Commentaries on *Benchmarks and Investment Management* [2003]

Theodore R. Aronson, Peter L. Bernstein, Barclay L. Douglas, Richard M. Ennis, Elizabeth R. Hilpman, M. Barton Waring, Arnold S. Wood, and Jason Zweig

Collected by Laurence B. Siegel as an addendum to *Benchmarks and Investment Management* [CFA Institute Research Foundation (2003)], but never published. Presented in alphabetical order by author.

Comments by Peter L. Bernstein[1]
Peter L. Bernstein, Inc., New York

Performance measurers seek benchmarks the way bees seek honey. Today, the benchmarks on which so much measurement depends have developed deep cleavages that raise serious questions about the relevance of these benchmarks. How representative of today's portfolios are the Dow-Jones Industrials, S&P 500, the Russell 2000, or the Wilshire 5000? How well do these indexes provide an answer to the perennial question: How's the market? How appropriate are these indexes for arriving at judgments about investment management? Most important, what kind of role *should* these benchmarks play when we try to arrive at judgments about investment management? Who really wants a portfolio that looks like today's S&P 500? Can any manager rationally track a benchmark containing over a thousand stocks when no active manager is going to hold anywhere near that number of issues? How crazy can you get?

Let me carry this line of argument one step further and ask the big question. *Should clients even consider using market indexes or so-called "normal portfolios" in making judgments about investment management*?

I am going to postpone a discussion of this last question, because it goes right to the heart of the linkage between portfolio measurement and the choice of investment objectives. I am also going to limit the analysis at this point to managers of marketable securities like stocks and bonds.

Clients do have to make some kind of judgment about how the managers of their money are doing. If the manager is active, the employment of a passive benchmark seems like a logical step to take. Most institutional clients appear to have taken that step. According to a survey in the February 2000 issue of *Institutional Investor*, 88% of the respondents use indexes or normal portfolios, and 8% use style indexes. Four percent still hang in with the performance of peer groups. Ninety-nine percent of U.S. equity portfolios, 91% of U.S. fixed income portfolios, and 87% of international equity portfolios are judged in this manner. Manager performance in small-cap

---

[1] From *Economics & Portfolio Strategy,* April 15, 2000.

stocks and other asset classes seems to be judged less frequently against these kinds of benchmarks. I was astonished to note that only one third of the respondents thought that there are too many different kinds of indexes, although the authoritative annual Prudential Securities study pioneered by Claudia Mott now lists 86 equity indexes, with any size cap you like and with or without value or growth combos.

One of the many troubles with using these indexes as benchmarks is that they are floating crap games. They are not passive portfolios in the true sense of the word. The membership of the leading equity index, the S&P 500, changes constantly, and not only when companies disappear through merger or other uncontrollable events. Standard & Poor's makes active decisions about the membership of this revered icon. Dow-Jones does the same with the Industrials. The composition of the other indexes, such as the Russell x000's or the style indexes, varies as stocks rise and fall and therefore gain or lose qualifications to be included among the Chosen. These benchmarks not only contain far more issues than any right-minded active manager would carry, but, even worse, this year's Russell x000 is not the same group of stocks as last year's. And every one of these changes in the indexes are put into effect *after* the fact, *after* the character of the market has changed, but never in anticipation of change in market character. As most of these indexes are market-weighted – and the same applies to normal portfolios - whatever is hottest in the market acquires the greatest weight in the index. In 1980, or in times like the present, when a very small number of stocks accounts for an outsize portion of the index, the index is in all likelihood riskier than many clients would be happy to tolerate. What should we be thinking about a manager whose benchmark is one of these indexes but whose portfolio is the same as last year's?

The managers tied to this bogey must make decisions about risk that may be totally irrelevant to the kinds of strategies they want to follow, or their conservatism can lead them to wider tracking error than clients are likely to permit. Consider the following advice from the March 2 issue of Goldman Sachs's bi-weekly *Portfolio Strategy Research*: "If portfolio managers believe the polarisation in market performance will continue (in either direction), active bets between the largest industry groups and the rest of the market should be reduced in order to contain tracking error within more 'normal' ranges."

Asking an active investment manager to manage against such unstable benchmarks is equivalent to cruelty to animals. But at least in horse races the finish line stands still while the race is under way. Human investment managers receive no such consideration.

