SURVIVE THRIVE
Natural Selection
Financial Services Forum
21-22 May 2018 • Hilton Sydney
Actuaries Institute

Metrics for Comparing Retirement Strategies: a Road Test

Prepared by Nick Callil, Hadas Danziger and Tom Sneddon

Presented to the Actuaries Institute
Financial Services Forum
21-22 May 2018

Institute of Actuaries of Australia
ABN 69 000 423 656
Level 2, 50 Carrington Street, Sydney NSW Australia 2000
t +61 (0) 2 9239 6100 **f** +61 (0) 2 9239 6170
e actuaries@actuaries.asn.au **w** www.actuaries.asn.au

Table of Contents

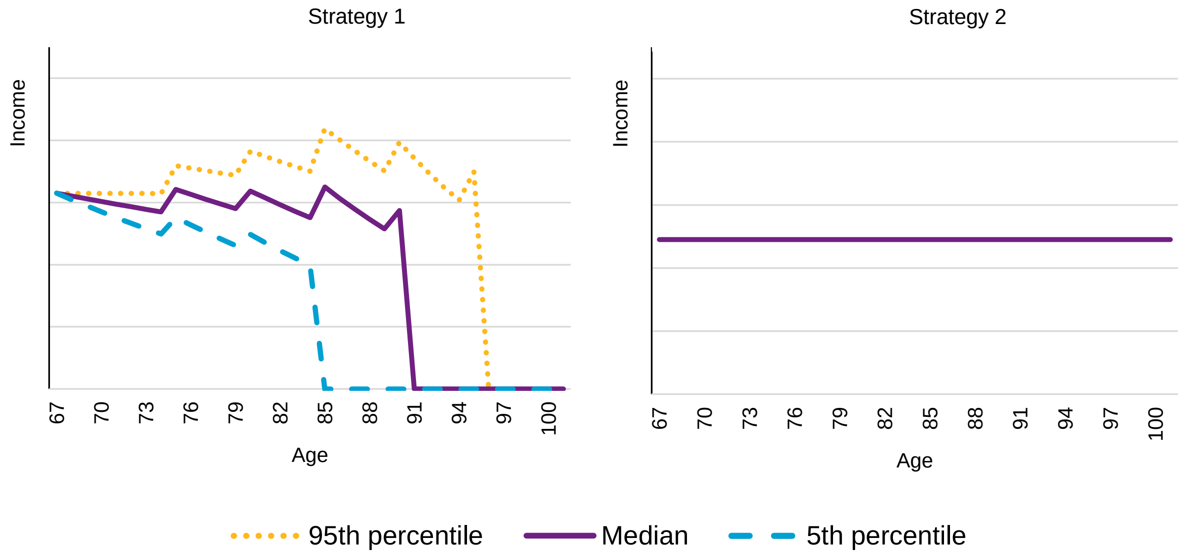
This page is intentionally blank

Section 1: Motivation - why we need retirement metrics

A simple motivation

Which pattern of income payments would you prefer in retirement, either for yourself or on behalf of one of your fund's retirees with an unknown lifespan?

Figure 1: Which income pattern would you prefer?



Even a brief consideration of this simple example raises the core challenge in comparing retirement strategies: determining the 'best' strategy involves weighing up a range of factors which are not directly comparable. In this case:

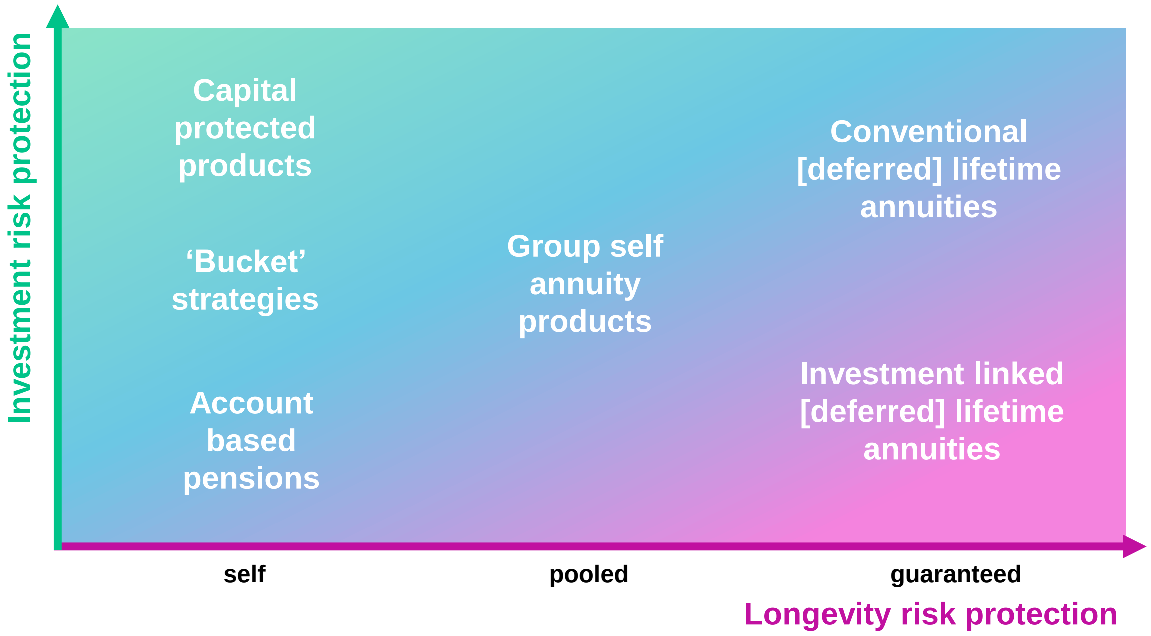
- Strategy 1 provides an income that is expected to be higher, but could be lower, than strategy 2;
- Strategy 1 provides a varying level of income, whereas strategy 2 provides a stable level; and
- Strategy 1 could run out while the retiree is still alive, whereas strategy 2 provides income for life.

The purpose of this paper is to survey and examine the properties of measuring tools or 'metrics' which allow these different properties of retirement strategies to be captured. A broader range, and better understanding of, retirement metrics will allow easier and more accurate comparison between potential strategies.

Retirement income products

The strategies in the example above (reflecting the income streams from an account-based pension and a lifetime annuity respectively) are of course only a sample of the range of retirement income solutions available to superannuation funds and their retirees. One way of viewing the range of strategies is to consider the extent to which they offer protection against investment and longevity risk. A snapshot of potential products across each of the risk dimensions is shown below.

Figure 2: The spectrum of retirement income products/strategies



Most, if not all, of these products are available in the Australian market. There is a growing awareness of the lack of longevity protection afforded by the most popular retirement product in Australia (the account-based pension) and hence of the need to enhance the take-up of strategies which include some longevity protection (whether via pooling or guarantee).

Combinations of the above products can be used to develop the proposed Comprehensive Income Products for Retirement (CIPR) which are intended to balance the need for income provision, risk protection and flexibility. With a requirement for superannuation fund trustees to prepare a Retirement Income Covenant, including offering a CIPR, now clearly on the horizon, we expect interest in building better retirement solutions to continue to grow, and as a result further product innovation to occur.

The challenge of comparing strategies

Traditional evaluation measures used by the investment industry such as expected return, volatility or the various ratios based on these, focus on *portfolio returns* rather than investment outcomes - or more specifically, *retirement outcomes*. In general, such tools are not useful in comparing and evaluating these more complex products. Nonetheless attempts are sometimes made to assess products using these familiar tools, resulting in misleading comparisons and potentially inappropriate product selection.

These difficulties are not well understood even within the investment and financial planning industry. The lack of appreciation is encapsulated in comments like 'annuities don't produce an adequate rate of return for retirees'. Implicitly, a returns-focused metric (rate of return) is being applied here to a product whose primary offering is risk (i.e. longevity) protection.

