Treasury Consultation Paper: Strengthening the Superannuation Performance Test

Opening Statement

CFA Society Australia welcomes the opportunity to comment on the Strengthening the superannuation performance test consultation paper. We are a member Society of CFA Institute, the global association of investment professionals. CFA Society Australia shares the CFA Institute mission of leading the investment profession with the highest standards of ethics, education, and professional excellence, for the ultimate benefit of society. As such, we do not represent any one part of the investment industry or have any commercial interest in the outcomes of this reform. Our comments are offered from the standpoint of measurement integrity, comparability, and the best financial interests of superannuation fund members.

CFA Institute sets globally recognised standards for professional excellence, ethics and education in investment management, and represents more than 200,000 charterholders worldwide across more than 160 markets. It champions ethical behaviour in investment markets and serves as the leading source of learning and research for the investment industry. With a belief in fostering an environment where investors' interests come first, markets function at their best, and economies grow, CFA Institute speaks on behalf of its members and advocates for investor protection and market integrity before standard setters, regulatory authorities, and legislative bodies worldwide.

CFA Institute maintains the Global Investment Performance Standards (GIPS®), the globally recognised ethical standards for calculating and presenting investment performance on the principles of fair representation and full disclosure. More than 1,700 organisations across 55 markets currently claim compliance, and a growing number of asset owners apply the GIPS standards to their own reporting. CFA Society Australia has consulted with CFA Institute and draws on this body of work, and on the expertise of Australian charterholders who have contributed to it, in the comments that follow.

CFA Society Australia supports the accountability that the performance test has brought to compulsory superannuation products, and the improved member outcomes that have followed. We consider, however, that certain unintended consequences have arisen, not from the concept of a test, but from the covered asset classes allowed and the benchmarks used.

Our submission makes four points:

Key point 1: The benchmarks, not the test itself, are the principal source of unintended consequences. The GIPS Standards Guidance Statement on Benchmarks for Asset Owners sets out the characteristics of a valid benchmark: several of the test's benchmarks, particularly for unlisted and alternative assets, do not meet them.

We support improving benchmark quality in accordance with this guidance statement and the further detail and analysis provided below.

Key point 2: Reform should refine the test, not dilute member protection. The test has materially reduced the number of members in underperforming products, and that outcome should be preserved while the approach is refined to support meaningful long-term results for members.

Key point 3: Rather than expanding the test to additional products, the expectation could be set that superannuation funds calculate and present performance in accordance with the GIPS standards, or claim compliance with them.

Key point 4: Members' ability to understand the information matters. The GIPS standards have been prepared with the oversight body and the investor in mind, and foster comparability, accuracy, and understanding of the strategy and results presented. Advancing financial literacy is a necessary complement to transparency.

Further information: The GIPS standards and benchmarks (Key point 1)
The GIPS standards require that the benchmark chosen for any total fund or composite must be an appropriate total return benchmark. A benchmark is appropriate if it reflects the total fund's or composite's investment mandate, objective, or strategy. There may be multiple benchmarks, however, that meet this single criterion, particularly for an additional composite or the components of a policy benchmark. Asset owners should therefore consider additional characteristics of valid benchmarks when selecting a benchmark.

A valid benchmark is one that is:

- **Specified in advance:** asset owners should select a total fund or composite benchmark prior to the evaluation period. *The Your Future Your Super (YFYS) benchmarks were selected and applied retrospectively.*
- **Relevant:** the benchmark reflects the investment mandate, objective, or strategy of the total fund or composite. *As the regulator is making the benchmark selection, it does not take into account specific investment mandates and strategies of individual superannuation funds.*
- **Measurable:** the benchmark is quantifiable. *Some benchmarks, such as the Unlisted Infrastructure benchmark, are not measurable ex ante as the benchmark figure is a median of a peer group and the median constituent changes each measurement period.*
- **Unambiguous:** the constituents of the investable universe can be clearly identified and priced.
- **Representative of current investment options:** the asset owner has current knowledge of the investable universe.
- **Accountable:** the asset owner selects the benchmark and holds itself accountable for any deviations from the benchmark.
- **Investable:** it is possible to forgo active management and simply hold the benchmark. *Many of the unlisted benchmarks are not investable as constituent funds are not open to new investors.*
- **Complete:** the benchmark provides a broad representation of the segment of the market to which it pertains. *Some of the unlisted benchmarks are limited to funds that can provide returns data within a short period post quarter end. Further, the benchmarks do not represent the investable universe.*

Benchmarks may not have all of these qualities, particularly benchmarks used for alternative asset classes. Asset owners should choose an appropriate benchmark that has as many of these characteristics as possible.

