MEKETA
INVESTMENT GROUP

Total Fund Benchmarking

WHITEPAPER
OCTOBER 2022

CONTRIBUTORS
Stephanie Sorg, CAIA
Frank Benham, CFA, CAIA

Benchmarking is a common and prudent practice in the investment industry. Many institutional investors spend significant time focusing on the benchmarks for their asset managers and how those managers perform relative to their respective benchmarks. However, it is arguably more important for those institutions to concentrate on and evaluate the success of their overall investment program. This is where total fund benchmarking comes into play.

There are a variety of methods that institutions use for developing total fund benchmarks. The different types of benchmarks reflect different criteria and philosophies for benchmark construction and evaluation. Yet, practical implementation among large institutional portfolios has led to several commonalities and trends in recent years. The following paper explores the various trends and challenges among total fund benchmarking practices.

The importance of benchmarks

In the broadest sense, a benchmark is a standard against which the performance of a portfolio is measured. Benchmarks allow fiduciaries to assess the success of their investment programs and help them understand allocation decisions as well as risk and return profiles. In other words, fiduciaries can use benchmarks to not only measure portfolio out/underperformance, but also understand whether this relative performance is to be expected and what it signifies moving forward. Therefore, a metric such as a benchmark can be extremely useful if it feeds back into the decision-making process.

Benchmark criteria

There are two widely accepted schools of thought for determining benchmark criteria: 1) the six characteristics outlined in the Bailey Criteria, and 2) the five characteristics from the CFA Institute. Both criteria have overlapping concepts.[1]

Additionally, the Global Investment Performance Standards ("GIPS®"), set forth by the CFA Institute, provide voluntary guidelines used by investment management firms throughout the world and include language around benchmarks. Within the GIPS fundamentals of compliance,[2] a benchmark is defined as a point of reference against which the composite or pooled fund's returns or risk are compared.

[1] Details of the characteristics pertaining to the Bailey Criteria and CFA Institute can be found in Appendix A.

[2] Global Investment Performance Standards for Firms, 2020.

According to these guidelines, a composite or pooled fund benchmark used in a GIPS composite report must reflect the investment mandate, objective, or strategy of the composite. The firm must not use a price-only benchmark in a GIPS composite report. For reference, a price-only benchmark (also referred to as a price return index), as opposed to a total return benchmark, captures only the capital appreciation aspect of index constituents and ignores the dividend payment component.

It is critical to note that many commonly used benchmarks lack one or more of these characteristics, and thus, the policy benchmark, made up of asset class benchmarks, will never be a perfect comparison for an institutional fund's diversified asset allocation. Strict adherence to these criteria are not required and may not be practical in the real world for the various kinds of evaluations that Trustees have to make, but they do pose a viable starting place and foundation to build out an applicable set of representative benchmarks.

Variety of total fund benchmarks[3]

[3] Total Portfolio Benchmarking, Meketa, 2019.

To identify commonalities and issues in total fund benchmarking practices, it is important to highlight the various composition methods of a total fund benchmark.

Total fund 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.

Generally, plan level benchmarks fall under three categories: 1) plan level (e.g., simple, static policy, dynamic), 2) peer group, and 3) institutional financial objectives (e.g., target returns).

Plan level benchmark | There are three generally accepted plan level benchmarks. While each has its pros and cons, it is common practice for institutional clients to utilize multiple benchmarks at the total plan level.

1. **Simple benchmarks** are typically based on a predetermined mix of relatively few public market investments, rebalanced at regular intervals (e.g., 60/40 stock/bond portfolio). The allocation determined for the simple benchmark may be based on a target or historical rate of return (commonly measured over a 20-year time horizon) or risk tolerance (could be measured by an annualized standard deviation figure).

2. **Static policy benchmarks** are typically based on an institution's policy asset allocation mix. Static policy benchmarks generally use a passive index for each asset class, weighted the same as the targets in their asset allocation policy, to calculate a portfolio's return as if it were passively invested at policy targets. It is important to note that this approach only applies if broad market indexes are applied across the board and exclude the utilization of real return or peer fund benchmarks.

This comparison includes not only the effect of active management versus a passive index but also incorporates the effect of having allocations different from an institution's targets. These differences may be deliberate, reflecting tactical positioning, or unintentional, such as when a portfolio drifts from its targets due to normal market movements and the institution either chooses not to, or is unable to, rebalance (especially in relation to illiquid asset classes). The asset class benchmarks are generally one of the following types of benchmarks depending on the asset class type:

- broad market (i.e., MSCI ACWI IMI index; Bloomberg Aggregate index);
- real return target (i.e., CPI + 3%); and
- peer fund universe primarily utilized for private markets (e.g., NCREIF ODCE; Cambridge Associates Private Equity index; HFRI Fund of Fund index).

3. **Dynamic policy benchmarks** use the actual weights of each asset class and uses passive indices to calculate the passive equivalent of the return achieved by the total portfolio. Thus, it is designed to focus on the cumulative added value solely from active management in an attribution analysis, whether that be in-house or outsourced to asset managers.

