

Attribution Challenges

Chapter 28 in the book *Noble Challenges: Investment Management in a Changing Milieu: a Reference Book for Finance Professionals*

Farragut, Jones & Lawrence

Ron Surz, PPCA, Ron@PPCA-Inc.com, 949/488-8339

Where misunderstanding serves others as an advantage, one is helpless to make oneself understood. Lionel Trilling, author (1905-1975)

Investment performance attribution determines why performance is good or bad, singling out what worked and didn't work. In the hierarchy of the search for manager talent, attribution ranks way above evaluation and is far more forward looking. The skillful can stumble and the unskilled can get lucky. We want to know the difference, and importantly we want to know how mistakes are being corrected and what proficiencies are being groomed. Some of the "Noble Challenges", the title of this book, in performance attribution are as follows:

- ✚ Differentiating not just between luck and skill, but between style, luck and skill. The relatively recent awareness of the importance of style goes a long way toward identifying true skill. It's easy to confuse style with skill but extremely difficult to make good decisions in the face of this mistake. Buying skill, not style, is akin to buying alpha not beta.
- ✚ Dealing with the active-passive trade-off. Use all the active managers you can find who have demonstrated skill, and complete the portfolio with passive investments to fill in parts of the market where talent has not been found. It should only matter that the manager adds value, not that value is added in a particular style box.
- ✚ Putting style boxes to good use. Insisting that a manager fit in a box is absurd. We miss too much talent that way, and end up with mostly index huggers. No offence to index huggers, but most skillful managers can't deliver under the constraint of living in a box. Rather, investment managers should be evaluated against custom style blends that reflect their people, process and philosophy. The due diligence process involves two central questions: (1) Do we like what this manager does? and (2) Does (s)he do it well? The answer to the first question shouldn't revolve around style boxes, rather blended boxes should be used to answer the second question.

- ✚ Regaining control of the assets. Financial consultants and institutional investors have relinquished control of their assets to investment managers, primarily through terrific sales and account management that manipulates the client in various ways, including creative performance reporting. Granted, investment managers are the smarter lot, but the assets are not theirs.
- ✚ Compensating investment professionals for delivering value added. Specifically, attribution determines which analysts are succeeding and failing, as well as the effects of the portfolio managers on overall performance. Knowing which players own which pieces of the performance puzzle, as well as who is contributing and who is not, is important for professional retention and morale. Compensation should be tied to contribution. Unfortunately, bonuses are typically based on ad hoc rules of thumb that ultimately make them fully discretionary. This creates a dynamic that rewards dominant personalities and pointy-haired bosses rather than talent.

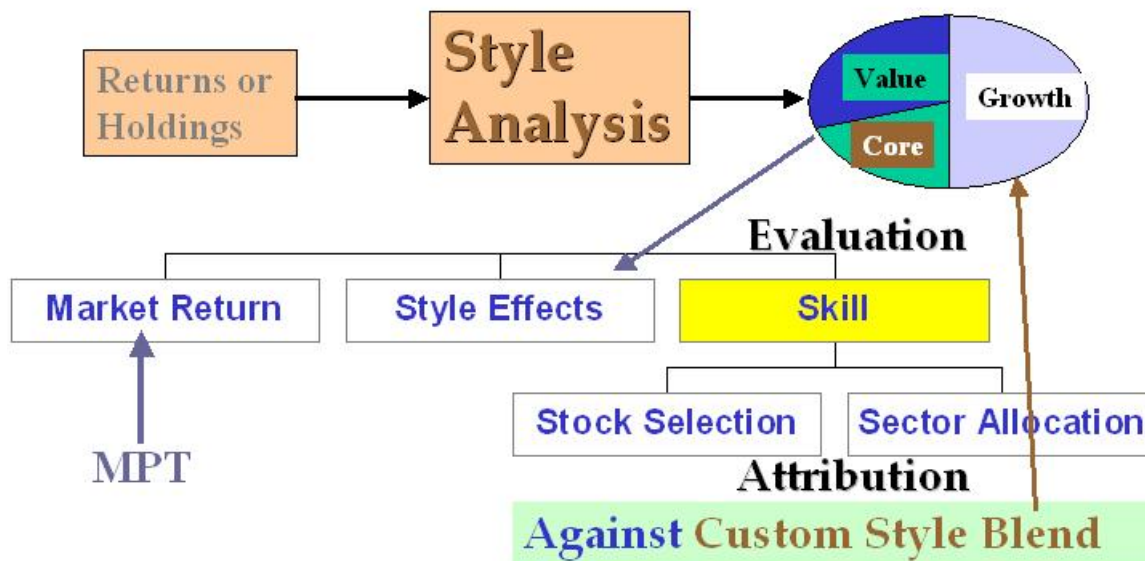
To address these challenges, performance attribution must take away all of the hiding places that managers have used for the past 40 years, which is the relatively short history of this profession. Fair is fair, and it's time for investors to get the real story; all of the cards in this poker game need to be dealt up. This chapter goes beyond the valiant efforts of the CFA Institute and its Global Investment Performance Standards (GIPs). The GIPs standards focus primarily on accurate measurements and reporting, but even the most accurate measurements can be misinterpreted when compared to faulty benchmarks, regardless of the intent. And once the benchmark is wrong, all of the analytics, including attribution, are wrong. It's the old garbage-in-garage-out, or GIGO, problem. So this chapter starts with a discussion of accurate benchmarking and then shows how solid attribution analysis uses the best benchmarks to remove the hiding places. The investor sees the real story, warts and all. What's good for the goose is good for the gander. Solid attribution analysis is also good for the investment management profession, especially when it comes to discretionary compensation, namely bonuses.

