return).

However, the median active equity mutual fund holds 75 stocks, which means dramatic overdiversification is common. This creates serious fund performance problems. As a manager moves down in best idea rank, stock alphas eventually turn negative; in other words, lower-ranked stocks morph into bad idea stocks. Indeed, in the typical fund portfolio, bad idea or low-conviction stocks outnumber good ideas by 3 to 1. So, in the mutual fund market, overdiversification is the rule and proper diversification is the exception.

---

**Figure 2**

**IMPACT ON FUND ALPHA OF INVESTING IN HIGH-CONVICTION BEST IDEA STOCKS**

Based on single variable, subsequent gross fund alpha regressions estimated using a data set of 44 million stock-month equity fund holdings over January 2001–September 2014.

**Figure 3**

**INCREMENTAL REDUCTION IN PORTFOLIO STANDARD DEVIATION**

Based on the following assumptions: Each stock added to the portfolio has the same standard deviation, has the same cross-stock correlation of 0.11, and is equally weighted in the portfolio.
This raises the perplexing question of why managers purchase so many stocks. Why not hold a small number of stocks, probably somewhere between 10 and 20, thus concentrating the portfolio in best ideas? Because the following industry structures and incentives drive active equity funds to oversimplify, buy more funds than they need.

**AUM-based compensation:** Funds are compensated based on AUM, which is an irresistible incentive to grow large. But as a fund grows, it becomes increasingly difficult to hold a small number of stocks. As revenues grow with fund size, funds find it beneficial to invest in a larger number of stocks and the gain due to increased AUM easily offsets the loss due to underperformance. Berk and Green (2004) argue that growing too large and then oversimplifying represents profit-maximizing behavior on the part of a fund attempting to extract the economic rents of superior stock-picking skill.2

**Style grid:** The current fund distribution system is based on the style grid, which also encourages funds to purchase many stocks. To fit into the style grid distribution system, a fund must be categorized in one of the style boxes, such as small-cap value. Once categorized, the fund is loath to style drift, which may lead to being excluded from profitable funds or institutional portfolios. Style drift is avoided by purchasing the stocks in the style index associated with the fund’s style box. Thus, in attempting to grow large by fitting into the style grid distribution system, the fund overdiversifies.

**Emphasis on volatility:** Investors, advisors, platform gatekeepers, and consultants place considerable importance on fund-level volatility, which encourages managers to purchase many stocks. Adding stocks reduces a fund’s standard deviation, maximum drawdown, downside risk, or other volatility measures, making the fund more attractive with respect to volatility. But adding many stocks in order to reduce volatility leads to oversimplification.

**Avoidance of criticism, legal problems:** To avoid criticism—or worse yet, legal problems—a fund manager may construct a portfolio to look like an index or other fund portfolios. Fund managers know that investor regret is much higher when a unique or different investment approach does not work than when a traditional approach does not work. This is reinforced by a legal system that at times requires overdiversification or leads to the same result by applying prudent man concepts. These rules and regulatory requirements mean that properly diversified portfolios (i.e., small concentrated portfolios) are more susceptible to regulatory criticism and lawsuits. Managers, often following the advice of legal counsel, oversimplify to avoid such potential problems.

Collectively these industry forces give rise to the closet indexing factory. Without these forces, managers would build small portfolios concentrated in their high-conviction best ideas and virtually all of them would generate positive alphas.

Collectively these industry forces give rise to the closet indexing factory. Without these forces, managers would build small portfolios concentrated in their high-conviction best ideas and virtually all of them would generate positive alphas. Instead, the vast majority of managers oversimplify to the point of underperformance. So, the good news is that equity managers are superior stock pickers, but the bad news is that the closet indexing factory wipes out the potential superior returns by encouraging investment in more bad ideas than good ones.

Industry forces encourage many but not all managers to oversimplify. The managers who resist oversimplifying can be identified in several ways, including the number of stocks held (the fewer the better), AUM (the smaller the better, generally less than $1 billion), best fit index R-squared (the lower the better), active share (the higher the better), style drift (the higher the better), and tracking error (the higher the better). Each of these measures is consistent with the notion that equity managers are most successful when they concentrate on high-conviction stocks. Those that limit oversimplification can produce superior returns.

The bottom line for investors is that it is very difficult to buy and hold active funds that generate superior returns because average fund returns diminish over time, driven by the incentives of the closet indexing factory. A successful truly active fund concentrates on best ideas, doesn’t grow larger than $1 billion in AUM, and doesn’t try to minimize volatility or tracking error, both of which encourage investing in low-conviction stocks.

**ACTIVE EQUITY OPPORTUNITY**

In recent years, however, many truly active funds, which adhere to the prescriptions described above, have underperformed as well. This begs the question of whether the returns-to-skill varies over time. Indeed, considerable evidence shows that stock picking is effective in certain market environments but not in others.

Various studies provide a basis for measuring how favorable or unfavorable a market environment is for stock picking (Gorman et al. 2010; Petajisto 2013; von Reibnitz 2017). They paint a picture in which the returns-to-skill rises in tandem with increased stock return cross-sectional dispersion and skewness, along with greater market volatility.
That is, high levels of cross-sectional and longitudinal volatility are the “fish-are-biting” scenario preferred by stock pickers.

I have proposed “active equity opportunity” or AEO to estimate the impact of market conditions on stock-picking returns (Howard 2018). AEO measures how active the emotional crowds are at driving individual stock-return dispersion.

