Private Equity Benchmarking for Asset Owners and Investment Managers

Emilian Belev and Dan DiBartolomeo

Emilian Belev
is head of ERM Analytics at Northfield Information Services Inc. in Boston, MA.
emilian@northinfo.com

Dan DiBartolomeo
is president of Northfield Information Services Inc. in Boston, MA.
dan@northinfo.com

KEY FINDINGS

- Analyze the ways the investment industry has evolved to tackle the challenges with private equity benchmarking in practice.
- Reduce the reliance on subjective measures like appraisal values and heuristic groupings, and focus on objective measures like cash flows, transacted values, and efficient benchmark portfolios.
- Evolve beyond static measures of performance to those that suggest the confidence intervals that differentiate between skill and luck.

ABSTRACT

This works focuses on the challenges posed by benchmarking private equity performance and the solutions that have evolved in practice. The paper puts the practical approaches used by investors into four general categories, and describes the relevant characteristics that distinguish each category, the typical circumstances in which they normally arise, and the areas in which each presents opportunities for improvement. The authors' perspective is neutral as to which method is preferred or superior, recognizing that each one presents particular advantages and challenges. The intent is to demonstrate the relationship among the different approaches that eventually should rest on similar underlying principles of objectivity, efficiency, and transparency. The intended audience includes institutional asset owners with significant allocations to private equity, that need to adequately measure the value added to their portfolio by direct and fund investment, as well as investment managers who need to track how the performance of their strategies and implementations compares to that of their peers.

TOPICS

Private equity, performance measurement*

*All articles are now categorized by topics and subtopics. **View at PM-Research.com**.

"One must not allow one's attitude to securities which have a daily market quotation to be disturbed by this fact or lose one's sense of proportion. Some Bursars will buy without a tremor unquoted and unmarketable investments in real estate which, if they had a selling quotation for immediate cash available at each Audit, would turn their hair gray. The fact that you do not [know] how much its ready money quotation fluctuates does not, as is commonly supposed, make an investment a safe one."

—John Maynard Keynes

Benchmarking past performance is a key element to understanding the potential that different assets, strategies, allocations, and managers have to add value to future investment performance. Having done well in the past suggests that given a similar environment, the particular investment choice has a higher chance to perform well in the future. Even if the environment is not expected to persist, observing past performance still can provide insights into how the evolving environment may impact future outcomes. In this sense, being able to identify or construct appropriate benchmarks for each of those cases is essential to the investment management process.

Historically, the concept of benchmarking has multiple disparate and sometimes conflicting purposes. The earliest investment benchmark was the Dow Jones Transportation Average, established in 1885, followed by the DJ Industrial Average in 1986. These performed a journalistic function of informing investors about the financial outcomes of a set of investments of interest that would otherwise be a burdensome task to ascertain. Another purpose of benchmarks is to act as a historic representation of an asset class for the purpose of strategic decisions such as asset allocation.

Benchmarks also act as a fair basis of comparison for managers working within a particular asset class. Using them in this way implies that the boundaries of the opportunity set are sufficiently well defined that we are not allowing spurious syllogisms (i.e., comparing the attractiveness of apples and fruit). There is a long list of academic literature such as diBartolomeo and Witkowski (1997), and Brown and Goetzmann (1997) providing evidence of "gaming" by asset managers to have themselves compared to inappropriate benchmarks and peer groups in order to be perceived favorably.

In the presence of a number of distinct objectives of benchmarking, private equity investors should have a clear understanding, and sometimes, a choice of which ones apply specifically to their case. Therefore, while the goal of this work is *not* to discuss the desirable properties of good benchmarks per se, some level of reflection on this perspective is necessary to understand the requirements posed by private equity as an asset class. First, the benchmark should be such that an investment manager is not put to an impossible task of comparing them to a target that lacks some of the real-world qualities and frictions experienced by anyone engaging in an investment transaction. Furthermore, the benchmark should be transparent, so that information asymmetry does not grant access only to a chosen few, to prevent narrow control of the narrative of the benchmark choice and interpretation. Likewise, the chosen benchmark should be objective, taking into consideration the level and types of risk the investment managers are allowed to take, and the constraints to which they are subject. Where the benchmark may be more conservative in terms of risks it assumes, it should be in order to recognize that there may be (and most likely is) a difference between the broad public perception of the risk involved and that of the manager, who may have unique insights or ways to mitigate such risks. The manager's job is to have unique (and thus confidential) investment insights, while the public prerogative is to have transparency; the comparison of the risks taken therefore should show how the manager's insights add value over the performance of a "naïve" portfolio that is efficient with respect to publicly available information.

Private equity presents challenges in each of these dimensions. While the asset class by itself is certainly investable, with total estimated AUM of trillions of dollars (Deloitte 2019), the investment is subject to long term contractual commitments. Any peer comparisons across performance of funds therefore do not possess the "investable" connotation in the sense of ongoing opportunity cost. Furthermore, any investment in a private equity fund is implemented over a period of time, where capital deployment speed is beyond the control of the investor, and the exits from the investment positions even less subject to the investor choice of time. This low liquidity is an obstruction to the arm's length measurement of a transaction fair value at the usual industry choice of periodicity. This in turn, gives rise of various appraisal based techniques that despite industry-wide efforts led by organizations like the AICPA (AICPA 2019), CFA Institute (CFA 2021), and IPEV Board (IPEV Board 2018) remain part of a largely subjective practice.

Moreover, by virtue of the asset class being "private," data sharing is deliberately limited by both portfolio companies and investment managers, significantly impairing the transparency aspect of benchmarking. Similarly, the scarcity of public information about private equity reduces the ability to construct what modern portfolio theory would consider to be an efficient investment portfolio. That limitation comes from two directions: the breadth of investment choices considered for the benchmark portfolio, and the confidence in the statistical properties of the performance of these investments choices.

Some additional practical challenges of private equity arise. For example, the choices of benchmark for many institutional investors are set by a Board of Trustees or Directors. For governance reasons, such bodies try to elicit a wide and inclusive representation, but do not always require that their members possess the independent investment technical sophistication to make an optimal choice for the investment toolset. Another issue is that there is relatively limited number of vendors of dedicated private equity benchmark indexes, in comparison to public asset classes. Additional complications arising from the vendor practices are various data source limitations. For instance, some service providers that may be a source of fund performance reference data, by virtue of becoming integrated in the communication channels between managers and investors in private equity funds, are effectively limited to the data pool for the specific funds they service on behalf of each investor. In contrast, vendors solely relying on publicly available data available through Freedom of Information Act disclosures by investment institutions tied to government organizations could have a majority of established funds and managers represented in the universe of investments they follow; that in turn may tilt their coverage away from venture capital funds and emerging managers. These investment segments are often the most obscure by definition, and, arguably, in most dire need of benchmarking.

The disparity with public assets in terms of benchmark availability is profound. Due to its transparent pricing and numerous data sources with standard formats and offerings, the collection of public asset classes has been the darling of quantitative investment research, spawning various tools and models that make construction of benchmarks a streamlined task. This also facilitated the establishment of a large number of well-recognized vendors of benchmark indexes with competitive offerings.

Investors and institutions investing in private equity have adapted to all these challenges with a number of approaches, relying on a particular data source or computational technique, or a combination, to overcome the void for readily available benchmark choices.

The following sections make an overview of four broad categories of methods for benchmarking private equity in the investment industry. As can be expected, there are varieties within each category. But to the extent that the features of the specific approach arise in similar circumstances and as a response to similar challenges and constraints, they can be naturally considered closely related, and associated with the same set of best practices and potential areas for improvement.

PRIVATE BENCHMARKS FOR PRIVATE ASSETS

A natural place for investors to look for benchmarks for private assets is at portfolios of other private assets. An example would be a private equity fund whose performance is measured against that of a portfolio of other private equity funds associated by some similar features, such as sector, vintage, strategy, or geographic or industrial focus.

The main advantage of this approach is that comparing likewise investments increases the likelihood that all appropriate benchmark requirements will be met. The choice of constituents of such benchmark portfolios represents a similarity of constraints that is one of the necessary conditions to objectively reference the performance of the asset. Such fundamental similarity does not require technical justification in front of any audience, which in itself reduces obscurity and provides a certain level of transparency.

