Review of Your Future, Your Super Measures

October 2022

Introduction

Frontier is pleased to provide this submission to Treasury's consultation on the Review of the Your Future, Your Super (YFYS) measures.

Frontier is one of Australia's leading institutional investment advisers. We have been advising Australian institutional investors as a trusted adviser for over twenty-five years. We provide advice on more than $600 billion of assets across the superannuation, charity, public sector, insurance and higher education sectors. The fact our advice is free of any product, manager or broker conflicts, means we can provide truly unconflicted advice aligned with our client's best interests.

The YFYS performance test policy aims are laudable:

- ensuring that members do not join persistently underperforming products
- providing members with information to choose better performing products
- encouraging underperforming products to improve their performance, including through potential mergers with higher-performing funds.

Frontier supports a well-designed performance test.

We provided submissions to both the initial consultation and the subsequent regulations consultation. In these submissions (available here and here), we provided commentary on where we believe the YFYS reform package could be enhanced to better achieve the stated policy aims. We have also written a number of research papers on the topic – The heat is on underperformance, No risk, no return, Learn to love the benchmark, and Super performance test.

In this submission, we reiterate our main concerns with the current performance test:

- It is a simple test based on past performance, which is a "weak and unreliable" predictor of future performance.
- It only assesses a small part of member outcomes, namely how well a fund has implemented its chosen strategy, not whether it is a good strategy. Importantly it does not incorporate a risk-adjustment outside of the SAA, such that funds can underperform the test if they reduce risk.
- It introduces a new regulatory risk, with real consequences for underperformers, resulting in the likelihood funds will adjust their investment portfolios to not fail the test in the short term. Given some strategies take time to pay off or require a full investment cycle, this means decisions can be made to the detriment of members' long term outcomes as funds would be moving away from well position strategies to meet their long-term objectives.

As an asset consultant, our submission concentrates on the performance test and to its interaction with the YourSuper comparison tool.

The current problem

Given superannuation plays an important role in supporting people in their retirement, Frontier supports legislation which helps ensure consumers are in good funds.

However, determining whether a superannuation fund is a "quality" fund (likely to perform well in the *future*) is complex. We firmly believe a purely quantitative approach is insufficient – our assessment of investment managers uses a substantially qualitative approach, where past performance forms a very small part of the assessment but is a useful tool in identifying areas for further investigation.

Arguably, such legislation already exists, particularly under SPS 515 which requires trustees to "regularly assess the outcomes provided to members and identify opportunities for improving these outcome".

In Frontier's view, any "bright-line" assessment based on a simple past performance test is problematic. There is a large body of academic research (including from ASIC) which highlights that determining future investment outcomes from past performance is difficult.

Additionally, as superannuation is a multi-decade investment, it is important that any test supports good long-term member outcomes. Funds which take a long-term approach can benefit from access to a broader investment opportunity set. If a performance test restricts the ability of funds to act as long-term investors, it is likely to be detrimental.

As the performance test represents a new regulatory risk, with significant consequences for underperformers, there is evidence that funds have adjusted their investment portfolios away from strategies to meet their long-term objectives in order to manage risk taking versus the test over a shorter time horizon. This could be to the detriment of members' long-term outcomes.

However, we agree it is important for member engagement that consumers can engage with the test results. This favours a simpler test.

This paradox of providing members with a simple comparison of a complex issue is vexing. The following options are possible:

- Elect for a simple solution, ignoring the complexity.
- Elect for a complex solution and attempt to present it simply.
- Elect for a two-tier solution – provide a simple solution to members and complex solution for regulators.

The danger with the first option is poor funds may slip through the test, and members receive poor outcomes in the future. Conversely, a good fund may fail the test and members miss good future outcomes. In addition, a simple test introduces real and significant unintended consequences – funds may consider actively managing to the performance test rather than concentrating on long-term member outcomes.

