Benchmarking Gulf Private-Market Returns Against Global Portfolios

How to measure whether a Gulf allocation has actually earned its place, net of fees, risk and the right benchmark

Chennakeshav Adya

Independent Researcher

June 2026

ABSTRACT

Every claim that a market or a manager has performed well rests, often implicitly, on a benchmark, and the choice of benchmark, and of the measure applied to it, frequently does more to determine the verdict than the underlying performance. This paper examines how the private-market returns of a Gulf allocation should be benchmarked against the global portfolio an institution would otherwise hold, so that the question "did the Gulf sleeve earn its place?" can be answered honestly. Drawing on the literature on private-equity performance measurement, the public-market equivalent, and risk- and fee-adjusted comparison, it advances five propositions concerning how a Gulf allocation should be assessed. Using a stylised, clearly-labelled framework, it shows how the same underlying performance can support very different conclusions depending on whether it is measured gross or net of fees, by internal rate of return or by a cash-multiple, against cash or against the investor's actual opportunity cost, and before or after adjustment for risk and illiquidity. The analysis finds that a fair benchmarking of a Gulf allocation, net of fees, measured by the public-market equivalent against the investor's real alternative, and adjusted for risk, supports a positive but more modest verdict than the headline gross figures suggest, and that the dispersion across managers and vintages is large enough that selection dominates the average. The paper sets out a disciplined benchmarking approach, discusses the data limitations that constrain it, and identifies avenues for further, evidence-based research.

JEL Classification: G11, G23, G24, C43, O53

Keywords: benchmarking, private markets, public market equivalent, performance measurement, Gulf allocation, net-of-fee returns, vintage dispersion, risk adjustment

1. INTRODUCTION

Behind every statement that an investment has done well lies a comparison, and the comparison is rarely neutral. To say that a Gulf allocation has performed is to say that it has performed relative to something, cash, a public index, the investor's other options, and the choice of that something, together with the measure applied to it, frequently determines the verdict more than the performance itself. An allocation that looks impressive measured gross and against cash can look ordinary measured net of fees and against the investor's actual opportunity cost. For an institution deciding whether a Gulf allocation has earned its place, and whether to add to it, the discipline of benchmarking, choosing the right comparator and the right measure, is therefore not a technicality but the substance of the judgement.

This paper examines how the private-market returns of a Gulf allocation should be benchmarked against the global portfolio an institution would otherwise hold. It is written for the chief investment officer and the investment committee that have allocated, or are considering allocating, to the region, and that now face the practical question of how to assess whether the allocation is working, honestly and on a like-for-like basis. Its purpose is not to assert that the Gulf has outperformed or underperformed, which depends on the specific allocation, but to establish the framework within which that question can be answered without the distortions that careless benchmarking introduces.

The motivation is that private-market performance is unusually easy to misrepresent, not through dishonesty but through the choice of measure. The internal rate of return, the headline figure managers report, can be flattered by the timing of cash flows and by leverage; gross returns ignore the fees that determine what the investor keeps; comparison against cash or bonds flatters any risk asset; and the wide dispersion across managers and vintages means the average tells an institution little about what it actually experienced. A Gulf allocation, expressed substantially through private assets and assessed by an institution unfamiliar with the region, is especially exposed to these distortions, which makes a disciplined benchmarking framework particularly valuable here.

The paper makes three contributions. First, it sets out how the same underlying Gulf performance can support different conclusions depending on the benchmarking choices, gross versus net, the measure used, the comparator chosen, and the treatment of risk and illiquidity, and it argues for the choices that yield an honest verdict. Second, it applies the public-market-equivalent approach to ask the question that matters to an institution: did the Gulf sleeve beat what the investor would otherwise have held? Third, it emphasises the dominance of manager and vintage dispersion over the average, with implications for how an institution should interpret any benchmark. Throughout, the figures are modelled and clearly labelled; they illustrate the mechanics of benchmarking rather than report the performance of any specific allocation, and they are not forecasts.

The remainder of the paper is structured as follows. Section 2 reviews the literature on private-market performance measurement, the public-market equivalent, and risk- and fee-adjusted comparison, and develops five propositions. Section 3 describes the data and the stylised methodology. Section 4 presents and discusses the results, covering gross versus net returns, the choice of measure, the public-market equivalent, vintage and manager dispersion, risk adjustment and the dependence of the conclusion on the benchmark. Section 5 reports robustness. Section 6 sets out implementation considerations. Section 7 offers concluding comments, including limitations and directions for further research. A statement of disclosures, the references and appendices follow.

2. LITERATURE REVIEW AND PROPOSITIONS

The analysis draws on three strands of the literature: the measurement of private-equity and private-market performance; the public-market-equivalent approach to benchmarking private against public returns; and the treatment of fees, risk and illiquidity in performance comparison. This section reviews each and develops the propositions the analysis examines.

2.1 Measuring Private-Market Performance

A substantial literature addresses the difficulty of measuring private-market performance and the limitations of the common metrics (Phalippou and Gottschalg, 2009; Harris, Jenkinson and Kaplan, 2014). The internal rate of return, the most frequently reported measure, is sensitive to the timing of cash flows and can be manipulated, for example through subscription lines that defer capital calls, while multiples such as total value to paid-in capital (TVPI) and distributions to paid-in capital (DPI) convey what was returned but not when, and gross figures omit the fees that determine the investor's net result. The consistent lesson is that no single metric captures private-market performance, that gross and net can diverge materially, and that an honest assessment requires several measures read together. For a Gulf allocation this means resisting the headline internal rate of return in favour of a fuller, net picture, which is the substance of the first proposition.

2.2 The Public-Market Equivalent

A central methodological advance in private-market benchmarking is the public-market equivalent (PME), which asks whether an investor would have done better or worse by putting the same cash flows into a public-market index instead (Kaplan and Schoar, 2005; Sorensen and Jagannathan, 2015). A PME above one indicates the private allocation beat the public alternative on a cash-flow-matched basis; below one indicates it did not. The PME's great virtue is that it benchmarks a private allocation against the investor's genuine opportunity cost, the public portfolio it would otherwise have held, rather than against cash or an arbitrary hurdle, and it accounts for the timing of cash flows that distorts the internal rate of return. For an institution asking whether a Gulf allocation earned its place, the PME against the institution's actual global portfolio is the most relevant single measure, which the analysis develops as the second proposition.

2.3 Fees, Risk and Illiquidity

A third strand concerns the adjustments required for a fair comparison. The literature on net-of-fee returns establishes that fees and carry consume a substantial and variable share of gross private-market returns, so that gross comparisons systematically overstate what the investor keeps (Phalippou, 2009). The literature on risk-adjusted comparison establishes that return must be assessed per unit of risk, and that private-market returns require care because reported volatility is understated by infrequent and smoothed valuations (Ang, 2014). And the literature on the illiquidity premium establishes that part of a private allocation's excess return is compensation for illiquidity rather than skill, which a fair benchmark should recognise (Ang, 2014; Ilmanen, 2011). Together these imply that a fair Gulf benchmark must be net of fees, adjusted for risk including the understatement of private volatility, and clear about how much of any excess is an illiquidity premium, which the analysis addresses as the third and fourth propositions.

2.4 Dispersion and the Limits of the Average

A final strand emphasises the wide dispersion of private-market returns across managers and vintages, and the consequent limits of any average (Kaplan and Schoar, 2005; Harris et al., 2014). Because the gap between top- and bottom-quartile managers and between strong and weak vintages is large and, in private markets, persistent, the average return of a market or strategy tells an individual investor little about the return it actually experienced, which depends overwhelmingly on its manager selection and entry timing. For a Gulf allocation, where the manager base is younger and dispersion may be wider still, this means that any benchmark of the

average must be read alongside the dispersion, and that the institution's own selection determines where in the distribution it lands. This is the fifth proposition.

2.5 Synthesis and Propositions

Taken together, these literatures imply that an honest benchmarking of a Gulf allocation must be net of fees, measured by an opportunity-cost benchmark such as the PME rather than against cash, adjusted for risk and illiquidity, and read alongside the dispersion that the average conceals. From this the paper develops five propositions:

Proposition 1. The verdict on a Gulf allocation depends heavily on whether performance is measured gross or net and by which metric; net measures and the PME give a materially more modest and more honest result than the headline internal rate of return.

Proposition 2. Benchmarked by the public-market equivalent against the investor's actual global portfolio, a well-selected Gulf allocation can show genuine outperformance, but smaller than gross-versus-cash comparisons imply.

Proposition 3. A fair comparison must be net of fees, which consume a substantial and variable share of gross private-market returns.

Proposition 4. Part of any Gulf excess return is compensation for illiquidity and risk rather than skill, and a fair benchmark must recognise this.

Proposition 5. Manager and vintage dispersion dominate the average, so the institution's own selection, not the market average, determines its experienced return.

2.6 Benchmark Gaming and the Politics of Measurement

A strand of the literature and of practitioner experience worth drawing out concerns the ways performance measurement can be gamed, deliberately or not, through the choice of benchmark and metric. Because managers are typically the ones who report performance, and because their compensation and fundraising depend on it, there is a structural incentive to present figures in the most favourable light: to quote gross rather than net, the internal rate of return rather than a cash multiple, performance against cash rather than against the investor's alternative, and the strongest vintage rather than the average (Phalippou, 2009). None of this need involve dishonesty; it is the natural consequence of allowing the measured party to choose the measure. The institutional response, well established in the literature on investment governance, is for the investor to define the benchmark and the metrics itself, in advance and independently of the manager, so that the assessment is conducted on the investor's terms rather than the manager's. For a Gulf allocation, where the institution is unfamiliar and may be inclined to accept managers' framing, this discipline is especially important: the institution must own the benchmark, not borrow it from the party being benchmarked. This is the governance counterpart to the methodological argument of the paper, and it is the reason the implementation section insists that the benchmark and measures be fixed before the result is known.

