

What Makes a Good Investment Policy Statement

A Best Practices Framework for Public Pension Funds

Companion to: Assessment of Investment Policy Statement Quality Across U.S. Public Pension Funds

Analysis Date: October 2026

Rubric Version: 1.3

Executive Summary

This paper builds on a structured assessment of Investment Policy Statement (IPS) documentation across 121 U.S. public pension funds, scored by a multi-lens LLM judge against the rubric's anchors, each read in full. The empirical analysis revealed strong governance fundamentals (a category average of 4.4 on the five-point scale) but systemic weakness in ESG integration (2.2), the risk framework (3.1) and portfolio-design detail (3.4). The overall mean of 3.7 masks substantial variance: about two and a half points separate the strongest document from the weakest. Across the four quality lenses the judge applies to every domain, enforceability — whether a provision is specific enough to bind a decision — is the systematic weak point (3.0), a full point behind structural consistency and trailing presence and governance process.

Drawing on the patterns observed among top-performing institutions and the underlying rubric of 14 domains and 102 fact elements, this paper articulates a normative framework for what a good IPS should contain. We identify six hallmarks of excellence, describe the markers of quality within each domain, set out in detail what an IPS needs on spending, liquidity and private markets — where the sector is weakest — catalogue the most common pitfalls, and present a structured roadmap that institutions can use to assess and improve their own policy documentation. The paper closes with a single summary of the best practices it recommends.

The central argument is straightforward: a good IPS is not a marketing document, an aspirational statement, or a description of past practice. It is an operational governance instrument whose value lies in its specificity, enforceability, and internal coherence. The institutions that score highest are those whose policy statements function as decision rules rather than narrative summaries.

1. Background and Motivation

The Investment Policy Statement is the foundational governance document for an institutional investor. It defines the entity, articulates objectives, sets risk boundaries, prescribes portfolio construction, allocates decision authority, and establishes the cadence of oversight. In the words of fiduciary scholarship, the IPS is both compass and constitution — directional in intent, binding in form.

For public pension funds, the stakes are particularly high. These institutions manage assets on behalf of millions of beneficiaries, operate under statutory fiduciary obligations, and answer to multiple stakeholders including legislatures, boards, plan members, and taxpayers. The IPS is the principal artifact by which governance quality is documented and defended.

Despite the importance of the document, practice varies widely. Our empirical assessment of 121 U.S. public pension funds found a range of about two and a half points in overall quality scores on a 5-point scale — meaning the gap between the strongest and weakest IPS documents in the sample is greater than the gap between an "Adequate" and a "Best Practice" rating. This variance is the motivating observation behind this paper: if some public pension funds produce IPS documents that meet best practice across the board, the question is what specifically distinguishes those documents — and what can the rest of the sector learn from them?

This paper answers that question in three steps. Part I summarizes what we found in the empirical assessment. Part II articulates the six hallmarks of a good IPS that emerge from the data. Part III provides a domain-by-domain framework drawing on the rubric used in the assessment, and closes with two deeper treatments: of spending and liquidity, the two weakest core domains, and of what an IPS needs to govern private markets. Part IV catalogues common pitfalls, among them the undated and unversioned document. Part V sets out a practical roadmap, and the paper ends with a single summary of the best practices it recommends.

2. Part I: Summary of the Empirical Assessment

2.1 What Was Measured

The assessment evaluated IPS documents using a 14-domain rubric (v1.3) covering 102 individual fact elements. The rubric was originally derived from the CFA Institute's guidance on Investment Policy Statements — specifically the elements and structural recommendations set out in *Elements of an Investment Policy Statement for Institutional Investors* (CFA Institute, 2010) — and extended to incorporate domains and fact elements relevant to contemporary public pension practice (notably liquidity stress testing, illiquid-asset treatment, and ESG / responsible investment integration). The 14 domains aggregate into six high-level categories:

Category	Domains
Governance & Oversight	Governance & Decision-Making; Monitoring, Reporting & Review
Objectives Clarity	Scope, Purpose & Coverage; Investment Objectives; Investment Philosophy
Risk Framework	Risk Tolerance; Time Horizon & Constraints; Liquidity & Cash Flow

Portfolio Design	Strategic Asset Allocation; Rebalancing Policy; Spending / Distribution
Implementation and Oversight	Implementation & Manager Oversight; Performance Measurement & Benchmarks
Responsible Investing / ESG	Responsible Investment / ESG

Each domain was scored on a 1-5 scale anchored at "Deficient" (missing or purely aspirational) at the low end and "Best Practice" (fully specified, enforceable, peer-aligned) at the high end. Within each domain the judge scores four quality lenses — presence/clarity, quantification/enforceability, consistency/structure, and governance/process — and averages the applicable lenses into the domain score. The headline overall score is the mean of a plan's core domains - 12 of the rubric's 15 for a public pension plan. Two domains are optional (investment philosophy and responsible investment / ESG): both are scored and shown throughout, but neither moves the headline nor raises a critical gap. Tax considerations is conditional on an institution type that does not apply here and is not scored at all. A critical gap is a core domain scoring 2 or lower.

2.2 How the Scores Were Produced: The LLM Judge

The scores in this paper were produced not by manual review but by a large language model (Anthropic's Claude) acting as a structured **judge** against the rubric. Because a reader's confidence in the findings depends on understanding how the scores were generated, we describe the procedure here.

