

Total Portfolio Benchmarking

WHITEPAPER

JANUARY 2018

Total portfolio-level benchmarking for institutional investment pools can be approached in a variety of ways, based on which goals fiduciaries are attempting to achieve with each comparison. There are different types of benchmarks that can be used to accomplish specific goals. Portfolio level benchmarks are typically based on (1) an institutional portfolio's asset allocation, e.g., an asset weighted policy or dynamic policy benchmark, (2) an investable, passively-managed index, such as a 60/40 equity and bond mix, (3) a comparison to institutional peers, and/or (4) an institutional financial objective of the portfolio's returns, e.g., a specified nominal or real (inflation-adjusted) rate of return.

CONTRIBUTORS

Timur Kaya Yontar, PhD
Frank Benham, CFA, CAIA

Each type of benchmark presents various challenges and requires some amount of investment knowledge to grasp the implications underlying each comparison. This can create communication gaps for fiduciaries and various interested parties, such as retirement plan participants and beneficiaries. If an investment out- or under-performs its benchmark over an appropriate measurement period, investors should understand why it happened, whether it was to be expected, and most importantly what it signifies for doing anything differently – since in the end a metric is only useful if it feeds back into a decision process.

In addition, an appropriate benchmark is not always available for every underlying asset class. The widely accepted Bailey Criteria¹ for benchmarks include six characteristics:

¹ Financial Analysts Journal, CFA Institute, 1992.

1. Unambiguous—identities and weights of constituents are well defined,
2. Investable—can own a portfolio of the benchmark's constituents,
3. Measurable—performance can be calculated at reasonable intervals,
4. Appropriate—consistent with the composition of the portfolio for which it is a benchmark,
5. Reflective of current investment options—represents the market of the asset class, and
6. Specified in advance—constructed before the evaluation period.

Many commonly used benchmarks fail one or more of these tests, and thus the policy benchmark, made up of asset class benchmarks, will never be a perfect comparison for an institutional portfolio's diversified asset allocation.

Types of benchmarks

Policy and Dynamic Policy Benchmarks

Policy and Dynamic Policy Benchmarks are typically based on an institutional portfolio's policy asset allocation. Policy Benchmarks generally use a passive index for each asset class, weighted the same as the *targets* in a portfolio's asset allocation, to calculate a portfolio's return as if it were passively invested at policy targets. Thus, this comparison includes not only the effect of active management versus a passive index, but also incorporates the effect of having allocations different from the targets. These differences may be deliberate, reflecting tactical positioning, or unintentional, such as when a portfolio is unable to rebalance into or out of an illiquid asset class.

A Dynamic Policy Benchmark uses the *actual* weights of each asset class, and uses passive indexes to calculate the passive equivalent of the return achieved by the total portfolio. Thus, it is designed to focus on the cumulative added value from active managers. However, as mentioned above, not every asset class (e.g., Private Markets, Hedge Funds) has a passive equivalent for purposes of comparison. In these cases, a peer group for the asset class, or an un-investable index of active managers are often used as a best proxy, even though they fail several of the Bailey criteria described earlier.

To summarize what these benchmarks measure:

*Portfolio performance - Dynamic Policy Benchmark performance =
total value added (deducted) by portfolio's active managers*

*Dynamic Policy Benchmark - Policy Benchmark performance =
total value added (deducted) by portfolio's actual asset allocation differing from policy
asset allocation*

Passive benchmarks

Passive benchmarks are typically based upon a predetermined mix of public market investments such as stocks and bonds, and are rebalanced on a regular schedule (e.g., monthly). Examples include a 60% allocation to the S&P 500 equity index and a 40% allocation to the Barclays Aggregate fixed income index. A global version might use the MSCI All Country World Index ("ACWI") and the Barclays Global Aggregate Index for equity and fixed income, respectively. The purpose of including such a benchmark is to compare how the investments, as constructed and as implemented, perform versus a simple blend, thus conveying the added value from the more diversified (and complicated) multi-asset portfolio. Ideally the simple blend should be chosen because it represents a risk level roughly comparable to that of the multi-asset portfolio.²

² Many types of risks matter to different investors and thus risk is multi-dimensional and cannot be fully captured by a single metric; however, the *volatility of returns* (as measured by standard deviation) is a useful starting point because it measures the unpredictability that concerns most investors. Alternatively one could compare risk adjusted returns using a metric such as the Sharpe ratio, which calculates excess return (over a risk-free rate) per unit of volatility, but such measures may not be intuitive.