2.7 Why Benchmarking Is Harder in a Younger Market

A final contextual point is that benchmarking is genuinely harder in a younger market such as the Gulf than in a mature one, and acknowledging this is part of an honest framework. In a mature market, established public indices, long return histories, and deep peer datasets provide robust comparators and dispersion statistics; in a younger market, public indices may be thin, return

histories short, and peer datasets sparse, so the comparators the framework requires are themselves harder to construct. This difficulty does not excuse careless benchmarking, it makes disciplined benchmarking more important, but it does mean an institution must work harder to assemble fair comparators and must hold its conclusions more tentatively where the data are thin. The practical responses are to use the best available global and regional comparators while being explicit about their limitations, to lean on cash-flow-based measures such as the public-market equivalent that are less dependent on a long index history, and to treat the resulting verdict as a reasoned estimate rather than a precise measurement. The maturation of the region's data infrastructure will ease this over time, and one contribution an institution can make, to its own future assessments and to the field, is to record its Gulf cash flows and outcomes rigorously so that a robust regional dataset eventually emerges. Until then, the honest position is that the framework specifies the right questions, while the data permit only approximate answers, a limitation the conclusion section addresses directly.

3. DATA AND METHODOLOGY

This study is a methodological and framework paper supported by stylised, clearly-labelled illustration rather than an empirical study of realised Gulf returns, which the region's still-developing data infrastructure does not yet support robustly. Its aim is to show how benchmarking choices drive the verdict and to argue for the choices that yield an honest one, using assumptions stated transparently. The figures illustrate the mechanics and relationships an institution should apply to its own data; they are not estimates of actual market performance and are not forecasts.

3.1 The Benchmarking Framework

The framework assesses a Gulf allocation through a sequence of increasingly honest comparisons: from gross return against cash, the most flattering and least meaningful, through net return, the choice of measure, the public-market equivalent against the investor's actual portfolio, and finally risk- and illiquidity-adjusted comparison. At each step the framework shows how the apparent performance changes, so that an institution can see how much of a headline result survives a fair assessment. The framework also requires that the average be read alongside the dispersion across managers and vintages, since the average alone is misleading.

3.2 Assumptions and Their Justification

The stylised inputs, indicative gross and net returns for Gulf real estate, private credit and private equity or venture, a representative global 60/40 portfolio as the opportunity-cost benchmark, and indicative PME, dispersion and risk figures, are calibrations informed by the private-market performance literature (Harris et al., 2014; Kaplan and Schoar, 2005) and by the general behaviour of these asset classes. They are representative rather than precise estimates and are intended to illustrate the benchmarking mechanics; Appendix A records them. The base-case assumptions are set out in Table 1.

Table 1. Base-Case Assumptions for the Benchmarking Framework

Parameter	Assumption
Opportunity-cost benchmark	Global 60/40 portfolio (~7.6% net)
Gulf real estate (gross / net)	12.0% / 9.5%

Gulf private credit (gross / net)	15.5% / 12.5%
Gulf PE / VC (gross / net)	16.0% / 11.5%
Public-market equivalent (PME)	1.18 to 1.30 (well-selected)
Top-to-bottom-quartile spread	~12 percentage points of net IRR
Illiquidity premium (share of excess)	Material; part of excess is not skill

Indicative, forward-looking assumptions used to illustrate benchmarking mechanics; not estimates of actual returns or forecasts. See Appendix A.

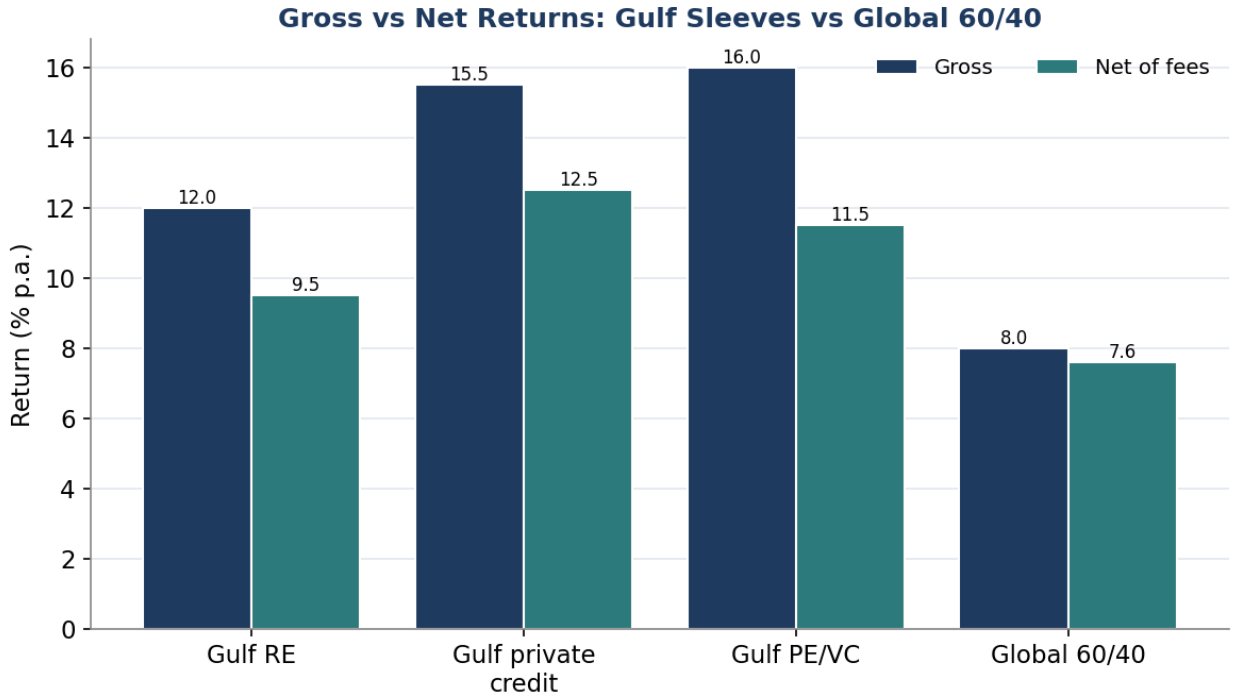
4. RESULTS AND DISCUSSION

This section presents the benchmarking framework and the supporting analysis in the order of the propositions: gross versus net and the choice of measure (Proposition 1); the public-market equivalent (Proposition 2); fees (Proposition 3); risk and illiquidity (Proposition 4); and dispersion (Proposition 5), together with the dependence of the conclusion on the benchmark chosen.

4.1 Gross versus Net, and the Choice of Measure

The first and most consequential benchmarking choices are whether to measure gross or net of fees, and which metric to use. Figure 1 contrasts gross and net returns for the Gulf sleeves against the global 60/40 benchmark.

Figure 1. Gross versus Net Returns: Gulf Sleeves versus Global 60/40

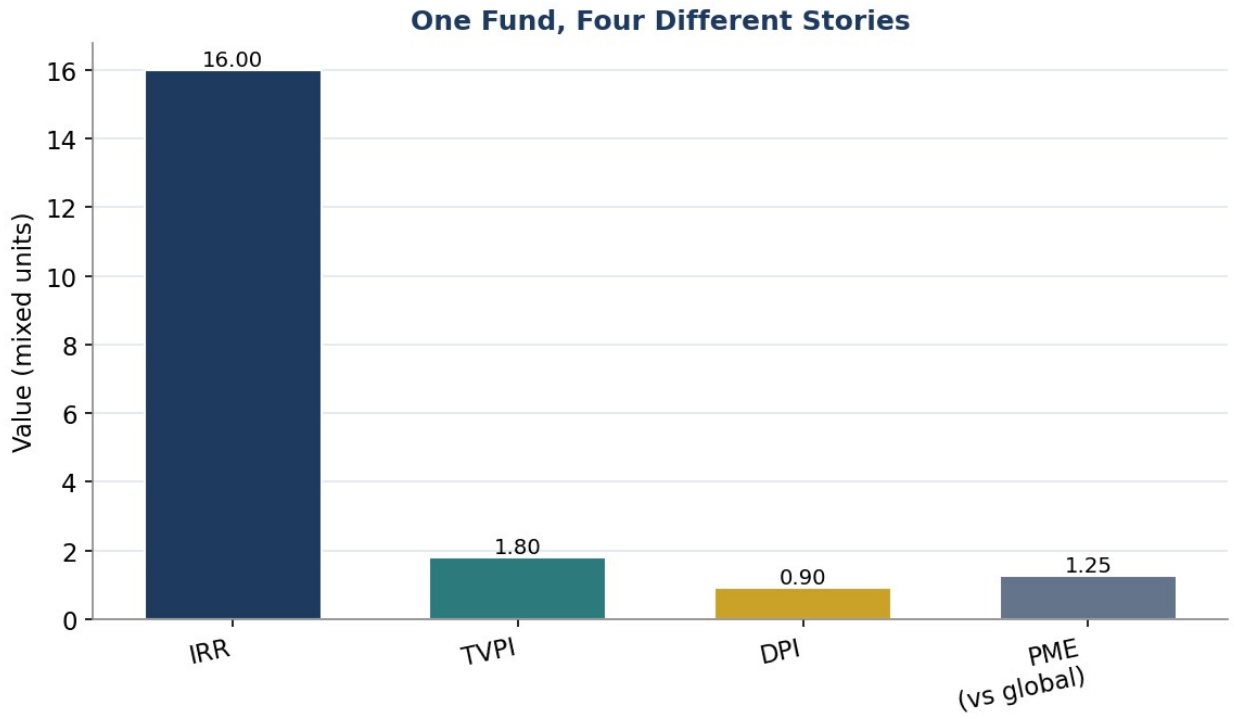


Indicative gross and net returns; the gap between gross and net is the fee load the investor does not keep.

