

Appendix 1

Your Future Your Super Performance Test

Constraints and Sustainable Tracking Error

12 October 2022

Author: David Bell

Modelling reviewed by Nick Callil (WTW)



Table of Contents

- 1. Introduction 3
- 2. Industry response to YFYS 4
- 3. Modelling sustainable tracking error 6
 - 3.1. Assumptions and results 7
 - 3.2. Sensitivities 8
 - 3.3. A comment on performance test buffer 8
 - 3.4. Opportunity Cost 9
- 4. Conclusion 10

1. Introduction

The results of the second round of the Your Future, Your Super (YFYS) performance test provided an interesting datapoint. Of the thirteen funds which failed in the first year, only four of nine (44%) failed a second time (four merged), yet different research pieces¹ suggested that in a controlled environment the expectation would be that six-to-eight of those nine remaining funds would fail.

One interpretation of this observed difference is that super funds are actively managing their funds (investment strategies and administration fees) to pass the YFYS performance test. This shouldn't be a surprise given the consequences of failure.

If we assume that all funds are accounting for the YFYS performance test in the design of their investment strategies, this motivates two important questions which are critical to the current review of the YFYS reforms.

The first matter is how well the performance test aligns with the objective of providing best outcomes for members. This is a complex issue which warrants significant attention. Hopefully the YFYS submission process will create a body of research on this issue.

The second issue is the degree to which the YFYS performance test constrains super fund investment strategies. Anecdotes suggest that the performance test is constraining the investment strategy of funds, but the degree of constraint is unknown. This is what we explore in this paper.

Through recent interview work and broader industry engagement we believe most funds track, or are on a pathway to tracking, their rolling performance and the likelihood of failing the performance test accounting for the estimated tracking error of their portfolios. Anecdotally we see many funds running at a tracking error of around 2% - 2.5%, figures sourced largely from funds with positive "performance test buffer". We suspect funds with little buffer are running at lower levels.

We previously researched this question when the YFYS performance test was first announced². This research suggested that an appropriate tracking error target for funds was around 1%.

Given the discrepancy between what appears to be happening at funds and our original research findings we thought it appropriate to revisit our original research on this subject. This afforded us the opportunity to extend our models to update for learnings and reflections. In particular we refined our definition of a sustainable investment strategy to be one where funds are at low risk of having to substantially reduce their performance test tracking error in response to the yearly performance cycle.

Overall we estimate that a sustainable level of performance test tracking error remains 1%. This level seems to be less than the level of tracking error many funds are operating at. The possible

¹ Research by Parametric (media report: SuperReview: "[20% of super funds may fail](#)") and CEM Benchmarking ("[What is the value of the Your Future, Your Super test](#)") was undertaken in a closed process environment.

² "[Your Future Your Super Performance Test Exploring the Impact on Super Fund Investment Strategies](#)"

scenarios from here are that (1) funds continue to reduce their performance test tracking error, (2) they run an elevated risk of experiencing a shock to their performance test buffer and have to sizably alter their investment strategy, or (3) the performance test is altered to reduce the degree of restriction. If pathway (2) is followed and funds obtain their tracking error through similar activities, then there is a risk that a cohort of funds experience difficulty at the same time.

Further, we find that existing positive buffer levels do not impact the sustainable level of tracking error – it simply affords the ability to take a short-term tactical bet. This means that the findings of our research are relevant to all funds.

Naturally, our estimate depends on assumptions. We consider a range of assumptions and make the model open source ([here](#)) to enable industry and policymakers to better understand this important issue.

Finally, we explore the opportunity cost of the constraints created by the YFYS performance test. Assuming that, over time, industry reduces tracking error towards the sustainable level identified in this paper, and assuming a modest reward for taking performance test tracking error, we estimate the opportunity cost to consumers, in the form of lower expected returns, is \$3.1b per annum.

2. Industry response to YFYS

There has been much interest in industry’s response to the YFYS performance test. The full impact may never be known as funds guard their strategy. Our research into this area³, which focused on funds which have performed well against the performance test, suggests that:

- While all funds are performance test-aware when designing their investment strategy, funds are at different stages of the integration process.
- Many funds have reduced their performance test tracking error.

The emergence and use of YFYS jargon highlights the impact of the performance test. Some common jargon which we refer to in this paper:

- Performance test tracking error: the tracking error of a portfolio calculated against the benchmarks used in the YFYS performance test. This has proven difficult to calculate, with issues such as unlisted assets and accounting for autocorrelation proving to be challenges.
- Performance test buffer: the accrued performance gap, against the performance test, over the past full period (e.g. 8 years). More detailed measures adjust for the performance year which will roll out of the buffer calculation.
- Limp mode: a fund which has little buffer which, to increase the likelihood of survival, needs to significantly reduce performance test tracking error. Unless a fund has one or two significantly bad years rolling out of the performance test calculation soon, recovering from

³ [“Assessing the impact of YFYS through interviews with CIOs of funds with performance “buffer””](#)

limp mode can take an extended period, as it is difficult to restore buffer without taking on performance test tracking error.

While fewer funds are failing the test, likely because they are actively managing their situation, we suspect some funds are operating in limp mode and that it may take years to emerge from that position. It is not a situation that can be publicised by affected funds.

Funds are at different stages of accounting for the YFYS performance test in their investment strategy. Broadly, we believe there are two specific factors which drive the degree of implementation:

1. Size, which informs overall resourcing and capability.
2. The degree of buffer, whereby funds with buffer have a need to monitor the test very closely.

