NAIC Economic Scenario Generator Field Test: VM-20 Quantitative Results

Scott O'Neal FSA, MAAA

May 18, 2023

NAIC NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS

Agenda

1. Background and Purpose
2. Limitations
3. Field Test Run Descriptions
4. Field Test Participation
5. High-Level Observations
6. Quantitative Results
   A. DR/SR Baseline Comparisons
   B. DR/SR Valuation Date Comparisons
   C. VM-20 Minimum Reserve Impact
7. Next Steps

Appendix 1: SERT Scenario Overview
Appendix 2: Treasury and Equity Scenario Overview

Background and Purpose

- The purpose of this presentation is to summarize quantitative information from the VM-20 field test participants to:
  - Understand the impact on reserves and capital,
  - Review the range of results across field test participants,
  - Compare the stability of results over time, and
  - Inform regulator decision-making on model and calibration choices.

Limitations

- The NAIC took steps to review the quantitative results for reasonableness, including reviewing qualitative survey responses, sending questions to participants, and asking participants to confirm that the NAIC compilations matched their intended result submission. However, the accuracy and reliability of the results are ultimately dependent on the quality of participant submissions.
- The field test analytics (average reserves, range of impacts, etc.) can be strongly dependent on a subset of the participants. Results shown today for the different field test runs will include varying numbers of participants corresponding to the levels of participation for that run. The lack of participation in some of the runs will limit their applicability to the overall industry.
- A number of comparisons between company-provided field test or baseline runs are made in the presentation. These comparisons are limited to the participation of whichever run had the least participation. For example, as Baseline 2 (as of 12/31/19 + 200 BP) had significantly lower participation than run 2A, many of the 2A results will not be included in the baseline comparison.
- Only three of the 15 companies made changes to their models to account for different features of the field test scenario sets (e.g. negative interest rates). Therefore, field test results may not be fully representative of company results post-implementation of the new GOES.
- Some companies mentioned that they would assess the need for changes to their assumptions prior to implementation of the new GOES but had not done so for the field test.
- Some of the field test SERT scenario sets contained errors, including the deterministic reserve (DR) scenario #12. Therefore, deterministic results cannot be shared for field test runs 5A, 5B, and 6.
- The VM-20 portion of the qualitative survey did not ask companies to specifically comment on the drivers of their results as was done for VM-21/C3 Phase II. Most companies did not comment on the drivers of their results.
- Variable and indexed products are included in the GOES field test VM-20 results, but isolating the specific impacts is challenging as some participants included those products with others in the same reserving category in one model (e.g. a model containing VULSG with ULSG). Further, we do not have data on the participants' separate account fund mapping.

Field Test Run Descriptions

Note: Bold = Required Run

Run #	Description	Purpose of Run
Baseline #1	Scenario set(s) the company used for 12/31/21 statutory reporting	Baseline used as comparative basis for 12/31/21 runs
Baseline #2	ESG the company used for 12/31/21 statutory reporting of reserves and RBC, but modified to produce scenario sets with a 12/31/19 yield curve modified using a 200 BP increase across all maturities	Baseline used as comparative basis for 12/31/19 + 200 BP runs
Test #1a	GEMS Baseline Equity and Corporate model scenarios as of 12/31/21, and Conning Treasury model calibration with generalized fractional floor as of 12/31/21	Tests Conning Treasury model w/ GFF and Baseline Equity at YE 2021
Test #1b	Same as Test #1a, but with Alternative Treasury model calibration with shadow floor as of 12/31/21	Tests Alternative Treasury model with shadow floor and Baseline Equity at YE 2021
Test #2a	Same as Test #1a, but with Equity, Corporate, and Treasury models with a 12/31/19 starting yield curve modified using a 200 BP increase across all maturities. All other initial market conditions are unchanged. The Equity model parameters would be adjusted from #1a so that the year 30 median Large Cap Equity gross wealth factors remain consistent with #1a.	Stresses the starting Treasury rates using the same calibration as 1a to evaluate whether the model produces appropriate results in different economic environments
Test #2b	Same as Test #2a, but with the Alternative Treasury model calibration with shadow floor instead of the Conning Treasury model calibration with generalized fractional floor	Same as 2a, but designed to stress the 1b calibration

Field Test Run Descriptions

Note: Bold = Required Run

Run #	Description	Purpose of Run
Test #3	Conning Treasury model calibration with generalized fractional floor as of 12/31/21, GEMS Corporate model as of 12/31/21, and GEMS Baseline Equity model corresponding to a 12/31/19 yield curve with a 200 BP increase across all maturities	Attribution analysis that will illustrate how much of the difference between runs #1a and #2a is driven by the equity model vs the Treasury and Corporate models
Test #4	Same as Test #3, but using Alternative Treasury model calibration with shadow floor as of 12/31/21	Same as #3, but with respect to runs #1b and #2b.
Test #5a	Same as #1a, but with Conning's original Equity model calibration that had significantly lower Gross Wealth Factor's (GWFs) than the AIRG Equity Model.	Tests Conning Treasury model w/ GFF and original equity model as of year-end 2021.
Test #5b	Same as #5a but using a 12/31/19 starting yield curve modified using a 200 BP increase across all maturities. The parameters of Conning's original Equity model are used without any adjustment.	Stresses the starting Treasury rates to understand the full impact of equity-Treasury linkage in Conning's original equity model
Test #6	Same as #1a, but with the ACLI's GEMS® Equity Calibration	Tests the ACLI's GEMS® Equity Calibration that assumes a constant mean equity return independent of rates and increases alignment with AIRG equity model GWFs

Field Test Participation

- The chart below shows the number of legal entities that submitted VM-20 results for the field test by reserving category and reserve component.
- Many companies submitted multiple products, and some submitted multiple model segments for a given reserving category. Other companies aggregated products with distinct risks (e.g. Variable Universal Life with Secondary Guarantee, vanilla Universal Life with Secondary Guarantee) into a single model segment (e.g. ULSG). Ranges of results shown in the presentation are reflective of a legal entity view, rather than a model segment view.
- There are two basic types of comparisons of the field test results in this presentation; 1) comparisons of field test runs to their respective baseline run, and 2) comparisons of field test runs across the two tested valuation dates. These comparisons are limited by the run with the least participation (e.g. comparisons to the baseline for the 12/31/19 + 200 BP valuation date are limited to Baseline 2 participation).

Participation by Legal Entity

VM-20 Reserving Category	SR/DR	Baseline 1*	Baseline 2	1A*	1B*	2A*	2B*	5A*	5B*	6
Term	DR	11	<5	11	10	9	9	10	10	<5
ULSG	DR	11	<5	11	11	11	11	9	9	<5
Other	DR	<5	<5	<5	<5	<5	<5	<5	<5	<5
Combined	DR	15	6	15	14	14	14	11	11	6
Term	SR	<5	<5	<5	<5	<5	<5	<5	<5	<5
ULSG	SR	9	<5	10	10	10	10	8	8	<5
Other	SR	<5	<5	<5	<5	<5	<5	<5	<5	<5
Combined	SR	11	<5	11	11	11	11	8	8	<5

Participation by Model Segment

Product	Number of Model Segments	Variable?	Indexed?
Term	15	N/A	N/A
ULSG	20	7	4
Whole Life	3	N/A	N/A
Universal Life	1	0	0

Valuation Dates: 12/31/21 | 12/31/19 + 200 BP

*Required Run

High-Level Observations

- When directly comparing baseline DR to field test DR results or baseline SR to field test SR results, there was a wide range of impacts across participating legal entities. Some legal entities saw large increases to their modeled reserves, and others experienced decreases. The range of results was in some cases greater when looking at a model segment level, with some model segments exhibiting much larger increases than were seen at a legal entity level. The range of modeled results by legal entity, however, was much smaller than it was for the VM-21/C3 Phase II GOES field test.
- While the range of modeled results was wide, the average increase to VM-20 minimum reserves by legal entity was muted given the domination of the NPR for many participants, even with large increases to modeled reserves. As VM-20 only became mandatory in 2020, the dominance of the NPR could be related to how recently the business was issued and may not be reflective of a mature block.
- Valuation date comparisons across baseline and field test runs were challenging given the limited participation in Baseline 2. For the DR considering all reserve categories combined, the field test runs were not, on average, more variable across valuation dates compared to the baseline runs. For SR, there was not enough participation in Baseline 2 to compare the change in valuation date results for field test runs to the baseline runs. However, for both DR and SR, the average change across valuation dates and the range of results were significantly smaller in magnitude than the results shown for VM-21.

