NAIC economic scenario reform: A model for VM-21 impact analysis

Collin Davidson
Zohair Motiwalla
Parker Henley

Milliman

Executive Summary

The National Association of Insurance Commissioners (NAIC) continues to evaluate a new economic scenario generator (ESG) for U.S. principle-based reserving (PBR). To help companies navigate these changes, this paper presents an illustrative framework for quantifying the potential impact of the scenario reform for a prototypical in-force block of variable annuities (VAs). Utilizing a few key economic scenario sets presented in the NAIC's first field testing, the analysis looks at VM-21 statutory reserve and C-3 Phase II capital impacts across a range of block characteristics, along with capital market sensitivities and measures of scenario dispersion in the tail. This paper is intended to lay the groundwork for developing an understanding of the complex and potentially material impacts that eventual ESG reform is poised to introduce, so that companies can monitor and plan around the altered risk profile that emerges.

Introduction

The NAIC has been engaged for several years in reforming the real-world ESG used in PBR for U.S. statutory requirements. The scope of impact from this change is expected to be significant, leading to reserve and capital impacts under the VM-20,[1] VM-21,[2] and C-3 Phase II[3] frameworks, as well as setting a precursor for eventual VM-22[4] and C-3 Phase I[5] reform. In 2022, the NAIC rolled out an initial ESG field test that evaluated several test scenario sets produced via the Conning GEMS[6] model. In early 2023, the NAIC released an update summarizing the initial findings from the first round of industry field testing and laid out a timeline that included a subsequent industry field test.[7] The timing for the latter is yet to be finalized but is expected to be rolled out toward the end of 2023.

While acknowledging that the effects of ESG reform will eventually span across practically all of life and annuity valuation in the United States, the focus of this paper is on the potential impact to VM-21 reserves and risk-based capital (RBC) C-3 capital requirements for VAs. This is no easy task, as there is substantial uncertainty still surrounding what the final economic scenarios could look like, and variable annuities introduce more complexity than other products due to capital market exposure, hedge modeling, and dynamic policyholder behavior. Given the importance of the ESG reform efforts, this paper seeks to provide a foundational model to review the VM-21 and C-3 impacts of the scenarios used in the initial round of industry field testing on some prototypical mixes of variable annuity business. We anticipate this analysis will be useful to companies as they navigate the impact of the Conning GEMS scenarios in both rounds of industry field testing.

[1] PBR framework for life products.

[2] PBR framework for variable annuity and registered index-linked annuity products. Only variable annuities are in scope for this paper.

[3] Capital framework for variable annuity and registered index-linked annuity products. Only variable annuities are in scope for this paper.

[4] Proposed PBR framework for non-variable annuities.

[5] Capital framework for fixed deferred annuities and payout annuities, which will be reformed after VM-22 is adopted.

[6] Conning GEMS Economic Scenario Generator. Retrieved July 9, 2023, from https://www.conning.com/-/media/marketingsite/documents/product-sheets/software-and-services/risk-solutions/gems-esg.pdf.

[7] NAIC (March 20-21, 2023). LATF Spring National Meeting: Supplemental Packet, Agenda Item 7. Life Actuarial (A) Task Force. Retrieved July 9, 2023, from https://content.naic.org/sites/default/files/national_meeting/LATF_Spring2023NM_SupplementalPacket.pdf.

When the NAIC released the initial report of findings from the initial field test, the results were instructive in terms of broad observations, but because of the heterogeneity in business mix and strategy across the VA industry it arguably had limited utility when trying to understand the relative impact to a typical VA block. This paper intentionally seeks to meet this need by generalizing on a few business dimensions to present illustrative impacts across key scenarios from the field testing. The goal is not to replicate the entirety of the industry field test, but to present some key findings when comparing the results produced by the American Academy of Actuaries (AAA) Interest Rate Generator, which currently serves as the default ESG, versus the results produced by Conning's GEMS models.

The goals of this paper are to evaluate VA reserves and capital between the AIRG and alternative economic scenarios with respect to the following:

- **Conditional tail expectation (CTE) dispersion**: Evaluating the distribution of scenario-level results, particularly focusing on the CTE70 and CTE98 metrics across scenario sets.
- **Hedge credit**: Calculating the level of hedge cost or credit under different scenarios when modeling an implicit hedging strategy, which is in line with VM-21 Section 9.B.3 guidance for a future hedging strategy. In this context, hedging is intended to insulate against the market risk introduced by the VA guarantees.
- **Market sensitivities**: Utilizing instantaneous equity shocks of +/-10% and +/-25%, as well as the implied impact of shocking interest rates by comparing scenarios produced using two different starting yield curves provided in the field testing sets (the yield curve of December 31, 2021, and the yield curve of December 31, 2019, plus 200 basis points [bps]).
- **Mix of business**: Reviewing new versus seasoned business, rider mix—Guaranteed Minimum Death Benefit (GMDB) versus Guaranteed Lifetime Withdrawal Benefit (GLWB)—fund mapping, and equity exposure.

The next two sections of this paper provide some background on the economic scenarios and the model used for the prototypical VA block. Following that, we analyze the impact of each of the various topics described above and then share some concluding thoughts.

Scenarios

In this paper, we review scenarios from three different real-world ESGs:

