SOCIETY OF ACTUARIES®

Article from

Retirement 20/20 Papers

June 2018

A Middle Ground for Public Plans

Rowland Davis, FSA

In this paper I propose two specific ideas for risk-sharing programs that I think could reasonably be implemented (together or separately) in the public plan space and would represent a middle-ground approach that might ease the current tensions between participants and sponsors:

1. **A program for accumulating funds during the active employment phase, which I call the Tracker Plan.** This is a defined contribution plan, but also includes a variable contribution feature that offers support for participants when investment performance falls short of expectations.
2. **A program for paying lifetime benefits to participants after retirement.** This is based on a collective risk-sharing fund that follows specific rules for payment of a base benefit amount, as well as bonus payments when the fund earns investment returns that build some surplus in the fund. The fund is designed to be fully sustainable for future years, without any liability or cost to the sponsor.

I will illustrate how these two ideas can be evaluated using a specific quantitative evaluation framework developed as part of a recent Society of Actuaries (SOA) research project. This framework is intended to aid in the evaluation of any new retirement program designs, with a specific goal of allowing direct comparisons of key results with other proposed designs.

1. Background Thoughts

A full-scale summary of the current public plan space is beyond the scope of this paper. Other groups have good summaries available (e.g., Pension Research Center, Center for State and Local Government Excellence). I only want to characterize my own personal view of the critical issues that motivate my thinking. The public plan space in the United States currently remains dominated by defined benefit plans—unlike the private sector. In recent years these defined benefit plans have been characterized by rising cost levels, triggered by unfunded pension obligations. State and local legislative bodies have seen pressure to reduce benefit levels and bring costs down. There is often a political dimension to the tensions: current participant interests are represented by unions and (generally) the political left, while cost control efforts are more strongly supported (generally) by the political right. In many situations it seems that middle ground, or compromise, solutions are never seriously considered. One side fights to preserve the status quo, while the other side often pushes the more extreme alternative of putting public plans on a path to conventional defined contribution designs.

Obviously, standard defined benefit plans and standard defined contribution plans represent the two ends on the spectrum of who bears risk. My view is that a middle ground solution should be the target, one where risks are shared to some degree between the participants and the sponsors, and also among participants collectively. I fully recognize the difficulty of changing the focus of opposing sides in the current environment, but clearly the starting point is to develop ideas for consideration and to illustrate the value of these ideas. My hope is that over time these ideas might take root and lead to implementation in one or more visible situations. Following successful implementation in a few cases, others can feel more comfortable taking the same path. I also think that opportunities are significant for innovation in the relatively new type of public plan set up by several states, under the so-called Secure Choice initiative, that cover workers in the private sector who lack coverage under an employer plan. Currently these plans are legally constrained during the accumulation phase, but these restrictions may fade over time—and the payout phase seems perfect for implementation of new designs.

2. Evaluation Framework: General Description

Because I intend to use a specific evaluation framework in the analysis of my design ideas, this framework should be outlined first. The Retirement 20/20 initiative started several years ago with a very open conversation about how innovative ideas could be developed and implemented in the evolution of the non–Social Security portion of the U.S. retirement system. A while later, a measurement framework was developed to aid in discussions and comparisons of different ideas. However, this measurement framework was qualitative in nature, and a need was recognized that a quantitative evaluation framework should be added to the overall measurement and evaluation process. In late 2015 a research project was initiated by the SOA for this purpose, and I was the lead researcher. While the project is not yet completely final, the technical aspects of the quantitative evaluation framework (QE framework) are now developed and will be used in this paper.

The QE framework includes the following general components:

- A set of well-defined metrics that should be determined for each proposed retirement system design
- The specific methods that should be used for the calculations
- The specific assumptions that should be used for the calculations
- Separate (but connected) models for developing metrics for the accumulation phase and for the payout phase, predicated on the concept that these two parts of any retirement program are fundamentally separable (i.e., any accumulation structure may be combined with any payout structure to form a complete retirement system)

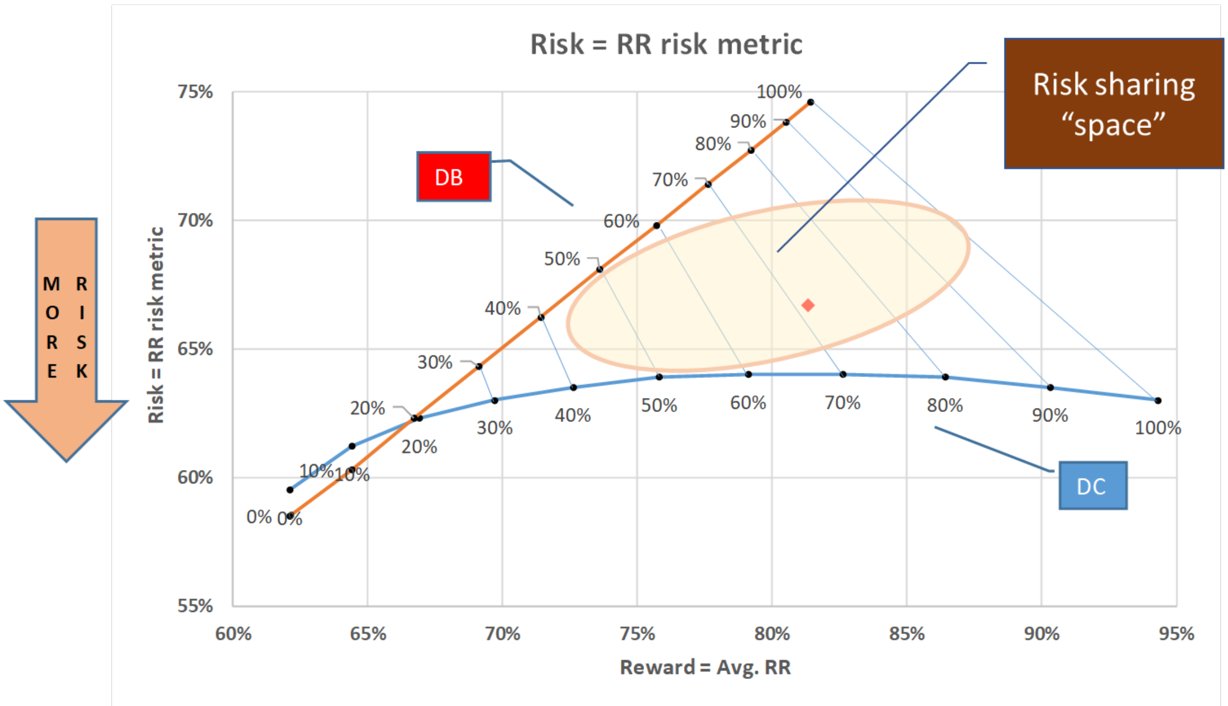
The standardized set of metrics, methods and assumptions allow direct comparison of quantitative results when analyzing different design ideas. In this paper I am not including full documentation of all the metrics, methods and assumptions, but this will be available on the SOA website when the research project is finished. In Appendix A to this paper I include some more information on the QE framework.

3. QE Framework for the Accumulation Phase

One of the goals of the QE framework is to clarify the risk-sharing nature of a proposed design—both the degree of risk sharing and the financial efficiency of the specific design. To help with this, the framework includes sets of baseline metrics for both a standard defined benefit (DB) plan and a standard defined contribution (DC) plan. Because the simplest risk-sharing design is some combination of a DB plan with a DC plan ("combo plan"), the framework can combine the baseline DB and DC metrics to obtain metrics for any combo plan. These combo plan metrics can then be used as a baseline for evaluation of any other type of risk-sharing design.