Getting the Benchmark Right

The central challenge in identifying investment talent is knowing who is winning and who is losing. We've failed to meet this challenge because we usually get the benchmark wrong. It's like evaluating Tiger Woods as a bowler, or the old joke about yesterday's football scores: 20 to 13, 34 to 5, etc. Style analysis goes a long way toward correcting this problem but we're currently stuck using flawed executions of an excellent idea.

This total performance evaluation and attribution picture is shown in Exhibit 1.

Exhibit 1: Sources of Return



The past couple of decades have taught us that Modern Portfolio Theory (MPT) does not work when it comes to evaluating investment performance. Specifically, the Capital Asset Pricing Model (CAPM) does not work when value investing is in favor because low beta stocks outperform high beta stocks in this environment. "Beta is dead" heralded the introduction of Arbitrage Pricing Theory (APT), and APT has morphed into our current use of style analysis. Beta is dead because style effects are so strong. Of course broad market effects remain

paramount in determining portfolio returns, but not beta adjusted market effects. In fact the largest component of return is usually the market return, followed by the style effect, and then a far distant third component is what might be attributable to skill. Note that while alpha, or skill, can be estimated using either holdings or returns, holdings are required to complete the picture with the components of skill, or attribution analysis. Note also that it is important that style be taken into account in both performance evaluation and performance attribution. As a practical matter, the search for skill ought to begin at the macro level with managers whose performance is good. Then due diligence can proceed with an understanding of the people, process and philosophy that produced the good performance. And then last, but not least, performance attribution confirms that the sources of this good performance are consistent with the people, process and philosophy. Throughout this process we keep in mind that the resultant decisions are all about the future, even though we use the past as a guide.

As shown in the exhibit, the benchmark used in this quest must be both market and style driven, and customization of the style component is important. A benchmark establishes a goal for the investment manager. A reasonable goal is to earn a return that exceeds a low-cost, passive implementation of the manager's investment approach, because the investor always has the choice of active or passive management. The relatively recent introduction of style indexes helps, but these need to be employed wisely, using blending rather than off-the-shelf style indexes. Before style indexes were developed, there was wide acceptance and support for the concept of a "normal portfolio," which is a customized list of stocks with their neutral weights. "Normals" were intended to capture the essence of the people, process, and philosophy behind an investment product. However, only a couple of consulting firms were any good at constructing these custom benchmarks. Today we can approximate these "designer benchmarks" with style analysis, sometimes called "the poor man's normals." While style analysis may not be as comprehensive as the original idea of normal portfolios, it at least makes it possible for many firms to now partake in this custom blending of style

indexes. Style analysis can be conducted with returns or holdings. Both approaches are designed to identify a style blend that—like normals—captures the people, process, and philosophy of the investment product.

Whether the returns or holdings approach to style analysis is used, the starting point is defining investment styles. The classification of stocks into styles leads to style indexes, which are akin to sector indexes such as technology or energy. It's important to recognize the distinction between indexes and benchmarks. Indexes are barometers of price changes in segments of the market. Benchmarks are passive alternatives to active management. Historically, common practice has been to use indexes as benchmarks, which works fine for index huggers, but there are many skillful managers who work best without the comfort of hugs. Style analyses have shown that most managers are best characterized as blends of styles. As a practical matter, we are no worse off with style blends, as the old practice is considered in the solution so there's always the possibility that the best "blend" is a single index. Managers feel compelled to complete the sentence "I manage to the _____ index". This is counterproductive and a convenience that helps no one unless again the manager is an index hugger. Also, the sentence is frequently completed with the index *du jour* for the RFP.