The usual manner in which benchmarking is performed in this setting is by comparing a chosen measure of average performance over time of an asset or a fund with the analogous figure for a published asset class index or a peer group of funds. A number of vendors (see the overview in Appendix A) provide subscription-based information on the aggregate performance of a universe of private fund they follow expressed as various standard measures: Internal Rate of Return (IRR), multiples to paid-in capital – Total Value to Paid-In (TVPI), Distributions to Paid-in (DPI), Residual Value to Paid-In (RVPI), and others. Most of the same vendors also provide various quantiles for these metrics—median, top, and bottom quartile—of their overall universe as well as for peer groups categorized by underlying assets, strategies, vintages, industry, or geographic focus. An industry-accepted standard measure of excellence of a manager or a fund is to rank above the cutoff of the top quartile for its respective peer group.

A notable difference of this comparison practice and the established benchmark process in public asset classes is the absence of the concept of tracking error (TE). The closest analogue to TE is the persistence with which a manager or a fund ranks in the top quartile historically. The proper measurement of tracking error is important to rigorously determine a manager's over-performance. Due to some of the statistical properties of the time series of return reported by private equity, accurate measurement of tracking error is a challenge. The nature of this problem is discussed later in this section, and we propose some analytical methods to more accurately measure tracking error against a rigorously derived benchmark.

Before focusing on such custom applications, a natural starting point is a discussion on the established commercial sources of benchmark data. To provide some substance, we conducted direct interviews with senior staff responsible for the benchmarking products, and made a summary comparison across some important dimensions of methodology (see Appendix A) for five of the major vendors of private capital benchmark data: Burgiss, Cambridge Associates, Pitchbook, Preqin, and State Street Associates (SSPEI$^{\text{SM}}$).

Several general observations can be made upon inspection of the information provided. First, while each of the vendors covers a representative, or at least substantial, sample of the overall fund universe, none has coverage for significantly more than a half of all 18,000 private funds in existence as of the time this article, as reported in the Congressional testimony by Gary Gensler, chairman of the US Security and Exchange Commission (Gensler 2021).

Secondly, in the absence of exchange driven arms-length transaction pricing familiar from the public markets, each of the vendors has to rely on various parties involved in the private asset investment process to self-report performance metrics. Such metrics include cash flow contributions, distributions, net asset values, IRR, and breakdown of fund portfolios by industry and geography, etc. While the authors do not intend to evaluate the merit of any one data collection approach vs. another or attest to the accuracy of any claims that cannot be tested statistically, it is worth noting some of biases perceived in the industry that may be affecting each of them. For example, the method of collecting direct Limited Partner sourced information, and then aggregating the data from each stakeholder to the fund level, introduces a layer of estimation, and potentially estimation error. Conversely, the method of collecting General Partner sourced information may reflect the potential bias of managers that would like to appear to be performing better than their peers, and use the timing of reporting and other financial figure assumptions at their disposal to effectively create an upward bias in the data. Also, when fund performance data is collected from public disclosures of large institutional Limited Partners, the resulting universe may be limited by the choice of funds following strategies involving a higher degree of secrecy not to fundraise from such investors. Large institutional Limited Partners also may choose to limit their investment in funds run by emerging managers.

Third, the persistence of historically reported data may have an impact on the reliability of the benchmark comparison itself. Most of the vendors restate historical results when such revisions and updates become available. Usually this becomes necessary when certain funds lag reporting for the latest quarter. The extent to which some vendors use automated transaction based feeds, rather than self-reporting, may mitigate the timing impact to some extent.

Finally, one common feature across all vendors is the custom benchmark creation capability. This important aspect of private capital analytics offers an opportunity to overcome some of the limitations of the existing benchmarking practice. The remainder of this section is focused on elaborating on the nature of these limitation and the techniques that can be used to address them.

In general, the custom private benchmark process for private asset is often challenged in practice for its lack of rigor. First, the selection of custom benchmark constituents depends on each one’s data availability. As the strategy or investment focus becomes more specialized, the presence of peer funds that are appropriately equipped with historical data is becoming more and more limited. Then the inclusion of any particular fund in the benchmark becomes less the result of analytical necessity than the aftermath of the chance of having access to a particular data set or vendor, or just being in a strategy that is “followed well.” To introduce even further bias, sometimes access to private funds data has to be permissioned or provided by fund services supplied by fund managers in which investors already invest. This renders the selection of benchmark elements akin to the night search of a lost key in the close vicinity of a street light because it is the area that is the best lit, rather than its proximity to the location of loss.

The potentially shallow pool of benchmark constituents for any combination of underlying, strategy, and vintage underlies an additional problem. An investor should pay for the active risk that a manger takes only when it is offset by superior performance. Having too few constituents in a benchmark portfolio almost guarantees that some active risk will be represented in the historical time series, which by extension will prevent isolating the active risk taken by the particular fund that the investor is benchmarking. The argument that some active risk in the benchmark portfolio is unavoidable, by simply having too few funds investing in a particular strategy or under the same constraints, is quite questionable considering the fact that the particular data set available to the investor can be too narrow.

A heuristic approach to reduce the impact of idiosyncratic risk in benchmarks is to use equal weighted portfolios. However, applying this approach is complicated by the way in which asset values are calculated. Unlike public assets where the fair value is readily available in the marketplace, as already mentioned, the values of private assets are based on appraisal and accounting guidelines that do not adjust as often and in full transparency with the general economic and market conditions.

Familiar from public asset management, one rigorous approach to constructing a benchmark portfolio that reduces the impact of active risk is to build minimum variance portfolios. This is usually achieved by mean variance optimization, using an asset covariance matrix constructed using a factor risk model. The risk model is required to reduce dimensionality of the problem in the presence of numerous assets and a finite number of observations. Such models filter out unusual past events that are unlikely to be repeated (e.g., two popular company CEOs are killed in the same plane crash) and thereby provide higher confidence in the expectation of future behavior.

The economic rationale for factor representation is to appropriately allocate the risk of an investment to systematic (common) and idiosyncratic risks. The Capital Asset Pricing Model (Sharpe 1964) and the Arbitrage Pricing Theory (Ross 1976) are examples of such models, but have limited direct applicability to private equity, in that both assume that financial markets have essentially infinite liquidity, which is clearly untrue for private equity. A paper by Buchner, Kaserer, and Wagner (2014) bridges this theoretical gap by attempting to derive the CAPM without the assumption of high liquidity. Their novel conclusion is that for private assets, idiosyncratic risk is priced (i.e., higher returns go with higher risk) while the traditional CAPM argues that while systematic risks are priced, idiosyncratic risks are easily diversifiable and hence not priced. Further evidence on this point is provided by Mueller (2011), which found that private companies had higher returns when management was exposed to higher degrees of idiosyncratic risk by having a higher proportion of their investment wealth concentrated in the firms they managed. The distinction between private equity and publicly traded equities is made even more stark by diBartolomeo and Kantos (2020) who found persistently negative returns associated with idiosyncratic risks in public equities, particularly when investors are presumed to be cognizant of the impact of “large events” such as pandemics or wars.

Applying similar approaches to private equity fund and asset portfolios is challenged by some of the same data and valuation methodology issues already discussed. The appraisal and accounting based valuations are notoriously smoothed, as is well documented in the literature (Ilmanen et al. 2020). Also, short histories make for statistically insignificant results when the same econometric techniques used to derive public asset risk models are used for private assets.

Alternatively, if raw time-series are used instead to construct a covariance matrix among the assets, two issues arise. First, with a greater number of assets under consideration for inclusion in the benchmark portfolio, the dimensionality of the covariance matrix increases, which quickly brings about matrix rank problems due to having more variables than historical observations. Secondly, the fact that the appraisal based time series are smoothed contributes to underestimating volatility and overestimating correlations. Underestimating volatility, including tracking error, will have an impact on determining the statistical significance of a manager’s skill. Overestimating correlations will cause the benchmark portfolio to appear over-diversified.