Total funds often use a policy benchmark. Asset owners may conduct an asset–liability study or another type of analysis to arrive at the asset classes and appropriate weightings (strategic asset allocation) to achieve an expected return and risk profile for a total fund. Asset class benchmarks are chosen to reflect each underlying asset class and typically represent the segment of the market from which investments for the strategy are drawn. These benchmarks may be an index or another type of benchmark, depending on the asset class. The asset owner then creates a policy benchmark that incorporates the strategic weightings of asset class benchmarks and is the default position. Investment decisions are made relative to benchmark weights, exposures, and risks.

Some asset owners measure performance results against a target benchmark (for example CPI + 5%) to determine if the total fund is meeting a minimum level of performance that is required to meet its liabilities.

According to the GIPS Standards Asset Owner Performance Survey Report, weighted benchmarks based on policy asset-class weights are the most common choice, used by 61% of asset owners. Target returns, weighted benchmarks based on actual asset-class weights, and market indexes are the next most commonly used benchmarks. Whereas 59% of respondents use only one benchmark, 41% use more than one. Some respondents indicated using three, four, or even five total fund benchmarks.

There are some considerations needed for when an asset owner calculates performance for a multi-asset class portfolio that includes alternative strategies and its benchmark(s).

Peer groups are often used as benchmarks for alternative strategies. Some common problems with peer group benchmarks include the following:

- Self-reporting bias: only some pooled funds choose to report performance data.
- Survivorship bias: historical returns of closed pooled funds are removed from the peer group benchmark.
- Inability to obtain returns for the same periods as the composite.
- Lack of investability: some pooled funds within a peer group benchmark are closed to new investors.

Also, peer group returns and risk statistics – which must be presented in a GIPS Asset Owner Report that includes time-weighted returns – are not necessarily meaningful. Because the peer group return is typically a median pooled fund return over a specific time period, the median fund could differ for each time period used in the peer group return and risk statistic calculations.

Although the use of peer groups as benchmarks is not considered best practice, in certain asset classes (for example private equity, real estate, alternatives) peer groups are widely used and generally considered the best available option. When using benchmarks that have limitations, such as peer group benchmarks, the asset owner should disclose these limitations. It is expected that the benchmark description would include whether the peer group benchmark is gross or net of fees.

Further information: Benefits of the GIPS standards (Key point 3)
Some of the benefits of the GIPS standards are:

- GIPS compliance helps ensure that asset owner's investment performance is complete and fairly presented regardless of whether assets are managed internally, externally, or both.
- Compliance requires establishment of robust investment performance policies and procedures, and it allows an asset owner to be confident that the data being presented to the oversight body is consistent and transparent.
- Compliance indicates a commitment to adopt the same performance standards often required of external investment managers the asset owner retains.
- Compliance demonstrates to legislative bodies, oversight bodies, and the general public a voluntary commitment to follow ethical standards.

There is a meaningful group of asset owners globally that have taken the step to not only require or ask the firms they hire to claim compliance, but also to claim compliance themselves. More information and an overview of the GIPS standards for asset owners is provided via the linked resources below.

Further input
CFA Society Australia are able to further support this important work and would be pleased to connect Treasury with Beth Kaiser Schwartz, CFA, CIPM, Director, Global Industry Standards at CFA Institute, as well as Australian practitioners and CFA charterholders who have significantly contributed to the GIPS standards.

We have provided the following supporting references:

- GIPS Standards Guidance Statement on Benchmarks for Asset Owners
- GIPS Standards Asset Owner Performance Survey Report https://www.gipsstandards.org/wp-content/uploads/2025/05/asset-owner-performance-survey-report-2024.pdf
- Introduction to the GIPS Standards for Asset Owners (for Boards and Decision Makers)
- introduction to the GIPS Standards for Asset Owners (Requirements)
- 2020 GIPS Standards for Asset Owners
- GIPS Standards Handbook for Asset Owners

Option 1.1 – Introduce a new emerging covered asset class

Q1 To what extent does the test constrain decisions to invest in certain asset types? Which asset types are most affected, and why? Provide evidence or examples to support your views.