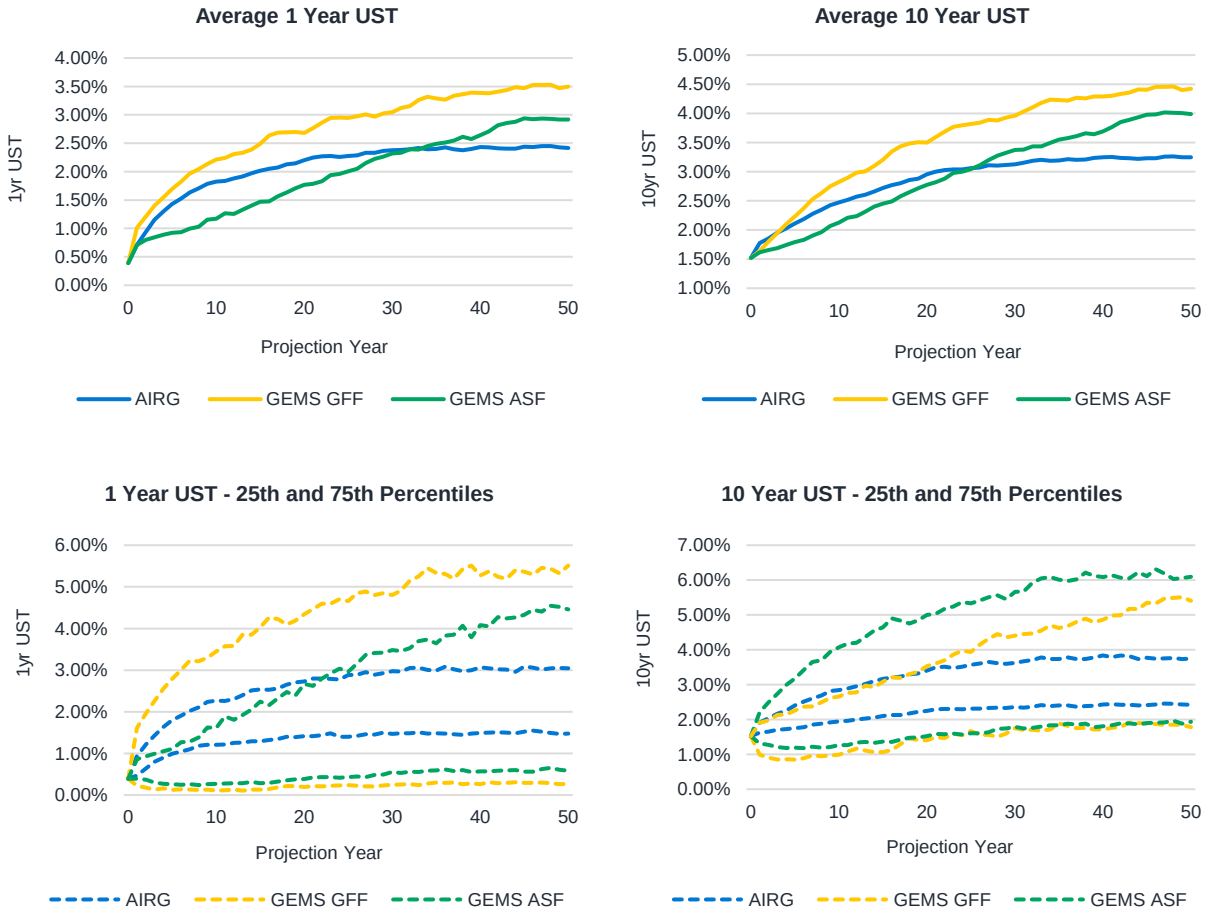
1. The AAA Interest Rate Generator (May 2021 release), referred to as "AIRG" in this paper.
2. Conning GEMS with Generalized Fractional Floor (corresponding to field test Sets 1a/2a), referred to as "GEMS GFF" in this paper.
3. Conning GEMS with Alternative Shadow Floor (corresponding to field test Sets 1b/2b), referred to as "GEMS ASF" in this paper.

Under current VM-21 guidance, the AIRG serves as the default ESG for PBR, making it an appropriate baseline for comparison of the new scenarios. (In particular, the AIRG is the only ESG listed above that is actively being used for financial reporting.) Each GEMS set was part of the NAIC's first round of industry field testing. These two sets present different approaches to interest rate modeling and reflect an equity risk premium methodology that sets equity returns as a function of the interest rate level. Ahead of the field testing, the NAIC released commentary explaining the differences between each GEMS model version and the motivations behind them.[8] Both GEMS model versions rely upon a risk premia model, meaning the simulated equity returns are the function of the interest rate level plus a risk premium. The GEMS GFF version utilizes Conning's calibration, which achieves NAIC acceptance criteria for a "low for long" scenario, along with a wide distribution of high interest rates and a flooring mechanism that limits how negative rates can get. The GEMS ASF version aims to address some of the extreme frequency and severity of rate levels from the GEMS GFF model without sacrificing the broad acceptance criteria and is both a tighter distribution of rates compared to GEMS GFF and closer to historical data.

[8] For more details, the NAIC deck can be found at https://azspcngcms.blob.core.windows.net/sitecoremedia/project/naic/files/latf-esg-exposure-3,-d-,17,-d-,2022/recommended-models-for-esg-field-testing_031722.pdf.

To help set the stage for the impact analysis in later sections, Figure 1 shows a few scenario statistics across the three sets used, as of December 31, 2021. Because VM-21 is focused on tail measures of risk, the averages are not intended to completely bridge the impacts but still serve as a useful reference point. We encourage the reader to review the materials made available at the joint NAIC-Conning website for a deeper dive into the scenario statistics and more technical details on the drivers of differences between scenario sets.[9]

FIGURE 1: INTEREST RATES



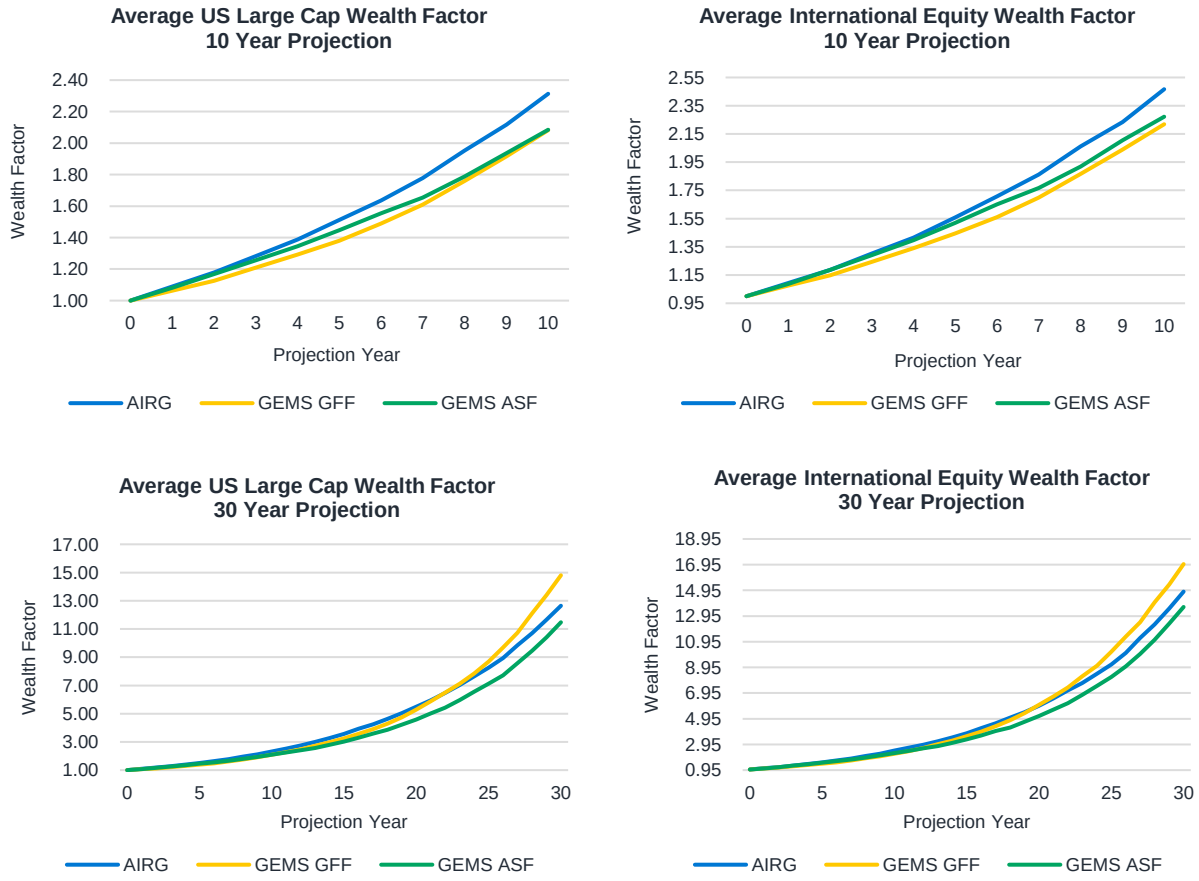
A few observations on these interest rate scenarios:

