The New "TRUST but VERIFY"

By Ron Surz

"Today's money managers say they compete with other money managers by generating the highest alpha. They denigrate the role of marketing. Yet each money manager has ready stories about other money managers with low alphas who snatched clients through clever marketing."
—Professor Meir Statman, Santa Clara University, Statman, 2004

The old "Trust but Verify" was all about the accuracy of performance results reported by money managers. We even distrusted custodians, so consulting firms performed custodial audits. The pressure worked. Today we have performance reporting standards promulgated by the CFA Institute, enlightened custodians, and a raft of computer checks and balances. So you'd think the past is behind us, but it would appear not. Service providers, especially performance attribution vendors, sell real time, or at least daily, product because it is more "robust." The pitch is that attribution is more accurate because the manager's calculated performance is more accurate. But do we really need to calculate manager performance over and over to achieve accuracy? Today we get performance numbers from managers, custodians, consultants, service bureaus, performance verifiers, etc. etc. Want to see a "robust" return? Run a daily attribution system and see if it produces anything different than the myriad other sources; if it does there's a good chance something is wrong with the attribution system. We're wasting our time, money and energy on the old "Trust but verify" and missing an important fact in the process: if the benchmark is wrong all of the analytics are wrong. Most attribution systems do not allow careful benchmark specification, unless the manager is an index hugger, which is generally pretty uninteresting. Performance attribution tells us why the manager has succeeded or failed relative to a benchmark, so if the benchmark is wrong we are completely misled. We pay a lot of money for flawed information. Even worse, we make bad decisions. It's the old garbage-in, garbage-out.

By contrast, the new “Trust but verify” is all about the spin. Professional sales people can turn mold into gold right before our eyes. The most accurate performance results are presented in the most favorable light, even when the money manager has failed. This subterfuge has gone on for so long that it is commonly accepted, hence the Meir Statman quote that introduces this article. We frequently hear “laughable” sayings like “I have never met a money manager who has performed below median” or “Every manager wins against the right benchmark.” But the clients aren’t laughing. Consultants need to take back control of the due diligence process by getting serious about the two key questions:

- What does this manager do?
- Does he/she do it well?

Both questions have everything to do with accurate benchmarking, and absolutely nothing to do with verifying reported manager performance, although answering these questions can reveal deceitful reporting as well. Bernie Madoff’s scam was there for all to see if they seriously addressed the two due diligence questions, but only a few did. We can trust the accuracy of reported performance, but not the story that is told about that performance. If the performance is cooked, we’ll discover that when we perform real due diligence, but we need to do it right. Complacency and laziness have allowed us to be fooled, and clients are demanding better. The active-passive debate has much to do with our inability to differentiate winners from losers; if you can’t tell the difference, you really should go passive – be honest now.

Stepping up the Due Diligence Process

Indexes and peer groups do not work for addressing the two key due diligence questions. In identifying what the manager does, most managers do not live in a style box; although most will agree to be benchmarked against whatever it takes to get the account. Style boxes are convenient shorthand that has gotten out of hand. As for the second question, performing well against a “representative” peer group tells us nothing. Peer groups are loaded with biases and sales people can always find a peer group that makes them look good, provided by a reputable source and that has a name that sounds like what they do. Don’t like your peer group ranking? Find another peer group. Clever sales people find clever ways to turn mold into gold.