The assessment runs in two deliberately separated stages. In the first, **extraction** stage, a Claude model reads each IPS and pulls out the facts relevant to each domain, citing a verbatim quote and page for each — it does not score at this stage. A faster, cheaper model (Claude Haiku) is used here because the task is retrieval. The extracted facts are frozen, together with a hash of the source text and the rubric, so the second stage is fully reproducible.

In the second, **judging** stage, a stronger Claude model (Claude Sonnet) scores each domain against its 1/3/5 anchors. Crucially, it does so along **four quality lenses** — presence/clarity, quantification/enforceability, consistency/structure, and governance/process — and the domain score is the average of the lenses that apply. This decomposition is what allows the assessment to distinguish a policy that is merely *present* from one that is *enforceable* — the distinction on which this paper's central argument turns.

Four conventions make the scores trustworthy. The judge runs deterministically (temperature zero), so the same document yields the same score. It returns its verdict through a forced structured-output schema with a fixed vocabulary of reason codes, so every score is well-formed and traceable to a stated reason. Every quote the judge cites is checked against the source text. And each scorecard records the exact model, rubric version, and prompt version, so the findings can be frozen to a precise configuration and regenerated on demand.

Finally, because the scores are model-produced, their credibility rests on **measured agreement with a human expert**, and that measurement has not yet been made: the scores in this paper are the judge's alone. The validation is designed and ready. A stratified sample of plans has been drawn, to be rated blind, domain by domain, against the full IPS text — with the judge's own score hidden until each human rating is locked in — and agreement will be reported as a quadratic-weighted Cohen's kappa, alongside per-domain error, directional bias, and the rater's own test-retest ceiling. That calibration is a measurement exercise, not an automated tuning loop: it will show where the judge agrees with expert judgement and where it does not. Its limitations — a

single expert rater, and evidence-checking that is reported rather than enforced — are stated plainly in the methodology companion to this paper.

2.3 Headline Findings

Metric	Value
Plans scored	121
Mean overall score	3.66 / 5.0
Median	3.71
Range	1.99 - 4.50
Standard deviation	0.42
Plans scoring 4.0 or higher	26 (21%)
Plans scoring 3.5 - 3.99	57 (47%)
Plans scoring below 3.5	38 (31%)
Plans with at least one critical gap	64 (53%)

The sector performs well at the governance fundamentals — defining who decides what, when reporting happens, and how the institution articulates its purpose and objectives. It performs less well at the operational and forward-looking dimensions: a quantified risk framework, liquidity stress testing, spending and rebalancing mechanics, and ESG integration. The judge's anchors are demanding: a fifth of the sample reaches the 4.0 "strong" tier, while a third scores below 3.5. The binding constraint is enforceability — most documents describe intentions clearly but stop short of the specific, testable provisions that would turn an intention into a decision rule.

2.4 Category Performance

Category	Mean	Median	Std Dev
Governance & Oversight	4.36	4.50	0.53
Implementation and Oversight	3.92	4.00	0.58
Objectives Clarity	3.83	3.83	0.37
Portfolio Design	3.37	3.42	0.50
Risk Framework	3.07	3.08	0.51
Responsible Investing / ESG	2.22	2.00	0.80

The same scores, decomposed by the four quality lenses the judge applies within every domain, locate the weakness precisely:

Quality lens	Mean
Consistency / Structure	4.04
Presence / Clarity	3.66
Governance / Process	3.51
Quantification / Enforceability	2.99

Documents are, on average, internally consistent and reasonably clear about what they cover. Where they fall down is enforceability — the gap between describing a policy and stating it in terms specific enough to bind a decision. Enforceability is the lowest-scoring lens in every weak category and the single most reliable correlate of a low domain score. This is the central empirical finding of the assessment, and the organising theme of the normative framework that follows.

Two further observations are worth emphasising. First, the absolute weakness in ESG documentation — 51% of the sample scored 2.0 or below — remains the widest single-domain quality gap in the sector. Second, the judge now finds a small number of ESG leaders: four documents reach the "strong" anchor, treating responsible investment as a policy in its own right, with objectives, a stewardship process and reporting. ESG retains the widest spread of any category (standard deviation 0.8), running from absent to fully specified; the models exist, but they are rare.

2.5 Where Critical Gaps Cluster

Across the 64 plans with critical gaps, the gaps concentrate in a small number of core domains (the two optional domains - investment philosophy and responsible investment / ESG - are scored but, by construction, never raise a critical gap, which is why the lowest-scoring domain in the sector does not appear below):

Domain	Institutions with Critical Gap
Spending / Distribution Policy	57
Liquidity & Cash Flow Management	22
Rebalancing Policy	6
Time Horizon & Constraints	5
Performance Measurement & Benchmarks	3
Monitoring, Reporting & Review	3
Risk Tolerance & Risk Philosophy	1
Strategic Asset Allocation	1

The pattern is consistent: the gaps are not random. They concentrate in exactly the domains that require quantitative specificity and operational detail — spending and liquidity mechanics, rebalancing triggers, and the horizon and risk-limit provisions that govern them. Two core domains dominate: spending / distribution policy (57 of the 64 affected plans) and liquidity & cash-flow management (22). Both are domains where

public-pension IPS documents tend to point at external plan documents or actuarial valuations rather than state the fund's own operational commitments — a gap the judge penalises precisely because the commitment is left unenforceable. These are the dimensions that most distinguish a governance document from a description of the status quo.