- **Ultimate level**: The average ultimate 10-year Treasury rate is approximately 3.25% for the AIRG (consistent with the mean reversion target) but is 75 to 125 basis points higher for the GEMS sets.
- **Mean reversion and curve shape**: The AIRG flattens out around long-term rate levels approximately 25 to 30 years into the projection, whereas all other test sets do not flatten out until around 45 years into the projection. The AIRG's faster mean reversion contributes to it exceeding the GEMS ASF over the first 25 projection years.
- **Short vs. long rate dispersion**: At the end of the projection period, the spread between the 10-year Treasury and 1-year Treasury grows to approximately 100 basis points for the AIRG. This pattern is closely matched by the two GEMS scenario sets.

[9] See https://naic.conning.com/scenariofiles.

- **Rate dispersion**: The dispersion of 1-year and 10-year Treasury rates is much narrower for the AIRG than the GEMS scenarios, and GEMS GFF exhibits the widest range. From the chart, the reader can observe that the difference for the ultimate 10-year Treasury between the 25th and 75th percentiles is less than 1.5% for the AIRG, but approximately 4% for the GEMS sets.
- **Minimum rates and curve inversion**: Not shown in the charts in Figure 1, the AIRG has an embedded floor of 1 basis point, which prevents interest rates from going negative. The other scenario sets introduce negative rates (to varying degrees) that widen the range of possibilities. Additionally, the other sets exhibit a greater likelihood of inverted yield curves than the AIRG model.

FIGURE 2: FUND WEALTH FACTORS AND RETURNS



In addition to the wealth factors, the table in Figure 3 provides supplemental statistics for several modeled funds.

FIGURE 3: SUPPLEMENTAL STATISTICS FOR MODELED FUNDS

METRIC	AIRG	GEMS GFF	GEMS ASF
U.S. LARGE CAP			
Average Return	8.77%	9.42%	8.85%
Volatility	16.20%	16.33%	15.97%
INTERNATIONAL EQUITY			
Average Return	9.29%	10.10%	9.60%
Volatility	18.59%	18.66%	18.50%
Correlation to U.S. Large Cap	0.5620	0.7552	0.7512
U.S. AGGRESSIVE			
Average Return	11.39%	13.79%	13.07%
Volatility	26.97%	24.76%	24.28%
Correlation to U.S. Large Cap	0.5823	0.8423	0.8366
U.S. SMALL CAP			
Average Return	10.13%	10.22%	9.71%
Volatility	22.09%	18.80%	18.45%
Correlation to U.S. Large Cap	0.7737	0.8505	0.8454
FIXED INCOME			
Average Return	2.82%	4.75%	4.06%
Volatility	4.16%	9.54%	7.03%
Correlation to U.S. Large Cap	0.2289	0.2360	0.3144

Some initial observations on the average wealth factors and key statistics across funds:

- **U.S. Large Cap wealth factors**: The AIRG outperforms the GEMS model over the first 25 years on average, but the GEMS GFF model eclipses the AIRG beyond that point, while GEMS ASF remains below the average AIRG path. The GEMS GFF return exceeding the GEMS ASF return is consistent with the pattern of higher rates and the risk premia model.
- **Fixed Income wealth factors**: The AIRG fund growth is significantly lower than the GEMS fund growth. This is consistent with the pattern of higher rates in the other test sets, along with higher average credit spreads. Despite having significantly higher interest rates compared to the GEMS ASF over the first 20 projection years, the AIRG wealth factor falls short of GEMS ASF's. This supports another feature present in the GEMS sets, which exhibit higher average credit spreads in comparison to the AIRG set.
- **Equity returns**: Average annual returns are generally comparable between AIRG and GEMS sets for the U.S. Large Cap and U.S. Small Cap funds, but the GEMS model exhibits higher returns for International and Aggressive equity funds.
- **Volatility**: Realized fund volatility is comparable for U.S. Large Cap and International Equity funds between the AIRG and GEMS models, whereas GEMS exhibits lower realized volatility for the Small Cap and Aggressive equity funds.
- **Correlation**: Correlations are notably higher between the U.S. Large Cap and other equity funds in the GEMS model, leading to increased portfolio volatility for a mix of funds and potentially reducing diversification benefit.

[10] The statistics in Figure 3 represent the average 50-year annualized return and realized volatility across all 1,000 scenarios.

Armed with this knowledge of the characteristics of the scenarios, it is worth considering how these points map onto changes in the total asset requirement (TAR) as we dive into the model and impact analysis between scenario sets. A summary of how key aspects of the scenario sets impact the TAR is outlined below:

- **Interest rates**: On average, interest rates are higher in the GEMS test sets, but the range is much wider and allows for more curve inversion and negative rates, contributing to more adverse discounting and new money rates. With the GEMS ASF model, the frequency and severity of negative rates is dampened, leading to the GEMS ASF set having an improved TAR relative to GEMS GFF.
- **Equity returns**: Despite variation on expected return levels across funds and time horizon, one consistent observation is the higher portfolio volatility among equity funds. This reduces the diversification benefit that can significantly reduce TAR at higher CTE levels.
- **Fixed Income returns**: The GEMS models embed higher credit spreads in the corporate bond modeling, which leads to the GEMS scenario sets having more favorable fixed income returns than AIRG.

Describing the model

The valuation model used in this analysis is kept consistent across runs used for each scenario set and is intended to reflect assumptions and methodology that are considered reasonably illustrative of existing VA business. Each set employs 1,000 scenarios, uses actuarial assumptions consistent with those prescribed in the VM-21 Standard Projection plus a 10% margin for adverse deviation for each assumption, and assumes all assets earned the net asset earned rate. The liability in-force used representative model cells; for GLWB riders, the modeled cells utilize the prescribed withdrawal delay cohort method outlined in the VM-21 Standard Projection for approximating the timing of withdrawal onset. Because the assumptions were defined to intentionally be more conservative than the Standard Projection, the VM-21 reserves calculation will assume that no Additional Standard Project Amount is required.