Baseline Reserve Comparisons: Term Reserving Category

Change in Deterministic Reserve by Legal Entity

Field Test	1A	1B
Average % Increase	29%	19%
# of Participants	11	10

Percentage Increase: Range and Percentile Statistics

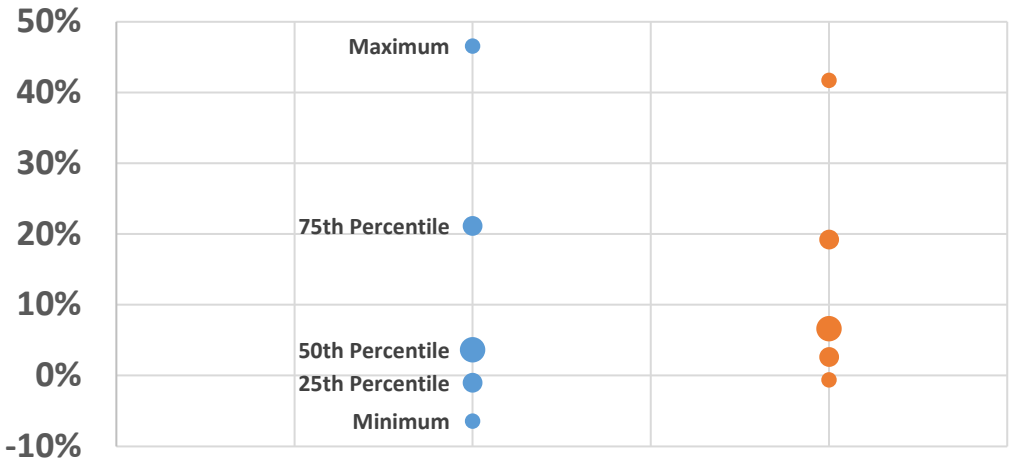
- Limited participation and SERT scenario errors did not allow for public sharing of DR baseline comparisons for 2A, 2B, 5A, 5B, and 6.
- Approximately half of the participant's Baseline 1 Term deterministic reserves were negative. Comparisons between relatively small negative values, or values that change signs between field test runs require adjustments to the standard (B-A)/A formula that typically is used for percentage change. The formula that was used was as follows: **Absolute Value** [(B-A)/A)] * IF(B<A, -1, 1)
- The 1A (Conning Treasury and Baseline Equity scenario set as of 12/31/21) average DR increase of 29% was significantly larger than the 19% average DR increase seen in 1B (Alternative Treasury with Baseline Equity parameters).
- Field test participants saw more variation in the field test 1A reserve impacts, with a higher maximum (105%) and lower minimum (-96%) than what was seen in 1B.
- For both 1A and 1B,
  - the maximum end of the range was from a positive baseline reserve increasing, and
  - The minimum end of the range was from a negative baseline reserve becoming more negative.

Baseline Reserve Comparisons: ULSG Reserving Category

ULSG Reserve Category: Deterministic Reserve (DR) Change from Baseline by Legal Entity

Field Test	1A	1B
Average % Increase	2%	6%
# of Participants	11	11

Percentage Increase: Range and Percentile Statistics

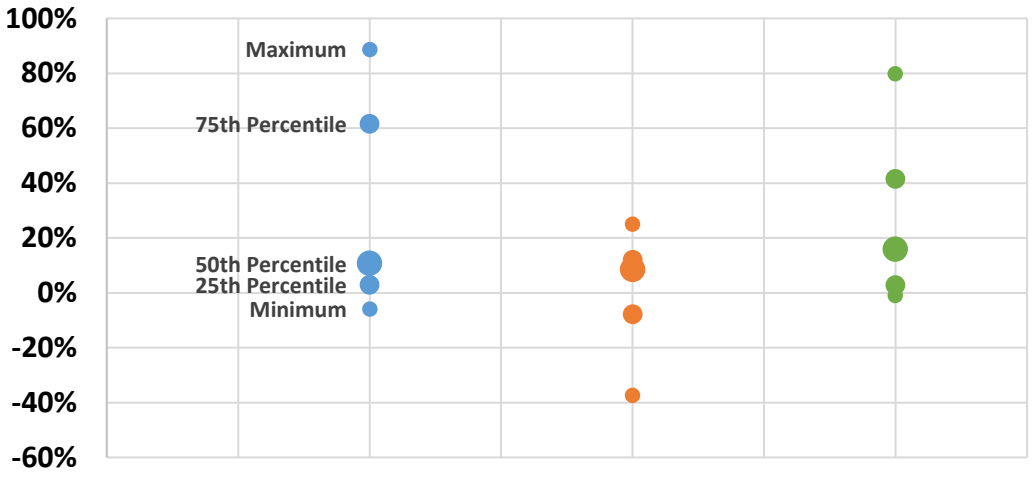


- Limited participation and SERT scenario errors did not allow for public sharing of DR baseline comparisons for 2A, 2B, 5A, 5B, and 6.
- The 1A (Conning Treasury and Baseline Equity scenario set as of 12/31/21) average DR increase of 2% was relatively smaller than the 6% average DR increase seen in 1B (Alternative Treasury with Baseline Equity parameters). A partial explanation for the higher average DR in 1B could be related to lower S&P 500 equity gross wealth factors (GWFs) present in 1B in later years of the projection compared to 1A.
- Field test participants saw more variation in the field test 1A results, with a higher maximum (47%) and lower minimum (-6%) than in 1B.
- Model segment level results fell within the legal entity level ranges for all but one of the participants.

Change in Stochastic Reserve by Legal Entity

Field Test	1A	1B	5A
Average % Increase	19%	11%	21%
# of Participants	9	8	7

Percentage Increase: Range and Percentile Statistics



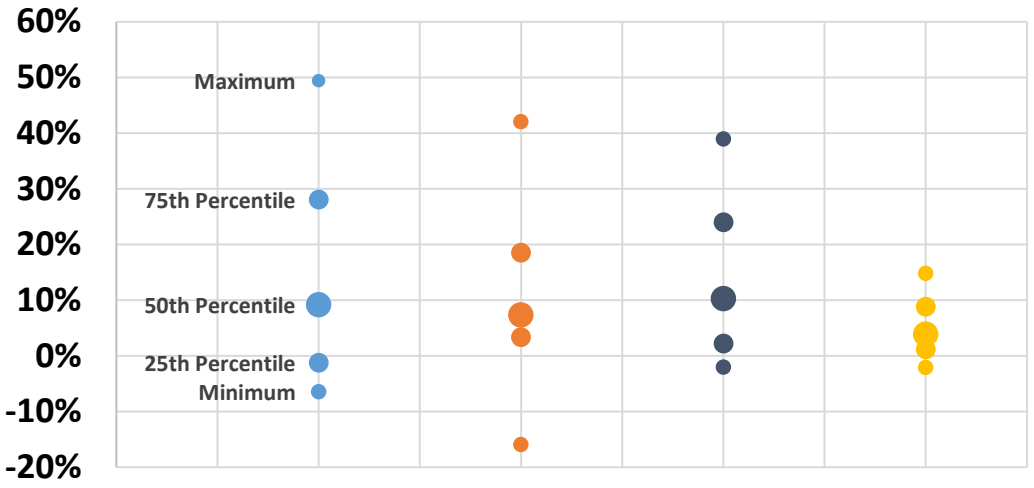
- Limited participation did not allow for public sharing of SR baseline comparisons for 2A, 2B, 5B, and 6.
- The 1A (Conning Treasury and Baseline Equity scenario set as of 12/31/21) average SR increase of 19% was significantly larger than the 11% average SR increase seen in 1B (Alternative Treasury with Baseline Equity parameters).
- Field test 5A (Conning Treasury and original Conning Equity calibration with lower equity GWFs) saw the highest average stochastic reserve increase. The treasury scenarios in 5A were the same as 1A, but the lower equity GWFs present in 5A resulted in larger reserve increases for indexed and variable life products in 5A compared to 1A.
- There was a higher maximum reserve increase in the field test 1A results compared to 1B, and 5A.
- When looking at the range of results at the individual model segment level, there were a number of reserve increases that were greater than those shown in the chart on the left. A company with one of these large model segment impacts noted that the increases would put their reserves higher than AXXX reserves.