2.6 What the Top Performers Share

The five strongest documents score an average of 4.9 on Governance & Oversight — but only 2.3 on Responsible Investing / ESG, the sector's weakest category. Excellence, as the judge measures it, is not the joint product of a culture that takes every dimension seriously. The leaders are the institutions whose documents are the most quantified and enforceable — explicit asset-class targets and ranges, numeric risk limits with named monitoring bodies, specific rebalancing triggers, defined review cadences — and several of them are nearly silent on ESG. What the top performers share is not breadth of coverage but depth of enforceability: they write provisions that bind decisions, and they do so consistently across the core governance and portfolio domains, whether or not they have yet engaged with ESG.

The implication for the rest of the sector is encouraging in one respect and sobering in another. Encouraging, because the quality that distinguishes the leaders — enforceable specificity — is a writing discipline, not a resource endowment, and can be adopted by any fund willing to do the work. Sobering, because best practice across *every* domain does not currently exist in a single document: even the strongest funds have an unaddressed frontier, most commonly ESG. The remainder of this paper develops the discipline of enforceability into an explicit normative framework.

3. Part II: The Foundations of a Good IPS

3.1 First Principles: What an IPS Is For

Before discussing what makes a good IPS, it is worth being explicit about what an IPS is for. The document serves five distinct functions, and a good IPS serves all five:

1. **A decision rule.** When a question arises — should we rebalance? what is the benchmark for this manager? who can approve this commitment? — the IPS should produce an answer without further interpretation.
2. **A constraint on discretion.** The IPS binds future boards and staff to commitments made by their predecessors. It protects beneficiaries from short-term pressure, novel theories, and individual judgment errors.
3. **A documentation of fiduciary care.** The IPS is the principal evidentiary

record by which a board demonstrates that it has discharged its duty of loyalty and prudence. Its specificity is its defensibility.

4. An onboarding instrument. New board members, new staff, and new external advisors should be able to read the IPS and understand how decisions get made at this institution.

5. A signal to the outside world. Beneficiaries, oversight bodies, the public, and counterparties learn from the IPS what the institution stands for and how it operates.

These functions create internal tension. A document that is comprehensive enough to constrain discretion is at risk of becoming too detailed to read. A document that is written for beneficiaries is at risk of being too high-level to operate. The art of writing a good IPS is the management of this tension.

3.2 The Six Hallmarks of Excellence

From the empirical patterns, six properties consistently distinguish the strongest IPS documents from the rest. These are not domain-specific — they apply across every section of the document.

Hallmark 1: Specificity

A good IPS specifies. It states ranges, thresholds, frequencies, percentages, and dates. It names committees and titles rather than vague constructs ("the investment team"). It defines terms that have technical meaning ("risk", "liquidity", "long-term") rather than relying on shared assumptions.

The diagnostic question is: could a reasonable reader, with no other context, extract a decision from this paragraph? If the answer requires conversation with the author, the language needs to be tightened.

Hallmark 2: Enforceability

A good IPS is enforceable. Each statement of policy implies a question of the form "what happens if this is not done?" A good IPS answers that question by naming the responsible party, the escalation path, and the consequences. Aspirational language — "we strive to", "we seek to", "we consider" — is appropriate in the philosophy section but corrosive when used for operational provisions.

Consider the difference between "the portfolio shall be rebalanced periodically" and "the staff shall rebalance whenever any asset class exceeds its target range by more than 200 basis points; the CIO shall report any rebalancing action to the Investment Committee at the next regular meeting." The first sentence describes a hope. The second sentence describes a process.

Hallmark 3: Internal Coherence

A good IPS is internally coherent. The investment objectives, risk tolerance, asset allocation, liquidity policy, and performance measurement framework all imply one another, and a good IPS makes those implications

explicit.

If the objective is a real return of 5%, the asset allocation should produce an expected return consistent with that target. If the risk tolerance permits a 20% drawdown, the asset allocation should be one whose stress-tested loss remains within that band. If the liquidity policy requires X months of cash on hand, the asset allocation should not be over-committed to illiquids inconsistent with that buffer.

These linkages are often the first casualty of incremental amendment. Each piece of the document gets updated by a different working group at a different time, and the cross-references stop holding together. A good IPS audits these linkages on every revision.

Hallmark 4: Completeness

A good IPS addresses every dimension of investment governance that the institution is responsible for. The 14 domains in the rubric used for the empirical assessment provide one such enumeration. Any IPS that omits one of these domains is implicitly delegating that decision to whatever person or process happens to be in the room when the question arises — which is the opposite of governance.

Completeness does not require length. A short paragraph that addresses liquidity stress testing is better than a long passage that omits it.

Hallmark 5: Reviewability

A good IPS specifies when and how it will be reviewed. It dates each version, identifies the body responsible for amendment, sets a cadence for routine review (typically annual or biennial), and creates a mechanism for ad-hoc amendment when circumstances change. The reviewability provisions are the mechanism by which the document remains relevant as the institution and the environment evolve.

Hallmark 6: Operational Realism

A good IPS reflects how the institution actually works. Provisions that nobody follows are worse than no provisions at all, because they create the appearance of governance without the substance and expose the institution to liability when the documented process is not the process actually used.

The institutions whose IPS documents we evaluated as strongest were those whose provisions were ambitious but realistic — extensions of practice rather than inventions on paper. Where the document and the practice diverge, one of them needs to change.