This framing is reflected in Figure 2 which stylises the different situations that funds may find themselves in.

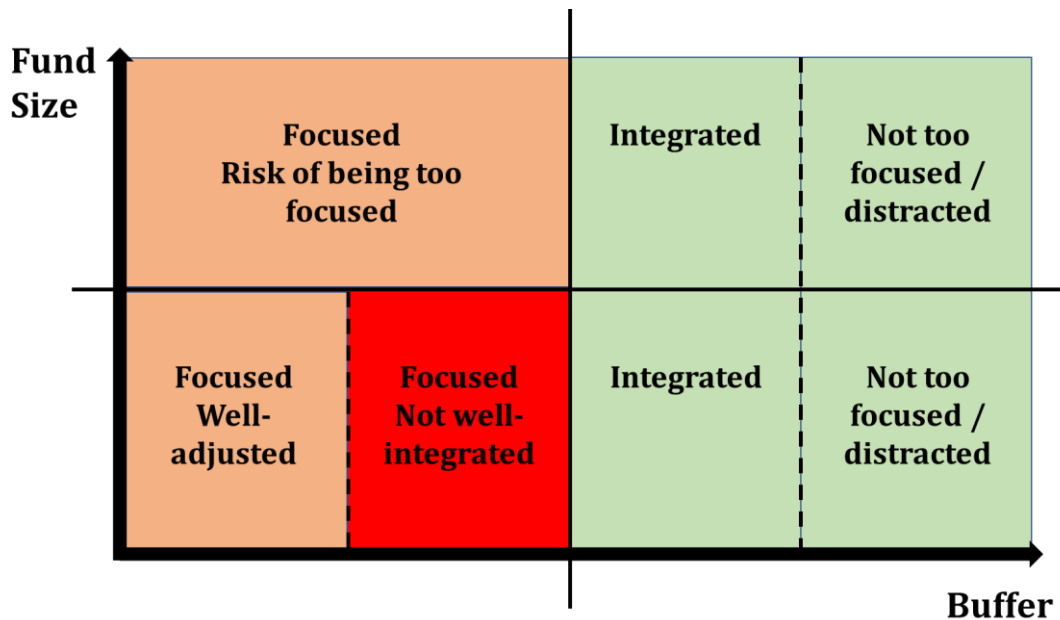


Figure 2: Super funds and approaches to integrating the YFYS performance test into investment strategy design. Stylised assessment by David Bell.

Observe in Figure 2 that there are two categories within three of the segments. We do this to broadly capture fund-level specific reasons (e.g. not well-equipped, don't want to change existing processes). Overall, our estimate of how the industry is presently positioned to account for the performance test:

- The majority of funds have a good handle of their rolling annual performance position, and account for the risk of failing in the next year incorporating an estimate of portfolio tracking error.
- Many funds don't look ahead more than the next year in determining their present year's tracking error.
- Most funds manage the risk of failure and aren't yet considering the risk of falling into limp mode.

Based on these estimates our concern is that funds may be targeting an unsustainably high level of performance test tracking error which does not account for future return series or the risk of falling into limp mode.

The outcome if funds target a relatively high a level of performance test tracking error is a higher possibility of falling into a limp mode scenario which can be costly (transaction costs, investment relationships) and long-lasting (since lower tracking may make it more difficult to re-build a buffer).

3. Modelling sustainable tracking error

We aim develop a model which can be used to estimate an appropriate level of performance test tracking error for super funds. We aim to incorporate relevant factors that a fund should consider in determining its investment strategy which. We frame our modelling through two steps.

The first step is a regular review process, detailed in Figure 3. This reflects the process of a fund which routinely undertakes a review of the appropriate level of performance test tracking error.

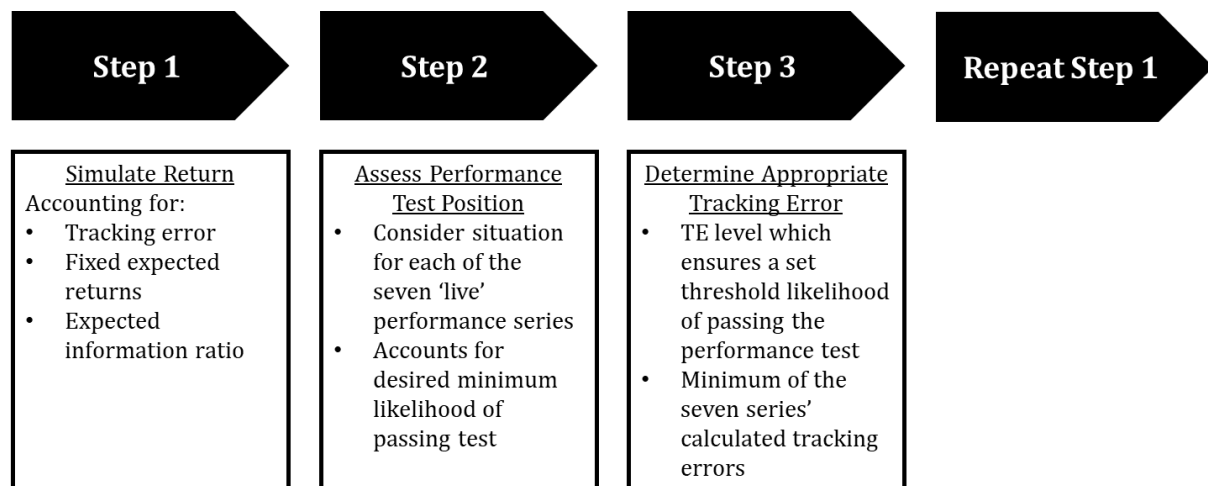


Figure 3: Model of the regular process for a fund reviewing the appropriate level of performance test tracking error.