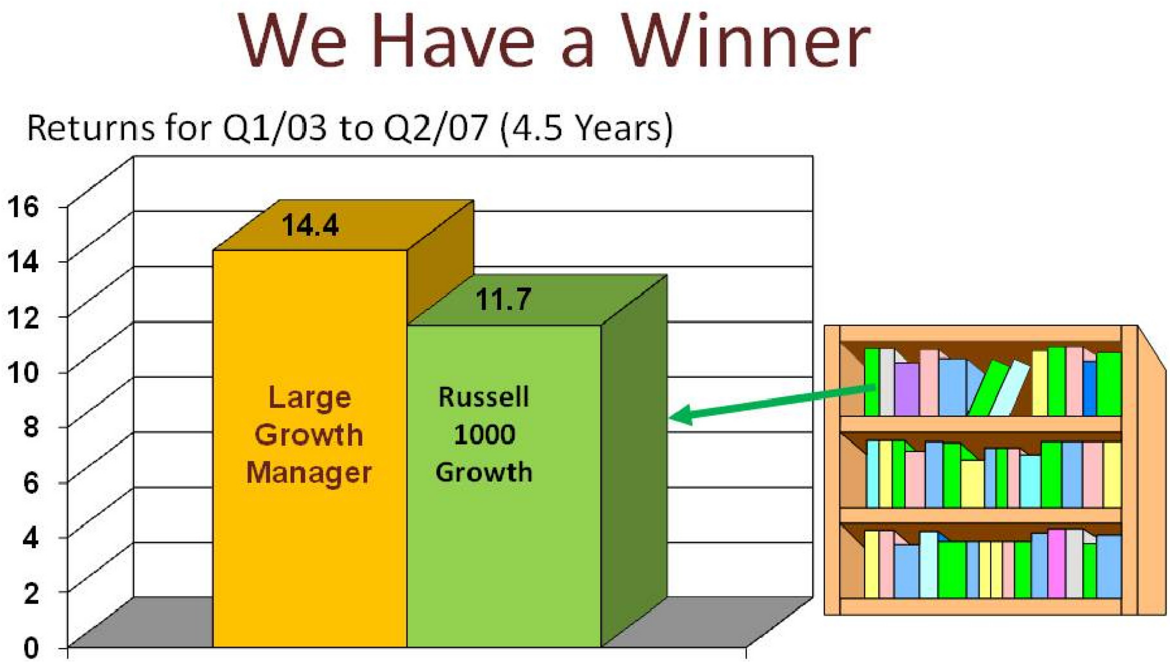
To properly answer the first due diligence question – What does this manager do? – We need tools that capture the people, process and philosophy (the three Ps) of the management firm. Importantly, we must allow for liberated managers who are not index huggers. Nothing against index huggers, except that the consulting industry seems to want to treat everyone like an index hugger, perhaps as a convenience or perhaps because they have swallowed the poppycock that tracking error is risk. One of the tools we have available to us is style analysis, in particular analysis that advocates custom blends of styles. We can use either performance or holdings to see how the three Ps manifest themselves. The best style analysis uses the best style indexes, and it has been shown that the best style indexes are mutually exclusive (no stock is in more than one index) and exhaustive (the collection of indexes spans the entire market).

With a custom benchmark in hand we can address the second question – Does the manager perform well? At this point we always have the choice of active or passive management. Even hedge fund managers can be inexpensively replicated. This second due diligence question is best addressed with hypothesis testing. We test the hypothesis that the manager can and has outperformed. But relative to what? Relative to all of the possible implementations of his approach. Hypothesis tests compare the actual outcome to all of the possible outcomes. If the manager’s performance is in the top decile of all the possibilities, he has had significant success. Note that this approach is bias free and customized to the manager, unlike traditional peer groups. Today’s computer technology makes it possible to simulate all, or a reasonably good sample of all, of the paths the manager could have taken through time, creating portfolios from stocks in the custom benchmark.

Does It Work?