The test creates real constraints for some investment decisions.

These constraints can come from benchmark quality. Where the benchmark for a covered asset poorly represents the underlying strategy, trustees face elevated and partly involuntary tracking error against an index that does not reflect what they actually hold. This may apply to unlisted infrastructure, unlisted and private property, private credit, venture capital, as well as values-based or transition-aligned strategies such as those with exclusions. Where unfamiliar, regulator-selected benchmarks are applied retrospectively, the tracking-error problem is compounded.

The GIPS Standards Guidance Statement on Benchmarks for Asset Owners sets out criteria that a well-designed benchmark should meet. A benchmark should be investable, unambiguous, measurable, appropriate to the investment mandate, specified in advance, and reflective of current investment knowledge. The current Alternatives benchmark (Appendix A, Item 24), which is a blended international equity and fixed income composite, does not meet the 'appropriate' criterion for most alternative assets, because the risk drivers of venture capital, private infrastructure, and social housing are different from listed global equities.

While the test does constrain decisions, there are also plausible alternative causes of changes to investment behaviour. For example, as firms consolidate to fewer, larger funds, these very large funds may not be able to deploy capital into venture capital at a scale that generates outcomes that justify the diligence and governance effort. In addition, current market conditions may result in some types of investments, such as social and affordable housing, carrying lower expected returns.

Benchmark hugging may also be a behavioural choice. The ten-year assessment horizon affords trustees considerable latitude, and improving benchmark quality would reduce hugging more durably than weakening the test.

Such benchmark hugging behaviour becomes more prominent should the probability of failing the test rise due to a period of underperformance. There are times where this behaviour would otherwise be imprudent because a period of underperformance in a particular active strategy can be followed by a rebound.

In addition, another form of benchmark hugging is the need to update strategic asset allocation (SAA) weights regularly to match actual unlisted asset allocation to manage the risk of underperformance. This is another form of short-term behaviour which is not in alignment with long-term objectives.

Benchmark hugging is, in part, a result of poor benchmark design rather than a fundamental flaw in the test concept itself.

Q2 To what extent would a new covered asset class with a CPI + X benchmark address investment constraints the test may impose? Provide evidence or examples to support your views.

A CPI + X figure is an absolute return hurdle rather than a benchmark in the accepted professional sense: it is not investable and has no defined constituents or weights, so it does not represent a passive opportunity cost that a trustee could actually have held.

A CPI + X benchmark may also be less suitable for certain types of assets. Inflation-linked assets would tend to exhibit lower tracking error, while assets with limited correlation to inflation may experience significant divergence.

This limits the extent to which performance attribution can be meaningfully conducted, since a CPI + X threshold does not represent an investment alternative against which a trustee's decisions can be fairly assessed in the way the existing benchmark portfolio does. It also creates outcomes that are sensitive to inflation cycles for reasons unrelated to investment skill. On balance, we consider Option 1.2 to be the more principled reform.

Q3.4 What allocation cap should apply, and why?

While the intent of an allocation cap is understood, this constrains trustees to allocate no more than (a proposed) 5% to the relevant asset class. Constraining allocations in this way would be undesirable and counterproductive to achieving optimal performance. Some asset owners operate within acceptable asset class ranges.

Q4 What are the key implementation and transition considerations for this option?

Introducing the CPI + X benchmark as an interim measure for venture capital during the J-curve is problematic. A newly established venture capital program is likely to deliver negative returns whereas the CPI + X benchmark is likely to deliver positive returns. The 10-year measurement period is a better way of addressing venture capital J-curves.

Option 1.2 – Improve the existing Alternatives covered asset class

Q5 To what extent would improving the Alternatives covered asset class address investment constraints the test may impose? Provide evidence or examples to support your views.

CFA Society Australia supports improving the Alternatives covered-asset class, over the introduction of a CPI + X benchmark.