The QE framework uses a variety of metrics for the accumulation phase, but the most important ones are the benefit metric (i.e., replacement ratio) and the cost metric (i.e., cost as a level percent of pay over the full career). Graphically, these baseline metrics can be used to show what might be called the "risk sharing" space where the results for any risk-sharing design will plot. This is most easily illustrated using the benefit metric graph in Figure 1.

Figure 1. Graph Using the Benefit Metric



Here are notes to understand the graph in Figure 1:

- The metrics are based on the replacement ratio ("RR") outcomes, including Social Security, produced at age 67 from a 10% of pay contribution (DC plan) or a 10% of pay average cost (DB plan). The reward metric on the horizontal axis is the average RR produced. The risk metric on the vertical axis is the average RR for the bottom quintile of outcomes (i.e., a conditional tail risk metric).
- The blue line shows the baseline DC plan results, and the orange line shows the baseline DB plan results. In both cases, the level of investment risk increases as you move from left to right. The data point labels indicate the percentage allocation to risk assets, with the balance in fixed income.
- The space in between the two lines is the risk sharing "space." The red diamond might indicate where the results from a specific risk-sharing plan proposal plot. For any such point, there will also be some combo DB-DC arrangement that will produce the same results. The QE framework provides a tool for determining this equivalent combo plan (for this hypothetical example, it is a 20% DB + 80% DC plan with a 69% risk asset allocation for each).
- Note that while the DB plan results seem to dominate the DC plan results when we look only at the benefit metrics, the reverse is true when we look at cost metrics. In this cost space, the cost risk metric is the average cost for the highest quintile of outcomes. For any standard DC plan, this will always be just 10% of pay (i.e., no cost risk), but for any standard DB plan the cost risk

will be higher than this, depending on the level of investment risk taken. The graph in Figure 2 shows the risk sharing space plotted using cost metrics, instead of benefit metrics.

Figure 2. Graph Using the Cost Metric

This procedure opens the door to an evaluation methodology for any proposed risk-sharing design—namely, use the equivalent combo DB-DC plan as a benchmark for comparison. This will be illustrated later in this paper when I evaluate the Tracker Plan.

4. The Tracker Plan Design

At the participant level, the major goals for the Tracker Plan are to:

- Provide an automatic path for participants to follow in accumulating the assets required to meet their retirement income needs.
- Control the risk of adverse outcomes, where assets are insufficient to meet needs.

At the macro program level, the major goals are to:

- Operate the plan(s) and manage the investments efficiently, professionally and at a low cost to the participants.
- Keep employer obligations, both financial and administrative, at reasonable and manageable levels, with a known upper limit on annual cost under worst-case conditions.
- Never have any unfunded obligations.

To meet these objectives, the Tracker Plan design is structured as a defined contribution plan, but with a variable add-on contribution feature (with a fixed maximum, or cap) designed to control downside risk for participants. More specifically, the Tracker Plan is designed to push some level of additional contribution into the accumulation process whenever negative investment outcomes result in a level of accumulation that falls below some specified target level. The idea is that this dynamic support mechanism can help limit the probability of inadequate benefit levels from a DC plan, while maintaining an acceptable upper limit on the cost.

With traditional defined contribution arrangements, two of the most common criticisms are that they are too risky for participants, and that participants lack the skills and training needed to make the critical financial and investment choices required for successful outcomes. The Tracker Plan meets these problems with a primary emphasis on risk control and simplicity:

- For each participant there is a single investment vehicle that gradually decreases risk over the course of a career (i.e., the target date fund concept is used).
- There is a standard contribution pattern to follow throughout the participant's career, designed to accumulate to the required target amount at retirement. This could be either a flat rate at all ages, or an age-graded schedule. For this paper I will just analyze a flat-rate approach. Participants could be given an option to contribute at a different rate, possibly subject to some limits.
- Progress toward the target is monitored (based on a hypothetical straw man for each cohort), and adjustments are made according to a fixed set of operational rules based on tracking error. If performance is adverse and the fund is tracking below the desired target path, then additional contributions may be triggered, up to a fixed maximum add-on.

Risk control is a critical objective, and specific measures and standards are needed to determine whether the amount of risk is contained within reasonable levels. My suggested standards are that (1) with about 90% confidence, the participant will (over a full career) meet or exceed the desired target asset accumulation, and (2) for those cases where the target is not met, the shortfall can be managed with relatively painless steps, which would include working no longer than one year beyond the regular retirement date. These specific standards became my benchmark test for each design option I analyzed with the Monte Carlo simulation model that is part of the QE framework toolkit. Through an iterative process I refined each of the design parameters to achieve the risk control standard. The following subsections describe the specific Tracker Plan model that resulted from this process. There are subjective calls made along the way, but mostly these were to maintain a balance between simplicity of design and achieving improvements in the risk control outcomes.

5. Retirement Income Target

The first parameter choice is to select a target level for retirement income. For this paper, I use a target 70% income replacement ratio at age 67, inclusive of Social Security, for a worker with career earnings at median U.S. levels. This means that at age 67 the total income available from Social Security benefits plus the Tracker Plan benefits will be equal to 70% of the gross income at the time of retirement.[1] The Tracker Plan benefits are based on annuitizing, with market-priced annuity factors, the accumulated funds at age 67. The annuity includes a post-retirement increase factor of 2.5% per year, and the pricing includes a 5% loading factor to approximate group product pricing. The Social Security benefit used is based on retirement at age 67 in 2053, and this produces a 39% replacement ratio for Social Security

[1] In an earlier paper on the Tracker Plan, I used a more generous 75% replacement ratio target at age 65. The actual target used for any design should be based on goals mutually agreed to by the sponsor and representatives of the participant group. The design parameters used for the earlier paper are also somewhat different than those used in this paper.

alone. To meet the 70% overall target, the Tracker Plan benefit should therefore replace 31% of pre-retirement income. More specifically, recognizing the risk control objectives stated in the previous section, the Tracker Plan benefit should equal or exceed 31% of final pay with about a 90% probability, and should almost never fall much below 25% of final pay (a 5%–6% shortfall is about what a worker can expect to recover by working to age 68 instead of to age 67).

6. Tracker Process and Design Parameters

The truly unique feature of the Tracker Plan is a set of automatic adjustments that will help keep accumulations on the desired path toward the required target. These adjustment provisions are a key part of the risk control process, and they facilitate a sharing of risk between workers and employers. The primary adjustment occurs when investment performance is adverse and the fund is tracking below the desired target path, which triggers additional contributions, up to a fixed maximum add-on (which ensures that this remains a DC plan with no sponsor liabilities). Optional adjustment features that might also be considered are:

- If a fund has better than expected investment performance and a positive tracking error above some trigger point is created, then a reduction in investment risk could be implemented. This would operate to protect the “surplus” in the fund (and in any scenario similar to 2008–2009 would have been very effective in cushioning losses for those near to retirement).
- A positive tracking error as described previously might also trigger a decrease in contribution rates, but limited to a “claw-back” of any previous positive additional contributions. This would be a way to reduce the cost for the variable contribution feature.

In this paper I will use a design model that has only the positive additional contribution feature.