To continue the equations from above:

Policy Benchmark performance - Passive Benchmark performance = total value added (deducted) from diversification beyond a risk-equivalent simple blended portfolio

Peer benchmarks

Institutional peer groups are also a method of total portfolio benchmarking that can be used when fiduciaries wish to compare their investment pool's performance, in terms of both return and risk (volatility), to that of other pools in the marketplace. The data included in peer groups is typically collected by vendors and published only on a quarterly basis, some 2-6 weeks after quarter-end. While this comparison can be meaningful when using peer groups with significant membership, the utility declines as the sample size of the peer group is reduced. At the same time, a large peer group is likely to comprise a diverse set of institutions, which may differ widely in objectives (e.g., target return), resources (institutional staff, access to capacity-constrained managers, etc.), and constraints (such as risk tolerances or financial situation). Thus, while stakeholders may find it interesting how their institution's performance has ranked versus peers, and seek to figure out what peers are doing differently (in case they should be emulated), the comparison may ultimately not lead to investment changes since their institution's circumstances are unique.

While Institutional Financial Objective benchmarks provide the ultimate measure of whether a portfolio is achieving its *raison d'être*, they are disconnected from what capital markets are delivering at any given time.

Institutional Financial Objectives

Institutional Financial Objective benchmarks may compare returns to a static target return, such as the actuarial assumed rate of return for a pension fund, or an annual real spending rate for an endowment.³ Investment performance for time periods over one year are most often presented as annualized numbers, so periods shorter than one year will be presented as a fraction of a one-year return when using this method.⁴ While Institutional Financial Objective benchmarks provide the ultimate measure of whether a portfolio is achieving its *raison d'être*, they are disconnected from what capital markets are delivering at any given time. They also display little-

³ To convert to a nominal target return the real spending rate is incremented by an appropriate inflation rate, which must be taken into account so the spending stream will preserve its purchasing power.

⁴ For instance, a one-quarter static return goal would be displayed as 1.71% for an annual target return of 7%, as it is calculated geometrically: $(1 + 7\%)^{1/4} - 1 = 1.71\%$.

to-no volatility. Thus they only offer useful long-term comparisons, such as over a full market cycle, whereas benchmarks that are composed of risk assets will permit short-term comparisons with institutional portfolios.

Challenges of alternative assets

Academic literature has presented various characteristics that need to be met for appropriate asset class level benchmarking. The most widely used list of such characteristics, as mentioned above, comprises the Bailey Criteria, all of which are easily met for public market benchmarks tracking efficient, liquid asset classes such as large cap domestic equity. However, these characteristics are not achievable for less liquid asset classes such as Hedge Funds, Private Equity, or Real Estate.

Due to the idiosyncratic and often illiquid nature of assets included in many “alternative” asset classes, as well as lack of transparency, replication of individual investments in these strategies is simply not possible. In the case of Hedge Funds and Private Equity, results from a large group of managers are compiled in indices such as the HFRI/HFRX set of benchmarks for Hedge Funds or the Cambridge Private Equity Composite for Private Equity funds. These benchmarks, while perhaps representing the best comparison available, introduce a set of upward biases due to the fact that the constituents self-report. These biases include sample selection bias (since funds have an incentive to report only when returns are good) and survivorship bias. Additionally, in the case of the Cambridge Associates Private Equity Composite, the benchmark includes performance for only a sample of funds that Cambridge Associates selects rather than for the broad universe of funds in the marketplace, which may make it harder to match or “beat.” This is especially the case for a new allocation whose returns are still in the “J-curve” whereas the benchmark will comprise funds of different vintages and hence more “mature” on average.⁵

⁵ The J-curve refers to the returns in the early years of a private markets investment. These are typically negative due to fees and setup costs being deducted with offsetting investment gains not having yet been reaped.

