



Annual Superannuation Performance Test- design options

Consultation paper

March 2024

© Commonwealth of Australia 2024

This publication is available for your use under a [Creative Commons Attribution 3.0 Australia](https://creativecommons.org/licenses/by/3.0/au/legalcode) licence, with the exception of the Commonwealth Coat of Arms, the Treasury logo, photographs, images, signatures and where otherwise stated. The full licence terms are available from <http://creativecommons.org/licenses/by/3.0/au/legalcode>.



Use of Treasury material under a [Creative Commons Attribution 3.0 Australia](https://creativecommons.org/licenses/by/3.0/au/legalcode) licence requires you to attribute the work (but not in any way that suggests that the Treasury endorses you or your use of the work).

Treasury material used 'as supplied'

Provided you have not modified or transformed Treasury material in any way including, for example, by changing the Treasury text; calculating percentage changes; graphing or charting data; or deriving new statistics from published Treasury statistics — then Treasury prefers the following attribution:

Source: The Australian Government the Treasury.

Derivative material

If you have modified or transformed Treasury material, or derived new material from those of the Treasury in any way, then Treasury prefers the following attribution:

Based on The Australian Government the Treasury data.

Use of the Coat of Arms

The terms under which the Coat of Arms can be used are set out on the Department of the Prime Minister and Cabinet website (see www.pmc.gov.au/government/commonwealth-coat-arms).

Other uses

Enquiries regarding this licence and any other use of this document are welcome at:

Manager
Media Unit
The Treasury
Langton Crescent
Parkes ACT 2600
Email: media@treasury.gov.au

Consultation Process	3
Request for feedback and comments	3
Executive summary.....	4
Background.....	5
Options for Reform.....	10
Design options	11
1. Status quo – SAA Benchmark Portfolio.....	13
2. Alternative single-metric test	17
3. Multi-metric test.....	24
4. Alternative frameworks	31
Broader considerations for reform	32
Scope of the test.....	32
Fees	37
Consequences.....	41
Barriers to consolidation	42
Appendix.....	43
Glossary	43
Current performance testing benchmarks	44
Performance testing metrics	47

Consultation Process

Request for feedback and comments

The Government seeks views on options to improve the sophistication of the Annual Superannuation Performance Test and the specific questions raised in this paper.

Closing date for submissions: 19 April 2024

Email	YFYS@treasury.gov.au
Mail	Superannuation Efficiency and Performance Unit Retirement, Advice and Investment Division The Treasury Langton Crescent PARKES ACT 2600
Enquiries	Enquiries can be initially directed to the above email address.

The principles outlined in this paper have not received Government approval and are not yet law. As a consequence, this paper is merely a guide as to how the principles might operate.

Executive summary

Since its introduction in 2021, the annual superannuation performance test (the test) has improved returns for members by holding trustees to account for investment performance and encouraging continual improvement or exit of poor performing products. However, there is evidence that the test may be influencing investment decisions to the detriment of member outcomes, including discouraging investment in asset classes that may otherwise be in the best financial interests of members.

Following the Your Future, Your Super Review the performance test was updated to ensure it could be extended to trustee-directed products (TDPs) in the choice sector and to fine-tune the test with changes that were feasible given the timeframe for the 2023 test. The purpose of this consultation paper is to allow stakeholders to provide feedback on options to improve the sophistication of the test in the longer-term.

In considering improvements to the performance test, the Government is focused on ensuring the test holds trustees to account for delivering the best financial outcomes for members. The test should be sector-neutral so that it does not create a barrier to invest in assets that provide strong returns for members. The test must be strong and encourage improvements to performance.

To help start the conversation, this paper includes four broad options for consideration in improving the operation of the test:

1. ***Status quo*** – retain the current testing framework but improve it.
2. ***Alternative single-metric*** – consideration of a different single-metric framework that would better assess performance. This paper includes three specific examples which are the Sharpe ratio, a peer comparison, and a simple-reference portfolio frontier.
3. ***Multi-metric framework*** – consideration of a multiple metric framework that provides a more fulsome assessment of performance. This option includes two specific examples which are a framework that aligns with the APRA heatmaps, and a targeted three-metric test.
4. ***Alternative framework*** – an opportunity for stakeholders to put forward an option they see as most fitting to improve the operation of the performance test.

While this paper canvasses the benefits, drawbacks and technical considerations of specific reform options, the Government is open to alternative proposals from stakeholders which continue to hold trustees to account and improve member outcomes. Your feedback on the future design of the test will assist in its evolution towards an enduring test that will facilitate better retirement incomes for Australians.

Background

The *Superannuation (Objective) Bill 2023* provides that the objective of superannuation is to ‘preserve savings to deliver income for a dignified retirement, alongside government support, in an equitable and sustainable way.’ To meet this objective, it is imperative superannuation funds are managed efficiently and trustees are held to account for how they assist their members in achieving a dignified retirement.

For most Australians, superannuation is now one of their largest assets.¹ 78 per cent of Australians have money in their superannuation accounts that will be used to support them in retirement.²

The annual superannuation performance test was introduced to protect Australians’ retirement savings by holding trustees to account for the investment performance they deliver and the fees they charge to members. The introduction of the test in 2021 followed a recommendation from the Productivity Commission’s Inquiry into the Efficiency and Competitiveness of Australia’s Superannuation System that products be subject to an annual performance test for the ‘right to remain’ in the system.³

The test is conducted by the Australian Prudential Regulation Authority (APRA) and assesses the performance of a superannuation product by comparing its:

- historical investment performance against a benchmark return, based on the product’s strategic asset allocation (SAA)
- most recent administration fees against the median fees charged by their peer group.

Products that fail the test are subject to clear legislated consequences: trustees must write to affected members notifying them that their product has failed the test and if a product fails the test two years in a row, it is closed to new members until it passes a future test.

In addition, funds that fail the test can expect to be subjected to heightened supervision from APRA to ensure that trustees are delivering better outcomes for their members.

How is it going?

The test has applied to MySuper products since 1 July 2021 and to TDPs, a subset of the choice accumulation sector, since 1 July 2023. In applying the test to MySuper products, the test covered 80 products and accounted for 14 million member accounts containing \$900 billion in assets. In expanding to TDPs, the test covered a further 805 products, accounting for a further 4 million member accounts and \$360 billion in assets.

The test was designed to remove underperformers from the sector by setting an objective minimum standard with clear consequences for not meeting those benchmarks.

¹ Australian Bureau of Statistics, [Household Wealth and Wealth Distribution, Australia 2019-20](#) [data set], Australian Bureau of Statistics, 2022, accessed 2022.

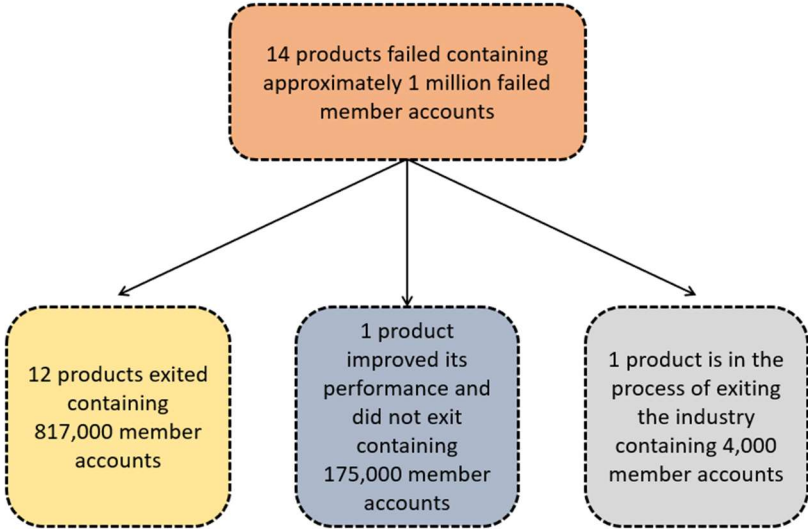
² Australian Prudential Regulation Authority, [Superannuation in Australia: a timeline](#), Australian Prudential Regulation Authority website, 2019, accessed 2022.

³ Productivity Commission, *Superannuation: Assessing Efficiency and Competitiveness – Inquiry report*, report to the Australian Government, 2019.

To date, 14 MySuper products have failed the test, of which 13 have exited the market or have announced plans to do so. This has resulted in over 800,000 member accounts merging with a better-performing fund.⁴ The remaining MySuper product has since improved its performance.

Each year of the test has had a lower failure rate of MySuper products, from 13 products failing in 2021 to five in 2022 and one in 2023. This demonstrates that the test has removed underperforming products in the MySuper sector, improving member outcomes, and enhancing transparency on the performance of their products.

Chart 1 – Status of MySuper products that have failed the test



Source: APRA Annual MySuper statistics back-series June 2014 - June 2022 and APRA Quarterly Superannuation Product Statistics, June 2023.

Since being extended to TDPs in 2023, the test found 12 per cent of TDPs were underperforming, with a higher failure rate of 25 per cent for platform TDPs (compared to only 4 per cent in non-platform TDPs). Members have been notified if their product failed the 2023 test, and failed products will now be subject to heightened APRA scrutiny and will be closed to new members if they fail in 2024 as well. As outlined above, the test has been effective at removing underperformers from the MySuper sector and is expected to drive similar outcomes for members in this segment of the market.

Without the test, affected members were unlikely to have known that they were in an underperforming product and would have remained there. APRA has monitored member movements from options that failed the test. They have seen the number of member accounts in products that failed the test fall by around 10 per cent in the five-month monitoring period following each of the tests, suggesting a positive impact on member engagement.

Further, data on the YourSuper comparison tool, which is a member focused online tool that aims to assist members to choose a well-performing MySuper product, shows a 30 per cent increase in traffic

⁴ Australian Prudential Regulation Authority, [APRA Annual MySuper statistics backseries June 2014 - June 2022 \(December 2022\)](#) [data set], Australian Prudential Regulation Authority, 2022, accessed 7 August 2023.

in the month of September compared to August, which aligns with the timing of members of underperforming products receiving notification letters.⁵

Heightened APRA scrutiny has also played a crucial role in driving member outcomes by placing an expectation on trustees to have a clear understanding of the drivers of their failure and to put in place an effective remediation plan. This includes consideration of whether transferring members to another product or fund would be in members’ best financial interests.

Concerns raised in Your Future, Your Super Review

In 2022, Treasury undertook a review of the Your Future, Your Super reforms (YFYS Review). The purpose of the review was to assess any unintended consequences and implementation issues arising from the YFYS legislation. The review covered all four elements of the YFYS measures.

During the YFYS Review, stakeholders raised concerns that the current test is a blunt tool that is leading to unintended consequences that affect the investment decisions of all funds (not just underperformers) and can potentially reduce long-term returns for members. The concerns raised by stakeholders are summarised in the table below.

Concerns raised by stakeholders	Summary of stakeholder views
Focusing on investment implementation over other measures of performance	<p>Performance delivered by trustees can be attributed to different factors. The current testing framework only assesses how well an investment strategy has been <i>implemented</i>. Stakeholders have raised concerns that the value delivered by a trustee’s option design and asset allocation decisions, a key factor in performance of the trustee, is not currently assessed.</p> <p>This means that a product with an investment strategy that is not suited for its membership and delivers low returns may pass the test if the implementation of the strategy is above the benchmark. Conversely, a product with a superior investment strategy that delivers good returns may fail the test if the implementation of the strategy is below the benchmark.</p>
Encourages short-term decision making	<p>Some superannuation funds have reported the test focuses their attention on the short-term impact that investment decisions will have on their next year’s performance test result. This detracts from the long-term outcomes that trustees should be considering when investing in assets.</p> <p>This risks funds prioritising investment in assets with more short-term certainty over assets that may provide superior long-term benefits to members.</p>
Incentive to hug benchmarks	<p>To manage the risks of failing the test, anecdotal evidence from stakeholders suggests that the test has created incentives for most superannuation funds, if not all, to ‘manage to the test’ and</p>

⁵ Australian Taxation Office, *Comparison tool statistics* [unpublished data provided to Treasury], Australian Tax Office, 2023.