To summarize what these benchmarks measure:

Peer group | Institutional peer groups are a common method of total portfolio benchmarking that can be used when investors wish to compare performance to that of similar investment pools by type and size. The data included in peer groups is typically collected by vendors and published on a quarterly basis. While this comparison can be meaningful when using peer groups with significant membership, the utility declines as the sample size of the peer group declines. At the same time, a large peer group is likely to be composed of a diverse set of institutions that may differ widely in objectives (e.g., target return), resources (e.g., institutional staff, access to capacity-constrained managers, etc.), and constraints (e.g., liquidity tolerance or financial situation).

Additionally, peer comparisons may be difficult to obtain, are often substantially lagged, and are subject to reporting issues (fee netting / asset classification variances), making them only marginally useful.

Institutional financial objectives: target returns | Institutional financial objective benchmarks compare realized returns against target returns (e.g., annual spending rate and assumed actuarial rate of return). While this type of benchmark provides the ultimate measure of whether a portfolio is achieving its objective, it is typically disconnected from what the capital markets are delivering over short periods of time (e.g., achieving a target return was all but impossible for most investors in 2008 given the significant downturn in risky assets that year). However, these benchmarks offer useful long-term comparisons, such as over a full market cycle.

Trends in public plan benchmarking

Identifying best practices in benchmarking is a common, though often unsatisfying, topic of conversation among institutional investors.

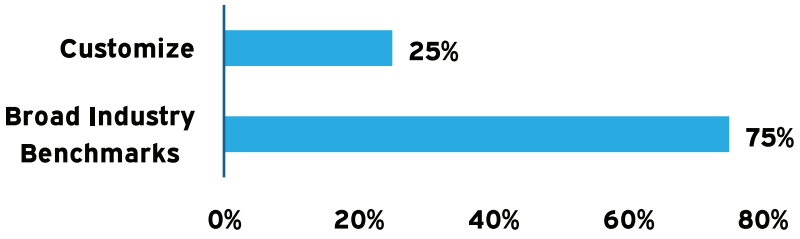
In the second half of 2021, Meketa conducted a survey that aimed to identify themes around the construction and processes of total fund benchmarking and incentive compensation. Participants included nine public pensions that were in the top 30 of the largest North American pension systems, with assets under management ("AUM") ranging from $64 billion to $750 billion. The survey asked five questions that address three themes: types of benchmarks, approval of benchmarks, and incentive compensation. Results from Meketa's 2021 Peer Benchmarking survey are detailed below.[4]

[4] Note that not every survey participant responded to each question. In some cases, eight of the participants responded, and in others, five of the participants responded.

Total fund benchmarking construction/composition

QUESTION 1

Does your organization customize benchmarks for plan-specific constraints/ preferences, or use broad industry benchmarks at the total fund level?



QUESTION 2

Are there multiple benchmarks at the total fund level?

Of the respondents, 75% used a broad-based approach to benchmarking at the total fund level, as opposed to customizing for plan-specific constraints. Broad-based benchmarks typically reference the performance of a wide opportunity set for a basket of investments and are more common across institutional portfolios. For example, the MSCI ACWI is a common index within the composition of a

static/dynamic total fund benchmark that would represent the opportunity set of global equity investments. On the other hand, if an investor is constrained from investing in such a broad opportunity set – for example, via the intentional exclusion of tobacco or fossil fuel companies – then selecting a benchmark that likewise excludes these stocks would be better aligned with the plan's objectives.

Regarding the number of benchmarks used, 63% of respondents implemented a single benchmark. In most cases, a static or dynamic policy benchmark was utilized to evaluate the portfolio's execution of a stated asset allocation policy and underlying asset class manager selection.

The CFA Institute conducted a case study in June 2021[5] that demonstrated the argument for utilizing a static/dynamic policy benchmark by comparing the California Public Employees' Retirement System's ("CalPERS") total fund rate of return with that of its static policy benchmark and an investable passive stock/bond split (79% US and non-US stocks and 21% US investment grade bonds). See Figure 1 below.

[5] https://blogs.cfainstitute.org/investor/2021/06/07/institutional-portfolio-benchmarks-slow-rabbits/

Fiscal Year Ending	CalPERS Total Fund (%)	Custom Benchmark (%)	Difference (%)	Passive Benchmark (%)	Difference (%)
2011	21.7	21.8	-0.1	23.6	-1.9
2012	0.1	0.7	-0.6	2.2	-2.1
2013	13.2	11.9	1.3	13.8	-0.6
2014	18.4	18.0	0.4	18.6	-0.2
2015	2.4	2.5	-0.1	3.8	-1.4
2016	0.6	1.0	-0.4	1.4	-0.8
2017	11.2	11.3	-0.1	13.3	-2.1
2018	8.6	8.6	0.0	9.2	-0.6
2019	6.7	7.1	-0.4	7.5	-0.8
2020	4.7	4.3	0.4	5.5	-0.8

FIGURE 1
Case Study Results for CalPERS Total Fund Return and Benchmarks
Source: https://blogs.cfainstitute.org/investor/2021/06/07/institutional-portfolio-benchmarks-slow-rabbits/

Annualizing this data over the 10-year time period, CalPERS' portfolio return (+8.54%) was within three basis points of the static policy benchmark (+8.51%) and lagged the investable passive benchmark by -1.14%. When measuring the correlation (or statistical fit) over the same time period, the total portfolio had a higher fit to the static policy benchmark (.995) than to the investable passive option (.991), though both correlations are quite high.