4. Part III: A Domain-by-Domain Framework

This section walks through each of the six high-level categories and the constituent domains, describing what excellence looks like at the domain level. For each domain we identify the markers of a 5/5 score (best

practice), the markers of a 3/5 score (adequate but with gaps), and the most common reasons institutions fall short.

4.1 Governance & Oversight

This is the category in which the sector performs best (4.4 average), but it is also the foundation of every other category. Weakness here propagates.

Governance & Decision-Making. Best practice requires explicit identification of all decision-making bodies (board, investment committee, staff, consultants, custodians), explicit allocation of authority among them, a stated standard of care, named conflict-of-interest provisions, an escalation procedure for matters beyond delegated authority, and a clear amendment process. Common gaps include generic references to "the board" without distinguishing the board from its committees, and absence of a stated standard of care (prudent person, prudent investor, or statutory variant).

Monitoring, Reporting & Review. Best practice specifies the frequency of performance reporting, the contents of each report, the review cadence for the IPS itself, and the trigger conditions for ad-hoc review. A document that says "the IPS will be reviewed periodically" without a cadence has not actually created a review obligation.

4.2 Objectives Clarity

The third-strongest category (3.8). The standard expectation here is hierarchical and measurable objectives.

Scope, Purpose & Coverage. A good IPS opens by defining the entity, enumerating the assets covered and excluded, stating the purpose of the document, naming the effective date, recording the adoption authority, and specifying the amendment authority. A surprising number of IPS documents leave these basic items implicit.

Investment Objectives. Best practice requires explicit primary objectives, a stated return metric (real return, nominal return, return relative to liability discount rate), an explicit time horizon, and a stated hierarchy when objectives conflict (e.g., "preservation of capital takes priority over return maximization in the bottom quartile of return scenarios"). Vague formulations such as "long-term growth consistent with prudent risk" fail the specificity test because they cannot be falsified.

Investment Philosophy. Best practice articulates the institution's view of market efficiency, its disposition toward active versus passive management, and its theory of how returns are generated — and then explicitly links those views to the asset allocation. A philosophy section that does not constrain subsequent decisions is decoration.

4.3 Risk Framework

The second-weakest category (3.1), and the clearest illustration of the enforceability problem. The weaknesses here are almost always failures of quantification.

Risk Tolerance & Risk Philosophy. Best practice defines risk in operational terms (volatility, drawdown, funded-status deterioration, real-return shortfall), states a quantitative limit, identifies the metric used to monitor the limit, references stress testing, names the body responsible for monitoring, and specifies the consequences of breaching the limit. The most common failure is a qualitative statement of risk appetite ("the fund has a long-term horizon and can tolerate short-term volatility") with no enforceable threshold attached.

Time Horizon & Constraints. Best practice states the investment horizon explicitly, enumerates liquidity constraints, identifies relevant legal and regulatory constraints, defines policy on leverage and derivatives, and addresses foreign currency exposure. For public pensions, statutory constraints deserve particular attention because they often impose binding limits that override stated risk preferences.

Liquidity & Cash Flow Management. Best practice specifies liquidity objectives in operational terms, identifies expected cash flows (benefit payments, contributions), defines a liquidity buffer, references a stress framework, sets a private-capital commitment policy that respects the liquidity profile, and names the party responsible for liquidity oversight. This is the domain most often forgotten when private allocations are increased, with predictable consequences.

4.4 Portfolio Design

A mid-tier category (3.4). The weaknesses are typically in rebalancing and spending policy — the latter the single most common critical gap in the sector.

Strategic Asset Allocation. Best practice defines each asset class, states target weights, defines allowable ranges, articulates the rationale for the allocation (linked back to objectives and risk tolerance), addresses the treatment of illiquid assets, defines an illiquidity budget, sets leverage limits, states derivative policy, and clarifies tactical-allocation authority if any. The minimum acceptable allocation table includes target, lower bound, and upper bound; benchmark; and rationale.

Rebalancing Policy. Best practice defines triggers (drift-based or calendar-based), identifies the authority to rebalance, specifies cash-flow treatment, sets a review frequency, addresses the treatment of illiquid drift (which cannot be rebalanced as easily), and identifies override authority. The single most common gap is silence on what happens when illiquid allocations drift outside target — a real problem in stressed environments when equity mark-downs leave private allocations at over-weight without any way to sell.

Spending / Distribution Policy. For public pensions, the distribution mechanics typically sit in separate plan documents or actuarial valuations rather than the IPS itself. Best practice in the IPS is to acknowledge the benefit obligation, address its liquidity implications, reference the external documents that govern distribution mechanics, and integrate distribution considerations with the investment strategy. A clean cross-reference is acceptable; silence is not.

4.5 Implementation and Oversight

The second-strongest category (3.9). The gap between strategy and execution remains a consistent weakness across the sample.

Implementation & Manager Oversight. Best practice describes the implementation model (internal, external, hybrid), states manager selection criteria, defines the monitoring process, specifies termination guidelines, addresses proxy voting policy, addresses securities lending policy, defines fee monitoring, references operational due diligence, and sets concentration limits. The diagnostic question is: if a new manager needs to be hired or an underperforming manager needs to be terminated, does the IPS produce the answer? In a majority of sampled documents, it does not.