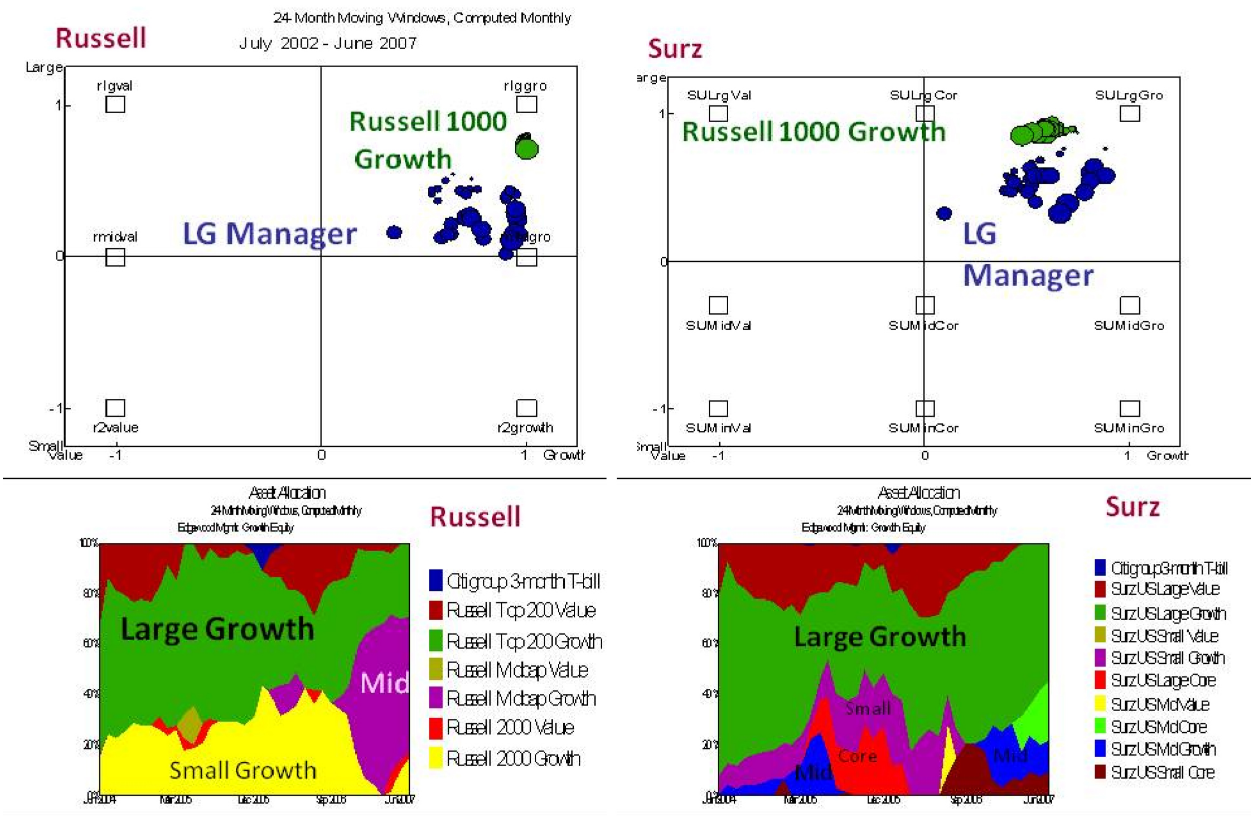
In May of 2005 I criticized Don Trone, then founder and president of the Center for Fiduciary Studies, for his due diligence criteria, which made extensive use of indexes and peer groups. Don responded with a challenge and an acknowledgement that "*Our due diligence criteria is designed for one purpose: to define the minimum due diligence process that should be applied by an investment fiduciary - it is not intended to be an industry best practice - that space is reserved for you!!.*" Don's challenge was to create competing multi-manager portfolios. Can best practices actually perform better than common practices? I won, & have the congratulatory e-mail from Don to prove it. By the way, it's been said that "Best practices that are not common practices are simply someone's unpopular opinion." How does that thought strike you?

Here's an up close look at best practices in action — a case study. In the exhibit below, We Have a Winner, a large cap growth manager has outperformed the Russell 1000 Growth index by 270 basis points per year. Would you hire this manager?



Before you decide, let's take a closer look. The exhibit on page 10 shows two returns-based style analyses, one using Russell indexes and another using Surz Style Pure® indexes. The Surz indexes are mutually exclusive and exhaustive while the Russell indexes are not. The two analyses agree that the manager is not pure large growth, although large growth is the predominant allocation. The analyses disagree however on the exposure to small companies, with Russell showing a 30% exposure and Surz showing very little. A closer examination of the holdings in the portfolio reveals that the manager's smallest holding had been $6 Billion, hardly a small company. Indexes that are not mutually exclusive and exhaustive make poor style palettes for returns-based style analysis.