Some may argue that the implementation of static policy benchmarking results in the portfolio "hugging the benchmark." According to the CFA Institute's case study,

"Institutional portfolios often exhibit close year-to-year tracking with their static policy. This results in part from how static policy benchmarks are revised over time. Sometimes revisions are motivated by a change in asset allocation, which may warrant adjusting the benchmark. Often, though, the revisions are more a matter of periodically tweaking the benchmark to more closely match the execution of the investment program." This revision process can be viewed as a reduction of the value of the benchmark as a performance gauge because the more a benchmark is tailored to fit the process being measured, the less information it can provide. The CFA Institute argues that, at some point, an overly tailored benchmark ceases to be a measuring stick altogether and becomes a mere shadow.

The case study illuminates an important notion when it comes to choosing a total fund level benchmark:

> in most cases, it is more of an art than science to construct and maintain, and there are a wide array of approaches to benchmarking

as each investor has a unique situation and characteristics surrounding determining factors such as funding and resources (e.g., staff, ability to manage complexity, knowledge and experience of the Board).

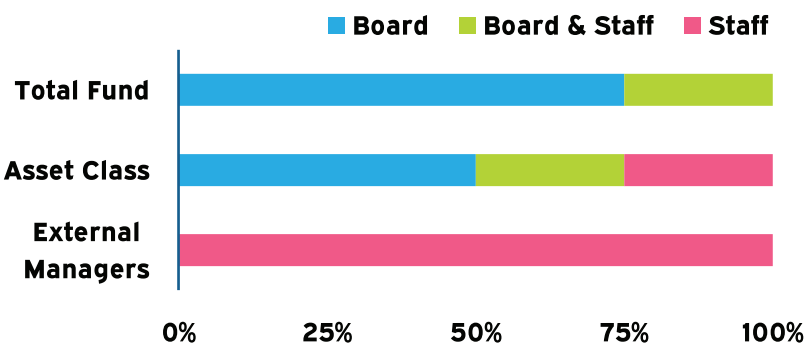
QUESTION 3
Are benchmarks reviewed on a periodic basis or in line with asset allocation reviews?

Regarding the ongoing oversight and maintenance of benchmarks, 63% of respondents typically re-evaluated their benchmarks in conjunction with their asset allocation reviews rather than on some other periodic basis.

Approval of benchmarks

QUESTION 4
Who approves the addition/modification/ removal of benchmarks at the various portfolio levels (total fund, asset class, external managers)?

Survey Results of who Approves Different Types of Benchmarks



On the topic of approval and adoption of additions/modifications/removals of benchmarks at the various portfolio levels, responses generally aligned with expectations associated with the knowledge-based requirements for each category. The public pension respondents' staff (denoted in yellow in Figure 3) unanimously approve benchmarks for the external managers. Justification stems from the fact that investment staff typically have the highest interaction with external managers and monitor both qualitative and quantitative aspects of the individual strategies on a continuous basis. Another reason for this approach relates to the increasing size of manager rosters across large public pension systems. In many cases, these portfolios have a high number of strategies that would entail a heavy lift for the Board to remain informed/educated on. With a large portfolio typically comes a large staff that has the ability to approve benchmark changes without Board consent.

There were mixed results for benchmarks at the asset class level as some plans delegate authority to staff; however, the survey did not distinguish between major and sub-asset classes, so the 50% share to the Board (in blue) could very well represent a delegation for a sub-set of the asset class grouping. In some cases, Boards maintain authority over the major asset classes and delegate sub-asset class benchmarks to staff for a variety of reasons (complexity, confidence in staff, increasing breadth and depth). Similar to external managers, the larger public systems with large investment staffs (teams of greater than 20 staff members) tend to divide their staff into teams of asset class specialists (instead of generalists). This likely gives the respective Boards confidence that their staff can independently identify and implement the appropriate benchmarks without Board approval. Additionally, with consultant oversight, Boards have additional checks and balances in place to further build confidence around delegation.

At the total fund level, the Board was involved in the benchmark approval process for all respondents. This is in line with the Board's purpose at a typical retirement system, which requires the approval of the system's asset allocation and investment policy, including the total fund benchmark.[6]

[6] https://www.nasra.org/files/Topical%20Reports/Governance%20and%20Legislation/NASRA%20governance%20overview.pdf

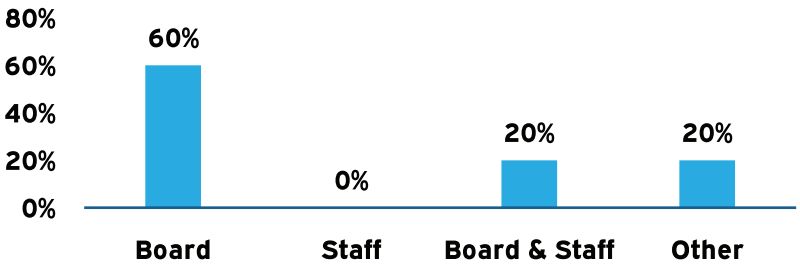
Incentive compensation

One of the growing topics across large institutional portfolios is the implementation of incentive compensation for investment staff members. Incentive compensation is a form of variable compensation in which an investment staff member's earnings are directly tied to the performance of some combination of the individual, team, and total portfolio. This additional component of pay is evaluated versus a benchmark that can contain both quantitative and qualitative components and are measured over a prescribed time period. The approval, implementation, and composition of these benchmarks vary across institutional plans and was a focus of the survey in order to identify trends and commonalities. to