To support a comprehensive review of the impact of various scenario sets, a variety of dimensions were considered, as summarized in the table in Figure 4. Within these dimensions, we have denoted what is considered the “baseline” assumed for comparison basis.

FIGURE 4: SUMMARY OF DIMENSIONS CONSIDERED

DIMENSION	OPTIONS CONSIDERED	CHOSEN BASELINE CASE
Average Block Age	– Newly issued: policy duration = 0, at-the-money – Aged block: policy duration = 10, 25% in-the-money[11]	Aged block
Rider/Guarantee Mix	– 100% GMDB only – 67% GLWB (with GMDB), 33% GMDB – 100% GLWB (with GMDB)	67% GLWB, 33% GMDB
Fund Mix	– 60% Equity Mix, 40% Fixed Income – 70% Equity Mix, 30% Fixed Income	60% Equity Mix, 40% Fixed Income
Hedging Approach	– Unhedged – Hedged with 10% E-Factor	Hedged with 10% E-Factor

[11] Moneyness is defined on a nominal basis, which is either (Benefit Base / Account Value – 1) for GLWB or (Benefit Balance / Account Value – 1) for GMDB. “Benefit Base” refers to the accumulated guarantee value used to determine the GLWB withdrawal amount and is fixed upon first withdrawal, whereas “Benefit Balance” refers to the remaining guarantee amount netted against realized withdrawals.

Below are a few additional notes to consider. These points are intended to provide additional context behind the definition and details of the options described in Figure 4.

- **Average policy duration**: The newly issued block assumes a seven-year surrender charge period and no immediate withdrawal onset for the GLWB, whereas the aged block is out of the surrender charge period and assumes that approximately 80% of the GLWB business has already begun taking regular withdrawals.
- **Rider/guarantee mix**: The assumed GLWB offers a 6% simple annual rollup rate, along with a 5% guaranteed withdrawal rate, and assumes the corresponding GMDB rider is the maximum of the return-of-premium or an annual ratchet (referred to as a "Weak DB"). The GMDB-only business is split equally between Weak DB riders and a GMDB that assumes the maximum of the Weak DB features or a 6% annual rollup guarantee.
- **Fund mix**: All scenario sets make a preset fixed income fund available, which is defined as 65% government bond and 35% corporate bond. The "Equity Mix" that we chose for the baseline case allocates a plurality to U.S. Large Cap, and the remainder across the International, Small Cap, and Aggressive equity funds. The Fixed Income fund embeds rebalancing, but the model does not assume rebalancing to target allocations between equity and fixed income, or among equity funds.
- **Hedging approach**: The implicit hedging approach that we have used assumes a strategy covering total equity and interest rate risk (or simply "full delta-rho"), based on the rider cash flows as the hedge target. The hedge cost is a risk-neutral valuation based on Milliman Guarantee Index® (MGI)[12] volatility as of December 31, 2021, assuming the 1-year constant maturity Treasury (CMT) forward curve as the risk-free rate. This hedge cost is amortized according to the release of rider fees. Rider fees are assumed to be 100% hedged, and 90% hedge effectiveness is assumed on the GLWB and GMDB claims. The hedge cost amount and hedge effectiveness levels are assumed to be the same across scenario sets.

In addition to the liability modeling assumptions, we also discuss a few key market sensitivities, using December 31, 2021, market conditions as the baseline. Equity shocks (+10%, -10%, +25%, -25%) are reflected as immediate shocks to equity-mapped account value. Given limited availability of the Conning GEMS scenarios, a singular interest rate sensitivity is constructed by looking at the difference in the TAR between scenario sets generated assuming the December 31, 2021, and the December 31, 2019, plus 200 bps starting curves. We independently generated AIRG economic scenarios using the same specified curves and can measure the degree of rate sensitivity of each scenario set based on the change in TAR between rate environments.

Impact analysis

BASELINE SUMMARY

The model runs assumed a $20 billion account value in-force book for the baseline runs. This section details the impacts to VM-21 reserves in excess of cash surrender value (CSV), as well as 400% RBC C-3 TAR (also in excess of CSV), which for the purposes of this paper we define as VM-21 reserves + (Pretax CTE98 – Pretax CTE70) x (1 – 21%).[13] Consistent with the VM-21 instructions, the reported reserve and TAR results assuming a future hedging strategy and 10% E-Factor are given by the following formula:

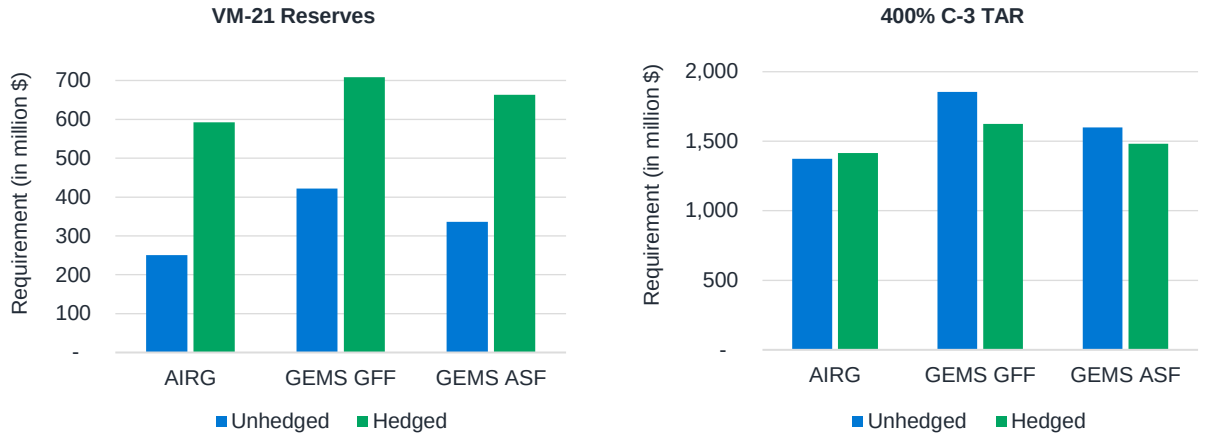
Reported CTE metric = Hedged CTE metric + 10% x max(0, Unhedged CTE metric – Hedged CTE metric)

The initial baseline case across each scenario set, as of December 31, 2021, is shown in Figure 5.

[12] See https://us.milliman.com/en/products/guarantee-index.