But there is a more important problem, which I like to refer to as bogey risk. Which index or normal portfolio do you choose as your benchmark out of the many that are available? Remember, whichever benchmark you choose, that is the one that the manager will manage against. But how does the client know which is the best index to choose? No matter which of the many indexes a client selects, there is a good

chance that some other index is going to outperform that choice. Just as in selecting managers, the client has no way of knowing ahead of time which choice is going to be the optimal one. Then the poor client is left with another nasty question: Which is better, a stable of managers outperforming benchmarks that underperform or a stable of managers underperforming benchmarks that outperform? And clients today think that they can use these kinds of benchmarks to make judgments about how their portfolios are doing?

Here is a variant on this theme, for which I am grateful to Russell Fogler.[2] Suppose a managed portfolio outperforms its benchmark by 400 basis points in some particular time period. As a result of this handsome achievement, the managed portfolio will now tend to have a P/E and cap size higher than the P/E and cap size of the benchmark. Subsequent performance from that point forward may appear to be due to factor bets by the manager on size and valuation when the manager has no such intent but is just holding on to positions that have been exceptionally profitable. How does the client make a proper evaluation of the manager's performance: will it be due to nothing more than these factor bets or to the manager's ability – or lack of it – at stock selection? You can get around this by rebalancing the benchmark portfolio after each performance measurement to match the active portfolio's size and valuation, but then you wipe out whatever information you might have on factor bets, which is just as important as information on manager skill.

Where You Stand Depends on Where You Sit

Theodore R. Aronson
AJO Partners, Philadelphia

I started in the industry in 1974 at Drexel Burnham. In those days, performance was routinely compared to the Dow industrials, with income included for the portfolio but not for the index. I soon figured out that starting out with a 6% lead (the yield on the Dow in the summer of 1974 with the Dow below 600) definitely helped relative performance! A few years later I started the Quantitative Equities Group, which used a full-blown Markowitz model to create efficient (ha!) portfolios. Things worked like a charm, until they didn't: October 1978, when all hell broke loose and our portfolios lost 11% more than the "benchmark" (now the S&P 500 total return). Small stocks were off 24% versus the 500's "mere" 9%, but who knew about small cap stocks? The real problem was, we ran a very public fund, The

[2] See H. Russell Fogler, "Normal Style Indexes—An Alternative to Manager Universes?" in *Performance Measurement: Setting the Standards, Interpreting the Numbers,* ICFA Continuing Education Series, CFA Institute, January 24, 1989, pp. 97-104, http://www.cfapubs.org/doi/pdf/10.2469/cp.v1989.n3.13

Revere Fund, which was the first active fund registered with the SEC which used MPT at its core. [Note to self: lose the Markowitz optimization nonsense.]

Fast forward a half dozen years to 1984 and the formation of Aronson + Fogler. No more overly-nuanced mean-variance optimization, brute force was all we needed: 250 equal-weighted stocks (about 20% of the buyable institutional market), with sector weights matched to the S&P 500. No stock-specific bets to worry about, no sector bets, fully invested,.... generally, a sensible way to run money. With some trips alone the way ('87 wasn't a great year to say the least), we added value (and clients) through the mid 90s. Then factors conspired against us -- small-cap bias turned substantially against us, value didn't help any, and there we were liking Exxon but betting against to the tune of 3%.

Although we had presented benchmarks like the equal-weighted S&P 500 (we called our work Mixed Cap and All Cap), everybody held our feet to the fire of the cap-weighted index. We jumped on Bill Sharpe's return-based attribution (actually, worked directly with him on it), just in time for the underlying sector weights in the style portfolios to whack us. It wasn't the perfect storm; it was just bad weather that went on for years. [Note to self: better pay more attention to underlying benchmarks. In fact, better to let the client move first and specify the bogey in advance.]

We are approaching our 10th anniversary of running money as we do now -- paying close attention to the benchmark (sector weights, individual stock weights, cap distribution within sectors, and other fundamentals) as well as statistical properties of individual stocks and the resulting portfolios. (Given that we now combine what I did in my prior incarnations, it's like an oldies radio station: The best of the 70s and 80s!) That we earned excess returns may be luck, but there is no denying that attribution in both good times and bad is now manageable and intellectually defensible. Amen.

Do we “waste” capital in terms of dead weight (thank you, John Freeman)? Yup. Is it worth it? Yup squared!

Warren Buffett doesn't have to compare his results to a benchmark (though he does), because of the magnitude and longevity of his outperformance. The rest of us mere mortals need a benchmark to justify our existence. It is a mathematical fact that, given the low level of alpha even good managers achieve and the variance of that alpha, it take many YEARS to prove anything but luck is involved.