Within what can be a complex topic, the ambitions of this paper are modest. We proceed as follows:

- We first introduce a number of metrics available to compare retirement products/strategies;
- We select a number of retirement strategies to examine, representing a sample of the product spectrum described in Figure 2;
- We set up a modelling framework to generate outcomes under each strategy (stochastically) for a specified strawman retiree, and apply the metrics introduced to each; and
- We compare and comment on the features of each metric as revealed by this analysis, and point to circumstances where each might be a suitable metric for use in developing a retirement strategy.

Section 2: Metrics considered

A review of academic literature and industry practice reveals a wide range of metrics used to compare retirement income strategies. We have classified the metrics examined in this paper into three categories:

'Entry level' metrics	Proportion measure metrics	Utility-based metrics
■ Probability of ruin ■ Average age at ruin ■ Probability of inadequacy ■ Duration and depth of income misses	■ Net present value of total retirement income / Money's worth ■ Desired Income Attainability ■ Goodness of Fit Index (GOFI)	■ Member's Default Utility Function (MDUF): ■ Risk-adjusted Income ■ MDUF Score

Clearly, additional metrics exist or could be developed. Further, there are variations on virtually all of the above metrics which can be considered.

Features of retirement income metrics

There is a range of differentiating features of retirement income metrics:

- The form the metric takes (e.g. a probability, a dollar value ($), or a proportion (%));
- Whether the metric assesses the strategy against a particular goal or 'target' income.

  A 'target income' approach assumes the retiree's objective is to draw an income (including age pension and any guaranteed income) equal to a defined target income each year. Nominating a target income to apply across strategies with varying income patterns can be a useful approach to assess alternatives on a consistent basis.
- Whether the metric makes any allowance for liquidity during retirement or a bequest upon death.

We summarise these features for each of the metrics considered in this paper below:

Metric Category	Metric	Form of output	Allows for target income?	Allows for bequest motive?
‘Entry level’	Probability and average age at ruin	probability	✗	✗
	Probability of income inadequacy	probability	✓	✗
	Depth and duration of income misses	$ and time horizon	✓	✗
Proportion Measures	NPV(total retirement income) / Money’s worth	$ / %	✗	✓
	Desired income attainability	%	✓	✗
	Goodness of Fit index	% between 0 and 100%	✓	✗
Utility-based	Risk-adjusted Income	numeric	✗	✗
	MDUF Score	numeric	✗	✓

In sections 4 - 6 we compare modelling results for the metrics in each category above, and select a ‘best in class’ within each category.

Section 3: Modelling approach

The approach used to assess retirement income metrics is as follows:

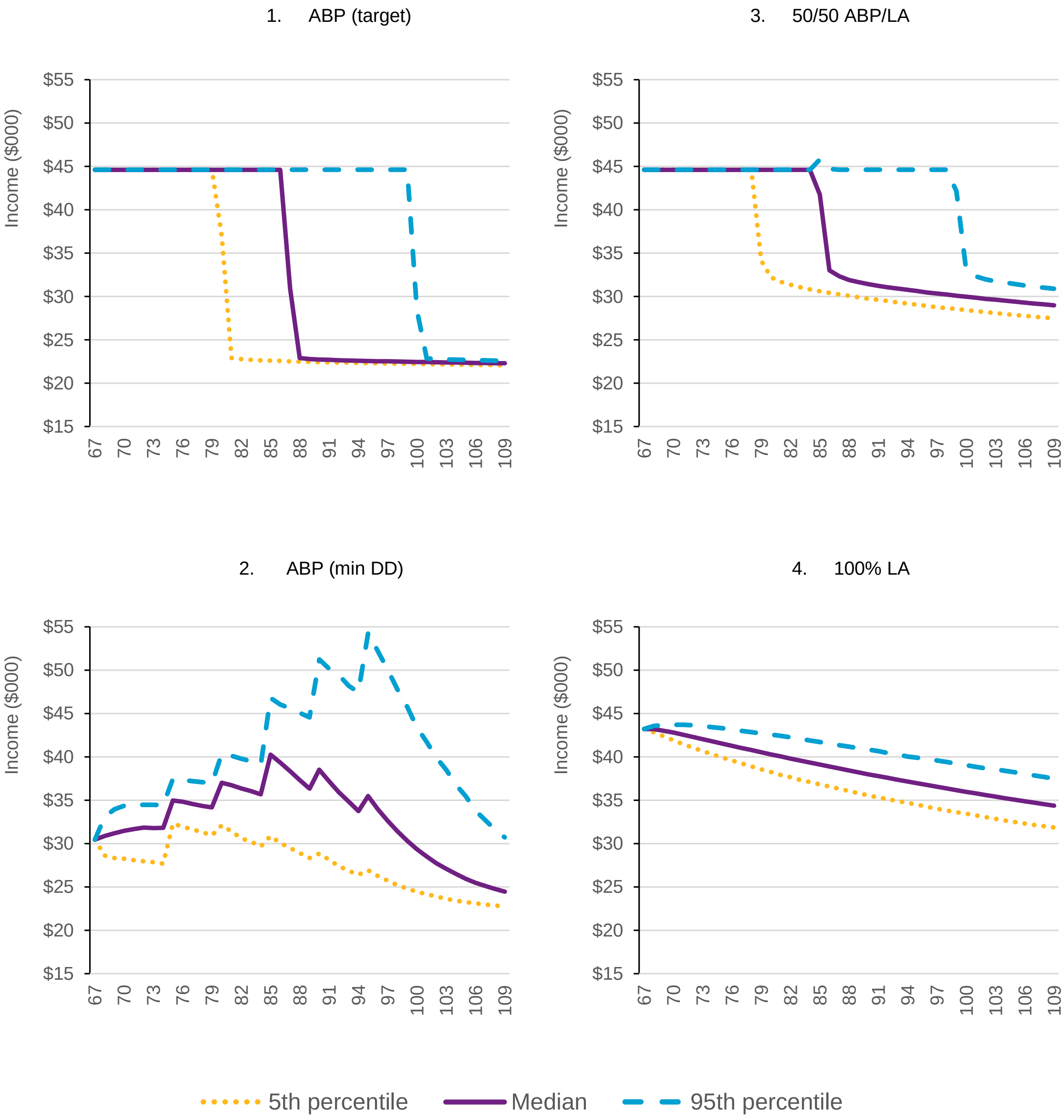
- We consider a ‘strawman’ retiree aged 67 with an initial retirement balance of $450,000.
- The income available to the retiree is made up of a combination of all or some of:
  - the amount drawn down from any account-based pension;
  - The prescribed amount of income produced from any lifetime annuity; and
  - The government age pension for which they are eligible subject to means testing of assets and income.
- Four retirement income strategies are modelled:

Strategy	Description
1. ABP (target)	100% investment in an account-based pension; retiree drawdowns are such that the aggregate income (including age pension) is a level real income expected to last to age 90
2. ABP (min DD)	100% investment in an account-based pension; retiree drawdowns are in line with the legislated minimum drawdown rates
3. 50/50 ABP/LA	50% investment in an account-based pension and 50% in a lifetime annuity; retiree drawdowns are such that the aggregate income (including annuity income and age pension) is a level real income expected to last to age 90
4. 100% LA	100% investment in a lifetime annuity

Further details of the modelled retirement strategies can be found in the Appendix.

The 90% confidence interval of the total income produced by each strategy from the stochastic modelling is set out below.

Figure 3: 90% Confidence interval of total income from modelled retirement strategies



The strawman and strategies chosen have been calibrated so that no strategy obviously dominates the others. This ensures that the metrics considered can be compared meaningfully and judged by their ability to recognise the favourable features of the strategies.

Section 4: ‘Entry level’ Metrics

Metrics considered:	**Probability of / age at ruin** **Probability of inadequacy** **Duration and depth of income misses**
Best in Class:	**Duration and depth of income misses**

In this category we examine some of the basic metrics used in retirement income modelling. These (as with the metrics considered in following sections) rely on a stochastic modelling framework – for example, the probability-based metrics are determined as the proportion of simulated scenarios where the relevant outcome (ruin or inadequacy) occurs.