The gap between the gross and net bars is the fee load, the share of the gross return the investor does not keep, and it is large enough that a comparison on gross figures materially overstates the advantage of the Gulf sleeves over the global benchmark. This supports the fee component of Proposition 1 and Proposition 3. The choice of metric compounds the issue: Figure 2 shows how

a single allocation can tell four different stories depending on whether it is described by internal rate of return, by a value multiple, by a cash-return multiple or by a public-market equivalent.

Figure 2. One Allocation, Four Different Stories



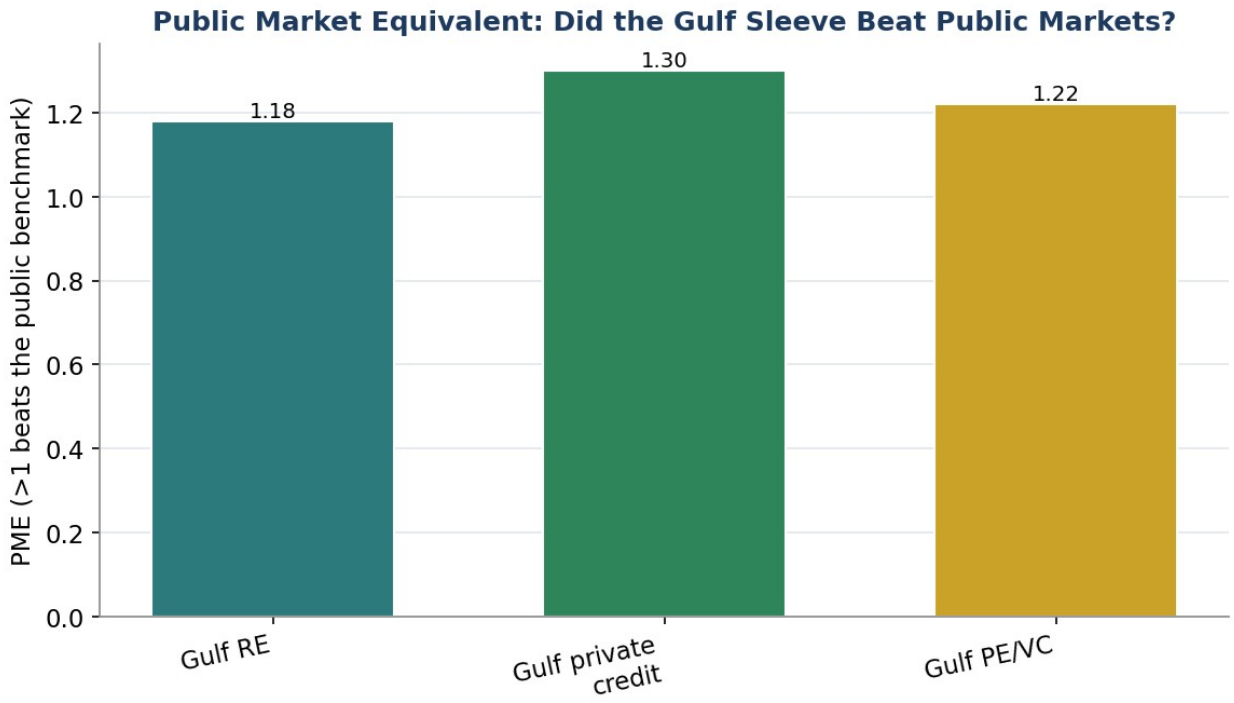
Indicative; the same allocation summarised by IRR, TVPI, DPI and PME conveys very different impressions.

The metrics tell genuinely different things. The internal rate of return is high but sensitive to cash-flow timing and leverage; the value multiple shows total value created but not when; the cash-return multiple shows what has actually been distributed, which is lower; and the public-market equivalent shows performance relative to the investor's alternative. An institution that anchors on the internal rate of return alone, the figure managers most readily report, will form a more favourable impression than the fuller picture supports. The discipline Proposition 1 requires is to read these measures together and to weight the net, distribution-based and opportunity-cost measures most heavily, because they best describe what the investor actually keeps and forgoes.

4.2 The Public-Market Equivalent: Did It Beat the Alternative?

The question that matters most to an institution is not whether a Gulf allocation produced a positive return but whether it beat what the institution would otherwise have held. The public-market equivalent answers exactly this, and Figure 3 reports it for the Gulf sleeves against the global benchmark.

Figure 3. Public-Market Equivalent: Did the Gulf Sleeve Beat Public Markets?



Indicative PME; a value above one means the sleeve beat the cash-flow-matched public alternative.

The PME values, above one for the well-selected sleeves, support Proposition 2: a well-chosen Gulf allocation can genuinely beat the investor's public-market alternative on a cash-flow-matched basis. But two qualifications matter. First, the outperformance the PME shows is more modest than the gross-versus-cash comparison implies, because the PME is net and benchmarked against the investor's real alternative rather than against cash; the discipline of the PME deflates the headline. Second, the PME shown is for a well-selected allocation, and, as Section 4.4 establishes, the dispersion is wide enough that a poorly selected allocation could show a PME below one, meaning it underperformed the public alternative. The PME is therefore the right measure precisely because it benchmarks against the genuine opportunity cost and on a net basis, and its verdict, positive but modest for good selection and potentially negative for poor selection, is the honest one that Proposition 2 anticipates.

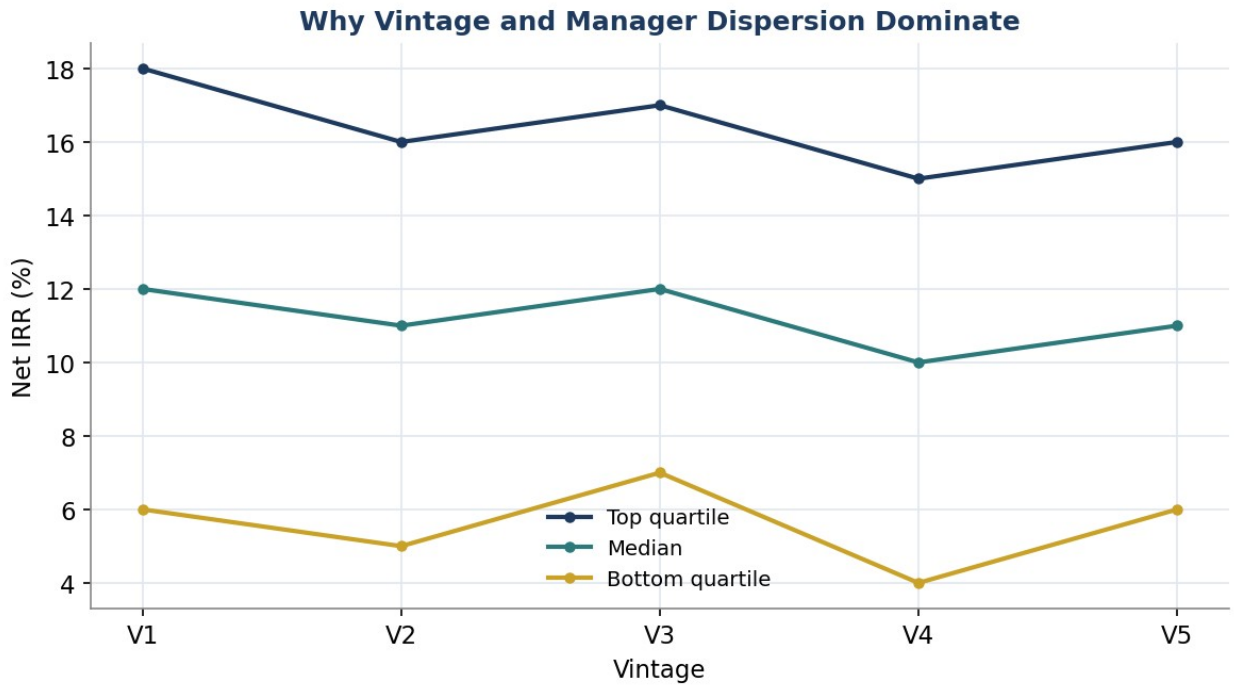
4.3 The Weight of Fees

The fee load deserves separate emphasis because it is the most controllable determinant of the net result and the one most often understated in benchmarking. The gap between the gross and net bars in Figure 1 represents fees and carry, which in private markets consume a substantial and variable share of the gross return (Phalippou, 2009). For a Gulf allocation the implication is twofold. First, any benchmarking must be conducted net, because the gross figure is not what the investor keeps and a gross comparison flatters the allocation. Second, because the fee load varies across managers and structures, two allocations with the same gross return can deliver materially different net results, which means fee negotiation and structure selection, not only gross performance, determine the benchmarked outcome. This connects benchmarking to the broader discipline of net-to-investor returns: the institution that benchmarks net, and that negotiates the fees and waterfall that determine the net, will both measure and achieve a better result than one that focuses on gross performance alone.