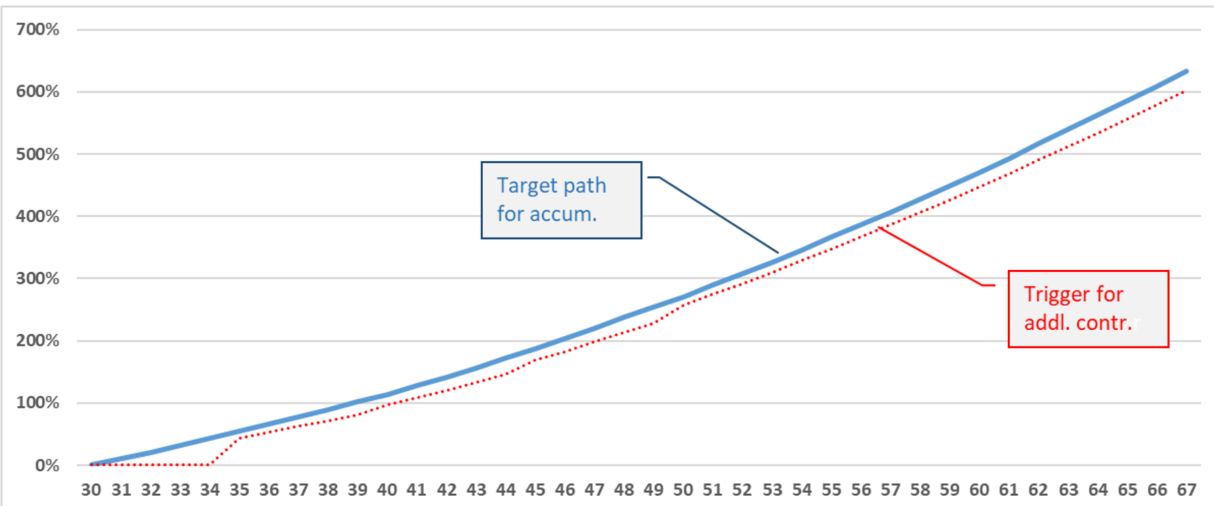
The tracking process does not need to be done at the individual participant level, as long as all plan features remain standardized. Separate tracker funds are established for each individual age cohort, or (subject to further testing) for cohorts in 3- to 5-year age groups. A hypothetical account can be tracked for each of these tracker cohort funds, based on the assumption of a worker making the scheduled contributions, and earning the investment returns actually realized by that tracker fund. The tracking error for this hypothetical account will be monitored, and on an annual basis the level of the tracking error will be used to trigger any needed automatic adjustments for all workers in that tracker fund. Within each tracker fund, workers will all be treated exactly the same way.

The next step in the design process is to develop the target accumulation path that will serve as the tracking benchmark. The contribution rate required for 90% of the scenarios to meet the 70% replacement ratio target at age 67 can be determined using assumptions about expected returns and inflation, and reflecting the uncertainty of these by using the Monte Carlo simulation model. An investment policy must be specified. The design analyzed for this paper uses a somewhat conservative target date fund for investments (the glide path for the risky asset allocation declines from 60% at age 35 to 20% at age 65). With this investment policy, a base contribution of 10.6% of pay is required to meet the replacement ratio constraints. At each age through the career the resulting accumulated values can now be calculated for all scenarios and expressed as a percent of pay. From the simulated range of pay multiples at each age, the median value is used as the tracking benchmark, or target accumulation, at each age. For each cohort tracker fund, the tracking error will be measured against this benchmark. If there is a negative tracking error that exceeds some threshold defined by the plan, then additional contributions would be triggered for all the participants in that tracker fund. The graph in Figure 3 illustrates the target path for the design parameters used in this paper. The dotted line shows the trigger level for additional contributions, which is based on an age-graded schedule of the tracking error limits (see Table 1). The tracking error threshold becomes tighter, in percentage terms, as participants approach retirement.

Table 1. Age-Graded Schedule of Tracking Error Limits

Age	*Trigger*
<35	NA
35 - 39	-20%
40 - 44	-15%
45 - 49	-10%
>=50	-5%

Figure 3. Graph of Target Path by Age for Accumulation (% of Pay)



The final design choice is how large the additional contributions should be, whenever they might be triggered. This could be defined in various ways, but I use three parameters to define the size of the additional contributions:

1. **Correction target.** The starting point for calculating any additional contribution is the tracking level, but this parameter determines whether the full tracking error is used or whether the corridor margin is subtracted first. Basically, the choice is between trying to get back to the target level itself or back to the corridor around the target. The net amount can be referred to as the "adjusted tracking error." The design choice I use for this paper is the use of the full tracking error (i.e., the goal is to move back to the target path).
2. **Correction speed.** This parameter would set the number of years, which is used to divide the "adjusted tracking error" to derive the additional contribution amount (before application of the cap). For example, if 3 years is used here, the additional contribution amount would be twice as large as compared with using 6 years—and the intended time frame for getting back on target would be correspondingly shorter. For this paper I am using a period of 5 years (i.e., the additional contribution would be equal to 20% of the tracking error).

3. **Cap on contribution adjustments.** This puts the upper limit on any additional contribution amount, which is what keeps the plan within the defined contribution family. For this paper I set the cap at 50% of the base contribution amount, or 5.3% of pay (= 50% × 10.6% base contribution rate).

This completes the basic definition of the Tracker Plan design being used for this paper. I have made design choices to illustrate how the Tracker Plan can work. Other choices for the design parameters are acceptable, and the choices made should be based on the goals and constraints specific to each situation. The QE framework includes a specific module that can be used by researchers to test alternative design options, as well as some analysis that has already been done on how results change with different design parameters.

7. Accumulation Phase Analytics

7.1 Cost

There are various ways to analyze the additional contribution feature of the Tracker Plan design just described. This section will review several of these. Probably the most important from the plan sponsor's viewpoint is the overall range of cost. The average total cost for this sample Tracker Plan is 11.5% of pay (= 10.6% base contribution [employee + employer], plus 0.9% average for the additional contributions). The maximum annual cost for any single cohort would be 15.9% of pay (= 10.6% base contribution [employee + employer] × 150%), but in any single year the employer would be funding multiple cohorts and this maximum amount would be extremely unlikely to occur for all cohorts at the same time. Table 2 is the range of total cost based on the simulation model:

Table 2. Range of Total Cost Based on Simulation Model

%tile	*Base*	*Addl.*	*Total*
99%	10.6%	3.1%	13.7%
90%	10.6%	2.1%	12.7%
75%	10.6%	1.4%	12.0%
50%	10.6%	0.7%	11.3%
25%	10.6%	0.2%	10.8%
10%	10.6%	0.0%	10.6%
1%	10.6%	0.0%	10.6%

Table 2 shows that the effective worst case produces a cost for the additional contribution feature of 3.1% of pay. The median level of cost is 0.7% of pay. There is about a 25% chance that the cost will be at, or very close to, zero.

From the participant's point of view, the focus is on how much they might benefit from the additional contribution feature. Using the sample employee from the QE framework (hired at age 30 and retired at age 67), there is about a 15% chance that they would never need any additional contributions because of favorable investment performance. Table 3 shows the distribution range for the number of years (over the 37-year career) where additional contributions would be paid, and the cumulative amount received (as % of pay):

Table 3. Distribution Range and Cumulative Amount Received for Number of Years With Additional Contributions

%tile	*# yrs.*	*% pay*
99%	22	110.6%
90%	15	77.2%
75%	10	50.9%
50%	5	25.4%
25%	2	8.3%
10%	0	0.0%
1%	0	0.0%
Mean	6.6	32.6%

The typical participant could expect additional contributions totaling between 25% and 33% of pay over their career, which is equivalent to another 2½ to 3 years' worth of contributions. In the worst quartile of investment outcomes, the additional contributions become equivalent to more than 5 years' worth of contributions.