Baseline Reserve Comparisons: Combined Reserving Categories

Combined Reserve Categories: Deterministic Reserve Change from Baseline by Legal Entity

Field Test	1A	1B	2A	2B
Average % Increase	3%	7%	8%	2%
# of Participants	15	14	6	6

Percentage Increase: Range and Percentile Statistics



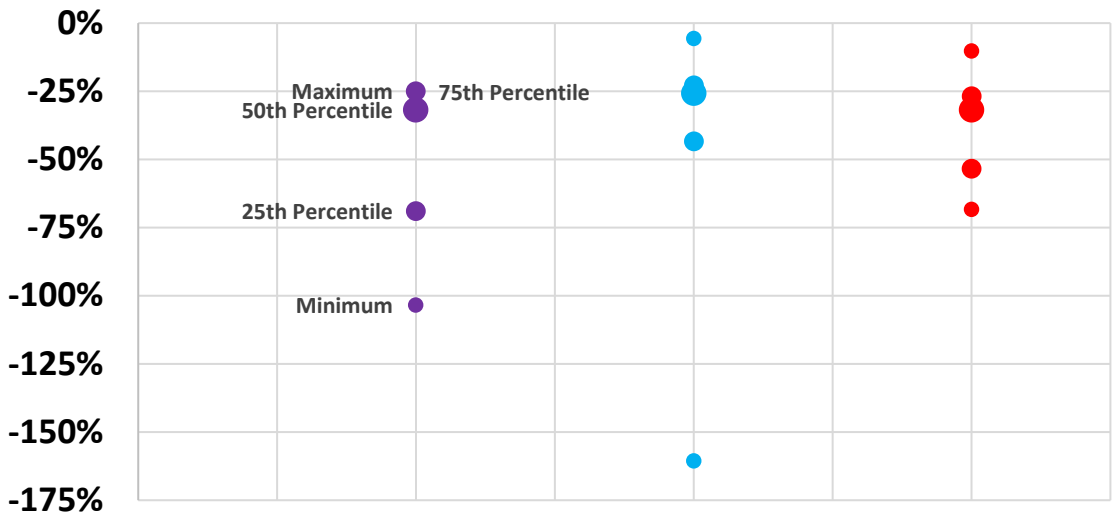
- The results shown on this page are reflective of the aggregated Term, ULSG, and Other (as applicable) model segment results by legal entity. Combining reserve categories increases the number of participants, allowing 2A and 2B results to be shared.
- Limited participation and SERT scenario errors did not allow for public sharing of DR baseline comparisons for 5A, 5B, and 6.
- ULSG represented over 97% of the Baseline 1 deterministic reserves in the combined category, and just over half of the model segments.
- The 1A (Conning Treasury and Baseline Equity scenario set as of 12/31/21) average DR increase of 3% was smaller than the 7% average DR increase for 1B (Alternative Treasury with Baseline Equity parameters). However, the relationship flipped for the 12/31/19 + 200BP field test runs shown, with a larger average DR increase of 8% for 2A compared to a smaller increase of 2% for 2B (both compared to Baseline 2).

Valuation Date Reserve Comparisons: Combined Reserving Categories

Change in Deterministic Reserve by Legal Entity

Field Test	B2 vs B1	2A vs 1A	2B vs 1B
Average % Increase	-29%	-22%	-28%
# of Participants	6	14	14

Percentage Decrease: Range and Percentile Statistics

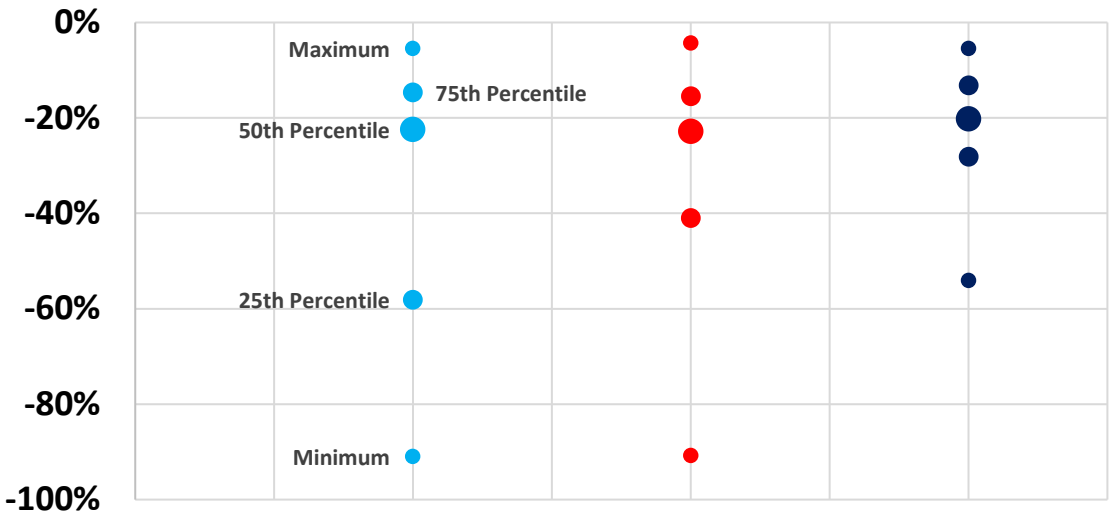


- Limited participation and SERT scenario errors did not allow for public sharing of DR valuation date comparisons for 5B vs 5A
- Across the baseline and field test runs, reserves significantly decreased in the 12/31/19 + 200 BP (higher starting interest rate level) runs compared to the 12/31/21 (lower starting interest rate) runs.
- The average percentage decrease was similar across the field test runs, although the comparison to the Baseline runs was challenging given the limited participation.
- The range of results was highest for the 2A vs 1A comparison. The largest decreases were driven by comparisons where the term DR was negative in both the 1A and 2A runs.

Change in Stochastic Reserve by Legal Entity

Field Test	2A vs 1A	2B vs 1B	5B vs 5A
Average % Increase	-24%	-22%	-22%
# of Participants	11	11	8

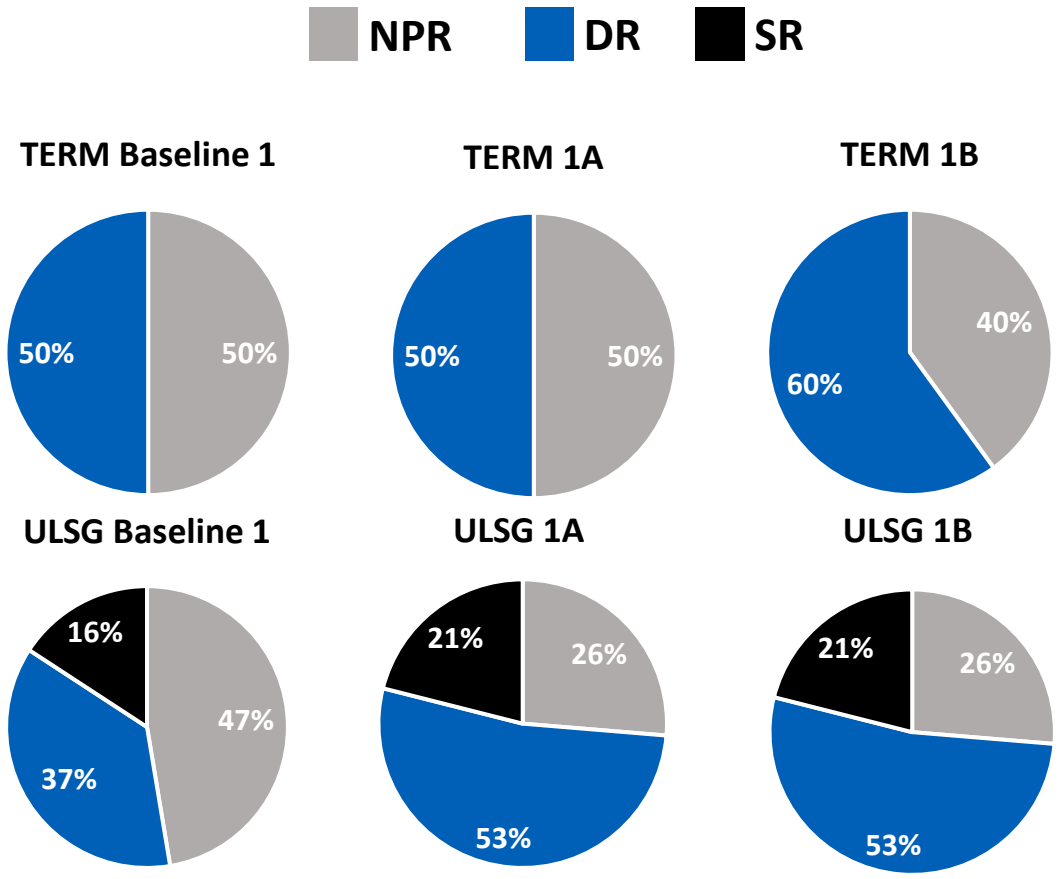
Percentage Decrease: Range and Percentile Statistics



- Limited participation did not allow for public sharing of SR valuation date comparisons for Baseline 2 vs. Baseline 1.
- Across the baseline and field test runs, reserves significantly decreased in the 12/31/19 + 200 BP (higher starting interest rate level) runs compared to the 12/31/21 (lower starting interest rate) runs.
- The average percentage decrease in the SR was similar across the different field test run comparisons.
- The large range of results was similar across the 2A vs 1A and 2B vs 1B comparisons, but somewhat narrower in the 5B vs 5A (same UST as 1A/2A, but Conning original equity model with equity Treasury Linkage) comparison. This result is somewhat counterintuitive, given the additional variation in the Equity GWFs between valuation dates present in the 5B vs 5A comparison. This can be partially explained by:
  - Some companies included variable, indexed, and/or "vanilla" ULSG in the same model segment making it challenging to isolate impacts,
  - Limited indexed and variable product participation, and
  - There were less participants in the 5B vs 5A comparison.