4.4 Dispersion: Why the Average Misleads

The single most important caveat in benchmarking a Gulf allocation is that the average return tells an individual institution little about its own experience, because the dispersion across managers and vintages is large. Figure 4 illustrates the spread.

Figure 4. Why Vintage and Manager Dispersion Dominate



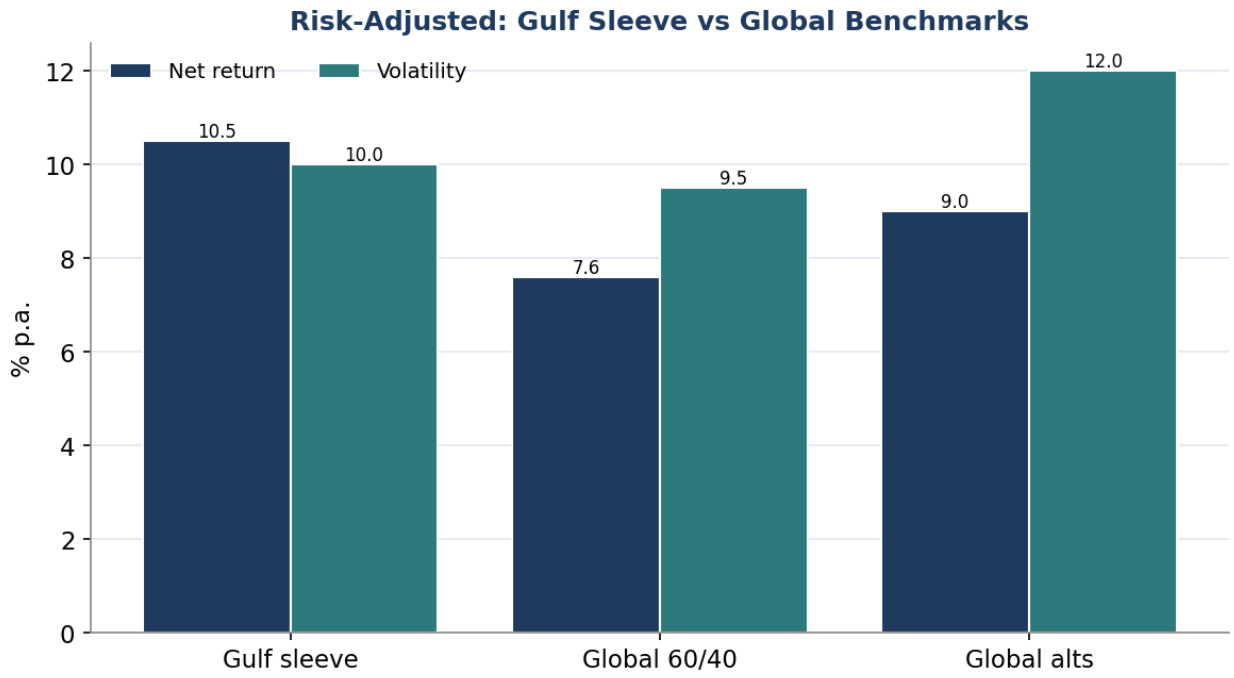
Indicative top-quartile, median and bottom-quartile net IRR across vintages; the spread is wide and persistent.

The dispersion supports Proposition 5 and qualifies every other result in the paper. The gap between top- and bottom-quartile managers and between strong and weak vintages is wide enough that an institution's experienced return depends overwhelmingly on its selection and timing, not on the market average. A benchmark of the average is therefore necessary but insufficient: it tells the institution how the market did, not how it did, and a poorly selected allocation can underperform even a strong market average while a well-selected one can substantially exceed it. The practical implication is that benchmarking must be read alongside the dispersion, and that an institution assessing its Gulf allocation should compare its own managers against the relevant quartiles rather than against the average, since the average is a description of the market and not of the institution's decisions. This also reinforces the central message of the companion analysis on first-allocation diligence: in a market with wide dispersion, manager selection is where the return is won or lost.

4.5 Risk and Illiquidity Adjustment

A fair benchmark must also adjust for risk and for the illiquidity premium, in line with Proposition 4. Figure 5 reports the Gulf sleeve against global benchmarks on a risk-adjusted basis.

Figure 5. Risk-Adjusted: Gulf Sleeve versus Global Benchmarks



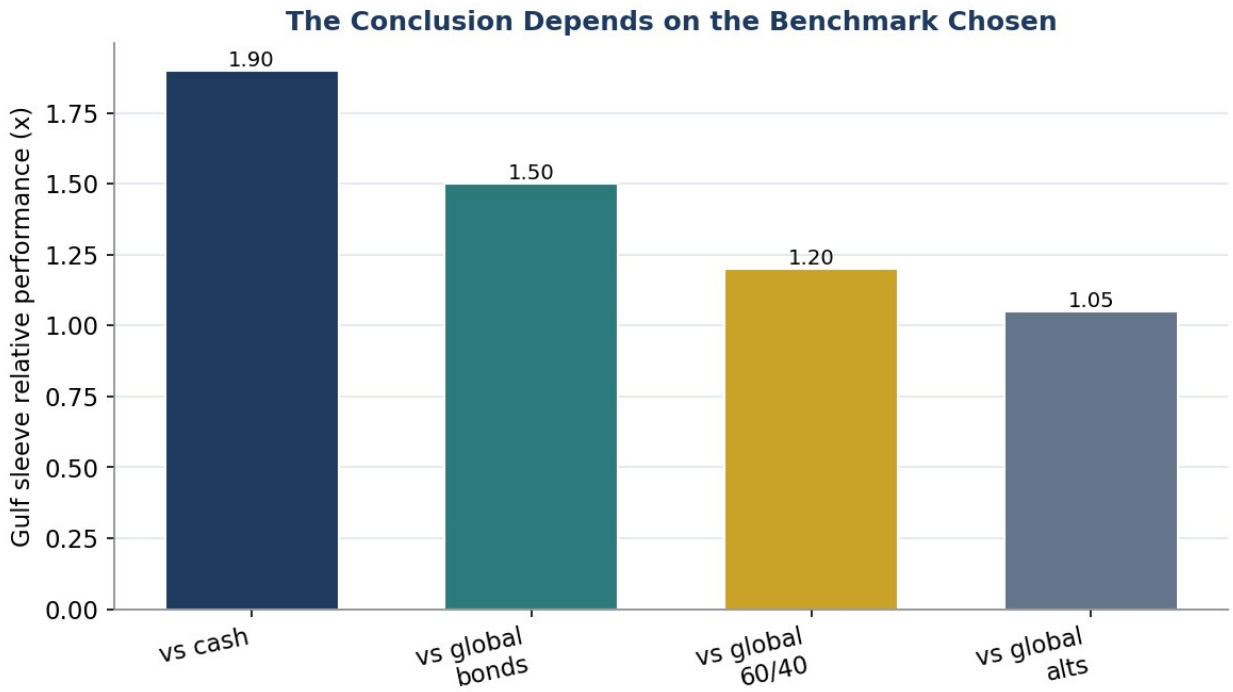
Indicative net return and volatility; private volatility is understated by smoothed valuations and should be adjusted.

Two adjustments matter. The first is that reported private-market volatility is understated by infrequent and smoothed valuations, so a naive risk-adjusted comparison flatters private allocations relative to public ones; an honest comparison should adjust private volatility upward toward its economic level (Ang, 2014). The second is that part of any Gulf excess return is compensation for bearing illiquidity rather than evidence of skill, so a fair benchmark should recognise that an illiquid allocation should earn an illiquidity premium and should not credit the whole excess to the manager or the region. After both adjustments, the Gulf sleeve's advantage over the global benchmark remains positive in the well-selected case but is more modest still, which is the honest conclusion. The point is not that the adjustments eliminate the advantage but that they right-size it: an institution that omits them will overstate the skill-based, risk-adjusted outperformance of its Gulf allocation, and may draw the wrong lesson about how much of the result to attribute to selection versus to compensated risk.

4.6 The Conclusion Depends on the Benchmark

Pulling the threads together, Figure 6 shows how the assessed performance of the same Gulf sleeve changes with the benchmark chosen.

Figure 6. The Conclusion Depends on the Benchmark Chosen



Indicative relative performance of the same sleeve against cash, bonds, a 60/40 portfolio and global alternatives.

Against cash the Gulf sleeve looks spectacular; against global bonds, very strong; against the investor's actual 60/40 portfolio, solidly positive; and against a global alternatives benchmark, only modestly ahead. The same performance, four verdicts, and the difference is entirely the benchmark. This is the central methodological message of the paper: the choice of benchmark is not a neutral technicality but the largest single determinant of the verdict, and an institution that wishes to assess its Gulf allocation honestly must benchmark it against its genuine opportunity cost, the portfolio it would otherwise have held, and against the closest comparable alternatives, rather than against cash or bonds that flatter any risk asset. The honest verdict, net, against the right benchmark, risk- and illiquidity-adjusted, is positive for a well-selected allocation but materially more modest than the flattering comparisons suggest, and dependent on selection.

