When Will Objectivity Enter the Active vs. Passive Debate?

August 19, 2008

The following is in response to Professor Tom Howard's article last week, which was in response to our article the prior week, Luck versus Skill in Active Mutual Funds.

In April 2003, I released a whitepaper entitled "Active vs. Passive: The Debate Continues" that highlighted exactly what has recently been happening in the saga of exchanges about skill versus luck in active management on *Advisor Perspectives*.

My original 2003 paper highlighted that there is a lot of marketing going on by BOTH "passive pundits" and "active advocates" instead of objective research. That paper demonstrated how it isn't objective science if the so called "research" uses essentially the same data (as the Wermers and Howard pieces on Advisor Perspectives, 2008 did) yet comes to opposite conclusions. True objectivity and science are not that forgiving. Marketing, on the other hand, is explicitly designed to spin such data into a story that sounds compelling, or at least plausible, free of the bounds of scientific objectivity.

At the highest conceptual level, Wermers' work is stating something simple, rational and mathematically provable. Essentially, all Wermers is saying is that just because a fund out performed (or under performed) for some period of time, that doesn't mean it was necessarily skill (or the lack thereof), it may have just been luck. He didn't say skill doesn't exist. He didn't say he could tell which funds had skill. His FDR (False Discovery Rate) is in essence a statistical attempt to estimate how much luck exists in the universe of funds, back out that likely statistical luck from the universe, and thus estimate how much skill remains across that universe. He even acknowledges that there were data problems because of the weakness of fund objective consistency over the entire time period.

Wermers is saying that if 10,000 monkeys were picking stocks, that some monkeys would out perform and some would under perform. In the case of monkeys it would be all luck, but in the case of money managers only some of them would be luck. Wermers is saying that it is erroneous to assume 100% of all managers that outperform are skilled, and he is making a statistical attempt to identify how much luck is present in a universe of funds and separating that out to identify how much skill remains in the universe.

Enter Howard, who responds to Wermers work but apparently thinks that all monkeys that outperform are “skilled” and completely discounts the notion that a lucky monkey could exist. All monkeys that out perform are skilled, according to Howard, at least according to his “research” methodology. There does not exist a lucky, unskilled manager in the Howard research. Surprisingly, in response to the Advisor Perspectives article about Wermers, even Ron Surz (whose PODS universes simulates all potential monkeys) chimed in as well with the standard industry bromides that are not measurable or provable (but are marketable) with “People, Process, and Philosophy.” (A.k.a. the “Three Ps” of investment manager research.)

In looking at the Wermers work, as a skeptical scientist, I’m wondering if there is enough statistical significance to necessarily draw all of the conclusions he did, especially since the classification criteria is fairly arbitrary and would introduce a significant amount of error that he - to his credit - objectively acknowledges as a potential weakness (i.e., classification as Aggressive Growth Funds, Growth, Growth & Income, etc.).

But, in looking at Howard’s response to Wermers, I do not find the same level of objectivity; I see marketing spin. Howard’s first exhibit in his Wermers response showed a trend line of average negative alpha of 2%, “trending” upward to almost positive 2%. I wonder what Howard was saying back in the 80’s and early 90’s when his average alpha was negative? He must have been an indexer back then and now has become more enlightened with the upward trending slope that shows all managers combined can beat the entire market combined…hmmm…doesn’t the market have to equal itself? (To be fair, Howard used average and not dollar weighted measurements. So, in theory, the market doesn’t have to equal itself with his method.)

In reality, Howard’s “alpha slope” is a slippery slope indeed. He is comparing all domestic equity funds (with very minor limitations) to the S&P500. Forgetting that there is a large universe of small cap, mid cap, value and growth funds (many of which have index benchmarks that outperformed the S&P500 over the last five years) in his domestic equity universe, he compounds the sleight of hand by attributing to all funds that outperformed the wrong benchmark a “growing alpha skill.”

The way Howard built his statistics, by basing it on the average domestic “any” fund relative to a recently poor performing large cap core benchmark, really shows nothing more than a large cap blend was hard to beat before and has been easier to beat more recently. Shazaam! Money managers are getting smarter and the markets must be getting less efficient!