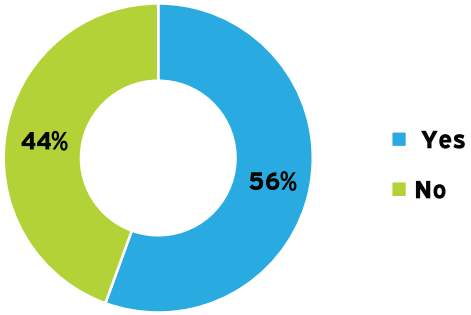
QUESTION 5

Who approves the implementation and periodic maintenance of the incentive compensation benchmarks?

Note: "Other" includes compensation consultants or other departments internally.

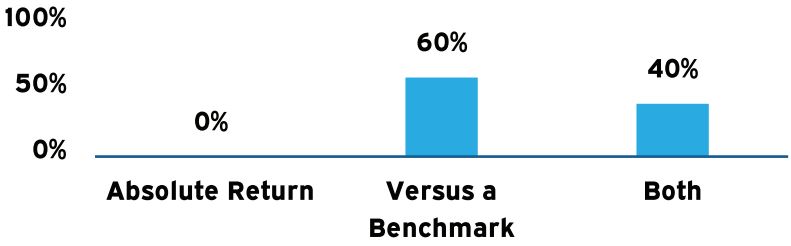


Do you use incentive compensation for staff?

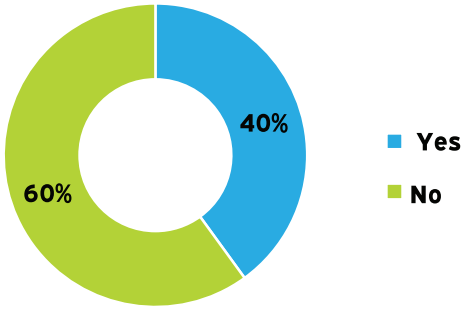


How do you determine incentive compensation for staff?

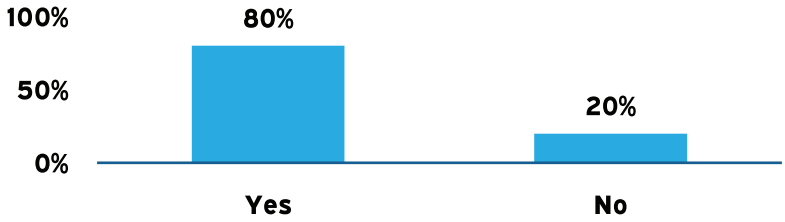
Within the 56% of respondents that utilize incentive compensation, 60% evaluate performance versus a benchmark, and the other 40% implement a combination of a benchmark and absolute return component when determining compensation.



Is incentive compensation based on a reference portfolio?



Do you have a compensation consultant to assist with structure and maintenance?



The majority of respondents (80%) employed a compensation consultant to help design the structure and oversee ongoing maintenance of incentive compensation. Likewise, the Board was involved in the approval process for 80% of respondents.

In many cases (and particularly with large pension systems), the composition of these specific benchmarks is highly complex, is not just quantitative, and is customized to each asset class. Some benchmarks compare just the level or return, while others may look at a risk-adjusted return. When looking under the hood of an incentive compensation benchmark, there could be multiple components with varying weights. Specifically, pension systems could include a qualitative component that is specific to either the individual or team as well as quantitative aspects that could apply to both the total fund and asset class levels.

As an example of the nuances across a portfolio, incentive compensation for a newer asset class may place a higher weight on a qualitative component given the short performance track record and incomplete manager roster. An asset class designed to provide downside protection (such as a risk mitigation program) would not necessarily implement a risk-adjusted return for the quantitative component of the benchmark (which would typically be used for a public equity or fixed income aggregate) and could instead utilize a relative return metric. These nuances underscore the complexities associated with compensation evaluation and illuminate the rationale behind why the majority of survey respondents lean on outside compensation consultants for implementation and oversight.

Common dilemmas in public plan benchmarking

The creation and maintenance of public plan total fund benchmarking does not come without its dilemmas. Whether it be through intentional shifts in the structure of the portfolio, inherent challenges with private market performance evaluation, or desire to benchmark against peers, each situation can be tackled in a variety of ways.

Benchmarking a Transitioning Portfolio

As investors adopt new asset allocation policies, a common issue/dilemma that arises is performance evaluation over the period it takes to transition the portfolio's assets (closer) to the newly approved policy targets. For significant shifts in allocations, across both public and private asset classes, the movement of assets does not occur overnight. Not only is this attributable to liquidity constraints among underlying strategies but also the time it takes for thoughtful reallocation among these managers thus underscoring the notion of asset allocation transitions being more complicated to implement than the average Trustee might assume. To mitigate harm to the portfolio and respective returns, investment teams and consultants often conduct thorough analyses to determine and implement a methodical transition of the assets. The transition period can range from a few quarters to several years, depending upon the investor's AUM and asset allocation mix (e.g., transitions for private market assets tend to be the lengthiest, given the nature of these assets).