Some practitioners and academics have responded to the smoothing phenomenon by attempting to de-smooth the time series data of private asset values and returns (Geltner 1991). The key idea in such attempts is to infer the underlying bias in the time series and reverse it, exposing the allegedly true, but unobservable, aberrations. In some cases this is done using “frequentist” statistical techniques, in others Bayesian inferences, but in all cases the rationale is the same.

While de-smoothing is designed to resolve the problem with non-arms-length valuations and returns, it comes with some hefty econometric baggage. In particular, like any empirical estimation, it brings along some estimation error. Estimation error is a natural part of regressions used in risk models, but those estimation errors are in the risk factor exposures rather than the risk factors themselves, and thus get diversified in a larger portfolio (Belev et al., 2014). In contrast, when time series that are de-smoothed are used to construct a covariance matrix, the estimation error is built into the covariance matrix and only accumulates as the portfolio becomes larger. If we take a theoretical example where one has confidence in any pairwise correlation between any two out of N variables of 90%, then assuming independence of the correlation estimates, when N = 30 we will have confidence of less than 5% that all estimates are correct. This, in turn, makes the mechanical construction of a covariance matrix in this fashion an exercise of dubious practical value.

A practical walk around the issue of non-transacted time series returns is to incorporate an analogous public market portfolio. This involves constructing a well-diversified public stock portfolio that has similar geographic, industry, and size characteristics as the asset being benchmarked. Given the illiquidity of private equity, one is reasonably expected to look for a way to incorporate the additional risk imposed by illiquidity. If a premium to expected return is estimated for private equity (Franzoni et al. 2012), then one can calculate the amount of additional volatility that would make a mean-variance optimal investor indifferent between investing in the public analogous portfolio and the illiquid portfolio with, otherwise, the same characteristics. To the extent that this approach is concerned with frequently priced public equivalents, it is concerned with "present value" risk—i.e., how the value of future cash flows from the private asset exchanged at any given day for a cash amount is the value that constitutes the payout from the investment.

An alternative to the public analogous portfolio approach is to recognize the natural way in which private assets and funds thereof yield investment returns. They do this through operational or divestiture cash flows that get passed on to investors either directly from private firms or as distributions from a fund. These distribution payouts by themselves do not require a marketplace for private asset stakes like the "present value" approach. In that sense the cash flow approach is essentially concerned with "future value" risk.

Gauging the risk-return characteristics of a fund requires several analytical capabilities. First the investor needs to be able to forecast the pacing of contributions to, and distributions form, the private asset or fund considered for inclusion in the benchmark. Second, they would need to forecast the expected growth and its volatility over the lifetime of the asset or fund, and generate statistical distributions of periodic cash flows (Exhibit 1). Third, they would need to be able to observe how each cash flow path aggregates in a future value outcome, most often through a simulation or sometimes analytically (Exhibit 2).

Since investors constructing the benchmark are concerned with the interrelationship among funds and assets to be included in the benchmark portfolio, they should additionally be able to calculate the multi-period covariance matrix. A natural way to approach this is to use a single period risk model of the cash flows, capture the beta among two assets, and aggregate to multiple periods using a simulation. The multi-period correlation is an algebraic outcome of combining two simulated future value distributions. The derivation and rationale of this approach is described below.

Sample two assets: **A** and **B**

Time horizon: $\boldsymbol{T}$ - the full (longer) fund maturity horizon for assets A and B over which their full Future Value realization will occur

$\boldsymbol{COV(B,A)_T}$: Covariance between assets A and B Future Values over time horizon T

$\boldsymbol{\sigma_{A,T}}$: Standard deviation of simulated Future Value statistical distribution of A over horizon T

$\boldsymbol{\sigma_{B,T}}$: Standard deviation of simulated Future Value statistical distribution of B over horizon T

$\boldsymbol{\sigma_{B\sim f(A),T}}$: Standard deviation of simulated Future Value statistical distribution of B over horizon T, where the path-wise simulated realizations of asset B are derived by multiplying the one-period beta of asset B to asset A by a simulated realization of asset A; the one period beta is derived by a one-period risk model

$\boldsymbol{\beta_{B-to-A,T}}$: Effective multi-period beta of asset B to asset A over horizon T

EXHIBIT 1

Periodic Non-Cumulative Cash Flow Distributions over Remaining Quarters

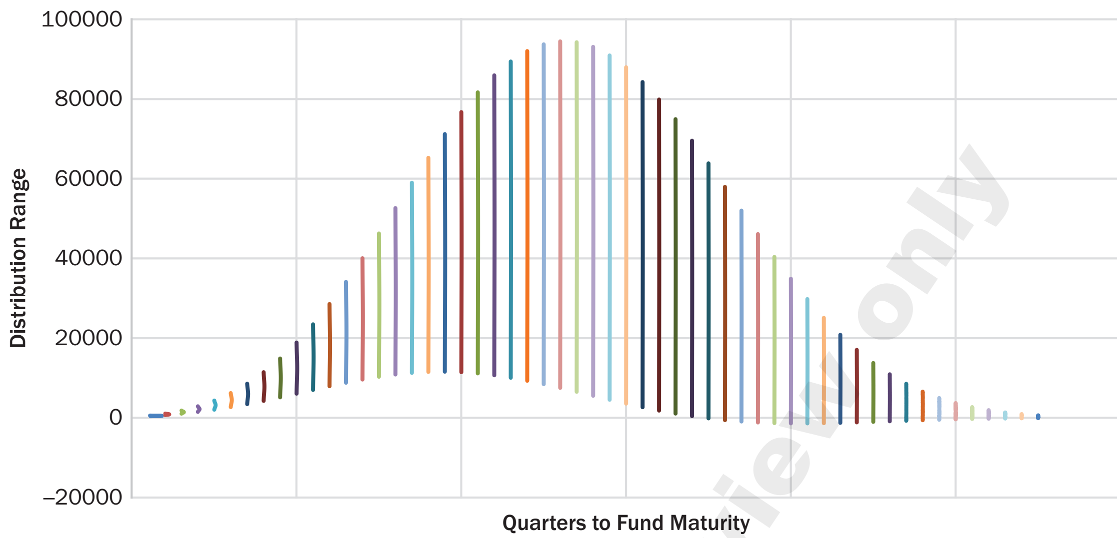
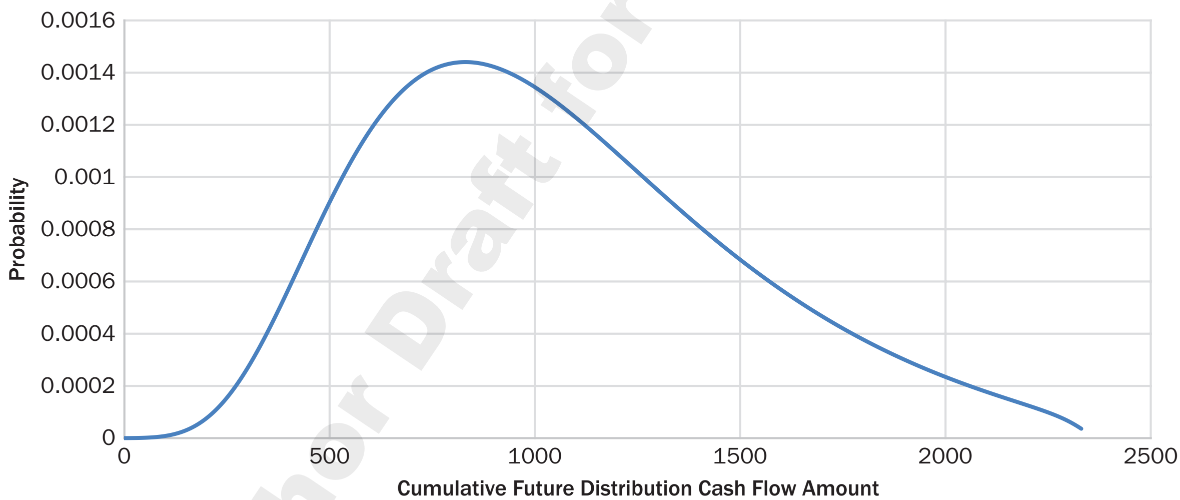


EXHIBIT 2

Cumulative Cash Flow Value Probabilities



$$COV(B,A)_T = \frac{COV(B,A)_T}{\sigma_{A,T} * \sigma_{B,T}} * \sigma_{A,T} * \sigma_{B,T} = \frac{COV(B,A)_T * \sigma_{A,T}}{\sigma_{A,T}{}^2 * \sigma_{B,T}} * \sigma_{A,T} * \sigma_{B,T}$$
$$= \frac{\beta_{B\text{-to-}A,T} * \sigma_{A,T}}{\sigma_{B,T}} * \sigma_{A,T} * \sigma_{B,T} = \frac{\sigma_{B\sim f(A),T}}{\sigma_{B,T}} * \sigma_{A,T} * \sigma_{B,T} = \sigma_{B\sim f(A),T} * \sigma_{A,T} \quad (1)$$

Or:

$$COV(B,A)_T = \sigma_{B\sim f(A),T} * \sigma_{A,T} \quad (2)$$

In other words, we can derive the multi-period covariance between any pair of assets A and B by running a series of simulations and computing the respective standard deviations of the resulting distributions.