The second option, a complex solution, has the benefit that it is more likely to identify poor funds. It is also less likely to unduly influence fund behaviour, as it is more difficult for funds to manage to a multi-faceted test. However, it faces the danger consumers do not engage with the complexity. As seen with the APRA heatmap, it is difficult to present multiple metrics in an easy-to-understand format.

The YFYS test is a variation of the first and third options – it has a simple comparison tool for members and a slightly more complex test. However, it satisfies neither criterion. It neither passes the "pub test" of matching with member outcomes nor is it a sophisticated assessment without unintended consequences.

We outline in Appendix more detail on our, and ASIC's, concerns about the efficacy of past performance to assess future performance. We also provide a summary of the Conexus Institute's Review of the YFYS Performance Test, which echoes our concerns with the current test.

Frontier's recommendation

The implementation of the performance test and the YourSuper comparison tool could be enhanced by making them operate more coherently. As it stands, the comparison tool can show a fund ranking highly based on its actual returns but fail the annual performance test (based on performance relative to a benchmark). The reverse can also occur – poor performance but meets the performance test. This would lead to consumer confusion and disengagement – the opposite of the desired outcome.

As outlined in our initial submission, we believe the performance test and the YourSuper comparison tool can better meet the policy aims with some minor enhancements. We base our recommendation on the following principles:

1. Directly link the performance test to actual member outcomes. This will also ensure the performance test and the comparison tool are aligned, so consumers do not receive conflicting information.

2. As a test with significant consequence of failure, we believe the performance test should be a comprehensive assessment. A test made up of a well-designed collection of multiple metrics is likely to be more effective and reliable than a single metric based on a narrow assessment of quality.

3. In line with the PC recommendation, we believe APRA should play a key role in working with funds which fail the performance test to improve or exit. This should provide more timely protection for consumers rather than relying on member engagement.

The YFYS legislation is trying to achieve two goals with the performance test and comparison tool:

- Empower members – providing simple and clear information to help them choose a high-performing and/or low-cost fund.

- Holding funds to account for underperformance – poor funds have no place in a compulsory system and members should be protected from persistently underperforming products.

We believe an integrated two-tier solution which separately addresses these goals should be adopted:

Test 1: Empower members – funds' net investment performance is ranked over 10 years compared to relevant peers in the YourSuper comparison tool. Funds which underperform the median by a prescribed threshold are then subject the second detailed test. Alternatively, the simple reference portfolio (as defined in the APRA heatmaps) would be an acceptable alternative.

Test 2: Holding funds to account for underperformance – this annual test assesses investment performance and fees using multiple metrics, and the regulator's prudential investigations, to better assess whether a fund has the "right to remain". If the fund fails this assessment, it would be subject to a 12-month period of remediation, or if remediation is not possible, the regulator would ensure an orderly exit for that fund. Funds which fail test 1 but pass test 2 would receive a single overall assessment (for instance as an "amber" rating), communicated to members on the YFYS comparison tool.

This is similar to the current approach with some important differences:

- The initial performance test is clearly linked to member outcomes and there is an explicit link between the YourSuper comparison tool and the performance test. When a member finds their fund near the bottom of the comparison tool, they will be confident the regulator is investigating. Those funds will need to justify their continued operation.

- The second test is more complex. No longer needing to drive member engagement, this test can reflect the complexity of making an assessment. For example, it could allow for different time periods, measures of risk and ESG considerations. Based on this second test, the regulator can assess whether the fund remains “on-watch” until the performance improves or instigates the fund/product’s orderly exit. This is the “right to remain” process the Productivity Commission recommended.

	Test 1	Test 2
Goal	Empower members – increasing competition and improving outcomes	Holding funds to account for underperformance
Focus	Members	Regulator
Test	Quarterly peer ranking of net investment performance (or simple reference portfolio)	Annual multiple metrics of investment performance, risk, fees, and sustainability
Failure	Member engagement + Test 2	Orderly exit

Alternatively, if member engagement with the test design was considered less important, Frontier would support a single test in line with test 2 above.