Over the course of the asset allocation transition process, an important consideration is how to proceed with performance evaluation against a total fund benchmark when a static policy is utilized. If the composition of the static policy benchmark is adjusted to the new policy targets at the time of asset allocation approval, the ensuing comparable performance metrics may not be reflective of the actual opportunity set and misrepresent out/underperformance.

- **Dynamic policy benchmark[7] |** One option is to implement a dynamic policy benchmark during the period when assets are being transitioned. The underlying benchmark components mirror the pre-existing policy composition but are weighted using actual portfolio asset values, thus muting the allocation effects of performance and focusing on underlying manager selection. A drawback of this approach is the limitation associated with not knowing whether the implementation process is on track. If there is an approved implementation track, it will not be reflected in a dynamic benchmark.

- **Phased policy benchmark[7] |** A phased policy benchmark is a pre-scheduled "step-up" approach that aims to align the composition of the policy benchmark with the transition of assets and ultimately ends up at the newly approved static policy targets. The benchmark would likely go through a couple iterations that mirror major moves in actual asset movements and would provide a comparable performance metric that highlights both manager selection and allocation effects. A drawback of this approach is the unanticipated delay in asset movements causing the unintentional misrepresentation in the allocation effects when reviewing performance attribution. This delay could also have a domino effect on the future phases of the policy benchmark.

- **Multiple benchmarks[7] |** In an effort to track the transition of assets, isolate manager performance, and/or view allocation effects in performance attribution, an investor could utilize multiple benchmarks at the total fund level for the purposes of viewing performance from various lenses during the asset allocation transition.

[7] Examples for each of these approaches can be found in Appendix D.

Benchmarking Private Market Assets in the Total Fund Policy Benchmark

The challenges of benchmarking a portfolio of private market assets have been prevalent since investors first started investing in venture capital. The highest hurdles for implementing a private markets benchmark include access, fit, and timeliness. The benchmarking options for private markets have been grouped into three main categories: peer benchmarks, public equity benchmarks, and inflation-linked benchmarks.

- **Peer benchmarks |** Several service providers, including Cambridge Associates ("Cambridge"), Burgiss and Preqin, aggregate industry fund performance to calculate a peer benchmark return typically broken out by asset class, geography, and/or vintage. The peer benchmark's quarterly "pooled returns" are generally used to calculate a trailing time-weighted return.

The biggest advantage of using peer benchmarks is the compatibility or the degree to which the benchmark matches with the investor's objectives, risk preference, and portfolio composition. An investor can use peer benchmarks to evaluate the success or drawbacks of their private markets portfolio through various market cycles relative to the opportunity set that was available to them. However, accessing peer benchmarks can be costly, the vintage and type constituency is often unknown (calling fit into question), and data arrival is very late (up to six months past period), which causes lagged or incomplete data. Further, the lag in reporting for private markets can present additional challenges. For example, any gap between two parties (e.g., custodian versus manager/staff) recognizing a change in valuation can result in discrepancies for return calculations.

→ **Public equity benchmarks |** In addition to or in lieu of peer comparisons, investors can use public indexes that are often similar to those used for benchmarking public equity portfolios. A "spread" is typically added to the public benchmark's annual return to account for the higher return that is typically anticipated for private markets investments. Each private markets asset class has a different public index and spread depending upon the risk and return profile.

Meketa believes the three largest benefits to an institutional investor utilizing a public equity benchmark to be the following:[8]

[8] Private Market Benchmarking, Meketa 2022.

- Timeliness | Returns can be calculated at month-end, and portfolios are not subject to the time lag present when utilizing a peer benchmark.
- Clarity | Public equity benchmarks provide perfect clarity into assets and weightings of all assets included in the index.
- Cost | As opposed to peer benchmark service providers that charge a notable fee for access to peer benchmarking information, public equity benchmarks can be constructed and maintained at little to no cost.

The largest drawbacks to implementing a public equity benchmark are tracking error and the notion of compatibility with the portfolio's respective private markets investment opportunity set. Private markets portfolios are likely to exhibit significant tracking error from public markets benchmarks in the short term. However, over long periods of time, the use of public markets benchmarks can help fiduciaries answer questions such as whether the choice to invest in private markets has been worthwhile. Compatibility refers to the degree to which the benchmark matches with the investor's objectives, risk preference, and portfolio composition. Public benchmarks are unable to mirror the unique dynamics of private investments.

→ **Inflation-linked benchmarks |** In some cases, investors choose to use a benchmark that is linked to the inflation rate (typically based on the consumer price index ("CPI") observed over a matching trailing period. This method is most commonly used for benchmarking the performance of private real assets portfolios. A spread (commonly 300-500 basis points) is also commonly added to the inflation benchmark. Advantages of implementing an inflation-linked benchmark include:

- Clarity | The subcomponents of the inflation benchmark (e.g., CPI) are publicly available.
- Timeliness | Returns can be calculated either at month-end (if the institution is willing to accept a one-month lag) or a couple weeks after the end of the month.
- Cost | The use of an inflation-linked benchmark is free.