Improving the Alternatives covered asset class would directly address investment constraints for funds with meaningful allocations to assets that are currently poorly benchmarked within that category. The extent of relief would be material but not complete. It would reduce involuntary tracking error for a defined set of alternative asset types but would not fully resolve benchmark hugging pressures for funds whose primary constraint lies in listed asset classes or values-based strategies.

The clearest example of the problem is renewable energy infrastructure. As noted in the consultation paper, investments in renewable energy projects are currently benchmarked against the MSCI Australia Quarterly Private Infrastructure Fund Index (Appendix A, Items 14 and 15), which is predominantly comprised of power, airports, and other transport infrastructure. A fund increasing its allocation to renewable energy, in members' long-term interests and consistent with the energy transition, will incur tracking error against this benchmark not because its investment decisions are poor, but because the benchmark does not reflect the asset. This is a benchmark design problem, not a performance problem. Improving the Alternatives covered asset class to better reflect the risk and return characteristics of such assets would remove this involuntary constraint.

Similar misspecification affects social and affordable housing, private credit, and certain climate-related strategies. In each case the existing benchmark proxies listed market exposures that bear little relationship to the underlying asset's risk and return drivers.

The return and risk profile of venture capital and private equity are so different from the listed market. The J-curve mechanism is the structural reason why the listed market benchmarks would be inappropriate. In addition, venture capital and private equity are often smaller companies than those held in the listed benchmarks and the industry composition of venture capital and private equity can be quite different to the industry composition of the listed market benchmarks.

The GIPS Standards Guidance Statement on Benchmarks for Asset Owners provides an appropriate framework for reform. Benchmarks should be investable, appropriate to the mandate, and specified in advance. Where a suitable investable index exists, even if imperfect, it is preferable to an absolute return hurdle. Where no suitable index yet exists, the routine review process proposed under Option 3 provides the mechanism to develop one over time. Any changes should apply prospectively only, consistent with the GIPS standards principle that benchmarks be specified in advance.

Q6 How could the Alternatives covered asset classes be improved?

Q6.1 What asset types should be eligible, and why?

Some eligible asset types should be:

- Venture capital.
- Private credit and direct lending.
- Renewable energy.
- Social housing.

Q6.2 What changes should be made to the benchmarks, and why?

The current benchmarks for alternative strategies do not bear any relationship to their actual return drivers. The current design of the benchmark creates a systematic bias as whenever listed equity markets perform strongly, an alternative strategy will more likely underperform its benchmark. Similarly, for private credit and direct lending, the benchmark is a fixed income index that does not reflect their illiquidity premium or credit risk characteristics.

Impact investments do not sit naturally within any existing listed market category due to a different return profile.

Alternatives are currently compared against the assumed fees of the underlying asset classes, which are materially lower than actual alternative investment fees. This creates an additional headwind for alternative strategies that is separate from the problem of having inappropriate benchmarks.

The GIPS standards require benchmarks to reflect the investment mandate and strategy of the portfolio. While a peer benchmark has challenges, it remains the most relevant benchmark for various alternative strategies.

Assets with an identifiable market index that represents their risk-return characteristics should remain in their existing covered asset class. Only assets with no appropriate index should be eligible for reclassification into Alternative Strategies.

Some other considerations with regard to benchmarks for multi-strategy portfolios that include alternative strategies are asset allocation transition, and lagged benchmark.

Asset Allocation Transition

When an asset owner changes an asset allocation, it is usually not possible for asset owners to move significant amounts of money from one asset class to another within a short period of time, especially when the asset allocation change involves illiquid asset classes. Buying and selling large amounts of assets within a short period of time may also have undesirable market impacts. For these reasons, asset owners typically transition from one asset allocation to another over an extended period of time according to a pre-defined transition schedule that eventually will reach the target policy benchmark allocation by a predefined date.

When transitioning from one asset allocation to another, asset owners may wish to calculate a benchmark that reduces tracking error between the total fund or composite and the benchmark during periods of rebalancing, and the rebalancing should be done using a predefined schedule. The asset owner would change weightings incrementally on a regularly scheduled basis (for example monthly or quarterly). For example, if a total fund’s hedge fund asset allocation is currently 15% and the new asset allocation is 20%, the asset owner may increase the allocation to hedge funds by 1% each month until reaching the 20% target. In this instance, the asset owner would increase the policy benchmark hedge fund allocation by 1% each month for five months.