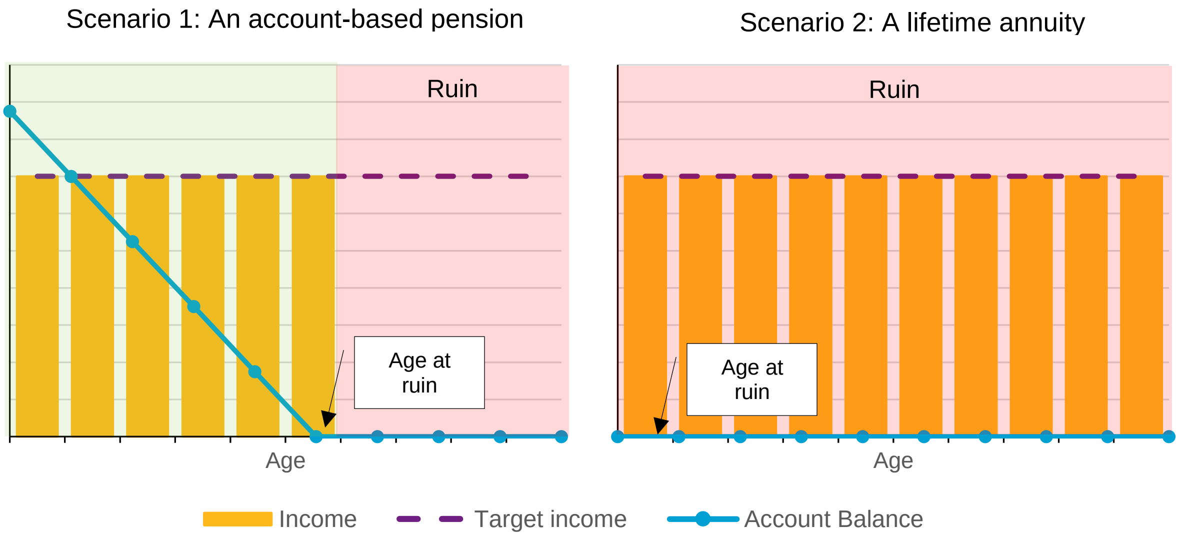
Probability of / age at ruin

Concept

The concept of ‘ruin’ refers to exhaustion of a retiree’s liquid sources of income. For strategies which include an account-based pension, ruin occurs when the account-based pension balance reaches zero and is no longer available to be drawn upon. The probability of ruin metric is the likelihood that a retiree’s liquid assets run out while they remain alive, or before a prescribed age. The age at ruin metric presents the distribution of ages at which the retiree’s liquid assets have reached zero.

For our strawman retiree, the possibility of ruin occurring exists only for strategies that contain a liquid component in the form of the account-based pension. As illustrated below, we can observe a point at which ruin occurs in the strategy containing an account-based pension. However, in the case where all post-retirement assets are invested in a lifetime annuity, ruin occurs immediately as all liquid sources of income have been exhausted at time of purchase.

Figure 4: Illustration of probability of ruin and age at ruin



Calculation

For N stochastic simulations and a retiree lifespan of t years, the probability of ruin is:

$$\Pr_t(ruin) = \frac{\sum_{n=1}^{N} I_n(t)}{N} \quad where\ I_n(t) = \begin{cases} 1, & LB_{t,n} = 0 \\ 0, & LB_{t,n} \geq 0 \end{cases}$$

Where $LB_{t,n}$ is the retiree's liquid balance at time t for simulation n.

Probability of ruin can be assessed to a particular age, or mortality weighted over all future possible retirement horizons $t \in [1, T]$. Below we show results assuming a time horizon to age 90 and over the whole of retirement on a mortality weighted basis as follows:

$$\Pr(ruin) = \sum_{t=1}^{T} \left(\Pr_t(ruin) \times \ {}_t p_x \ \times q_{x+t} \right)$$

The age at ruin is calculated as the age at which the retiree's account-based pension balance reaches zero. For each simulation within a retiree timespan of T, this is merely:

$$Age(ruin) = t\ such\ that\ LB_t = 0, 0 \leq t \leq T$$

We are then able to observe the distribution of ages when ruin occurs. The usual metrics that would be presented based on this distribution are the median or average/expected age at ruin. Below we also show the tail (5th percentile) value of this metric.

Comment

Probability of ruin and age at ruin metrics work well for fully liquid strategies. In such cases, exhaustion of the liquid account coincides with the cessation of any income (other than age pension), and so these metrics fully measure the capacity of the strategy to sustain income.

Conversely, the probability of ruin and age at ruin metrics are of little value when non-liquid products (e.g. annuities) are included in the retirement strategy, as they fail to recognise that the retiree still has retirement resources even when liquid assets are exhausted. The metrics also cannot provide information about the adequacy of any income provided, only whether or not liquid income is available. This is shown in the annuity example above, where ruin occurs at the start of retirement regardless of the presence of income throughout retirement. The next metric attempts to address this issue.

Probability of inadequacy

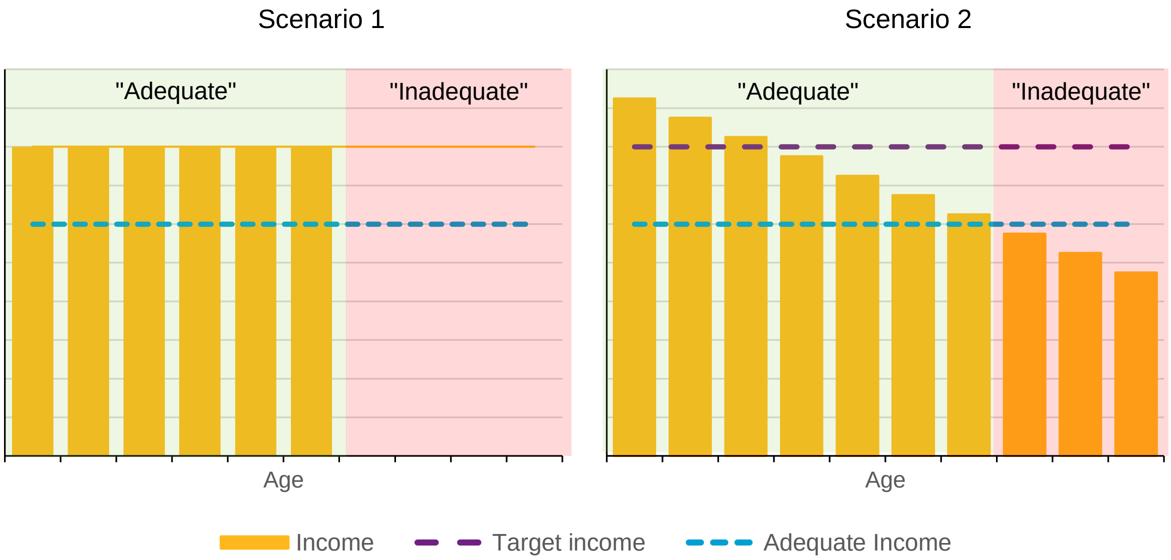
Concept

We extend our focus from looking at the depletion of account balances in isolation (ruin) to incorporating a view of whether the income generated by a strategy is sufficient to meet some pre-determined level of adequacy (in this context we might view ‘adequacy’ as a level of income to cover basic subsistence). Reaching ‘inadequacy’ refers to the point at which a strategy’s generated income drops below the defined ‘adequate’ level.

The probability of inadequacy is the likelihood of a strategy’s income falling below the ‘adequate’ level while the retiree remains alive, or before a prescribed age.

For an account-based strategy with a target income drawdown approach, ‘ruin’ and ‘inadequacy’ represent the same event, as illustrated in scenario 1 below. In these cases the ruin and inadequacy based metrics are identical. However, more generally ‘inadequacy’ can occur even when a strategy has not reached ‘ruin’, as illustrated in scenario 2 below where the retiree receives an income from a liquid source throughout their retirement but the level of income falls to below ‘adequate’ prior to the account being exhausted.