The end result from constructing the covariance matrixes from the "present value" and the "future value" approach is that using optimization, one is able to construct a minimal variance portfolio of peer private funds to use as a benchmark. While the investor can still continue to compare the aggregate measures of return of a specific fund—IRR, Paid-in multiples, etc.—to such a custom benchmark, as with any private equity benchmark already in place, the minimum variance benchmark serves a distinct and rigorous analytical purpose. This is to ascertain the statistical significance of a manager's over-performance by a) providing an objective basis for comparison that has minimal active risk, and b) deriving the basis to calculate tracking error with minimal impact of estimation error. To calculate tracking error one must still resort to de-smoothed fund and benchmark returns derived using the family of techniques pointed out earlier. In this setting, however, de-smoothing does not represent an estimation error challenge, since it is used only on a single pair of time series rather than numerous time series used to construct a covariance matrix.

The future value approach also provides the key to resolving another subtle challenge posed by private equity performance measurement in general. An important issue in benchmark index design is the *assumption* that the cross-section of returns used to form the index is normal. Empirical studies such as Kaplan and Schoar (2005) suggest that the cross-section of private equity returns has strong positive skew. Even stronger evidence of positive skew was reported by CalPERS for its venture capital portfolio (Sague 2017). Certain types of private financing activities (e.g., feature films) are commonly acknowledged to also have this kind of payoff distribution, although this effect is expected to lessen as new outlets (such as streaming) diversify the revenue sources for entertainment productions. To the extent that investors cannot achieve the high degree of diversification implied by a broad index, the *median return* of the investments making up an index is a more realistic benchmark than the typically published *mean return*. After rigorously constructing a benchmark that assures the maximum level of diversification in the manner described in this section, one can reliably compare a portfolio with a chosen measure of aggregate return, whether that is the cross-sectional average or the cross-sectional median for each data point in the historical time series. This step makes it even more likely that the impact of "active skew" is minimized by design on the benchmark side.

PUBLIC BENCHMARKS FOR PRIVATE ASSETS

After comparing private assets with private benchmarks, the next logical possibility is to consider a benchmark consisting of public investments or portfolios thereof. This choice certainly would not be surprising, considering the challenges the investor in private assets faces when identifying the appropriate weights to use to construct a private benchmark. Public assets are liquid, possess observable pricing, and offer an extensive risk model methodology in both academia and industry to construct minimum variance portfolios. Another advantage of public benchmarks is that they have been in existence for a very long time in the public asset space, providing familiarity and, to some extent, less room for subjectivity in the choice of constituents.

Normally, such a public benchmark is applied in conjunction with a set margin added on top of a chosen public index (Aksia 2020 and MassPRIM 2020). The reason for the margin application is the assumed higher expected return due to an illiquidity premium. The presence of such an illiquidity premium is discussed in the works of Franzoni et al. (2012), Ang (2013), Harris et al. (2014), and Anson (2017). The Franzoni and Harris studies found it to be on average approximately 3%; Anson found that it varies between 1% and 8%, with a long term average of 3.7%; and Ang estimated it to be between 2% and 6% ,depending on the holding time horizon. Given that these studies gravitate around an average estimate of approximately 3%, this is what most institutional consultants recommend for the level of the applied margin.

While the public equity index as a benchmark is certainly investable, the application of the illiquidity premium warrants a special note. The compensation for illiquidity as a source of systematic risk, unlike other sources of systematic risk, cannot be captured passively due to the various rebalancing constraints posed by the asset class; it instead requires being able to pick the best managers. Nevertheless, due the proximity to being an investable benchmark, the approach is chosen by asset owners in addition to or instead of a private fund constructed benchmark.

Even if this approach offers some advantages, it comes with it is own set of challenges left to the investor to resolve. First, in the absence of a precise estimate of the illiquidity premium, the choice remains subjective within the range of the studies that measure it. While practically all the previously mentioned research efforts attempted to account to various sources of systematic risk to isolate the illiquidity component, using different datasets (e.g., private fund cash flows, as in Franzoni et al. 2012, or yields of mezzanine illiquid debt, as in Anson 2017) and different econometric techniques (e.g., regression on continuous time returns as in Ang and Sorenson 2012 or analysis on de-lagged time series as in Anson 2013) in combination eventually produces a range of estimates that leaves room for a subjective judgement.

Another issue is the fact that while public index returns are reported at potentially very high frequency, private asset returns are reported quarterly, at the most frequent. To make matters worse, as already alluded, the reported returns of the private assets are smoothed as opposed to the returns of the benchmark, which being a market-based index, are not. Furthermore, Czasonis, Kritzman, and Turkington (2019) documented that due to time lags in reporting of valuations by general partners, the valuations themselves and the derived returns are often biased with respect to public market performance that occurred after the end of the measurement period. All these discrepancies necessitate making certain adjustments before return performance measurement is executed.

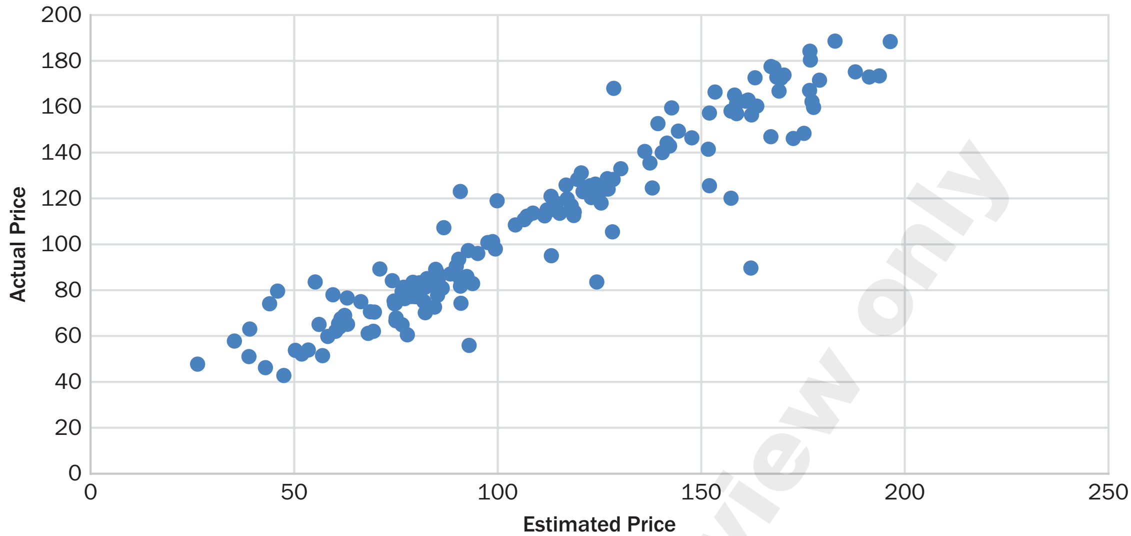
One way to assure that a naturally unsmooth benchmark like a public asset index is comparable to a smoothed private asset is to smooth the public benchmark time series. This can be done in a number of ways, but the most usual is to take a moving average over a certain number of periods, usually comparable to the time span that the private asset is believed to be smoothed. The periodicity challenge can be addressed by taking private benchmark returns that reflect the same frequency as the private assets. If the private asset reports quarterly returns, then the public index returns can be stringed into quarterly returns, and then smoothed over a number of quarters before being used as a benchmark. Obviously, while making the two time series comparable, this approach also obscures the true statistical significance of the tracking error due to both compared time series being smoothed. Additionally, this approach does not in any way deal with the issue of subjectivity of the margin on top of the benchmark.