VM-20 Minimum Reserve Impact

Proportion of Reserve Category/Model Segments by Dominant Reserve Type

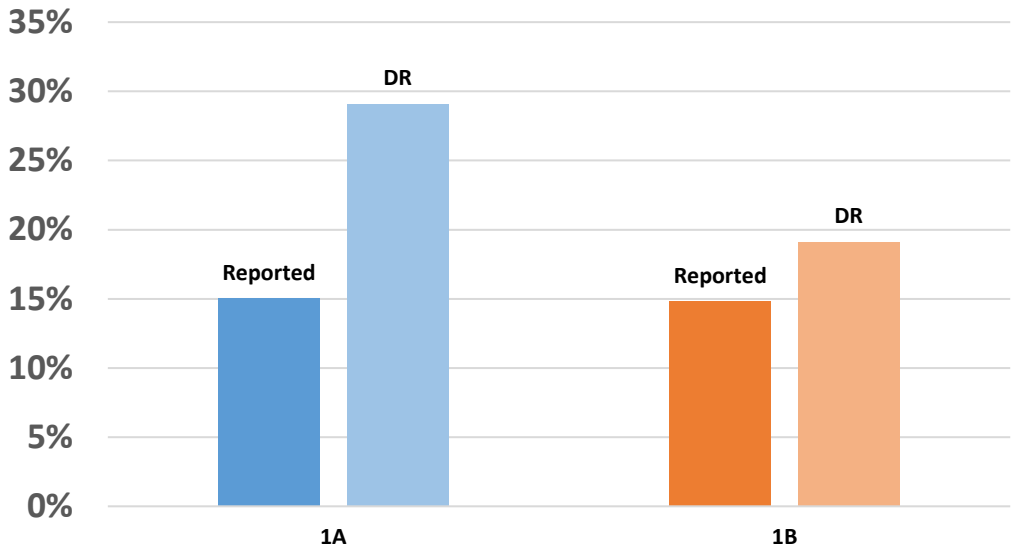


- A partial survey of 2021 PBR Actuarial Reports indicated that:
  - Of 99 companies that included Term results, 63% held NPR, 35% held the DR, and the remaining 2% had the stochastic reserve as the dominant reserve, and
  - Of 68 companies that included ULSG results, 57% held the NPR, 31% held the DR, and the remaining 12% held the SR as the dominant reserve.
- For the term reserving category, approximately half of the participants held negative deterministic reserves for their Baseline 1 submission.
- While the chart for Term 1B seems to indicate a switch from NPR to DR, the change in proportion of NPR/DR is entirely due to less participation in 1B.
- Almost half of the participant ULSG products held a net premium reserve as their minimum reserve for Baseline 1. For field tests 1A and 1B, there was a large shift to the deterministic reserve and a smaller shift to the stochastic reserve as the dominant reserve.
- Although the proportions of winning NPR, DR, and SR are the same across ULSG 1A and 1B, there was movement in the winning reserve type for some model segments between 1A and 1B.

Term Reserve Category: VM-20 Minimum Reserve Change from Baseline

Field Test	1A	1B
# of Participants	11	10

Average % Reserve Increase: Minimum and DR

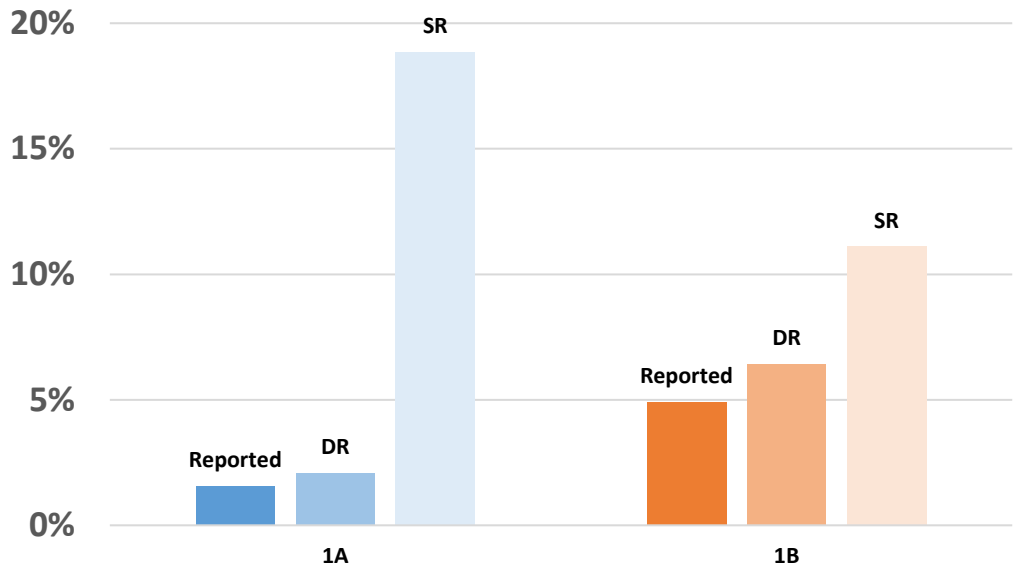


- The graph on the left shows average percentage increases in the VM-20 minimum reserve and DR for the Term Reserving Category.
- Despite reserve increases for many of the participants for their field test modeled reserve runs (DR), the effect on the legal entity level minimum reserve was muted due to the net premium reserve still dominating in many cases.
- Field test 1A saw a larger increase to DR than 1B, but the change to the average reported (minimum) reserve was very similar due to:
  - There were no companies that switched dominant reserves from their Baseline 1 result to either the 1A or 1B for the Term Reserving Category. For the companies where the NPR was the dominant reserve, the change in reported reserve was zero.
  - When the DR was the winning reserve, some companies had larger increases in 1A and others saw larger increases in the 1B run.
- The dominant reserve may change throughout a product's lifecycle. PBR only became mandatory in 2020, so all of the business was recently issued. Therefore, these results may not be applicable to business that is in a more mature phase.

ULSG Reserve Category: VM-20 Minimum Reserve Change from Baseline

Field Test	1A	1B
# of Participants	11	11

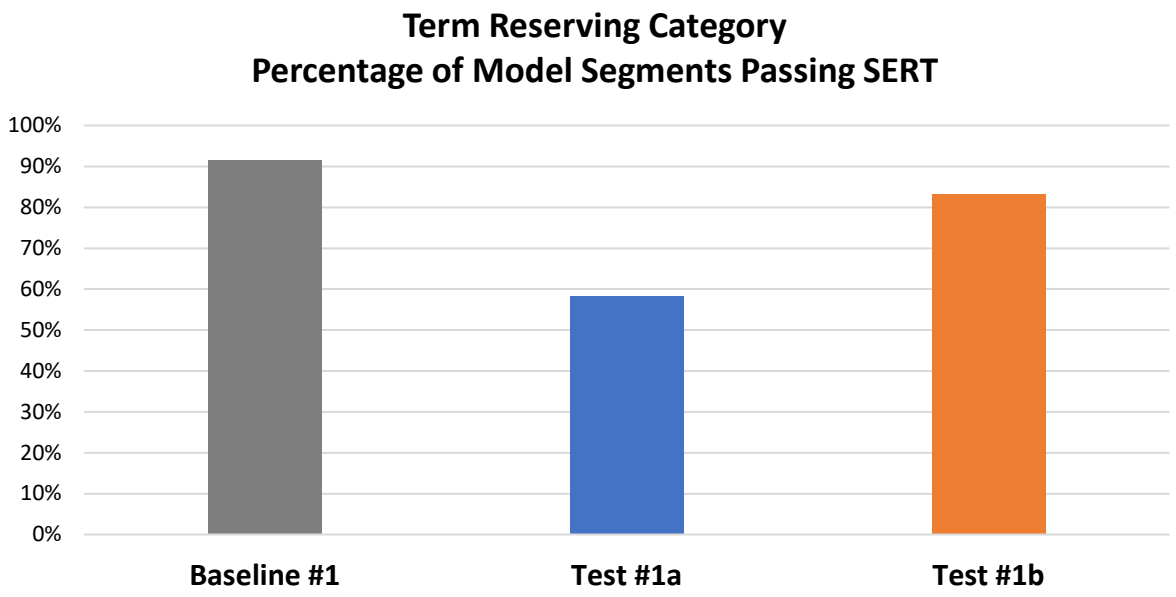
Average % Reserve Increase: Minimum, DR, and SR