[13] For the capital component, it is assumed that the amount of nonadmitted deferred tax asset (DTA) allocated to the VA line is zero so that the after-tax statutory/tax reserve difference is capped at zero.

FIGURE 5: INITIAL BASELINE CASE



As an alternative view, Figure 6 shows a summary of the 400% C-3 TAR, presented as a percentage of initial account value.

FIGURE 6: 400% C-3 TAR AS A PERCENTAGE OF INITIAL ACCOUNT VALUE

400% C-3 TAR	AIRG	GEMS GFF	GEMS ASF
Unhedged	6.9%	9.3%	8.0%
Hedged	7.1%	8.1%	7.4%
Hedge Impact	0.2%	-1.2%	-0.6%

These views illustrate that, on an unhedged basis, the GEMS scenario sets produce materially higher TAR than the AIRG. However, once hedging is reflected, the pattern generally still holds, but the increase in TAR is dramatically muted. Consequently, the impact of hedging is more significant in the GEMS tail scenarios, as we can see by comparing the hedged versus unhedged TAR for each scenario set.

It is worth pausing to emphasize the baseline scenario impacts for our prototypical case, based on a $20 billion account value (AV) block of business. Further, we can observe the 400% C-3 capital requirement change, defined as TAR in excess of VM-21 reserves.

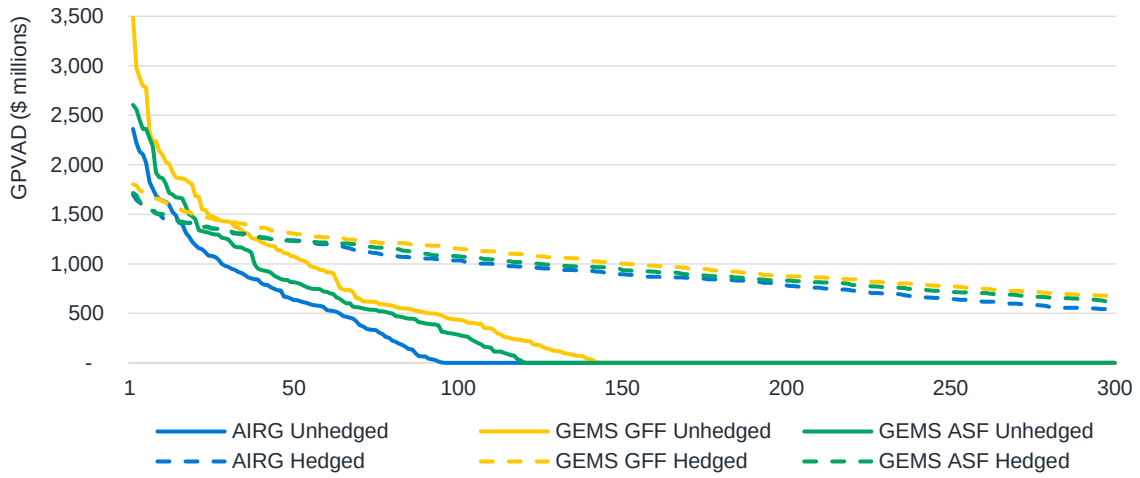
- If unhedged, TAR increases $479 million, or +2.4%, for GEMS GFF. This includes a 27% increase in capital.
- If hedged, TAR increases $210 million, or +1.0%, for GEMS GFF. This includes an 11% increase in capital.
- If unhedged, TAR increases $225 million, or +1.1%, for GEMS ASF. This includes a 12% increase in capital.
- If hedged, TAR increases $68 million, or +0.3%, for GEMS ASF. This includes a -0.4% change in capital.

As we get further outside the tail, there are some scenarios within the CTE70 set for which the CSV floor is binding on an unhedged basis. The GEMS GFF model is the only scenario set that produces a requirement on an unfloored reserves basis, and both GEMS scenario sets exhibit a smaller effect of flooring than the AIRG, consistent with earlier observations of higher TAR requirements for GEMS-based scenario sets.

SCENARIO DISTRIBUTION

Beyond the initial impacts suggested, the distribution of the scenario-level greatest present value of accumulated deficiencies (GPVAD) also varies. The chart in Figure 7 plots the ranked GPVAD for the worst 300 (i.e., CTE70) scenarios of each scenario set. Note that scenarios with a negative GPVAD are floored at zero in Figure 7, whereas in the actual calculation the scenario reserve would be floored at the binding CSV floor.

FIGURE 7: RANKED WORST 300 SCENARIOS (WITH GPVAD FLOORING)



Based on Figure 7, we note the following observations:

- **Hedging impact**: The dashed lines indicate how effectively hedging tightens the distribution as each case embeds the convergence of rider cash flows toward the risk-neutral hedge cost, thus removing substantial variability in dispersion across scenario sets.
- **Scenario flooring**: On an unhedged basis, the AIRG begins experiencing scenarios that would be floored around rank 100, or approximately CTE90. In contrast, GEMS ASF hits that point around CTE88 and GEMS GFF hits it around CTE85.
- **Scenario distribution**: Driven by a few particularly adverse scenarios, GEMS GFF unhedged results exhibit a much wider range in GPVADs than the AIRG or GEMS ASF sets. GEMS ASF is fairly consistent with the degree of AIRG dispersion but is shifted toward slightly worse GPVADs across the distribution.

An alternative way to view the distribution of scenarios might be to consider what the equivalent CTE level is under each set compared to the AIRG. For instance, if capital requirements are generally considered about CTE98-level risk as a measure of solvency, then the CTE98 might reflect a very different degree of conservatism across scenario sets. The table in Figure 8 shows the approximate equivalent CTE level across each scenario set, which applies to both the hedged and unhedged distribution of results.

FIGURE 8: CTE LEVELS

	EQUIVALENT CTE LEVEL	
SCENARIO SET	**RESERVES**	**CAPITAL**
AIRG	70	98
GEMS GFF	61	95
GEMS ASF	65	97

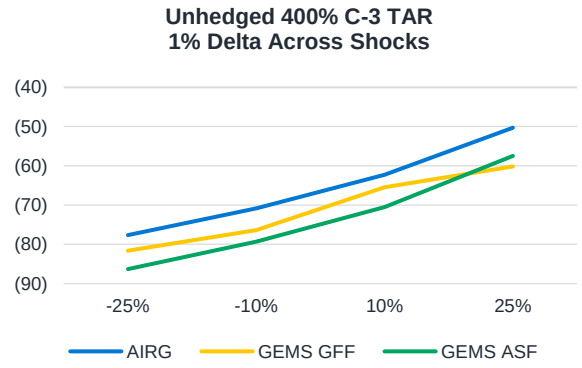
MARKET SENSITIVITIES

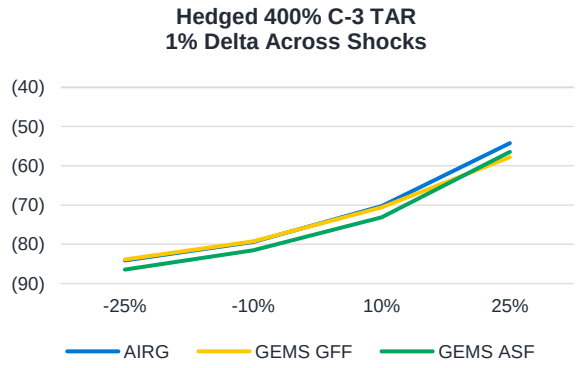
We also explore the potential impact to market sensitivity profiles between scenario sets. For companies that are interested in hedging their statutory balance sheet to maintain stable RBC multiple levels, this can be a difficult task to manage because the PBR mechanics in statutory reporting still exhibit some dynamics that may behave differently than their respective economic hedge target metrics.