Disadvantages associated with the utilization of an inflation-linked benchmark are the lack of investability and compatibility. From an investable standpoint, investors do not have the option of earning the return of an inflation benchmark. Though an inflation-linked benchmark can be modified (via adjusting the spread) to better match a desired risk profile, compatibility-wise, it cannot capture equity/credit returns over a given period.

Asset Class	Benchmark
Private Equity	MSCI ACWI IMI + 120 – 300 bp
Private Credit	Credit Suisse Leveraged Loans + 0 – 250 bp or Bloomberg High Yield Index + 0 – 250 bp
Real Estate	NCREIF ODCE EW Net (one quarter lagged) + 0 – 150 bp, DJ Global Select RESI + 0 – 150 bp, or CPI + 300–500 bp
Natural Resources	S&P Global Natural Resources + 0 – 200 bp or CPI + 300 – 500 bp
Infrastructure	DJ Brookfield Global Infrastructure + 0 – 150 bp or CPI + 300 – 500 bp
Real Assets	CPI + 300 – 500 bp or blend of underlying asset class benchmarks (NR, IS, RE)

FIGURE 2
Common Public Benchmarks for Private Markets Asset Classes
Source: Meketa Investment Group.

Composition and calculation of the policy benchmark

When employing a static policy benchmark at the total fund level, inclusion of underlying constituents is a common dilemma for large and complex institutional portfolios. As a starting point, most portfolios mirror the asset classes reflected in the Investment Policy Statement ("IPS"); however, several issues can arise in practice.

The first dilemma is the mismatch between the benchmark opportunity set reflected in the asset class benchmarks (as outlined in the IPS) and the sub-asset class benchmarks that may not be included. For example, if a fixed income asset class is benchmarked to the Bloomberg Aggregate index, but the portfolio has allocations to sub-asset classes such as high yield bonds and bank loans that are not included in the Bloomberg Aggregate, there will be a mismatch as the benchmark does not reflect the opportunity set of the actual portfolio investments.

One way that portfolios have overcome this issue is by creating a custom asset class benchmark that is inclusive of the sub-asset class benchmarks. Potentially outlined in the IPS or determined separately, sub-asset class target allocations could be implemented in the calculation and weighted according to the target implementation to each.

For example, the fixed income custom benchmark could be based upon a 50% weight to the Bloomberg Aggregate, a 30% weight to the Bloomberg High Yield index, and 20% to the S&P Leveraged Loan index. In the event that sub-asset class target weights are

not defined or the portfolio is in transition (e.g., phasing one of the sub-asset classes in or out), the fixed income custom benchmark could be dynamically weighted based upon actual asset values. However, this does segue into the second dilemma that arises in the composition of a total fund policy benchmark.

Employing a dynamically weighted benchmark based upon actual sub-asset class weights for an asset class but utilizing static weights in the calculation of the total fund policy benchmark does create a mismatch. When using dynamic weights, the allocation effects in an attribution analysis are not considered, and manager performance is isolated as the only factor affecting relative results. When evaluating overall performance at the total fund level, the dynamic weights of the sub-asset class benchmarks roll up into the calculation of the static policy benchmark and could create difficulty when assessing the implementation of a policy. In turn, the investor's ability to implement and execute a policy cannot be determined based off relative performance.

The inconsistent methodology throughout the components of the policy benchmark may or may not be an intended design of the portfolio. If intended, it should be carefully noted when looking at performance metrics and thus taken into consideration when conducting an attribution analysis.

Utilizing peer benchmarks at the total fund level

In addition to existing total fund benchmarks (whether they be static/dynamic policy, target return or simple stock/bond split), some institutions (especially public funds) use peer group benchmarks as an additional metric. A peer group is usually defined in a couple different ways:

- **Plan type** | public defined benefit, endowments, foundations, Taft-Hartley, defined contribution, etc.
- **AUM** | Any threshold of assets; for example: > $5 billion, between $1 billion - $5 billion, < $50 million
- **Combination of plan type and AUM** | Example: public defined benefit plans > $100 million

Benchmarking a portfolio to a peer group is a means for comparing to institutions with presumably similar resources and objectives. Peer institutions are navigating the same economic and market environment, could be enduring similar dilemmas, or capitalizing on similar performance drivers. A peer benchmark can also shed light on commonalities and differences in asset allocations. For example, in a bull equity market, it would not be surprising to see that the majority of portfolios that rank in the top quartile (i.e., top 25% percent) had a higher-than-average allocation to public equities.

There are a handful of challenges with peer group benchmarking, notably the lack of transparency and the presumption of similar objectives and resources. In reality, pension plans of similar size may have very different target rates of return and tolerance for risk. Likewise, two similar sized endowments may have very different governing bodies, cultures, staff, access, and resources. Efforts to deduce these details will likely be unsuccessful, as details of the underlying constituents are typically limited to the plan type and AUM, trailing returns and asset allocations are aggregated, and individual

identifiers such as plan name are anonymized. There is also an inconsistency across the variety of peer group providers (e.g., Public Defined Benefit Plans > $1 Billion peer groups can be offered by Investment Metrics, State Street or Wilshire Trust Universe Comparison Service ["TUCS"]) and often exhibit a wide dispersion in performance and portfolio information as the underlying constituents vary and cannot be identified given the lack of transparency. For instance, some peer groups are smaller than others, composed of large plans with larger than normal private markets allocations, and report mostly gross of fees.