7.2 Benefits

The Tracker Plan design is built around the goal of mitigating downside risk for the participants, in a cost-effective manner. The sample plan discussed in this paper set out very specific goals for the control of downside benefit risk, so the analysis of benefit results is focused on:

- Does the plan design meet the goals?
- How much is participant risk reduced, relative to a standard DC plan?

Table 4 shows the full range of replacement ratios for the sample Tracker Plan design and compares them with the range of results using standard DC plans with contributions set at three different levels: (1) the 10.6% base contribution rate for the Tracker Plan, (2) the 11.5% expected average contribution rate for the Tracker Plan, and (3) a 12% contribution rate equal to the 75th percentile cost for the Tracker Plan. Downside, or shortfall, risk probabilities are also shown for replacement rate targets of 70% and 65%. All results are for the 37-year career sample employee used in the QE framework.

Table 4. Replacement Ratios for Sample Tracker Plan

		--- DC Plan w/ contr. rate: ---		
%tile	*Tracker*	*10.6%*	*11.5%*	*12.0%*
99%	94.2%	93.0%	97.6%	100.2%
90%	84.8%	83.5%	87.2%	89.3%
75%	79.8%	77.7%	81.0%	82.8%
50%	75.6%	73.1%	76.0%	77.7%
25%	72.4%	69.3%	71.8%	73.2%
10%	69.8%	66.0%	68.3%	69.6%
1%	66.0%	61.3%	63.2%	64.3%
Mean	76.6%	74.0%	77.0%	78.6%
Shortfall risk:				
70% RR	11.0%	30.4%	16.2%	11.3%
65% RR	0.2%	6.8%	1.8%	1.8%

Table 4 shows a significant reduction in shortfall risk versus a DC plan with a contribution rate equal to the 11.5% average Tracker Plan cost, and even a noticeable reduction when the contribution rate is raised to the 75th percentile cost level of 12%. This illustrates the efficiency of a design like the Tracker

Plan, with a dynamic variable contribution targeted specifically to mitigate poor investment performance.

7.3 Sensitivity to Investment Policy

The results analyzed so far are based on a fairly conservative target date fund approach (60–20 glide path), which is equivalent to a flat 38% risk asset allocation (i.e., both produce the same mean DC plan accumulation value at retirement). In this section the same results will be produced using a more typical target date fund (90–50 glide path), which is equivalent to a flat 68% risk asset allocation. This will show the sensitivity of results to the level of investment risk in the allocation policy selected.

With the increased risk asset allocation, the base contribution rate required to meet the replacement ratio goal (70% RR with 90% confidence) drops to 9.5% (from 10.6% previously). The average expected cost for the additional contribution feature increases to 1.2% of pay (from 0.9% previously). Table 5 shows the full range of cost using the higher risk asset allocation.

Table 5. Full Range of Cost Using Higher Risk Asset Allocation

%tile	*Base*	*Addl.*	*Total*
99%	9.5%	3.4%	12.9%
90%	9.5%	2.6%	12.1%
75%	9.5%	1.9%	11.4%
50%	9.5%	1.1%	10.6%
25%	9.5%	0.4%	9.9%
10%	9.5%	0.1%	9.6%
1%	9.5%	0.0%	9.5%

Table 6 shows the full range of replacement ratios using the higher risk asset allocation and compares them with the range of results using standard DC plans with contributions set at three different levels: (1) the 9.5% base contribution rate for the Tracker Plan, (2) the 10.7% expected average contribution rate for the Tracker Plan, and (3) a 11.4% contribution rate equal to the 75th percentile cost for the Tracker Plan. Downside, or shortfall, risk probabilities are also shown for replacement rate targets of 70% and 65%. All results are for the 37-year career sample employee used in the QE framework, as before.

Table 6. Full Range of Replacement Ratios Compared With DC Plans

		--- DC Plan w/ contr. rate: ---		
%tile	*Tracker*	*9.5%*	*10.7%*	*11.4%*
99%	129.8%	127.3%	138.5%	145.0%
90%	103.4%	100.2%	108.0%	112.5%
75%	90.8%	87.0%	93.1%	96.6%
50%	81.6%	77.0%	81.8%	84.6%
25%	74.8%	69.8%	73.7%	75.9%
10%	69.4%	64.4%	67.6%	69.4%
1%	62.2%	57.3%	59.6%	61.0%
Mean	84.3%	79.9%	85.1%	88.1%
Shortfall risk:				
70% RR	10.5%	26.3%	14.5%	11.1%
65% RR	3.2%	11.2%	7.0%	5.0%

The overall summary of the changes relative to the Tracker Plan using a lower risk asset allocation is as follows:

- The base rate drops by 1 percentage point, and the total expected cost drops by 0.8 percentage points, using the percent of pay cost measure.
- The mean replacement ratio increases by 7.7 percentage points, or about 10%.
- Shortfall risk is essentially the same at the 70% RR target, but the probability of falling below a 65% RR increases from 0.2% to 3.2%.

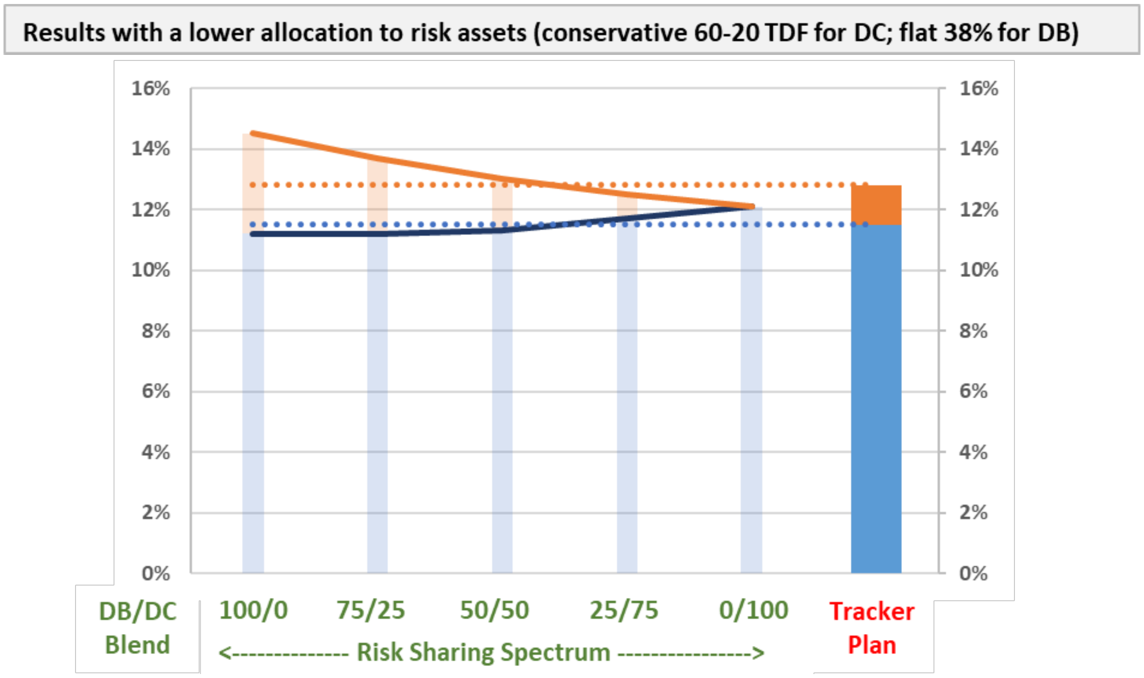
7.4 Comparative

The QE framework is designed to facilitate comparison of different plan designs, especially risk-sharing designs. The framework will be used to compare the Tracker Plan with other risk-sharing designs. I previously noted that for any design, an equivalent combo DB + DC plan can be determined, which will produce the same key benefit metrics. The equivalent combo plans can then serve as benchmarks to evaluate different designs.