An approach that deals with all three challenges of the public benchmark used for private asset performance measurement involves two valuation steps. One of them is determining the "now-casted" value of the private asset, and the other re-valuation of the public market index using risk aversion assumptions adapted from the private markets.

Now-casting the private asset value is necessary in order to introduce the most recent market conditions and information into the current value of the investment. A current "fair value" for it is most likely missing externally, or rather stale from the last quarter end. And, due to the subjective way quarter end private asset values are reported, it is most likely smoothed along with prior reported values. According to the AICPA valuation guidelines, the fair value should incorporate calibrating the pricing

EXHIBIT 3

Risk Neutral Valuation: Estimated vs. Actual, Russell 3000 ETF



	Coefficients	Standard Error	*t*-Stat	P-Value
Intercept	11.5	2.8	4.2	0.00
Slope	0.9	0.02	37.5	0.00
R-Squared	0.90			

method based on the most recent transaction of an asset that is closest in character to the one being valued. AICPA prescribes a number of valuation approaches in two general categories: market multiples and discounted cash flow methods (DCF). Each adopts a different way for calibration on the most recent market transaction. For market multiples, one calibrates by taking the proportional relationship embedded in the most recent transaction's EBTTDA-to-price, or sale-to-price, ratio. DCF methods span a wide variety, from a simple expected cash flow discounting to more complex models like the Probability Weighted Expected Return Method (PWERM) and Option Pricing Method (OPM). In those cases the calibration is embedded in the discount rate or in the adjustment of probabilities of the projected various investment outcomes. The adjustment of probabilities of outcomes is known from option theory as risk-neutral pricing. PWERM and OPM, or a hybrid method thereof, present the highest level of flexibility to accommodate the full range of outcomes for an investment.

In order to assure comparability, the two valuation steps—for the private asset and the public index—should be performed with the same method. The following illustration of implementing now-casting incorporates a hybrid method of OPM and PWERM valuations, where the probabilities of outcomes are adjusted to risk neutral pricing based on the most recent reference transaction. The example in Exhibit 3 demonstrates how this valuation approach reproduces the values of the Russel 3000 index monthly from January 2007 to June 2020. (Source of traded prices: https://www.ishares.com/us/products/239714/ishares-russell-3000-etf.) The transparency that arms-length public markets provide is the ultimate measure of how well this approach performs.

The specific process for this test is as follows. At time ***t* – *1***, the statistical properties of the dividends are projected over the expected lifetime for the average company in the index. Then the probabilities of outcomes are shifted to more adverse outcomes to assure that the average of the present values using the risk free rate under each outcome reflects the latest transacted price of the index. This is the central idea behind risk-neutral pricing (Cox et al. 1979). Then the statistical distribution of cumulative dividends at time t is calculated, and the probabilities of this range of outcomes are modified by a shift implied by the calibration at time $t - 1$. The probability-weighted average at time t is taken and compared with the actual transacted price, and the results are plotted as a test of the quality of the valuation procedure.

While the results give us a baseline idea of the performance of the hybrid valuation model, the actual valuation steps mentioned earlier imply a different reference point of calibration. In order to incorporate the same level of risk aversion embedded in the private asset valuation, we have to first infer it from the most recent transacted price of the same or a similar private asset. This involves building the distributions of the cumulative cash flows from the distributions of the transacted private asset, and finding the shift in probabilities that will make the average of the outcomes equal to the most recent transacted calibration price. Using the risk aversion embedded in the probability shift derived from the most recent private market transaction, the investor can calculate the price of the public index that reflects the same level of risk aversion. Calculating this price under a private risk aversion will then reveal what the public index is worth as if it is traded as a private investment. When the proportional difference from the actual price of the index is taken, the size of that difference will be a rigorously derived margin on top of the public index in order to benchmark the private equity asset.

Based on the risk aversion calibrated from the latest private transaction, one can price the private asset at an immediately subsequent point in time. In essence, this step ensures that the latest market conditions are reflected in how the cumulative cash flows of the private asset are related to the now-casted fair value. If it is necessary to do “now-cast” values historically, then one has to derive risk aversion adjustments for each most recent transaction prior to the time point they want to compute fair-value, to make sure the valuation is performed with the facts and circumstances knowable at the time and avoid lookback bias.

Once both tasks are accomplished—the margin over the public index is derived and the now-casted values of the private asset are calculated with the same periodicity of the public benchmark—traditional performance measurement can be performed relative to the benchmark.

This approach to benchmarking is also a gateway to other investor and market capabilities. It is well known that benchmarks can be used as the basis of traded derivative contracts. While this application is not yet well developed in the private equity world, there is clear potential to add hedging of PE investments, and to add to functional liquidity. In concept the foregoing applications of benchmarks can by addressed by appropriately constructed indexes of private equity investments that are transparently derived by a particular public equity portfolio plus an appropriately computed margin on top of its realized return. Such indexes form the basis of informed evaluations about the absolute returns of private equity as a broad class and about the relative returns of specific investments.

We conclude this section with a note on the general motivation for *re-valuation* of public portfolio as an appropriate reference point for illiquid private equity. The comparison of specific investments to a benchmark index leaves the potential to judge a poorly performing investment as satisfactory only because the broad index performed even worse. Foreclosing this possibility requires a different framework for benchmarking, that compares returns achieved to the return expectations of some asset pricing model. To the extent that many invest in illiquid assets with expectations of an illiquidity risk premium, private equity returns must compensate for the loss of "the option to do something else." The cost of the loss of flexibility has been studied in the similar problem of hedge fund "lockups" by Ang and Bollen (2008). A process for defining of "risk adjusted return" as the benchmark for both illiquid and liquid assets in a unified framework is presented in diBartolomeo (forthcoming, *The Journal of Performance Measurement*, 2021).

PUBLIC MARKET EQUIVALENT BENCHMARKING

The public market equivalent (PME) performance measurement is an approach to benchmarking that is specific to private asset investing. It recognizes that the returns calculated using appraisal and accounting based methods have limited merit when benchmarking actual investment performance. The PME methodology has evolved in a number of different applications—PME (Long and Nickels 2006), PME+ (Rouvinez 2003), mPME (Cambridge Associates 2013), and PME-KS (Kaplan et al. 2005)—which each deals slightly differently with the issue of finding a proxy of residual value to calculate investment returns of private equity. The common feature among them is that, instead of trying to calculating periodic returns, they use private asset cash flow compounded by the return on a market index across time to simultaneously synchronize cash flows that are contributed to or distributed by the private investment, and to incorporate the opportunity cost from the alternative to invest in a public market index.

In this section we will deal specifically with the version of PME developed by Kaplan and Schoar (2005), since it is largely devoid of the need to calculate a *residual value* at any point in the life of the private asset. Unlike cash flow distributions, contributions, and public market realizations that have actually taken place, it is a *subjective* construct. PME-KS also lends itself to a transformation in a convenient measure of continuous time "alpha," an approach developed by Gredil, Griffiths, and Stucke (2014).

PME-KS is defined as:

$$PME_{KS} = \frac{FV(D)}{FV(C)} \tag{3}$$

where:

$$FV(D) = \sum_i \left(Distribution(t) * \frac{I_T}{I_t} \right) \tag{4}$$

$$FV(C) = \sum_i \left(Contribution(t) * \frac{I_T}{I_t} \right) \tag{5}$$

and:

I_T: is the value of the benchmark public equity index at the time of PME measurement T

I_t: is the value of the benchmark public equity index at the time ***t***

It can be observed that the performance of the public market index "*I*" is embedded in the calculation of the future value of the cash flows, which reflects the opportunity cost of investing the contributions used to invest in the private asset, and the value of the distributions from the private asset compounded up to point in time T, assuming that they are reinvested in the public market index. Should the private investment offer higher than cumulative benchmark return over the period, the ratio will be higher than 1.