	<p>seek to minimise their tracking error against the regulated benchmarks.</p> <p>This means that trustees focus on passing the test each year above seeking strong long-term investment returns for members. In some instances, they may be passing on opportunities to deliver better returns in lieu of more closely hugging the benchmark – and this includes trustees that have a track record of outperforming.</p> <p>Ultimately, this leads to superannuation funds viewing these benchmarks as a constraint on what they can invest in, which may have the long-term effect of dampening performance.</p>
<p>Lack of investment flexibility</p>	<p>There is consistent stakeholder feedback that the current test is not sector neutral and the benchmarks are better suited to investment in traditional asset classes (e.g., commercial property), but do not properly cater to possible investment opportunities (e.g. residential property in Australia). To avoid being identified as a poor performer, a superannuation fund may choose to invest in asset classes that are clearly covered under the benchmarks.</p> <p>For some trustees this has been taken as a barrier to particular investment strategies which would deliver strong outcomes for members. For example:</p> <ul style="list-style-type: none"> • Discouraging investment in assets that are not well-represented in the benchmark indices, including emerging asset classes such as those associated with the climate and energy transition, and housing. • Not supporting values-based investing, including faith-based or ESG focused investments, where the composition of the benchmarks do not align with these values of investing. This prevents investment strategies that can both deliver good returns and support member choice where members have made a deliberate decision to choose such products.
<p>Reduced choice, diversification, and active management</p>	<p>The influence of the test provides incentives for superannuation funds to avoid active investment management by encouraging passive benchmark hugging. This in turn may reduce diversification, as funds seek to divest from (or not direct new investment into) assets or investment strategies that are not well represented in the benchmark indices. However, trustees may prioritise passive investment strategies to lower fees, even if this provides lower net returns in the long-term.</p> <p>Ultimately, members could be left with less choice in where to invest their retirement savings, and funds may end up investing in the same assets which reduces diversification and increases systemic risks.</p>

In response to issues raised during the YFYS Review, the Government amended the test regulations with improvements that could feasibly be implemented in time for the 2023 test to:

- include additional benchmark asset classes
- amend key indices of concern
- increase the testing period.

These changes are expected to improve some of the unintended consequences identified in the YFYS Review, such as encouraging longer-term decision making, and supported the extension of the test to TDPs. However, these changes will not address *all* the identified potential unintended consequences.

More information on the YFYS Review, consultation process and outcomes can be found on the Treasury website at <https://treasury.gov.au/consultation/c2022-313936>.

Options for Reform

In considering potential reform options, the primary focus is on improving member outcomes through an enduring test that supports and incentivises trustees to achieve the best retirement outcomes for their members. Existing superannuation obligations, such as the best financial interests duty, the sole purpose test, and member outcomes assessments, are designed to provide guardrails around trustee decision-making.

To achieve this outcome, in developing the options outlined in this consultation paper, consideration has been given to the following key principles. They will be used when assessing potential changes to the testing framework.

Principle	Description
Improves member outcomes	<p>Consistent with the proposed objective of superannuation, any performance testing regime should be focused on delivering income for members' dignified retirement.</p> <p>As recommended by the Productivity Commission, the test protects members from being offered objectively poor products by setting a minimum bar or a 'right to remain', but in addition it should not inhibit funds from seeking better risk-adjusted returns.</p> <p>Any test should place the obligation on trustees to make decisions about what is in the best interests of their members and not create barriers to invest in assets that deliver good returns.</p>
Effective and efficient	<p>To be effective, the test must continue to be objective and have clear consequences for failure. This provides clarity as to when a product has failed and allows for efficient and timely regulator action.</p> <p>The test should seek to be effective in identifying underperformers, without constraining or misidentifying well-performing funds.</p> <p>The test should also be efficient and timely to administer from the perspective of both APRA and superannuation funds.</p>
Widely applicable and transparent	<p>To achieve a level of certainty, it is preferable to have a consistent and comparable measure of performance across all superannuation products.</p> <p>Ideally the test should be applied to as many product types as possible, without unfairly punishing products using particular investment mandates or strategies.</p> <p>The results of the test should be clearly communicated to members and the industry in a timely fashion.</p>
Enduring	<p>Superannuation is a long-term asset and, to remain consistent with this outlook, a testing framework should remain appropriate to apply consistently over time. The test should aim to remain applicable as markets change without requiring regular changes to the framework. This is important to ensure industry have certainty about the long-term testing regime, and that members can see long term performance that is not impacted by changes to the test that reset the rules or baseline.</p>

Consultation Question

1. Do you agree with these principles? Are there any other principles that should be considered?

Design options

The Government is inviting views on a more sophisticated and objective performance test. The options presented below provide alternative approaches which are designed to guide and prompt stakeholders. The preferred design should be guided by the principles set out above. For the purposes of this consultation paper, the options are grouped into four categories:

Option	Objective	Measure of Performance
Option 1 - Status quo		
1. Current Test	Assesses how well a trustee has implemented their investment strategy , based on SAA.	<p>Product Performance Long-term annual average investment performance <i>less</i> most recent administration fees.</p> <p>Benchmark Benchmark portfolio of indices based on the product's individual SAA <i>less</i> median administration fees of relevant peers.</p>
Option 2 - Alternative single metric		
2a. Sharpe ratio	Assesses how effectively the trustee delivers risk-adjusted investment returns above that of the risk-free rate .	<p>Product Performance Long-term average investment performance less the risk-free rate. Result is then divided by volatility of investment.</p> <p>Benchmark Multiple options, including a prescribed number (such as 1), peer comparison, or Sharpe ratio of a benchmark portfolio.</p>
2b. Peer comparison of risk-adjusted returns	Assesses whether a product is providing competitive risk-adjusted returns compared to peers .	<p>Product Performance Long-term average investment performance (net of administration fees) against its exposure to growth assets (as a proxy for risk).</p> <p>Benchmark A linear trendline based on results for the relevant product cohort (e.g., MySuper).</p>
2c. Risk-adjusted returns relative to Simple Reference Portfolio (SRP) frontier	Assesses whether a product provides superior investment returns relative to a simple benchmark portfolio that bears a similar level of risk.	<p>Product Performance Long-term average investment performance (net of administration fees) relative to volatility (standard deviation).</p> <p>Benchmark A line that reflects the risk-adjusted returns of a simple reference portfolio, of bonds and equities, for all levels of risk.</p>
Option 3 - Multi-metric framework		
3a. Heatmap	Assesses the performance of a product against multiple metrics, similar to the APRA heatmaps, to provide a fulsome performance assessment .	<p>Product Performance Utilises eight metrics contained within the APRA heatmaps (investment performance (3), fees (2), and sustainability of member outcomes (3)).</p> <p>Benchmark Varies depending on metric but includes benchmark portfolios and peer comparisons.</p>

3b. Targeted three-metric	Assesses the performance of a product against a smaller set of metrics to provide a more fulsome assessment of performance relative to the current test, but is simpler than 3a.	Product Performance Three independent metrics measuring performance, such as risk-adjusted returns , implementation of promises to members, and cost to members. Benchmark Varies depending on metric, but could include peer comparisons, disclosed targets and/or benchmark comparisons.
Option 4 – Alternative Framework		
4. Alternative metrics	This option is an opportunity for stakeholders to put forward an alternative framework that addresses concerns with the current test and the principles outlined in this paper. Options one to three are only examples of test frameworks that could be used and feedback on the detail of these options is welcomed.	

A list of alternative metrics previously proposed by stakeholders has been compiled in the Appendix.

Cohorts of products

The current test separately assesses three cohorts of products by assessing the administration fees against the median administration fee of different peer groups. These peer groups are: MySuper, platform TDPs and non-platform TDPs. The decision to test administration fees separately for these cohorts of products accepts the premise that these cohorts provide materially different services to members and therefore should be assessed as such. For example, platform TDPs are expected to offer a higher level of member services and so members may be prepared to pay a higher administration fee.

Later in this paper consideration is given to options to expand the scope of the test to other superannuation products, such as externally-managed, single-sector and retirement products. Whilst the underlying principles outlined above are applicable to all types of superannuation products, this does not necessarily mean a one size fits all approach to the test is the right outcome and different product cohorts could have different testing considerations (such as the metrics they are tested against, consequences etc.). Throughout this paper, consideration should be given to how different cohorts could be treated under different design options to ensure that the test remains widely applicable and enduring.

The Government’s initial focus is to get the test right for products in the accumulation phase. There is separate work underway to refine the superannuation retirement phase settings (see ‘Retirement phase of superannuation - Discussion paper’ released December 2023). There is a role for Government and regulators in creating an environment that supports funds to deliver on members’ retirement needs and meet their retirement income covenant obligations. However, as set out in the recent discussion paper, trustees need to do more to understand their members’ retirement needs, set a vision for their members’ retirement outcomes, and provide well-rounded retirement products.

1. Status quo – SAA Benchmark Portfolio

The current test has been effective in removing underperforming products from the superannuation system and has encouraged funds to reduce the fees that they charge. However, there are several shortcomings of the current test that may limit its effectiveness in improving member outcomes in the longer term.

This option considers further adjustments to the current test's methodology to address those concerns and ultimately asks whether the current framework can be adjusted to sufficiently address the unintended consequences of the test.

Current Methodology

The existing performance test has been adapted through consultation with industry from the methodology proposed by the Productivity Commission and APRA's heatmap methodology.⁶ The test assesses a product's *net investment returns* against an objective benchmark portfolio tailored to its strategic asset allocation (SAA) and assesses its *administration fees* against its peers. The methodology involves calculating a *performance measure* for each product tested:

$$\text{Performance measure} = (\text{Actual return} - \text{Benchmark return}) + (\text{Benchmark RAFE} - \text{Actual RAFE})$$

There are two components to calculating the performance measure. Both have effectively a 50 per cent weighting in the current test, with an overall 50 basis points buffer applied.

First, the net investment return of a product over the past 10 years (*actual return*) is compared to a *benchmark return* over the same period.⁷ The benchmark return is a passive investment portfolio of indices tailored to the product's reported SAA. The benchmark portfolio was intended to act as a counterfactual set of assets that members could have been passively invested in, based on the investment strategy of the product over the same period. This provides an assessment of the value a trustee has added for their members for the fees that they charge.

Second, the product's representative administration fees and expenses (*actual RAFE*) for the most recent financial year is compared to the median RAFE (*benchmark RAFE*). RAFE is calculated to represent a member who has an account balance of \$50,000. The benchmark RAFE is calculated separately for MySuper products, platform TDPs, and non-platform TDPs.

Options to improve the current methodology

The current methodology has several variables that feed into the result which could be amended. The Government is interested in stakeholder views on what changes could be made to the methodology to ensure the test encourages trustees to improve outcomes for members and allows trustees to invest in assets that provide strong returns, regardless of the sector. These changes could include:

- the lookback period for assessment of actual return and actual RAFE
- the number of asset classes and selection of corresponding indices used in the test

⁶ Australian Prudential Regulation Authority, [Superannuation heatmaps](#) [data set], Australian Prudential Regulation Authority, 2023, accessed 10 July 2023.

⁷ Gross investment return net of fees for products is used instead of net investment return for products that do not report this, such as those on platforms; the lookback period is currently nine years but will increase to 10 years in 2024. A minimum of seven years will be required to be subject to the test.

- calculation of a product's RAFF (such as the assumption of a \$50,000 member balance, which could be set at a different amount)
- calibration or weighting of investment returns and RAFF, and the corresponding benchmarks, including the overall 50 basis points buffer.