Given the fact that most peer universes are self-reported, fee netting (or lack thereof) is not verified. If a public plan evaluates performance net-of-fees but the majority of the constituent data in the peer group is reported gross-of-fees, performance evaluation may be misleading and likely cannot be diagnosed due to the lack of transparency. Availability is another common dilemma as universe data is typically released 30-60 days after a quarter-end. This lag is due to the time needed for underlying universe members to report quarterly performance information to the universe provider. For this reason, peer group benchmark availability for inclusion in performance evaluation is dependent upon release dates and portfolio deadlines, or it may need to be reported on a lagged basis.

Conclusion

Benchmarking a portfolio at the total fund level has numerous considerations, iterations, and dilemmas that could leave one wondering, "Is this the best benchmark to use?" The first overarching takeaway from this paper is the notion that total fund benchmarking is more of an art than science. There is no single "right way" to evaluate performance. Moreover, the benchmark used today is unlikely to remain the same into perpetuity. The opportunity set of an investment portfolio is ever-changing, and objectives will evolve as a portfolio grows and matures. As such, the benchmark(s) will need to be updated to adapt to the opportunities and objectives at hand.

The second takeaway is that it is common to have multiple benchmarks at the total fund level, as various benchmarks/compositions express different things about performance and serve as guideposts for multiple objectives. Employing multiple benchmarks could provide a more well-rounded lens, whether it be for comparisons versus a peer group, a target return hurdle, a simple stock/bond split or policy allocation.

The third takeaway is that it is prudent to regularly review total fund benchmark(s). Whether it be in line with the asset allocation review or at some other regular interval, reviewing the composition and appropriateness of the benchmark will better serve the total portfolio and all parties involved.

Finally, the composition of each benchmark should be understandable, and the role of each should be clear to the plan's fiduciaries. Consistent oversight and effective utilization are crucial components of successful benchmarking.

Appendix A | Benchmark criteria

Bailey benchmark characteristics:[9]

[9] Source: Financial Analysts Journal, CFA Institute, 1992

- → **Unambiguous |** The components in a benchmark should be clearly identifiable.
- → **Investable |** It must be possible to replicate and hold the benchmark to earn its return (gross of fees).
- → **Measurable |** It must be possible to measure the benchmark's return on a frequent and timely basis.
- → **Appropriate |** The benchmark must be consistent with the manager's investment style or area of expertise.
- → **Reflective of current investment option |** The manager should be familiar with the securities that constitute the benchmark and their factor exposures.
- → **Specified in advance |** The benchmark must be constructed prior to the evaluation period so that the manager is not judged against benchmarks created after the fact.

CFA benchmark characteristics:[10]

[10] Source: CFA Institute

- → **Investable |** It is possible to forgo active management and simply hold the benchmark. That is, investors can effectively purchase all securities in the benchmark.
- → **Accessible |** Difficult-to-produce benchmarks should be avoided.
- → **Transparent |** Understanding the underlying constituency of a benchmark is critical to understanding its suitability for a particular manager.
- → **Independent |** A manager's performance should not impact the prescribed benchmark return.
- → **Relevant |** Spurious correlation exists between many random sets of data over various time periods. High correlation or low tracking error to a particular benchmark is not enough to conclude the benchmark is appropriate for a particular manager.

Appendix B | Private market benchmark service providers

Cambridge is a privately owned firm that provides investment advice and research, including private equity consulting services, to institutional investors and private clients. The firm publishes performance indices and statistics covering approximately 8,500 global private equity funds[11] (with services dating back to 1981). Information is sourced from Cambridge consulting relationships (endowments, foundations, public and private pension funds, family offices, and other small institutions) and from self-reporting private equity managers. Benchmark data is published approximately 85-90 days after quarter-end, and there is no detail provided on underlying funds in the benchmark.

Burgiss is a privately owned firm that provides software applications for monitoring, reporting, measuring performance, and benchmarking investments in alternative assets. Burgiss tracks approximately 10,400 private equity funds[11] (dating back to 1978), and sources information from limited partner cash flows and valuation data from its client base. Data is contributed from a variety of institutional investors, including public pension funds and corporate pension funds, endowments, foundations, and advisers. One challenge exhibited by clients that utilize this provider for peer benchmarking purposes is the publishing timeframe for quarterly data, which is approximately 75-90 days after quarter-end. For clients that receive monthly/quarterly performance reports and/or have monthly/quarterly investment committee meetings, the 75-90-day time frame typically exceeds the deadline necessary for inclusion. Additionally, there is limited detail provided by Burgiss on underlying funds in the respective benchmarks utilized (i.e., constituents).