The rebalancing would be included as part of the custom benchmark description, which includes the benchmark components, weights, and rebalancing process.

Lagged Benchmark
Private market investments are often valued using the last available historical price or preliminary estimated values. Asset owners should consider how private market investments are valued when determining the appropriate benchmark for these assets, in order to prevent a mismatch of information between the total fund or composite and the benchmark. For example, information for the equity and fixed-income asset classes within a total fund and the related policy benchmark is as of 30 June, but the information from the private market investments and hedge funds is as of 31 March. In this instance, the asset owner should use the private equity benchmark return from the previous quarter (31 March) to match the timing of the return for the private equity asset class but use the equity and fixed-income benchmarks as of 30 June. The asset owner would disclose that there is a difference in valuation dates for those asset classes. The portfolio and benchmark timing should be synchronised to prevent differences due to timing mismatches. In addition to matching the timing, lagging the benchmark also addresses an availability issue. Private equity peer universe benchmarks are often not available until three months or more after the valuation date. It is not feasible to use the current period private equity benchmark in a timely fashion.

If an appropriate asset class benchmark is not available in a timely fashion for an additional composite, the asset owner may choose not to include an asset class benchmark return in the GIPS Asset Owner Report and must disclose why no asset class benchmark is presented. The asset owner may label the benchmark returns as preliminary and then update the GIPS Asset Owner Report when the benchmark returns are available. If the total fund policy benchmark has an asset class benchmark that is not available in a timely fashion, the asset owner might choose to use the actual return of that asset class or the target return of that asset class when calculating the total fund benchmark return. When using the asset class return or the target return as a component of the total fund benchmark calculation, the asset owner would disclose this fact as part of the benchmark description. The asset owner may label the GIPS Asset Owner Report as preliminary and then update the GIPS Asset Owner Report when the asset class benchmark returns are available.

Option 2 – Introduce an assessment of risk-adjusted returns

Q8 To what extent would introducing a simple reference portfolio address investment constraints the test may impose? Provide evidence or examples to support your views.

Treasury's paper considers replacing the existing strategic asset allocation benchmark with a simple reference portfolio.

Introducing a simple reference portfolio would be a step toward better risk-adjusted performance evaluation. The current performance test focuses on actual returns vs a prescribed benchmark return, overlooking the differences in risk or strategy.

A simple reference portfolio would address the benchmark-hugging problem more than Option 1. It removes the SAA benchmark portfolio entirely and replaces it with a single comparator. Trustees can invest in any asset classes without worrying about accumulating tracking error at each asset class.

There are also other considerations:

- The greater flexibility would make it harder to identify areas of underperformance.
- The proposed reference portfolio approach considers risk (as measured by volatility) but **not** returns per unit of risk. A range of other indicators to consider includes alpha, Sharpe ratio, and Sortino ratio as they measure returns earned per unit of risk taken, depending on the type of risk being measured.
- While volatility (measured by standard deviation) is a universal definition of risk, other risk measures that are widely used in the industry include downside deviation and drawdowns (maximum drawdown).

Q9 How could the simple reference portfolio be designed?

Q9.1 What asset classes and market indices should be included in the reference portfolio, and why?

The Reference Portfolio of New Zealand Superannuation (NZSF) offers a good example. It is a notional, low-cost portfolio designed to meet the long-term risk and return objectives of the fund. The current Reference Portfolio has 75% global equities, 5% NZ equities and 20% fixed income, and 0% property, currently reflecting an 80% growth and 20% defensive structure. In the past, property had 5% weight.

The design principles[1] for the NZSF Reference Portfolio are simple. The Reference Portfolio should be a simple, low-cost portfolio, be diversified, be appropriate in risk and return, be relevant to a NZ based investor and be an equilibrium construct.

The Reference Portfolio for Australian superannuation funds' products should also be a simple, low-cost portfolio. The assets are diversified and should represent the typical asset structure of the super funds to meet their long-term investment horizon and risk profile. The aim of the super funds, as active investors, is to add more value – after all fees and costs – to the fund than the reference (passive) approach would do.

Q9.2 Is volatility an appropriate proxy for risk? If not, is there an alternative approach?