Performance Measurement & Benchmarks. This is the most fact-element-dense domain in the rubric (18 elements), which makes complete coverage demanding even where the essentials are present. Best practice defines a total-fund benchmark, asset-class benchmarks for each asset class, the construction method for custom benchmarks, an approach for private-market benchmarking (which cannot use the same machinery as public markets), performance objectives linked back to the investment objectives, risk-adjusted metrics, calculation methodology, the official book of record, reconciliation processes, valuation policy for private assets, reporting frequency, peer-group comparators, an attribution framework, and (for pensions) a liability-relative benchmark. The most revealing question to ask of an IPS is whether the asset-class benchmark section actually lists a specific benchmark for each asset class. In a surprising number of cases, it does not.

4.6 Responsible Investing / ESG

The weakest category by a wide margin (2.2) and the one undergoing the most rapid evolution.

Responsible Investment / ESG. Best practice in the IPS itself (not in a separate document) requires a dedicated section, an articulation of ESG objectives, an explicit acknowledgment of ESG-related financial risks, a stewardship or engagement approach, a description of the ESG program or governance structure, and a proxy-voting approach. Stronger documents address climate risk explicitly and outline an ESG performance or reporting framework.

The argument for documenting ESG in the IPS, even where the institution's substantive position is limited, is the same as the argument for documenting any other consideration: undocumented positions are not governance, and absence of documentation creates exposure both ways. An institution that does not intend to integrate ESG should say so explicitly; an institution that does should describe how. Silence is the worst option.

Four documents in the sample reach the "strong" anchor on this domain, and none the best-practice level. Below them the upper tail consists of adequate sections that acknowledge ESG-related financial risk and outline a stewardship approach without rising to a fully specified, enforceable ESG framework. The four are the closest available reference points, not finished templates; the frontier of best practice in this domain remains open.

4.7 Spending and Liquidity: Closing the Two Deepest Gaps

Spending / distribution policy and liquidity & cash-flow management are the two weakest core domains in the sector, and between them they account for most of the critical gaps in Section 2.5. They are also the gaps that revision does not close: the companion assessment finds that when plans rewrite their IPS, the scores that move are risk tolerance, asset allocation and benchmarks, while spending and liquidity barely change. Plans

revise what the document already covers. These two domains have to be put on the agenda of a revision deliberately.

The reason so many documents fall short is understandable. A public pension's benefits are set by statute and its contributions by an actuarial valuation, so drafters treat distribution as somebody else's document. But the investment consequences of those decisions belong in the IPS, because the fund has to meet the payments from the portfolio the IPS governs.

What the spending / distribution section should say. The IPS cannot set benefits or contributions, but it can and should:

1. **State the cash-flow profile.** The expected net cash flow — benefit payments

and expenses less contributions — as a share of assets, over a stated horizon, with the party that projects it (the actuary or staff) and how often the projection is refreshed.

2. **Link the policy to the actuarial framework.** Name the assumed rate of

return, who sets it, and how the expected return of the policy portfolio is compared with it — and what the board does when the two diverge.

3. **Name the contribution risk.** Say how a shortfall in contributions against

the actuarially determined amount is monitored and reported, and how the investment policy responds to a sustained shortfall.

4. **Say where the cash comes from.** The order in which benefit payments are

funded: contributions and income first, then rebalancing sales, with the asset classes that serve as sources named.

5. **Cross-reference cleanly.** Cite the statute or plan document that governs

benefits and the valuation that sets contributions, by name and date.

6. **Report against it.** An annual comparison of actual against projected cash

flows, presented to the board.

An illustrative provision (not drawn from any one document): *"Net benefit outflows are projected by the Fund's actuary at approximately 3% of assets a year over the next ten years. Staff shall maintain liquid assets sufficient to meet twenty-four months of projected net outflows and shall report actual against projected cash flows to the Board annually."*

What the liquidity section should say. Liquidity governance is the frontier of the sector: in the companion assessment's scan of the full texts, only a handful of documents — a minority even of the strongest — set liquidity tiers or a liquidity stress test. Best practice requires:

1. **An operational objective.** For example: meet benefit payments and capital

calls without selling assets below fair value, and keep the capacity to rebalance in a falling market.

2. **Liquidity tiers.** Every asset class assigned to a tier by time to cash —

same day, within a week, within a quarter, over years — with a minimum share of the fund held in the first two tiers.

3. **A coverage ratio.** Liquid assets divided by twelve to twenty-four months of

projected net outflows plus the unfunded commitments expected to be called over the same period, with a floor the board sets.

4. **A stress test that sizes the buffer.** A defined scenario — a sharp equity

decline, delayed contributions, accelerated capital calls and slowed distributions together — run at least annually, with the buffer sized to survive it.

5. **The other claims on liquidity.** Collateral and margin for derivatives and

overlays, recall of securities on loan, and any credit lines, each with a limit.

6. **Responsibility and escalation.** Who monitors the tiers and the ratio, how

often, which level triggers a report to the board, and what action follows.

4.8 Private Markets: What the IPS Needs

Public pension funds now hold a substantial share of their assets in private equity, private credit, real estate and infrastructure, and the median target allocation to private equity alone is about a tenth of the fund. Most IPS documents still govern those assets with the machinery built for public markets: a target, a range and a quarterly benchmark. That machinery fails in three predictable ways. Capital is committed long before it is invested and returned long after, so the exposure cannot be set by a single decision. The assets cannot be sold to rebalance, so a fall in public markets pushes the private allocation over its range with no way back. And the reported values lag the market, so short-horizon performance comparisons are meaningless. In the companion assessment's text scan, about a quarter of documents mention a commitment pacing plan, about one in seven addresses unfunded commitments, and fewer than one in four says what happens to illiquid drift.