Concerns with current test seeking to be addressed

Measurement of performance

Stakeholders during the YFYS Review expressed concern that this single-metric using a benchmark portfolio only captures the *implementation* of an investment strategy, not the decision to set that strategy or account for risk, a key source of value-add which should be captured. Several suggested that this may impact the accuracy of the test, leading to some performing products failing the test (reducing member choice) and some underperforming products passing (reducing member outcomes).

The Productivity Commission acknowledged this concern, noting that benchmarking relative to asset allocation would not provide a direct measure of whether the investment strategy itself is appropriate for the member cohort. However, the Productivity Commission recognised that this aspect of performance would continue to be assessed through related elements of the regulatory framework, such as requirements to justify to APRA and the Australian Securities and Investments Commission (ASIC) that the investment options are appropriate for members (such as target market determinations, prudential standards, and legislated investment covenants).

Benchmarks

The original Productivity Commission recommendation was for a simple test where products would be compared to a limited listed investment benchmark portfolio tailored to their asset allocation. However, after consultation with industry, the test was disaggregated into more asset classes, increasing the number of indices that products are tested against. There are currently 20 indices across 26 asset classes in the test, compared to the Productivity Commission's recommended nine across 16 asset classes.⁸

Notably, the test was intended to only benchmark against *listed* indices, but stakeholders recommended the inclusion of unlisted assets to better reflect how they invest. However, unlisted assets are difficult to benchmark and there is less stakeholder consensus on the appropriate indices for unlisted assets.

While it is possible to include additional indices within the test, several factors may limit the benefit of this option:

- The test would continue to only assess the implementation of an investment strategy and would not provide any assessment of the strategy itself.
- Consensus on the right indices is difficult to achieve.
- For some emerging asset classes appropriate indices do not exist or are not as developed.

⁸ Productivity Commission, *Superannuation: Assessing Efficiency and Competitiveness, Inquiry Report no. 91*, Technical Supplement 4, report to the Australian Government, 2019.

- For some emerging asset classes, particularly ESG products, taxonomies are yet to be created that would underpin the population of products subject to different indices. Although ESG indices have developed since the introduction of the test in 2021, for many indices there is still a lack of sufficient historical data to allow a 10-year lookback history.⁹
- Additional benchmarks generally require new data collections by APRA, which will take time to implement.
- Trustees typically need to purchase the indices from private providers, the costs of which can be expensive and/or material, with the cost ultimately paid by members through their fees.
- More detailed benchmarks provide greater opportunity for trustees to ‘choose’ their benchmark, increasing the risk of trustees gaming the test.
- Where benchmarks need to be updated to keep pace with changing investment landscapes, trustees will have a lack of certainty about whether there is potential for other changes to the testing framework in the future.

Benchmark ‘hugging’

Many stakeholders during the YFYS Review also suggested that there is a strong incentive to ‘hug’ the benchmarks. This reduces the risk of failing the test by investing in the same assets that make up the benchmarks. Similarly, stakeholders have raised concerns that products which employ an investment strategy that has material sectoral differences relative to the benchmark face additional risk in failing the test due to higher levels of tracking error.¹⁰ An example of this is environmentally sustainable investment strategies that do not invest in fossil fuels, or high-emission assets.

This is not good for member outcomes. The current framework affects the investment decisions of all funds (not just underperformers). Instead of being neutral towards asset classes, the test may be unintentionally distorting investment choices and creates a disincentive to invest in any assets that are not well-represented in the benchmarks, even if they deliver superior investment returns to members.

One option to improve the test’s accuracy and reduce the risk of benchmark hugging is to recalibrate and disaggregate the existing set of benchmarks to include more asset classes, such as commodities, private equities, and longer duration bonds. However, as outlined above, this option has some limitations, particularly in relation to data availability.

An alternative option to improve the current test would be to simplify the test by reverting to the Productivity Commission’s original recommendation that the benchmarks only contain listed indices, or even further to a simple-reference portfolio comprised of only bonds and equities. While this may solve some issues, such as the cost of acquiring indices, this will come at the cost of increased tracking error.

⁹ The Government has recently consulted on the Sustainable Finance Strategy, which can be viewed at <https://treasury.gov.au/consultation/c2023-456756>.

¹⁰ Tracking error refers to the difference between the return generated by an investment and the benchmark.

Administration fees

The current performance measure reflects administration fees over the most recent 12 months, meaning administration fees charged over the most recent 12 months account for half of the test's formula. This has created a strong driver for underperforming funds to reduce fees to improve their results, which was an intended outcome.

There are divided opinions about the appropriate timeframe for assessing administration fees in the test. Some stakeholders prefer a longer lookback period for administration fees (8-10 years to align with the investment performance period) as the current approach does not reflect some trustees historical fees and may result in them shifting fees to other areas.

Further discussion on the appropriate testing regime for administration fees is considered in the *Fees* section.

Benchmark for performance

A product fails the test if the performance measure is lower than -0.005 (or -0.50 per cent / 50 basis points) below the benchmarks. This number was chosen as a buffer to account for short-term volatility and other factors outside the trustee's control. However, some stakeholders have argued that this number is arbitrary and could be better calibrated based on evidence. However, other stakeholders would see increasing this buffer as a weakening of the test and not in the interests of members.

Consultation Questions

2. Is assessing the implementation of a strategy, as opposed to assessing the choice of strategy itself, a strength or weakness of the current framework?
3. Can the existing methodology be materially improved, such as by further calibrating benchmarks, to largely address unintended consequences? How could these improvements overcome the incentive to benchmark hug, and remove barriers to invest in emerging asset classes?
4. What asset classes do you consider require better coverage in the test? What asset classes are covered well by the existing test?
5. Do you consider additional indices covering additional asset classes should be added to the test? If so, please provide the following details for **each** of your recommendations:
 - a. Description of asset class
 - b. Name of recommended index covering the above asset class, including the length of time data is available on the index
 - c. Details of appropriate fee and tax assumptions for such an asset class
 - d. Explanation of why you consider this index is appropriate for inclusion
6. How should the test cater for new asset classes in the future?
7. Should the threshold for failure be recalibrated for some products? What evidence supports the need for a different threshold? How could a different threshold deliver better long term returns to members?
8. Would retaining the current framework but moving to a simpler structure, such as a simple-reference portfolio of only bonds and equities, address some of the concerns with the current test?

2. Alternative single-metric test – Risk-adjusted returns

Rationale

Some stakeholders have argued that an alternative single-metric test could better assess performance and would have the benefit of being a simpler and more comprehensible framework when compared to a multi-metric test (discussed below). However, a key challenge lies in identifying an appropriate metric and benchmark that more effectively detects underperformance across a range of product types and improves member outcomes.

All options in this section focus on risk-adjusted returns, as this measures trustees providing a good return for members for a given level of risk. This measure is sector neutral but will benefit portfolios that are well-diversified, which include a range of unlisted assets.

Three alternative single-metric options are set out below – the first is an absolute measure (a numerical outcome) of a trustee’s ability to achieve risk-adjusted returns beyond the risk-free rate, the second is a peer comparison of risk-adjusted returns, and the third is a measurement of risk-adjusted returns relative to a simple-reference portfolio (SRP) frontier. It should be noted that the proposed measurements of performance in the second and third options could be paired with either of the benchmarks for performance proposed in those options (a peer comparison or SRP frontier).

2a. Sharpe ratio

Stakeholders have argued that the primary role of a fund is to provide good risk-adjusted returns for their members. The Sharpe ratio is a simple approach that seeks to compare the excess return (that is, the net return generated above the risk-free rate) against the volatility of the returns. This single-metric could be used as a broad assessment of the fund’s ability to generate good risk-adjusted returns for their members. Arguably, this could be a more appropriate testing metric than the current framework, as it provides a broader assessment of performance, not just SAA implementation.

Measurement of performance

Box 1 – Sharpe ratio

The Sharpe ratio, proposed by William Sharpe in 1966, is a widely used measure of investment performance. The Sharpe ratio seeks to calculate the risk-adjusted performance of an investment option by dividing the excess returns of an investment by its risk:

$$\text{Sharpe ratio} = \frac{(R_p - R_f)}{\sigma_p}$$

R_p = return of the product

R_f = risk free rate

σ_p = standard deviation of the product’s returns

‘Excess returns’ is used in the numerator as the measure of the product’s returns. It is the amount of return the product delivers above the risk-free rate (for example, the return of a long-term Government Bond).

The Sharpe ratio uses standard deviation of the product’s returns as a proxy for risk. Standard deviation is calculated by measuring the volatility of the product’s returns over time. Products with returns that are volatile (achieve periodic returns that are higher or lower than the long-term average return of the product) have a higher standard deviation.

Benchmark for performance

A key challenge for the Sharpe ratio is calibrating the bar for failing the test. Therefore, consideration needs to be given as to whether this will make for an enduring test that will be widely applicable to a range of products, particularly single-sector products that do not have as much influence over the volatility they experience. There are a few options for how the threshold for failure could be calibrated:

- The simplest is if a product falls below a certain margin of a benchmark ratio (either as an absolute figure or percentage of the benchmark ratio to account for a buffer) which would be possibly based on an historical average or other agreed calculation.
- A more complicated approach is if a product falls below a certain margin of a reference benchmark, such as relevant peer group or equivalent SRP. This option is covered in more detail in 2b below.

Benefits and drawbacks

The Sharpe ratio has several possible benefits over the current performance test:

- It is a reasonably simple calculation requiring limited data, particularly if using data which is already collected by APRA.
- It does not include any indices (beyond the risk-free rate), avoiding any benchmark hugging concerns and therefore does not create a barrier to investing in any particular asset class.
- It focuses trustees on only one metric – the unit price (or equivalent measure) of their product – which is ultimately the most important metric for members.
 - The test incentivises trustees to maximise the value of the unit price through whatever strategy they see fit, but also remain focused on minimising the volatility experienced in delivering that return.

Like any single metric, the Sharpe ratio has some drawbacks in its application:

- The Sharpe ratio treats upside volatility (that is returns that outperform the long-term average) and downside volatility in the same way, meaning the test may disadvantage a product with strong upside volatility.
- The use of standard deviation of the unit price of the product in the test will likely sharpen trustees' focus on how (and how often) assets are valued and how the unit price is calculated. This may incentivise funds to shift investment away from listed assets, which can have more volatility, to unlisted assets that may demonstrate less volatility.
- A consistent source of data will be required, particularly relating to the frequency of returns data and volatility. Depending on the frequency and composition of data, this may require trustees to report more data, or on a more frequent basis. Calculating volatility on a monthly or quarterly basis may be pragmatic.
- The Sharpe ratio's standard calculation generally does not include administration fees, so it would have to be amended to continue to provide transparency on fees as part of the performance test.
 - There is currently limited *net returns* data reported to APRA, outside of MySuper.

- The use of a simple benchmark for performance/failure (either an absolute figure or percentage) may not capture broader market forces. This could lead to ‘good’ products failing in down years, even if they outperform their peers.
- While appropriate for diversified products, it may be more difficult to use the Sharpe ratio for some more niche investment strategies or single-sector products where the trustee may not have as much influence over the volatility they experience.
- It would need to be designed carefully to ensure it’s accessible to members.

Consultation Questions

9. Would the Sharpe ratio be a more appropriate testing approach than the current framework? Would this lead to better member outcomes?
10. How should the benchmark for performance be calibrated?
11. What data should be used to estimate the Sharpe ratio, and how frequently?

2b. Peer comparison of risk-adjusted returns

Another single-metric approach could assess a risk-adjusted return metric against relevant peers to test whether a product has achieved competitive risk-adjusted returns. This approach could overcome a shortcoming of the Sharpe ratio – namely, determining the appropriate benchmark for products to pass – by setting the benchmark with reference to the performance of peer products.

Measurement of performance

This metric could measure performance by plotting the investment returns of a product against its exposure to growth assets, which would be a basic proxy measure of risk. This is the approach taken in one of the metrics used within the APRA heatmaps. This would provide a graphic representation of the risk-adjusted returns generated by all relevant products (example in Chart 2 below).