In the ideal world Jason Zweig and Peter Bernstein envisage, smart managers aren't constrained by policy benchmarks or style boxes, they do what's right and profitable, and investors are rewarded. Then how come the only place were managers outperform is in their marketing materials? Have you seen any reputable studies were less than 75% of us have underperformed almost every benchmark in

sight? And those studies are of time-weighted returns -- I shudder to think of what dollar-weighted returns would show!

Only in a world were investors reflect Buffett's favorite holding period (forever) and invest against the tide of short-term returns and have managers who stick to their knitting (without selling out, taking on too much dough, losing nerve) does the benchmarkless world of investing lead to better returns for the owners of the capital. Otherwise, benchmarks are a good thing -- heck, if only because they give Jack Bogle ammunition for blowing the rest of us out of the water!

Comments by Barclay L. Douglas
Criterium Advisers, Medfield, MA[3]

Benchmarks are indispensable and mark progress in our industry. As indices, benchmarks help us understand asset class return characteristics, serve as the low cost alternative to active management, and help sponsors structure and monitor total fund exposures. As performance objectives, they are essential in the analysis of manager skill. For a variety of reasons, most managers have periodic complaints about benchmarks, and some are quite vocal in their contempt. While it is noteworthy that many of these same firms use benchmarks as an internal tool for compensation and/or will trot out benchmark analytics with clients to explain periods of disappointing results, some practices related to the use of benchmarks may produce handicaps in the pursuit of sound investing.

These handicaps include the use of overly specific mandates that restrict potential sources of value added, methods of manager evaluation that generate counterproductive hiring and firing, manager fees that are too high for services rendered, a lack of clarity between managers and clients in defining risk, and, in extreme cases, a distortion of the capital market process of resource allocation and pricing.

Benchmarks have contributed to the demand for and packaging of portfolios that are specialized to an absurd level. As recognition grew that some managers emphasized or biased portfolios toward a non-core philosophy, style and sub-sector indices were constructed to better explain the cyclical performance. Specialized indices legitimized style boxes and manager pigeonholes. As acceptance of these indices became the norm, two questionable practices emerged. First, medium and large asset management firms developed a “product” culture. They rushed to fill their shelves with specialized product, similar to other businesses creating mini skirts, mood rings and SUV’s. A manufacturer can make a quality pair of shoes to consumer specifications whether or not he likes the design. Investing is different.

[3] At the time this was written, Mr. Douglas was president of Baring America Asset Management (Boston).

For a manager to deliver a mid cap growth product, personal beliefs and interests should be in sync with management style. Second, as sponsors re-configured equity and fixed income structures to employ multiple specialty managers in a rigid mosaic, their investment sources of value added became focused on intra-mandate (benchmark) selection, reducing or eliminating the opportunity to exploit perceived relative value differentials between mandates. While I am conceptually a supporter of minimizing the non-productive "moving parts" within a total fund, I fail to understand why relative inefficiencies would exist only at the security level and not between styles or asset classes. I sympathize with the manager who has been told to stay in his box – what Warren Buffett referred to in his 1994 shareholder letter as a client accusation of intellectual cross dressing. "You cannot own Samsung because it is not in a developed country, you cannot hold Dean Foods because its price to book is too low, you must sell Pactiv Corporation because the cap size is 500 million over your limit." These artificial distinctions, chosen in the name of simplicity and specialization, will trouble true investors.

Now, the easy critique. With so much at stake, a decision-maker's natural inclination is to seek precision. Associated with precision is a desire for the quick fix and the short time horizon. The nature of markets confounds this objective. A quick review of studies comparing manager results in successive five year periods shows a low to negative correlation between consistency of success or lack of success. Manager-client partnerships that are contractually or conceptually tied to returns surpassing a target benchmark fail to recognize the cyclicality of manager returns and the importance of qualitative factors such as personnel stability and being right for the right reasons.

The big questions: Is it a market of securities or a securities market? Am I an investor or an asset class participator? Am I after alpha or the risk premium? All investors are ultimately after real returns. Most institutional sponsors have accepted the risk premium construct across asset classes, which is the financial equivalent of a belief in God. It is the long-term safety net. Acceptance of this construct generates the policy portfolio and the validity of measurement tools such as tracking error and information ratio. The problems: Who made the benchmark sacred? Are my tracking error tools accurate? Why am I sacrificing investment merit for requisite factor and sector exposures?