An IPS that governs private markets properly sets out the following.

What the IPS should say	
diversification, inflation sensitivity); a target and range for each private class; an overall illiquidity budget covering both invested assets and unfunded commitments	
d by the board, built from the target, expected calls and distributions and fund growth; limits on commitments per year, per manager and per vintage	
unded commitments counted in liquidity planning and in the coverage ratio; a cap as a share of the fund or of liquid assets	
that an overweight caused by falling public markets does not force sales; how the allocation is brought back (slower pacing, secondary sales) and over what period	Silen
on basis (manager marks, frequency, audit), the reporting lag, and how the lag is treated in total-fund reporting and rebalancing	
s — a public-market equivalent, a lagged public index plus a premium, or a peer universe — and an evaluation horizon of five to ten years; IRR and multiple reporting	
s; fee and carried-interest transparency (for example, the ILPA reporting template); alignment of interest; key-person and governance rights	
and per fund; policy on co-investments, separate accounts and secondaries; the commitment size above which board approval is needed	

Limits on, or reporting of, subscription lines, NAV facilities and fund-level leverage	
h flows against the pacing plan, exposure by vintage, strategy and manager, and the expected shape of returns over a fund's life	
Who may authorise the sale of fund interests on the secondary market, and on what terms	

Two principles hold the table together. First, private-market exposure is a flow to be managed, not a stock to be set: the pacing plan, not the target, is the operative decision, and the IPS should say so. Second, the private-market sections must be written together with the liquidity section of Section 4.7. The coverage ratio and stress test are where unfunded commitments, lagged valuations and benefit outflows meet, and a document that treats them separately will find out how they interact only in a crisis.

5. Part IV: Common Pitfalls

Across the assessment, certain weaknesses recurred with such frequency that they merit explicit treatment. An institution writing or revising an IPS should treat the following as a pre-flight checklist of items to avoid.

5.1 Aspirational Language in Operational Sections

"The Fund seeks to achieve long-term growth consistent with the preservation of capital" is appropriate as the opening sentence of an investment philosophy. It is not appropriate as a risk tolerance statement, an asset allocation rationale, or a rebalancing rule. Operational sections require operational language: thresholds, frequencies, triggers, and authorities.

5.2 Missing Quantification

The single most common gap is the absence of numbers in places where numbers belong. Asset allocation tables without ranges. Risk tolerance statements without volatility or drawdown limits. Liquidity policies without buffer percentages. Rebalancing rules without trigger bands. Performance measurement sections without reporting frequencies. In each case the fix is the same: add the number.

5.3 Undefined Roles

"The investment team" is not a role. "The CIO and the Investment Committee, acting under delegation from the Board" is a role. The diagnostic question is whether a reader unfamiliar with the institution could identify the named person or body and verify that the named party has the authority described.

5.4 No Treatment of Illiquids

Public pension funds have increased private-market allocations substantially over the past two decades. Many IPS documents have not kept pace. The governance of illiquids requires explicit attention to commitment

pacing, liquidity stress testing, rebalancing in the presence of un-saleable allocations, valuation policy, and benchmarking. An IPS that treats private markets as if they were liquid public markets is misaligned with the actual portfolio.

5.5 Benchmark Silence

A surprising number of IPS documents list asset classes without listing benchmarks for them. A surprising number list a total-fund benchmark without describing how it is constructed. A surprising number describe performance reporting without specifying calculation methodology. The benchmark section is the location where investment objectives are operationalized; silence here is silence on whether the institution is meeting its goals.

5.6 ESG Silence

By far the most common critical gap in the sample. Even institutions that have decided not to integrate ESG into investment decisions should document that decision and the reasoning. Silence is interpreted by readers — including beneficiaries and oversight bodies — as either omission or evasion. Neither is helpful.

5.7 Stale Review Provisions

An IPS that has not been reviewed in five years has likely fallen out of alignment with the portfolio it is supposed to govern. A review cadence is itself a substantive provision, and the absence of a cadence is the absence of an institutional commitment to keep the document current.

5.8 Internal Inconsistency

Documents accrue amendments over time. Each amendment may be reasonable on its own terms, but the cumulative effect is often a document whose risk tolerance, asset allocation, liquidity policy, and performance measurement framework no longer fit together. A periodic coherence audit — does the asset allocation still produce the targeted return at the stated risk? does the liquidity policy still cover the private-capital commitments? — is the only protection against this drift.

5.9 Undated and Unversioned Documents

A surprising number of IPS documents do not say when they took effect. In the companion assessment's scan of the full texts, only about a third state an effective date, and fewer than one in six carries a revision history. Many give the date the policy was first adopted — sometimes decades ago — on the cover, with later amendments recorded, if at all, in a list at the back or not at all. Sponsors also commonly post a new version at a new web address while leaving the old file in place, unlabelled.

The cost is not cosmetic. A board member, a beneficiary or an auditor cannot tell which version governs a decision taken last year. An outside monitor — including the one that produced the companion assessment — cannot tell that a policy has changed, or reads the original adoption date as the date of the current text. And the

document's own review cadence cannot be checked against anything.