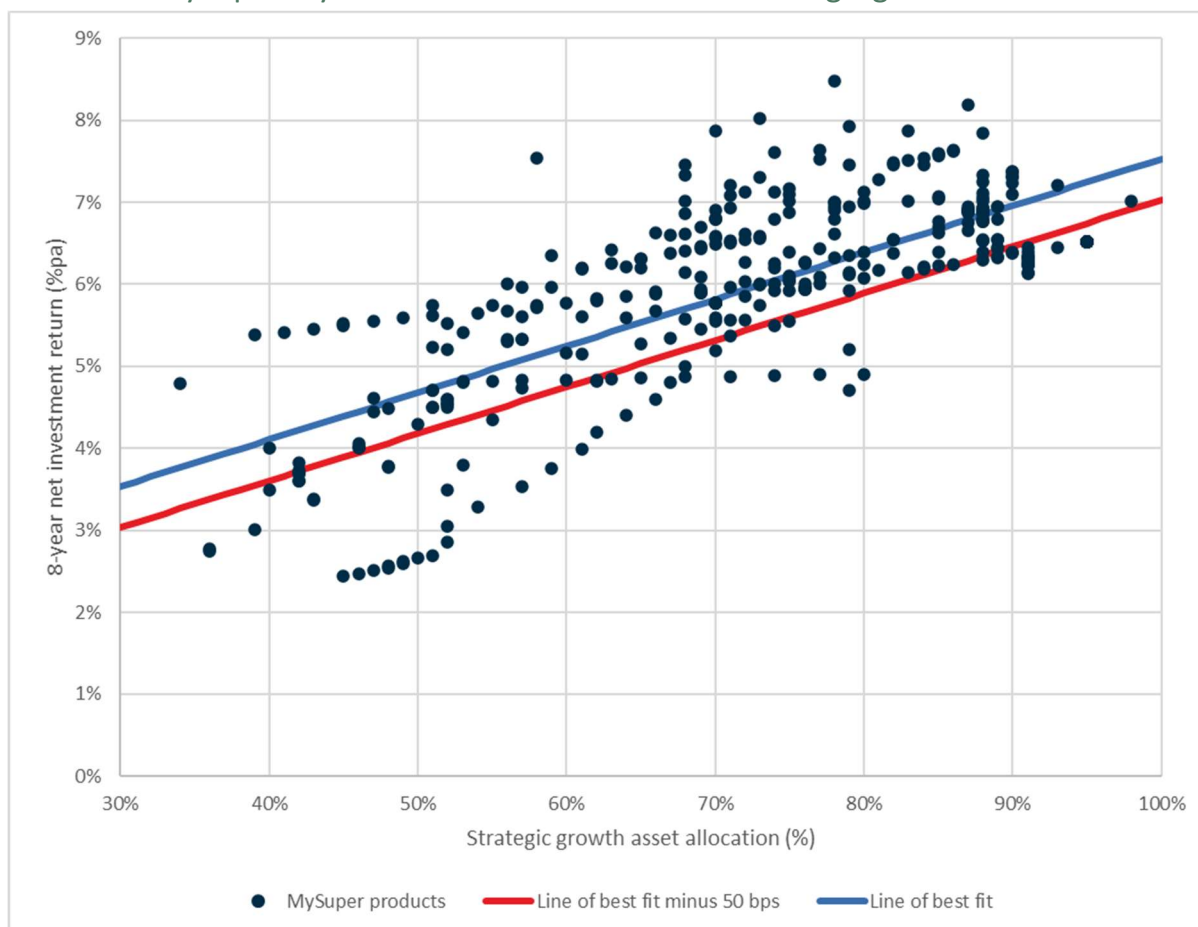
Benchmark for performance

By plotting the risk-adjusted returns of all products in a peer group, a linear line of best fit could then be generated across the expected returns within the peer group. Consideration would have to be given to determine the appropriate peer groups. Products that fall below a specified threshold of the trendline will have failed to deliver competitive returns for their members relative to their peers.

As an example, Chart 2 represents an assessment of products based on the returns they achieve relative to their exposure to growth assets. A linear line of best fit has been derived, representing the returns across the industry. The benchmark for performance could be set at 50 basis points below this line. Any products that chart below the benchmark line would fail the performance test.

Based on APRA heatmap data for MySuper products in 2021 and 2022 the failure rates would have been 19 per cent and 10 per cent respectively, assuming a failing margin of 0.50 per cent. These rates were higher than the failure rates of the current performance test, although the performance test also includes administration fees, which is not reflected in the below calculation.

Chart 2 – MySuper 8-year net investment return vs strategic growth asset allocation



Note: a linear line of best fit is used. Analysis also shows multiple results for lifecycle products as each life stage is represented separately.

Source: APRA 2022 MySuper Heatmap.

Benefits and drawbacks

This option has several benefits when compared to the current test:

- Focuses trustees on maximising risk-adjusted returns.
- Limited reliance on indices, which should mitigate incentives to hug certain indices.
- Using a peer comparison benchmark may facilitate greater competition amongst superannuation funds.
- The bar for passing the performance test can shift over time to reflect the market, resolving one of the key drawbacks of a simple Sharpe ratio and more accurately testing the value-add of the trustee.

However, there are drawbacks that should be considered when applying such a framework:

- There is no agreed industry view on what constitutes a ‘growth’ or ‘defensive’ asset, which is a limitation of using allocations to growth assets as a proxy for risk. Trustees may seek to game this benchmark by amending their SAA, particularly around assets that do not clearly fit within the definitions. However, using volatility as a proxy for risk (as outlined below) may mitigate these concerns.

- A peer comparison trendline may result in some value-adding products failing the test if they are performing worse relative to their peers, particularly as the industry returns increase over time as poor products exit the market.
- Values-based and other products that increase consumer choice but do not track traditional investment markets, and therefore their peers, may be at a disadvantage.
- Additional complexity might result in less digestibility of results for members, compared to simpler approaches.

2c. Risk-adjusted returns relative to simple-reference portfolio (SRP) frontier

Another single metric could incorporate elements of both the Sharpe ratio and the peer comparison metrics. This could measure the risk-adjusted returns against a benchmark that is calculated using an SRP frontier (discussed in detail below). Like the peer comparison option, this approach has an advantage over a simple absolute benchmark (such as a Sharpe ratio) because it provides greater weight to the value-add of the trustee, rather than performance being strongly influenced by overall market conditions.

Measurement of performance

Like the Sharpe ratio, this metric could focus on the net investment returns that a product has generated relative to their risk, reflected by the standard deviation of the investment returns (example in Chart 3). This addresses the concern in Option 2b of needing to define growth assets and thereby mitigates the risk of trustees gaming the test by changing their SAA.

Benchmark for performance

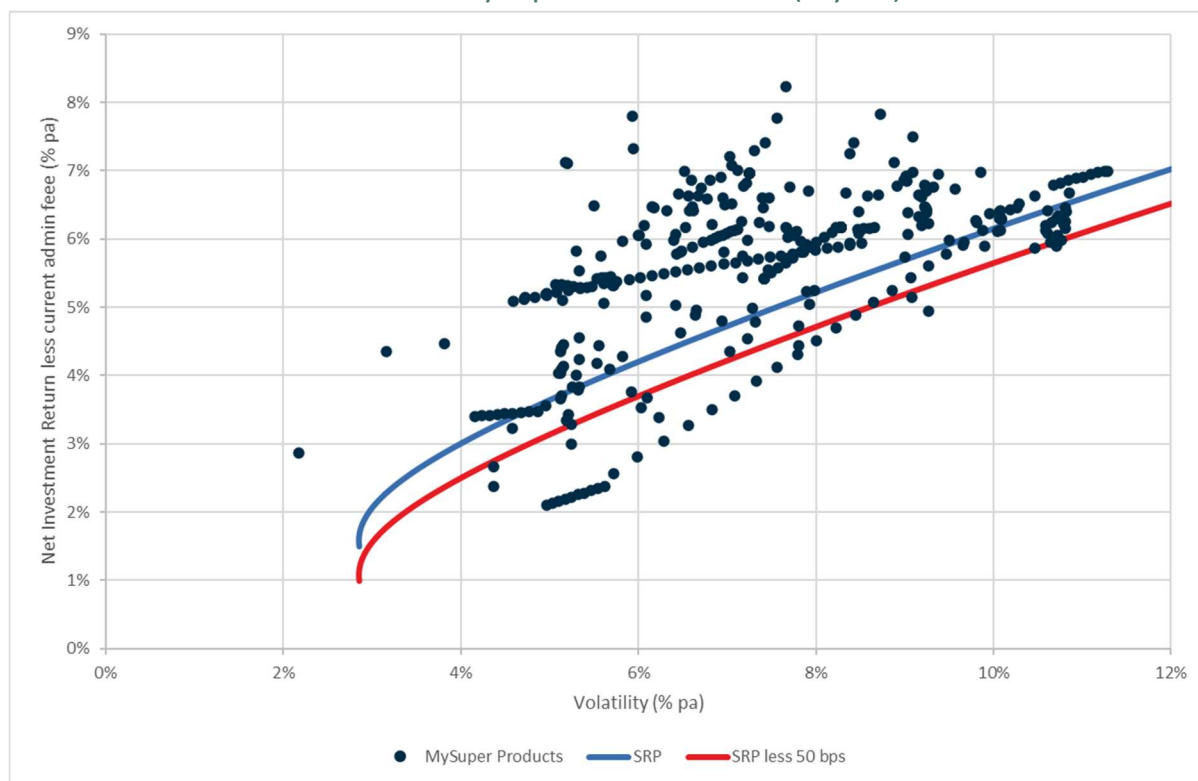
The relative benchmark for performance could involve comparing a product's performance against the hypothetical risk-adjusted return of an equivalent simple-reference portfolio (SRP). This would see the risk-adjusted returns compared against what the corresponding SRP with a similar risk profile would have achieved (the 'SRP frontier').

Instead of deriving a peer trendline, the SRP frontier reflects what a product invested only in bonds (defensive) and equities (growth) would have delivered at each point of the hypothetical SRP, stretching from 100 per cent bonds and 0 per cent equity (the left hand side of the benchmark) to 100 equity and 0 per cent bond (the right hand side of the benchmark). The threshold for failure could be set at the SRP frontier or a margin below the SRP frontier, and any products below this line would fail. This has the benefit of effectively establishing a counterfactual risk-adjusted return which can then demonstrate the value-add of the trustee's decision-making.

Chart 3 is a representation of a framework where MySuper products are assessed based on their net returns relative to the volatility (standard deviation) experienced.

This approach would have the advantage that the threshold for failure would be well known (assuming access to relevant equities and bonds indices) and understood by industry in advance of any APRA calculation and not be influenced by peer performance.

Chart 3 – MySuper Return vs Risk (8-year)



Note: Analysis also shows multiple results for lifecycle products as each life stage is represented separately.

Source: data sourced from Chant West

Benefits and drawbacks

This option has all the same benefits as the peer comparison approach in 2b, but includes the following additional benefits:

- This approach does not rely on a product's SAA, meaning it would not be subject to 'gaming' the reporting of the SAA.
- It will be possible for trustees to track and pre-empt their results, using the relevant bonds and equities indices.

However, there are drawbacks that should be considered when applying such a framework:

- Additional complexity around the concepts of volatility and the 'SRP frontier' might result in less digestibility of results for members, compared to simpler approaches.
- Products that do not out-perform simple equity and bond market indexes may be at a disadvantage.

Interactions between peer comparison and SRP frontier examples

The examples presented in options 2b and 2c offer multiple approaches to measuring performance through the mechanism to adjust returns to account for risk (growth allocations or standard deviation) as well as the benchmark for performance (peer trendline or SRP frontier). While these examples focus on two specific combinations, a different combination could be considered.