We provide further details on performance metrics which could be used as part of test 1 and 2 in Appendix B.

Consultation questions

Performance test

Test methodology

1. **Does the measurement of actual return using strategic asset allocation affect risk taking behaviour by superannuation trustees?**

Yes. It introduces a new regulatory risk that only takes into account risk taking embedded in SAA settings, with significant consequences of failure. As such all funds need to assess their risk relative the performance test.

We know the superannuation clients we advise have all assessed the YFYS risks they are taking. They have evaluated the expected additional returns from taking on these risks and, particularly those which are closer to failing the test, have considered how confident it will pay-off over the short term.

The public quote from Rest's CIO Andrew Lill[1] encapsulates the decision process most funds will have undertaken:

"You need to be aware that any decision you take to move an asset allocation or level away from your SAA is effectively at both an opportunity and a risk, and introduces tracking error to the portfolio compared to the features of the benchmark."

"But we do have a lower tolerance for total portfolio tracking error, so we have to decide which of the investment views [should] be added at an asset allocation level, as a portfolio tilt or at the active stock selection level," he says. "You need to be very cognizant of your overall portfolio and where your biggest active risk opportunities are."

In particular, funds have reassessed risk-reducing (rather than return-enhancing) strategies that are made separately to the high level SAA settings using APRA defined asset classes (these may be within asset class risk settings, or new or sub-asset classes not defined by APRA). Because the test does not recognise the benefit of risk reduction, funds will now be less likely to take risk reducing position away from their YFYS benchmark. This is particularly the case for strategies such as equity option protection which are not recognised by the test.

In addition, funds may be less willing to invest in new asset classes which are not covered by the performance test due to the additional regulatory risk even where it is in members best interest. This may stifle innovation.

2. **Does the current set of indices used to calculate benchmark returns unintentionally distort investment decisions or reduce choice for members? If so, is there a way to adjust the benchmark indices while maintaining a clear and objective performance test?**

The current set of indices defines the least-risk position for funds, any decisions away from these indices increases a fund's YFYS risk, and effectively sets a benchmark risk level for each asset classes. Importantly, this includes some asset classes where the risk positioning and exposures can vary substantially depending on the preference of the investor and the role it has defined that asset class to take in the overall portfolio. As a result, it can constrain the asset classes and investments in which funds will invest, even when they may produce better long-term outcomes.

In addition, it can constrain funds from investing in line with their ESG and net-zero principles. While funds typically expect ESG criteria will benefit members, the timeframe for their payoff is longer than the eight years used in the YFYS test.

We endorse the industry view that the time period for the test should be extended to ten years.

[1] Investment Magazine, July 2021. Frontier is not an adviser to Rest.

Frontier would rather see the test redeveloped such that the choice of indices becomes less important, rather than an extension of the current list of indices.

3. Does the calculation of actual RAFE and benchmark RAFE discourage non performance related product features that members may value (such as customer service or platform products)? If so, can this be addressed without diminishing the test's focus on performance?

The impact of changes in administration fees on test results are proportionally higher than improvements in funds' performance which has a disconnect with actual member outcomes. A ten basis point reduction in administration fees, a ten basis point reduction in investment fees and a ten basis point increase in fund performance all have an equal impact on member outcomes. However, the administration fee impact improves test results by ten basis points (as current year administration fees are used for the full eight year test period) while the investment fee reduction and the performance increase are diluted over the eight years.

We are also concerned that funds can arbitrage the test by reducing administration fees and increasing investment fees resulting no change in member outcomes but an improvement in the test result.

4. What are the longer-term impacts of the performance test on market dynamics and composition? How will these factors impact on long-term member outcomes?

The YFYS performance test has been a success if the desired outcome was to reduce the number of superannuation funds. However, it negatively impacted member outcomes for many members who choose to switch away from those funds. A number of funds which failed the initial test produced market leading investment returns in the subsequent year.