The second step is to account for the concept of a sustainable investment strategy. Here, we believe a sustainable investment strategy is one where the investment strategy does not incur significant disruption, as we consider this detrimental to long-term outcomes⁴.

⁴ The potential adverse consequences of substantial change include:

- Higher transaction costs.
- The need to reduce exposure to illiquid assets which are difficult to sell.

To account for this consideration the model simulates many outcomes and calculates the percentage which experienced substantial through-time change in investment strategy.

3.1. Assumptions and results

We set the base case assumptions for a fund as:

1. Fixed expectation of outperforming its YFYS tailored benchmark by 20bp pa (e.g. associated with cost advantages).
2. Expected information ratio of 0.2, meaning an additional outperformance expectation of 0.2 times the performance test tracking error target.
3. Desire to maintain a 95% likelihood of passing the performance test over any rolling 8-year period, starting from now (t=0).
4. Desire to be 95% sure of maintaining a sustainable investment strategy over a rolling 8-year period, starting from now (t=0).
5. Defined threshold level for a sustainable investment strategy is not having to reduce performance test tracking error by more than one-third (cumulative, over time).

Note that assumption (3) is dominated by assumption (4) and doesn't drive the sustainable level of performance test tracking error calculation. In the model the one number is used for both assumptions.

When we apply these assumptions to our model, we find that 1% to be approximately the sustainable level of performance test tracking error. Figure 4 presents a set of simulations of the ongoing tracking error management process to illustrate the variability in the investment strategy over an 8-year cycle.

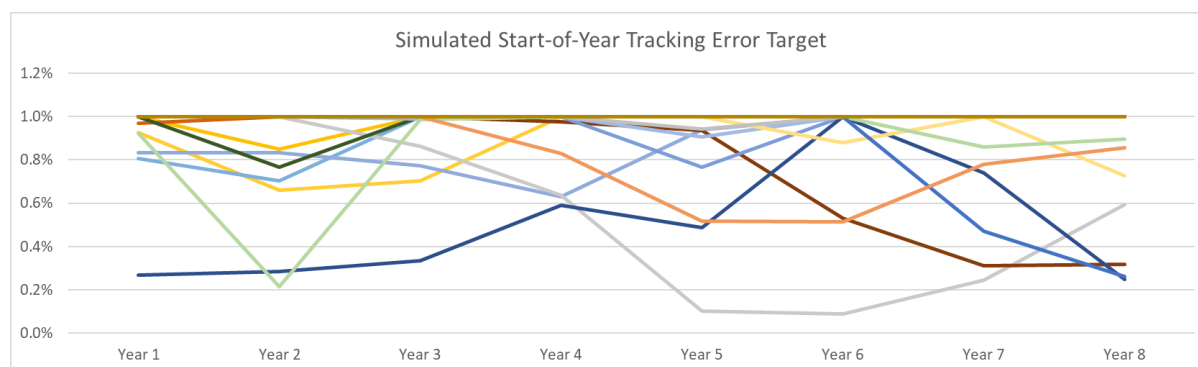


Figure 4: Simulated through-time tracking error management for a scenario using base case assumptions.

- Not being able to manage risks to member outcomes to the degree which the Trustee would like or committed to. Notable examples include managing ESG risk, portfolio overlays and diversification strategies.
- Funds may potentially be 'squeezed' out of positions at a time when they are undervalued.
- Funds may impair their standing and relationships with external fund managers.

3.2. Sensitivities

In Figure 5 we consider the impact of changing each assumption on the sustainable level of tracking error. The base case assumptions and the ranges considered are detailed in Table 1.

	Base Case	Range
Expected return	0.2%	0% - 0.5%
Information ratio	0.2	0 - 1.0
Threshold level of certainty (of both passing text and of not experiencing substantial modifications beyond threshold)	95%	90% - 99%
Sustainability threshold (maximum permitted cumulative reduction in tracking error)	33.3%	20% - 50%

Table 1: Base case assumptions and ranges considered for sensitivity analysis of sustainable level of tracking error.

