Superannuation Performance Test Regulations 2023

Submission

May 2, 2023

David Bell & Geoff Warren

About The Conexus Institute

The Conexus Institute is an independent, not-for-profit research institution focused on improving retirement outcomes for Australian consumers. Philanthropically funded, The Institute is supported by the insights of a high-quality advisory board, who work on a pro-bono basis. The Institute adopts a research-for-impact model and frequently collaborates with researchers from academia, associations, and industry. Where possible research is made open source to assist industry and create transparency and accountability. Further information here.

About David Bell

Dr David Bell is Executive Director of The Conexus Institute. Bell's career has been dedicated to the investment and retirement sector. He has worked with both commercial and profit-for-member firms, and ran his own consulting firm. Bell worked with APRA in the development of the APRA Heatmap. Academically, Bell taught for 12 years at Macquarie University and in 2020 completed his PhD at UNSW which focused on retirement investment problems. Full bio here.

About Geoff Warren

Dr Geoff Warren is an Associate Professor at the Australian National University and Research Director with the Consensus Institute, as well as a member of various investment and research advisory boards. Warren's research focuses on investment-related areas specially including superannuation and retirement, and is widely published in leading journals. He has a prior career in the investment industry spanning over 20 years. Full bio here.

***** The authors are willing and able to participate in further consultation. *****

Executive summary

The submission focuses on the Your Future, Your Super (YFYS) performance test component of the proposed amendments. The changes as detailed in the Exposure Draft and Explanatory Statement are generally sensible. They should have a modest positive impact while incurring no negative impacts (apart from some additional administrative complexity for APRA). Treasury should be acknowledged for navigating such a complex topic with many divergent opinions in a short period of time.

In our opinion, the adjustments to the YFYS performance test do not go far enough to significantly improve consumer outcomes. As it stands, the performance test continues to focus on only part of the investment process for multi-asset options, specifically implementation of a strategic asset allocation rather than its formulation, while doing so on a backward rather than forward looking basis. Our research indicates that the implicit costs of this approach continue to exceed the explicit benefits. We recommend leaving the possibility of a major overhaul to the test on the table for the future. In this regard the positive sentiment expressed by Minister Jones for further review is encouraging.

For a review to be effective at protecting consumers from future underperformers while reducing the implicit costs of the present design, we make two recommendations:

1. The review should begin in 18 months. This would follow four rounds of annual performance testing, which should make it possible to assess benefits and costs, both implicit and explicit.

2. The review should be undertaken by independent investment experts. Exploring the implicit costs requires expert understanding of specialised areas such as portfolio construction, asset allocation and risk management.

One substantial idea we would like to see considered by such a review is whether performance testing could be taken out of legislation and transferred to APRA, allowing them to implement a holistic assessment that combines qualitative assessment alongside a range of metrics. This would result in an evergreen framework that can readily adjust to future challenges as they arise.

The structure of this submission is as follows. We initially discuss specific changes, noting the outcomes that we anticipate from the changes. We close by detailing our recommended solutions for future consideration.

1. Detailed modifications

1.1. Extended assessment timeframe

Extending the timeframe of the performance test from 8 years to 10 years will provide a largely cosmetic benefit, with the true benefit being modest. Statistically lengthening the timeframe of the test notionally improves the statistical power of the performance test through reducing possibility of failing to identify “poor” funds as poor performers, as well as of “false positives” where ‘good funds’ are mistakenly identified as ‘poor’ funds. However, the degree of this statistical improvement is small (detailed in the Appendix). In addition, extending the timeframe of the test lengthens the exposure to a past underperformance ‘event’ that may have been subsequently addressed by a fund trustee through instigating changes to address any issues.

There will be winners and losers from extending the assessment timeframe, both in terms of funds and, ultimately, consumers. Winners are those funds with good performance in the earlier part of the sampling period, as they will carry their good performance numbers for longer. Losers are those funds with weak performance earlier on, which will also be carried for longer. Portfolio management flexibility for these 'loser' funds will remain constrained for longer, which could work to the detriment of their members as they prioritise managing their YFYS test results over innovating or pursuing non-benchmark assets that might deliver better outcomes for members.

