Explainer

Simple reference portfolio (SRP) performance test as proposed by Treasury

Working version at 4 June 2026

David Bell and Geoff Warren

'Option 2' for reform of the Your-Future-Your-Super (YFYS) performance YFYS test in the recent Treasury consultation paper is to replace the existing test with a test based on performance versus an SRP. We explain what an SRP test entails, outline various considerations for super funds in managing under the test and highlight the design and implementation issues. Appendix A presents a simple model of the impact of holding mid-risk assets, while Appendix B overviews SRP metric currently produced by APRA.

How an SRP metric works

The SRP metric as envisaged involves tracing out a risk-return frontier of portfolios comprising various mixes of listed equities and fixed income against which fund performance is compared. It is envisaged that risk would be measured by the standard deviation (SD) of fund returns (although this is open for debate). Essentially the SRP test compares fund (net) returns against portfolio of listed assets that members could notionally access themselves passively at low cost of the same SD.

The chart below, taken from Treasury's 2024 consultation paper on design options for the YFYS test, illustrates the basic workings. The blue line represents the SRP frontier, which is constructed by varying the equity and fixed income weights. The frontier slopes upwards as both return and SD increase as the equity weight increases. The test failure threshold is set at a certain percentage below the frontier and is represented by the red line. Each dot represents the position of a fund based on its return and SD.

MySuper Return vs Risk (8-year)

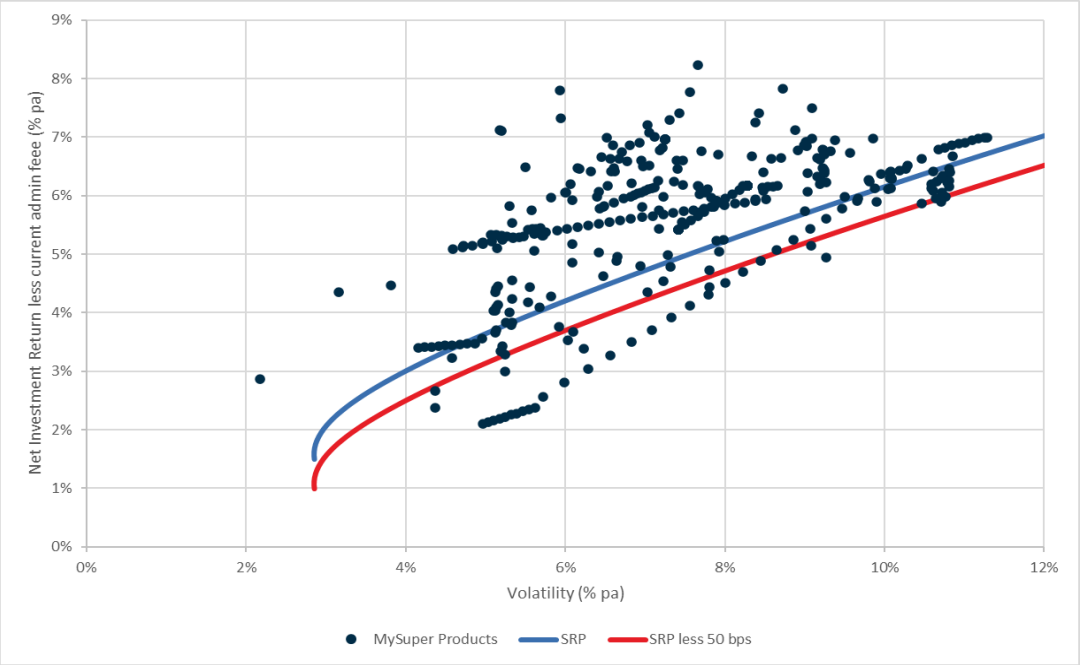


Chart sourced from *Annual Superannuation Performance Test - design options: Consultation paper*, Treasury, March 2024 (page 22). Data sourced from Chant West.

Test tracking error

A key consideration for super funds is test 'tracking error' (TE), which reflects the potential for performance to deviate from the SRP return benchmark. TE matters as it provides a guide to the probability of test failure.

Under the existing YFYS test, TE largely arises from deviations within asset classes relative to the asset indices. There can also be impacts from differences in FX hedging and interim deviations from declared SAA arising from tactical positions and asset rebalancing. There is also no risk adjustment under the existing test.

Under an SRP test, TE would reflect two components:

(a) Deviations in the asset class weights versus the SRP (reflecting strategic asset allocation, or SAA). Any asset classes outside of those in the SRP, most notably alternative assets to equities and fixed income should raise TE.

(b) Deviations from the SRP indices within the asset classes, i.e. equites and fixed income. This is similar to the existing test but will be narrower in scope as the asset indices outside of equities and fixed income disappear from the mix.

Deviations associated with (a) are likely to swamp those from (b) in determining to TE. Funds will thus be incentivised to pay closer attention to SAA weights relative to the SRP than the make-up of their equity and fixed income portfolios, although the latter remains relevant.