4.7 The J-Curve and Interim Benchmarking

A specific difficulty in benchmarking a Gulf allocation, or any private-markets programme, in its early years is the J-curve: the tendency for a programme to show negative or weak returns in its early life as capital is called and fees charged before investments mature, with the positive returns arriving later. The implication for benchmarking is that interim performance, measured a few years into a programme, can look poor even for an allocation that will ultimately do well, and conversely that an allocation flattered by early paper gains may disappoint later. An institution that benchmarks too early, or that judges its Gulf allocation on interim marks against a benchmark of realised public returns, will draw misleading conclusions. The disciplined approach is to benchmark with the J-curve in mind: to use measures such as the public-market equivalent that account for the timing of cash flows; to compare against vintage-matched peers rather than against a contemporaneous public index; and to defer firm conclusions until the programme has matured sufficiently for realised distributions, not just paper marks, to carry weight. For a first-time Gulf allocator especially, patience in benchmarking is itself a discipline, because the

temptation to judge the allocation prematurely, and to act on that premature judgement, is strong and the early signal is weak.

This difficulty interacts with the smoothed-valuation problem discussed earlier. Because private-asset valuations are appraisal-based and infrequent, interim performance is not only affected by the J-curve but reported with a lag and a smoothing that understates volatility and can flatter or flatter-then-disappoint the interim picture. The combination means that interim benchmarking of a private Gulf allocation should be treated as indicative rather than conclusive, and that the institution should weight realised, cash-based measures, distributions to paid-in capital and the public-market equivalent on realised flows, more heavily as the programme matures. Building this understanding into the institution's benchmarking from the outset prevents the premature, marks-driven judgements that lead institutions to abandon sound allocations or to add to weak ones on the strength of an unreliable interim signal.

4.8 Attribution: Skill, Beta, Leverage and the Premium

Beyond asking whether a Gulf allocation outperformed, a thorough benchmarking asks why, because the source of any outperformance determines whether it is likely to persist. The literature on performance attribution distinguishes the components of a return: the market or beta component, which the investor could have obtained more cheaply elsewhere; the leverage component, which amplifies returns but adds risk; the timing component, which may be luck; the illiquidity premium, which is compensation for a risk borne rather than skill; and the residual, true skill or genuine market access. A Gulf allocation that outperformed principally through beta and leverage is a different proposition from one that outperformed through manager skill or genuine access to an under-intermediated market, even if the headline outperformance is identical, because the former is fragile and replicable while the latter is durable and scarce. A disciplined benchmarking therefore decomposes the result, crediting to skill only the residual after beta, leverage, timing and the illiquidity premium are accounted for. For the Gulf, where part of the return reflects an illiquidity and complexity premium and part reflects genuine access to a growing, under-intermediated market, this attribution is essential to forming a view on whether the performance will continue, and to deciding whether to add to the allocation. An institution that benchmarks only the headline, without attributing it, learns whether it did well but not whether it will, which is the more important question for a forward-looking decision.

4.9 Benchmarking by Sleeve, Not in Aggregate

A practical refinement that materially improves the honesty of the assessment is to benchmark each Gulf sleeve, real estate, private credit, private equity or venture, against its own appropriate comparator rather than benchmarking the Gulf allocation in aggregate against a single global index. The reason is that the sleeves have different risk, return and liquidity profiles and therefore different fair benchmarks: Gulf private credit should be judged against a comparable credit alternative, Gulf real estate against a real-estate or real-asset benchmark, and Gulf venture against a venture or growth benchmark. Aggregating them and comparing the whole against a single 60/40 portfolio blurs the assessment, because a strong result in one sleeve can mask a weak result in another, and because the aggregate comparison does not tell the institution which sleeve actually earned its place. Benchmarking by sleeve, against sleeve-appropriate comparators and on a net, opportunity-cost basis, yields a far more actionable verdict: it tells the institution not only whether the Gulf allocation as a whole worked but which parts of it did, which is precisely what it needs to know to decide where to add and where to retrench. This sleeve-level discipline is

more demanding than an aggregate comparison but it is the standard a serious benchmarking should meet, and it connects to the broader point that the average, whether across managers, vintages or sleeves, conceals more than it reveals.

Sleeve-level benchmarking also guards against a subtle error in which a diversified Gulf allocation is credited with the diversification benefit and the return of its best sleeve simultaneously, double-counting the contribution. By assessing each sleeve against its own comparator and then considering the diversification of the combination separately, the institution avoids conflating the two and reaches a cleaner verdict on each component and on the whole. This separation, between the sleeve-level performance question and the portfolio-level diversification question, is a hallmark of disciplined benchmarking and a frequent casualty of the aggregate, single-benchmark shortcut.

4.10 The Asymmetry of Benchmarking Errors

It is worth considering the asymmetry between the two kinds of benchmarking error, because it bears on how cautious an institution should be. An institution can err by judging a sound allocation harshly, through an unfairly demanding benchmark, and thereby abandon or under-size a genuinely good position; or it can err by judging a weak allocation generously, through a flattering benchmark, and thereby retain or add to a genuinely poor position. Both errors are costly, but they are not symmetric in their drivers: the flattering error is the more common, because the incentives, managers' presentation, decision-makers' desire to validate their choices, the natural appeal of the impressive gross figure, all push toward generosity, while few forces push toward unfair harshness. The practical implication is that an institution's benchmarking discipline should be calibrated primarily against the flattering error, by insisting on net, opportunity-cost, risk-adjusted measures that strip out the sources of false comfort. This is not a counsel of pessimism; it is a recognition that the gravitational pull of performance measurement is toward flattery, and that the discipline must lean against that pull to land at an honest verdict. An institution aware of this asymmetry will treat impressive headline figures with particular scepticism and modest fair figures with appropriate respect, which is the correct posture given where the errors actually come from.

4.11 Benchmarking and the Decision to Add, Hold or Exit

The ultimate purpose of benchmarking is to inform a decision, to add to the allocation, hold it, or exit, and it is worth making explicit how the framework supports each. A decision to add should rest on a fair verdict that the allocation has genuinely outperformed its opportunity cost, net and risk-adjusted, and on an attribution showing that the outperformance reflects durable sources, skill and genuine access, rather than fragile ones, beta and leverage. A decision to hold is appropriate where the verdict is positive but the allocation is still maturing, so that the J-curve and interim valuations counsel patience before a firmer judgement. A decision to exit, or not to re-up, should rest on a fair verdict of underperformance that persists after accounting for the J-curve, or on an attribution showing the result depended on sources unlikely to continue, or on a finding that the institution's managers sit in the wrong part of the dispersion. In each case the framework supplies not a number but a structured basis for judgement, and it guards against the two errors that careless benchmarking invites: adding to a flatteringly measured weak allocation, and exiting a harshly measured sound one. By tying the add-hold-exit decision explicitly to a fair verdict, an attribution and the dispersion, the institution ensures that its most consequential

decisions about the allocation rest on honest assessment rather than on the impression created by a chosen benchmark.

4.12 Currency, the Peg and Cleaner Comparison

One feature of the Gulf simplifies benchmarking relative to other emerging markets and deserves note: the currency peg. Benchmarking an allocation to a floating-currency emerging market is complicated by the need to decide whether to compare in local or home-currency terms and how to treat the volatile currency effect, which can dominate and obscure the underlying asset performance. For a Gulf allocation, the dirham peg removes most of this complication: because the currency is effectively the dollar, a dollar-based comparison is clean, and the underlying asset performance is not obscured by a large, variable currency effect. This makes the benchmarking of a Gulf allocation, in this one respect, easier and more reliable than that of a typical floating emerging market, and it is the reason the sensitivity analysis found the currency treatment to be the least influential of the benchmarking choices. The practical benefit is that an institution can compare its Gulf allocation against a dollar-based global benchmark with confidence that it is comparing like with like on the currency dimension, and can concentrate its benchmarking effort on the choices that matter more, the net-of-fee treatment, the opportunity-cost comparator, the risk and illiquidity adjustments and the dispersion. The peg, examined at length in a companion analysis as a risk-reducer, is thus also, more quietly, a benchmarking simplifier, and it removes one of the sources of distortion that make the assessment of other emerging-market allocations so fraught.

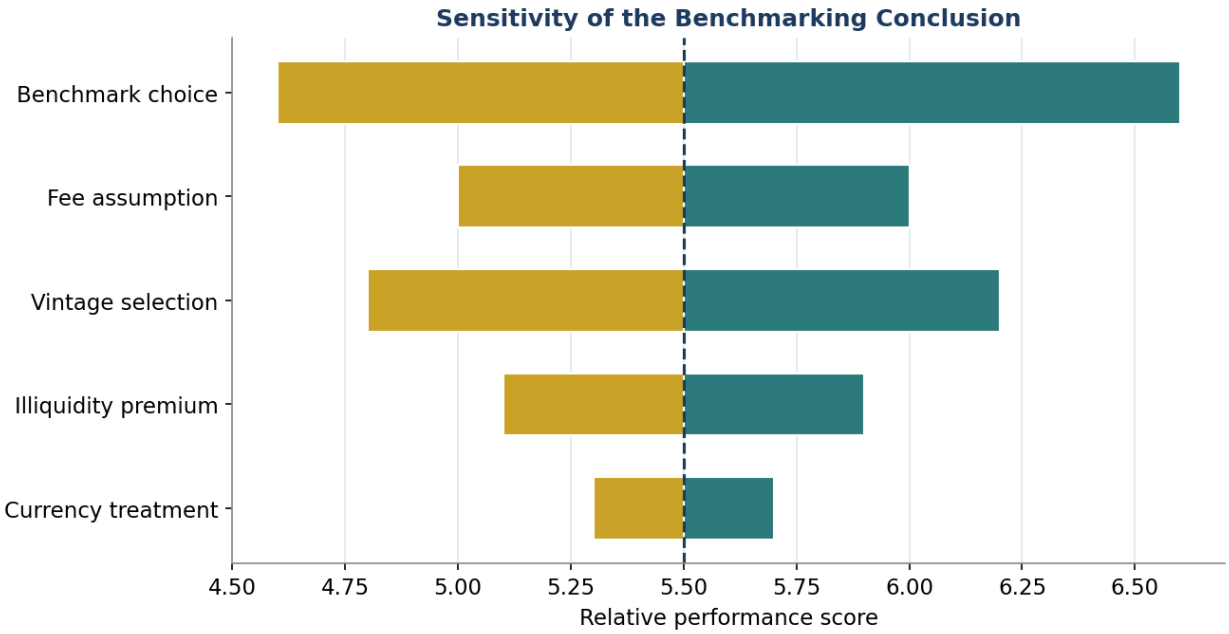
5. ROBUSTNESS AND SCENARIO ANALYSIS

A benchmarking framework must itself be tested: how sensitive is the verdict to the choices the framework makes, and how do different evaluators, applying different standards, reach different conclusions from the same data? This section reports a sensitivity analysis and a comparison of three benchmarking approaches.