Consultation Questions

12. Are either of these approaches better than the existing test methodology (Option 1) or a simple Sharpe ratio (Option 2a)? Are there any other considerations that make this a better or worse option?
13. Are there any other alternative single-metrics that would be superior in addressing the principles set out in this paper? How would they provide a better testing framework? What net benefits do they provide over other proposed metrics?
14. What incentives would these alternative single-metric options provide trustees, and what would be the consequence of this for member outcomes?

3. Multi-metric test

Rationale

The rationale for a multi-metric test is the ability to obtain a more fulsome picture of a product's value to members in order to identify underperformance more effectively. This framework provides an opportunity to assess multiple aspects of a product's performance through multiple metrics and variables.¹¹ Therefore, a multi-metric framework contains multiple assessments of performance that are independent from one another.

A desirable outcome of a multi-metric test would be a reduction in unintended behavioural consequences, such as benchmark hugging or disincentives to invest in emerging asset classes. Further, this testing framework should be more accommodating to a wider range of products, such as values-based, given the impact of any one metric on the test outcome will be proportionate to the number of metrics.

However, designing a multi-metric test comes with additional complexity. The additional complexity of the framework may also make the test less digestible, and therefore less transparent, to members. Depending on the metrics used, this may involve additional data collections by APRA which increases the reporting burden on trustees.

A multi-metric test provides countless options in terms of constructing the most appropriate testing framework – two specific approaches are set out below.

3a. Alignment with the APRA heatmap

Some stakeholders have advocated for the test to adopt the format of the APRA heatmap which assesses multiple metrics, over multiple time periods and member balances.¹²

Measurement of performance

The performance test could assess performance through some, or all, of the metrics contained within the APRA heatmaps. The APRA heatmap has previously included eight metrics:¹³

- the implementation of the investment strategy against benchmark portfolios, through the SAA metric
- the implementation of the high-level investment strategy based on exposure to growth and defensive assets, through the SRP. While assessing implementation of the strategy, it provides more room for trustee value-add over the SAA metric
- a peer comparison of investment performance, through the net investment return (NIR) growth/defensive trendline metrics
- the cost of the product to members, through total fees and administration fees metrics

¹¹ For clarity, this paper makes a distinction between metrics and variables. A metric can be viewed as a method of assessing performance, whereas a variable is an input to the metric, such as a lookback period or representative member balance.

¹² Australian Prudential Regulation Authority, [Superannuation heatmaps](#) [data set], Australian Prudential Regulation Authority, 2023, accessed 10 July 2023.

¹³ Note net investment returns are coloured on the heatmap by distance away from a growth/defensive trend line. Net return metric has been excluded as it did not feature in the choice heatmap.

- the sustainability of members outcomes, through the Registrable Superannuation Entity (RSE) adjusted total accounts growth rate, the RSE net cash flow ratio, and the RSE net rollover ratio.

A number of these metrics are assessed against different variables. For example, the SRP and SAA are assessed against multiple time periods (three, five and eight years) and the fee metrics are assessed against differing member balances (\$10,000, \$20,000, \$50,000, \$100,000 and \$250,000).

Benchmark for performance

For the purposes of the heatmaps, APRA applies a 0.50 per cent margin to all their investment performance metrics to indicate the worst performance (i.e., darkest shade of red). This is consistent with the margin for failure used in the current performance test.

A testing framework that follows the APRA heatmaps approach would need to consider:

- How many of the eight metrics should be used?
- What margin for failure should apply to each metric?
- How are the results for each metric combined to a single outcome? Should it use a majority fail methodology (i.e., each metric has equal weighting), a more complex weighting mechanism, or a hierarchical methodology (discussed in more detail below)?

Benefits and drawbacks

The heatmap was raised as a possible alternative performance test during the YFYS Review, as it has the following benefits:

- It provides a more well-rounded picture of a product's performance.
- Industry is already familiar with these metrics.
- Providing many different metrics may reduce the ability and incentives for funds to hug particular benchmarks.
- Additional time periods can assist in demonstrating positive actions to address historical performance issues, such as changes in governance or staffing.

However, the number of metrics included in such a framework poses the following drawbacks:

- Several of the metrics contained in the heatmaps are highly correlated with one another, meaning that they are assessing the same outcome. This might increase complexity without providing additional insights.
 - For example, the SRP and SAA benchmarks both assess the implementation of an investment strategy, with SRP also assessing the overall investment strategy itself.
- Multiple time periods may increase short-termism, as more recent periods have a higher weighting than longer timeframes, due to appearing duplicate times.
 - Stakeholders have raised that independent time periods could adjust for this concern.
- It is difficult to combine multiple metrics into a digestible overall outcome and set the benchmark for failure.

The table below demonstrates the rate of failure for MySuper products based on the common investment performance metrics used in the heatmaps, using a failing margin of -0.50 per cent or worse.¹⁴ For comparison, we have included the rate of failure under the current test, as well as overall fail rates should a multi-metric test be based on failing the majority of these metrics, or failing all eight metrics.

Table 1: MySuper failure rates by metrics

	2021 (7 year lookback period)	2022 (8 year lookback period)
Current Performance Test Results	16%	6%
Net investment return relative to trendline	19%	10%
Net investment return relative to SRP benchmark	18%	7%
Net investment return relative to SAA benchmark	18%	9%
<i>Fail the majority of the net investment return metrics</i>	<i>16%</i>	<i>7%</i>
<i>Fail all three net investment return metrics</i>	<i>9%</i>	<i>6%</i>
Administration fees (\$50,000 account balance)	1%	7%
Total fees (\$50,000 account balance)	1%	1%
RSE adjusted total accounts growth rate (3 year average)	20%	17%
RSE net cash flow ratio (3 year average)	15%	13%
RSE net rollover ratio (3 year average)	16%	14%
<i>Fail half (or more) of the APRA heatmap metrics</i>	<i>3%</i>	<i>0%</i>
<i>Fail all eight metrics</i>	<i>0%</i>	<i>0%</i>

Source: APRA MySuper Heatmap 2022

Consultation Questions

15. Would greater alignment to the APRA heatmaps improve the sophistication of the test?
16. Would it reduce incentives to benchmark hug and improve member outcomes?
17. Is correlation between metrics an issue? If so, how should this be addressed?
18. Should the test capture all the metrics in the heatmap? If not, what metrics?
19. How would the benchmark for performance be calibrated for chosen metrics? How would these metrics combine to determine overall pass/failure of the test?
20. What costs would be associated with aligning the test to the heatmap? What would be the benefits?

¹⁴ Australian Prudential Regulation Authority, [Superannuation heatmaps](#) [data set], Australian Prudential Regulation Authority, 2023, accessed 7 November 2023.

3b. Targeted three-metric test

Another approach could involve testing a smaller set of metrics focused on separate indicators of performance that can be given equal standing, over a single lookback period. Many metrics allows for an assessment of more aspects of performance but comes at the cost of complexity and a three-metric test may provide a balance of performance assessment and comprehension.

Measurement of performance

Stakeholder views have varied on the most appropriate indicators of performance. Some agree with the current test that measures the implementation of an investment strategy, noting that it has been beneficial in providing members with a realistic promise and not an aspirational target. Many have instead advocated for an assessment of the design of the investment strategy. Other stakeholders have argued to test the absolute return of the product, fees charged to members, or performance relative to peers. In choosing the three most important indicators of performance, it is important to consider the degree in which they are assessing different aspects of performance and how the metrics fit together to create an assessment of a product's overall performance.

While there may be other aspects of performance, the three key aspects of performance that this option considers (as they are reasonably separate and able to be tested) are:

- The overall **return** the member receives – this aspect seeks to measure how well the trustee has delivered returns to members.
- How well the trustee has implemented its investment **promise** to members – this aspect seeks to measure how well the trustee has delivered on its promise to members.
- How much the service has **cost** members – this aspect seeks to measure how much the service the trustee has provided cost members.

Possible metrics

Once the aspects of performance are settled, it is important to consider the most appropriate metric to assess that indicator of performance. Ideally, these metrics will work independently and have a low correlation to one another (i.e., not measure the same outcome).

Overall returns

While the simplest measure of returns to members would be absolute returns, this approach does not reflect the level of risk a trustee has taken to deliver that return. As such, a risk-adjusted measure of returns could be considered. Possible metrics (which would exclude administration fees) include:

- Sharpe ratio – as outlined above, the Sharpe ratio is a measure of risk-adjusted returns that could be used for a diverse range of investment approaches.
- Peer comparison – as outlined above, broadly assess the overall return a product achieves relative to their peers. This will need to occur on a risk-adjusted basis by means such as comparing products based on their growth/defensive exposures, or volatility experienced.

Implementation of promise to members

There are several metrics that could be used to measure how well the trustee has implemented its promise to members. These will depend on what the trustee has promised members, which may vary between funds. Possible metrics include:

- Benchmarking against their own investment strategy – testing how well a trustee has implemented the investment strategy that is disclosed to their members, such as the current performance test (investment component only), or an SRP approach comprised of only bonds

and equities based on the SAA exposure to growth and defensive assets. Noting this option would be for the investment component only, as administration fees are considered under a separate metric below.

- CPI+X or Cash+X – it is common for trustees to make disclosures to members that identify a particular target range of returns above the rate of inflation. While this is potentially the simplest metric, it may be necessary to mandate clearer disclosures from trustees on their promise to members or include these promises on the ATO comparison tool. This would be to ensure that the CPI+X / Cash+X target is not deliberately low, as all products could move to a lower and easier promise. Under current ASIC guidance, these targets are often set as returns they are *more likely than not to meet*, instead of average or expected returns, and therefore very few funds may fail such a metric.

Cost

Normally a measure of costs to members would focus on the level of administration fees, as investment fees would typically be included in the investment measures. When utilising a multi-metric approach, a stand-alone metric on administration fees would generally be preferred. How to measure costs/fees will be an issue to consider under any of the options outlined in this paper, as well as in a multi-metric test. A detailed consideration of how fees can be measured, with options, is outlined later.

Benchmark for performance

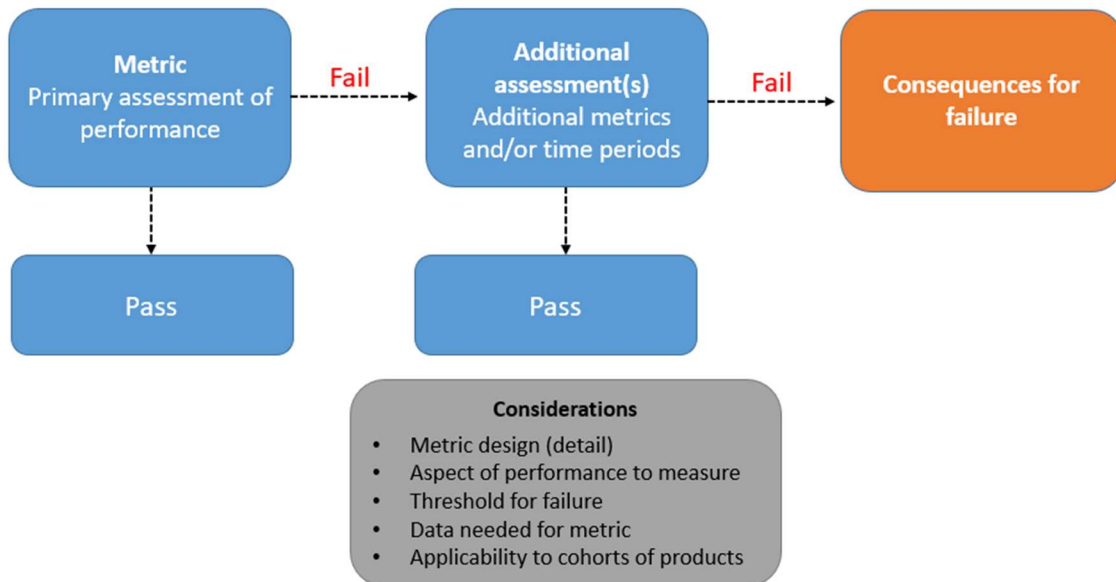
A decision will need to be made on the calibration of the threshold for failure for each metric. This threshold will be dependent on the nature of the selected metrics. However, a multi-metric framework introduces an additional consideration on how these metrics should combine to determine the overall threshold for failure.