Many of the comparative metrics in the QE framework are based on results with a standardized mean *cost* of 10% of pay for each program, which then allows for direct comparison of the benefit results and cost risk results. However, another approach in the QE framework is to standardize the plan results based on a specific *benefit* objective. One of those standardized benefit objectives is to derive the cost for each plan that would be required to produce a 70% replacement ratio (including Social Security) with a 90% confidence level (i.e., no more than about 10% of the simulation results for a full career sample employee would fall below the 70% replacement ratio target). This is the same benefit objective that I used to derive the specific Tracker Plan design discussed in this paper.

With this fixed benefit objective, a simple graph can be used to illustrate for any proposed plan design the degree of risk sharing, and the cost efficiency relative to the combo DB-DC plans. Figure 4 shows the graph for the Tracker Plan, using the more conservative asset allocation policy (a target date fund with a 60–20 glide path). The blue portion of each bar indicates the expected average cost level, and the orange portion indicates the “cost risk metric” level (i.e., the average cost in the worst quintile of outcomes). The large bar on the right is for the plan under evaluation (here the Tracker Plan), while the smaller set of bars to the left show results for various combination blends of a standard DB plan and a standard DC plan—which spans the full range of risk-sharing possibilities.

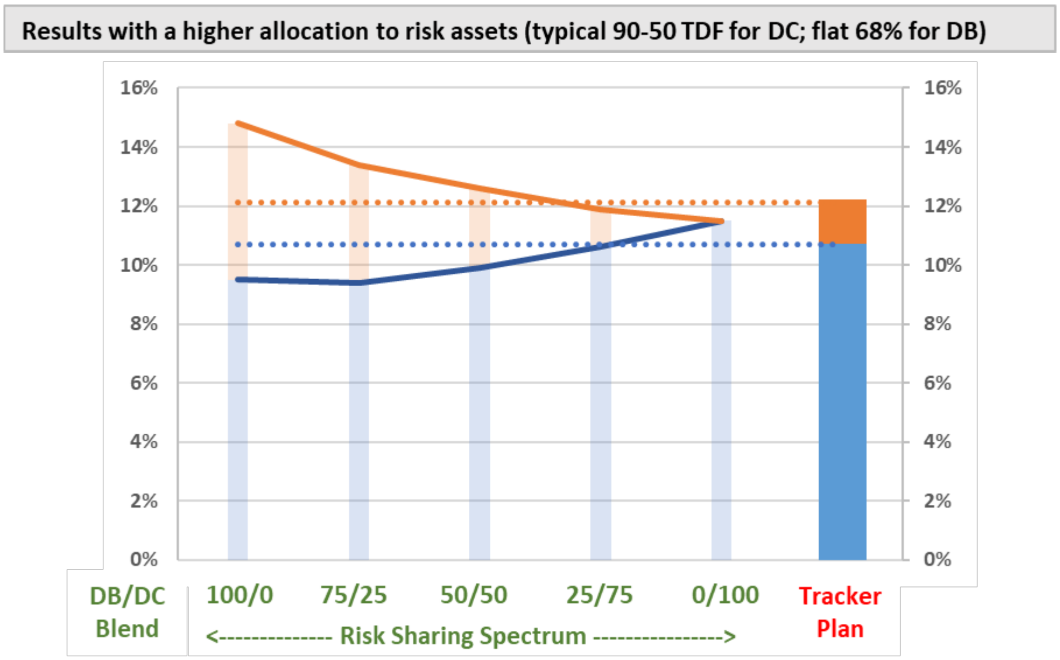
Figure 4.



The dotted blue line maps into the risk-sharing space at a point between the 50% DB–50% DC combo plan and the 25% DB–75% DC combo plan, indicating a significant degree of risk sharing for this Tracker Plan design. We then can see that the dotted orange line maps to a very similar location, indicating that the Tracker Plan has a cost risk almost the same as the equivalent combo plan. The specific results here are that the equivalent combo DB-DC plan is a 35% DB–65% DC blend, and the cost risk metric for this equivalent combo plan is 12.7% of pay. This compares to a cost risk metric for the Tracker Plan of 12.8% of pay. They both have a mean cost equal to 11.5% of pay. Since the Tracker Plan is a hybrid DC plan with an absolute cap on sponsor cost and no sponsor benefit liabilities, this seems to be a very favorable outcome—essentially the same cost risk as a combo program that includes a significant DB plan component.

Typically, within the QE framework these types of comparisons would be done at various levels of investment risk. Figure 5 is the same graph, but now using a higher allocation to risk assets (a target date fund with a typical 90–50 glide path).

Figure 5.



Now the Tracker Plan maps into the risk-sharing space more toward the DC end of the spectrum. The equivalent combo DB-DC plan here is a 20% DB–80% DC blend, and the cost risk metric for this equivalent combo plan is 11.8% of pay. This compares to a cost risk metric for the Tracker Plan of 12.1% of pay. They both have a mean cost equal to 10.7% of pay. While the relative efficiency for the Tracker Plan is now somewhat less than with the lower risk investment policy, the result still looks good for a plan with no sponsor benefit obligations.

To further illustrate how the QE framework can be useful in comparing various plan designs, I have included in Appendix B the same type of analysis for one other risk-sharing design currently used by at least one large public plan—namely the State of Kentucky, which recently adopted a form of cash balance DB plan design.

8. Payout Phase: A Collective Risk-Sharing Program

The previous sections described the Tracker Plan—a program designed to accumulate funds prior to retirement, with a special focus on the control of downside risk for participants. Now I want to focus on the payout phase,[2] where the accumulated funds at retirement need to be converted to a sustainable income stream for the lifetime of the retired participant. The concept I propose is a collective fund, with both investment and mortality risk shared among the group of retired participants, and also across different cohorts of retired participants (i.e., intergenerational risk sharing). Such a fund needs to be sustainable for all future years, and there should be no recourse to sponsor financial support (i.e., no liabilities or cost for the sponsor). To accomplish this, the fund must rely on some set of rules for dynamically adjusting the level of benefits depending on the level of funding relative to some measure

---

[2] My prior paper on the Tracker Plan concept dealt only with the accumulation phase.

of future benefit payments (i.e., funded status). I believe this structure provides a way to offer important features to retiring participants:

- Pooling of mortality risk on a very cost-efficient basis
- Guaranteed lifetime benefits
- Modest equity exposure to enhance long-term returns, with intergenerational risk sharing to mitigate short-term volatility
- Much more stable pricing for conversion of the lump sum accumulation into an annuity stream
- A natural way to manage aggregate, long-term longevity risk through the variable benefit features of the collective fund

This type of collective payout arrangement is not constrained to the Tracker Plan design. It could be used for any traditional DC plan or for a cash balance DB plan.