These are not attacks; they are sincere concerns. I bristle at the high equity turnover in July when the Russell indices are reconfigured and my reaction to a 2000 managers' report that they had lowered risk in the portfolio by increasing technology exposure was an immediate desire to be "less sophisticated." Yet I do not wish to preside over a fund whose equity exposure is 25 stocks regardless of the fundamental justifications, or more strongly, I do not want the equity and fixed income exposures of American retirement pools to contain risks so specific they essentially guarantee a large percentage of those pools will produce unacceptable returns. Advocating more opportunistic strategies will not increase average fund returns; it will, however, expand the range between good and bad.

Regarding managers, Grinold and Kahn say they are much more averse to the risk of deviation from the benchmark than they are averse to the risk of the benchmark. This is appropriate, but based on a broad generalization. Categorize managers as follows:

1. Passive
2. Enhanced index. 1.5% tracking error or below.
3. Benchmark driven -- Active managers who focus solely on active positions and are comfortable holding neutral weights in big benchmark securities with no positive or negative conviction.
4. Benchmark aware -- These managers hold only investments they like but will utilize risk control tools to understand ranges of prospective return and assure that conviction is aligned with active exposure.
5. Benchmark oblivious -- These managers focus on returns expectations for each individual position without regard to relative or asset class returns.

Two questions spring out of the differences in these two categories. First, is a benchmark driven firm, institutional or retail, really earning a full active fee? The second is more detailed. Consider the problems that arise when a benchmark oblivious manager has been retained in a role where the client actually seeks a benchmark-driven firm or vice versa. The two parties will think and speak about risk from two entirely different perspectives.

Benchmark oblivious managers will likely feel shackled by clients who expect benchmark like returns and utilize short time periods for evaluation. Consider why managers behave so differently in their actions regarding diversification. The manager whose top 100 ideas all have relatively equal conviction, embraces diversification; however, the manager with large differences in the degree of conviction between ideas #1, #5, #15, and #50 should favor high active exposures (caveat – conviction is not always a guarantor of results). I first noticed this differential in 1980 when reading the newly published book *The Money Masters* by John Train. Mr. Train profiled nine different legendary investors. One stated that he seeks to own 12-15 stocks, and, if forced to own more, would produce a riskier portfolio because the additional holdings would not be as well understood. One hundred and sixty pages later, another investor said if forced to own a subset of his traditional 80 stocks, he would be scared ****less.

I have a few observations regarding index construction. The differences in performance between the major indices (S&P 500 v. Russell 3000) and the enormous differences between the performance of specialized style indices produced by competing consultants, e.g., Russell and Wilshire, suggest a modified view on benchmarks. The passive benchmark is simply a portfolio with its own unique characteristics.

The cap-weighted characteristics of benchmarks present a particularly troubling characteristic for me. While I will concede cap-weighting is essential in meeting the "investability" criteria for a benchmark, its momentum oriented characteristics seem far too destructive. And for those who attempt to gauge the success of active management by the position of the index in a manager universe, consider that the equal weighted tendency of managers hurts them in markets when the biggest stocks do well and helps in markets when large companies under-perform (see attachment).

Finally a comment on what is potentially the biggest problem. Because other issues can be solved by better individual behavior, this problem is less controllable by the individual and has the potential to distort markets. When a critical mass of managers, both retail and institutional, adopt an emphasis on tracking error (benchmark driven), positive or negative cash flows naturally move in and out of the largest cap stocks without regard to valuation or fundamental prospects. While it appears 95% of our industry readily accepts that large-cap U.S. stocks are the most efficiently priced public assets, I believe a reasonable assessment of market behavior over the past five to ten years will suggest otherwise. The emerging emphasis on tracking error controls has generated at least part of this opportunity.

In conclusion, benchmarks are fine. The problem arises in how they are used. Consider the following suggestions.

1. De-emphasize performance measures in manager retention decisions and try your best to lengthen time horizons.
2. Reconsider your belief in the risk premium construct, then determine where you seek to be on the spectrum between investor and asset class participation. Do this separately for each asset class, and incorporate your thoughts regarding the prospective levels of excess return for each asset class.
3. Study your benchmarks. Much work can be done here. Some are structural, focusing on questions related to reducing the cap weighted characteristics or the potential for optimized benchmark. If I seriously considered the benchmark to be my safe harbor, I would be actively involved in its approval. Figuring out ways to lower exposure to Japan in the early 90's or Nortel from the Canadian indices more recently were sound ideas and they worked. Seek out sound modifications.
4. Don't pay full fees to benchmark driven managers. As Richard Ennis has said, encourage them to get the hamburger helper out of the portfolio. Don't pay managers active fees to hold neutral conviction positions. Incorporate the study of portfolio dead weight into your manager analysis.
5. Force mutual funds to provide portfolio deadweight statistics as they have recently been forced to present after tax performance results. Understand how their diseconomies of scale or their business risk concerns impact their investing.