Figure 5: Illustration of probability of inadequacy



Calculation

For N stochastic simulations and a retiree lifespan of t years, the probability of inadequacy is:

$$\Pr_t(inadequacy) = \frac{\sum_{n=1}^{N} I_n(t)}{N} \; where \; I_n(t) = \begin{cases} 1, & shortfall_{t,n} > 0 \\ 0, & shortfall_{t,n} = 0 \end{cases}$$

Where

- $shortfall_t = max\left(AI_t - \ income_{t,n}, 0\right);$
- $income_{t,n}$ = retiree income at time t for simulation n; and
- AI_t = 'adequate' income at time t.

Probability of inadequacy can be assessed to a particular age, or mortality weighted over all future possible time horizons $t \in [1, T]$. Below we show results assuming a time horizon to age 90 and over the whole of retirement on a mortality weighted basis as follows:

$$\Pr(inadequacy) = \sum_{t=1}^{T} \left(\Pr_t(inadequacy) \times \ {}_t p_x \ \times q_{x+t} \right)$$

Comment

The probability of inadequacy metric improves upon the probability of ruin and age at ruin metrics in that it directly addresses whether a retirement strategy provides the retiree with adequate income. Also, it is applicable for all retirement strategies, not just wholly account-based pension strategies.

However, as with ruin, the probability of inadequacy metric suffers from the shortcoming that it fails to measure the extent of any income shortfalls, nor the number of years of income inadequacy where these periods are intermittent. We attempt to address these issues with the next metric.

Duration and depth of income misses

Concept

The duration of income misses metric is the *number of years* in retirement where the retiree’s income is below a defined target income level. Similarly, the depth of income misses metric provides the *average amount* by which the retiree’s income is below the target level in those years where it falls short.

In the illustration below, the duration of income misses is the number of shortfalls to target income in retirement, as shown by the blue bars in the charts. The depth of income misses is the average size of the shortfalls, which is the total dollar value of the blue bars divided by the number of blue bars.

Figure 6: Illustration of duration and depth of income misses

Calculation

For a retiree lifespan t, the metric calculations are:

$$Duration\_income\_misses_{t,n} = \sum_{i=1}^{t} I_n(i) \;\; where \;\; I_n(i) = \begin{cases} 1, & TI_i - income_{i,n} \geq 0 \\ 0, & TI_i - income_{i,n} = 0 \end{cases}$$

$$Depth\_income\_misses_{t,n} = \frac{\sum_{i=1}^{t} \max(TI_{i,n} - income_{i,n}, 0)}{Duration\_income\_misses_{t,n}}$$

Where

- $income_{i,n}$ = retiree income for simulation n at time $i \in [0, t]$; and
- TI_i = target income at time $i \in [0, t]$.

Generating these metrics over the stochastic simulations, we are then able to observe the distribution of each. The usual metrics that would be presented based on these distributions are the median or average/expected duration and depth of income misses. In the results below we also show the tail (5th percentile) value of these metrics. In the table below, we have presented results to age 90.

Comment

The duration and depth metrics encompass whether income is being provided; how that income compares to a target; and how long that target is or isn't met. As such it is a broad and comprehensive metric pair suitable for comparing strategies with and without a liquid component.

Comparison

The table below provides the stochastic results for each metric in this category for each of the four products chosen:

Metric Name		ABP (target)	ABP (min DD)	50/50 ABP/LA	100% LA
Probability of ruin	To age 90	59%	0%	68%	100%
	Mortality weighted	48%	0%	54%	100%
Age at ruin	Median	89	109	88	67
	5th percentile	82	109	81	67
Probability of inadequacy	To age 90	59%	64%	9%	0%
	Mortality weighted	48%	64%	13%	0%
Duration of income misses (to age 90)	Median	8	24	8	24
	5th percentile	23	24	22	24
Depth of income misses (to age 90)	Median	$17,038	$9,940	$11,456	$3,935
	5th percentile	$21,603	$13,086	$13,583	$5,461

Our observations on these modelling results are:

- The ABP (min DD) strategy has a zero probability of ruin as, by definition, the retiree draws a percentage (less than 100%) of the account balance so that it is never exhausted. On the other hand, the 100% LA strategy results in 'ruin' occurring immediately at time of purchase, as the account is exhausted (the income the annuity generates is ignored under this metric). The probability of ruin results for the other two strategies lie between 0% and 100% and are broadly informative - albeit with the weakness that the ruin probability does not give credit to the ongoing annuity income under the 50/50 ABP/LA strategy.

- The age at ruin metric reflects a similar picture to probability of ruin: the ABP (min DD) and 100% LA strategies have age at ruin metrics at the extremes of the retirement period for the same reasons as explained above. Age at ruin is an informative metric for the other strategies, which have broadly similar ages at ruin at the median and 5th percentile levels.

- While ‘ruin’ and ‘inadequacy’ (of income) represent the same event for the ABP (target) strategy, this is not the case for the other strategies where income can fall below the defined ‘adequacy’ level even where the account-based pension is not exhausted. For example, the 100% LA strategy provides adequate income (at a defined level of $30,000) across the whole of retirement despite the account being ‘exhausted’. Similarly, a partial investment in annuities is able to provide greater certainty of achieving adequacy over the other strategies.

- The duration and depth of income misses metrics are best looked at as a pair. Although the 100% LA strategy has the most incidences of falling below target (duration), the severity of the shortfalls (depth) is lowest of all the strategies. At the other end of the spectrum, the ABP (target) strategy has fewer instances of falling below target, but when it does fall below target, this coincides with the account balance having been exhausted and the retiree receiving age pension only going forward, resulting in the highest depth of income misses. Although the ABP (min DD) will generate income that is often below target, the continued presence of any income at all (as the balance is never exhausted) reduces the depth of income misses. The 50/50 ABP/LA strategy is similar to the ABP (target) strategy as to how long it is below target but the presence of the annuity income dampens the depth of the shortfalls when they occur. The additional information provided by the 2 parts to this metric means that it can better reflect the features of the diverse strategies considered.

Best in class: Duration and depth of income misses

When examined against the desirable features of a retirement income metric, we judge that the duration and depth of income misses pairing best captures the relevant features of the strategies tested. This metric pair can allow the assessment of retirement strategies by reference to a retiree’s preference between the frequency of shortfalls they are willing to experience, and the severity of shortfalls they are willing to tolerate when they occur.

Section 5: Proportion measures

Metrics considered:	**NPV lifetime income / Money's worth** **Desired income attainability** **Goodness of Fit Index (GOFI)**
Best in Class:	**Goodness of Fit Index (GOFI)**

In this category we examine 'proportion measures', by which we mean metrics which represent the retirement outcome as a proportion (or percentage) of a benchmark outcome (e.g. achievement of the target income), or of the initial purchase price.

We consider three potential metrics in this category:

Net present value (NPV) of lifetime retirement income / Money's worth

Concept

The net present value of lifetime retirement income metric is the present value of the retiree's total income throughout retirement, including any bequest available at death.

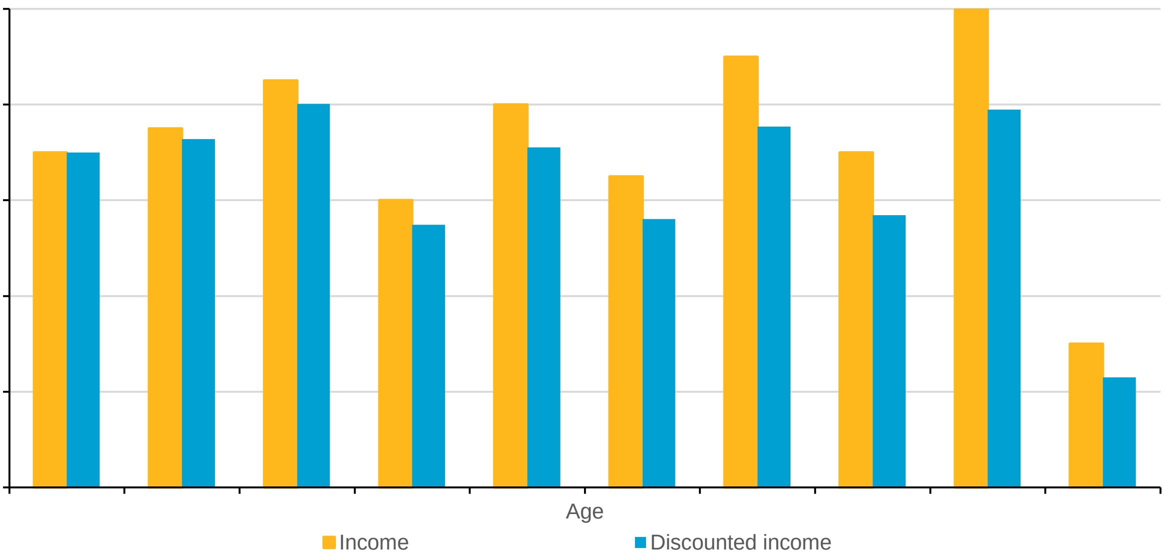
An extension: Money's worth

A variant on the traditional NPV is the 'Money's worth' metric, which is simply a scaled version of the NPV by expressing the NPV as a percentage of the retiree's initial retirement balance.