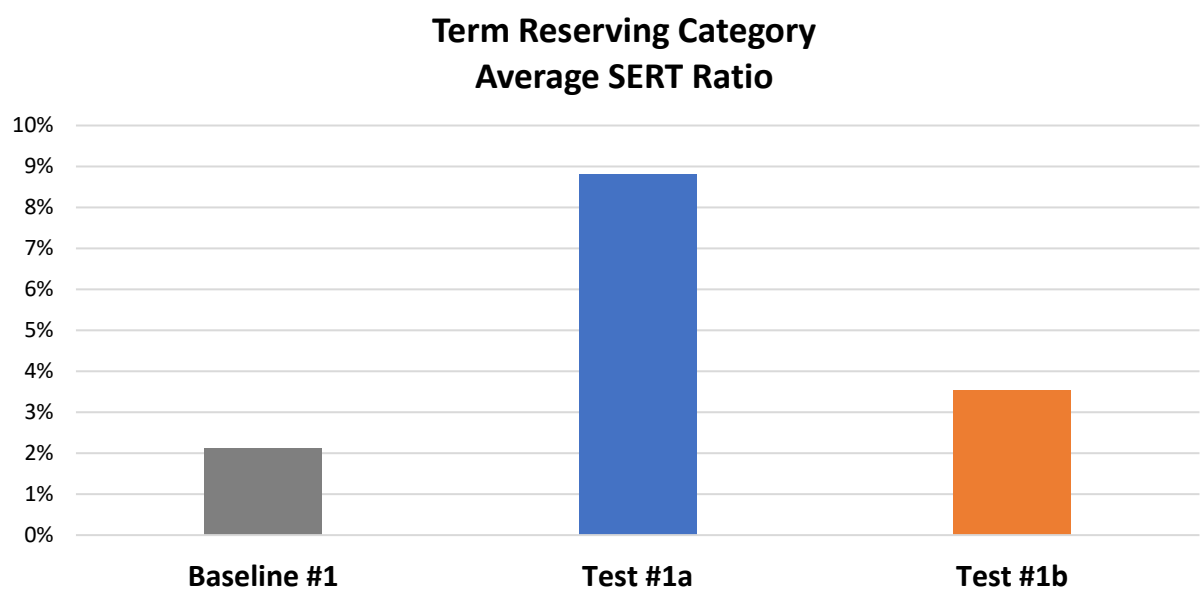


- The graph on the left shows average percentage increases in the VM-20 minimum reserve, DR and SR for the ULSG Reserve Category.
- Despite reserve increases for many of the participants for their field test modeled reserve runs (DR and SR), the effect on the legal entity level minimum reserve was muted due to:
  - the net premium reserve still dominating in many cases, and
  - several of the largest increases to modeled reserves did not end up being the winning reserve.
- The dominant reserve may change throughout a product's lifecycle. PBR only became mandatory in 2020, so all of the business was recently issued. Therefore, these results may not be applicable to business that is in a more mature phase.

Stochastic Exclusion Ratio Test (SERT) Scenario Results

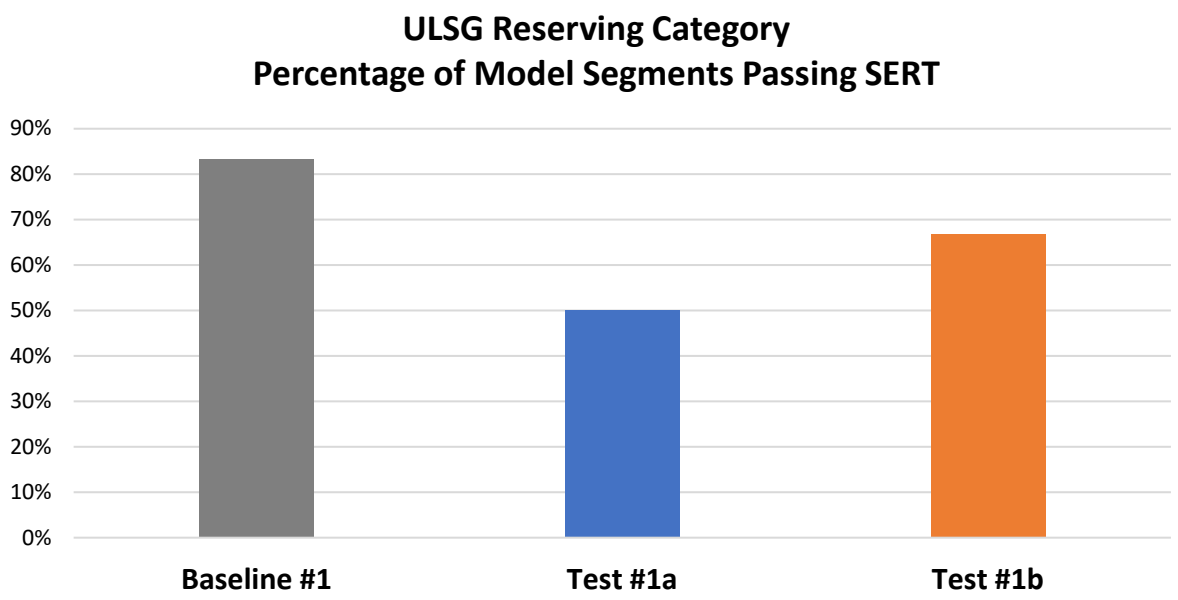
Field Test SERT Results - Term

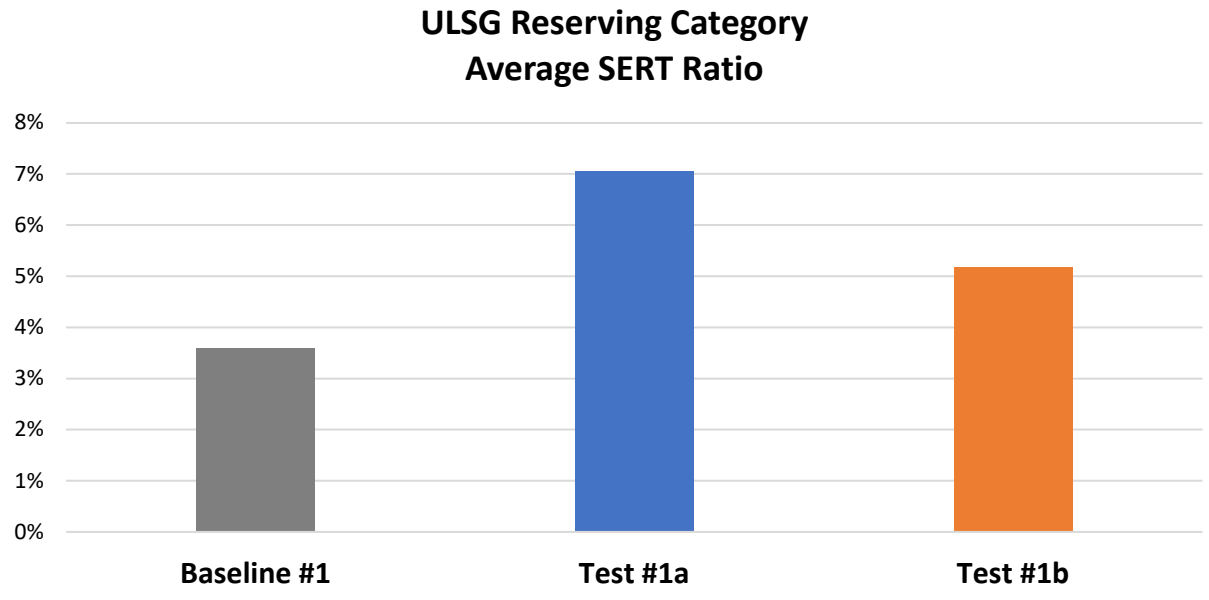




- As compared to company Baseline #1 results, less of the field test run term model segments passed the SERT, with the biggest drop-off seen for the Conning Calibration w/ GFF (1A).
- The average (non-weighted) SERT result for term model segments increased for the field test runs compared to Baseline #1. Average SERT ratios increased the most for the Conning Calibration w/ GFF (1A).
- For the term model segment, the “b” largest adjusted DR scenario was mostly consistent for a given model segment between the different field test runs. However, across model segments/legal entities, different “b” SERT scenarios were constraining.