6. Reconsider sources of value added for your portfolio – Should some mandates be combined to allow the manager greater flexibility in exploiting the relative value between sub styles or asset classes?
7. High tracking error managers with high perspective excess returns should be free from concern (re: their place in a total fund structure.) If deemed important, the sponsor can control benchmark risk at the aggregate asset class level.
8. Recognize the limitations of our quantitative tools. Balance judgment with data. Do not seek precision beyond its limitations.

In Defense of the Lowly Benchmark (Or, Don't Shoot the Messenger)
Richard M. Ennis[4]

We grumble about benchmarks. We talk of improving them, de-emphasizing them, and dumping them altogether. But maybe benchmarks aren't the problem, or much of a problem. Maybe the problem is us.

Following Sharpe's "Arithmetic of Active Management," should we not expect the majority of active managers to underperform properly constructed benchmarks? If a properly constructed benchmark equates to passive implementation of a strategy and the cost of active management exceeds the cost of passive investment, the answer would seem to be *yes*. This unhappy fact appears to be at the heart of our benchmark discontent.

Most investors appear to be in a state of denial regarding their performance prospects. Outwardly confident they can select managers who in turn can exploit security mispricing, they proceed to closet index with alacrity, hiring several domestic equity managers with "complementary" styles, when one or two would do. They fuss over so-called style drift. They put managers on watch lists after a relatively brief period of poor performance, even though the manager exhibits no discernable weakness, e.g., a change in key personnel, process, or ownership. In the name or risk control, they embrace "enhanced" indexing, the eat-your cake-and-have-it-to form of active investment. Is it any wonder that investors that act like this wind up grumbling about benchmarks?

Most should consider indexing a larger fraction of their portfolios. The appropriate indexing percentage is the one that affords them the peace of mind necessary to allow them to hire a very few of the smartest investment managers they can locate. The overriding criterion for selecting managers should be skill – certainly not style.

---

[4] Richard Ennis, who is now retired, was with Ennis, Knupp & Associates when this was written.

And then the investor must resolve to stay the course until something other than interim performance undermines their conviction in the manager.

So the prescription for curing benchmark angst has two parts: (1) only bet what you can afford to lose, and (2) play the game smartly.

Comments by Arnold S. Wood
Martingale Asset Management, Boston

The "eye of the beholder" issue may not be technically relevant, but one which may deserve some honorable mention. Managers, clients, prospective portfolio buyers, and consultants – have different issues, ambitions, and intentions when it comes to using benchmarks. I would venture that the majority of prospects – dare I suggest, unsophisticated masses – are more concerned with "hot dot returns" than how those dots got there. Performance expressed in returns has always been the driving motivation for most buyers. For most buyers, benchmark tracking error, IRs, etc. are not that relevant when making hiring decisions. It may be less so for firing decisions. It's sort of like Reality TV – sex (returns) sells, not brains (portfolio engineering)! If I have a compelling 3-year compound return, an outrageously high tracking error, and a good, plausible story as to why will returns will continue, I will outsell the "propeller heads" who have realistic returns, a tight tracking error and a good story. A .6 IR with 3-4% tracking will lose to a .8 IR and tracking of 8-10% every time. Winners often have unreasonable deviation away from the benchmark – lucky or skillful? We see small cap value competitors 20-30% invested in the Russell 2000 Value and the balance in cash and the Russell 1000. Return per unit of risk – how many funds sell this in the newspapers? Returns sell, benchmark tracking doesn't; that's what I am suggesting. Maybe I am just covetous!

One question I keep coming back to is, "Is tight benchmark tracking error risky in the sense that it doesn't inhibit artsy managers? The question is grounded in role of affect (the feeling of risk) not Elizabethan standards of calibration – probabilities, regression, etc. We all tend to get caught up in normative practices probably because it's what we have been taught to do. Sometimes the visceral sense of risk needs to be weighed in with the math when making decisions. The question is when and by how much? Survival instincts (affect heuristic) are natural, not something we spend much time trying to figure out. Many animals avoid extinction by fleeing danger with out calibrating. Now that's real return. Those that wait to assess the danger often become a meal. That's real risk. Some managers achieve IRs with a touchy feely approach; the problem is that it's only one or two in a hundred (I may be overestimating) and the chances that those apparent investment savants will keep producing is not only low, it's random.