The Money's worth metric effectively illustrates whether, for a given time horizon, a strategy gives the retiree more or less value in income than their initial investment. A Money's worth value below 100% indicates that the purchase price exceeds the value received. Conversely, a value above 100% suggests that the strategy represents 'good value' for the retiree. In current practice, the Money's worth metric is primarily used to measure the value of annuities, but is applicable to any retirement strategy.

The chart below illustrates the NPV calculation: expected income cash flows (and any bequest payments) are discounted and summed to produce the metric value.

Figure 7: Illustration of NPV lifetime retirement income / Money's worth



Calculation

The net present value of total retirement income over a horizon period is the summation of the discounted values of each year of retirement income including the remaining liquid assets available at the end of the horizon period. For retirement timespan t, the NPV of lifetime income can be presented as:

$$NPV(lifetime\ income_t) = \sum_{i=1}^{t}\left(income_{i,n} \times (1+r_{i,n})^{0.5} \times \prod_{i=1}^{t}(1+r_{i,n})^{-1}\right) + liquid\_assets_t \times \prod_{i=1}^{t}(1+r_{i,n})^{-1}$$

Where, for simulation n,

- $income_{i,n}$ = the income from the strategy at time i;
- $liquid\_assets_t$ = the amount remaining in liquid assets at time t available for a bequest upon death; and
- $r_{i,n}$ = the discount rate applicable for year $i \in [0, t]$.

The Money's worth metric over the horizon period is calculated as the ratio of the net present value of lifetime income to the starting retirement balance at time 0 = RB_0:

$$MoneyWorth_t = \frac{NPV(lifetime\ income_t)}{RB_0}$$

The most common usage of this metric is the actuarial present value with mortality-weighting over the whole retirement period T as follows:

$$MoneyWorth = \sum_{t=1}^{T} \left(MoneyWorth_t \times {}_tp_x \times q_{x+t} \right)$$

Comment

The NPV/Money's worth metric reflects the aggregate amount of income (including bequest) received over retirement. However it does not benchmark against a target outcome, nor does it directly assess year-on-year income adequacy.

Desired income attainability

Concept

The Desired Income Attainability metric gives the proportion of a retiree's total desired/target retirement income needs met throughout their retirement. This metric is suited to retirement strategies where the income is continuous and is always at or below the target income level, but the metric is misleading where the retiree receives income which may unavoidably exceed this level.

As illustrated below, the metric consists of adding up the total income (all of the yellow bars), adding up the total desired income (all of the green bars), and dividing the first number by the second. Clearly, for the annuity example (scenario 2), this can potentially yield a result above 100% despite some years of income inadequacy.

Figure 8: Illustration of Desired Income Attainability

Calculation

The Desired Income Attainability figure is calculated as the ratio of the summed total of all retirement income received to the summed total of target income throughout retirement. Mathematically, for retirement timespan t this is calculated as:

$$DIA_t = \frac{\sum_{i=1}^{t}(income_{i,n})}{\sum_{i=1}^{t} TI_i}$$

Where

- $income_{i,n}$ = retiree income for simulation n at time $i \in [0, t]$; and
- TI_i = target income at time $i \in [0, t]$.

Desired Income Attainability can be assessed to a particular age, or mortality weighted over all future possible time horizons $t \in [1, T]$. The mortality weighted formula is

$$DIA = \sum_{t=1}^{T} \left(DIA_t \times {}_t p_x \times q_{x+t}\right)$$

Below we show Desired Income Attainability results evaluated to age 90.

Comment

The Desired Income Attainability metric compares a total desired income throughout the retirement period to an aggregate of all retirement income achieved.

Hence, it provides for benchmarking against a desired or target outcome. However, the metric has the drawbacks of ignoring the timing of income relative to the target income level in any individual year, and arguably unnecessarily rewarding overachievement of income adequacy even if this excess comes at the expense of income shortfalls in other years.

While the Desired Income Attainability metric improves on the Money's worth metric by allowing benchmarking against target, it ignores the 'shape' or incidence of any shortfalls. We explore the impact of the 'shape' of when shortfalls occur with the next metric.

Goodness of fit index

Concept

The Goodness of Fit Index, or 'GOFI' metric is a substantially new metric that aims to measure how well a chosen strategy delivers retirement income in line with a target. The GOFI calculation involves the sum of the square of the shortfalls (measured on a year-by-year basis) to the target income and hence applies a 'square penalty' to shortfalls. The example below shows two scenarios with the same overall level of shortfall:

Figure 9: Income patterns with the same shortfall – which is the better fit?

The GOFI measure will result in a heavier penalty being applied to scenario 2, which delivers larger shortfalls over a shorter period, compared to scenario 1, which delivers smaller shortfalls over a longer period.

That is, a product or strategy which results in a large shortfall in a single year produces a lower GOFI than a strategy that delivers a smaller shortfall across a number of years, even if the aggregate shortfall is the same for each. In this way, GOFI reflects the presumed risk aversion of retirees on a year-by-year basis and effectively introduces the concept of utility of outcomes, without employing a formal utility function (as covered in the next section).

Calculation

The formula to calculate the GOFI is quite complex. For evaluation over a retirement timespan t it is:

$$GOFI_t = D_t \times \frac{A_t}{B_t}$$

The GOFI measure comprises two parts:

- The *Delivery Ratio* (D_t) reflects the ratio of the total shortfall provided by a strategy relative to the total target income up to time t. The Delivery Ratio is only concerned about the total level of shortfall and not when the shortfalls occurred.

$$D_t = 1 - \frac{\sum_{i=1}^{t} \max(TI_i - income_{i,n}, 0)}{\sum_{i=1}^{t} TI_i}$$

- The *Shortfall Sequence Ratio* (A_t/B_t) reflects, for a given Delivery Ratio, how well a particular sequence of shortfalls from a strategy performs relative to the best possible sequence of shortfalls up to time t. The Shortfall Sequence Ratio is only concerned about the sequence of the shortfalls and not the total level of the shortfalls relative to the target income.

Where

- A_t (the *Actual Squared Shortfall Ratio*) is the average of the actual squared shortfall ratios produced by the strategy up to time t

$$A_t = \frac{1}{t} \times \sum_{i=1}^{t} \{1 - \left(\frac{\max(TI_i - income_{i,n}, 0)}{TI_i}\right)^2\}$$

- B_t (the *Best Squared Shortfall Ratio*) is the average of the optimal squared shortfall ratios to time t, i.e. given a total shortfall relative to target income up to time t, this measures the best squared shortfall ratio possible based on spreading the shortfalls evenly across all years to time t.

$$B_t = \frac{1}{t} \times \sum_{i=1}^{t} \{1 - \left(\frac{\left[\frac{1}{t} \times \sum_{t=1}^{t} \max(TI_{i,n} - income_{i,n}, 0)\right]}{\left[\frac{1}{t} \times \sum_{i=1}^{t} TI_{i,n}\right]}\right)^2\} = 1 - (1 - D_t)^2$$

Where

- $income_{i,n}$ = retiree income for simulation n at time $i \in [0, t]$; and
- TI_i = target income at time $i \in [0, t]$.