In our www.fundgrades.com data, if we go back over the last six years where this supposedly surging trend of alpha occurred, and look at the total number of funds that fit into one of our thirteen domestic equity categories (also using Thomson data) we find there were 5,745 funds that had at least 70% in equities, and an R-squared relative to at least one of our thirteen domestic equity benchmarks of at least 67.25% (a correlation coefficient of only .82). Only 618 (10.8%) of those funds fell into large blend as their best fits. Small, micro, mid and even total domestic equities would have all outperformed the S&P500 over this recent period…IT MUST BE SKILL!!! Thus, based on a potentially very weak methodology, all Howard's "alpha trend line" really shows is the S&P500 was a tough benchmark to beat in the 80's and 90's (back when his trend line said active managers were stupid and produced negative alpha) and recently became a benchmark just about any indexer of a sub class could beat (or the now brilliant managers that "have become more skilled at creating alpha").

All of this is caused by comparing his universe of domestic equity funds to a large blend benchmark, and is thus doing nothing other than attributing "alpha skill" to potentially 89.2% of the universe where most of the sub class indexes would have also outperformed his easy benchmark. (The data for just the first three of the six years ending June, 2008, shows a similar relationship with 5,882 funds fitting a domestic equity sub class criteria and only 557 of those fitting into large blend, thus 90.5% of the domestic funds were not large blend. Also, the second three year period of the last six showed 87.0% of the 7,058 funds best fit was something other than large blend. Since we include expenses in our grading routine we do not group share classes of funds and grade each share class individually.)

Ok, so Wermers is somewhat more objective than Howard; and Surz missed a marketing opportunity for the money manager monkeys Surz simulates in his objective PODS universes by reciting the active advocate mantra of the three P's.

I cannot say that alpha doesn't exist. It is provable to exist with mere luck as Wermers work and Surz PODS demonstrate. I cannot say that skill doesn't exist, and statistically one would think someone out there would have it. Intuitively therefore, I assume there is skill.

But, in my clients' interests I have to objectively assess whether the risk of attempting to identify that skill for *potential* out performance is worth the risk of underperforming I have the choice to avoid. I also cannot just accept a marketing bromide that sounds good but is not provable. I need to be able to understand how the odds are stacked to know whether making the bet makes sense. Objectively, it is a bet with at least some of the odds being knowable.

For example, there is essentially zero chance that an index fund will even equal the index and, if it is well managed to minimize tracking error, it is nearly certain that it will under perform the benchmark by something close to the expenses. But with this nearly certain small underperformance of the index fund, also comes *nearly complete certainty that it will not materially under perform the benchmark.* The certainty of the avoidance of the risk of material under performance has value, despite what active advocates might claim. There is value to avoiding that risk.

As a skeptic worried about my clients' best interests, I need to understand whether the payoff for the bet of potential out performance is worth the risk of potential under performance, AND the odds of either occurring. Knowing the odds is the tough part. If you are objective and thus not either a passive pundit or active advocate, you must acknowledge the facts. Index funds are nearly certain to underperform by their expenses and have essentially no chance of either outperforming or materially underperforming their benchmark. This is fact, not marketing, so all of you active advocates should face it. Also, objectively, one would also have to acknowledge that any active bet introduces a risk of potentially materially underperforming, a risk that one can have nearly complete certainty of avoiding by indexing, yet the active bet also introduces a chance of potential out performance that does not exist with the index fund. Passive pundits, face it!

This is not that complicated if you are objective and filter out the marketing noise and pseudo research of the passive pundits and active advocates. And Wermers' and Surz' works go along way toward measuring the luckiness that exists in the historical data; with both (sometimes) objectively acknowledging it isn't necessarily predictive.

But studies like Howard's that compare apples to oranges and assume that ANY out performance is automatically skill, that lucky unskilled managers do not exist and that the money management industry was stupid but is now becoming smart do nothing for clients' interest. But it is good marketing.

While I put no weight on past relative performance, since I have yet to obtain a time machine that would make past performance useful, I couldn't help but test Howard's surging alpha "evidence" in our www.fundgrades.com database.

This is not proof of anything; it is just data. However, it is interesting to see how it contrasts with Howard's "research" on surging alpha skill and declining standard deviation.