Additionally, because one of the purposes of including Real Assets in a portfolio is the inflation hedge potential, real assets are often benchmarked to inflation-related measures. One example would be to compare these investments to the Consumer Price Index (CPI) plus an extra 3% to 5%. The problem with this comparison is that, while investments included under the Real Assets umbrella can have positive or negative returns, CPI is very rarely negative or strongly negative, even when capital markets experience stressed periods and investment returns are negative. Hence there can be significant mismatch in the short term, especially during periods of amplified market volatility.

The same problem befalls other “plus-a-spread” benchmarks, e.g. absolute return assets may be measured against cash (3-month T-bills) plus 3% to 5%, or private equity against public markets (S&P 500 or ACWI) plus 2% to 3%. The intent of using these is to account for some premium that investors are seeking from the alternative investments, such as compensation for illiquidity, or expectation of manager alpha. Yet all of these fail the second Bailey criterion, investability, because one cannot own a portfolio composed of instruments that deliver these returns, so tracking error between the investments and the benchmarks is likely to be substantial.

After a new target allocation to privates is adopted, it can take several years for the pool to reach that target.

Further, benchmarking for private market assets presents a unique set of challenges largely due to illiquidity, especially early on in a portfolio’s allocation to the asset class. After a new target allocation to privates is adopted, it can take several years for the pool to reach that target. During that transitional period, an interim benchmark can be appropriate, both for the asset class and for its component in the total portfolio benchmark. The interim approach recognizes that a significant portion of assets committed to private markets may not yet have been called, and remained “parked” in their public market proxies. It would measure against a private benchmark only that portion of the private allocation that has actually been invested, with the remainder still in public proxies being benchmarked against the public market equivalent.

Finally, illiquid investments complicate the picture still more because the most accurate way to assess their performance calculates “dollar-weighted returns,” whereas liquid investments use “time-weighted returns.”⁶ To combine the two in a single metric is mathematically meaningless, so a total portfolio benchmark would need to use time-weighted returns for all of its components including illiquids. Supplementing this top-level metric, a separate performance report for the private market assets would show detail about the fund-level IRRs.⁷

Conclusion

Due to the intricacies and diverse composition of total portfolio benchmarks, no single total portfolio benchmark can provide a perfect comparison for all time periods. Because of this limitation, Meketa Investment Group recommends that institutional investors utilize two or more total portfolio level benchmarks, while being aware of the structure (and flaws) of each.

⁶ In illiquid investments the investor commits a certain amount of capital but the manager controls when and how much of that capital is called (invested) or distributed, while for liquid investments the investor determines the timing of investments and redemptions. Although a time-weighted return can be calculated for illiquid investments, dollar-weighted returns (also called internal rates of return or IRRs) more properly account for their more complicated asset flows.

⁷ These are properly compared to a “public market equivalent” benchmark which takes the same cash flows, and call and distribution dates, and calculates a dollar-weighted return for how those flows would have performed in public markets.

Fiduciaries should understand why each benchmark performs the way it does in different capital market environments, and this understanding adds context to the investment pool's total return. For instance, a portfolio may lag a policy benchmark due to an active manager's underperformance and tactical positioning, while it only lags a dynamic policy benchmark due to manager underperformance. A passive index provides insights into the usefulness of a diversified asset allocation and active management at a given level of risk. Peer group benchmarks may provide insight into varying asset allocation differences and trends in the marketplace and the resulting impact on performance. Finally, it may be appropriate to use different benchmarks for different time horizons: over the short term, consider comparisons to weighted averages of market indexes or to peers, but to measure whether a pool is achieving its long-term goals, fiduciaries may prefer to focus on static or target benchmarks tied to Institutional Financial Objectives.