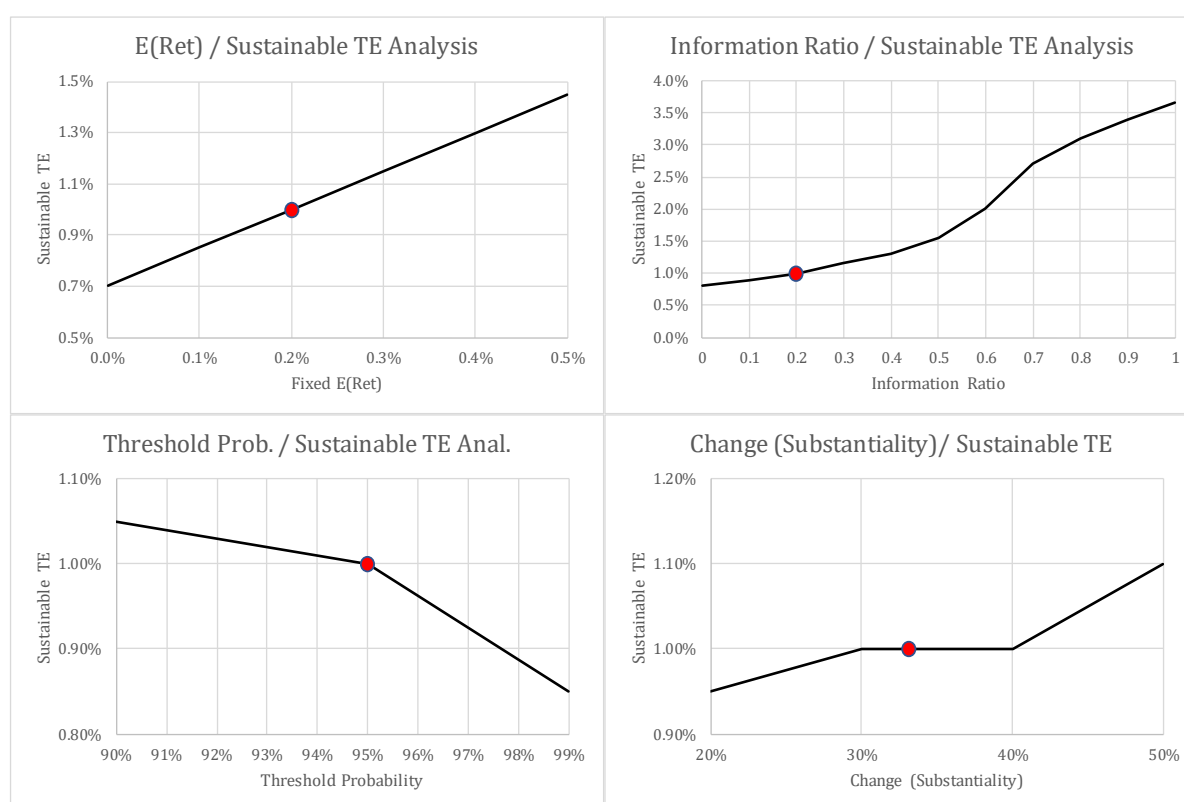


Figure 5: Sensitivity analysis of sustainable tracking error. Red dots reflect the base case assumptions. One parameter is explored at a time, holding the other three variables fixed at their base case assumption levels.

From Figure 5, we identify that assumptions relating to information ratio have greatest impact on the level of sustainable tracking error.

3.3. A comment on performance test buffer

A common anecdote is that funds with positive performance test buffer can operate a higher tracking error strategy. We thought this warranted further consideration because of the way it could distort investment strategies between super funds (between those with and without buffer).

We manipulated our models to consider this. We considered different levels of cumulative performance buffers (i.e. accrued outperformance) ranging from 1% to 5%. The sustainable tracking error did not change in any of these scenarios.

This result may surprise but it makes sense once our modelling is explained. The model searches for the sustainable level of tracking error. This is a long-term consideration, and the model only explores the second 8-year window once performance track records for a first 8-year window have been simulated (effectively a seasoning process).

Funds with higher performance test buffer could elect to run at a tracking error above the sustainable level. In effect, this is a tactical decision. Each year of performance is reflected in eight performance series, so higher short-term tracking error creates potential for a future limp mode situation.

3.4. Opportunity Cost

From this analysis it is possible set out a basic estimate of the expected opportunity cost of the YFYS performance test constraints. This is based on an assumed industry-level information ratio and the reduction in performance test tracking error.

The estimated annual cost of the constraint can be estimated by:

Assumed information ratio	x	Size of impacted asset pool	x	Reduction in performance test tracking error
---------------------------	---	-----------------------------	---	--

Applying the following assumptions:

0.2 (assumed information ratio)	x	\$1.55t (capital in non-public sector non-SMSF, non-DB super funds ⁵)	x	1% (assumed reduction in performance test tracking error)
------------------------------------	---	--	---	--

On the basis outlined, the estimated opportunity cost to consumers, in the form of lower expected returns, is \$3.1b per annum. This is slightly lower than the estimate in our previous (March 2021) work (\$3.3b pa)⁶. The differences are based on an improved insight into the level of performance test tracking error being taken by super funds.

⁵ We assume that 50% of DB super assets are held in public sector funds.

⁶ [“Your Future Your Super Performance Test: Estimating the Opportunity Cost to Consumers”](#)

4. Conclusion

In this paper we estimated a sustainable level of performance test tracking error for super funds to be 1%. This accounts for the concept of a sustainable investment strategy, meaning a strategy which won't require an unworkable degree of change through time. All our modelling is open-source so that alternative assumptions can be explored.

Further, we explored the case of performance test buffer. Our modelling suggests buffer affords a tactical tracking error decision but has little impact on the sustainable level of YFYS tracking error. It seems that many funds with positive buffer are running at higher levels of tracking error than the sustainable level identified in this paper. Unless their return expectations (particularly information ratio assumptions) are justifiably higher than ours, they run a reasonable likelihood of experiencing strategy impairment (limp mode) in the future. If these funds obtained their tracking error through similar activities, then there is a risk that a cohort of funds experience difficulty at the same time.