We start with a universe of 12,039 funds and share classes that have six years of data ending June of 2008 to focus on Howard's recent alpha surge. Instead of

arbitrary labeling of “domestic anything” funds relative to the S&P500, we apply two criteria to benchmark the funds to help avoid some of the inherent misclassification and arbitrary mislabeling. One is the macro holdings: i.e. under our rules, an equity fund must have at least 70% in equities, a balanced fund must have at least 25% in both bonds and stocks, and a fixed income fund must have at least 70% in bonds, etc. Granted, this will eliminate alpha created by radical asset allocation skill AND luck, but Howard eliminates those too. So give me a pass on that since I objectively admit this is nothing other than useless past data. Second, we benchmark the fund against its best fit of 31 sub asset classes in an attempt, for example, to avoid giving false kudos to a mid cap fund against the S&P500. This isn’t rocket science, just rather basic common sense. It is benchmarking against the best fit style if the macro asset class holdings fit. The fundgrades website lets you grade funds against any of the 31 sub asset classes if you don’t like how our screening criteria benchmarks the fund.

What does the six years of data show? Not surprisingly (unless you are fooled by Howard’s easy bar to beat) 70% of the funds underperformed their best fit sub asset class. So what. Keep in mind this is all funds against 31 sub asset classes. Of the mere 618 funds and share classes whose best fit was the S&P500, 76.7% underperformed. Now there is some surging alpha!

Capturing the “newer fund” alpha supposedly identified by Howard’s “research,” of 12,039 funds in their first three years, 70.43% underperformed their best fit benchmark. And, the second three year period of the entire six years clearly demonstrates Howard’s surging “alpha skill trend” with a broader universe of 15,255 fund share classes having only 66.14% underperforming their best fit benchmark.

On the risk side, many funds have less risk than their benchmarks. This is to be expected since no benchmark includes cash and nearly every fund has some cash tempering at least a micronic amount of volatility with our precise measures. In fact, over the six year period, 54% of all funds had less standard deviation than their benchmark (kudos go to the 59.2% of the large blend funds that had less standard deviation for holding a little cash to manage redemptions…What skill!). It is interesting that the average standard deviation was 100.25% of the benchmark standard deviation and the median was 99.32%. This is not statistically meaningful. But, as one would anticipate, all funds together, when classified appropriately against a broad universe of benchmarks, on average have about market risk.

Is there really anything surprising in this data? Not yet. Also not surprisingly, those funds that outperformed their benchmark had more risk relative to their benchmark, and those that underperformed had less risk. What a shocker here! Of the funds that outperformed their benchmark, they averaged 109% of the

standard deviation of their benchmark and 59% of those funds had more standard deviation versus the 46% of all funds that had more risk. Those that underperformed averaged 96% of their benchmark's standard deviation with 60% of those under performers having less risk. Those funds that did beat their benchmark had two and a half times more risk (26% versus 11%) of having a standard deviation greater than 115% of the benchmark standard deviation than the under performers (our criteria for a relative risk grade of "F"). Still no surprise here.

You might have skill. Or, you might just think you do. There is a big difference. You might assume that all excess results are caused by skill. You might assume luck is a figment of every winning gambler's imagination. Winning gamblers also often falsely attribute their luck to skill, a betting system, or some other secret method, much like active advocates in money management.

I am fairly confident that skill exists, although I do not have very good evidence for it. For now, it remains an unprovable intuition. I suspect that such skill, should it exist, is somewhat rarer than the percentage of funds that happen to outperform because of luck, as I'm equally confident that luck also exists. The existence of luck is a lot easier to mathematically prove, but like Wermers' work one cannot tell which funds were just lucky.

Beyond this, I have to think about the bet being made relative to odds that are knowable in my clients' interests. I don't know the odds of skill existing, how common it might be, or how much value can be obtained. To an objective scientist, the attempt to try to identify skill in historical data samples is not sufficiently provable when weighed against the odds that are knowable and how skilled one must be to make up for risks introduced of material under performance in an attempt to out perform. Marketers won't sell this because real statistics are not marketable. However, your clients can benefit by keeping your level of skepticism high enough to avoid being fooled.