9. Collective Payout Program: Summary of Plan Provisions

Here I describe the key features of a prototype design that I have tested using the QE framework for payout programs. Obviously, this is just one illustration of a fund that fits into the collective risk-sharing family. Alternative choices for the plan design parameters are plausible (subject to testing for sustainability).

- There is a single investment pool, invested 35% in risk assets (U.S. and non-U.S. equities) and 65% in a core-type fixed income fund. For the most part, I would assume index funds are used to minimize expense charges. (I assume an expense charge of 0.25% each year, including administrative and investment expenses.)
- At retirement the lump sum is transferred into the collective payout fund, where it is used to purchase a base annuity income. Pricing would be based on a conservative estimate of the long-term expected return on the portfolio. The pricing structure would remain fixed from year-to-year, but the board would have authority to change it if there are significant changes in future return expectations or for periodic updates of mortality assumption. Any change could be phased in over a period of years. For the plan discussed in this paper, I use a 5.5% interest rate for the annuity prices, which is approximately the 25th percentile net return expected over a 30-year period in the QE framework simulation model. If the 10-year Treasury yield is ever higher than 6%, then the annuities would be priced using the 10-year Treasury yield plus 50 basis points to approximate market pricing in the high-yield environment.
- The annuity will include a 15-year certain period, which provides something very close to a "return of principal" guarantee for retirees.
- The base annuity benefit would be increased by a fixed 2.5% COLA factor each year after retirement, subject to the adjustment features described later.
- The fund would most often be in a surplus position relative to the liability for the base benefit, using a 5.5% discount rate. (More details on funded ratios will be shown in a later section.) If the funded ratio exceeds 110%, then onetime bonus payments would become payable for the following year, based on a published schedule. The schedule I used is based on this formula, where FR is the funded ratio at the beginning of the year: Bonus % = FR – 100% + max(0, FR – 130%). As an example, if the funded ratio was 141%, then the bonus percent would be 52% (applied to the base benefit amount). The bonus percent is capped at 100%.

- These bonus payments are for a single year only—they do not become part of the future base benefit income. However, the board always has full discretion to make special ad hoc decisions. If the funded position of the plan is very strong, the board could decide to issue some of the bonus in the form of an increase in the base benefit, increasing the liabilities of the plan.
- If the funded ratio falls below a specified trigger level for two out of the preceding three years, then the 2.5% COLA is suspended for the following year. Once the funded ratio has exceeded a specified second trigger for two out of the preceding three years, the COLA is reinstated. For this paper, I set the first trigger at 85% initially, increasing to 90% after 20 years of fund operation, and then to 95% at year 30. The second trigger, to restore the COLA, is set at the first trigger plus 10 percentage points.
- The board would always reserve the right to reduce annuity benefits in emergency situations, to maintain sustainability. (In my testing, this type of adjustment was never required.)
- Finally, collective risk-sharing programs generally have something like an evolutionary process, as the fund rules operate over time to build toward the target level, or range, of surplus assets. To support this process, I have assumed that for the first 5 years of operation the fund would pay no bonuses, nor would there be any COLA suspensions.

10. Payout Phase Analytics

10.1 Funded Status

With the previously described plan and using the QE framework simulation model, I ran a multicohort forecast of results to ensure the sustainability of the fund. Table 7 shows the range of funded ratios at various times after the fund starts operation. The "liability" is measured as the present value of ***e***xpected base benefits, including future COLAs (even though these are not guaranteed), using a 5.5% discount rate (the same rate used to price the annuity conversion factors).

Table 7. Range of Funded Ratios After Fund Start-up

	------------- Yrs. After Fund Start-up -------------					
	5	*10*	*20*	*30*	*40*	*50*
Percentile range:						
95%	111%	115%	127%	136%	142%	150%
75%	103%	105%	114%	124%	129%	133%
50%	98%	99%	106%	114%	120%	125%
25%	94%	93%	98%	106%	111%	116%
5%	87%	83%	87%	91%	99%	103%
Shortfall probabilities:						
<100%	61.8%	53.8%	31.1%	14.5%	6.4%	2.4%
<90%	10.8%	16.9%	8.4%	4.2%	0.8%	0.4%

While the evolution is slow, the fund ultimately stabilizes with a median funded ratio of about 125%, and only rare occurrence of less than a 100% funded ratio (and almost never below 90% funded.) Even during the process of building surplus levels, however, the fund operates very effectively, as I will show in the next section.

10.2 Benefits

Table 8 shows some statistics for the bonus payments that would be payable across the lifetime for various retiree cohorts. These results also illustrate the evolutionary process of a collective fund like this.

Table 8. Bonus Payments Payable Across the Lifetime

	------------ Cohort # -------------			
	1	*11*	*21*	*31*
Average % of years with bonus	20%	44%	68%	80%
Average bonus when paid	16%	23%	27%	31%

Even the very first retiree cohort under the program could expect to receive significant bonus payments, on average, over their lifetime. The bonus impact grows substantially as the plan reaches a more mature level of expected funded status. Although the initial cohort cannot expect to benefit as much from bonus payments, the value they receive from the favorable annuity pricing is still very substantial (as will be illustrated later).

Table 9 summarizes some results for COLA suspensions over the first 25 years of retirement.

Table 9. COLA Suspensions Over First 25 Years of Retirement

	------------ Cohort # -------------			
	1	*11*	*21*	*31*
No COLA suspensions	72%	70%	78%	88%
Between 1 and 5 suspensions	25%	26%	19%	10%
More than 5 suspensions	3%	4%	3%	1%

For all cohorts, there is at least a 70% expectation that they will not experience any COLA suspension during the first 25 years of retirement. Ultimately, that expectation is close to 90% for later cohorts. Experiencing more than five COLA suspensions is rare.

10.3 Comparative

The QE framework has a relatively straight-forward approach for comparative analysis of different payout program designs. First, a set of *baseline* benefits are determined from the simulation model:

- Initial accumulated fund balances are set at age 67 for each scenario equal to the standard DC plan accumulation using a typical target date fund and a 37-year accumulation period.
- These lump sums are then converted to lifetime benefits using fixed-price annuities that include a CPI COLA, and which are priced with a 5.85% discount rate (the expected return from a fund with a 30% risky asset allocation, as in a conservative DB plan). There is no load in the pricing, and no death benefits are assumed.

The intention is that these *baseline* benefits represent something like the best-case scenario from the participant's point of view.