Assuming that, over time, industry reduces tracking error towards the sustainable level identified in this paper, and assuming a modest reward for taking performance test tracking error, we estimate the opportunity cost to consumers, in the form of lower expected returns, is \$4b per annum.

Appendix 2

Assessing the impact of YFYS through interviews with CIOs of funds with performance “buffer”

26 July 2022

David Bell

1. Introduction

Twelve months on from the introduction of the Your Future, Your Super (YFYS) performance test, this research explores the impact on those super funds with performance test “buffer”. The basis for this research was confidential interviews with CIOs of ten super funds.

Most, but not all, of these ten funds expect returns will be adversely impacted over the medium to long term (because they will take less active risk) and some believe it will be harder to manage risk. Investment time horizons are shorter for most, and most funds could readily identify where they are restricted. A number had already made portfolio changes.

Two specific challenges were constantly highlighted. Both relate to “choice” or trustee-directed options: one, the incorporation of fiduciary decision-making into conservative-style options; the other, the management of SRI options where member feedback has informed expectations around exclusions.

Support for a performance test to protect consumers in default options was unanimous. But there was no clearcut solution to the problems identified. Most agree that a purely benchmark-based approach is flawed and that it is important to account for asset allocations and risk-adjusted outcomes. Many believed that a qualitative oversight would create a more forward-looking outcome. A solution for choice options is an imperative. There seems potential in an APRA-based solution.

Arguably, this research provides a picture of how the entire industry is impacted by the performance test. In short, buffer informs the performance test tracking error that a fund can tolerate. It would therefore be expected that funds with little or no buffer will be more adversely impacted by the test than the funds who participated in this research project. We believe that a sizable portion of the industry is in “limp mode” or not far off, and question whether this is the best foundation for an industry to manage portfolios for long-term outcomes.

Thank you to the staff of the ten funds involved, both preparation and meeting time.

In keeping with the confidential nature of this survey, direct quotes from participants used in this paper are non-attributed.

2. Research approach: why a confidential interview approach?

The performance test has been the subject of much opinion and research. Most of this has been opinion-based; with limited data it has been difficult to undertake ex-post analysis. Some notable examples of ex-ante research have explored tracking error¹.

The lack of hard data will remain an issue, making it important to utilise other non-empirical research techniques. Funds also are careful with what they will say publicly on policy and regulation. They are naturally guarded about sharing intellectual property around their investment strategies.

This is where a confidential, qualitative interview approach can work well. For example, Geoff Warren (ANU) has used this technique to explore various issues including active management², MySuper default design³ and in-house management⁴.

The ten funds selected covered different characteristics, notably size, degree of buffer, fund model (profit-for-member, commercial), and investment model aspects (degree of internal versus external management).

2.1. Jargon

Some jargon used in this paper:

- **Performance test tracking error:** the tracking error of a portfolio calculated against the benchmarks used in the YFYS performance test. This has proven difficult to calculate, with issues such as unlisted assets and accounting for autocorrelation proving to be challenges.
- **Performance test buffer:** the accrued performance gap, against the performance test, over the past full period (e.g. 8 years). More detailed measures will adjust for the performance year which will roll out of the buffer calculation.

¹ In "[Exploring the Impact on Super Fund Investment Strategies](#)" David Bell explored the appropriate level of tracking error for super funds. Researchers from Parametric (Jennifer Sireklove, David Post and Josh Mckenzie) explored the tracking error associated with different ESG implementation approaches in "[Can ESG Investing Survive Your Future, Your Super?](#)".

² "[Interviews with Institutional Investors: The How and Why of Active Investing](#)" by Doug Foster and Geoff Warren.

³ "[Design of MySuper Default Funds: Influences and Outcomes](#)" by Adam Butt, Scott Donald, Doug Foster, Susan Thorp, and Geoff Warren.

⁴ "[In-House Asset Management in the Australian Superannuation Industry](#)" by David Gallagher, Tim Gapes, and Geoff Warren.

- **Limp mode:** a fund which has little buffer which, to increase the likelihood of survival, needs to significantly reduce performance test tracking error. Unless a fund has one or two significantly bad years rolling out of the performance test calculation soon, recovering from limp mode can take an extended period, as it is difficult to build buffer without taking on performance test tracking error.

3. How funds with buffer are accounting for the YFYS performance test in their investment strategies

The broad industry framing of priorities for investment strategy priorities in a YFYS performance test world is depicted in Figure 1.

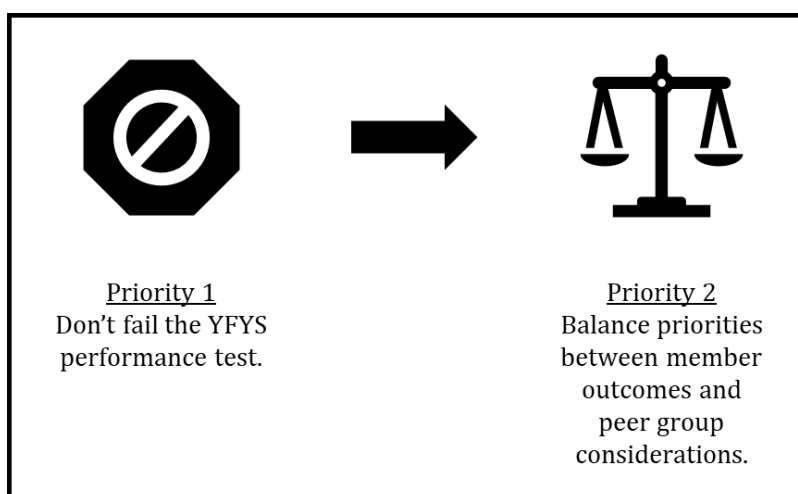


Figure 1: Framing of investment strategy priorities in a YFYS performance test world.