Considerations for super funds

A fund might consider four elements in varying their portfolios from the SRP benchmark and hence how much TE to accept:

- *Expected return on off-benchmark asset classes* – Returns are important but quite uncertain, which may reduce the extent to which they are given credence where a fund wishes to limited TE.
- *Impact on SD* – Diversifying into off-benchmark assets may reduce portfolio SD depending on SD is measured especially how private assets are treated, i.e. how much risk reduction yielded by return smoothing. Sliding the fund down the SRP frontier should lower the benchmark return required to pass the test, although the strength of this effect will depend on the slope of the SRP frontier[1].
- *Failure threshold* – Set at 50 basis points under the existing test but could (and should) be reviewed under the SRP test.
- *Buffer* – This is the distance between fund returns and the failure threshold. Lesser buffer will tend to raise the probability of test failure, which can create an incentive for a fund to converge their portfolio towards on the SRP. Here funds may consider what returns will remain or drop out of the 10-year history looking forward.

Appendix A presents some modelling of the impact of a fund holding mid-risk assets such as infrastructure and property. Appendix B reports data on existing buffer across the industry under APRA's version of the SRP metric.

Design and implementation issues

A number of issues would arise in designing and implementing an SRP test:

- **Assets included** – One issue is whether only equities and fixed income are used, or other asset classes that are passively available should be included (e.g. listed property and infrastructure). Another issue is which classes of equities (e.g. emerging markets included or excluded) and fixed income (e.g. cash proxy, government bonds or broader indices) should be included in the SRP.
- **Indices used** – Indices will need to be chosen to represent each asset class. There is also a choice over whether indices or passive funds or ETFs are used for estimating returns. There will be multiple options available.
- **Adjusting asset weights to form the SRP frontier** – One approach would be to form equity and fixed income sub-portfolios and vary the sub-portfolio weights to trace out a frontier. This requires specifying asset weights within the sub-portfolios. If assets outside of

[1] Changes in SD will lower the benchmark return by more as the slope of the SRP frontier becomes steeper. Perversely the opposite will occur if equities underperform bonds so that the SRP frontier slopes downwards, in which case diversifying to lower SD will raise the benchmark return.

equities and fixed income are included in the SRP, the question arises over how they are allocated or if another sub-portfolio is formed.

- **FX hedging assumption** – The hedging assumption will be embedded in the mix of hedged versus unhedged international asset classes within the SRP. This is an important decision as FX hedge ratios have potential to be a source of significant TE, and so may influence how funds set their hedge ratios in practice.
- **Rebalancing assumptions** – A decision needs to be made on whether the SRP is assumed to be rebalanced yearly, quarterly, monthly or even daily for the purpose of estimating the SRP return and SD over the course of the assessment period. Another issue is whether to allow for notional rebalancing costs.
- **Failure threshold** – The 50bps threshold under the existing test may not be appropriate under an SRP test. The 'right' failure threshold will be tricky to calibrate. Consideration may need to be given to whether a larger threshold might be seen as a weakening of the test.
- **Estimation of SD** – It is assumed that SD would be estimated with respect to the assessment period (i.e. 10 years). A key choice is the data interval, i.e. whether monthly, quarterly or yearly return observations are used. There is a trade-off between having more observations and diluting the distortion from assets with smoothed returns, e.g. private assets. Another option might be to apply a model to account for the impact of any serial correlation in the data.
- **Transition to an SRP test** – The SRP test might be introduced on an immediate, deferred or transition period basis, with the latter potentially involving running the SRP and existing test in parallel for a period.

The Treasury consultation paper asks for inputs on some of the above issues, mentioning:

- asset classes and market indices
- whether SD is an appropriate risk measure, and if there are any alternatives
- implementation and transition considerations

APPENDIX A

Impact of diversifying mid-risk assets

We undertake notional modelling to gauge the incentive to diversify into mid-risk assets, such as unlisted infrastructure and property. We assume that mid-risk assets will reduce the SD of the overall portfolio but also offer lower expected returns. Our focus is the 'breakeven' reduction in expected return versus equities where the fund would expect no detriment from adding a 20% weighting to mid-risk assets under the SRP test relative to a 70/30 equity/fixed income portfolio. We assume the allocation to mid-risk assets is funded 15% out of equities and 5% out of fixed income. Key asset return assumptions appear in the table below.