1.2. Additional asset classes

YFYS performance test tracking error is incurred in two ways:

1. Intended tracking error incurred through taking on active risk as measured against relevant benchmarks; or,

2. Unintended tracking error (or noise) incurred through the mismatch created by benchmarking a particular exposure against an index that does not accurately reflect the underlying nature of the exposure, e.g. the prior benchmarking of hedge funds against a 50/50 equities/fixed income benchmark regardless of their equity market exposure.

For the benchmarking-based YFYS performance test to work efficiently, it needs to reflect (1) and not (2). A set of benchmarks that is smaller and less relevant to the investment activities undertaken by funds will increase the degree of unintended tracking error.

In this context, the incorporation of benchmarks for additional asset classes or sub-classes will improve the efficacy of the performance test through reducing unintended tracking error.

Regarding the three major sectors where changes have been made:

1. ***Global equities***: The distinction between developed and emerging markets is sensible.

2. ***Fixed income and credit***: Facilitating greater distinction between investment activities within this broad sector is also sensible. That said, we believe there are some shortcomings in the particular indices proposed:

   a. For Australia, the proposed credit index reflects high-grade and largely fixed-rate credit exposure. We understand that, increasingly, this will not reflect the investment activities of Australian super funds in Australian credit. For instance, significant activity has developed in private credit, which is generally floating rate. We suggest that Bloomberg AusBond Credit FRN 0+ Yr Index (BAFRN0) would be a more appropriate benchmark for these activities.

   b. For both Australian and international geographic sectors there is further opportunity for improvement through additional indices that differentiate exposures across the credit spectrum, between low and high duration, and separate out inflation-linked exposure as appropriate.

3. ***Alternatives***: The allowance for defensive and growth alternatives is well-merited. The challenge with the documented benchmarking approach for this sector remains referencing equities and fixed income as benchmark assets, when these are the exposures that alternative strategies are often attempting to diversify away from. We acknowledge there is no easy solution to this shortcoming. The alternative of "cash+" type benchmarks is more intuitive, but has the potential to be gamed.

Our previous analysis suggests that *at least 50* additional indices would be required to make the performance test an accurate assessment of past performance for all super funds. The benefit of this greater efficacy from expanding the benchmarks under an ultimately backwards looking test that focuses on one element of performance needs to be weighed against the extra administration requirements for APRA and industry. We suggest that an assessment of the potential for distortion and the prominence as an asset or asset sub-class in super fund portfolios may guide benchmark selection. From this perspective, we consider the following to be the most important investment sectors that may be worthwhile considering:

- Australian small cap equities
- Australian private equity
- Global private equity
- Tail risk protection
- Commodities
- ESG indices

Apart from Australian small cap equities, each area is highly difficult to benchmark accurately. We hence recommend adding Australian small cap equities in the current round of amendments, and considering the other investments listed under future reviews.

1.3. RAFE

The following text is taken from the Explanatory Memorandum:

New subregulations 9AB.4A(3) to (6) clarify that the RAFE for a lifecycle Part 6A product is the lifestage that incurs the largest RAFE from the set of lifestage RAFEs.

We are confused by this text, which should be tidied up. At first instance it appears to treat lifecycle strategies unfairly. However, some in the industry have said that lifecycle strategies are required to charge a consistent fee across all life stages (which would potentially create cross-subsidisation challenges). If correct, this would make this text appear redundant.

1.4. Currency hedging ratio

The Exposure Draft and Explanatory Memorandum detail modifications to the assessment of currency exposure. To our reading and interpretation, we can see changes to determine a more accurate measurement of gross foreign currency exposure. To our understanding, the underlying currency exposures are currently assumed to be the same as the currency basket represented in the MSCI World index. This index may be a poor proxy for the currency exposures and hence hedging activities of a fund. We recommend that efforts to improve the measurement of aggregate foreign currency exposure needs to be matched by improved ability to account for specific country currency mixes in performance calculations.

1.5. Application of Performance Test to Trustee Directed Products

The Explanatory Memorandum references ensuring that the performance test *"is fit for purpose when it is extended to trustee directed products"*.