Next the benefit stream generated from any payout program design can be calculated for each scenario using the simulation model. The ratio of these benefits to the baseline benefit for each year of each scenario are calculated, and these ratios are used as the basis for analysis. Various metrics are available, but the primary ones for comparative analysis are developed as follows:

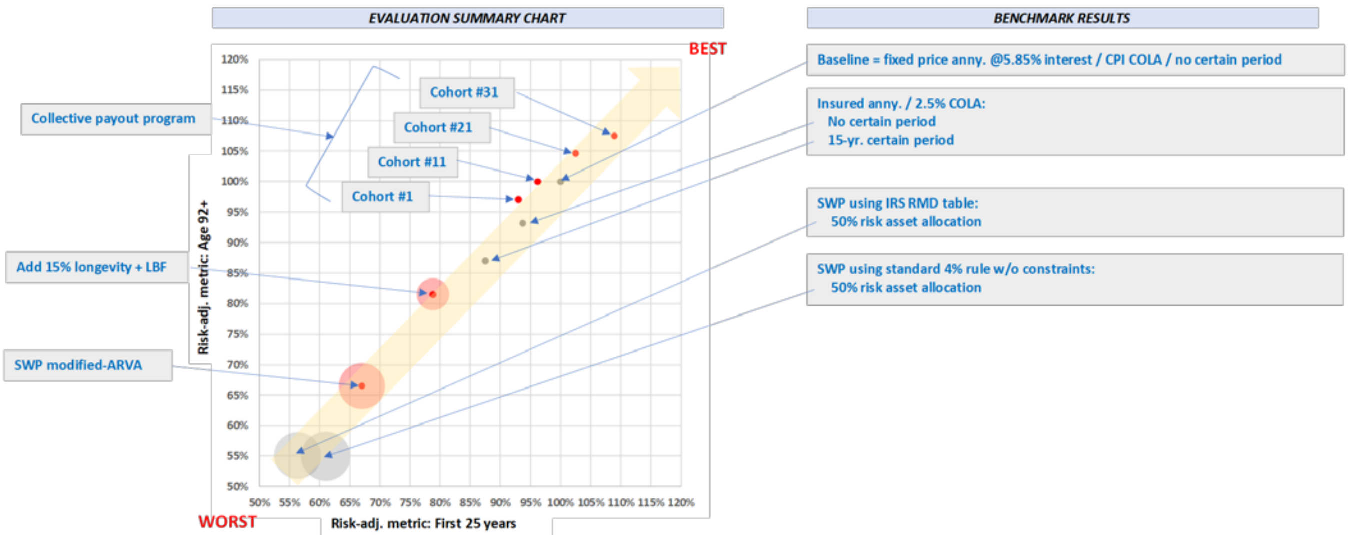
- Develop a risk-adjusted version for each ratio, reflecting the basic goal of maintaining a mostly level income stream, adjusted for inflation. Since the baseline benefits are fully CPI-indexed, the goal is for the benefit ratio measure to remain relatively uniform throughout the retiree's lifetime. If later years show significantly inflated values relative to the inflation-adjusted starting value, these results will receive a reduced weight. Conversely, if values in later years show significantly decreased values relative to the inflation-adjusted starting value, these receive a higher weight.
- Split the retiree lifetime into the first 25 years after retirement (approximately the life expectancy), and all years after that.
- For each of these periods, determine the average value of the risk-adjusted metric described previously.

The results can then be plotted in a graph, as shown in Figure 6. This graph also contains the results for some important benchmark designs, including insured annuities as well as several structured withdrawal plans (SWP)

- The standard "4% Rule" (initial payment is 4% of lump sum accumulation, with later payments increased by CPI).
- Withdrawals based on the IRS required minimum distribution life expectancy factors.
- Withdrawals based on an enhanced SWP using the annually recalculated virtual annuity (ARVA) methodology (with specific modifications). This is not a standard benchmark, but it is shown here to help illustrate the value of the collective program relative to a good SWP that includes no annuitization, or mortality risk pooling.
- The enhanced ARVA / SWP above, but with 15% of the accumulated balance used to purchase a longevity annuity starting at age 85, with a longevity bridge fund (LBF) to fund benefits up to age 85. This is also not a standard benchmark, but helps show the value of mortality risk pooling.

Where a shaded circle appears around a plotted result, that indicates liquidity at age 85 (accessible wealth as a multiple of pay). Results for the collective payout program are plotted for various retiree cohorts.

Figure 6. Evaluation Summary Chart



The graph shows how well the collective program performs, relative to all of the other plotted results. This is true even for the first retiree cohort. The values improve noticeably as the program reaches a more mature state.

11. Transition Issues

Navigating the transition process is a major issue in moving to a new plan design. Every situation will be unique, and legal issues vary from state to state. However, new designs should offer at least the financial flexibility to accommodate various types of transition arrangements. With the Tracker Plan, cohorts are generally established based on years to retirement, with a funding target set for the cohorts. In a transition arrangement, there could be special cohorts established for midcareer employees and the funding target for those cohorts could be determined with recognition of the value of currently expected levels of benefits under the existing defined benefit formula. This would lead to contribution levels that, with the support of the tracking process, will keep employees on a path to maintain their current expectations at retirement—albeit within a defined contribution plan, instead of a defined benefit plan. We have seen how the Tracker Plan can be designed around specific targets, along with the desired confidence level for achieving the target. Obviously, early retirement benefits and other ancillary benefits will also need recognition in any negotiation of the value trade-off decisions.

12. Conclusion

I believe that public programs will eventually move away from being mostly defined benefit plans, but I hope that they do not follow the private sector and move to standard defined contribution designs. There are many good designs that will work better for participants—some already proposed, others yet to be discovered. These are middle ground designs where risk is shared between the sponsor and the participants, and also collectively among the participants. The quantitative evaluation framework provides a set of tools and a methodology that should facilitate the acceptance of these new designs.

Appendix A: The Quantitative Evaluation (QE) Framework

The QE framework includes the following general components:

- A set of well-defined metrics that should be determined for each proposed retirement system design
- The specific methods that should be used for the calculations
- The specific assumptions that should be used for the calculations
- Separate (but connected) models for developing metrics for the accumulation phase and for the payout phase, predicated on the concept that these two parts of any retirement program are fundamentally separable (i.e., any accumulation structure may be combined with any payout structure to form a complete retirement system)

The standardized set of metrics, methods and assumptions allow direct comparison of quantitative results when analyzing different design ideas. The calculations and metrics use a stochastic economic simulation model to capture the full range of risk and reward for both benefit levels and for cost levels.

Table A1 shows some of the key economic assumptions (i.e., either the mean or median value from the full set of stochastic scenarios). The simulation model is specifically designed to model an economy "in equilibrium," to avoid the need to constantly adjust relative to current market conditions. The stability of the assumption set is critical to the process of comparing results that might be derived from design studies done at different times. While this type of simulation model is not appropriate for some purposes (i.e., actual pricing or liability valuations of any specific product or plan), it is completely appropriate when confined to comparative analysis purposes.

Table A1. Key Economic Assumputions

	Baseline assumption
Price inflation	2.50%
Real wage growth	0.50%
Nominal wage growth	3.00%
Real yield on 10-yr Treas	1.80%
Nominal yield on 10-yr Treas	4.30%
Credit/duration spread for for "Core Fixed Income"	0.20%
Nominal "Core Fixed Income" returns	4.50%
Risk premium for risky assets	3.50%
Nominal risky asset returns (gross, long-term compounded)	8.00%

Any retirement system includes both an accumulation phase (where benefits are accrued and assets are accumulated and invested to support the payouts that commence at retirement) and a payout phase (where accumulated assets are converted to some form of payout stream). The QE framework is predicated on the concept that accumulation structures and payout structures are fundamentally separable, such that any accumulation structure may be combined with any payout structure to create a complete retirement system. Because of this feature, the framework provides specific metrics (and supporting analysis) for each of these two separable parts of any system.

Since the primary goal of this framework is to facilitate comparisons on a consistent basis for various retirement system designs, the metrics are developed using a highly standardized model, with well-defined sample employees, calculation methods, and assumptions (both demographic and economic). The framework is based on a stochastic simulation process, as this is the only way to properly quantify the various risk and reward metrics that are needed for an adequate analysis.