5.1 Sensitivity of the Verdict

Figure 7 reports the sensitivity of the benchmarking conclusion to the principal choices, each varied around the base case.

Figure 7. Sensitivity of the Benchmarking Conclusion



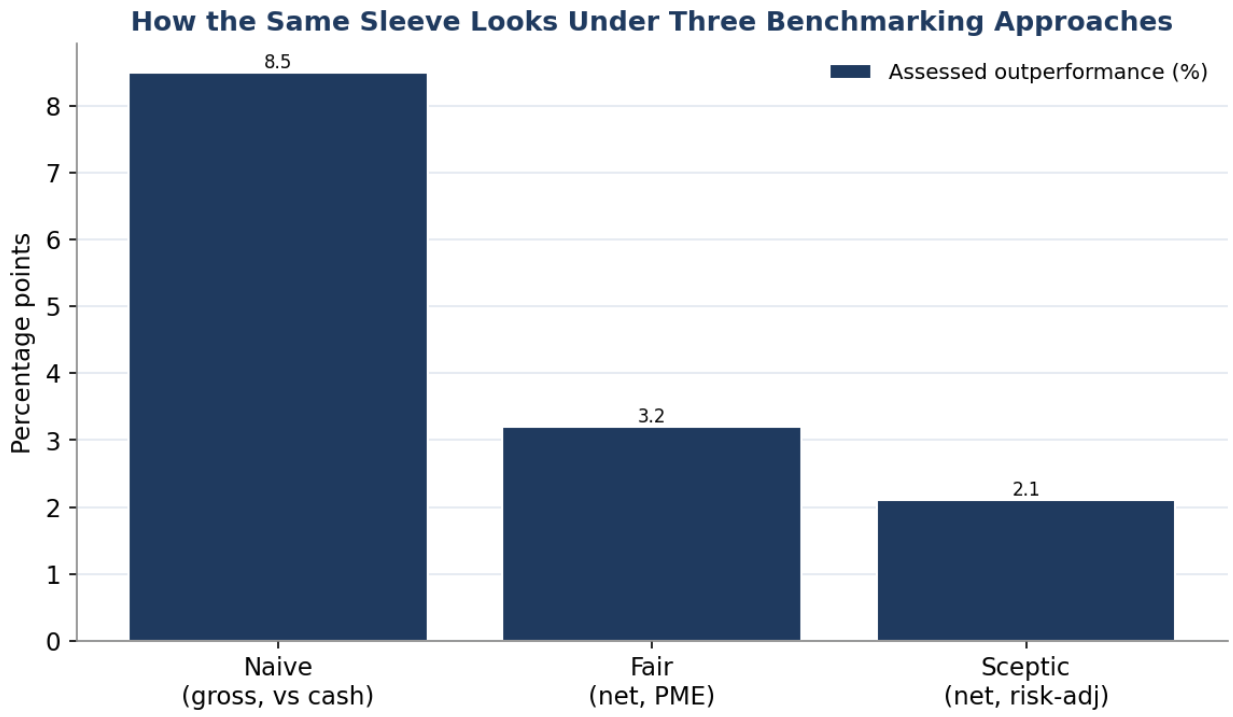
Each bar shows the swing in the relative-performance verdict from varying one benchmarking choice.

The benchmark choice and the fee assumption move the verdict most, confirming the central message that how the allocation is measured matters as much as how it performed. Vintage selection and the illiquidity-premium assumption follow, and the currency treatment matters least, because the peg removes most of the currency variability that would otherwise complicate the comparison. The ordering directs an institution's scrutiny: when assessing a benchmarking claim about a Gulf allocation, the first questions to ask are which benchmark was used and whether the figures are net of fees, because those two choices drive the conclusion more than any other. A claim of strong outperformance that rests on gross figures against cash should be treated with the scepticism the sensitivity analysis warrants.

5.2 Three Evaluators, Three Verdicts

Figure 8 makes the dependence on approach concrete by showing how three evaluators, applying different standards to the same allocation, reach different verdicts.

Figure 8. How the Same Sleeve Looks Under Three Benchmarking Approaches



Indicative assessed outperformance under a naive, a fair and a sceptical benchmarking approach.

The naive evaluator, measuring gross against cash, records large outperformance; the fair evaluator, measuring net by the public-market equivalent against the investor's portfolio, records modest but genuine outperformance; and the sceptical evaluator, additionally adjusting for risk and the illiquidity premium, records a smaller margin still. None is dishonest; they simply apply different standards, and the difference between their verdicts, all from the same underlying allocation, is the paper's thesis in a single figure. The fair and sceptical verdicts are the ones an institution should rely on, because they reflect what the investor actually keeps relative to its genuine alternative and after accounting for the risk and illiquidity borne. The naive verdict, though the most flattering and the most often quoted, is the least informative, and an institution that governs by it will both over-state its success and mis-allocate on the strength of a comparison that means little.

5.3 Building a Benchmarking Policy

The practical embodiment of this paper's argument is a written benchmarking policy, adopted before the allocation is assessed and ideally as part of the original mandate. Such a policy specifies the opportunity-cost benchmark against which the Gulf allocation will be judged, the primary measures, net returns, the public-market equivalent and a risk- and illiquidity-adjusted comparison, the treatment of the J-curve and interim performance, and the requirement to read the average alongside the dispersion. By fixing these in advance, the policy removes the discretion that allows a benchmark to be chosen after the fact to suit the result, and it ensures consistency across allocations and over time. A benchmarking policy is to performance assessment what an investment policy is to allocation: the discipline that prevents ad hoc, self-serving judgements and that makes the institution's assessments comparable, defensible and honest. For an institution building a Gulf programme, adopting such a policy at the outset is among the most valuable governance steps it can take, because it determines how every future assessment of the programme will be conducted.

5.4 Communicating Performance Honestly to Stakeholders

A benchmarking framework serves not only internal decision-making but the institution's communication with its own stakeholders, trustees, beneficiaries, regulators and, for a fund-of-funds, its investors. Honest benchmarking protects the institution from the reputational and governance risk of having presented a flattering picture that later proves misleading, and it builds the credibility that comes from assessing one's own decisions by a fair standard. The temptation, particularly for a novel allocation that decision-makers championed, is to present the flattering gross-versus-cash comparison; the discipline is to present the fair, net, opportunity-cost-benchmarked picture, including the dispersion and the attribution, so that stakeholders understand both what was achieved and why. An institution that communicates performance honestly, even when the honest verdict is more modest than the flattering one, is better governed and more trusted than one that does not, and it is better placed to make sound decisions about the allocation's future because its own understanding is undistorted. The benchmarking framework is therefore as much a tool of governance and communication as of measurement, and its adoption signals an institution that intends to assess itself by the same standard it would apply to a manager.

5.5 A Benchmarking Checklist

The practical guidance of this paper can be distilled into a checklist an institution can apply when assessing a Gulf allocation. Is the assessment net of all fees and carry rather than gross? Is the comparator the investor's genuine opportunity cost, the portfolio it would otherwise have held, rather than cash or bonds? Is the public-market equivalent used as a primary measure, on realised cash flows? Is each sleeve benchmarked against its own appropriate comparator rather than the aggregate against a single index? Is the result adjusted for risk, including the understatement of private volatility, and for the illiquidity premium? Is the average read alongside the dispersion, with the institution's own managers compared against the relevant quartiles? Is the result attributed to its sources, beta, leverage, timing, premium and skill, rather than credited wholesale to the manager or the region? And was the benchmark and were the measures fixed in advance, independently of the result and of the managers being assessed? An assessment that can answer yes to each of these is an honest one; an assessment that cannot is, to that extent, unreliable, and a verdict resting on it should be discounted accordingly.

5.6 Frequency and Horizon of Benchmarking

A practical question is how often, and over what horizon, a Gulf allocation should be benchmarked. The answer follows from the J-curve and the long horizons of private investment: benchmarking should be conducted regularly for monitoring, annually is typical, but firm conclusions about performance should be drawn over horizons long enough for the programme to have matured and for realised distributions to carry weight. An institution that benchmarks frequently but interprets each interim reading as a verdict will be whipsawed by the J-curve and by smoothed, lagged valuations; one that benchmarks regularly but reserves judgement for the mature horizon will track the allocation without over-reacting to noise. The framework therefore distinguishes monitoring benchmarking, frequent, indicative and used to watch for problems, from assessment benchmarking, conducted at maturity and used to judge whether the allocation earned its place and whether to continue. Confusing the two, treating an interim monitoring reading as a mature assessment, is a common error that leads institutions to act prematurely, and

separating them is a simple discipline that improves both the monitoring and the eventual judgement.