The application of the performance test to trustee directed products (TDPs) will provide a degree of consumer protection. However, it will also exacerbate all the shortcomings and costs

associated with the existing performance test structure. Applied to MySuper defaults, the lived experience suggests that the performance test has delivered some explicit benefits in the short term (like admin fee reductions), while implicit costs which will grow over the long term. We expect the same outcomes should the test be applied to TDPs. Acknowledging that the Government has committed to a more detailed review of YFYS in the future (which will hopefully lead to reduced implicit costs), we support the extension of the performance test to TDPs.

A special case is socially responsible investment (SRI) options. The Conexus Institute has undertaken research in conjunction with FTSE Russell[1] which demonstrates that the investment strategy of some of these products will be difficult to sustain as they generate sizable YFYS performance test tracking error, hence 'business' risk. For some funds it will be difficult to wind back the commitments made, such as exclusions. Here, we are aware of strong stated preferences of members in these investment options. There is no easy solution to this problem without major changes to the testing framework. We believe Treasury and Government are suitably aware of this issue, and their decision to test SRI products as part of TDPs is an informed one.

2. Impact of changes on the YFYS performance test

Consumers deserve the protection of a high-quality, effective performance test – especially those in default options. Unfortunately, the current YFYS performance test is deeply flawed. It provides some strong performance benchmarking signals to trustees, but also gives rise to a range of implicit costs that potentially exceed the identified benefits[2].

The changes outlined in the Exposure Draft and Explanatory Memorandum improve the existing test. To summarise our previous analysis:

1. The extended assessment timeframe improves the statistical accuracy of the test, but only by a small amount, while creating a winners and losers effect. It also entrenches the fact that the test does not allow for structural changes in the way that funds are managed.

2. The additional indices support a more accurate assessment of benchmark-relative performance and remove some portfolio management restrictions. This is where we see the main benefit of the changes.

Nevertheless, the scale of improvement is modest and constrained by the broad form of the test and its shortcomings. Specifically:

- The test remains backwards-looking and does not account for important qualitative considerations that may inform whether past performance outcomes are likely to persist;

[1] Your Future, Your Super Performance Test: Constraint on ESG, Sustainability and Carbon Transition Activities by David Bell and Trista Rose (Nov 2022). Full paper here.

[2] Your Future Your Super Performance Test: Estimating the Opportunity Cost to Consumers (here) and related update Your Future Your Super Performance Test: Constraints and Sustainable Tracking Error (here).

- The test utilises a single metric, which assesses only a component and not all important sources of performance, in particular the formulation of the asset mix or investment exposures at the total portfolio level from which member outcomes primarily stem;
- The test creates a range of distortions that carry implicit costs, such as constraints on portfolio management and benchmark-gaming.

3. Recommendation for improved consumer outcomes from performance testing

We acknowledge that the existing YFYS performance test was a product of the environment at the time. This included the direction provided by the Productivity Commission's work, the subsequent decision to adopt a 'bright lines' test to precipitate change, and the resulting implementation in 2021 by the previous Government that gave the test its current form. The aim of ensuring greater accountability and protection for consumers was laudable.

Our view is that the performance test has subsequently delivered some explicit benefits. These include a reduction in fees (particularly administration fees), a sharper focus of fund governance structures and exit from the industry of some funds whose right to exist might be questioned.

However, we have now entered a phase where the industry is becoming increasing adept at managing to ensure they pass the test. This is giving rise to implicit costs, largely due to imposing constraints on the manner in which portfolios are managed.

From here, we believe that the performance test should not be removed, but rather recommend that the nature of the test undergo significant review at an appropriate time. Consumers deserve the protection of a high-quality, effective performance test. A well-designed test could capture many of the identified benefits but incur fewer costs with less unintended consequences.

Regarding a more substantial review of the YFYS performance test, we make two recommendations:

1. The review should begin in 18 months. Four rounds of annual performance testing would then have been undertaken, making it possible to for an informed assessment of the benefits and costs, both implicit and explicit.