Some quotes:

“YFYS is like the qualifying event for the main tournament – you have to qualify”

“You can’t maximise outcomes for your members if you aren’t in existence”

For many interviewed funds, the focus on peer grouping was significant. Arguably, the focus on peer groups has increased as a result of the impacts of various components of the YFYS reforms. The YFYS Comparison Tool, despite its shortcomings, provides an avenue for performance comparison by consumers, while stapling and consolidation are likely to add to competition, particularly for default awards.

The intense focus on peer group behaviour was surprising. A natural concern is whether this is to the detriment of member outcomes. Reconciliation is reached through two lines of reasoning. The first is the observation that there are many candidate portfolios which provide a threshold probability of achieving a targeted member outcome. The second is the acknowledgement that the science of maximising returns is far from exact in the presence of significant forecasting uncertainty.

The YFYS-integrated investment strategy (i.e. integration of the elements mentioned in Figure 1) of the funds interviewed could be broadly separated into two groups, detailed in Table 1.

Approach 1: YFYS-integrated	Approach 2: Not too focused / distracted
<p>The YFYS performance test has been integrated into the overall investment framework, impacting activities from determining the SAA through to implementation activities by asset class teams. These funds can estimate their performance test tracking error and communicate the likelihood of failure to their board / investment committee.</p>	<p>These funds take the view that it was their existing investment processes which produced this buffer. They have concerns that formally integrating the performance test into their processes may negatively impact on investment team culture and outcomes. Some of these funds feel there is the chance to further distance themselves ahead of peers by continuing to take risk and apply their process.</p>

Table 1: Framing of investment strategy priorities in a YFYS performance test world.

A distinguishing characteristic was buffer: Approach 2 was a less viable option for funds with small buffer. Indeed this proved a central reflection of this research: in a YFYS world many aspects of investment strategy and management are buffer-dependent. Some of the interviewed funds would correctly argue that their process reflects elements of both approaches. For some funds their prior-YFYS process only required moderate tweaking to fall into Approach 1. This would be where a central focus of the overall investment strategy was implementation.

4. Impacts of the performance test on investment strategy and implementation

The majority of the identified impacts of the performance test would be categorised as negative. This needs to be considered separately from views on the YFYS performance test itself and its industry-level impacts. As detailed later, all funds were supportive of a performance testing framework as a valuable consumer protection. Further, funds could identify positive system level outcomes.

Most of the identified impacts have been previously mentioned (or indeed, predicted) in other YFYS analysis and commentary. A contribution of this research is confirming that these are issues faced by super funds with performance test buffer. Issues identified:

- **Reduced investment horizon:** a number of interviewed CIOs viewed that the YFYS performance test has shortened their investment horizon. In most cases it wasn't to manage the risk of failing the performance test, rather to manage the risk of finding themselves in limp mode.

“Longer horizon investing was a real advantage, but YFYS takes that away”

“Are we genuine long-term investors? No more”

“This test makes asset owner capital less stable... which is important for unlisted partnerships”

- **Reduced portfolio management levers.** Many CIOs noted investment opportunities which were now much more difficult to apply in size. These included opportunities in fixed income

(low and high duration bonds, credit, inflation-linked bonds), public equities (low volatility strategies, small caps, and emerging market equities), alternatives (insurance-linked strategies, hedge funds, alternative risk premia), private assets (j-curve risk in new private equity exposure, characteristics of unlisted property and infrastructure).

Some funds have already made portfolio changes, notably reducing equity exposure in emerging markets.

“YFYS is not just a measurement exercise,... it is changing the way that funds invest”

“The ability of the industry to participate in PPP's and nation building style transactions is hampered if those deals aren't necessarily part of the benchmarks”

- Constraints in managing risk in a fiduciary-style manner. Specific situations include:
 1. Conservative-style portfolios where the conservative interpretation was applied not just to the strategic asset allocation but to asset classes as well, for instance fixed income (via low duration bonds) and equities (via low volatility strategies). Many of these options across industry experienced prolonged underperformance until the recent significant market sell-off.

“The industry was so lucky that fixed income sold off in the months before the performance test was applied to conservative options... I feel sorry for those funds who got squeezed out of those positions early”

2. Managing SRI (socially responsible investing) options and climate risks. Here the opposite situation occurred, where recent events represent a perfect storm for these options: a large rally in energy and materials.

Exclusion-based SRI options pose a unique challenge as the exclusions are based on the demands of their membership. SRI options have been one of the fastest growing sectors in superannuation.

Some CIOs believe that offering SRI options in their current format is unsustainable in under the YFYS performance test.

“We are trapped between meeting the stated demands of our members or risk being on the front page of the mainstream newspapers”

- The performance test is having the impact broadly predicted by most. As depicted in Figure 2, there was general agreement that most focus would be on engagement, as this activity generated no performance test tracking error. Exclusions are difficult to incorporate, even to the point where some CIOs questioned the ongoing ability to exclude tobacco (*“it just chews up a chunk of our limited tracking error budget”*). Investments in dedicated impact investments will be limited. Here, some CIOs acknowledged that their firms would seek to maximise the full benefit of such investments through brand association.