One form of style analysis is returns-based style analysis (RBSA). RBSA regresses a manager's returns against a family of style indexes to determine the combination of indexes that best tracks the manager's performance. The interpretation of the "fit" is that the manager is employing this "effective" style mix because performance could be approximately replicated with this passive blend. Another approach, called holdings-based style analysis (HBSA), examines the stocks actually held in the investment portfolio and maps these into styles at points in time. Once a sufficient history of these holdings-based snapshots is developed, an estimate of the manager's average style profile can be developed and used as the custom benchmark. Note that HBSA, like normal portfolios, starts at the individual security level and

that both normal portfolios and holdings-based style analysis examine the history of holdings. The departure occurs at the blending. Normal portfolios blend stocks to create a portfolio profile that is consistent with investment philosophy, whereas HBSA makes an inference from the pattern of point-in-time style profiles and translates the investment philosophy into style.

The choice between RBSA and HBSA is complicated and involves several considerations. Although RBSA has gained popularity, this doesn't necessarily mean that it's the best choice. The major trade-off between the two approaches is ease of use (RBSA) versus accuracy and ease of understanding (HBSA). RBSA has become a commodity that is quickly available and operated with a few points-and-clicks. Some websites offer free RBSA for a wide range of investment firms and products. Find the product, click on it, and out comes a style profile. Offsetting this ease of use is the potential for error. RBSA uses sophisticated regression analysis to do its job. As in any statistical process, data problems can go undetected and unrecognized, leading to faulty inferences. One such problem is multicollinearity, which exists when the style indexes used in the regression overlap in membership. Multicollinearity invalidates the regression and usually produces spurious results. The user of RBSA must trust the "black box," because the regression can't explain why that particular blend is the best solution. In his article that introduced RBSA, Nobel laureate Dr. William Sharpe [1988] set forth recommendations for the style indexes used in RBSA, known as the "style palette":

"It is desirable that the selected asset classes be:

- Mutually exclusive (no class should overlap with another)
- Exhaustive (all securities should fit in the set of asset classes)
- Investable (it should be possible to replicate the return of each class at relatively low cost)
- Macro-consistent (the performance of the entire set should be replicable with some combination of asset classes)."

The mutually exclusive criterion addresses the multicollinearity problem, and the other criteria provide solid regressors for the style match. The only indexes that currently meet all of these

criteria are provided by Morningstar and Surz. Morningstar is available for U.S. stocks, while Surz indexes are provided for U.S., international, and global stock markets. Using indexes that do not meet Dr. Sharpe's criteria is like using low octane fuel in your high-performance car. See Picerno [2003 and 2006] for an extensive discussion of a proper style palette.

Holdings-based style analysis (HBSA) provides an alternative to RBSA. The major benefits of HBSA are that the analyst can both observe the classification of every stock in the portfolio as well as question these classifications. This results in total transparency and understanding, but at a cost of additional operational complexity. HBSA requires more information than RBSA; that is, it needs individual security holdings at various points in time, rather than returns. Since these holdings are generally not available on the Internet, as returns are, the holdings must be fed into the analysis system through some means other than point-and-click. This additional work, sometimes called "throughput," is not that difficult and is well worth the effort. Like RBSA, HBSA also requires that stocks be classified into style groups, or indexes. Dr. Sharpe's criteria work for both RBSA and HBSA; i.e., for consistency purposes, the same "palette" should be used for both types of style analysis. Note that the "mutually exclusive" and "exhaustive" criteria are particularly important in HBSA as it is highly desirable to have stocks in only one style group and to classify all stocks.

In certain circumstances, deciding between RBSA and HBSA is really a matter of Hobson's choice. When holdings data is difficult to obtain, as can be the case with some mutual funds and unregistered investment products such as hedge funds, or when derivatives are used in the portfolio, RBSA is simply the only choice. RBSA can also be used to calculate information ratios, which are style-adjusted return-to-risk measures. Some researchers are finding persistence in information ratios, so they should be used as a first cut for identifying skill. Similarly, when it is necessary to detect style drift or to fully understand the portfolio's actual holdings, HBSA is the only choice. Holdings are also required for performance attribution analysis that is focused on differentiating skill from luck and style--an important distinction.