For the example in Figure 9 above, where the shortfall totals are identical, we calculate GOFI for each scenario as follows to arrive at differing GOFI metric results despite identical total levels of income shortfall:

	Scenario 1	Scenario 2
Total Target Income	\$500,000	\$500,000
Total Shortfall	\$80,000	\$80,000
Delivery ratio (D)	81.0%	81.0%
Actual Squared Shortfall ratio (A)	95.0%	82.2%
Optimal Squared Shortfall ratio (B)	96.4%	96.4%
GOFI	79.8%	69.1%

GOFI can be assessed to a particular age, or mortality weighted over all future possible time horizons $t \in [1, T]$. Below we show results over the whole of retirement on a mortality weighted basis as follows:

$$GOFI = \sum_{t=1}^{T} \left(GOFI_t \times {}_{t}p_x \times q_{x+t}\right)$$

Comment

Intuitively, the GOFI can be regarded as the 'average' proportion of target income delivered over retirement allowing for downside (but not upside) differences, and with allowance for the greater utility of smaller shortfalls to target income. Conveniently, as a scaled index, GOFI lies between 0 and 100%; 100% indicates a perfect fit to the target income, 0% indicates no income (so that the shortfall equals the target income at all times).

Comparison

The table below provides the median and 5th percentile of each metric on a mortality weighted basis in this category for each of the above products:

Metric Name		**ABP (target)**	**ABP (min DD)**	**50/50 ABP/LA**	**100% LA**
NPV (Lifetime income)	Median	$1,007,192	$1,011,488	$999,428	$942,774
	5th percentile	$810,877	$791,756	$829,179	$793,208
Money's worth	Median	224%	225%	222%	210%
	5th percentile	180%	176%	184%	176%
Desired Income Attainability (to age 90)	Median	93%	78%	94%	91%
	5th percentile	79%	71%	85%	88%
GOFI	Median	92%	75%	94%	92%
	5th percentile	81%	70%	88%	89%

Our observations on these modelling results are:

- The ABP (min DD) strategy displays the highest Money's worth metric value in the median scenario as it provides an income each year and never fully exhausts the account. While under current pricing, a 100% LA strategy ranks the lowest on Money's worth, the 50/50 ABP/LA strategy displays a Money's worth metric value close to both of the ABPs in median scenarios. However, it ranks higher on 5th percentile outcomes as it contains a guaranteed portion to dampen the impact of fluctuating outcomes compared to the pure ABP strategies, which are highly contingent on investment returns.

- By contrast, the ABP (min DD) ranks last on the Desired Income Attainability metric as, although by design it ensures the account-based pension is never exhausted, it does so at the expense of achievement of the target income. While the other 3 strategies are ranked similarly by this metric at the median level, the 100% LA wins at the 5th percentile level as when investment markets are poor, the income received is protected from deviations away from target.

- The GOFI metric more heavily penalises larger shortfalls to target income during retirement. Hence we see the ABP (min DD) strategy ranking well below the other strategies as this provides less than target income at younger ages when the retiree is still likely to be alive. On the other hand, the 50/50 ABP/LA strategy ranks relatively higher on this metric, as the combination of an annuity ensuring that income never falls to age pension level, plus the account-based pension to meet target income where available, is valued by this metric, particularly at the 5th percentile level where the partial annuity cushions the impact of adverse markets.

Best in class: GOFI

We judge the Goodness of Fit Index metric to be the best performing metric in this category. The key contributing factor in this judgement is the ‘square penalty’ applied to shortfalls so that it favours strategies which maintain, as far as is possible, income close to the target relative to those where income falls to zero (or age pension) level, even for shorter periods. This is evidenced by the comparison of the GOFI results with those of the other metrics in the comparison table above. GOFI captures the diversification benefit of the 50/50 ABP/LA strategy by assigning the highest score to this strategy at both a median and 5th percentile level.

GOFI reflects the presumed relative risk aversion of retirees on a year-by-year basis, albeit on an implicit basis. The next section considers metrics which include an explicit utility function to reflect retiree preferences.

Section 6 : Utility-based measures

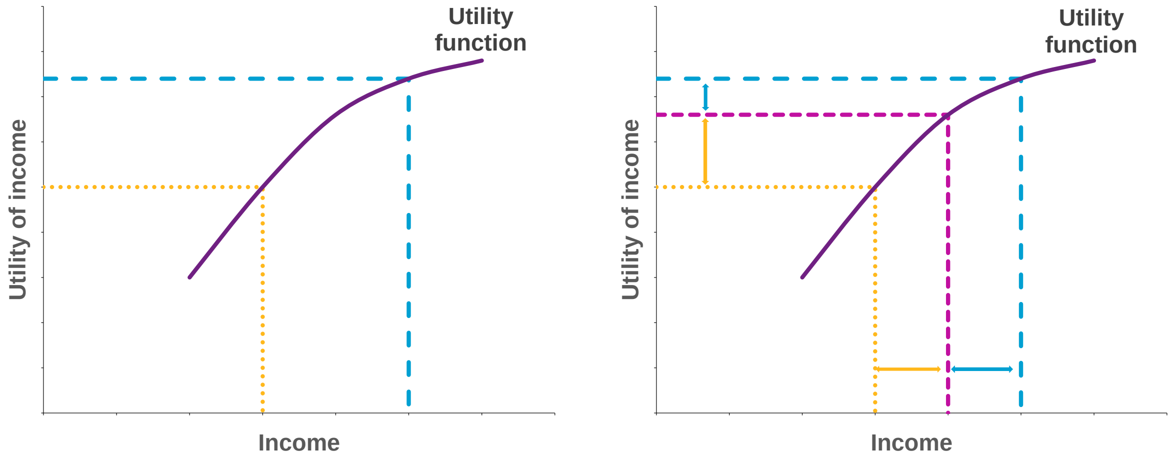
Metrics considered:	**Risk-adjusted Income** **MDUF Score**
Best in class:	**MDUF Score**

In this section we consider metrics where there is explicit inclusion of a utility function in the metrics used to evaluate strategies. A utility function is a mathematical expression of preferences. Utility-based metrics are therefore those which provide a numerical output linking the outcome of a particular retirement income strategy to a value placed upon the satisfaction ('utility') derived by the retiree from that outcome. Utility-based metrics are used widely in academic literature but, until recently, less so by industry practitioners.

However in 2015 and 2016 the team at Mine Super led by David Bell developed a framework for allowing for utility in assessing retirement options[1]. The core concept which emerged was that of a Member's Default Utility function version 1 ('MDUFv1' or simply in this paper, 'MDUF'). (By way of full disclosure, the lead author of this paper was a member of a working group which provided input and review to this project).

The core concept of utility within the MDUF framework is simple: higher income produces greater utility, but retirees are economically risk averse with the result that additional income generates less extra utility than the loss caused by an equal drop in income. This is illustrated below.

Figure 10: Income utility functions



The MDUF project developed its framework from first principles, with reference to relevant academic work and empirical studies both on the form of the utility function and the parameter values to be used. In our view, MDUF represents a significant step forward in the application of utility to retirement strategy assessment, and reflects a wide body of knowledge on this topic.

---

[1] See <http://membersdefaultutilityfunction.com.au> for a number of papers documenting the development, concepts and formulae associated with the MDUF framework.

Therefore, unlike the approach in the previous sections in this paper, we have not sought to explore and compare a range of utility measures outside the MDUF family.

The depth of the MDUF framework means that, even after restricting our scope in this way, a range of 'MDUF metrics' are available. Our approach therefore is to consider and compare different measures drawn from within the MDUF family.

Risk-adjusted Income

Concept

Risk-adjusted Income, within a utility framework, represents the constant level of income which delivers an equivalent level of income utility to the retirement strategy being evaluated. The metric is assessed by taking the expected utility function for income (the raw income utility score) and 'reversing' this to derive a Risk-adjusted Income.