Preqin is an independently owned provider of information and intelligence for the alternative assets industry. The provider tracks approximately 7,400 private equity funds[11] with operations dating back to 1980. Unlike Cambridge and Burgiss, information is sourced primarily from public institutions via Freedom of Information Act (FOIA) requests as well as from self-reporting private equity managers. There are a few drawbacks associated with this method of information gathering. The first is capitalization amounts for private equity data is not provided; as such, some composition data is unavailable. In addition, Preqin receives cash flow data for approximately one-third of the funds in its quartile benchmarks, which results in a different and much smaller universe for its pooled benchmarks. Benchmark information is published approximately five months after quarter-end (the longest of all service providers), posing the biggest challenge for utilizers of this information.

[11] As of 2021.

Appendix C | Private market benchmarks using public market indices and spreads[12]

[12] Private Market Benchmarking, Meketa 2022.

Private Equity: MSCI ACWI IMI + 150-300bps

MSCI ACWI IMI is suggested as a widely used global public equity benchmark. The Russell 3000 is not optimal for typical private equity portfolios (with meaningful international exposure) but may be a good option for clients almost entirely focused on domestic exposure within private equity portfolios or if benchmarking domestic private equity exposure separately.

Private Credit: Credit Suisse Leveraged Loans + 0-250bps or Bloomberg High Yield Index + 0-250bps

Real Estate: NCREIF ODCE EW Net (one quarter lagged) + 0-150bps, DJ Global Select RESI + 0-150bps, or CPI + 300-500bps

As an aggregation of open-end private real estate fund returns, the NCREIF ODCE is not a public equity benchmark but has advantages of peer benchmarks due to the transparency into underlying funds comprising the index as well as being more investable than peer return benchmarks. Additional spread is recommended as NCREIF ODCE represents core exposures and will be used to benchmark portfolios of both core and non-core exposures.

Natural Resources: S&P Global Natural Resources + 0-200bps or CPI + 300-500bps

Infrastructure: Dow Jones Brookfield Global Infrastructure + 0-150bps or CPI + 300-500bps

Real Assets: CPI + 300-500bps or blend of underlying asset class benchmarks (NR, IS, RE)

As with each of the underlying real asset strategies, an inflation benchmark may be more appropriate for some portfolios, while a weighted public equity benchmark will work best for others depending on whether the primary objective of the real assets program is real returns or outperformance of public equity alternatives. That said, a CPI (+ spread) benchmark is the most common real assets benchmark among institutional allocators.

Appendix D | Examples of benchmarking to a transitioning portfolio

Dynamic benchmarking | A large public pension that prefers to only have one benchmark at the total fund level adopts a new asset allocation mix that materially shifts the policy targets for the majority of the asset classes. The anticipated time frame for this transition is 18 months and will occur gradually. Currently the benchmark at the total fund is a static policy benchmark that utilizes underlying asset classes and their respective target policy weights.

During this transition, the Board and staff decide not to benchmark the portfolio to the newly adopted policy targets due to the fact that relative performance will be indicative of how assets are invested throughout the 18 months. Instead, a dynamic total fund benchmark is implemented that uses the actual weights of the asset classes at the beginning of each month and isolates underlying strategy performance within each asset class. Once assets are fully transitioned, the policy benchmark will change back to a static composition and will utilize the new policy targets into perpetuity.

Phased benchmark | A public pension plan that currently utilizes a static policy benchmark adopts a new asset allocation that incorporates sizeable increases in private equity and private real assets. Given the portfolio's pacing plan for each of these asset classes, the implementation process will take several years for the portfolio's allocations to align with the "long-term" targets. As a way to evaluate performance over the next few years, "interim" targets are adopted to illustrate the transitioning nature of the portfolio as capital is committed and called in a programmatic fashion and used to create "interim" static policy benchmarks.

Asset Class	Current Exposure	Interim Targets	Long-Term Targets
Public Equity	50	48	45
Private Equity	5	7	10
Public Credit	10	8	6
Investment Grade Bonds	18	17	15
TIPS	5	5	5
Real Estate	5	7	10
Natural Resources	4	4	4
Infrastructure	3	4	5

FIGURE 3
Example Interim Target Structure

Source: Meketa Investment Group.

Note: Does not represent an actual investment portfolio. Asset allocation mix and target percentages created for illustrative purposes.

Once the private asset classes have grown beyond the "interim" policy targets and closer to the "long-term" targets, the asset allocation and policy benchmark are re-evaluated to determine whether a new "interim" allocation should be implemented or if the "long-term" policy targets can be fully adopted.

Multiple benchmarks | Similar to the example illustrated above with phased benchmarks, a public pension plan that utilizes a static policy benchmark (encompassing of underlying asset classes and respective targets) adopts a new asset allocation that incorporates sizeable increases in private equity and private real assets. Given the portfolio's pacing plan for each of these asset classes, the implementation process will take several years for the portfolio's allocations to align with the "long-term" targets. In the meantime, multiple benchmarks are utilized to identify and evaluate different objectives. One benchmark is denoted as the "tactical" benchmark that mirrors the phased benchmark illustrated in the prior example. The other metric implemented is a "strategic" benchmark that allows the Board to see what the portfolio's returns might look like when the private markets pacing plans are fully executed. This benchmark is the static "long-term" policy.

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.