This level of analysis must use holdings because performance must be decomposed into stock selection and sector allocation. Returns cannot make this distinction.

Custom benchmarks developed through either RBSA or HBSA solve the GIGO problem, but statisticians estimate that it takes decades to develop confidence in a manager's success at beating the benchmark, even one that is customized. This is because when custom benchmarks are used, the hypothesis test "Performance is good" is conducted across time. An alternative is to perform this test in the cross-section of other active managers, which is the role of peer group comparisons. We'll discuss this alternative as it is integrated into performance attribution, discussed in the next section.

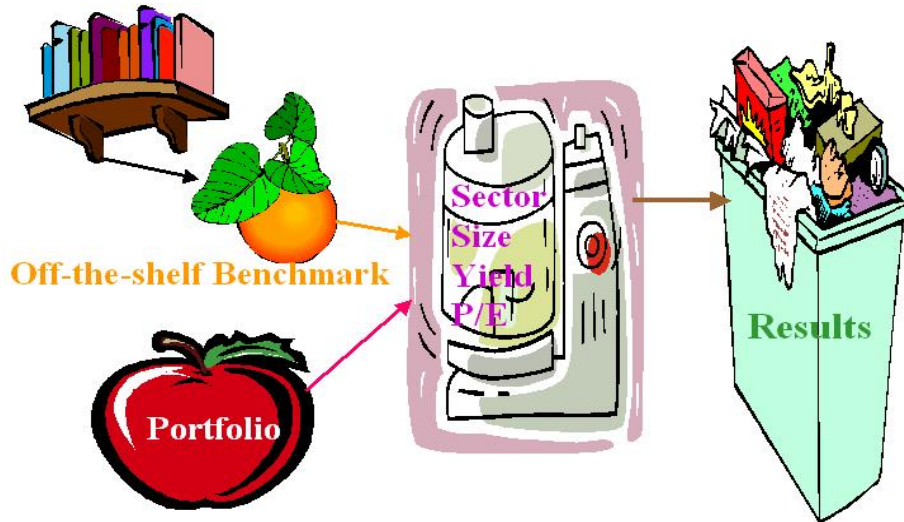
Performance Attribution

There has been an evolution in performance attribution. Much of the attribution analyses that had been used until recently were developed back in the 1980s, when we were only beginning to understand that there was more to life than MPT, and had moved on to APT. We knew back then that characteristics like capitalization, price/earnings ratio, and dividend yield mattered, but hadn't figured out how to best integrate these factors into attribution analysis.

Consequently, we wrote "slicers and dicers" that segmented the portfolio and the benchmark by whatever characteristic we liked. Want to see how the segment of your portfolio with high P/Es fared against the comparable segment of stocks in the S&P500? No problem. Just draw the P/E line wherever you want, and voila. The problems with these old approaches are standardization and benchmark inflexibility. If you draw the P/E line at 15 and I draw it at 20, we'll each get different insights. Also, as described above, we'd like to use a custom style blend as the benchmark, but we can't do so with the 1980s technology because it doesn't provide the ability to customize the benchmark as a blend of indexes. So with the old

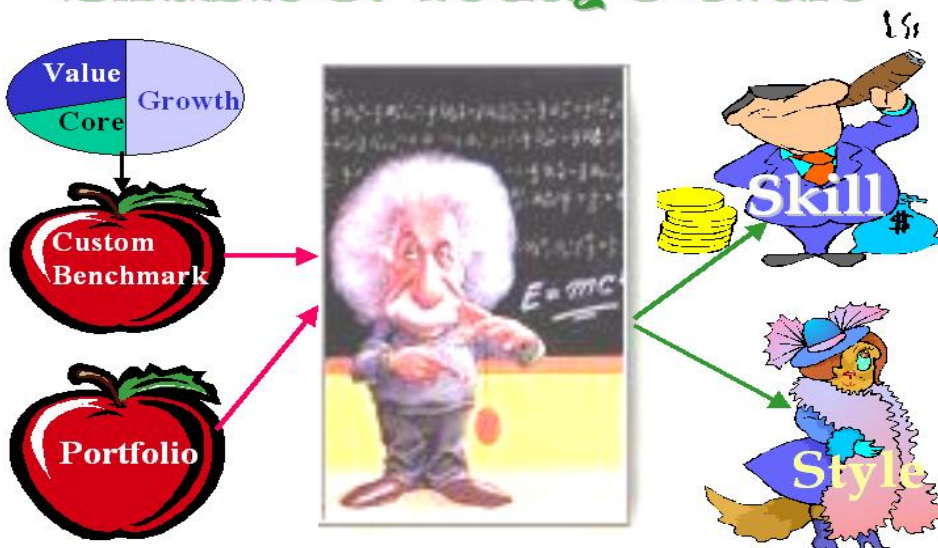
technology we can peel the apple like an orange or slice the orange like an apple, but all we've got to show for it is fruit salad. Exhibit 2 describes this mistake.