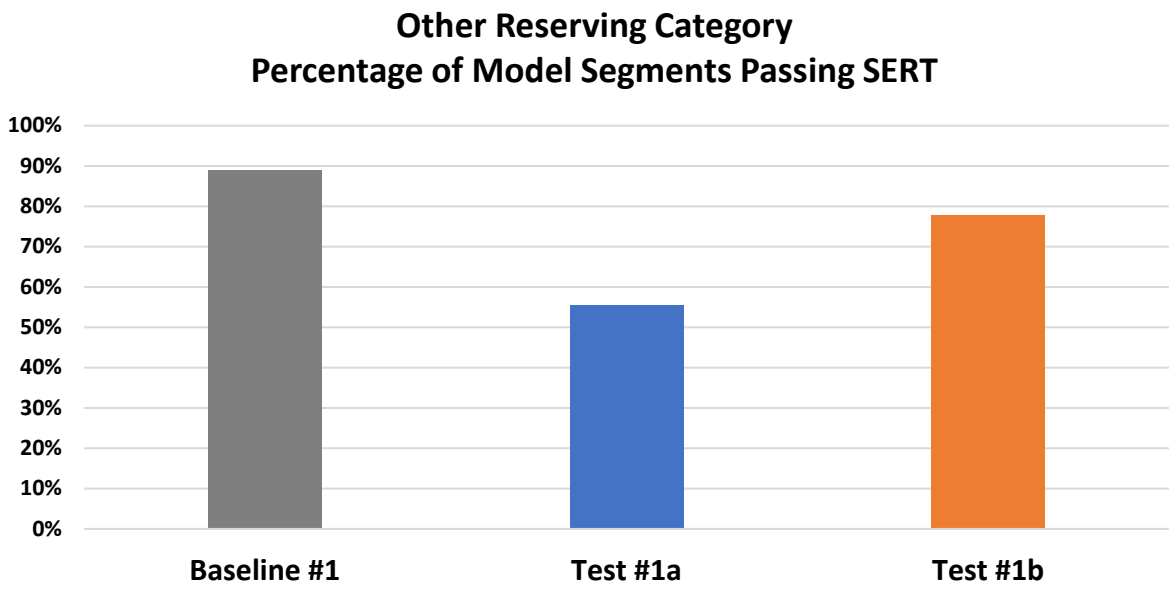
Field Test SERT Results - ULSG

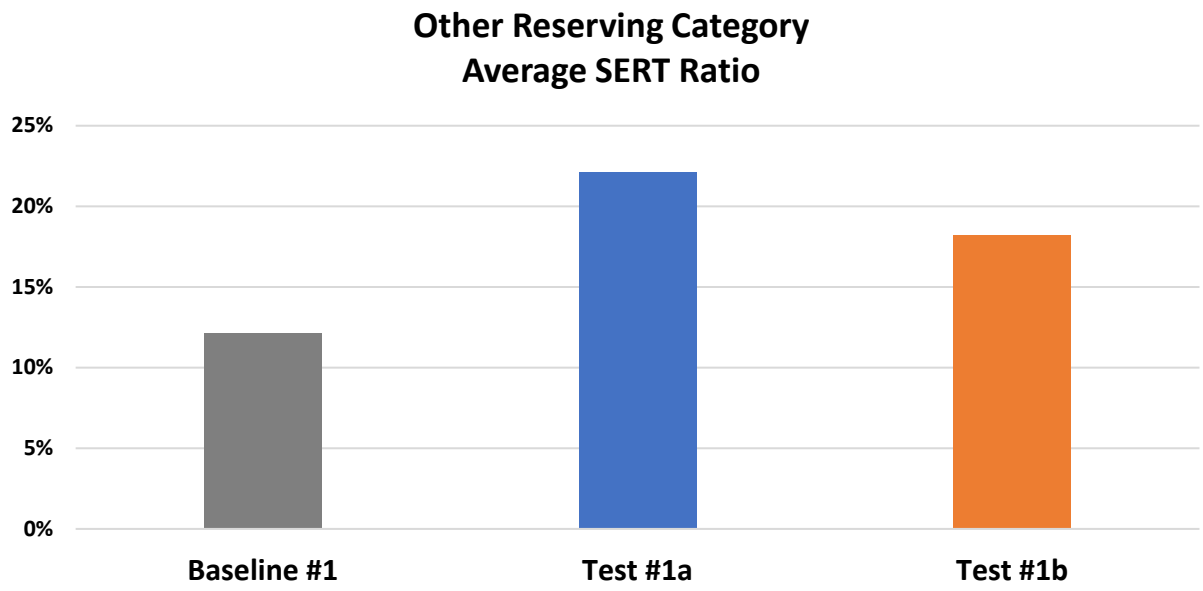




- As compared to company Baseline #1 results, less of the field test run ULSG model segments passed the SERT, with the biggest drop-off seen for the Conning Calibration w/ GFF (1A).
- The average (non-weighted) SERT result for term model segments increased for the field test runs compared to Baseline #1. Average SERT ratios increased the most for the Conning Calibration w/ GFF (1A).
- The “b” scenario in the SERT calculation fluctuated between field test runs for some ULSG model segments but was stable in others

Field Test SERT Results - Other





- As compared to company Baseline #1 results, less of the field test run ULSG model segments passed the SERT, with the biggest drop-off seen for the Conning Calibration w/ GFF (1A).
- The average (non-weighted) SERT result for term model segments increased for the field test runs compared to Baseline #1. Average SERT ratios increased the most for the Conning Calibration w/ GFF (1A).
- For the Other model segment, the “b” scenario frequently changed between the baseline and field test runs. Of those that change, most switched to a pop-down UST SERT scenario. Across model segments/legal entities, different “b” SERT scenarios were constraining.

Next Steps

- The NAIC will look to present economic scenario generator field test results for the C3 Phase I in late June. Additional time for follow-up discussions may be necessary.
- Regulators will continue to work with interested parties in economic scenario generator drafting groups to continue progress on reserve/capital framework specific implementation tasks.
- The Life Actuarial (A) Task Force will engage with the American Academy of Actuaries and other interested parties to decide on stylized facts and acceptance criteria ahead of a recalibration of the economic scenario generator and a second field test.

Appendix 1: Stochastic Exclusion Test Ratio (SERT) Scenario Overview

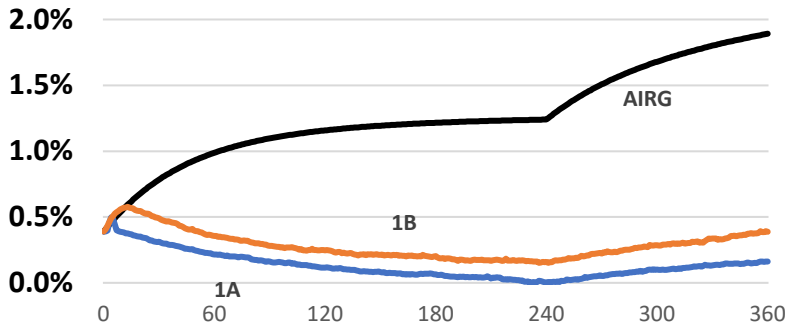
Deterministic Reserve 12/31/21 Scenario Statistics

- SERT Scenario 12 (the DR scenario) has significantly lower UST rates for 1A/5A/6* and 1B compared to the AIRG. Lower and longer interest rates can tend to increase VM-20 reserves due to, for example, challenges with companies being able to reinvest in assets with enough yield to support minimum crediting rates and/or a lower discount rate on future claim payments.
- The deterministic reserves for variable insurance products with direct investment in equity funds and indexed products are also impacted by equity scenarios. The table below shows the Gross Wealth Factors (GWFs) for the 12/31/21 AIRG and field test runs. 1A, 1B, and 6 have similar GWFs to the AIRG, but the 5A field test run that utilized the original Conning equity calibration with the equity-Treasury linkage had significantly lower GWFs given the low starting interest rate environment.