The “Direct Alpha” approach recognizes that the public index return is effectively a “beta” return. It uses that relationship in continuous time to isolate the effect of over-performance of the private asset in order to assign it to a time period of a particular size rather than the full time horizon T. The continuous time alpha of the private asset is calculated in two steps:

$$a = IRR\left(FV(D),\ FV(C),\ NAV_T\right) \tag{6}$$

$$\alpha = \frac{\ln(1+a)}{\Delta t} \tag{7}$$

where:
NAV_T: is the assumed net asset value at time T
Δt: is the time interval of interest for which “alpha” is to be computed

While the PME and Direct Alpha methods of benchmarking circumvent some of the major issues with appraisal based returns inherent in the first two benchmark methods, they present some of their own challenges.

First, similar to the approach where a public index return series is used to benchmark private asset return series, the choice of public index is often a subjective and imprecise exercise. This is because the policy and constraints guiding the private equity mandate may be different from those reflected in public equity indexes. The investor should carefully inspect such discrepancies and adjust for them to the best extent possible—constructing, for example, public index benchmarks that are combinations of other public indexes weighed in a particular fashion that reflects the required adjustment in investment focus and constraints. For example, if the mandated private equity strategy has a focus on emerging managers in particular industries and geographies, the proper implementation of PME measurement entails identifying indexes that focus on public investments in the same industries and geographies, and weighting their time series returns in a similar fashion as the weights represented by the private equity mandate.

Importantly, once again, the concept of “tracking error,” or the uncertainty of over- or under-performance, is missing in the way the PME measures are constructed. This is an important component of the benchmarking process, because the probability ranges associated with having a positive or negative “alpha” speak directly to whether the over-performance is due to chance or manager skill.

One Can

It is not possible to gauge the statistical significance of skill, but one can instead simulate the PME measures and take their point estimate derived from history. As evident from the way the PME measures are constructed—using cash flows from the private asset—this, once again, requires simulating future cash flows from the private asset. The uncertainty and the value of manager choice pertain to the future and, hence, the exercise of getting the probability range of these cash flows and PME measures should be forward looking. The alternative to looking at a cross-section of manager PMEs in the past lacks merit, in the sense that it does not measure the uncertainty of the skill of the particular manager but rather a collection of different managers operating in the same investment class. This is hardly helpful when one has to select a particular one out of the group.

The instrumentality of the cash flow simulation is very similar to what was discussed in the sections dedicated to the first two benchmarking approaches. This time, however, instead of computing the cumulative cash flows themselves, we are concerned with the outcomes for the specific PME ratio under each realization path for contributions and distributions.

LIABILITY DRIVEN BENCHMARKING

While goal of all the previous approaches to benchmarking private equity is a comparison of the performance of a private asset with a peer group of similar assets or managers, liability driven benchmarking has a different objective. Instead of answering the question, "Which investment out of a comparable pool performed best?", it is asking, "Which investment has the highest likelihood of meeting my mandated investment objectives?" Under some circumstances there can be a definitive overlap of investments that fit both objectives, but there are is a distinction in terms of the timing of the two types of performance projections. For example, if an investor is concerned with meeting a liability occurring in the next two to four years, picking a manager who makes the best active bets does not guarantee the optimal results with respect to the liability objective. The best-performing manager might be taking too much active risk to assure confidence in meeting the liability target under the worst case outcome.

In this sense, the liability driven benchmarking approach is the most directly objective for most investors with identifiable liabilities. Most investors fall under this category and, understandably, prefer to like to maximize their long term portfolio value subject to being able to meet their liabilities.

By definition, liability driven benchmarking is concerned with having liquid assets at a certain point in time in the future. Naturally, liquid assets can be available to the investor from two sources: liquid publicly traded assets and proceeds from private assets. In turn, proceeds from private assets also can come from two sources, liquidation of interest in private assets, and positive cash flows produced by private assets that the investor continues to hold. Since liquidation discounts of private equity in the secondary market can be quite significant (Nadauld et al. 2019), the rational choice for a non-distressed investor is to plan to cover liabilities with the cash flows from private equity investment vehicles they intend to hold to maturity.

The implication is that liability driven investors should be concerned with maximizing the value of a combination of resources—the future market value of public asset holdings and the future value of the cumulative cash flows produced by private assets. A liability benchmarking process constructed in this fashion has a number of distinctive features:

- It is concerned with the standalone performance of both public and privates buy-and-hold investments, as well as their interaction over time.
- Under certain mandates it is possible for the investor to be concerned only with the cash flow performance of private assets. That can occur when a portfolio of private assets is required to maintain a self-sustainable balance of cash flows from distribution from some private asset investments and cash flows going out to contributions to other private asset investments. The liability driven benchmarking then still is concerned with the interaction effects among the cumulative cash flow distributions of the different private assets that will be used to meet capital call liabilities.

- It is an ex-ante analytical problem, not ex-post traditional benchmarking. Having not met liability targets in the past has a binary survivorship implication to the present moment and almost no relevance to the future ability to do so.
- Almost definitely, liabilities must be met not just over a single time horizon, but over a number of horizons. For this reasons, the concept of "tracking error" is not applicable. Its closest conceptual analogue in this setting is the confidence level at which the investor is willing to accept "statistical certainty" that liabilities will be met.

Based on these distinctive features, it is evident that the ability to perform liability driven benchmarking rests on the same analytical approach of forecasting the cash flow statistical distributions of private assets over the projected liability horizons. In a multi-asset class portfolio this also requires the forecast of the contemporaneous performance of public assets. As previously described, all these forecasts also should incorporate the correlation properties of the assets involved. Given the multi-period and path-dependent nature of the variables forecasted this can only be accomplished in practice by a simulation.

To avoid redundancy at this point, we restate that one can construct a covariance matrix of "liquid value" assets within the portfolio for each time horizon. In conjunction with the asset class weights, this allows the calculation of the mean and variance of the liquid portfolio value at each time horizon using:

$$\mu_{Liquidity\ \ Level,T} = \sum_{i=1}^{N} E[X_{i,T}] \tag{8}$$

$$\sigma^2_{Liquidity\ \ Level,T} = \sum_{i=1}^{N}\sum_{j=1}^{N} COV(X_{i,T}, X_{j,T}) \tag{9}$$

where:

$\mu_{\boldsymbol{Liquidity\ Level,T}}$: Is the mean of the market value of liquid resources available at time T from the multi-asset class portfolio that is equal to the sum of liquid public asset values, and the cumulative private asset cash flows generated by private assets

$\sigma^2_{\boldsymbol{Liquidity\ Level,T}}$: Is the variance of the market value of liquid resources available at time T from the multi-asset class portfolio

$\boldsymbol{X_{i,t}}$: Is a specific realization of a source of liquid resources "***i***" at time T

Once the mean and variance statistics are calculated for each time horizon, one can determine the quantiles of asset value that correspond to the investor level of "statistical certainty," e.g., 90%. Then this quantile should be compared with the amount of the liability to be covered in the particular period, and determination made whether the specific portfolio meets the criteria required by the liability benchmarks.

It should be noted that when performing the simulation, the rational investor will choose to use the "Future Value" liquid resource view of the portfolio. Considering the "Present Value" view of portfolios assets when simulating the private asset values over different horizons would cause the investor to perform fire sales of these assets in order to cover mandated periodic liabilities. This is not an optimal choice, especially when planning is performed well in advance specifically with the intent to avoid the need for such fire sales.

By definition, the liability-driven benchmark approach is specific to the needs and circumstances of a particular institution. For pension funds, for example, the broad cross-section of assumed rates of return that determines current and future contributions means that each plan sponsor faces different odds of meeting future liabilities, even if overall asset allocations are the same. Similarly, the extent to which a sponsor of an endowment or a foundation relies on the proceeds of its portfolio for routine operational needs, or a specific contribution schedule, will define a unique schedule of liabilities it needs to meet. Insurance companies investing in private assets also would face various actuarial estimates and assumptions defined by their product mix, requiring a custom approach to constructing an asset portfolio designed to maximize investment performance while ensuring a certain level of confidence for meeting liabilities.