5.7 The Role of Independent Performance Assessment

A governance practice that materially improves benchmarking quality is to separate the assessment of performance from the advocacy for the allocation, just as operational due diligence is separated from investment advocacy in the selection process. Where the people who championed and manage a Gulf allocation are also the ones who assess and report its performance, the structural incentive toward flattering measurement, choosing the comparator and metric that validate the original decision, is strong and largely unconscious. Assigning the performance assessment to an independent function, or at least subjecting it to independent review, counteracts this incentive and produces a more honest verdict. The independent assessor applies the benchmarking policy, net, opportunity-cost, risk-adjusted, attributed and read against the dispersion, without the stake in the conclusion that the advocates hold. This is standard practice in well-governed institutions for good reason, and it is especially valuable for a novel allocation that decision-makers are invested in seeing succeed. The cost is modest and the benefit, an assessment the institution can genuinely trust, is large, because a benchmark produced by the party with a stake in its conclusion is worth little regardless of the rigour of the underlying method.

5.8 Avoiding Survivorship and Selection Bias in Peer Data

A technical but important implementation caution concerns the peer datasets an institution uses to construct its comparators and dispersion statistics, which are prone to biases that flatter the apparent performance of the asset class. Survivorship bias arises when failed funds drop out of a dataset, leaving a sample skewed toward successes and overstating the median and the quartiles; selection or self-reporting bias arises when managers choose whether to report, and the better performers are more likely to do so. Both biases inflate the benchmark of the peer group and, paradoxically, can make an institution’s own well-performing allocation look merely average against a flattered peer set, or can lead it to over-estimate how easy strong performance is to achieve. In a younger market with thinner data, these biases can be more severe because the surviving, reporting sample is smaller and less representative. The disciplined response is to be aware of the biases, to prefer datasets that mitigate them where available, and to treat peer benchmarks as indicative rather than precise, weighting the cash-flow-based opportunity-cost comparison, the public-market equivalent against an investable index, which is less vulnerable to these biases, more heavily than the peer-quartile comparison. Acknowledging the limits of peer data is part of honest benchmarking, and an institution that treats biased peer statistics as ground truth will reach distorted conclusions however carefully it applies the rest of the framework.

6. IMPLEMENTATION CONSIDERATIONS

The framework translates into a small number of practical commitments for an institution benchmarking a Gulf allocation.

6.1 Choose the Benchmark Before the Result

The first commitment is to define the benchmark, the investor’s genuine opportunity cost, and the measures, net, PME, risk- and illiquidity-adjusted, before assessing the allocation, so that the standard is not chosen after the fact to suit the result. Defining the benchmark in advance, ideally

in the original mandate, removes the temptation to select the comparator that flatters and ensures that the allocation is judged against the alternative the institution actually forwent.

6.2 Benchmark Net, Against the Real Alternative

The second commitment is to benchmark net of all fees and against the portfolio the institution would otherwise have held, using the public-market equivalent as the primary measure. This is the comparison that answers the question that matters, did the allocation beat the alternative?, and it is the comparison most resistant to the distortions that flatter private-market performance. Cash and bond benchmarks should be avoided as primary comparators, because they flatter any risk asset and tell the institution nothing about its real choice.

6.3 Read the Average Alongside the Dispersion

The third commitment is to interpret any benchmark of the average alongside the dispersion, comparing the institution's own managers against the relevant quartiles rather than the mean. Because selection dominates the experienced return, an institution should assess whether its managers are top-quartile or median, and should attribute its result to its selection rather than to the market. This also disciplines future decisions: an institution that recognises its result as the product of its selection, good or bad, learns the right lesson, whereas one that credits or blames the market average learns the wrong one.

It is worth drawing out a connection that runs through the analysis: that honest benchmarking and good decision-making are the same discipline viewed from two angles. The institution that benchmarks its Gulf allocation honestly, net, against its real alternative, risk-adjusted, attributed and read against the dispersion, is the institution best placed to decide whether to add to it, hold it or exit it, because it understands what the allocation actually achieved and why. The institution that benchmarks carelessly understands neither, and its decisions about the allocation's future will be correspondingly poor. Benchmarking is therefore not a backward-looking scorekeeping exercise but the foundation of forward-looking judgement, and the rigour an institution brings to assessing the past directly determines the quality of its decisions about the future. This is the deepest reason to invest in the discipline: not to produce a number, but to see clearly enough to decide well.

A final reflection concerns the relationship between honest benchmarking and the credibility of the broader case for the Gulf. It might be thought that a rigorous, deflating benchmarking framework works against the case for allocating to the region, since it shows the outperformance to be more modest than the flattering comparisons imply. The opposite is true. A case for the Gulf that rests on flattering, gross-versus-cash comparisons is fragile, because it collapses the moment a sceptic applies a fair benchmark; a case that survives honest benchmarking, net, against the real alternative, risk-adjusted, is robust precisely because it has been tested by the hardest fair standard. By providing that standard, this paper strengthens rather than weakens the credible case for the Gulf: it separates the genuine, defensible outperformance of a well-selected allocation from the illusory outperformance of a flatteringly measured one, and it equips the institution to pursue the former with confidence. The honest benchmark is the friend of the allocator who has chosen well and the enemy only of the one who has measured carelessly, and that is exactly as it should be.

A broader implication of the analysis is that benchmarking discipline is a marker of institutional quality that extends well beyond the Gulf. An institution that benchmarks honestly, that defines

its comparators in advance, measures net and against its real alternative, adjusts for risk and illiquidity, attributes its returns and reads the average against the dispersion, is an institution that thinks clearly about all of its investments, not only its Gulf allocation. The Gulf case is a particularly clear illustration of the principle because it is unfamiliar and prone to flattering presentation, but the discipline it calls for is universal. An institution that adopts the framework for its Gulf allocation will find it improves the assessment of every other private-markets commitment it holds, and an institution that already benchmarks its other allocations rigorously needs only to extend that rigour to the Gulf rather than to invent a new discipline. In this sense the paper's contribution is both specific, a fair way to judge a Gulf allocation, and general, a reminder that honest benchmarking is among the defining habits of a well-run institutional investor.

It bears restating, finally, that the framework is designed to be fair in both directions. Its purpose is not to deflate the Gulf, nor to inflate it, but to measure it honestly, which means crediting genuine, durable, net, risk-adjusted outperformance where it exists and withholding credit where the apparent outperformance is an artefact of a flattering benchmark, gross figures, or uncompensated risk and leverage. An institution applying the framework should expect it to validate a well-selected Gulf allocation and to expose a poorly selected one, and it should welcome both outcomes, because each is information it needs. The framework is, in this sense, neutral as to the Gulf and committed only to honesty, and that neutrality is precisely what makes its verdicts worth acting on. An allocator who wants the truth about its Gulf allocation, whatever that truth turns out to be, will find in this framework the means to obtain it; an allocator who wants only reassurance will find the framework uncomfortable, which is itself a sign that it is working.

In closing, the practical takeaway for an institution is a short discipline that this paper has built up in stages: define the benchmark and measures in advance and independently; assess net of fees, against the genuine opportunity cost, using the public-market equivalent; benchmark each sleeve against its own comparator; adjust for risk and the illiquidity premium; attribute the result to its sources; read the average against the dispersion and the institution's own managers against the quartiles; respect the J-curve by distinguishing monitoring from mature assessment; guard against peer-data biases; and subject the whole to independent assessment. An institution that follows this discipline will know, honestly and on a fair basis, whether its Gulf allocation has earned its place, and will make sound decisions about its future on that basis. An institution that does not will be at the mercy of whichever benchmark happens to be quoted, and its decisions will be correspondingly arbitrary. The difference between the two is not access to better data, which the region does not yet fully provide, but the discipline of asking the right questions of whatever data exist, which is available to any institution willing to adopt it.

The argument of this paper can be reduced to a single proposition that an institution should carry away: a Gulf allocation has earned its place only if it beats, net of fees and after adjusting for risk and illiquidity, the portfolio the institution would otherwise have held, and whether it has done so cannot be known from the headline figures that are most readily quoted. Everything else in the paper, the public-market equivalent, the sleeve-level comparison, the attribution, the dispersion, the treatment of the J-curve and of data biases, the independent assessment, is in service of answering that one question honestly. An institution that answers it honestly will allocate its capital where it genuinely earns its keep and withdraw it where it does not, which is the whole of

good capital allocation; an institution that accepts the flattering answer will misallocate in both directions. The discipline is demanding but not esoteric, and its reward is not merely a more accurate scorecard but better decisions, which is why honest benchmarking, though it appears to be about measuring the past, is in truth about investing well in the future. For an institution building a Gulf programme, adopting this discipline at the outset is the surest way to ensure that, whatever the region ultimately delivers, the institution will know it, and will act on the knowledge rather than on an impression.