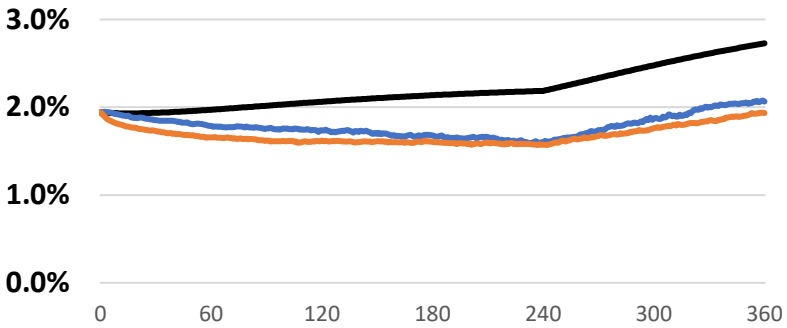
Large Cap (S&P 500) Equity Gross Wealth Factors

	12	60	120	240	360
AIRG	1.04	1.22	1.48	2.19	4.52
1A	1.03	1.16	1.38	2.01	4.29
1B	1.04	1.19	1.40	2.00	4.04
5A	1.03	1.09	1.17	1.36	2.47
6	1.06	1.27	1.56	2.29	4.77

1-Year UST Yield

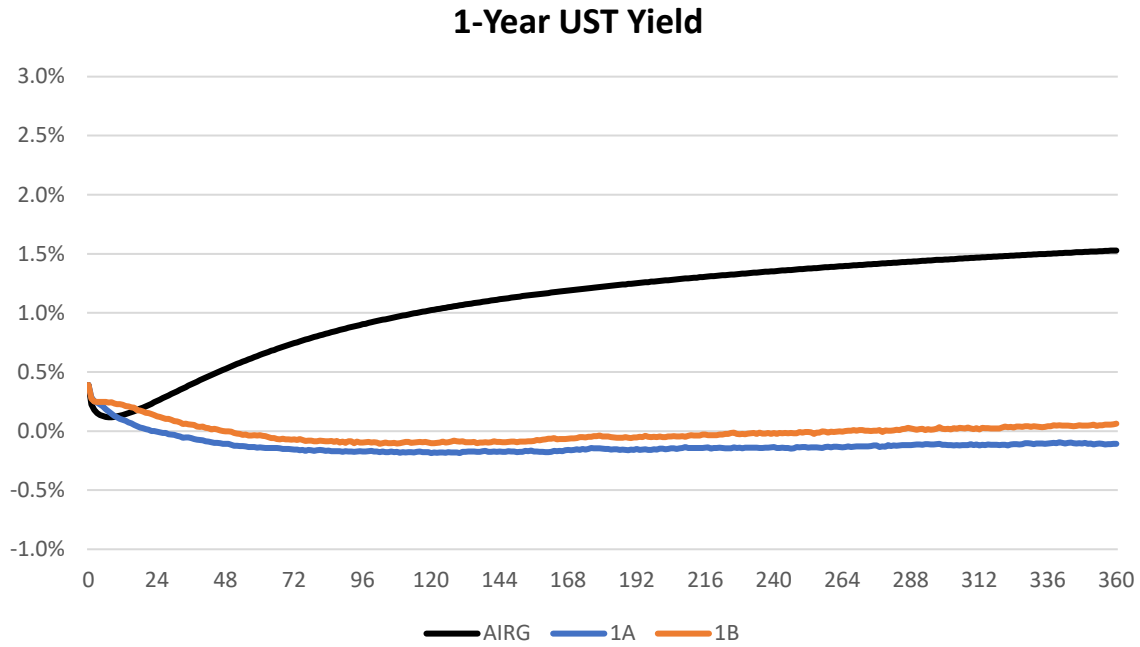


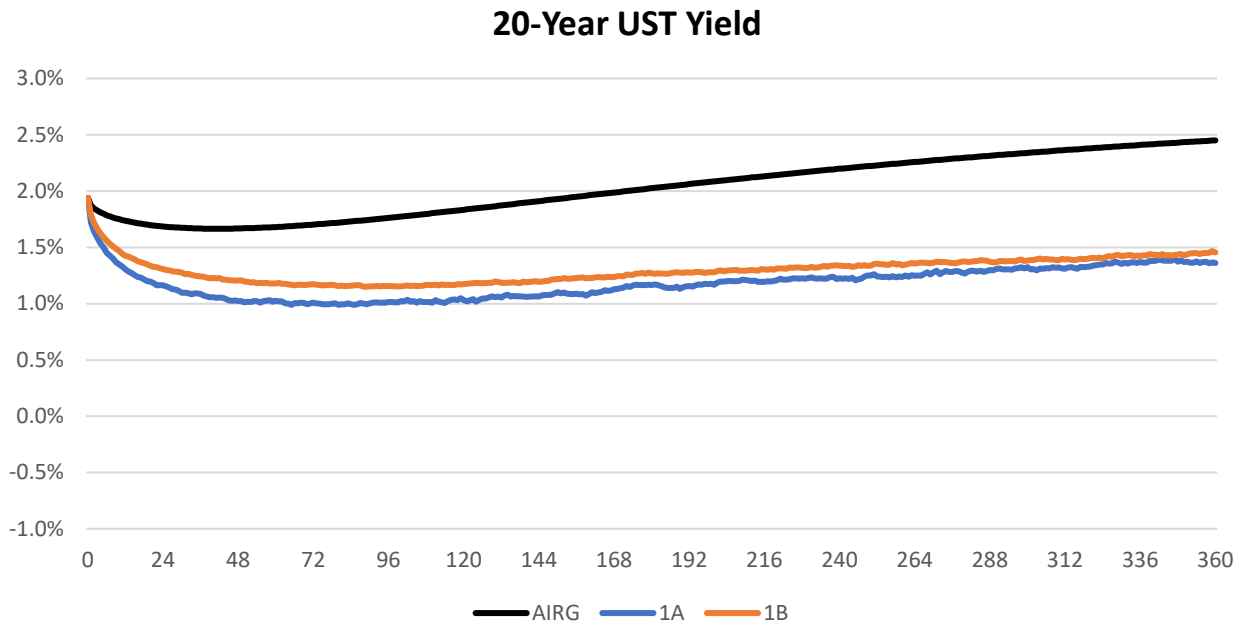
20-Year UST Yield



***Note: 5A and 6 have the same UST scenarios as 1A.**

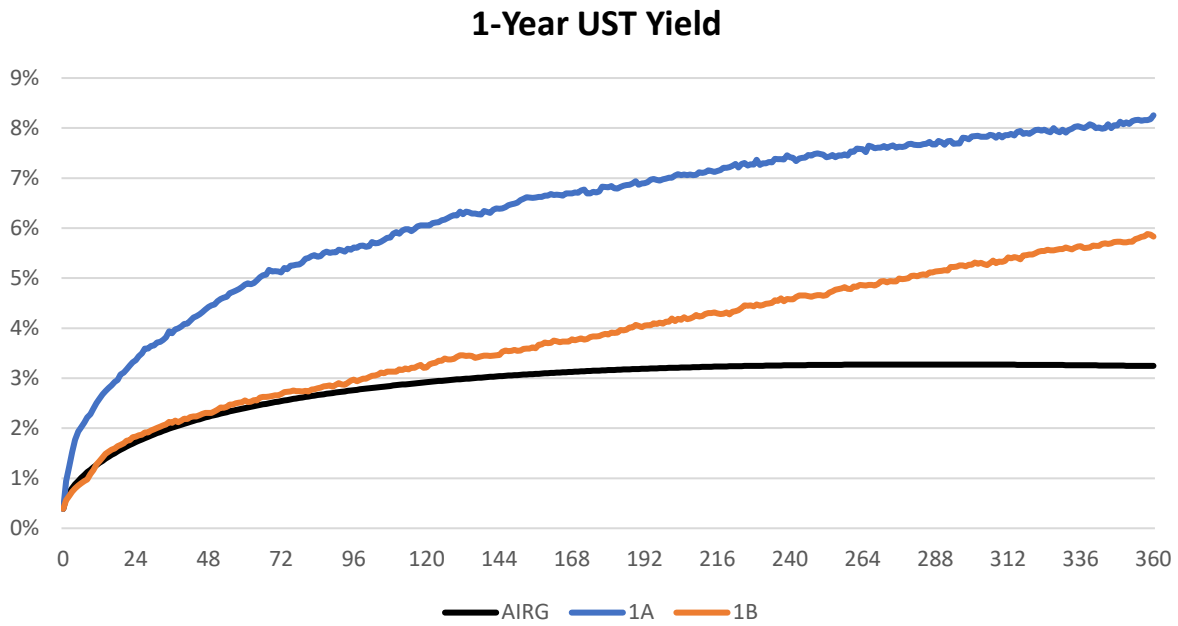
UST SERT Scenario 3 (Pop-down) at 12/31/21