The use of a liability driven benchmarking method can have benefits beyond what transpires in its name. As noted, the use of benchmarks is closely aligned with the assessment of "skill" by managers who undertake to invest in a particular asset class. An important consideration in assessing the skill of private equity managers is the distinction between external and internal managers. External managers (GPs) can control the timing of inward cash flows from investors by deciding when to open a particular fund (partnership). Internal managers within a large asset owner may receive incremental allocations of cash as the result of asset allocation decisions that are driven by considerations unrelated to private equity. Internal managers of illiquid assets therefore will argue that investing incremental funds promptly when faced with a finite set of opportunities represents a distinct form of skill that is not measured by conventional indexes. The likelihood is higher that internal manager cash flows will be more deterministic as they are driven by the investor organizations. Provided that cash flow pacing models on the contributions side can be calibrated to both internal and external manager circumstances, the two types of *existing* portfolios can be compared on an equal footing with the clear objective of hitting the same cumulative distribution targets over comparable time horizons in the future. Then the tradeoff between the exogenous constraint to stay fully invested at any given time can be compared as a presumed cost, with the benefit of possibly lower management fees of internal funds.

CONCLUSION

Benchmarking private equity is wrought with challenges posed from both the ambiguous identity of the benchmarking process itself and the inherent obscurity of the asset class. In the face of these significant challenges, there is an opportunity to redefine the objective and the process, and to outline all alternative solutions that may resolve different aspects of the problem and result in a richer and more informative result. A convenient starting point is to analyze the ways the investment industry has evolved to tackle such problems. By exploring the motivation and the circumstances in which these techniques arise, the investment professional can contrast with the fundamental measurement objectives and identify potential areas of improvement in each method. One overarching principle in each case is to reduce the reliance on subjective measures like appraisal values and heuristic groupings, and focus on objective measures like cash flows, transacted values, and efficient benchmark portfolios. Another is to evolve beyond static measures of performance to those that can suggest confidence intervals that truly differentiate between skill and luck.

APPENDIX A

The information below has been self-reported by private equity benchmark providers in interviews conducted by the authors. Authors disclaim any assertions on the accuracy of the information. The intent of this survey was to give each vendor an opportunity to provide a fair representation of its coverage and methodology, which then can be compared across the group.

Several notes should be taken into consideration when reviewing the data. First, the latest reported date slightly varies due to the timing of the scheduled update cutoff dates for each vendor.

Second, we treated fund-of-funds and secondary fund assignment by aggregating them with the General Private Equity category. This was done for two reasons. The first—a semantic one—was the argument that funds holding equity in private companies that happen to be legally formed as limited partnerships can legitimately be considered a form of private equity fund in their own right. The second, on the practical side, is that each of the five vendors assigns funds according to a somewhat different granularity of categories. In order to allow drawing parallels across them, the authors had to opt for more aggregate breakdowns than originally provided.

Third, some vendors, in addition to a fund type count breakdown, also provided a breakdown of fund type by total monetary commitment weights. Since not all vendors provided this data, for comparability, only the fund type count breakdown is presented here. As expected, the general effect of weighting fund types by total commitments instead of counts would be that types of funds that naturally have smaller monetary size, like venture capital funds, would reflect a reduced representation.

Finally, different vendors chose to word certain responses with a small variation. While drawing complete parallels in this sense may require some level of interpretation, the authors preferred the original wording because it may imply features that the vendors consider differentiators to their offering.

Burgiss

Total Fund Covered In Benchmarks: *10,930 Private Capital Funds* (*as of fourth quarter 2020*)

Breakdown of Number of Funds by Type

General Private Equity: 36% (3,966), Venture Capital: 28% (3,011), Private Debt: 10% (1,105), Real Assets: 19% (2,066), Other: 7% (782)

Sub-Benchmarks: *Available to reflect a breakdown by underlying asset, strategy, vintage, geographic, and industry focus*

Source of Data: *Cash flow transactions and net asset values reported by Limited Partners holding stakes in private funds. Total cash flows and net asset values of each fund are found by scaling up the individual Limited Partner stakes.*

Major Reported Metrics: *Total and horizon IRRs, Multiples to Paid-in Capital, PME, Direct Alpha*

All performance metrics are net-of-fees and carried interest

Restating Historical Results: *Restates historical time series if corrected information for historical performance of certain funds comes through its data sources*

Ability to Create User-customized Benchmarks: *Yes, through a user interface*

Aggregation and Drilldown Options: *Reports equal-weighted average benchmark performance of constituent funds*

Reports median, top, and bottom quartiles cutoff return levels

Provides breakdown of each benchmark by industry, geography, vintage content

Cambridge Associates

Total Fund Covered In Benchmarks: *8,243 Private Capital Funds (as of third quarter 2020)*

Breakdown of Number of Funds by Type

General Private Equity: 42% (3,447), Venture Capital: 30% (2,506), Private Debt: 6% (512), Real Assets: 22% (1,778)

Sub-Benchmarks: *Available to reflect a breakdown by underlying asset, strategy, vintage, geographic, and industry focus*

Source of Data: *Cash flow transactions and net asset values reported by General Partners and Managers of Private Funds in quarterly and annual audited statements to Limited Partners*

Major Reported Metrics: *Since Inception and horizon IRRs, Multiples of Paid-in Capital, Modified PME (mPME), K&S PME, Direct Alpha*

Benchmark returns are "net to LP"—i.e., what an investor would realize after paying management fees, expenses, and carried interest.

Restating Historical Results: *Restates historical time series if corrected information for historical performance of certain funds comes through its data sources*

Ability to Create User-customized Benchmarks: *Yes, through a user interface*

Aggregation and Drilldown Options:

Also, reports equal-weighted and average return, as well as median, top, and bottom quartiles cutoff return levels

Using CA's online application, one can have numerous cuts of any given benchmark sample: by capitalization, fund count, geographic focus, strategy, vintage, industry focus, etc.

Other features in the online application include the organization of any given benchmark sample by different dimensions: sole/emerging/established managers, quartiles, liquidated/active funds, capitalization buckets, etc.

Pitchbook

Total Fund Covered In Benchmarks: *6,310 Private Capital Funds (as of third quarter 2020)*

Breakdown of Number of Funds by Type

General Private Equity: 52% (3,297), Venture Capital: 18% (1,130), Private Debt: 9% (555), Real Assets: 21% (1,328)

Sub-Benchmarks: *Available to reflect a breakdown by underlying asset, strategy, vintage, geographic, and industry focus*

Source of Data: *Public disclosures of cash flow transactions and net asset values reported by institutional Limited Partners holding stakes in private funds, as well as information reported by General Partners managing private funds*

Major Reported Metrics: *Total and horizon IRRs, Multiples to Paid-in Capital, PME, Direct Alpha*

Fund performance data is reported net of fees and carried interest

Restating Historical Results: *Restates historical time series if corrected information for historical performance of certain funds comes through its data sources*

Ability to Create User-customized Benchmarks: *Yes, through a user interface*

Aggregation and Drilldown Options: *Reports equal-weighted and average benchmark performance of constituent funds*

Reports median, top, and bottom quartile, as well as top and bottom deciles cutoff return levels

Provides breakdown of each benchmark by industry, geography, vintage content

Preqin

Total Fund Covered In Benchmarks: *12,120* (*as of April 2021*)

Breakdown of Number of Funds by Type

General Private Equity: 44 % (5,313), Venture Capital: 20% (2,415), Private Debt: 10% (1,272), Real Assets: 26% (3,120)

Sub-Benchmarks: *Available to reflect a breakdown by underlying asset, strategy, vintage, geographic, and industry focus*

Source of Data: *Public disclosures of commitments and individual fund performance metrics by institutional Limited Partners holding stakes in private funds, as well as fund performance metrics reported by General Partners and Managers of Private Funds*

Major Reported Metrics: *Net IRRs, Net Multiples to Paid-in Capital, Called (%), DPI (%), RVPI (%), PMEs including Direct Alpha*

Restating Historical Results: *Restates historical time series if corrected information for historical performance of certain funds comes through its data sources*

Ability to Create User-customized Benchmarks: *Yes, through a user interface*

Aggregation and Drilldown Options: *Based on constituent funds in the benchmark, aggregation and drill down options for Net IRR and Net Multiple include: median, weighted, mean, standard deviation, top and bottom quartiles among others*

Reports median, top, and bottom quartile cutoff level per each metric (Called%, DPI%, etc.)