Exhibit 2: 1980's Slicer & Dicer



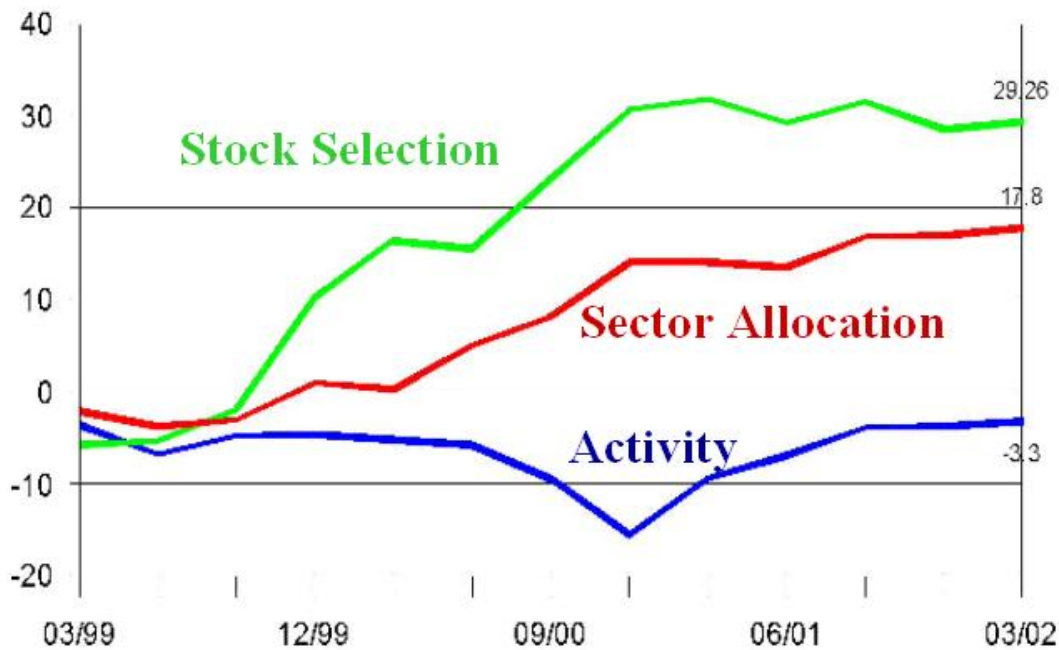
By contrast, contemporary technologies encourage the use of custom style-blended benchmarks, and standardize style definitions so there is comparability across managers. In this way a manager's stock selection and sector allocation skills are not confused with his style. Exhibit 3 summarizes this evolution.

Exhibit 3: Today's Graft



In the search for skill, we look for persistence in the reason(s) for good performance, and for confirmation of the People, Process, and Philosophy. The following Exhibit 4 shows a real life manager who has consistently added value through stock selection, although the amount of value added has slowed somewhat in the recent past. This particular manager is a bottom up stock picker, and the attribution analysis confirms his skill in this endeavor. Sector allocation has also added some value, which is consistent with bottom up stock picking. Only Trading Activity has had a modest negative effect on performance. Trading Activity measures the intra-period effects on performance of transactions executed during the period. If this manager were looking for ways to improve performance a place to start would be the trading desk.