For the accumulation phase, the metrics focus primarily on the range of benefits provided at the point of retirement and the range of possible cost for the sponsor. At a basic level, the accumulation phase is a completely scalable entity, in that the level of benefits for any structure (inclusive of a specified investment strategy) can be scaled up or down simply by adjusting the contribution inputs. Therefore, to facilitate comparisons, a standardized scale is needed, and this is done by specifying that total contributions (employee, plus expected sponsor) are assumed to equal 10% of pay over the active career.[3] The analysis can then focus on the expected level of benefits produced from this contribution, and the range of uncertainty for both the benefits and the sponsor cost. This 10% of pay assumption should not be considered an integral part of any design proposal, but merely a way to standardize results for comparative purposes. For any specific implementation of a design, the final level of contributions/costs (and how they may be split between participants and sponsor) should be based on the evaluation of desired benefit levels relative to adequacy standards, and cost constraints that may exist.

The accumulation phase metrics are based on a benchmark employee (or age cohort, for collective structures) with a starting age of 30 and retirement at age 67 following a continuous 37-year period of participation. The pay level and the career pay progression are designed to be representative of U.S. median income levels. The key benefit metric is the replacement ratio at age 67 (lifetime benefits, including Social Security, divided by final pay at retirement). For funds accumulated in a defined contribution plan, the lifetime benefit is determined by applying market-priced annuity factors for each scenario (i.e., the annuity price is a function of the market interest yields at retirement under the stochastic simulation model).

For the payout phase, the metrics show how benefits are distributed across the period of retirement, with separate metrics for income payments and for death benefits, and with the range of uncertainty shown for all of these. To standardize the payout phase metrics for comparative purposes, the starting accumulation amount at retirement is set (for each scenario in the stochastic model) at the age 67 balance under a standard defined contribution plan using a typical target date fund for investments and

---

[3] In addition to being a round number, the 10% of pay contribution is deemed a reasonable choice for two reasons: (a) current levels of total contributions (employee + employer) into U.S. 401(k) plans for private employers are about 10% of pay (in Vanguard 2015 data for participating employees, the median rate is 8.8% and the average rate is 9.5%), and (b) for the benchmark employee used in the framework, a 10% of pay contribution into a standard defined contribution plan using a typical target date investment fund over the full career produces a total replacement rate (annuitized plan benefit, plus Social Security) that will generally (i.e., with a probability greater than 50%) exceed 70% of final pay, which can be viewed as something of a minimal target to maintain the preretirement standard of living for a median income worker.

an assumed 10% of pay contribution level. A *baseline* set of annual income payments for each scenario in the stochastic model is calculated by applying a fixed-price annuity factor to the age 67 balance (i.e., the annuity is priced with a fixed 5.85% interest rate, including a CPI-based COLA priced at an assumed 2.5% inflation rate). The benefits generated by the payout design under analysis are then divided by the baseline benefits to develop a ratio, and this ratio (% of baseline benefit) is the basis for the metrics used in the payout phase of the QE framework.

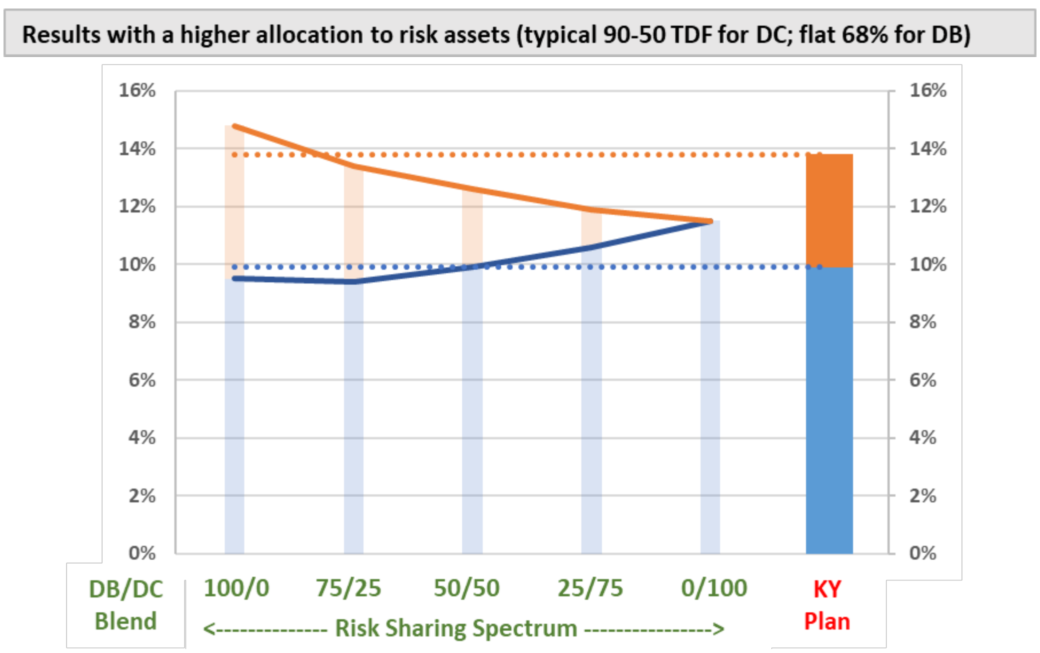
Appendix B: Analysis of Kentucky Plan

In this appendix I will also use the QE framework to analyze the risk-sharing design currently used by at least one large public plan—namely the State of Kentucky, which recently adopted a form of cash balance DB plan design. I describe this plan as a "participating cash balance" design, since the cash balance credits each year will depend on actual fund investment returns. This type of plan has been presented as an option to a number of other large public plans, and the design specifics can be modified to meet goals/constraints for benefits or cost. The specific formula used in the Kentucky plan to determine the annual investment credit is as follows:

- The credit rate is equal to 4%, plus 75% of the excess fund return over 4%—based on the 5-year moving average return
- Employees who terminate before retirement only receive the flat 4% credit rate

Figure B1 is the same graph for the State of Kentucky plan as previously shown in Figure 5 for the Tracker Plan—showing the cost for the plan to provide 90% probability of achieving a 70% total replacement ratio (inclusive of Social Security). For this illustration we use a 68% risk asset allocation (i.e., this would be comparable to the second Tracker Plan graph in the paper).

Figure B1.



The average cost is 9.9% of pay (compared to 10.7% for the Tracker Plan with this level of investment risk), which matches the average cost for a combo plan with a 50% DB and 50% DC split. But now the cost risk metric is much higher at 13.8% of pay (compared to 12.7% for the Tracker Plan, and 12.6% for the equivalent 50/50 combo plan). This participating cash balance plan design is less efficient at controlling cost risk, and since all cash balance plans are in the defined benefit category, there is also exposure to sponsor liabilities and uncapped cost levels. In this case, the advantages of the Kentucky plan (relative to a 50/50 combo plan) would have to be based on other features—perhaps packaging/communication (the plan in effect provides a 4% guaranteed return), but this graph helps illustrate the implicit cost of any guaranteed return feature in a plan design. A more complete analysis would require comparison of the full range of benefit outcomes, which is beyond the scope of this paper.

Rowland Davis, FSA, is the former president of RMD Consulting in Chicago. He can be reached at *rwlnddvs@gmail.com.*