Majority fail

Under a multi-metric approach, each of the three metrics chosen could have equal influence and a product will fail the test overall if they fail most metrics, which would be run independently. It is worth noting that in the proposed three-metric option, administration fees would count towards only one of the three metrics, meaning products could not pass the test by lowering fees in isolation – they would also need to pass at least one of the investment-related metrics.

Hierarchical fail

An alternative could be a hierarchical approach to passing or failing the performance test. This would essentially see a 'primary' assessment of performance, followed by supplementary assessments should a product fail the original test. This could include subjecting a product to additional metrics, or the same metric over a different period.



Benefits and drawbacks

The targeted three-metric approach has several benefits if implemented as a performance test:

- It allows for the assessment of multiple outcomes that both trustees and members may consider important in the overall performance of a product.
- Metrics that aim to measure different outcomes should have lower levels of correlation which should aid in reducing the incentives to game a particular metric.
- Having a smaller number of independent metrics testing separate aspects of performance, ideally an odd number with equal influence in the overall outcome, provides for a simpler framework for passing or failing the test. In this example, a product will fail the test if they fail most of the metrics (i.e., fail two out of the three metrics).

Like all testing frameworks, this example has some drawbacks to consider:

- Limiting the number of metrics means it will be difficult to achieve industry consensus on what the most important aspects of performance might be, and which metrics should test them.
- Using a single time period will not allow trustees to show demonstrable improvements to address historical performance issues.
 - Although having additional metrics may alleviate the need for additional time periods, compared to a single-metric.
- Using a hierarchical failure model will likely increase the complexity of the test as two separate tests will be run, making it difficult for APRA to complete the test in a timely manner.
- It would need to be designed in a way that doesn't lower accountability of trustees.

Consultation Questions

21. Would this framework improve the sophistication of the test? Would it reduce incentives to hug benchmarks and improve member outcomes?
22. Would this approach be more, or less, favourable than the heatmap approach?
23. What would the costs of implementing this approach be? What would the benefits be?
24. Are these the right measures of performance or are there other more important indicators of performance that should be measured in addition to or instead of those outlined? What metric should be used to assess these indicators?
25. How should the benchmark for performance be calibrated?

4. Alternative frameworks

The Government also welcomes stakeholder feedback on any alternative performance test frameworks that may more effectively address the principles set out in this paper than the options outlined. In providing feedback, it would be useful to consider the following elements when constructing a framework:

- What outcome(s) should be assessed and what metric(s) achieves this?
- What should the metric(s) be compared against to assess performance?
- How many metric(s) are required to form an adequate assessment?
- How does the framework determine whether a product passes or fails?

Consultation Questions

26. How would an alternative framework be constructed according to the elements outlined above? Please provide specific details.
27. How would this framework more effectively advance the principles outlined in this paper?
28. What would be the costs and benefits associated with this framework, compared to the current test and any other alternatives?

Broader considerations for reform

Further to considering options to reform the existing test, the Government is also inviting stakeholder views on a range of broader considerations that need to be taken into account should the Government seek to make future changes to the test. These include:

- the scope of the test;
- the assessment of fees;
- consequences for failing the test; and
- barriers to product consolidation.

Each of these considerations are important to ensure that the test is robust to changing circumstances and meets the key principles outlined earlier in this paper.

Scope of the test

Current product coverage

The test currently applies to MySuper products and TDPs. TDPs are accumulation phase investment options that have allocations to more than one asset class, and the trustee (or a connected entity) influences the investment outcomes by designing or implementing the investment strategy.

The diversified nature of MySuper and TDPs mean they are the most conducive to the current testing framework, which covers over 40 per cent of the total superannuation population (and 60 per cent of the APRA-regulated superannuation population).¹⁵ Trustees of these products have a level of influence on the outcomes of the product through activities such as designing the investment strategy.

Products to consider in an expanded test

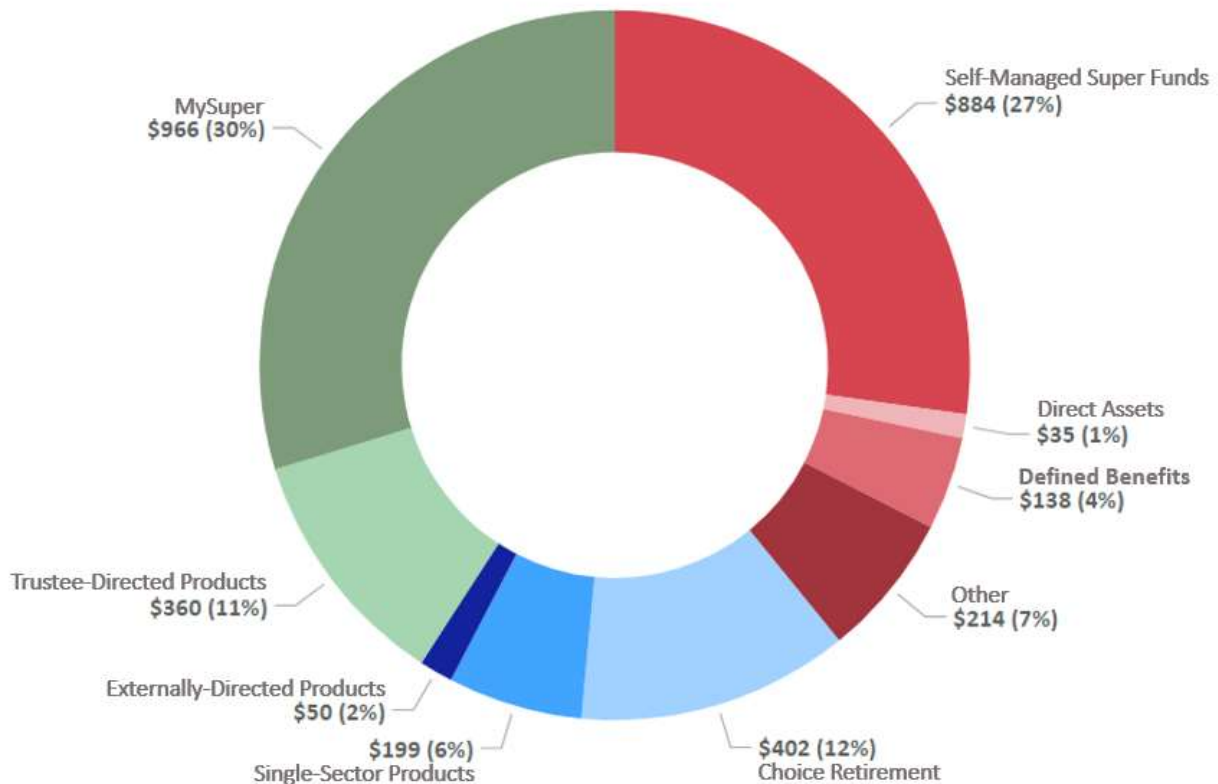
While the test is not currently legislated to extend further than the existing product coverage, many stakeholders have advocated for the test to apply to all APRA-regulated superannuation products. This has been on the basis that members in these products also have the right to be afforded the same transparency and protection from underperformance as those currently covered by the test. Extending the test may also uphold the integrity of the superannuation system by reducing incentives for trustees to:

- restructure products to avoid the test
- encourage members to move into products that are not tested
- increase fees on untested products to cross-subsidise tested products.

The current test was only designed to apply to multi-sector accumulation products. When considering changes to the test framework, it would be desirable that the test has the flexibility to remain fit for purpose for a wider range of products so that the Government could increase product coverage without a further redesign, giving industry stakeholders certainty about the future test design. Products that are not currently captured by the test, but could be considered in the future, can be roughly sorted into three categories – externally-managed, single-sector and retirement.

¹⁵ Excludes exempt public sector schemes; Australian Prudential Regulation Authority, [Quarterly Superannuation Industry publication \(June 2023\)](#), Australian Prudential Regulation Authority, 2023, accessed October 2023.

Chart 4 – Total assets in the superannuation system for June 2023 (in \$ billions)



Green represents assets already covered by the test, **blue** represents assets that could be tested and **red** represents assets that are impractical to test.

Other includes single member ADFs, small APRA funds, exempt public sector superannuation schemes and balance of life office statutory funds.

Source: APRA Quarterly Superannuation Industry Publication June 2023 and APRA Quarterly Superannuation Performance Statistics June 2023.

Externally-managed products

Externally-managed products are those where the trustee, or their connected entities, are not involved in the design or implementation of the investment strategy. There is around \$50 billion in assets in externally-managed products in the accumulation phase.¹⁶

Externally-managed products have not been tested to date, as it is less clear the role the trustee plays in influencing the outcomes of the product and there are additional technical nuances that would need to be considered in extending the test to these products. However, superannuation trustees charge members for administration services and have a duty to act in their members' best financial interests. This is likely to extend to ensuring that trustees only offer performing products, regardless of the level of control they maintain over the investments.

Administration fees must be carefully considered when testing externally-managed products, as the associated administration services are provided by the superannuation entity (separate to the responsible entity of the investment product). Therefore, administration fees will vary across superannuation funds that offer these products and will not be influenced by the responsible entity of the investment product. Consideration needs to be given to how these products could be tested,

¹⁶ Multi-sector choice accumulation products net of TDPs and MySuper, based on APRA-regulated entities with more than six members; Australian Prudential Regulation Authority, [Quarterly Superannuation Industry publication \(June 2023\)](#), Australian Prudential Regulation Authority, 2023, accessed August 2023.

given the member experience could vary significantly from one super fund to another and the impact of the consequences will be felt by multiple entities.

Single-sector products

Single-sector products are those that have a strategic asset allocation to only one asset class (or only two, where one has an allocation of less than 10 per cent). There is around \$200 billion invested in single-sector accumulation products, across nearly 6 million member accounts.¹⁷

Performance testing these products would be different to testing multi-sector products, as the ability of the trustee to value-add is limited to asset selection and fees charged. Trustees have a greater ability to influence the outcomes of multi-sector products through diversification and the design of the investment strategy. However, the Productivity Commission recommended that single-sector products be included in the test and noted that asset class returns exceeded benchmark returns for most asset classes.

Retirement products

The Government released a discussion paper to canvass stakeholder views on ways the superannuation system can best provide the security and income Australians need as they live longer and healthier lives in retirement. The discussion paper examines three key areas:

- Supporting members to navigate the retirement income system
- Supporting funds to deliver better retirement income products and services
- Making lifetime income products more accessible.

The Retirement phase of superannuation discussion paper can be accessed at <https://treasury.gov.au/consultation/c2023-441613>.

The following section of this options paper considers technical issues specific to extending a performance test to products in the retirement phase, noting that options to refine the retirement phase policy settings are being considered separately.

Scale of retirement products

Retirement products are offered to retired members in the drawdown phase and make up around \$400 billion in assets across 1.2 million member accounts within APRA-regulated funds with more than six members (as at 30 June 2023).¹⁸ Members over age 55 represent over half of total APRA-regulated assets and over the next 10 years, an estimated 2.5 million Australians will move from the accumulation to the retirement phase of superannuation.¹⁹

While a performance testing regime may be applicable to account-based pensions, it would be important to calibrate the test design and scope in such a way as to not discourage the innovation of retirement products that may better manage the use of retirement assets, including longevity risks. Account-based pensions are a flexible retirement income stream product, whereas a longevity

¹⁷ APRA-regulated entities with more than six members, excluding retirement products. Note that the single sector definition in regulations is different to APRA's definition where these values come from; Australian Prudential Regulation Authority, [Quarterly Superannuation Industry publication \(June 2023\)](#), Australian Prudential Regulation Authority, 2023, accessed August 2023.