Impact Investing	Opportunistic Investing	Exclusions	Engagement
<ul style="list-style-type: none"> • Assumed to be in private markets. • Introduces significant tracking error into the portfolio. • Potential for J-curve effect if unlisted. • Likely to be small scale (if at all). 	<ul style="list-style-type: none"> • Assumed to be in liquid markets. • Introduces tracking error, but this can be measured and controlled. • Likely to be small scale. 	<ul style="list-style-type: none"> • Can introduce significant tracking error if applied in a coarse manner. • Quantitative techniques can be applied to limit tracking error impact. • Likely to remain a practice for small components of benchmark. 	<ul style="list-style-type: none"> • Supplement implemented investment strategy with a range of engagement strategies. • Likely to remain commonplace.

Figure 2: Assessed impact of the YFYS performance test on different sustainability / ESG implementation approaches. Reproduced from [“Moving forward with YFYS: Super funds”](#).

- Nearly all funds have changed their operational practices to report an SAA that is much closer to their actual asset allocation. Some funds have changed or are in the process of changing their benchmarks to match those used in the YFYS performance test.

“Industry has definitely been beefing up investment operations”

- The focus on and resourcing of investment governance has ramped up significantly. This has largely been viewed as a positive. The majority of funds found that their boards / investment committees trusted the internal team for education on YFYS. Downsides mentioned were the degree of reporting, and the heightened connection between brand and performance (especially as it relates to choice option challenges such as SRI and conservative options).

“For our trustee it is death by ‘peralysis’ (performance analysis)”

“Branding now attaches to our investment risk appetite”

- On the potential for gaming the performance test, every fund noted awareness of all the obvious gaming opportunities, but none want to go down the gaming path as it starts to distort their existing investment process. Some CIOs believed that funds with low / no buffer will find it more difficult to resist the temptation to game.

“Perversities and gamification will undoubtedly become issues”

4.1. Overall impact on member outcomes

The most important question for policymakers was whether the performance test would reduce member outcomes. The answer was generally yes (especially over the medium to long term), but not unanimously.

- From a risk perspective, most funds said they were constrained in how they could manage risk, notably within sector risk (e.g. low volatility equities or low duration bonds), through diversification (e.g. alternatives), and risks relating to climate and ESG.

“Yes, my portfolio will be less diversified, but I feel there are sufficient levers to manage risk in my default fund. Choice options are more of a challenge.”

- From a return perspective, a number of funds noted that controlling performance test tracking error means a reduction in expected active returns. Many noted that their investment time horizons are shortened, moving them into a market space where they feel they have reduced insight.

“I have reduced my performance test tracking error, ergo my expected active returns are lower”

5. Solutions

It is important to acknowledge some of the beneficial impacts of the YFYS reform package:

- Reduction in administration fees across the industry. The inclusion of administration fees in the YFYS performance test was a significant factor.
- Considerable industry consolidation, which should flow through to reduced costs over time. Here YFYS was a significant factor, complemented by APRA’s more determined approach to fund sustainability.

Support for a performance test to protect consumers in default options was unanimous.

“Consumers, especially default investors, deserve the protection of a high-quality performance test”

Some funds reflected that relying on members to be the enforcement agent is flawed given low levels of engagement. There was sentiment that a better quality test could underpin stronger consequences for underperformance.

“Empowering members to leave will never be the answer”

However, most but not all, viewed the current performance test as deeply flawed. Purely through the lens of investment performance, most CIOs were unclear whether the YFYS performance test would deliver a net benefit to consumers in aggregate.

The small minority of interviewed funds which supported the performance test in its present form believed implementation performance to be a sound foundation (minimum standard) for protecting consumers.

In terms of solutions, and putting aside the small minority who supported the performance test in current form, four common observations emerged:

1. Any solution based on benchmarks would always be flawed, and the wins from any benchmark changes (whether additional benchmarks or improvements to existing benchmarks) would be modest.

“YFYS tests the ingredients and not the cake”

“Any enhancement to the existing benchmark-based approach would just be papering over the cracks”

2. Focusing on the full investment strategy was critical. This includes accounting for the asset allocation decisions of funds and both the return and risk outcomes that they deliver. This led most surveyed CIOs to reflect positively on a multi-metric approach.
3. A number of funds suggested the need for a mechanism which acknowledges and assesses the forward-looking impact of changes which funds make to enhance their investment model. The risk of a limp mode environment for a sizable portion of the industry is a concerning system outcome.
4. There is a need to revisit the philosophy around performance testing of choice products. This includes both SRI products and non-default products like conservative options. Given the active choice, hence engagement, associated with these products, should the bright lines nature of the performance test be tempered compared to how it is applied to default products?

However, no clear solution emerged from the interviews. Many of the CIOs agreed that a framework combining multiple metrics and a qualitative assessment would be a significant improvement, if implemented well. Here, CIOs broadly agreed that hindsight reflects well on the pre-YFYS APRA operating model (multiple metrics via the Heatmap and a qualitative assessment). The question marks were whether the YFYS performance test had permanently raised APRA's stance to its currently perceived strong level, and whether APRA's frontline team would be able to assess a large throughput of complex qualitative assessments.

6. Completing the picture: how the YFYS performance test has impacted super funds

This research piece allows the completion of an anecdotal assessment of how the YFYS performance test has impacted industry. A clear observation was the direct linkage between performance test buffer and performance test tracking error.