AEO is estimated using four components, listed here from most to least important: individual stock cross-sectional standard deviation, individual stock cross-sectional skewness, Cboe Volatility Index (VIX), and expected small stock premium. The higher the AEO, the more active the crowds are, so the greater the returns-to-skill; the lower the AEO, the less active the crowds and the lower the returns-to-skill. Active equity managers who build a strategy for harnessing a specific set of return factors prefer a higher level of AEO, because it is more likely their high-conviction picks will outperform. On the other hand, a low AEO foretells a period in which it will be difficult for even the most talented to beat their benchmarks.

Monthly AEOs for December 1998 through December 2018 are presented in figure 4. The average AEO over this time period is 40, which means values greater than 40 signal a better environment for stock picking and lower values signal a worse environment. During this 20-year sample, 1998–2006 and 2008–2010 favored stock picking. Of particular interest is that since 2010, AEO has mostly been below average, declining to a low of 18 in mid-2017. It since rebounded to a high of 41 in February 2018 and was 28 in December 2018.

Based on cross-sectional dispersion going back nearly 50 years, von Reibnitz (2017) concluded that the mid-2017 AEOs were among the lowest in a half-century. So, going on nine years, stock pickers have faced a difficult environment in which to succeed. That is, fish have been biting less and less in recent years.

The shaded areas in figure 4 represent NBER (National Bureau of Economic Research, the official arbitrator of business-cycle turning points) recessions. Although there seems to be a relationship between recessions and higher levels of AEO, von Reibnitz (2017), based on a longer 1972–2013 fund sample, concluded the following:

Overall, these results suggest that periods of elevated dispersion have a positive effect on alpha for the fund sample as a whole, beyond that coming from recessions. Further, the positive relation between fund activeness and performance is driven by return dispersion, as opposed to business cycle fluctuations.

In other words, the impact of AEO on stock picking is largely independent of where we are in the business cycle. Thus, trying to time portfolio changes in active equity fund holdings based on the business cycle alone is ineffective.

### AEO IMPACT ON ACTIVE EQUITY PERFORMANCE PREDICTORS

Howard (2016) presents criteria for identifying funds that are likely to outperform in the future: high-conviction positions, asset size, and tracking error. Over the entire period shown in figure 5, each of these three measures is predictive of future fund performance as evidenced by the positive average return advantage at the average AEO of 40. That is, funds with higher conviction,
smaller AUM, and more tracking error outperform those at the opposite end of each scale. It is also the case that each return advantage varies with AEO, with performance prediction ability increasing as AEO increases and vice versa. Because each of these predictors is driven by stock-picking skill, the observed variation is the result of AEO proxying for returns-to-skill.

As AEO approaches 60 in figure 5, each predictor shows strong intra-decile spreads, ranging from 2.0 percent for high-tracking-error funds versus low, to 5.0 percent for the highest-conviction funds. That is, the highest-decile conviction funds outperform the lowest-decile-conviction funds by 5.0 percent when AEO is 60.

On the other hand, each of the three measures’ predictive ability reverses at some point as AEO drops. At AEO = 20, less convicted, larger AUM, and index-tracking funds outperform their opposite counterparts. So, what works for predicting fund performance when AEO is high is reversed when AEO is low. Again, this makes sense if each measure proxies for stock-picking skill; and as AEO declines, so does returns-to-skill. The 20–60 range in figure 5 captures 85 percent of monthly AEO values over this 20–year period.

MORE-DIRECT MEASURE OF BENEFITS OF SKILL
Fund conviction, AUM, and tracking error discussed above represent proxies for stock-picking skill because they capture the impact on equity fund performance and not, per se, the returns to the stocks held by a fund. A stock rating system, which I refer to as the best ideas of the best funds, provides a more direct measure of the benefits of skill. Howard (2010a) provides details about how the best ideas of the best funds are identified.

Figure 6 presents the impact of AEO on stock alpha of the best ideas of the best funds. All active equity mutual funds that existed in any month during January 1998–May 2017 are included in the sample, resulting in more than 230,000 fund-month observations. AUM and tracking error ranges, the latter measured using benchmark R-squared, are estimated by subtracting decile 10 alpha from decile 1 alpha, reverse of the method used for calculating conviction alpha differences, so that all three measures display a positive AEO relationship. Month-ahead alpha differences are calculated during January 1998–May 2017 and regressed on the beginning of the month AEO in order to estimate the results displayed.
Stock-picking skill benefits are the difference between best idea alpha and worst idea alpha for each value of AEO.

As AEO declines, even truly active funds struggle to beat their benchmarks, confirming the notion that stock picking is effective in certain markets but not in others. However, when AEO is low, it still does not make sense to invest in closet indexers, because although they outperform their truly active counterparts, their average alpha is negative. On the other hand, there may be a benefit to allocating less to stock picking and more to market exposure as AEO declines.4

ENDNOTES

1. Relative portfolio weights are used to rank a manager’s best ideas. See Cohen et. al. (2010) for details.

2. Berk and Green (2004) argue that this is the expected outcome of competition among an unlimited number of investors and a limited number of skilled stock pickers. Because investors are performance chasers, they move their money to high-performing funds, leading to excessive AUM growth that makes it difficult for the fund to consistently pursue the fund’s strategy.

3. Conviction is measured as the portion of the portfolio invested in the top-10 relative-weight stocks (see endnote 1 for how relative weights are calculated). For this exercise, the conviction range displayed is determined by sorting, each month, the fund universe into 10 conviction deciles, then taking the decile 10 conviction (highest conviction) average fund alpha (net of fees and benchmark return) minus the decile 1 (lowest conviction) average fund alpha.


REFERENCES


CONTINUING EDUCATION

To take the CE quiz online, www.investmentsandwealth.org/IWMquiz

© 2019 Investments & Wealth Institute, formerly IMCA. Reprinted with permission. All rights reserved.