So, benchmarks are a necessity. Otherwise so many will diverge so far that the clients will lose. Relying on an affect approach to investing can lead to disaster. Losing other peoples' money is a serious business and those who think they can make money in relatively efficient and transaction cost-laden markets are deceiving themselves and their clients. An old gambling proverb is "You've got to stay in to be in it."

There are clients who claim that staying close to benchmarks is simply a ploy to reduce money manager business risk. Are they really suggesting – "Go ahead, try to shoot the lights out with our money." Then again, maybe clients are in the best position to optimize the overall aggregate risk/return issue by controlling the managers. If this is true, why are they constraining their portfolios by constraining shorting? We can talk more about this complicated aside if you think it's a point worth accentuating.

Interview with Elizabeth R. Hilpman
Barlow Partners, New York

Larry Siegel (LBS): Liz, you've had one of the most diverse backgrounds of any practitioner in investment management. You've worked at a large, multi-strategy investment management firm; you've been a plan sponsor; and you've worked in several hedge fund organizations, including Barlow Partners, the hedge fund of funds where you are a partner now.

In the hedge fund world, there's no underlying asset-class return or "benchmark," unless one counts hedge-fund "benchmarks" that are more like peer groups. How should hedge fund managers be measured? How should their risk-taking be assessed?

Liz Hilpman (ERH): The hypocrisy is that we in the hedge fund world all say we're an absolute-return game, but we all invest in securities, and the alternative to investing in hedge funds is long-only equities and bonds, so it's really a relative-return game.

Our objective is to preserve capital so we don't want a negative return, but when we're down 2% and the market is down 30% we want to show our superior performance! A hedge fund's primary objective is to not lose money, but the alternative to hedge funds is equities and fixed income which can make a lot of money, so you have to make a relative comparison as well as an absolute one, or else single digit returns every year would look fine because they preserve capital and are positive numbers.

The absence of real hedge fund benchmarks also makes it harder to decide whether one hedge fund is better than another hedge fund. You need to evaluate how much leverage they take, what kind of instruments they invest in, what market or factor exposures they have, how much diversification is involved, how much style drift the fund has experienced or is likely to experience. You can't do what

evaluators of long-only managers do – measure alpha and tracking error – because there's no realistic benchmark.

LBS: Some index providers *have* begun to construct hedge fund "benchmarks." Unlike the long-only world, where the benchmarks consist not of other managers but of the underlying asset choices, hedge fund benchmarks do consist of other hedge funds. Are such benchmarks any good? Do you use them? Does their existence cause any behavior that you think is regrettable, or desirable?

ERH: Hedge fund benchmarks are being designed, not for performance measurement, but for asset gathering. They're being created as the basis for index funds, which makes them poorer, not better, benchmarks. To be in a benchmark, a hedge fund manager has to provide 100% transparency, capacity to the index entity, and liquidity to the index entity. That makes them potentially less rewarding investments.

Having said that, I do think there's useful peer group information in hedge fund "benchmarks." They are a collection of 100 or more hedge funds, and unlike the much larger hedge fund databases, the returns are carefully documented and effort is taken to remove biases. Just because they're not really benchmarks doesn't mean they're useless.

LBS: Moving to the long-only world, much has been made of the idea that benchmarking is harmful to the process – and, ultimately, costly to the client – because it forces managers to take "real" risk to avoid "apparent" risk, or tracking error. What do you think of that? How should one measure the performance of long-only managers?

ERH: It's the whole dead weight concept that doesn't make sense to me. Why should you own a stock that you think is going bankrupt? Why should you own WorldCom? If your analysis says that the company is actually bankrupt and that we're just waiting for the market to recognize it, is it really "risk" to not own WorldCom if you're an active manager? But it's tough because you need a benchmark to measure performance. The resolution is to use benchmarks but to allow an active manager more freedom to maneuver than many plan sponsors current do – you should allow them more tracking error.

LBS: Finally, let's move to the plan sponsor's world of investment policy. Let's say there's a new asset class – TIPS or non-U.S. corporate bonds. Do you change the policy benchmark to reflect the changing opportunity set ? Or it is more sensible to have simple, stable benchmarks and then to take risk relative to the benchmark to make desired changes in the portfolio?

ERH: Either has problems. If there's a long-term change in the investment opportunity set, change the benchmark. I like live benchmarks that evolve with the world. It depends on whether the benchmark is to see what I get paid, or as a starting point and guide and a check on investment behavior. If it's used for my pay,

it should be an evolving benchmark. If it's just a guide, a simpler and more stable benchmark.