Exhibit 4: Sources of Value Added -- Persistence

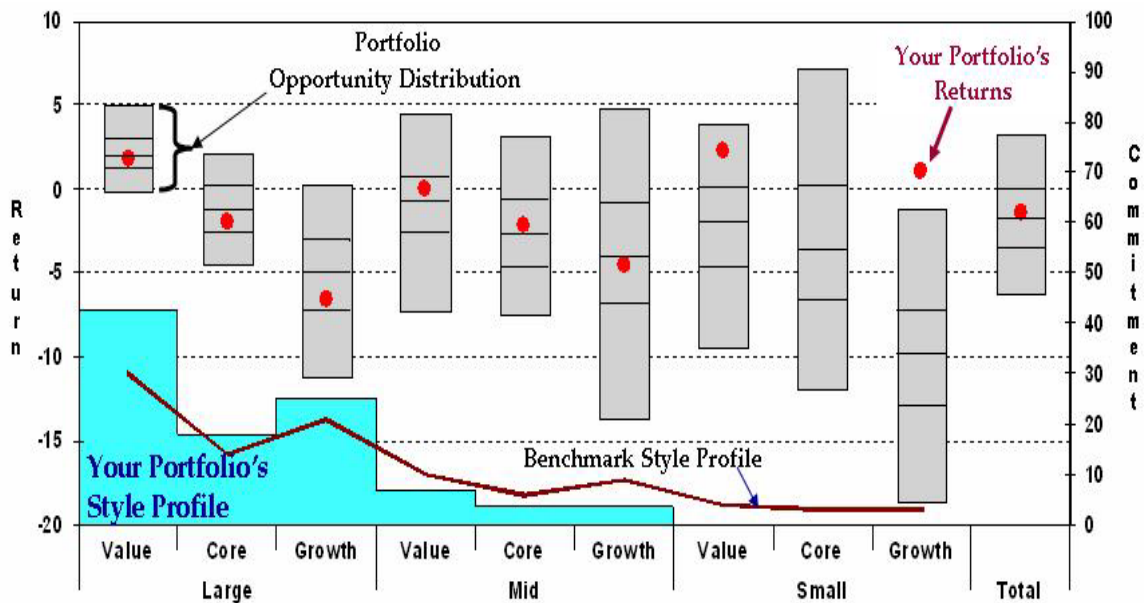


These relatively new tools give the professional evaluator the insights needed to determine whether good performance is likely to continue into the future. To rely on this analysis we need to be confident that the benchmark is correct. This is achieved by careful examination of

the style profile of the manager as well as the period by period attributions against a custom style profile. Let's take a closer look at these style attribution details.

Exhibit 5 shows a portfolio's composition and performance broken out by style. There are 2 scales on this graph – the scale on the left is for return and applies to the dots and floating bars, and the scale on the right is for allocation and applies to the shaded area and line on the bottom of the exhibit.

Exhibit 5: Style Attribution



Let's start with allocations. The shaded area shows the portfolio's allocations to styles, and the line shows the custom benchmark's allocations. In this case we see that a style bet has been made in overweighting large value companies while underweighting mid and small size companies. The floating bars in the exhibit tell us that this bet paid off because large value was in favor – the opportunities in this style, while narrow, tended to exceed the opportunities in other styles for this period. Then looking at the dots in the exhibit we see how well or poorly this portfolio performed relative to the opportunities in each style. Note that performance in

large companies was near or below median while performance in mid and small companies tends to be above median. This manager did a better job of selecting smaller companies. We can run this analysis for any time period we choose, and examine the details by style and by individual stock.

The opportunity sets shown in the exhibit are special peer groups. Traditional peer groups come with a boatload of biases that render them totally useless. But we can use the notion of a peer group to solve the waiting time problem mentioned above with custom benchmarks, namely we need to wait decades for regression analyses against custom benchmarks to produce statistically significant results. This is because we are testing the hypothesis “performance is good” across time. To solve this waiting problem we structure the test in the cross-section of all possible portfolios that could have been held when selecting stocks in the benchmark. It’s classical statistics using Monte Carlo simulations. The hypothesis is tested by comparing the actual performance outcome to all of the possible outcomes, so if the observed return exceeds 90% of the possible returns we say the result is a statistically significant success at the 90% confidence level.

When the evaluator is confident that style attribution is using an accurate style benchmark, this attribution can be extended to sector and country attribution, which can be used for analyst and manager compensation, in addition to taking away all of the old hiding places. Let’s focus on the positive application of manager compensation.

Attribution for Analyst and Manager Incentive Compensation

A fair incentive formula should incorporate the following elements:

1. **Responsibility.** What part of the overall investment process does this employee influence?

2. *Expectation*. Some measure of the employee's performance must be agreed upon, and a barometer of what is expected must be established.
3. *Impact*. The significance of the employee's contribution, beyond expectations, should be specified as an independent variable in the bonus formula. "If you add x you can expect a bonus of y ."