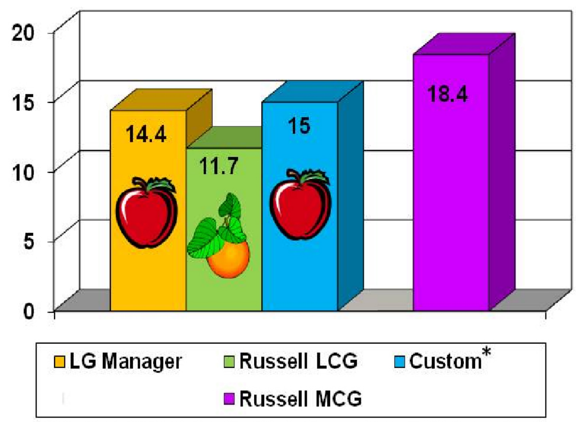
Russell and Surz agree manager is NOT Russell 1000 Growth but disagree on Small Growth Exposure



So now we can re-assess the manager's success or failure using a custom benchmark that is more precise about what this manager actually does. The exhibit on the right [Evaluate Skill not Style], shows that the manager has actually underperformed by 60 basis points per year. In other words, had we not customized the benchmark we would have evaluated style, not skill – a costly but common mistake. So now would you hire this manager? Before you decide, let's do some attribution work.

Evaluate Skill not Style

Returns for Q1/03 to Q2/07 (4.5 Years)