In the future, with the test now built into funds investment objectives, we expect fewer funds to fail. However, because the test is now an additional constraint on fund's investment strategies, we do not expect this to improve long-term member outcomes and will reduce outcomes if funds respond to the test by investing more short-term to limit the risk of underperforming the test in future years.

Consequences of failure

5. Is there evidence to indicate that the notification and website publication requirements have been effective at encouraging members to consider, and switch to, alternative products? Are there ways this could be improved?

The notification process depends on member engagement, which is acknowledged to be lower than preferred. Where a poor quality fund has been identified, engaged members will leave. This can lead to even poorer outcomes for the remaining members as the fund needs to liquidate investments to meet the cash outflow.

We believe the regulator should play a more central role in the process rather than relying on member engagement.

6. Have the consequences been effective at encouraging trustees to improve their performance or merge with better performing funds? Are there ways this could be improved?

The 2021 performance test results have clearly encouraged fund mergers, with 10 of the 13 failed funds either merged or are working through merger discussions. However, we are concerned a negative test result will hinder mergers in the future. Good performing funds may become reluctant to merger with poor funds; it will certainly increase the level of due diligence the accepting fund will need to undertake.

We believe the regulator should play a more central role in the process.

7. Are the measures in place to resolve underperformance sufficient given the potential for members to be stapled to these products? How can the system best support members in underperforming products?

We believe members should not be stapled to funds which do not pass a (more rigorous) performance test.

Product coverage

8. Are there any significant issues to be expected when the test is extended to TDPs? If so, how could these issues be addressed?

9. What would be the impact of extending the current performance test to other Choice products (such as single sector or retirement products)? How could any issues be addressed?

The Productivity Commission report found the default (MySuper) segment outperforms the system on average. Therefore it is more important that the choice segment is subject to a performance assessment including both returns and fees.

Perversely, the problem identified by the Productivity Commission (proliferation of little used and complex products) has resulted in problems with implementing the test for choice products.

We acknowledge that APRA is undertaking a superannuation data transformation program to improve their ability to regulate these products. It is unfortunate that this data collection program was not sufficiently coordinated with the YFYS test such that there are asset classes for which no YFYS benchmark has been determined.

If the current test is implemented for non-TDF products, it may require a proliferation of benchmark indices to reflect the complex products available. As an interim/alternative solution, APRA could rely on the product providers to assess these products against the pre-existing product benchmarks and providing results to APRA.

Appendix A

Past performance

ASIC Report 22 – A review of the research on the past performance of managed funds undertakes an extensive review of the academic literature on the "persistence" of managed fund performance[2]. The academic studies look at whether funds' past performance is related to their future performance.

This report reaches a number of relevant conclusions:

- Performance comparisons can be quite misleading if not done properly.
- Returns are only meaningful if adjusted for risk/volatility or comparing "like with like".
- Good past performance seems to be, at best, a weak and unreliable predictor of future good performance over the medium to long term. About half the studies found no correlation at all between good past and good future performance.
- Where persistence was found, this was more frequently in the shorter-term, (one to two years) than in the longer term. The longer-term comparison may be more relevant to the typical periods over which consumers hold managed funds.
- More studies seem to find that bad past performance increased the probability of future bad performance.
- Where persistence was found, the "out-performance" margin tended to be small. Where studies found persistence, some specifically reported that frequent swapping to best performing funds would not be an effective strategy, due to the cost of swapping.

Furthermore, in Regulatory Guide 53, ASIC warns “it may be **misleading** to imply that reliance on simple past performance figures would be a good way to select a financial product or service.

- Promotions have a higher risk of being misleading if they focus on past performance as a sole or dominant method of selecting a financial product or service.
- The issue here is an implication about the significance of simple past performance figures, without any reference to how the returns were achieved or their relevance to future circumstances.
- If a promotion implies that some aspect of past performance should be the sole or dominant method of selecting a financial product or service, the promoter should have evidence to substantiate the implication.”