7. CONCLUDING COMMENTS

This paper has examined how the private-market returns of a Gulf allocation should be benchmarked against the global portfolio an institution would otherwise hold, and has argued that the benchmarking choices, more than the underlying performance, frequently determine the verdict. The findings are consistent across the propositions. The verdict depends heavily on whether performance is measured gross or net and by which metric, with net measures and the public-market equivalent giving a more modest and more honest result than the headline internal rate of return (Proposition 1). Benchmarked by the public-market equivalent against the investor's actual portfolio, a well-selected Gulf allocation can show genuine outperformance, but smaller than gross-versus-cash comparisons imply (Proposition 2). A fair comparison must be net of fees, which consume a substantial and variable share of the gross return (Proposition 3). Part of any excess is compensation for illiquidity and risk rather than skill, which a fair benchmark must recognise (Proposition 4). And manager and vintage dispersion dominate the average, so the institution's own selection determines its experienced return (Proposition 5).

The implication for the institution is that benchmarking a Gulf allocation is a discipline to be designed in advance and applied honestly, not a number to be produced after the fact. The institution that defines its benchmark and measures before allocating, assesses net of fees against its genuine opportunity cost, adjusts for risk and illiquidity, and reads the average alongside the dispersion, will reach a verdict it can trust and act on. The institution that allows the benchmark to be chosen to suit the result, or that anchors on the flattering gross-versus-cash comparison, will mis-judge its allocation and may either abandon a sound one on a misleadingly modest verdict or add to a weak one on a misleadingly strong one. Honest benchmarking, in short, is not only how an institution measures the past but how it governs the future of the allocation.

The deeper message connects benchmarking to the broader case for a Gulf allocation. Companion analyses argue that the region is under-owned, that its currency risk is removed by the peg, and that its risk-adjusted profile is closer to developed than emerging markets. This paper supplies the discipline by which those claims, and the performance of any actual allocation, can be tested honestly. A reader sceptical of the broader case should welcome the framework here, because it provides the means to verify or refute the performance claims on a fair basis; and an institution persuaded by the broader case should adopt the framework, because it ensures that the allocation it makes is judged by a standard that neither flatters nor unfairly penalises it. Honest benchmarking is the friend of the genuine opportunity and the enemy only of the overstated one.

The study is subject to limitations that also define its contribution. It is a methodological and framework paper supported by stylised illustration rather than an empirical study of realised Gulf returns, because the region's data infrastructure does not yet support a robust empirical benchmarking; the figures show the mechanics, not the market's actual performance. The

benchmarking choices it recommends are well established in the private-markets literature but their application to the Gulf specifically awaits better data. These limitations point directly to the most valuable further research: the construction of robust, investable return series for Gulf private and public assets; the empirical estimation of public-market equivalents for Gulf allocations against global benchmarks; and the measurement of manager and vintage dispersion in the region. As that data matures, the framework set out here can be applied to it, and the honest verdict on the Gulf, which this paper has shown how to reach but not, on current data, how to settle, can be established. In the meantime, the contribution is the discipline itself: a way of asking, and eventually answering, whether a Gulf allocation has earned its place, that neither the flattery of gross-versus-cash comparison nor the scepticism of an unfair benchmark can distort.

DISCLOSURES

The author is an independent researcher and transaction advisor who advises institutional and family-office investors on matters including capital allocation, performance assessment and structuring related to the markets discussed in this paper, and may have professional or commercial interests in transactions involving the Gulf region. This paper is provided for information and educational purposes only; it does not constitute investment, tax or legal advice, an offer or solicitation, or a recommendation regarding any manager, security or strategy. The figures presented are modelled and illustrative of benchmarking mechanics, are not estimates of actual market returns and are not forecasts, and should be independently verified. Readers should obtain advice appropriate to their own circumstances and jurisdiction before acting.

References

[1] Ang, A. (2014). Asset Management: A Systematic Approach to Factor Investing. New York: Oxford University Press.

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

[3] Ilmanen, A. (2011). Expected Returns: An Investor's Guide to Harvesting Market Rewards. Chichester: Wiley.

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

[5] Phalippou, L. (2009). Beware of Venturing into Private Equity. Journal of Economic Perspectives, 23(1), 147-166.

[6] Phalippou, L., and Gottschalg, O. (2009). The Performance of Private Equity Funds. Review of Financial Studies, 22(4), 1747-1776.

[7] Sorensen, M., and Jagannathan, R. (2015). The Public Market Equivalent and Private Equity Performance. Financial Analysts Journal, 71(4), 43-50.

[8] Harris, R. S., Jenkinson, T., Kaplan, S. N., and Stucke, R. (2023). Has Persistence Persisted in Private Equity? Journal of Corporate Finance, 81.

[9] Cambridge Associates (2024). Understanding Private Investment Benchmarks and the Public Market Equivalent.

[10] International Monetary Fund (2024). Regional Economic Outlook: Middle East and Central Asia. Washington, DC: IMF.

About the Author

Chennakeshav Adya is an independent researcher with more than 25 years of industry experience as both an operator and an investment banker, spanning the United Arab Emirates, the United

Kingdom, India and beyond, over the course of which he has originated and advised on transactions across corporate finance, private capital, real estate and alternatives. He holds a Bachelor of Engineering from Visvesvaraya Technological University (VTU), an MBA from London Business School, and is pursuing a Master of Laws (LLM) at University College London (UCL). His research focuses on translating practitioner experience into rigorous, accessible frameworks for institutional allocators and capital seekers across the Gulf and the United Kingdom, with particular emphasis on real estate, private credit, alternatives and cross-border capital.

APPENDIX A. BASE-CASE ASSUMPTIONS

Table 2. Parameters Used in the Benchmarking Framework

Parameter	Indicative value
Global 60/40 benchmark (net)	7.6% p.a.
Gulf real estate (gross / net)	12.0% / 9.5%
Gulf private credit (gross / net)	15.5% / 12.5%
Gulf PE / VC (gross / net)	16.0% / 11.5%
Public-market equivalent (well-selected)	1.18 to 1.30
Top-to-bottom-quartile net IRR spread	~12 percentage points
Naive verdict (gross vs cash)	~8.5 pp outperformance
Fair verdict (net, PME)	~3.2 pp outperformance
Sceptical verdict (net, risk-adjusted)	~2.1 pp outperformance

Indicative figures illustrating benchmarking mechanics; not estimates of actual returns or forecasts.

APPENDIX B. GLOSSARY

The following terms are used in this paper. Definitions are provided for the general reader and are not intended as formal definitions.

Benchmark The comparator against which performance is judged; the choice of benchmark strongly influences the verdict.

Carry (carried interest) The manager's share of fund profits, typically above a hurdle; a component of the fee load that separates gross from net.

Cash-flow matching Comparing a private allocation to investing the same cash flows, at the same times, in a public index; the basis of the PME.

Distributions to paid-in (DPI) Cash actually returned to investors relative to cash invested; a realisation-based performance measure.

Dispersion The spread of returns across managers or vintages; wide and persistent in private markets.

Gross return Return before fees and carry; overstates what the investor keeps.

Illiquidity premium The additional expected return compensating investors for holding assets that cannot be readily sold; part of private-market excess return.

Internal rate of return (IRR) The annualised return implied by a series of cash flows; sensitive to timing and leverage, and the most-quoted but most manipulable measure.

Net return Return after all fees and carry; what the investor actually keeps.

Opportunity cost The return the investor forgoes by choosing one investment over its best alternative; the right basis for a benchmark.

Public-market equivalent (PME) A measure of whether a private allocation beat investing the same cash flows in a public index; above one means it did.

Quartile A quarter of a ranked distribution; top-quartile and bottom-quartile managers differ widely in private markets.

Risk adjustment Assessing return per unit of risk; for private markets, requires correcting for understated, smoothed volatility.

Smoothed valuation Infrequent, appraisal-based private-asset valuations that understate true volatility and flatter risk-adjusted comparisons.

Subscription line Short-term borrowing a fund uses to defer capital calls, which can inflate the reported internal rate of return.

Total value to paid-in (TVPI) Total value (distributed plus remaining) relative to capital invested; a multiple measure of total performance.

Vintage The year a fund begins investing; performance varies widely by vintage, so diversifying across vintages spreads timing risk.

Volatility Variation in returns over time; in private markets, reported volatility is understated by smoothed valuations.

Attribution The decomposition of a return into its sources: market (beta), leverage, timing, illiquidity premium and residual skill.

Beta The portion of return attributable to the market's movement rather than to skill.

Benchmarking policy A written, advance specification of the benchmark and measures by which an allocation will be judged.

J-curve The pattern of early negative and later positive returns in a private-markets programme, complicating interim benchmarking.

Opportunity-cost benchmark The return on the alternative the investor would otherwise have held; the fair basis for judging an allocation.

Sleeve-level benchmarking Assessing each component of an allocation against its own appropriate comparator rather than the whole against a single index.

Realised return Return based on actual cash distributed, as distinct from paper or marked valuations.

Vintage-matched peers Comparators drawn from funds of the same starting year, for a like-for-like comparison.

Double-counting Crediting an allocation with both a sleeve's return and the diversification benefit simultaneously, overstating its contribution.

Monitoring benchmarking Frequent, indicative performance comparison used to watch for problems, distinct from mature assessment.

Assessment benchmarking Performance comparison conducted at maturity to judge whether an allocation earned its place.

Benchmarking error Misjudging an allocation through an unfair benchmark, either too generous (the more common) or too harsh.

Comparator The specific index, portfolio or peer set against which performance is measured.

Add-hold-exit decision The choice, informed by benchmarking, to increase, maintain or reduce an allocation.

Independent performance assessment Evaluation of an allocation's performance by a function separate from the one that advocates for and manages it.

Re-up A commitment to a manager's successor fund, a decision that should rest on a fair performance verdict.

Durable outperformance Excess return attributable to skill or genuine access, which is more likely to persist than beta- or leverage-driven outperformance.