Current compensation schemes use indexes and/or peer groups to address these three requirements, but there is now science that combines the better characteristics of benchmarks with those of peer groups to create a superior new compensation formula. Monte Carlo simulations serve as reference points for determining not only the success or failure of the professional, but also the significance of this success or failure. Here are the details.

Responsibility for a research analyst is the list of companies from which recommendations can be made – the eligible list. Ideally, neutral weights are assigned to the list to create a normal portfolio for the analyst. Responsibility for a portfolio manager is the roll-up of analyst responsibilities, weighted by neutral allocations to each analyst's economic sector. In other words, custom benchmarks are created for each analyst and aggregated to create a custom manager benchmark. An alternative approach is to begin with a custom benchmark for the manager and to disaggregate it into economic sectors, like utilities and technology. The returns on these custom benchmarks represent expectations that need to be exceeded to earn a bonus. It's extremely important that these benchmarks are customized, or at least that the benefits of customization have been considered. Processes for establishing custom benchmarks are presented above.

The measure of actual performance for a portfolio manager is straightforward; it's the return on the total portfolio. Analyst results can be tracked in two ways, as the actual return on the analyst's economic sector and as a paper portfolio return representing the

analyst's recommendations. The difference between these two portfolios is of course a reflection of the analyst's ability to sell recommendations to the manager.

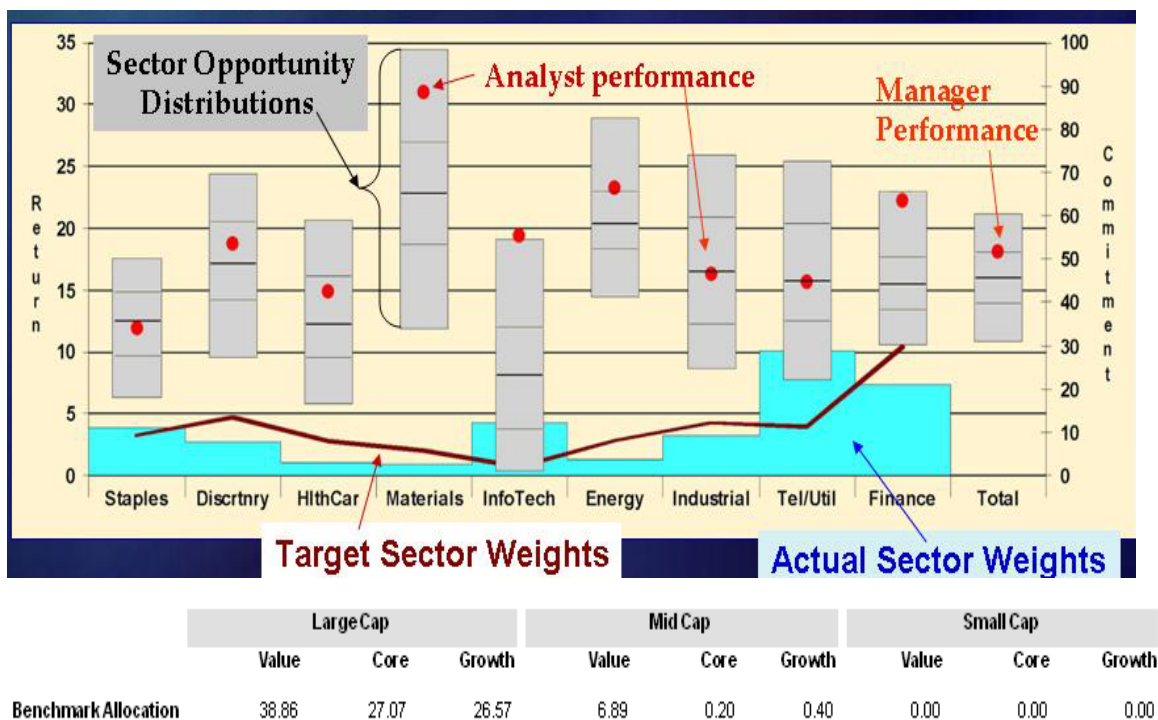
At this point we have sufficient data to calculate attribution measures for both managers and analysts, conducted at the security level, i.e. holdings-based attribution analysis. As a result we know how much each professional has succeeded and why. We're now ready to determine impact, or the significance of success. Current common practice is to set the same threshold for everyone, like 2% above the benchmark. This is neither fair nor realistic. 2% in some environments may be extremely easy or hard, plus some economic sectors are more volatile than others so 2% is easier for some analysts than others. What we really need is some sense of the opportunities for each professional, a custom peer group for each. This is a great concept but one that cannot be implemented in practice. Peer groups cannot work despite frequent attempts to make them work.