The remedy is simple and costs nothing:

1. **Date the document where a reader looks first.** The effective date and the adopting body on the cover, and the version and effective date in the footer of every page.
2. **Keep a revision history.** A table at the back recording, for each version, the date, the body that adopted it, the sections changed and the reason.
3. **Say what it supersedes.** "This policy supersedes the version adopted [date]" removes any doubt about which text governs.
4. **State the next review date,** not only the review cadence.
5. **Date the schedules.** Where the asset allocation, benchmarks or manager guidelines sit in exhibits, each exhibit carries its own effective date and the IPS refers to it by date.
6. **Publish one current version.** Post the current policy at a stable address, and keep superseded versions in a clearly labelled archive rather than beside it.

6. Part V: A Roadmap for Improvement

The empirical assessment placed each institution into one of three tiers. The appropriate next step depends on where an institution sits.

6.1 For Institutions Scoring Below 3.5 (Lower Tier)

These institutions face systemic weakness and benefit from comprehensive revision rather than incremental amendment. Recommended sequence:

1. **Baseline against the rubric.** Score the current IPS against the 14 domains and 102 fact elements to identify every gap.
2. **Address mandatory domains first.** Investment Objectives, Risk Tolerance, Strategic Asset Allocation, Performance Measurement, and Governance are the highest-priority targets.
3. **Adopt a peer template.** The top-performing institutions identified in the empirical assessment (one plan, one plan, one plan, one plan, one plan) have published IPS documents whose governance and portfolio sections can serve as structural templates — though these documents lead on enforceability, not on ESG, which must be sourced separately.

4. **Commission a comprehensive revision.** Targeted patches to a weak document typically produce internal inconsistencies. A full rewrite is often faster and yields a better result.
5. **Establish a regular review cadence.** Annual review with substantive amendment every three years is a defensible default.

6.2 For Institutions Scoring 3.5 to 4.0 (Middle Tier)

These institutions have a workable IPS but identifiable weak spots. Recommended sequence:

1. **Identify the two weakest domains.** Concentrate improvement effort where the marginal return is highest.
2. **Address ESG explicitly.** Given the sector-wide weakness and the regulatory and stakeholder trajectory, an explicit ESG section is the single highest-leverage addition for most institutions in this tier.
3. **Quantify the risk framework.** Replace qualitative risk language with specific volatility, drawdown, or funded-status thresholds and identify the monitoring body.
4. **Add operational detail to implementation.** Manager selection criteria, monitoring process, and termination guidelines should each be specified.
5. **Audit internal coherence.** Verify that objectives, allocation, risk, and liquidity provisions still imply one another.

6.3 For Institutions Scoring 4.0 and Above (Top Tier)

These institutions face a different problem: how to maintain excellence as the environment evolves. Recommended sequence:

1. **Establish a continuous improvement loop.** Subscribe to peer-institution amendments and incorporate proven innovations.
2. **Lead on emerging dimensions.** Climate risk integration, stewardship metrics, and AI / data-driven manager evaluation are areas where leading institutions are extending the frontier.
3. **Share publicly.** Top-tier IPS documents are public goods for the sector. Publication, peer presentation, and explicit invitation of replication raise the floor for everyone.
4. **Test enforceability under stress.** Tabletop exercises in which the

document is consulted under simulated stress scenarios (liquidity crisis, major drawdown, manager fraud) reveal which provisions are operational and which are decorative.

6.4 The Diagnostic Checklist

A short list of questions that an institution can ask of its own IPS as a quick health check:

- Does every asset class have an explicit target, range, and benchmark?
- Is the risk tolerance stated as a number, with a named monitoring body and a stated consequence for breach?
- Does the document specify rebalancing triggers, authority, and treatment of illiquids?
- Does the document name a specific party responsible for each major decision domain?
- Is there a dedicated ESG section, even if brief?
- Is the review cadence stated, with a date for the next scheduled review?
- Does the cover state the effective date and version, and is there a revision history?
- Is there a liquidity coverage ratio and a stress test that sizes the buffer?
- Does the document set a commitment pacing plan, count unfunded commitments, and say what happens when private assets drift over their range?
- Are projected benefit cash flows stated, and compared with actual flows each year?
- Does the document address proxy voting, securities lending, and fee monitoring?
- Are the investment objectives, asset allocation, and risk tolerance consistent with one another?
- Can a new board member read the document and understand how decisions get made?
- Would a reasonable reader, with no other context, be able to apply the document to a real decision?

A "no" answer to any of these questions identifies a high-leverage area for revision.

7. Conclusion

The empirical assessment that accompanies this paper found that U.S. public pension funds produce IPS documents of widely varying quality, with average scores well above the floor but well below the achievable ceiling. The gap is not random. It concentrates in the operational, quantitative, and forward-looking dimensions of investment governance — exactly the dimensions that most distinguish a governance document from a narrative summary.

A good IPS, on the framework set out in this paper, is specific, enforceable, internally coherent, complete, reviewable, and operationally realistic. It addresses every dimension of investment governance for which the institution is responsible. It produces decisions rather than describing intentions. It binds future actors to commitments made in the present. It documents the exercise of fiduciary care. And it remains accurate to the institution it governs because it is reviewed, amended, and audited for coherence on a defined cadence.

The institutions that score highest on the empirical assessment are not those with the longest or most sophisticated documents. They are those whose documents have been written, and re-written, with the question "what does this provision actually require of us?" in front of every paragraph. The discipline of writing for enforceability is the discipline of governing for accountability. The two are the same discipline.