Provides breakdown of each benchmark by strategy, geography, and vintage content

State Street Associates SSPEI℠

Total Fund Covered In Benchmarks: *3,361* (*as of fourth quarter 2020*)

Breakdown of Number of Funds by Type

General Private Equity: 49% (1,646), Venture Capital: 37% (1,260), Private Debt: 14% (455)

Sub-Benchmarks: *Available to reflect a breakdown by underlying asset, strategy, vintage, and geographic focus, among others*

Source of Data: *Custody and administrative service data feeds of daily cash flow transactions and quarterly net asset values prepared for Limited Partners holding stakes in private funds. State Street receives and maintains the commitment percentage of each limited partner client for each private equity fund in which it has invested. It will then choose one representative LP invested in the fund and gross up its cash flows by its commitment percentage in the relevant partnership to derive the entire private equity fund's data*

Major Reported Metrics: *Total and horizon IRRs, Multiples to Paid-in Capital, PME, Direct Alpha (labeled Private Alpha), Modified Dietz, Cash flow ratios on monthly basis*

All performance metrics are net of fees and carries, to reflect the true experience of LPs/investors

Restating Historical Results: *Does not restate historical published results unless there is significant or material error. Each quarterly index database is locked after a publication and no changes can be made afterwards*

Ability to Create User-customized Benchmarks: *Yes, through a web-based analytical platform*

Aggregation and Drilldown Options: *Reports pooled or any custom weighted, including equal-weighted, benchmark results, at both aggregated level or detailed granular level.*

Reports median, top, and bottom quartiles cutoff return levels

Provides breakdown of each benchmark by various sorts of classifications and characteristics: vintage year, geography, focused sector, fund size, etc.

REFERENCES

AICPA. *Accounting and Valuation Guide: Valuation of Portfolio Company Investments of Venture Capital and Private Equity Funds and Other Investment Companies*. Hoboken, NJ: John Wiley & Sons, 2019.

Aksia, LLC. 2020. "Private Equity Performance Report as of June 30, 2020." *Los Angeles City Employee Retirement Systems*. https://www.lacers.org/sites/main/files/file-attachments/private_equity_performance_as-of_june_30_2020.pdf?1606953495.

Ang, A. 2013. "Illiquid Asset Investing." Columbia Business School Research Paper No. 13-2. Available at SSRN: https://ssrn.com/abstract=2200161.

Ang, A., and N. P. B. Bollen. 2008. "Locked Up by a Lockup: Valuing Liquidity as a Real Option." https://ssrn.com/abstract=1291842 or http://dx.doi.org/10.2139/ssrn.1291842.

Ang, A., and M. Sorensen. 2012. "Risks, Returns, and Optimal Holdings of Private Equity: A Survey of Existing Approaches." *Quarterly Journal of Finance* 2 (3).

Anson, M. 2013. "Performance Measurement in Private Equity: Another Look at the Lagged Beta Effect." *The Journal of Private Equity* 17 (1): 29–44.

——. 2017. "Measuring Liquidity Premiums for Illiquid Assets." *The Journal of Alternative Investments* 20 (2): 39–50.

Brown, S. J., and W. N. Goetzmann. 1997. "Mutual Fund Style." *Journal of Financial Economics* 43 (3): 373–399.

Buchner, A., C. Kaserer, and N. Wagner. 2014. "Private Equity Funds: Valuation, Systematic Risk and Illiquidity." Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1102471.

Cambridge Associates. 2013. "New Method for Comparing Performance of Private Investments with Public Investments Introduced by Cambridge Associates." Cambridge Associates press release, https://www.cambridgeassociates.com/press-releases/new-method-for-comparing-performance-of-private-investments-with-public-investments-introduced-by-cambridge-associates/.

CFA Institute. 2021. "Private Company Valuation." *Curriculum*. https://www.cfainstitute.org/en/membership/professional-development/refresher-readings/private-company-valuation.

Cox, C. J., S. A. Ross, and M. Rubinstein. 1979. "Option Pricing: A Simplified Approach." *Journal of Financial Economics* 7 (3): 229–263.

Czasonis M., M. Kritzman, and D. Turkington. 2019. "Private Equity Valuations and Public Equity Performance." *The Journal of Alternative Investments* 22 (1): 8–19.

Deloitte. 2019. "The Growing Private Equity Market." *Deloitte Insights*. https://www2.deloitte.com/us/en/insights/industry/financial-services/private-equity-industry-forecast.html.

diBartolomeo, D., and C. Kantos. 2020. "How the Pandemic Taught Us to Turn Smart Beta into Real Alpha." *Journal of Asset Management* 21 (7): 581–590.

diBartolomeo, D., and E. Witkowski. 1997. "Mutual Fund Misclassification: Evidence Based on Style Analysis." *Financial Analyst Journal* 53 (5): 32–43.

Franzoni, F., E. Nowak, and L. Phalippou. 2012. "Private Equity Performance and Liquidity Risk." *Journal of Finance* 67 (6): 2341–2373.

Geltner, D. 1991. "Smoothing in Appraisal-Based Returns." *Journal of Real Estate Finance and Economics* 4 (1): 327–345.

Gensler, G. 2021. "Testimony Before the Subcommittee on Financial Services and General Government, U.S. House Appropriations Committee." *US Security and Exchange Commission*. https://www.sec.gov/news/testimony/gensler-2021-05-26.

Gredil, O., B. E. Griffiths, and R. Stucke. 2014. "Benchmarking Private Equity: The Direct Alpha Method." Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2403521.

Harris, R. S., T. Jenkinson, and S. N. Kaplan. 2014. "Private Equity Performance: What Do We Know?" *Journal of Finance* 69 (5): 1851–1882.

Ilmanen, A., S. Chandra, and N. McQuinn. 2020. "Demystifying Illiquid Assets: Expected Returns for Private Equity." *The Journal of Alternative Investments* 22 (3): 8–22.

IPEV Board. 2018. "International Private Equity and Venture Capital Valuation (IPEV) Guidelines." *Valuation Guidelines*. http://www.privateequityvaluation.com/Valuation-Guidelines.

Kaplan, S. N., and A. Schoar. 2005. "Private Equity Performance: Returns, Persistence, and Capital Flows." *Journal of Finance* 60 (4): 1791–1823.

Long, A. M., and C. J. Nickels. 2006. "Method for Calculating Portfolio Scaled IRR." *US Patent and Trademark Office*. https://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PALL&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.htm&r=1&f=G&l=50&s1=7,058,583.PN.&OS=PN/7,058,583&RS=PN/7,058,583.

MassPRIM. 2020. "Comprehensive Annual Financial Report, Fiscal Years Ended June 30, 2020 and 2019." *Pension Reserves Investment Trust Fund*. https://www.mapension.com/wp-content/uploads/2020/12/CAFR_Fiscal_Year_2020.pdf.

Mueller, E. 2011. "Returns to Private Equity: Idiosyncratic Risk Does Matter." *Review of Finance* 15 (3): 544–574.

Nadauld, T. D., B. A. Sensoy, K. Vorkink, and M. S. Weisbach. 2019. "The Liquidity Cost of Private Equity Investments: Evidence from Secondary Market Transactions." *Journal of Financial Economics* 132 (3): 158–181.

Ross, S. A. 1976. "The Arbitrage Theory of Capital Asset Prices." *Journal of Economic Theory* 13 (3): 341–260.

Rouvinez, C. 2003. "Private Equity Benchmarking with PME." *Venture Capital Journal*, August 1.

Sague, G. 2017. "Understanding the Nature of Venture Capital Returns." Medium.com. https://medium.com/@guillemsague/understanding-the-nature-of-venture-capital-returns-95846c65c049.

Sharpe, W. 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance* 19 (3): 425–442.