We're always looking for reference points. The S&P 500 is not an ideal portfolio, but it has familiarity and predictability, so it's a benchmark in the truest sense of the word. A policy benchmark that stays the same while markets change is the same way. It's a stable reference point and it has history and accumulated understanding.

LBS: To conclude, note that Peter Bernstein has suggested that if the purpose of an asset pool is to pay a liability, or to fund spending needs, then a liability-related or absolute-return benchmark is what is needed. Plan sponsors should then be free to search for whatever asset mix best achieves the liability or spending goal at a given point in time, and should be encouraged to change the asset mix in response to changing conditions, rather than rebalancing to a fixed asset mix. Does this suggestion resonate with you?

ERH: Peter's idea is too hard to implement. There are too many differences of opinion to figure out how to get started investing and to stay well-behaved without an asset-mix benchmark. Using a liability benchmark, or an absolute-return benchmark, makes the investment process into a contest to prove that your asset class or strategy will earn, say, 10%. You need the discipline of a starting point. So while it's intellectually appealing to say the liability is the benchmark, there's no asset that pays that rate of return. You have to buy assets that exist, so we need an asset-mix benchmark.

At The Common Fund, which manages endowment assets of colleges and universities, we spend hundreds of hours building a forecasting model. In the end, however, we used the output of the forecasting model to modify pre-existing asset weights, weights that are in a given school's benchmark. Without a starting asset mix, we would have spent thousands instead of hundreds of hours on forecasting, but the forecasts would not have been any better, and we would have just wound up investing in the assets with the biggest upside error in the forecasts.

LBS: Thank you very much.

Are Benchmarks a Good Thing?
M. Barton Waring[5]

While today's investment management practice doesn't acknowledge it, the truth is that the liability is the ultimate benchmark for a pension plan's investments. The goal of the investments is to maximize the growth of the surplus while controlling the volatility of the surplus. (Humor me; I do know that most plans are actually in

---

[5] At the time this was written, Mr. Waring was chief investment officer for investment policies and strategies at Barclays Global Investors (now BlackRock) in San Francisco.

deficit right now, but a deficit is just a negative surplus!) Bottom line: low volatility of surplus is a good thing.

It is interesting to note that other types of organizations that hold assets in order to fund liabilities—banks, insurance companies, etc.—typically take very little risk relative to the liability. Banks have raised this to an art form, and insurance company regulation is all over this. Why is it then, that pension plans are invested 70-30 equities-bonds, while the liability is close to being all bond, for argument sake let's call it 20-80 equities-bonds? No wonder plan sponsors don't like the risk that they are now experiencing. There is a huge gap between the liability benchmark and the investment policies typically being used.

So the ultimate policy benchmark is the liability. But there is another level of benchmarking that we deal with more regularly, and that is really more at the heart of the question. This is at the level of implementing the investment policy, once it has been set and at whatever level of surplus risk that it has been set. Are the investments, *as implemented*, beating the asset allocation policy benchmark(s)?

Benchmarks are much more important than just their role as a performance yardstick. There is a fundamentally important fact about investments that is highlighted by using benchmarks. Benchmarks represent *market* risk exposure. I use this term in its technical sense—systematic risk, diversified risk, beta risk. A value benchmark is fully diversified, and captures the market risk in the value domain. A manager will have a beta with respect to one or more benchmarks. For readers familiar with returns-based style analysis, you know that it is the norm rather than the exception for traditional active managers to have betas (style weights) in more than one style/size category, more than one market. Managers also take on *residual* risk, aka diversifiable, unsystematic, idiosyncratic, or *active* risk, and investors hope that the manager will show a positive residual return as a result. This residual return, a return that is relative to the beta-adjusted benchmarks representing the markets or sub-markets that the manager is invested in, is a "true" alpha.

So with this we can complete the discussion of why benchmarks are not only valuable, but necessary. Market risk or beta risk exposures, theory says, are inherently rewarded over time. These are the stuff of the upward-sloping security market line. But active risk can be diversified away without cost, so it is *not* inherently rewarded. It is only rewarded conditioned on skill; only the specially skillful can win in the great zero sum game. We don't hire active managers to get their beta exposure. This is apparent, because if we were hiring managers to get market return we would instead hire index managers, who will gladly deliver market returns at lower cost and with lower risk than will traditional active managers. The only reason left for hiring active managers in the face of index competition is to get true alpha—return over the beta-adjusted benchmark. Since the search for alpha is a negative sum game after fees and costs, it is important to look at it clearly. Only the specially skillful will be successful at generating it other

than through luck. But luck isn't consistent over time, and can't and shouldn't be depended on. You pay active managers well only because you think they have that special skill. And if you don't think so, you shouldn't hire them. Active management skill is a very special thing, and the active managers you hire should be exalted and respected. And deserving of good pay. Or they shouldn't be employed.