For the sector as a whole, the path forward is clear. The top performers have already demonstrated what enforceable, decision-binding policy looks like across the core governance and portfolio domains. No single document yet leads on every dimension — ESG in particular remains an open frontier even for the strongest funds — but the writing discipline that produces excellence is visible, and the frameworks, templates, and substantive provisions for the core domains exist. The question is whether institutions will adopt them. The case for doing so — fiduciary, regulatory, operational, and reputational — has only strengthened over time.

The Investment Policy Statement is the document by which a pension fund tells itself what kind of institution it intends to be. The institutions that take that document seriously will be better positioned, in every dimension that matters, than those that do not.

8. Best Practices in Summary

The practices this paper recommends, gathered in one place. Each is a provision a board can adopt; together they describe a policy that produces decisions rather than describing intentions.

The document

- Every operational provision states a number, a responsible party and the

consequence of a breach; aspirational language is confined to the philosophy.

- The effective date, adopting body and version appear on the cover and in the footer; a revision history and a "supersedes" statement appear at the back.
- The review cadence and the date of the next review are stated, and the linkages between objectives, risk, allocation and liquidity are audited at every revision.
- Schedules and exhibits carry their own dates, and one current version is published at a stable address.

Governance

- Every decision-making body is named, with its authority, its delegations and the standard of care it applies.
- Conflicts of interest, escalation beyond delegated authority and the amendment process are set out.

Objectives and risk

- The primary objectives are measurable, with a stated return metric, horizon and a hierarchy for when they conflict.
- Risk tolerance is a number — volatility, drawdown, funded status or shortfall — with a named monitor and a stated response to a breach.
- A tail-risk limit sits beside the volatility target, and an active-risk (tracking-error) budget governs active management.
- Leverage is defined, capped and approved by a named body, or prohibited in terms.

Asset allocation

- Every asset class, or every role, has a target, a range and a benchmark.
- Where the allocation is organised by role, each role states what it is for, which assets may fill it and the test that shows whether it worked.
- The policy portfolio is measured against a simple passive reference portfolio, to test whether its complexity pays.
- Rebalancing triggers, authority and cash-flow use are stated, including the

treatment of illiquid assets that cannot be rebalanced.

- Temporary deviations from a range carry an end date.

Spending and liquidity

• The projected net cash-flow profile, its source and its refresh cycle are stated, and actual flows are compared with it each year.

• The link to the assumed rate of return and the response to a contribution shortfall are set out.

• Assets are assigned to liquidity tiers, a coverage ratio over benefit payments and expected capital calls has a floor, and a stress test sizes the buffer.

Private markets

• The purpose of each private allocation and an overall illiquidity budget, including unfunded commitments, are stated.

• A board-approved pacing plan, with limits per year, manager and vintage, is the operative decision.

• An overweight caused by falling public markets does not force sales, and the route back to target is stated.

• Valuation lag, long-horizon benchmarks, fee and carry transparency, concentration limits, fund-level leverage and secondary-sale authority are each addressed.

Implementation and measurement

• Manager selection criteria, monitoring, termination triggers, fee monitoring, proxy voting and securities lending are specified.

• Every asset class has a named benchmark, the total-fund benchmark's construction is explained, and the book of record and calculation method are stated.

Responsible investment

• A dedicated section states whether and how ESG risk enters decisions, the stewardship and proxy-voting approach, who is accountable and how it is reported — or states, with reasons, that the fund does not integrate it.

Appendix A: Cross-Reference to the Empirical Assessment

This paper is intended as a companion to the empirical assessment titled *Assessment of Investment Policy Statement Quality Across U.S. Public Pension Funds* (research_paper.md). Readers seeking the underlying data, institutional rankings, score distributions, and category-level visualizations should refer to that document. The two papers are designed to be read together: the empirical paper describes the state of the sector, and this paper articulates what the sector should aspire to.

Appendix B: Rubric Reference and Provenance

This paper uses the same rubric framework as the empirical assessment: ips_rubric_v1_3, comprising 14 domains and 102 fact elements aggregated into six high-level categories. The scoring scale runs from 1 (Deficient — missing or purely aspirational) to 5 (Best Practice — fully specified, enforceable, peer-aligned). Domain definitions, required fact elements, and scoring anchors are documented in config/rubrics/ips_rubric_v1_3.yaml.

Provenance. The initial structure of the rubric — including the identification of the core domains, the required-fact taxonomy, and the scoring anchors — was built from the CFA Institute's *Elements of an Investment Policy Statement for Institutional Investors* (CFA Institute, 2010), which remains the most widely cited industry reference on IPS construction. Subsequent revisions (v1.1 through v1.3) extended the CFA-derived framework in three directions: (i) deeper treatment of liquidity, illiquid-asset governance, and private-market benchmarking, reflecting the increased private-asset allocation typical of contemporary public pensions; (ii) substantially expanded treatment of responsible investment and ESG integration, which the original CFA guidance addresses only briefly; and (iii) institution-sensitive scoring for spending / distribution policy, recognising that public pensions typically house distribution mechanics in separate plan documents and actuarial valuations rather than within the IPS itself. The hallmarks of excellence articulated in Part II of this paper — specificity, enforceability, internal coherence, completeness, reviewability, and operational realism — are consistent with, and substantially derived from, the CFA Institute's design principles.

End of Paper