To evaluate the market sensitivity, we first looked at hedged and unhedged TAR sensitivity across a few key equity shocks. The table and charts in Figure 9 present these impacts as the sensitivity per 1% equity change, effectively implying the TAR delta around both the baseline TAR level and across shocks.

FIGURE 9: HEDGED AND UNHEDGED TAR SENSITIVITY (IN MILLIONS)

	BASE 1% TAR DELTA ($M)	
SCENARIO SET	**UNHEDGED**	**HEDGED**
AIRG	(67)	(75)
GEMS GFF	(71)	(75)
GEMS ASF	(75)	(77)



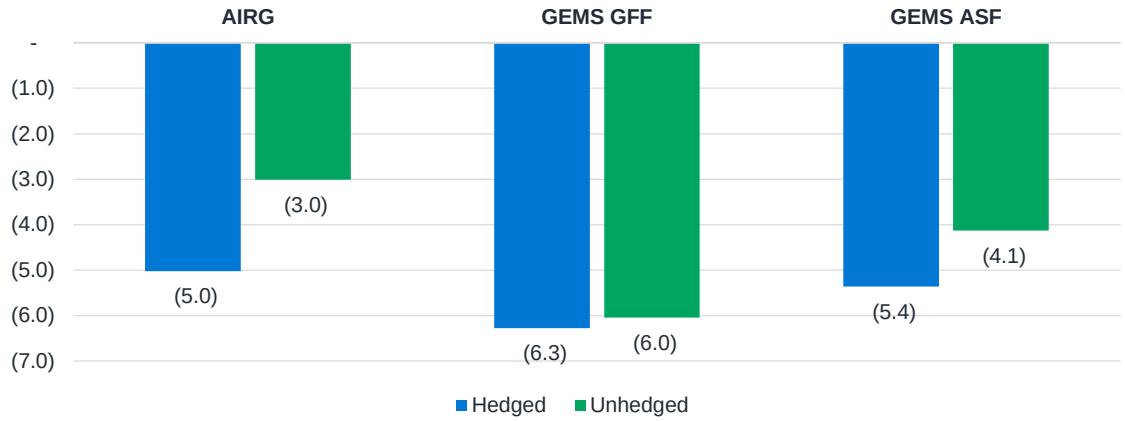


Based on these charts, we note some observations below. Complete explanations for this behavior can be challenging because of the complex interactions of equity and rate behavior at various shock levels on CTEs. Companies should contemplate a robust grid of equity, rate, and cross-shocks to bolster understanding of these dynamics and how a new scenario set could influence block sensitivities.

- **Hedging effect**: Even though the modeled hedge strategy only covers the rider, the presence of hedging significantly converges the level of delta across scenario sets.
- **Unhedged sensitivity comparison**: On an unhedged basis, the GEMS sets exhibit more equity sensitivity than the AIRG across shocks. In particular, the GEMS ASF sensitivity exhibits the greatest delta until the +25% shock, where GEMS GFF begins to level off.
- **GEMS +25% shock**: Moving from the +10% to +25% shock, the AIRG and GEMS ASF deltas reduce at a faster rate than GEMS GFF. This is attributable to the emergence of scenario flooring, which mutes the sensitivity of those sets, whereas GEMS GFF is not impacted by scenario flooring, thus realizing more TAR sensitivity.
- **GEMS GFF vs. GEMS ASF**: Ignoring the +25% shock due to flooring, if we compare the lower sensitivity of GEMS GFF versus GEMS ASF, the observation is consistent with principles of option theory, where higher volatility tracks with less delta.
- **GEMS ASF vs. AIRG**: GEMS ASF is approximately a parallel shift toward more equity sensitivity when compared directly to the AIRG delta across shocks, and the two sets exhibit a similar rate of change in delta between shocks. One possible explanation for the increased sensitivity with GEMS ASF (and GEMS GFF to a lesser extent) is the level of rate variation and increased correlation between equity and fixed income returns.

The ability to review rate sensitivity is limited due to the reliance on new scenarios being generated for any initial rate curve shock. With that in mind, we used the additional valuation curve provided in the field testing sets that represented the Treasury curve of December 31, 2019 plus 200 basis points. Compared to December 31, 2021, this meant a +240 basis point change in the 10-year Treasury rate, and a +320 basis point change on the 1-year Treasury rate. Based on conventional key rate exposure for GLWBs, we assumed that the 10-year Treasury rate movement was a reasonable proxy to normalize the TAR rate sensitivity against. Consequently, we divided the TAR sensitivity from the rate curve by 240 to present the TAR change per 1 basis point in the chart in Figure 10.

FIGURE 10: 400% C-3 TAR, 1 BP RHO (IN MILLIONS)



Whereas the story on equity sensitivity may be more complex, the characteristics of the interest rate models across scenario sets support a cleaner explanation of the patterns we observe above. The reader should also keep in mind that it is common to see the sensitivity of a risk-neutral valuation exceed the rate sensitivity for TAR valuations based on real-world scenarios, because the mean-reverting anchoring nature of real-world scenario generators can dampen the effect of any rate shock, while risk-neutral scenarios would experience the effects of curve shocks over the entirety of a projection.