Of course active managers, at least all those typical managers that are net long in security positions, do in fact have market exposure. But if that is all that you want, you can get their betas better and cheaper elsewhere, usually in the form of index funds. The only reason to keep such a manager is if he or she really is expected to deliver alpha at a high level of forward-looking probability. This means that they have to be well above average in their level of skill.

The appropriate retort to any manager that disdains benchmarks is to ask them if they are exposed to any market risk or beta risk factors. You'll get some humor out of this, as most won't know what you're talking about. But any that do will of course have to acknowledge that they are "exposed" to one or more market sectors. This is an equivalent statement to saying that some component of their returns is systematic and some other component is pure active, and this in turn is equivalent to acknowledging that the manager is a relative return manager appropriately evaluated by reference to a benchmark! There really is no escape; there are always two components of returns and risks. *There is no such thing as a manager that doesn't produce relative returns.*

And while you're on this topic with your manager, ask him whether he thinks you should pay him an active fee simply for being exposed to the market? His attempt to answer ought to be priceless (and his actual answer should be, but won't be, nearly so). Managers should be paid for their skill in beating the market, not for being exposed to the market, and not for taking unskillful bets around the market.

It can be hard to identify a manager's true benchmark or beta exposures—especially when the manager makes a lot of market timing bets, switching between benchmark components. The historic returns will reveal unstable betas in a regression against the multiple components among which they rotate. Yet even if it is difficult to accurately define, there is always some "normal portfolio" representing the manager's normal level of market risk exposures. In extreme cases it is best for the analyst to simply default and use the closest general benchmark as a good-enough proxy—the S&P 500 for example. But a better job can usually be done simply by understanding the manager's articulation of its process. A value vs. growth style rotator, for example, might acknowledge favoring value; the investor might estimate his normal portfolio as 60% large value and 40% large growth.

What if the manager argues that his or her style rotation isn't explicit or intentional, but simply "falls out," an accident of the manager's stock selections? We would respond that such a manager is taking many uncompensated sector rotation bets,

and would generate more consistent returns if he or she were to use risk controls to dampen them.

Of course, any time a manager shows such heavy levels of sector timing activity, the logical question to ask is whether you believe the manager has *skill* at that activity. Market timing and style rotation are legitimate active management disciplines, but because of their low breadth they require greater skill than security selection approaches require if they are to deliver similar information ratios.

So benchmarks are important with respect to both the policy-setting decision and to implementation with active managers. Benchmarks really represent the inherently rewarded market risk of the overall policy or of any manager; so they provide structure allowing us to separate market risk and active risk. Such structure is badly needed given that active management comes to us with market risk and active risk all entangled. We want to make good decisions about who to hire in the very special role of active manager; without benchmarks we'd just be guessing. The fact that we use benchmarks for performance analysis is just a happy consequence of our ability to use benchmarks to sort out market and active risks.

Comments by Jason Zweig[6]
from *The Intelligent Investor,* by Benjamin Graham and Jason Zweig [2003]

Investing isn't about beating others at their game. It's about controlling yourself at your own game. The challenge for the intelligent investor is not to find the stocks that will go up the most and down the least, but rather to prevent yourself from being your own worst enemy – from buying high just because Mr. Market says "Buy!" and from selling low just because Mr. Market says "Sell!" ...[Y]ou must also refuse to judge your financial success by how a bunch of total strangers are doing. You're not one penny poorer if someone in Dubuque or Dallas or Denver beats the S&P 500 and you don't. No one's gravestone reads "HE BEAT THE MARKET."

I once interviewed a group of retirees in Boca Raton, one of Florida's wealthiest retirement communities. I asked these people – mostly in their seventies – if they had beaten the market over their investing lifetimes. Some said yes, some said no; most weren't sure. Then one man said, "Who cares? All I know is, my investments earned enough for me to end up in Boca."

Could there be a more perfect answer? After all, the whole point of investing is not to earn more money than average, but to earn enough money to meet your own needs. The best way to measure your investing success is not by whether you're beating the market but by whether you've put in place a financial plan and a behavioral discipline that are likely to get you where you want to go. In the end,

[6] Jason Zweig is personal investing columnist for *The Wall Street Journal*, New York and the author of *The Devil's Financial Dictionary.*

what matters isn't crossing the finish line before anybody else but just making sure that you do cross it.