We believe ASIC’s concerns are valid.

[2] As superannuation fund options are a subset of managed funds, we believe similar conclusions would be drawn from analysing superannuation fund performance.

Performance test efficacy

A similar view was detailed by the Conexus Institute 2020 YFYS Working Group (of which Frontier was a contributor). “The statistical effectiveness of the YFYS performance metric at identifying poor performing funds is found to be weak. The YFYS performance metric faces three major challenges:

1. timeframe (8 years may not be sufficient),
2. it focuses on one (likely minor) component of performance (implementation) rather than investment performance in total, and
3. the benchmarking process generates inaccuracies which create benchmark ‘noise’.”

The report’s conclusion noted “there is merit in protecting consumers with a performance test. However, it needs to be an effective performance test with limited undesirable outcomes. Unfortunately, our analysis suggests the YFYS performance test does not meet these goals: we have strong concerns that the YFYS performance test will be ineffective at identifying poor performing funds while introducing a range of undesirable outcomes. We are concerned that the detriments of the YFYS performance test could outweigh the benefits.

Our concerns relate to:

- what the metric actually assesses (a small component of total investment performance),
- the limited ability of backwards-looking performance measures to predict future performance,
- the statistically weak effectiveness of the YFYS metric for distinguishing between ‘good’ and ‘bad’ funds, and
- the anticipated impacts on fund behaviour, consumer outcomes and industry structure.

In summary we have reservations about the effectiveness of the YFYS performance test and believe it risks significant undesirable outcomes.”

Appendix B – Investment Metrics

Assessing investment performance

When evaluating investment performance, each fund option has its own unique membership, return objectives and risk profile, which will influence its investment strategy and, ultimately, the investment returns that are achieved for members. The comparison of performance across products should consider the investment objectives and level of risk for each product to enable a like-for-like assessment of performance.

There are of different metrics which can be used to assess performance:

- performance versus fund objectives
- ranking of performance versus peers
- simple reference portfolio
- implementation test
- risk-adjusted performance.

We provide some brief pros/cons of each metric on the following page.

In addition, the time period over which the performance assessed is important:

- a longer period, such as ten years or more, provides more statistical significance that the results are not just done to randomness (aka luck). Ten years also matches with the product dashboard requirements.
- a shorter period, such as three years, provides relevancy especially as investment processes change over time. A fund with good recent returns but a poorer longer term track record is arguably better than the reverse situation.

A well-designed collection of metrics measured across multiple time periods and measures of investment risk, all else equal, is superior to an individual metric. Any individual metric will have shortcomings, and these can be reduced through the judicious use of additional metrics.

Frontier's recommendation would be to use a test based on multiple metrics and multiple time periods, similar to the APRA heatmaps. However, we would suggest an overall result be determined from these metrics – this could be as simple as a "pass" is if more than half of the metrics individual show an above threshold result.

Frontier would welcome the opportunity to participate in an industry working group to agree this test.