* Custom Determined with Returns-Based Style Analysis

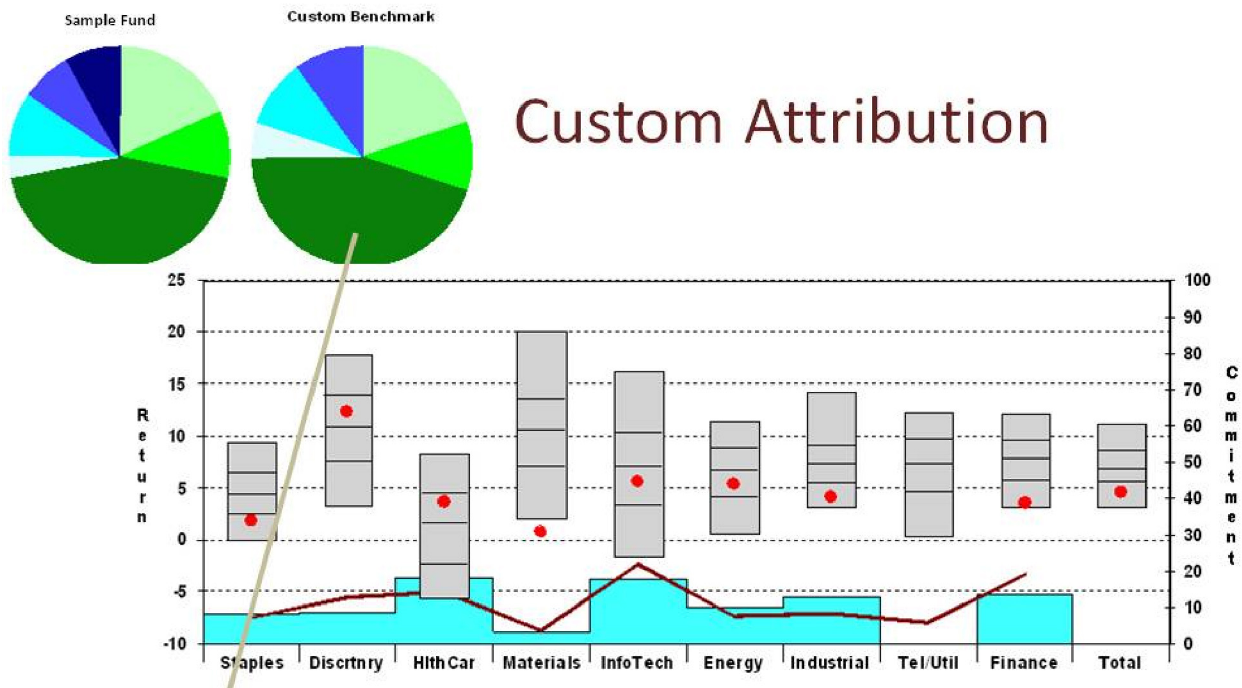
The New Attribution

In keeping with the new "Trust but verify," the new attribution is most concerned about benchmark accuracy although of course it's also important to have an accurate manager return. The new attribution takes linked monthly buy-and-hold returns and the most accurate actual return we can find, and uses these two pieces of information to calculate "activity" as a measure of value added or subtracted by intra-month decisions. Activity is the difference between the most accurate actual return and the linked monthly buy-and-hold return. This provides an additional insight that is frequently quite interesting.

Most importantly, the benchmark in the new attribution is customizable in a variety of ways, as follows:

- Customizing the benchmark in the new attribution
- Blend of styles, sectors or countries
- Normal portfolio, which is a list of stocks with their neutral weights
- Any fund, like a competitor or an index fund

When the benchmark is right we can rely on the sources of success or failure revealed by attribution. Also, we can dispense with the slicers and dicers that segment the portfolio and the benchmark by characteristic, such as capitalization and price/earnings ratio. These segmentations serve only to demonstrate that the benchmark is wrong, and to complicate an already complex analysis.



	Staples	Discrtnry	HlthCar	Materials	InfoTech	Energy	Industrial	Tel/Util	Finance	Total
Fund Return	1.81	12.30	3.69	0.71	5.58	5.39	4.12		3.59	4.52
Fund Commitment	8.15	8.41	17.98	3.39	17.83	10.06	13.02	0.00	13.11	
Benchmark Return	4.50	10.87	1.67	10.65	7.15	6.77	7.41	7.47	7.99	6.96
Benchmark Commitment	7.10	12.80	14.12	3.70	21.92	7.49	8.22	5.67	18.98	

	LargeCap			Mid Cap			Small Cap		
	Value	Core	Growth	Value	Core	Growth	Value	Core	Growth
Benchmark Allocation	20.00	10.00	45.00	5.00	10.00	10.00	0.00	0.00	0.00

The exhibit on page 11 [Custom Attribution] exemplifies the new attribution and its incorporation of hypothesis testing. The floating bars in the exhibit represent all of the portfolios that could have been formed in sectors of the customized benchmark. In this way, we can see significance straight away. A dot plotted in the top or bottom decile is statistically significant at the 90% confidence level. The dots are the manager's actual returns in the sector, and in total.

As you can see in the bottom of the exhibit, the benchmark in this case has been specified as a custom blend of styles. Accordingly the rankings of performance within sector are customized to this style blend, and employ hypothesis testing. Also shown in the exhibit are allocation "bets". The blue shaded area at the bottom of the exhibit is the portfolio's average allocation to sectors and the red line is the custom benchmark's average allocations.

Since we can rely on the sources of success or failure reported in the new attribution, we can question the manager about both. We'd like to confirm that the successes are likely to continue and that the failures have been identified and are being corrected. After all, due diligence is all about the future.

Conclusion

Just as the Bubonic plague is no longer a concern, we need not worry anymore about the accuracy of reported returns because there are plenty of cross-checks in place. We need to move on to what really matters, namely accurate benchmarking, especially in the critical and expensive area of performance attribution.

If the benchmark is wrong, all of the analytics are wrong.

◆◆◆

Ron Surz is President PPCA Inc Ron@ppca-inc.com (949)488-8339

REFERENCES

Statman, Meir. "What Do Investors Want?" ***Journal of Portfolio Management***, 30th Anniversary Edition 2004, pp. 153-161

Surz, Ronald J. "Getting All the Pieces of the Puzzle." ***The Sortino Framework for Constructing Portfolios***, edited by Dr. Frank Sortino, 2010 Elsevier Inc.