Calculation-

The raw income utility score U_C, being the utility arising from income/consumption only, for a retirement timespan T is:

$$U_C = \sum_{t=0}^{T} \beta^t \ {}_tp_x \frac{income_t^{(1-\rho)}}{1-\rho}$$

Where

- ${}_tp_x$ = the probability of the retiree being alive at $x+t$ conditional on being alive at age x;
- β = the subjective utility discount factor that captures a person's time preference for near versus far-dated income; and
- ρ = the level of risk aversion.

The Risk-adjusted Income S_C is then calculated as:

$$S_C = \left[U_C \times \frac{1-\rho}{[\sum_{t=0}^{T} \beta^t \ {}_tp_x]} \right]^{\frac{1}{1-\rho}}$$

Comment

The Risk-adjusted Income metric isolates the utility derived from income received when comparing strategies with different patterns of income. The metric is a dollar figure representing a constant retiree income, which has an appealing simplicity in an area which can be complex to explain. This metric is useful for assessing the relative worth of different retirement strategies where it is agreed that no bequest motive is appropriate. We extend this metric to allow bequest motive to be incorporated in the next section.

MDUF Score

Concept

The MDUF score extends the previous metric by including in the utility function the bequest motive. MDUF score is the constant level of income (considering the trade-off of income against residual benefit) which delivers an equivalent level of expected utility to the retirement strategy being evaluated. The metric is assessed by looking at the expected utility function (the raw utility score) and ‘reversing’ this to derive a MDUF Score.

Calculation

The parameter ϕ represents the relative strength of the retiree’s residual benefit motive. The higher ϕ, the stronger the residual benefit motive i.e. the more the retiree is willing to trade-off income received today in exchange a higher residual benefit. The utility score U_B, for the residual benefit component, for a retirement timespan T is:

$$U_B = \sum_{t=0}^{T} \beta^t \times {}_{t-1|}q_x \frac{bequest_t^{(1-\rho)}}{1-\rho} \left(\frac{\phi}{1-\phi}\right)^{\rho}$$

Where ${}_{t-1|}q_x$ is the probability of dying between age $x + t - 1$ and $x + t$ conditional on being alive at age x.

This utility function isolates the residual benefit component of utility and looks at the Risk-adjusted Residual Benefit S_B, being the constant level of residual benefit which delivers an equivalent level of residual benefit utility as follows:

$$S_B = \left[U_B \times \frac{1-\rho}{\sum_{t=0}^{T} \beta^t \, {}_{t-1|}q_x \frac{\phi}{1-\phi}} \right]^{\frac{1}{1-\rho}}$$

The overall utility score U_O is the sum of the income utility and residual benefit utility i.e.

$$U_O = U_C + U_B = \sum_{t=0}^{T} \beta^t \left\{ \, {}_t p_x \frac{income_t^{1-\rho}}{1-\rho} + {}_{t-1|}q_x \frac{bequest_t^{1-\rho}}{1-\rho} \left(\frac{\phi}{1-\phi}\right)^{\rho} \right\}$$

The overall MDUF score, applying a monotonic transformation of the utility for the constant level of income (considering the trade-off against residual benefit), which delivers an equivalent level of expected utility for a simulation n of timespan T is:

$$MDUF\_Score = \left[U_O \times \frac{1-\rho}{\left[\sum_{t=0}^{T} \beta^t \left\{ \, {}_t p_x + {}_{t-1|}q_x \frac{\phi}{1-\phi} \right\}\right]} \right]^{\frac{1}{1-\rho}}$$

Comment

Allowance for a bequest motive in retirement utility metrics is controversial. The working group panel for MDUF decided after some discussion that the bequest motive should be given weight in the MDUF – in particular, reversion to a surviving spouse. However there are applications where allowance for bequest may not be considered critical or appropriate. For example, if a fund has knowledge of its retirees that suggests that spouses generally have separate, adequate, provision for retirement, then inclusion of a bequest motive may in fact result in ‘reversionary overprovision’ and be detrimental to the primary objective of income during the member’s retirement.

The MDUF Score is an extension of the Risk-Adjusted Income which allows for bequest motive. If MDUF Score is used, then setting the bequest motive ϕ to 0 in any particular case produces the Risk-adjusted Income.

Comparison

The table below provides the median and 5th percentile of the MDUF score with and without the bequest motive:

Metric Name		ABP (target)	ABP (min DD)	50/50 ABP/LA	100% LA
		No bequest motive (ϕ = 0)			
Utility (raw score)	Mean	-1.72E-31	-7.64E-32	-2.90E-32	-1.74E-32
	5th percentile	-3.71E-31	-1.18E-31	-5.29E-32	-2.33E-32
Risk-adjusted income	Mean	$30,881	$32,764	$38,201	$40,333
	5th percentile	$25,967	$30,573	$34,294	$38,557
Risk-adjusted bequest	Mean	N/A	N/A	N/A	N/A
	5th percentile	N/A	N/A	N/A	N/A
MDUF Score	Mean	$30,881	$32,764	$38,201	$40,333
	5th percentile	$25,967	$30,573	$34,294	$38,557
		Bequest motive (ϕ = 0.83)			
Utility (raw score)	Mean	-9.53E-26	-3.64E-26	-1.73E-25	-6.67E-25
	5th percentile	-1.09E-25	-8.02E-26	-1.89E-25	-8.61E-25
Risk-adjusted income	Mean	$30,881	$32,764	$38,201	$40,333
	5th percentile	$25,967	$30,573	$34,294	$38,557
Risk-adjusted bequest	Mean	$4,740	$6,803	$4,346	$3,552
	5th percentile	$4,589	$4,797	$4,246	$3,418
MDUF Score	Mean	$5,080	$7,241	$4,658	$3,808
	5th percentile	$4,921	$5,144	$4,552	$3,665

Our observations are as follows:

- The 100% LA strategy has the highest MDUF score when the retiree does not have a preference for a bequest motive as it provides a level of certain income. This is followed in rank by the 50/50 ABP/LA strategy, which contains a guaranteed income component, then the ABP (min DD) strategy, under which income lasts for life by construction. The ABP (target) is ranked last in this case.
- Including a (fairly high) preference for a residual benefit swings the MDUF Score in favour of the ABP (min DD) strategy, which provides the opportunity for income and bequest at each point in time. This is followed by the ABP (target), which also has a higher prospect for bequest, followed by the 50/50 ABP/LA strategy where the bequest potential is diluted by the annuity component. In last place is the 100% LA strategy, which provides no bequest[1].

Best-in-class: MDUF Score

Both MDUF metrics considered are ‘income scaled’ (i.e. the metric value is in annual income units), and hence are easier to work with than the raw score expected utility which often produces difficult to work with numeric values as showing in the table of results.

The Risk-adjusted income has the appealing quality of representing an ‘equivalent’ annual income to the retirement strategy being considered. The same is not true of the MDUF Score due to the inclusion of the bequest motive. Nonetheless, as Risk-adjusted income is effectively a special case of the MDUF Score, for the purposes of this paper we rank the latter as ‘best in class’ as it allows full flexibility in the inclusion (or otherwise) of bequest motive as may be appropriate to the application.

---

[1] In accordance with the working group’s calculations, we apply a floor on the residual benefit equal to the age pension entitlement. Hence an MDUF score with a $\phi = 0$ will be different to that where $\phi > 0$ even when the bequest amount is zero.

Section 7: Comparison of ‘best in class’ metrics

In this section we collect the modelling results of each of our ‘best in class’ metrics and view them alongside each other.

As the units of the various metrics are different, which makes comparison difficult, we have re-presented the earlier modelling results to show the ranking (1 - 4) of the strategies under each metric rather than the metric value. For ease of comparison, rankings are colour-coded as indicated below.