For long-term investors quarterly valuations for unlisted assets are appropriate and reflect genuine illiquidity characteristics. However, as these assets are valued less frequently than listed equivalents, their measured volatility will be lower. This is not because they carry less risk, but because the valuation process smooths return variability over time.

A framework that uses volatility as a proxy for risk would therefore make unlisted assets appear relatively more attractive on a risk-adjusted basis, potentially incentivising additional allocation to these assets to reduce measured portfolio volatility. This creates a systematic bias in the test that is unrelated to investment merit.

Q9.3 What is an appropriate failure buffer to apply under this approach, and why?

The current failure buffer is minus 50 basis points under the SAA benchmark methodology. Under the reference portfolio approach, the buffer would need to be recalibrated. The decision on the buffer should be made based on empirical evidence rather than an arbitrary level. Treasury should model the historical distribution of total portfolio returns relative to the proposed reference portfolio across all products currently tested and set the buffer at the level that identifies material underperformance.

Option 3 – Introduce a routine review of the benchmarks

Q11 Would introducing a routine review support the ongoing effectiveness of the benchmarks? Provide evidence or examples to support your views.

CFA Society Australia supports a routine benchmark review. As investment markets evolve, the benchmarks used in the test must keep pace, not through ad hoc adjustments but through a structured, transparent, and predictable process.

The GIPS Standards Guidance Statement on Benchmarks for Asset Owners recommends that benchmarks be reviewed whenever there is a significant change in investment mandate or market structure. A formal review cycle would embed this principle in the regulatory framework, providing trustees with greater certainty about how and when benchmarks may change.

Changes arising from a review should apply prospectively only, with sufficient notice (we would suggest a minimum of 12 months) so that trustees are not exposed to retrospective performance penalties for investment decisions made under prior benchmark settings. This is consistent with the

[1] The 2020 Reference Portfolio Review

GIPS standards requirement that benchmarks be specified in advance. In addition, the review body should include independent technical experts alongside industry representatives, Consideration should also be given to allowing trustees greater flexibility in selecting benchmarks that are appropriate to their specific investment strategy.

We also note that if benchmarks are to be reviewed routinely, consideration should be given to whether specifying them in primary legislation remains appropriate. A more flexible mechanism, such as a disallowable instrument or APRA standard, would allow timely updates while maintaining transparency and parliamentary oversight.

The 2020 GIPS Standards for Asset Owners contains more information regarding changes to benchmarks under Provision 2.C.26.

Option 4 – Test externally directed accumulation products / Single-sector and retirement products

Q16 Are there specific products for which performance testing would be unsuitable? If so, why? Provide evidence or examples to support your views.

CFA Society Australia supports retaining a performance test for MySuper default products, and the trustee-directed Choice products currently covered by the test. In the case of MySuper default products, this has demonstrated protective benefits for members who have not actively engaged in selection of their superannuation fund.

Rather than extending the test to additional products we would instead recommend that the regulator sets out what is expected in presentation of performance. The Global Investment Performance Standards could be used by funds to fulfil the need for fair representation and full disclosure of investment performance, and comparability of reporting across funds to support investors in making informed choices. Superannuation funds could use the GIPS standards to calculate and present performance or take an additional step and make the claim of compliance themselves.

Critical to this, is supporting members to understand the information provided. The GIPS standards address this through the required disclosures, which mandate that funds provide clear information about the investment strategy, how performance was calculated, which benchmarks were used, the fees and costs deducted, and the time periods covered. This creates a standardised disclosure framework that allows for comparison of performance across funds on a consistent and meaningful basis, something the current binary pass/fail test does not achieve.

For products outside the current scope of the test, including externally directed products, single-sector products, and retirement products, disclosures consistent with the GIPS standards would provide accountability and comparability without the implementation difficulties of extending a mechanistic pass/fail mechanism to products where trustee control over investment outcomes is limited.

The benefits of the GIPS standards in this context are well established. Compliance helps ensure that an asset owner’s investment performance is complete and fairly presented, whether assets are managed internally, externally, or both. It requires robust performance policies and procedures, giving the oversight body confidence that the data presented to it is consistent and transparent. It demonstrates to legislative bodies, oversight bodies and the public a voluntary commitment to ethical standards. A meaningful group of asset owners globally have taken the step of claiming compliance themselves, not only requiring it of the firms they hire.