Appendix

Asset Class	Strategic Policy Benchmark	Notes
Equity (40%)	75% Public, 25% Private	Represents portion actually invested in (not committed to) privates
Public Equity	40% Russell 3000, 30% MSCI World (ex. U.S.), 30% MSCI EM	Represents geographic weights
Private Equity	MSCI ACWI + 200 bp	Represents global opportunity set, illiquidity and alpha premia; not "investable"
Rate Sensitive (20%)	50% Barclays Long-Term Gov't, 25% Barclays Corporate, 25% Barclays US TIPS	
Credit (10%)	50% US, 50% Non-US	Represents geographic weights
U.S. Credit	80% Barclays US High Yield, 20% S&P/LSTA Leveraged Loan	
Non-U.S. Credit	60% Bloom/Bar EM Local Gov't, 20% Bloom/Bar EM Hard Sov, 20% Bloom/Bar EM USD Agg Corp	
Real Assets (20%)	75% RE, 25% NR/IS	
Real Estate	60% NCREIF ODCE (1 qtr lag, gross) 40% FTSE EPRA/NAREIT Developed (net)	Represents portion actually invested in (not committed to) privates
Natural Resources & Infrastructure	CPI-U + 500 bp (capped at 10%)	Represents inflation-linked nature of assets, illiquidity and alpha premia; not "investable"
Absolute Return (10%)	3-Month T-bill + 400 bp	Represents alpha premium; not "investable"
TOTAL PORTFOLIO	→ 12% Russell 3000 ⁸ , 9% MSCI World ex U.S., 9% MSCI EM → 10% MSCI ACWI + 200 bp → 10% Barclays Long-Term Gov't, 5% Barclays Corporate, 5% Barclays US TIPS → 4% Barclays US High Yield, 1% S&P/LSTA Leveraged Loan → 3% Bloom/Bar EM Local Gov't, 1% Bloom/Bar EM Hard Sov, 1% Bloom/Bar EM USD Agg Corp → 9% NCREIF ODCE, 6% FTSE EPRA/NAREIT Developed → 5% CPI-U + 500 bp → 10% 3-Month T-bill + 400 bp	
SIMPLE BLEND (RISK EQUIVALENT)	→ 55% MSCI ACWI → 45% Barclays Global Aggregate	
INSTITUTIONAL FINANCIAL OBJECTIVE	→ CPI + 4% (spending rate)	

TABLE 1
Illustrative Asset Class and Total Portfolio Benchmarks

⁸ E.g., 12% = 40% equity * 75% public * 40% U.S. (benchmarked to Russell 3000).

Disclaimers

This document is for general information and educational purposes only, and must not be considered investment advice or a recommendation that the reader is to engage in, or refrain from taking, a particular investment-related course of action. Any such advice or recommendation must be tailored to your situation and objectives. You should consult all available information, investment, legal, tax and accounting professionals, before making or executing any investment strategy. You must exercise your own independent judgment when making any investment decision.

All information contained in this document is provided "as is," without any representations or warranties of any kind. We disclaim all express and implied warranties including those with respect to accuracy, completeness, timeliness, or fitness for a particular purpose. We assume no responsibility for any losses, whether direct, indirect, special or consequential, which arise out of the use of this presentation.

All investments involve risk. There can be no guarantee that the strategies, tactics, and methods discussed in this document will be successful.

Data contained in this document may be obtained from a variety of sources and may be subject to change. We disclaim any and all liability for such data, including without limitation, any express or implied representations or warranties for information or errors contained in, or omissions from, the information. We shall not be liable for any loss or liability suffered by you resulting from the provision to you of such data or your use or reliance in any way thereon.

Nothing in this document should be interpreted to state or imply that past results are an indication of future performance. Investing involves substantial risk. It is highly unlikely that the past will repeat itself. Selecting an advisor, fund, or strategy based solely on past returns is a poor investment strategy. Past performance does not guarantee future results.