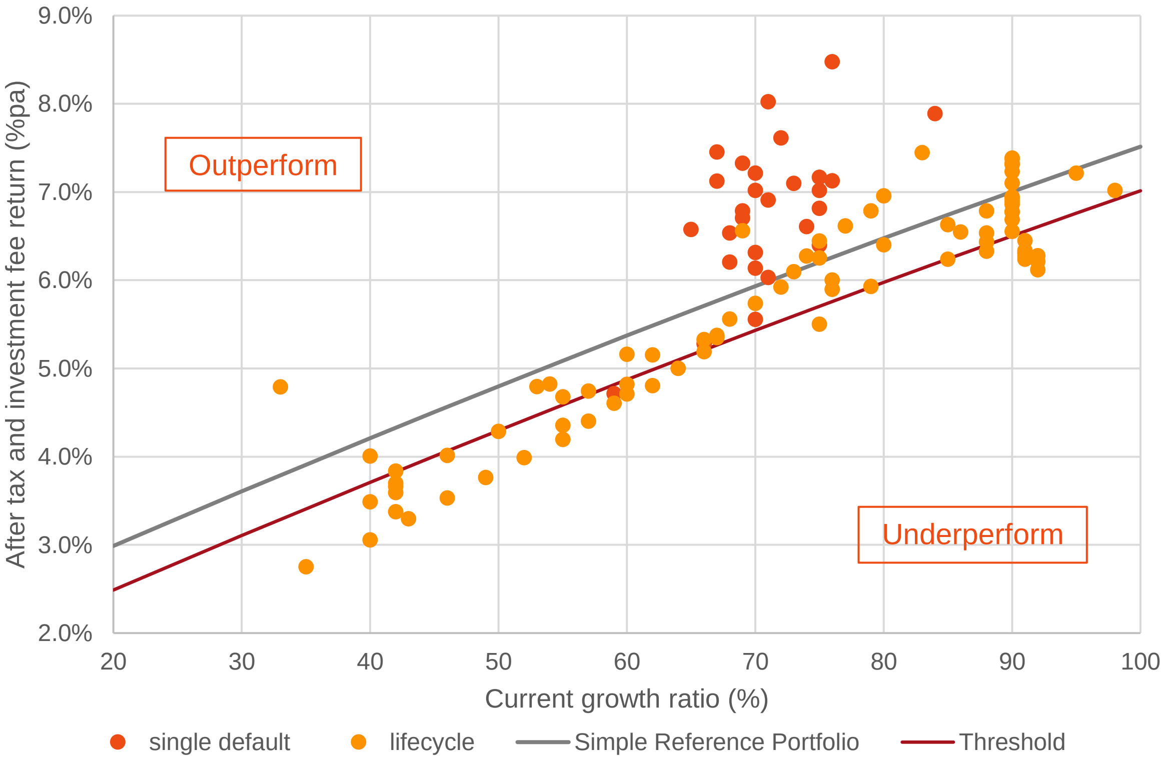
Performance metrics

Metric	Fund objectives
Calculation	Return versus investment objective (e.g. CPI + x%)
Period	Option dependent
Pros	Most logical assessment, as it is the actual target funds are aiming to meet. Required by SPS530 and Product Dashboard. Used by the Future Fund.
Cons	Has less efficacy in differentiating between ‘good’ and ‘bad’ funds. In the shorter term, whether a fund meets its objective will largely be determined by the investment markets rather than the skill of the fund.
Sophistication	Low

Metric	Actual returns
Calculation	Actual returns, unadjusted for risk. A specified proportion of the lowest performing funds would ‘fail’ (e.g. lowest 10%).
Period	Multiple
Pros	Aligns with actual member outcomes and easily understood by consumers. Most funds track their performance against their peers. One of the metrics in the APRA Heatmap.
Cons	Need to compare with “peer” funds. For this reason, the survey providers categorise funds into ‘universes’ based on each fund’s exposure to growth assets (as a proxy for risk). Promotes risk taking.
Sophistication	Low

Metric	Simple reference portfolio
Calculation	Comparison of actual returns with a naïve benchmark of equity, bond, and cash allocations. Benchmark weights can be the same for all funds (e.g. 65/35) or based on the risk level of each fund. If the same weights are used for all funds, then the rankings will be the same as using the actual return
Period	Typically 10+ years
Pros	Measures fund's value-add over a passive benchmark. One of the metrics in the APRA Heatmap.
Cons	Need to measure return-enhancements and risk-reduction. Promotes risk-taking if single benchmark chosen. Otherwise need to determine basis on which to determine fund-specific weights. Reliance on a simple growth/defensive classification to proxy investment risk which can also promote risk taking.
Sophistication	Moderate