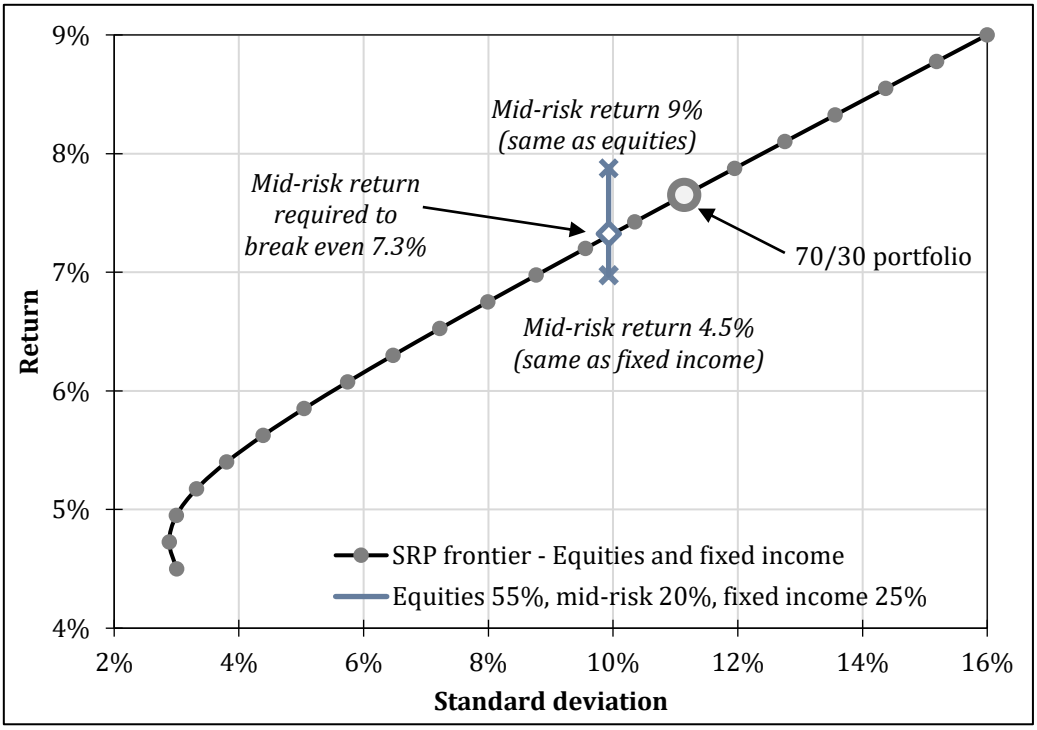
SRP and mid-risk assets: Modelling assumptions

	Fixed income	Mid-risk	Equities
Return	4.5%	*Imputed*	9.0%
SD	3.0%	10.0%	16.0%
Correlations			
Fixed income	1.00		
Mid-risk	0.10	1.00	
Equities	-0.10	0.50	1.00

The outcome from the analysis is presented in the chart below. Adding mid-risk assets reduces the standard deviation of the portfolio and accordingly lowers the benchmark return by shifting the portfolio down the frontier relative to a 70/30 equity/fixed income mix. The breakeven expected return on mid-risk assets to match the benchmark return is imputed at 7.3%, or -1.7% below the assumed equity return of 9.0%. A fund might thus be encouraged to allocate to mid-risk assets if they are sufficiently confident that they beat the 7.3% hurdle. If not, they might be encouraged to hug the SRP benchmark. The chart also shows the range where the mid-risk assets offer a return equal to equities and equal to fixed income.

Bear in mind that the breakeven expected return estimate depends on the modelling assumptions. Our analysis does not consider confidence in achieving the expected return relative to equities (and fixed income). If the fund is not confident in the relative expected returns, it may be more inclined to limit exposure to mid-risk assets to limit the risk of test failure.

Impact of mid-risk assets under the SRP test metric



APPENDIX B

Results under APRA's application of the SRP metric

APRA calculates performance versus an SRP under its Comprehensive Product Performance Package (CPPP). APRA's methodology differs from that proposed under the SRP test by assigning a fund-specific SRP based on growth/defensive mix. Nevertheless, it gives some hint about the potential SRP buffer that currently exists across super funds.

The table extracts SRP-relative performance for MySuper funds to June 2025 over a range of periods. A majority of funds underperformed APRA's SRP metric over all time periods, with the underperformance largely sourced from the three years to June 2025.

We surmise this spate of underperformance reflects poor performance from certain private assets and security selection (i.e. active management) relative to public market indices. These trends appear to have continued into FY2025-26. APRA's estimates hint that super funds may enter a SRP test regime with a period of 3-4 years of substantial underperformance that will remain within the history for years to come. While shifting to SD as a risk measure may trim back some of the underperformance, we doubt it will nullify the full extent of the poor performance over recent time.

This situation could boost the perceived probability of failing the test going forward for both individual funds and the industry at large, especially if funds focus on the fact that the spate of poor performance will remain in the history .

Performance of MySuper funds vs. the SRP to June 2025

	Count	Performance versus SRP				Imputed for year t-10 to t-4
		10-years	7-years	5-years	3-years	
Single strategy	28					
Average		-0.16%	-0.44%	-0.34%	-1.55%	0.44%
Median		0.04%	-0.44%	-0.43%	-1.58%	0.73%
% below SRP		43%	75%	71%	96%	
Lifecycle (average of stages)	15					
Average		-0.51%	-0.66%	-0.43%	-1.34%	-0.16%
Median		-0.59%	-0.59%	-0.30%	-1.22%	-0.32%
% below SRP		87%	100%	87%	100%	
All funds	43					
Average		-0.18%	-0.46%	-0.35%	-1.58%	0.43%
Median		-0.25%	-0.53%	-0.40%	-1.51%	0.29%
% below SRP		58%	84%	77%	98%	

Data source: APRA CPPP, MySuper Product Performance