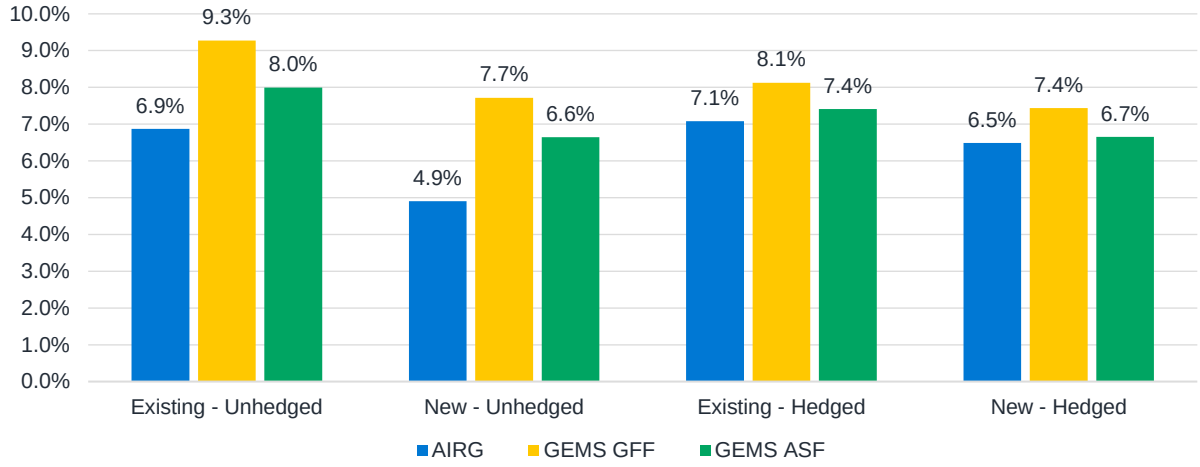
Some observations:

- **Overall trend**: GEMS scenario sets exhibit greater rate sensitivity than the AIRG, at least partially driven by their risk premia model for equity returns.
- **Hedging effect**: All scenario sets converge toward greater rho when hedging is reflected, in line with the interest rate sensitivity for the risk-neutral scenarios underlying the hedge cost. However, the degree of difference between the hedged and unhedged TAR sensitivity varies significantly across scenario sets.
- **GEMS GFF hedging impact**: The relationship of hedged versus unhedged rho is quite comparable for the GEMS GFF scenario set, which signals that the progression of interest rates in this scenario set is comparable to how the curve may evolve along the forward path.

NEW VERSUS EXISTING BUSINESS

The age of the block could contribute to how sensitive it is to the underlying scenario sets. To evaluate this, the chart in Figure 11 reviews the TAR across existing business versus newly issued business. For the existing business, the majority of the GLWB policies are assumed to be withdrawing as of the valuation date, and the block is assumed to be 25% in-the-money (ITM), whereas the new business reflects no assumed withdrawers at issue and initially is at-the-money.

FIGURE 11: 400% C-3 TAR (AS % OF AV) – EXISTING VS. NEW BUSINESS



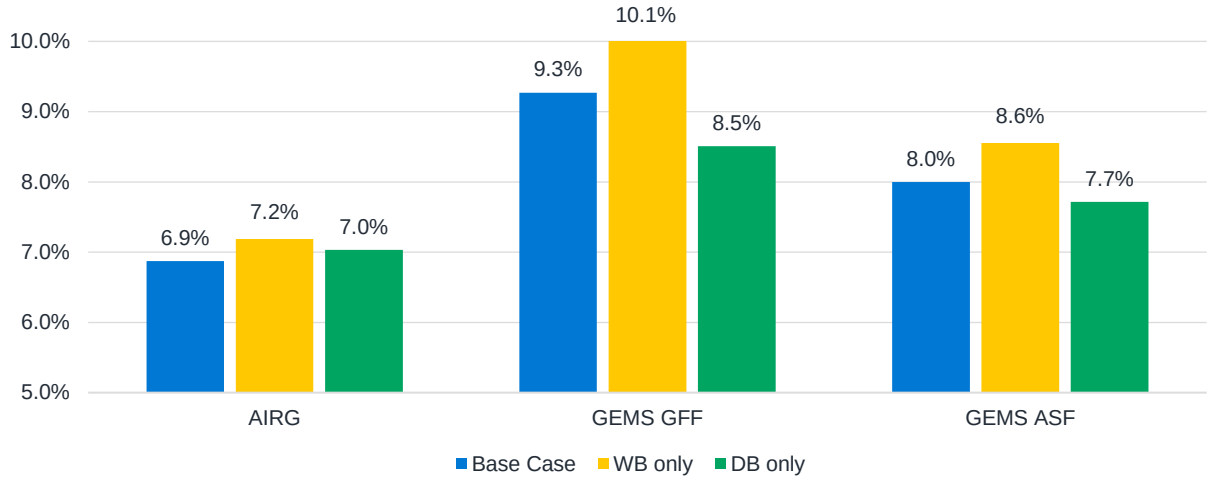
New business exhibiting a lower TAR is in line with expectations, because it reflects the reserving of the future fee income that is typically collected earlier in the block's life cycle, whereas the existing business has likely already passed that stage and is experiencing withdrawal activity at the onset. Relatedly, we observe that the new business reflects more scenarios being floored at CSV in the CTE70 reserve calculation, although it does not skew the findings based on the 400% C-3 TAR metric. With a focus primarily on the interaction of the block age with scenario sets, we observe the following:

- **Overall trend**: There is a general consistency in the TAR across scenario sets between existing versus new business—specifically, the GEMS scenarios exhibit higher TAR on both unhedged and hedged bases compared to the AIRG.
- **Impact of new business**: The TAR impact is lower for the GEMS scenarios than it is for the AIRG set. We observe approximately a 1.5% reduction in TAR for new business under GEMS scenarios, relative to an approximate 2.0% reduction for the AIRG scenarios. This could be attributable to the lower diversification benefit (and therefore more uncertainty being reserved for) driven by the GEMS equity model. The additional reserve held for this is being released more slowly over time.
- **Market sensitivities**: Although not displayed in Figure 11, the delta and rho of the new business also shows a relationship across scenario sets that is comparable to the existing business.

RIDER MIX

Within the baseline run set, the assumed two-thirds weight on GLWBs and one-third weight on GMDBs capture an illustrative mix of riders that reflects natural longevity versus mortality risk diversification. The richness of the respective GLWBs and GMDBs plays a role, but generally the GLWB's longevity risk is at least partially offset by the GMDB's mortality risk. The average durations of GLWB and GMDB claims vary, and due to the dispersion differences across the scenario sets, along with the pace of the accumulation factors described earlier, the impact of the scenarios could vary according to the mix of rider guarantees. Figure 12 summarizes the TAR compared to the GLWB/GMDB mix in the baseline set.