1 = Highest	2	3	4 = Lowest

Metric Name		ABP (target)	ABP (min DD)	50/50 ABP/LA	100% LA
Depth of income misses (to age 90)	Median	4	2	3	1
	5th percentile	4	2	3	1
Duration of income misses (to age 90)	Median	1	4	1	4
	5th percentile	3	4	1	4
GOFI	Median	2	4	1	2
	5th percentile	3	4	2	1
MDUF Score (with bequest motive; $\phi = 0.83$)	Mean	2	1	3	4
	5th percentile	2	1	3	4

The pattern of rankings demonstrates that, even among the (so-called) ‘best in class’ metrics, rankings differ across the different metrics, and in particular no single strategy ranks highest or lowest on all metrics.

This is to be expected; as the discussion of the individual metrics demonstrates, metrics will give different weights to different features of retirement strategies.

Our final observations on the above results, and on the use of retirement metrics generally, are:

- Both the duration and depth of income misses metrics and GOFI assess the strategies against a goal or target income. While metrics which include a target allow more explicit assessment against an income goal, any comparison of strategies will itself be dependent on the choice of that target. It will be important, when using target based strategies, to consider results over a range of target incomes includes those applicable for ‘high’ and ‘low’ income retirees to ensure the needs of different cohorts of retirees are considered.

- Neither of the above metrics however include an allowance for bequest on death. The MDUF Score, alone of these metrics, allows for bequest, as is reflected in the lowest ranking of the 100% LA strategy by this metric. Again, in practice it will be important to test strategies using different parameter values – if the bequest motive is removed the rankings are quite different as the results of section 6 showed.

- While the above results show a high degree of correlation of median/mean and 5th percentile results, this is not always the case – some strategies will perform differently in the ‘tail’ compared to the central or expected results, and a range of metrics may be needed to fully explore tail outcomes. Tail results are particularly important in the context of developing a ‘soft default’ retirement strategy, as – arguably – a lower level of engagement by retirees implies that funds have a greater responsibility to ensure the retirement strategy produces acceptable outcomes in adverse markets.

Appendix: Modelling framework

To compare the various metrics in this paper, we modelled the retirement period of a strawman across a retirement period of 42 years, from retirement age 67 to 109. We then calculated the metric output of this projection across 5,000 potential economic outcomes to establish distributions of the output of each metric.

Modelling basis

The details of the strawman are provided below:

Aspect	Input
Gender	Male
Retirement balance ($)	$450,000
Sustainable income target age	90
Retirement age	67
Part of couple for age pension purposes?	No
Homeowner for age pension purposes?	Yes
Target income	$44,621
Adequate income	$30,000

For the purposes of the paper, the modelled strawman is assumed to experience mortality consistent with the Australian Life Tables (ALT) 2010-2012, adjusted for mortality improvement on the 25 year basis since 2011 onwards.

It is assumed for the purposes of the modelling in this paper that the strawman is a single homeowner. The age pension rules[1] and rates[2] employed in the modelling are provided below:

Age pension payment rates	Rules as applied to strawman
Full Age Pension	$814 per fortnight
Pension Supplement	$66.30 per fortnight
Minimum Pension Supplement	$35.70 per fortnight
Clean Energy Supplement	$14.10 per fortnight
Age pension income test rules	
Income test full age pension threshold (ITFAPT)	$168 per fortnight
Rate of Reduction	$0.50 per $1 over the ITFAPT

---

[1] The age pension rules employed in the paper for the lifetime annuity are those outlined in the most recent Department of Social Services position paper ‘Means Test Rules for Lifetime Retirement Income Streams’ (7 February 2018) <https://engage.dss.gov.au/wp-content/uploads/2018/02/updated_position_paper_means_test_rules_for_lifetime_retirement_income_streams_february_2018.pdf>.

[2] The age pension rates employed in the paper are those current as at 20 September 2017 – see <https://www.humanservices.gov.au/individuals/enablers/payment-rates-age-pension/39901> for the latest rates.

Age pension payment rates	Rules as applied to strawman
Age pension assets test rules	
Assets test full age pension threshold: Homeowners (ATFAPT)	$253,750
Rate of reduction for assets exceeding ATFAPT	$3 per $1,000 over the ATFAPT
Deeming rules	
Financial investments threshold	$50,200
Assumed annual return on assets up to threshold	1.75% pa
Assumed annual return on assets above threshold	3.5% pa

Economic variables modelled

For ease of explanation and modelling, the paper assumes that the strategic investment allocation of any financial product is able to be modelled by reference to its investment in two generic categories of asset class: 'growth' assets (consisting of shares, property, infrastructure; and growth alternatives) and 'defensive' assets (consisting of fixed interest and cash investments). The summary statistics for the economic variables employed to derive the modelling results presented in this paper are as follows:

Variable	Price inflation	Wage inflation	Growth asset investment return	Defensive asset investment return
Mean	2.0% pa	3.3% pa	7.6% pa	4.2% pa
Standard deviation	2.3% pa	3.2% pa	16.6% pa	4.1% pa

The economic variables were projected using the Willis Towers Watson Global Asset Model, based on its 'lower for longer' calibration as at 30 September 2017 (the latest calibration available at the time of writing). These summary statistics are based on a 42 year time period (between ages 67 and 109), the maximum of the modelling range.

The model outputs results in 'today's dollars' deflated with stochastic wage inflation. Present values are discounted at a nominal rate of 2.2% pa as a proxy for the current bond yield environment.

Retirement products

This paper assesses the performance of a variety of metrics across retirement strategies involving two main retirement income products which realistically could be available to an Australian retiree in the current economic climate:

- an account-based pension alone; and
- a 'pure' lifetime annuity with no bequest element.

The product assumptions and fee structures of each product are provided below:

Account-based pension strategic asset allocation and fee structure

The account-based pension individually and in combination with the lifetime annuity have differing fee structures. This accounts for the fact that the longevity products are theoretically equivalent to defensive assets, so the account-based pension when in combination with a longevity product should be invested in a slightly more aggressive option to increase its equivalence for comparison purposes to the account-based pension on its own.

This paper models the fee structures for representative realistic ‘balanced’ and ‘growth’ account-based pension investment options. The strategic asset allocation and fee structures of these products are provided below:

Product element	Alone as sole retirement income source	In combination with longevity products
Growth assets allocation	70%	80%
Defensive assets allocation	30%	20%
Asset-based investment fee	0.4% pa	0.5% pa
Administration flat fee	$78 pa	$78 pa
Asset-based administration fee	0.1% pa	0.1% pa
Asset-based investment maximum fee cap	$750 pa	$750 pa

These fees are calculated and applied to the account-based pension balance of the strawman throughout modelling on an annual basis at the middle of each year of retirement.

Lifetime annuity pricing structure

The lifetime annuity modelled in this paper has no explicit fees throughout retirement, but income from the product is net of fees and is calculated on the basis of price at purchase. This price is presented as annual dollars of income per $100 of annuity investment at purchase/retirement age.

The lifetime annuity pricing assumption applied within this paper is based upon the average pricing of the ‘enhanced income (no liquidity)’ product provided by Challenger (currently the largest annuity provider in Australia) over a period to April 2018[1]. The pricing provided below is that relating to full price inflation protection, the age 67 annuity price being an interpolation of the age 65 and age 70 prices available:

Age at purchase	Annual income per $100 of purchased annuity (Male)
67	6.075

---

[1] For annuity pricing purposes, we relied upon the average of enhanced income (no liquidity) full inflation protection rates between 4 September 2017 and 15 April 2018 - see <https://www.challenger.com.au/products/rates.asp> for the latest Challenger annuity prices.

Sustainable income derivation approach

Within the strawman assumptions above, a target income was calibrated for the strawman assuming the strawman is fully invested in the account-based pension earning average investment returns. The target income was calculated as the annual level of the sustainable income (including age pension entitlements) such that the strawman's account-based pension is fully exhausted by age 90.

The setting of the sustainable income in this manner means that the drawdown from the account-based pension in each year of retirement varies according to the strawman's age pension entitlements and retirement product entitlements in that year so that their sustainable income remains constant in real terms. Further, the sustainable income is subject to the legislated minimum drawdown rules. This target income is then measured against in the stochastic framework.