We divide the industry into categories based on two dimensions: buffer and fund size. This is reflected in Figure 3, and we assume, for simplicity, that both size (small or large) and buffer (little / no or sizable) are binary measures.

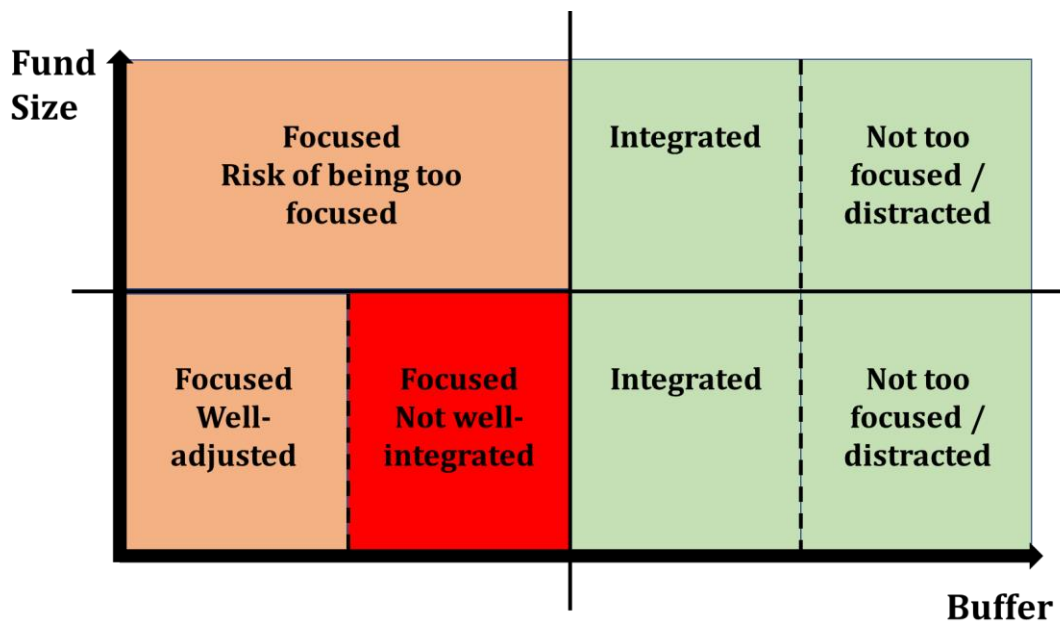


Figure 3: Assessed impact of YFYS performance test on different ESG implementation approaches.

Noting that the right-hand side of Figure 3 was addressed in Table 1, we focus on the remaining three categories:

1. *“Focused, well-adjusted”* small funds have integrated the performance test into their processes, typically with the help of consultants. They have made necessary portfolio adjustments and can keep their board / investment committee informed. Nonetheless, they are in limp mode and their lack of scale means they will likely face ongoing regulatory pressure to merge up.
2. *“Not too focussed / distracted”*: these funds broadly have large buffer and take the view that it was their existing investment processes which produced this buffer. They have concerns that formally integrating the performance test into their processes may negatively impact on investment team culture and outcomes. Some of these funds feel there is the chance to further distance themselves ahead of peers by continuing to take risk and apply their process.
3. Some large funds with little buffer are arguably *“too focused”* on the performance test. This could be to the detriment of member outcomes, whereby asset allocation decisions and active risk decisions are highly constrained. They also consign themselves to operating in limp mode for an extended period. Competition may catch up with them if they underperform peers.

Industry observers have been debating the merits of a scenario of a highly concentrated universe of super funds with non-differentiated investment strategies. In a strange way the YFYS performance test serves to increase dispersion in investment strategies across industry as dispersion in buffer amongst funds drives significant ongoing differences in investment strategy.

In the short to medium term, the investment strategy amongst funds may be more diverse than it ever has been. But over the longer term we expect the consolidation towards a smaller number of very large funds remains the core scenario.

7. Closing comments

An interview-based research approach allows us to confirm that the YFYS performance test has impacted those funds with healthy levels of performance test buffer. For most of the funds interviewed the performance test has the same impacts as those with little or no buffer, except scaled down.

For funds with buffer there appears two different models for how the performance test is being integrated into investment strategies. This research completes the picture of how the entire industry is impacted by the performance test. In short, buffer informs the tracking error that a fund can tolerate. It also informs investment time horizons. We believe that a sizable portion of the industry is in “limp mode” or not far off, and question whether this is the best foundation for an industry to manage portfolios for long-term outcomes.

“That scenario of small buffer: it’s like turning up to a gladiator event with a butter knife!”

Overall most funds with buffer acknowledged that long-term expected outcomes can only be lower in the presence of the performance test (reduced performance test tracking error creates an expectation of lower active returns).

Particular pain points felt right across the industry are management of choice options such as conservative and SRI options. Here the consensus was strongly that the bright lines benchmark-based test is ill-equipped.

Support for a performance test to protect consumers in default options was unanimous. But there was no clearcut solution to the problems identified. Most agree that a purely benchmark-based approach is flawed and that it is important to account for asset allocations and risk-adjusted outcomes. Many believed that a qualitative oversight would create a more forward-looking outcome. A solution for choice options is an imperative. There seems potential in an APRA-based solution.

Thank you to the staff of the ten funds involved, both preparation and meeting time.