Enter Monte Carlo simulations. Virtual peer groups are created from each custom benchmark by creating a reasonable representation of all possible portfolios that the professional might have held. This requires the specification of portfolio construction rules, like industry and security constraints, and the number of stocks. Impact is then determined as the ranking within this virtual peer group. This is fair because the professional's actual decisions are compared to all of the decisions that might have been made instead: the cost or benefit of untaken paths. Value added is placed into perspective as its statistical significance within this framework. The ranking in a Monte Carlo opportunity set is the statistical distance of actual performance away from expectation.

An example will help to clarify this new approach. The following Exhibit 6 shows a one year attribution for an investment product with 9 sector analysts. In the far right bar the

portfolio manager is ranked against a total custom opportunity set. The line running along the bottom of the exhibit is the benchmark, or neutral, allocation to each sector and the shaded area is the actual average allocation. We can see sector by sector where this manager has over- and under- weighted sectors, as well as the ranking of each sector's performance. These sector rankings belong primarily to the responsible analysts, and, as suggested above, the analysts' paper portfolios could also be located in this framework. In this example we have started with a total custom benchmark, and disaggregated it into sectors, but you could just as easily take the opposite approach of aggregating custom sector benchmarks.

Exhibit 6: Example of a New Compensation Approach



Here's how to read the exhibit. The Materials analyst has delivered a 31% return, dwarfing the InfoTech return of 19% and Finance return of 22%. But the virtual peer group rankings of the InfoTech and Finance analysts are better than that of the Materials analyst, and the allocations to InfoTech and Finance are higher than that to Materials. The Energy analyst performed in the top quartile, even though performance

was only 2% above benchmark. The performance of the remaining 5 analysts is in line with expectations.

Let's develop a sample bonus formula to see how each professional would be compensated. Using the convention that the top of the distribution is a rank of 100, a bonus could be specified as 1% of additional pay for each 1% that the ranking is above 55%. Here's how everyone would be compensated on this basis, or at least where the bonus discussions would start:

Sample Bonus Formula Using the New Approach

Consumer Discretionary Analyst	2%
Healthcare Analyst	15
Materials Analyst	35
Information Technology Analyst	45
Energy Analyst	20
Financials Analysts	40
All Other Analysts	None
Portfolio Manager(s)	20

This new virtual peer group approach is fair, accurate, timely, objective and easy. It's fair because compensation is tied to the significance of achievement, instead of relying on some arbitrary thresholds. It's accurate if good custom benchmarks are employed, which I recommend. It's timely because simulations can be run anytime during the year on demand. It's objective because the rules are straightforward – a 1% improvement in ranking earns an "x" percent bonus. And it's easy to understand and monitor.

Conclusion

The search for investment manager talent puts a lot of emphasis on recent past performance. Unfortunately, in evaluating past performance, style is routinely confused with skill. The retirement industry is particularly notable for making this mistake as a group. Perhaps it's because the test of fiduciary prudence rests partially with what other fiduciaries have done.

Solid performance attribution is very careful about getting the benchmark right so we avoid the GIGO problem. It also drills down into style, sector and country details so there is no place for the manager to hide. Some investment management marketing and relationship personnel benefit from the current obfuscation. They are among the "others" in the quote at the beginning of this article who are served by misunderstanding. These folks are some of the best spin meisters in the world, and are well compensated for their talents. Like the affably charismatic lawyers on the TV show *Boston Legal*, good spin meisters always win even when they should not. But we can stop the confusion and manipulation, and replace it with accuracy and fairness. "Live and let live" is a nice mantra, but not if it betrays client trust or breaches fiduciary duty. Fiduciaries have the duties of both prudence and care. Prudence argues for common practice, but care trumps prudence and requires best practice.

REFERENCES

Picerno, James. "In the Quest to Identify Investment Skill, Ron Surz Claims He Has the Better Mousetrap." *Bloomberg Wealth Manager*, June 2003, pp 80-82.

_____. "A Style All His Own." *Wealth Manager*, September 2006, pp 64-66.

Sharpe, William F. "Determining a Fund's Effective Asset Mix." *Investment Management Review*, December 1988, pp 59-69.