FIGURE 12: UNHEDGED 400% C-3 (AS % OF AV), BY RIDER



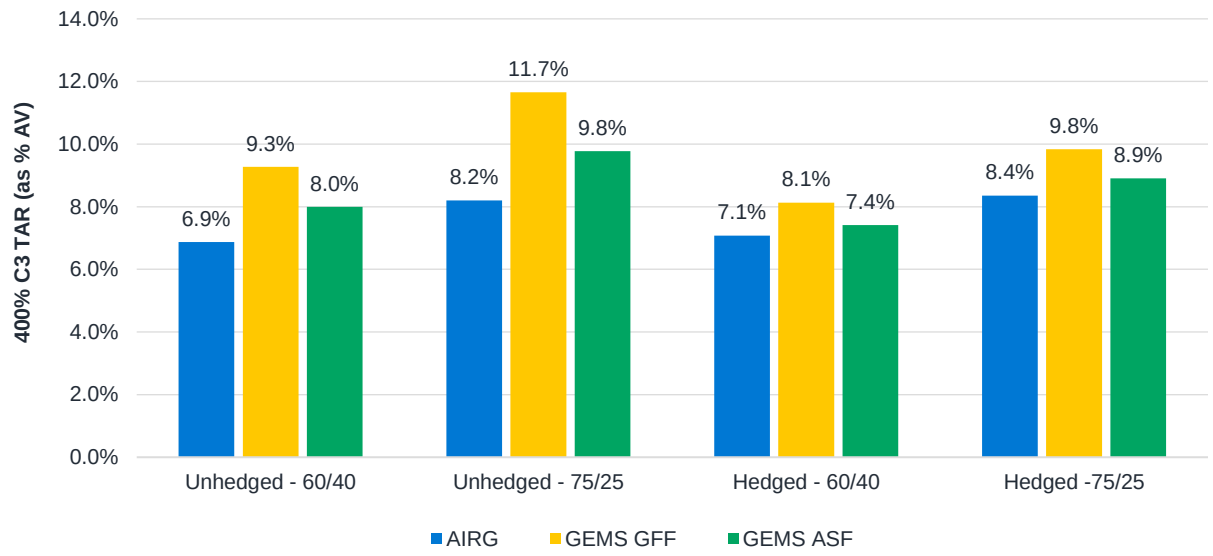
By looking at a weighted average of the standalone GLWB and GMDB TAR compared to our base case, we can infer the amount of diversification benefit stemming from the rider mix. This turns out to be consistent at 0.27% of TAR (as a percentage of AV), or approximately $53 million on a block of $20 billion AV. From the unhedged TAR in Figure 12, we note the following observations.

- **GLWB TAR**: The GLWB rider requires a larger TAR on a standalone basis than the base case rider mix across all scenario sets. This difference is most notable for GEMS GFF, where GLWB-only TAR is +0.8% more than the base case, whereas the AIRG is only +0.3%.
- **GMDB TAR**: The standalone GMDB TAR is lower than the base case rider mix for the GEMS scenarios, but is larger than the base case for the AIRG. After further review, this pattern is a function of the assumed 25% ITM and differences in GMDB versus GLWB sensitivity. Tests at 50% ITM then show that the standalone GMDB TAR aligns with the pattern observed on GEMS scenarios, where it is less than the base case rider mix.
- **GLWB vs. GMDB mix**: Even though the scenario sets exhibit comparable diversification benefit levels, the relationship of GLWB versus GMDB TAR is more substantial in the GEMS scenarios. This may suggest more sensitivity to GEMS scenarios for business with a higher share of living benefit guarantees and longevity risk, although other factors in the GLWB or GMDB product features may also contribute to varied sensitivity.

FUND MIX

A review of the fund mix can help to understand how the equity versus bond allocation of a block may drive the TAR impact from the new scenarios. Based on general observations of the equity and fixed income returns, the changes to equity returns introduced in the GEMS model leads to higher TAR, whereas changes to the fixed income returns in the GEMS model leads to lower TAR. The chart in Figure 13 shows TAR with increasing the equity allocation from 60% to 75% of the total AV.

FIGURE 13: FUND MIX: 60% VS. 75% EQUITY



Recall that, for this illustrative block, 1% of AV is $200 million. As one might expect, greater equity allocation reflects more risk, so the increases in TAR are reasonable. We observe the following:

- **GEMS equity mix sensitivity**: The GEMS scenarios, particularly the GEMS GFF set, experience a significant increase in TAR from the additional equity exposure. In contrast, the AIRG set has the lowest total equity volatility, and is less impacted.
- **Hedging impact**: Hedging has a material effect on the GEMS scenarios to reduce sensitivity to the fund mapping, despite the increased cost of hedging reflected from higher equity allocation. Hedging has minimal effect on the AIRG set, and this observation could be attributable to the relative level of volatility in the AIRG scenarios being lower than risk-neutral.

Conclusion

The current scenario sets reviewed as part of the NAIC's scenario reform work stream present a vast opportunity for quantitative analysis on variable annuity reserves and capital under the PBR framework. This paper serves to review several dimensions in that type of analysis by using a prototypical block in a manner that is transparent and consistent across quantifications.

Based on the impact analysis across the four scenario sets considered, there are some common themes in TAR levels. We consistently observe that the ranking from lowest to highest TAR is the AIRG, GEMS ASF, and GEMS GFF. The GEMS scenarios exhibit higher interest rates and fixed income returns, but also higher interest rate and equity volatility, in comparison to the AIRG. There is also some variability in equity return patterns that suggests a mixed view on GEMS versus the AIRG.

In addition to these broad themes, the analyses performed highlighted the following:

- The impact across individual companies will vary significantly depending on rider mix, moneyness level, fund mix, block age, and other factors.
- Robust dynamic hedging is the most effective way to insulate business from potential scenario reform impacts.
- Rate sensitivity is likely to increase, potentially bridging a statutory balance sheet closer to a more "economic" sensitivity profile. Equity sensitivity may be more complex but could also increase.

These points may hint at the potential benefits and consequences of companies already hedging movements in their statutory liability. We will continue to monitor developments in the NAIC's scenario reform efforts and in future papers we will seek to refresh our analysis with new iterations of scenarios that are released through a second field test. Additionally, we may contemplate what companies can do to best prepare for the anticipated reform.

Lastly, it is important to note that the analysis shared in this paper is illustrative and may not be reflective of what any single company might observe when testing the scenario sets using its unique combination of product design, actuarial assumptions (particularly dynamic policyholder behavior), and mix of in-force business. We encourage readers to carry out their own testing and analysis of the scenarios to understand the potential implications on the reserve and capital requirements for their variable annuity portfolios.

Milliman

Milliman is among the world's largest providers of actuarial, risk management, and technology solutions. Our consulting and advanced analytics capabilities encompass healthcare, property & casualty insurance, life insurance and financial services, and employee benefits. Founded in 1947, Milliman is an independent firm with offices in major cities around the globe.

milliman.com

CONTACT

Collin Davidson
collin.davidson@milliman.com

Zohair Motiwalla
zohair.motiwalla@milliman.com

Parker Henley
parker.henley@milliman.com