2. The review should be undertaken by independent investment experts. The design underpinnings of the performance test are motivated by economic approaches. Explicit benefits have been achieved related to addressing underperforming funds. However, implicit costs are being incurred related to impacts on fund behaviours that may have adverse unintended consequences for members going forward. Our concern is that future implicit costs may be substantial and may ultimately outstrip the benefits to consumers. Exploring these implicit costs requires an expert understanding of specialised areas such as portfolio construction, asset allocation and risk management.

Throughout the Government's review of the YFYS performance test we have proposed that the performance test is transitioned from a legislative process with bright lines determination of pass versus fail, to a regulatory process housed by APRA that considers multiple metrics and accounts for qualitative considerations. This would improve the quality of the test and ensure it has evergreen properties that account for developments in policy, industry and markets. The focus could then switch from being backward-looking to considering forward-looking prospects. The

assessment could appropriately account for the nuances of different product categories (MySuper, TDP's and Choice), and structural changes within funds that might have a significant influence on future member outcomes. The burden of proof would be for funds to demonstrate that an apparent failure on the metrics does not justify being branded an 'underperforming fund', but could be structured to prevent legal challenges by super funds. We suggested that a transition of this magnitude could take place over three years. This proposal would be best considered by the independent review we have recommended.

The trap with aiming to improve the existing performance test is being drawn into a black hole of trying to develop better performance metrics including increasingly more granular benchmarks. This will only increase the administrative burden, and possibly widen the scope for benchmark-gaming. Focusing on past performance and the benchmark by which it is evaluated provides the perfect diversion from the elephant in the room: the fact that the test itself is flawed as does not qualitatively account for a range of factors that are important for future member outcomes, such as governance structures to investment capability to resourcing. A qualitative overview informed by performance metrics has the potential to provide substantial consumer benefits.

Appendix – Analysis of assessment timeframe

Our analysis of increasing the assessment timeframe from 8-years to 10-years applies the YFYS performance test as it stands in a controlled environment. We ignore other sources of performance such as strategic asset allocation, as well as flaws such as failure to allow for risk. We estimate the impact of different timeframes on the probability of accurately identifying poor funds as poor and falsely identifying good funds as poor. Our prime focus is the possibility for mis-identification due to randomness in investment outcomes at an assumed tracking error of 2.5%. The latter is a proxy for the level of tracking error at which many funds seem to be operating, informed by insights gained through confidential interviews with super fund CIO's (here). We also account for differing levels of expected relative performance.

Result 1: Effectiveness

The table below reports the probability of identifying a 'poor' fund in a controlled testing environment assuming an expected return (in excess of benchmark) of either -0.5% or -1.0%. We can observe that increasing the timeframe of the performance test has benefits in terms of increasing the likelihood of identifying very poor performers with a -1.0% expected return, although the improvement is only marginal with the probability rising from 71.4% to 73.6%. However, there is no impact on the probability of identifying poor performers with a -0.5% expected return, which remains at 50%[3].

		Test timeframe				
		6yrs	8yrs	10yrs	12yrs	20yrs
Expected return	-0.5%	50.0%	50.0%	50.0%	50.0%	50.0%
	-1.0%	68.8%	71.4%	73.6%	75.6%	81.4%

Result 2: False positives

The following table reports the probability of mistakenly identifying a 'good' fund as a 'poor' fund (i.e. delivering a return of less than -0.5% versus benchmark) for funds with an expected return of between 0% and +1.5%. The reduction in the probability of such false positives in extending the timeframe form 8-years to 10-years are quite modest, ranging from 0.6% to 2.6% lower.

		Test timeframe				
		6yrs	8yrs	10yrs	12yrs	20yrs
Expected return	0.0%	31.2%	28.6%	26.4%	24.4%	18.6%
	0.5%	16.4%	12.9%	10.3%	8.3%	3.7%
	1.0%	7.1%	4.5%	2.9%	1.9%	0.4%
	1.5%	2.5%	1.2%	0.6%	0.3%	0.0%

[3] Our results also highlight that the test embeds a significant probability of failing to identify poor performers using a cut-off of -0.5%, with an expected return in the range of -0.5% and -1.0% leads to a probability of non-identification that sits between 26.4% and 50%.