¹⁸ Australian Prudential Regulation Authority, [Quarterly Superannuation Industry publication \(June 2023\)](#), Australian Prudential Regulation Authority, 2023, accessed October 2023.

¹⁹ Ibid.; Treasury modelling.

product provides the member with a guaranteed regular income for a defined number of years, or the rest of their life, such as an annuity.

Issues to consider for the retirement phase

Members who are in or are approaching the retirement phase should be similarly protected from poor fund performance as members in accumulation. However, there are a range of different risks which trustees are required to manage in the retirement phase, such as longevity and liquidity risks, which may impact performance. These need to be considered when designing a fit for purpose test to assess the performance of retirement phase products.

Longevity protection

Longevity risk protection was raised in the Retirement Income Review as important as it allows people to confidently draw down assets to fund their retirement. Management of longevity risk is a key consideration for trustees when formulating their strategy as required under the Retirement Income Covenant. Without longevity protection, concern about running out of savings contributes to retirees underspending and experiencing lower standards of living in retirement. The test would either need to ignore longevity components of products or have a suitable testing framework that takes this into account, given these products generally offer a guaranteed source of income in lieu of maintaining a reserve of assets that are invested. Furthermore, these products often do not allow members to switch once they have commenced. While the Productivity Commission proposed that the retirement phase should be tested, they noted that annuity products should be excluded from this.

Retirement products – innovation and liquidity

A testing framework for retirement phase products should seek to improve performance in the retirement phase, but not discourage innovation. If a performance test was introduced in the retirement phase, it should consider that this is still an emerging market.

Retirement products have higher liquidity and member services needs compared to accumulation products. This is because members in the retirement phase require ongoing income to be drawn from their product in addition to the processing of lump sum withdrawals which may make managing liquidity more challenging. Therefore, it is important that the material difference in fees and liquidity needs are considered when applying a performance test to retirement products.

Performance criteria

To account for the issues identified above, performance testing the retirement phase could take the following forms for different products.

- Account-based pensions: similar test to the accumulation phase test but with due consideration to unique factors such as liquidity, management of downside risk and income stability.

Annuity/longevity products: This could take the form of expanding the YourSuper comparison tool to provide a simple way for members to compare the rate of return and fees at the time of purchase. It would not assess whether a product is underperforming but instead give members a way to easily compare products. This may be sufficient to drive competition among this cohort of products.

The test would need to consider how to treat market-linked annuities and other guarantee products where the trustee offers insurance against some risks, but the member ultimately bears other risks.²⁰ For example, market-linked annuities might provide the member with longevity protection, but the member bears the market risk.²¹

Alternatively, a performance test for the retirement phase could aim to link more closely with trustee obligations under the Retirement Income Covenant. A paper by the Conexus Institute in conjunction with the Australian National University offered quantitative assessments of retirement income products against a summary scorecard.²² For a product, this reports the risk-adjusted income and a score between one and five for each Retirement Income Covenant objective.²³

Consultation Questions

29. What are the most important considerations for performance of retirement products?
30. If the test were to expand to retirement products, would they require a different test to the accumulation phase? Would the test differ for different retirement products?
31. How could longevity products be most appropriately assessed? How could the products be compared?

Other products

The intent of a performance test is to assess the performance of an investment vehicle where an agent other than the member is responsible for the asset allocation or investment decisions. Other products not captured by the test include self-managed superannuation funds (SMSFs), direct assets purchased through platforms, balance of life office statutory funds and defined benefits (all outlined in red in Chart 4).

These products are generally used by more engaged members that exercise a higher level of control over their retirement savings or have unique features which mean the test could not be administered or the consequences for failing the test would be inapplicable. For example, SMSFs are not regulated by APRA (who administers the test), data and record keeping is inconsistent with that of APRA-regulated entities, and it would be inappropriate to apply the same consequences of failure to individuals who are investing their own retirement savings.

Consultation Questions

32. Do you agree that retirement phase, single-sector and externally-managed products are suitable for testing? Why or why not?
33. Should different assessment methods be applied to different cohorts of products?
34. Do you agree that the 'other products' outlined above are unsuitable for testing? If you think the 'other products' (or a sub-section of these products) are suitable for testing, how could they be appropriately tested?
35. Under each design option, how could the test accommodate cohorts that are suitable for testing? For example, using different metrics or benchmarks for performance for different cohorts.

²⁰ Market-linked annuities are annuity products where the member receives a regular guaranteed income stream, but the payments can increase or decrease based on the performance of the investments (market-risk).

²¹ Market risk refers to the possibility of losses due to impacts of the financial market.

²² Conexus Institute, [Quantitative Assessment of RIS](#), Conexus Institute website, 2023, accessed November 2023.

²³ Ibid.

Fees

The performance test was introduced to hold trustees to account for the investment performance they deliver as well as the fees they charge to members. Whilst fees are important in compensating for the service that funds provide, they also can erode members' returns, especially over the long term, if they are too high.

How fees are tested should be considered under each of the design options discussed above, as a balance needs to be struck between creating downward pressure on fees and ensuring that members are receiving an appropriate level of member service.

Currently fees are tested through the option's representative administration fees and expenses (actual RAFE) for the most recent financial year, which is the administration fees as calculated using a representative member who has an account balance of \$50,000. The actual RAFE is compared to the median RAFE of all the products contained in the relevant peer group (benchmark RAFE). This has successfully placed pressure on funds to reduce their fees to improve their performance test result. The current test assesses fees separately for MySuper, platform TDPs and non-platform TDPs.

In 2023, while non-platform TDPs had a benchmark RAFE of 0.27 per cent, which was similar to MySuper products 0.26 per cent, there were still several products that had particularly high administration fees. Since 2021, the MySuper benchmark RAFE has decreased from 0.33 per cent to 0.26 per cent in 2023, representing an approximate 20 per cent decrease in median administration fees. This demonstrates that fees have been successfully lowered in the MySuper sector. However, it is important to ensure that members are receiving good value for money, as there is only so far fees can reduce before member services may be negatively impacted.

Stakeholders have raised various concerns with the way fees are assessed under the current test. These issues can be categorised as the: testing period, level of assessing fees and representative member balance.

Testing period

Several stakeholders suggested that the RAFE could be assessed over the entire lookback period to better reflect member outcomes and reduce the overall weighting that one year worth of administration fees plays on performance test outcomes. Others suggested that the existing one-year RAFE should be retained as it is more reflective of future fees and has successfully encouraged trustees to reduce fees.

A key issue in changing to a longer-term assessment of fees is the availability and quality of data, as APRA do not have administration fee data for choice products prior to 2021. Therefore, any changes in this space will require additional data reporting from all superannuation trustees. Another option is that as more years of fee data become available, these years are considered under the test. This would not create an additional reporting burden on trustees as trustees would not be reporting historical years of data. Under this option, a greater weight could also be given to the current year of fees.

Level of assessment

In extending the test to TDPs, some stakeholders raised concerns with testing administration fees at the investment option level, given some of the complex business and fee arrangements that occur in the choice sector. Further, some members may invest in multiple investment options within their account, a practice that is more common in the choice segment, particularly platforms. While some have advocated for administration fees to be tested at a higher level (such as the overall fund or

investment menu), others have argued it should be tested at the more granular pathway or even at the individual member account level. These are the following levels at which administration fees could conceivably be tested:

- RSE – an assessment of total administration fees collected by an RSE on an adjusted basis for either assets under management, or members.
- Superannuation product – Similar to the above but distinguished where there may be multiple products within an RSE.
- Investment menu – to account for the different administration fees members would pay based on the investment menu they access.
- Investment option – to account for the administration fees the populations of members would pay that are invested in the option. This is consistent with the current approach to assessing fees, which currently sees fees tested for 869 different MySuper products and TDPs.
- Investment pathway – to account for the unique fee arrangement members would pay to access an investment option through the relevant superannuation product and investment menu specific to them.
- Individual member basis – to account for the exact administration fees that any given member incurs. APRA does not have individual member data, meaning a significant overhaul to data reporting would be required, and the number of tests required would increase to 21.8 million to suit the number of member accounts in APRA-regulated MySuper and choice products.

In determining the right level to assess administration fees, it is important to strike a balance between being reasonably representative of the member experience and remaining practical to implement. Assessing fees at the investment option level is consistent with the current approach to fees and the broader assessment of the investment performance of options. Moving the assessment to a more granular level may better represent the member experience but this comes at the cost of exponentially increasing the number of assessments, impacting APRA’s ability to implement the test on a timely basis. It may also reduce transparency and comprehension of the test, for example where one member within an investment option fails but another member in the same option passes simply based on the investment pathways used.

Table 2: Non-platform TDP administration fee distribution at different reporting levels (\$50,000 member balance)

Level	No. of observations	Minimum Fee	Median Fee	Maximum Fee
RSE	58	0.00%	0.26%	0.88%
Superannuation Product	110	0.00%	0.28%	1.07%
Menu	90	0.00%	0.28%	1.07%
Option	454	-0.06%	0.27%	1.51%
Pathway	618	0.00%	0.28%	1.51%
Member Account	3,862,200	N/A	N/A	N/A

Note: Negative minimum fees can occur where benefits are attributed to member accounts which exceed the costs taken out, such as a rebate of administration fees paid in previous periods.

Source: Unpublished APRA data, SRF 606.0 RSE Profile, Table 4, as at June 2023.

Table 3: Platform TDP administration fee distribution at different reporting levels (\$50,000 member balance)

Level	No. of observations	Minimum Fee	Median Fee	Maximum Fee
RSE	15	0.00%	0.53%	1.00%
Superannuation Product	41	0.03%	0.70%	2.42%
Menu	46	0.03%	0.74%	1.06%
Option	280	0.01%	0.54%	1.24%
Pathway	578	0.01%	0.69%	2.71%
Member	196,246	N/A	N/A	N/A

Source: Unpublished APRA data, SRF 606.0 RSE Profile, Table 4, as at June 2023.

Representative member balance

Many stakeholders have raised concerns that the current \$50,000 representative member balance is not fit-for-purpose for particular products or members and will only become less representative over time as superannuation balances increase. The basis of the \$50,000 representative member balance was to be consistent with broader disclosure frameworks, such as product disclosure statements, product dashboards and the YourSuper comparison tool. This value was initially chosen for broader disclosure frameworks as it reflected the average MySuper balance at the time.

One option could be to amend the representative member balance used within the test to better represent members within different product cohorts, and as the market shifts over time. The APRA heatmap uses the approach of multiple representative member balances for fee assessments.

Another option would be to use the average or median balance for a relevant cohort of products. This will allow the balance to vary for different product types and shift as the market evolves over time. For example, MySuper products could have a different representative balance to retirement products.

Table 4: Benchmark administration fees (BRAFE) for different cohorts based on member balance

Product Type	BRAFE (\$50,000 representative member)	Median Member Balance (nearest \$100)	BRAFE (median balance)
MySuper	0.26%	\$76,600	0.23%
Non-platform TDP	0.27%	\$86,000	0.24%
Platform TDP	0.54%	\$79,900	0.43%

Source: Unpublished APRA data, SRF 606.0 RSE Profile, Table 4 as at June 2023.

Consultation Questions

36. How should fees be measured under each design option?
37. Should fees be measured at the current option level, or should they be measured on a different level? How would this be achieved?
38. Are the current assumptions made in comparing fees acceptable? For example, should the \$50,000 representative member balance be adjusted based on the median member balance for a product cohort?
39. Is a peer comparison of fees the best way to measure fees? Is there a better approach to benchmarking fees? If so, how should this work?
40. What product cohorts should be considered? How should different cohorts be defined where products could meet multiple cohort definitions, such as single-sector retirement products?
41. How many years of fees data is appropriate to test? Should a greater weighting be given to certain years?