Simple Reference Portfolio – MySuper funds 8 years to June 2022

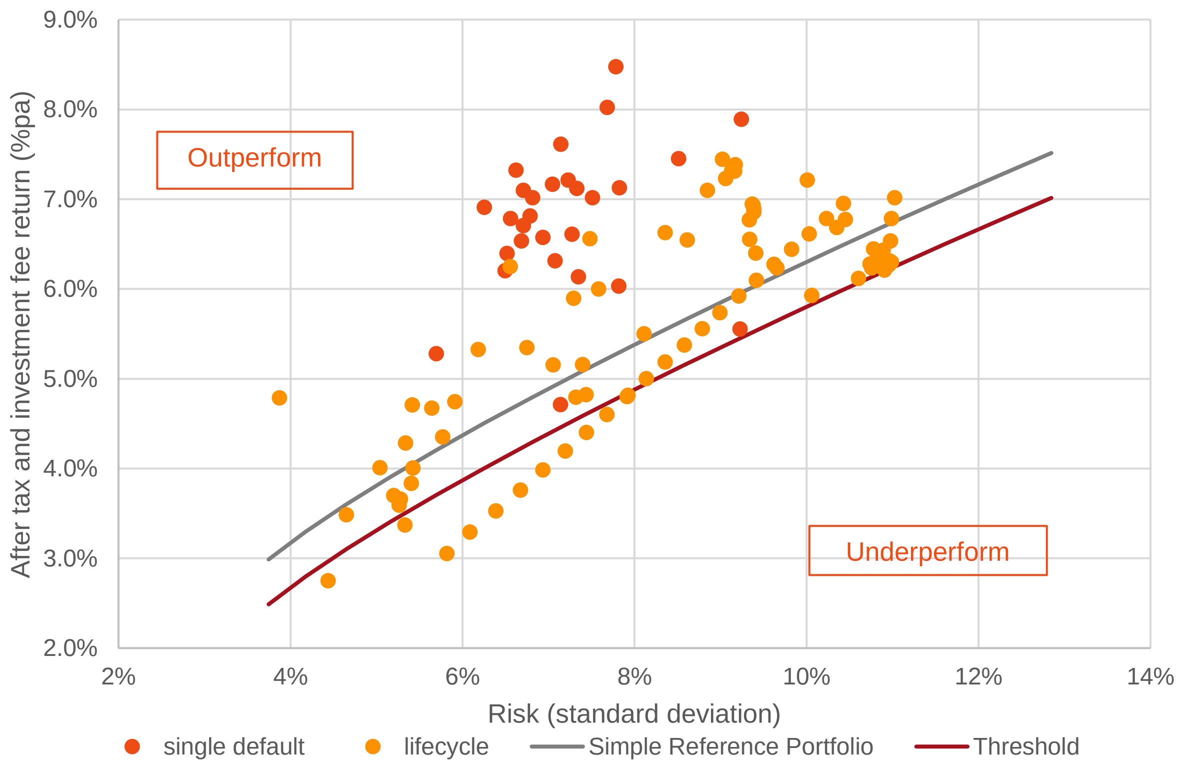


Source: Frontier, SuperRatings data

Metric	Implementation benchmark
Calculation	Comparison of actual returns with a benchmark which reflects each fund’s SAA. The current performance test approach.
Period	8 years
Pros	Measures a fund’s implementation value-add. The current performance test approach.
Cons	Doesn’t assess the effectiveness of the SAA.
Sophistication	Moderate

Metric	Risk-adjusted performance
Calculation	Ratio of actual returns and volatility, e.g. Sharpe ratio
Period	Typically 3, 5 and 10 years
Pros	Measures a fund’s risk-adjusted performance. Accounts for both return enhancement and risk reduction, including in more difficult to benchmark asset classes that are challenged by the current performance test approach. Assesses the value of SAA and the total portfolio outcome in risk-adjusted terms.
Cons	Volatility of returns understated for unlisted and liquid alternative assets. Typically calculated with monthly (rather than quarterly) returns.
Sophistication	Higher

Risk/Return – MySuper funds 8 years to June 2022



Source: Frontier, SuperRatings data