Consequences

Clear consequences of failing the test have been important in making it effective and efficient. They have also been important in leading to products either exiting the market or improving their performance and will continue to play an important role in cleaning up the tail of underperformance in TDPs over the short-term.

A balance is required to ensure that the consequences are substantial enough for the test to remain meaningful, but not in a way that detracts from member outcomes. As a starting point, calibrating the test metrics correctly reduces the impact of unintended consequences.

During the YFYS Review, stakeholders argued for several changes to the consequences:

- Additional discretion for APRA in applying the consequences of the test
- More severe consequences for persistent underperformance
- Removing the consequence of closure to new members for consecutive failures for choice products
- Applying secondary tests to determine reasons for failure before applying consequences.

Given how diverse the superannuation sector remains, there may be certain edge cases that are not conducive to performance testing (such as annuity retirement phase products or niche single-sector products). Changes to the consequences in these instances could be considered so that the test is not eliminating products from the market for reasons beyond the control of the trustee.

However, introducing discretion into the consequences will impact the efficiency of the test by subjecting the regulator to challenge by trustees, resulting in delays, additional regulator costs and ultimately higher costs to members through a higher industry levy. Greater discretion could also lead to less meaningful improvement by trustees, as they shift focus from improving performance to 'justifying' their failure.

Clear consequences have been effective in eliminating underperformance; however, many stakeholders have advocated for the test to also encourage funds to improve across the board and not only punish underperformers.

Consultation Questions

42. Should the consequences be adjusted to improve outcomes for members? How would this need to be tailored for the different options for performance testing?
43. How should the consequences be amended to better account for edge cases or different cohorts that fail the test for reasons beyond the trustee's control?
44. How could these provisions be effectively ring-fenced so that it applies only to the edge cases and not failures at large?
45. How could this be achieved without subjecting the regulator to undue challenge and impacting the efficiency of the regime?
46. What other remediation processes could occur?

Barriers to consolidation

When a product fails the test for a second consecutive time, the product is closed to new members until it passes a subsequent performance test. In MySuper, this consequence has proven to be an extremely effective tool in encouraging underperforming products to exit the market. However, given the performance test only first applied to TDPs in 2023, we are yet to see the impact of these consequences on the consolidation of underperforming products in the choice sector. It is also noted that there are already many products closed to new members in the choice sector, and therefore this consequence may have limited impact where these products fail the test on consecutive occasions.

Stakeholders have raised that there are barriers to consolidating these closed and underperforming products. Trustees have a clear obligation under the law to act in the best financial interests of their members, including taking action to improve performance of an investment option, or taking steps to move members into better performing products. Products that fail the performance test will continue to be subject to heightened supervision from APRA.

Consultation Questions

47. Are there any key barriers to consolidating closed and underperforming products? What quantitative evidence is there of these barriers? How do these weigh against other reasons a person may choose to remain in a product?
48. What evidence do trustees use to demonstrate that remaining in a closed and underperforming product is in the best financial interests of members, compared to moving to a performing product?
49. What is the process or criteria that trustees use when deciding on what product they will transfer members to when consolidating underperforming products?
50. Should APRA receive increased regulatory powers to direct superannuation trustees to consolidate underperforming products?

Appendix

Glossary

Account-based pension (account-based income stream or allocated pension): An account-based pension is a flexible retirement income stream product. Retirees who are in an account-based pension are required to draw a minimum pension payment amount each year or elect to draw an amount of pension payment above the required minimum amount. They can withdraw some or all the account balance and the income is tax free from age 60.

Annuity (lifetime income product or a fixed-term pension): An annuity provides the member with a guaranteed regular income for a number of years (fixed-term annuity) or for the rest of their life (lifetime annuity). Common features of annuities include guaranteed regular payments regardless of how the share market performs, indexation of payments in line with inflation, reduced flexibility to move capital held in the annuity or to opt out and/or a guaranteed death benefit to dependants. Annuity providers bear both longevity and investment risks while the member can be sure of their future income.

Benchmark: what a product's performance is measured against, for example comparing an asset class to an index.

Externally-managed product: multi-asset products where the trustee or a connected entity does not have control over the design or implementation of the investment strategy of the option.

Failing rate: the level that indicates a product has failed a test.

Metric: a method of assessing the performance of a product.

Non-platform TDP: a TDP that is offered through only one investment menu.

Performance indicator: an aspect of the product that demonstrates whether it is successful or not.

Platform TDP: a TDP that is offered through one or more investment menus of the platform type.

Test: a way to assess the performance of a product in regard to a performance indicator using one or more metrics and variables.

Threshold: value at which performance is benchmarked against. This value will determine whether a product passes the test.

Trustee-directed product (TDPs): multi-asset products where the trustee or a connected entity has control over the design of the investment strategy of the option.

Variable: an input to a metric, such as a lookback period or representative member balance.

Current performance testing benchmarks

Covered asset classes, assumed indices, assumed annual fees and assumed tax rates for quarters starting on or after 1 July 2014

Item	Description	Assumed index	Assumed annual fee	Assumed annual rate of tax
1	Australian Equity	ASA52 S&P/ASX 300 Total Return Index	0.05%	0.00%
2	International Equity (hedged; international economy type not specified or not applicable)	DE725341 MSCI All Country World Ex-Australia Equities Index with Special Tax (100% hedged to AUD)	0.11%	14.00%
3	International Equity (hedged; emerging markets)	DA725342 MSCI Emerging Markets with Special Tax (100% hedged to AUD)	0.16%	14.00%
4	International Equity (hedged; developed markets)	DA750700 MSCI World ex Australia with Special Tax (100% hedged to AUD)	0.10%	14.00%
5	International Equity (unhedged; international economy type not specified or not applicable)	DN714533 MSCI All Country World Ex-Australia Equities Index with Special Tax (unhedged in AUD)	0.09%	14.00%
6	International Equity (unhedged; emerging markets)	NA714531 MSCI Emerging Markets with Special Tax (unhedged in AUD)	0.14%	14.00%
7	International Equity (unhedged; developed markets)	NA714532 MSCI World ex Australia with Special Tax (unhedged in AUD)	0.08%	14.00%
8	Australian Listed Property	ASA6PROP S&P/ASX 300 A-REIT Total Return Index	0.12%	14.00%
9	International Listed Property	RAHNSAH FTSE EPRA Nareit Developed ex Aus Rental 100% Hedged to AUD Net Tax (Super) Index	0.22%	14.00%
10	Australian Listed Infrastructure	FDCICSAH FTSE Developed Core Infrastructure 50/50 100% Hedged to AUD Net Tax (Super) Index	0.26%	14.00%

Covered asset classes, assumed indices, assumed annual fees and assumed tax rates for quarters starting on or after 1 July 2014

Item	Description	Assumed index	Assumed annual fee	Assumed annual rate of tax
11	International Listed Infrastructure	FDCICSAH FTSE Developed Core Infrastructure 50/50 100% Hedged to AUD Net Tax (Super) Index	0.26%	14.00%
12	Australian Unlisted Property	MSCI/Mercer Australia Core Wholesale Monthly Property Fund Index – NAV-Weighted Post-Fee Total Return (All Funds)	0%	14.00%
13	International Unlisted Property	MSCI Global (Excl. Pan-Europe and Pan-Asia Funds) Quarterly Property Fund Index (Unfrozen) (Net Total Return; AUD fixed)	0%	14.00%
14	Australian Unlisted Infrastructure	MSCI Australia Quarterly Private Infrastructure Fund Index (Unfrozen) – 50th Percentile Post-Fee Total Return (All Funds)	0%	14.00%
15	International Unlisted Infrastructure	MSCI Australia Quarterly Private Infrastructure Fund Index (Unfrozen) – 50th Percentile Post-Fee Total Return (All Funds)	0%	14.00%
16	Australian Fixed Income	BACMO Bloomberg Ausbond Composite 0+ Yr Index	0.10%	15.00%
17	Australian Fixed Income Excluding Credit	BAGVO Bloomberg Ausbond Govt 0+ Yr Index	0.08%	15.00%
18	Australian Credit	BACRO Bloomberg Ausbond Credit 0+ Yr Index	0.15%	15.00%
19	International Fixed Income	LEGATRAH Bloomberg Global Aggregate Index (hedged AUD)	0.10%	15.00%
20	International Fixed Income Excluding Credit	BTSYTRAH Bloomberg Global Treasury Index (hedged AUD)	0.08%	15.00%
21	International Credit	LGCPTRAH Bloomberg Global Aggregate Corporate Index (hedged AUD)	0.15%	15.00%
22	Australian Cash	BAUBIL Bloomberg Ausbond Bank Bill Index	0.04%	15.00%
23	International Cash	BAUBIL	0.04%	15.00%

Covered asset classes, assumed indices, assumed annual fees and assumed tax rates for quarters starting on or after 1 July 2014

Item	Description	Assumed index	Assumed annual fee	Assumed annual rate of tax
		Bloomberg Ausbond Bank Bill Index		
24	Alternatives		0%	0%
25	Defensive Alternatives		0%	0%
26	Growth Alternatives		0%	0%

Performance testing metrics

Stakeholders have proposed many performance testing metrics, particularly during the Your Future, Your Super Review. Below is a list of proposed performance testing metrics with a brief explanation:

Metric	Explanation
Simple reference portfolio (SRP)	A notional portfolio of passive, low cost and liquid investments. An SRP portfolio is usually comprised of a small number of benchmarks, usually equities and bonds, based on the product's exposure to growth and defensive assets.
Risk SRP	A notional SRP is derived based on the amount of risk (volatility) that a product experiences. With the same level of risk, the test assesses the return that the SRP generates compared to the product.
Strategic Asset Allocation (SAA) benchmark	This would measure the product against a benchmark strategic asset allocation which would assess the implementation of the product's investment strategy. In contrast to the SRP, an SAA benchmark has a more expanded list of benchmark indices.
Naïve reference portfolio	An alternative to a simple reference portfolio. An example could be creating a portfolio of a 70 per cent/30 per cent split between growth and defensive assets using index exposures. A product's performance would then be assessed against this portfolio.
Sharpe ratio	A measure of the risk-adjusted performance. This metric assesses the return a product generates for the volatility (risk) that is endured.
Volatility	This is assessed through the standard deviation of a product's returns. Could use this in either a peer comparison of volatility or have a lower threshold which products must reach.
CPI+	Compare the return of the product to CPI. This would determine whether the product is returning a real return. This can align with the investment objectives that funds must disclose to members.
Heatmap metrics	Use the measures already calculated in the heatmap (SRP, SAA, net return, etc.) to assess a product's performance in the test.
Use multiple time periods	Assess metrics across multiple time periods.
Use multiple member balances	Assess metrics across multiple representative member balances.
Net returns/actual returns	Peer comparison of net returns or have a threshold that products must pass. This measure incorporates fees and returns and is the actual return that members receive.
Real costs	Measure of the real costs that members face in the product compared against peers.

Types of services offered by the fund	Measure of the number and performance of services offered by the fund.
Peer comparison	Compare the performance of a product against its peers based on one or a variety of metrics. Failure can occur to the bottom cohort of products, or those below prescribed thresholds depending on the metric.
	Compare the actual return of the product against a benchmark portfolio. Adjust results for risk and fail the bottom 10 per cent of products, or those below a prescribed threshold.
	Plot the returns from a product versus their volatility. Calculate a trendline based off the industry and then if a product falls below a threshold buffer amount below the trend line they fail the test.
Dashboard of risk metrics	Multiple risk metrics could be used to combat members being concerned about different types of risk and having differing tolerances for those risks.
Individualised member test	Apply investment option returns to members' own projected retirement balances. Collate into cohorts and report on projected retirement balances.