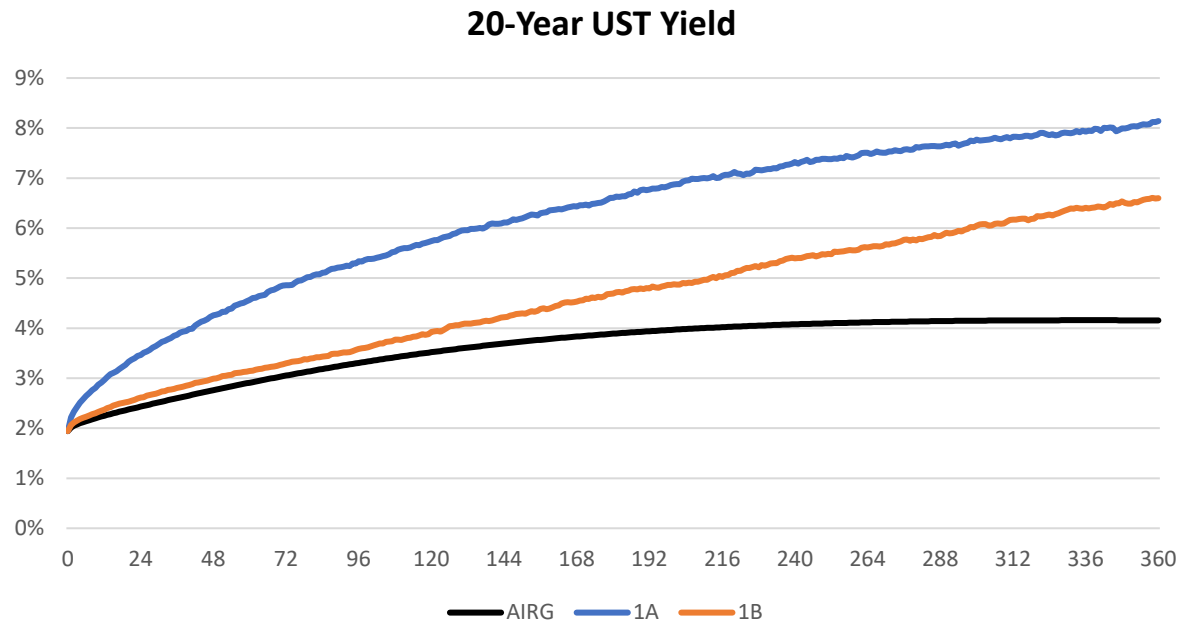




- The pop-down UST scenario for field test runs 1A and 1B are significantly lower than those produced by the AIRG
- **Pop-down description:** Interest rate shocks are selected to maintain the cumulative shock at the 10% level (1.282 standard errors).

UST SERT Scenario 1 (Pop-up) at 12/31/21





- The pop-up UST scenario for field test runs 1A and 1B are significantly higher than those produced by the AIRG. However, in the pop-up scenarios, field test 1A is also materially higher than field test 1B
- **Pop-up description:** Interest rate shocks are selected to maintain the cumulative shock at the 90% level (1.282 standard errors).

Appendix 2:
Treasury and Equity Scenario Overview

Field Test 1A: US Treasury Overview

- Field Test 1A (as of 12/31/21) included a recalibration of the Conning GEMS® US Treasury model that was designed to meet the regulator's acceptance criteria related to low for long, the prevalence of high interest rates, upper and lower bounds, initial yield curve fit, and yield curve shape. The frequency and severity of negative interest rates were controlled using a generalized fractional floor.
- The 1A UST scenario set as of 12/31/21 had a much higher prevalence of low UST rates, including negative interest rates, compared to the scenarios produced by the AIRG as of 12/31/21, which is floored at 1 BP.
- The 1A UST scenario set also included greater and more frequent high UST rates, with maximum UST rates greatly exceeding that of the AIRG. While a floor was employed in all of the field test UST scenario sets, no cap was employed on how high rates could get.

1A: 10,000 1-yr UST Scenario Percentiles by Projection Month as of 12/31/21

Percentile	12	60	120	240	360
Min	-0.49%	-0.97%	-0.94%	-0.91%	-0.93%
1%	-0.17%	-0.51%	-0.58%	-0.56%	-0.56%
10%	0.10%	-0.14%	-0.19%	-0.13%	-0.11%
25%	0.25%	0.14%	0.14%	0.19%	0.25%
50%	0.62%	0.84%	1.18%	1.61%	2.09%
75%	1.63%	2.83%	3.59%	4.39%	4.93%
95%	3.15%	6.14%	7.78%	9.35%	10.38%
99%	4.32%	8.86%	11.38%	13.53%	14.47%
Max	7.93%	14.36%	19.89%	25.18%	26.72%

AIRG: 10,000 1-yr UST Scenario Percentiles by Projection Month as of 12/31/21

Percentile	12	60	120	240	360
Min	0.01%	0.01%	0.01%	0.01%	0.01%
1%	0.01%	0.21%	0.33%	0.32%	0.32%
10%	0.27%	0.66%	0.87%	0.98%	0.99%
25%	0.47%	0.96%	1.22%	1.41%	1.45%
50%	0.69%	1.35%	1.68%	1.99%	2.10%
75%	0.92%	1.78%	2.27%	2.74%	2.90%
95%	1.29%	2.57%	3.40%	4.29%	4.66%
99%	1.59%	3.37%	4.75%	6.17%	6.31%
Max	2.31%	5.82%	10.94%	13.22%	12.76%

1A-AIRG: 10,000 1-yr UST Scenario Percentiles by Projection Month

Difference	12	60	120	240	360
Min	-0.5%	-1.0%	-0.9%	-0.9%	-0.9%
1%	-0.2%	-0.7%	-0.9%	-0.9%	-0.9%
10%	-0.2%	-0.8%	-1.1%	-1.1%	-1.1%
25%	-0.2%	-0.8%	-1.1%	-1.2%	-1.2%
50%	-0.1%	-0.5%	-0.5%	-0.4%	0.0%
75%	0.7%	1.0%	1.3%	1.7%	2.0%
95%	1.9%	3.6%	4.4%	5.1%	5.7%
99%	2.7%	5.5%	6.7%	7.4%	8.2%
Max	5.6%	8.5%	8.9%	12.0%	14.0%

Field Test 1A: Equity Overview

- The 1A equity scenario set used a calibration that targeted the median gross wealth factor (GWF) produced by the AIRG at the end of 30 years. This recentering of the equity return distribution with changes to the starting interest environment partially mitigates the impact of the GEMS® equity-Treasury linkage functionality.
- While the GWF's between the AIRG and field test 1A are consistent at the 50th percentile at the end of the 30th projection year, the 1A scenario set generally has somewhat lower GWFs in the lower percentiles and earlier projection years compared to the AIRG.
- In the later durations and higher percentiles, the 1A GWFs are greater than those produced by the AIRG.

1A: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.50	0.28	0.24	0.39	0.39
1.0%	0.71	0.59	0.59	0.83	1.17
2.5%	0.77	0.68	0.75	1.06	1.60
5.0%	0.82	0.78	0.87	1.34	2.11
10.0%	0.87	0.89	1.05	1.69	2.86
25.0%	0.97	1.09	1.40	2.54	4.88
50.0%	1.07	1.35	1.88	4.01	8.99
75.0%	1.16	1.64	2.57	6.49	16.98
90.0%	1.25	1.96	3.41	10.26	31.70
95.0%	1.31	2.20	4.04	13.67	47.46
97.5%	1.35	2.45	4.70	17.57	66.83
99.0%	1.41	2.77	5.65	23.45	101.58
Max	1.81	4.53	13.89	55.97	457.07

AIRG: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.41	0.32	0.26	0.35	0.38
1.0%	0.70	0.62	0.66	0.83	1.22
2.5%	0.76	0.72	0.77	1.10	1.69
5.0%	0.82	0.81	0.92	1.41	2.25
10.0%	0.89	0.93	1.12	1.83	3.09
25.0%	0.98	1.16	1.51	2.74	5.11
50.0%	1.09	1.45	2.09	4.27	8.84
75.0%	1.19	1.81	2.88	6.80	15.35
90.0%	1.30	2.22	3.81	10.15	24.98
95.0%	1.37	2.48	4.44	12.92	34.25
97.5%	1.44	2.72	5.17	15.65	45.88
99.0%	1.52	3.06	6.18	20.49	60.45
Max	1.92	4.77	11.86	66.94	235.95

1A/AIRG: GWF Ratios by Projection Month

SP500	12	60	120	240	360
Min	123%	90%	94%	113%	101%
1.0%	101%	96%	90%	100%	95%
2.5%	102%	95%	97%	96%	95%
5.0%	100%	96%	95%	95%	94%
10.0%	99%	96%	93%	92%	92%
25.0%	98%	94%	92%	93%	96%
50.0%	98%	93%	90%	94%	102%
75.0%	98%	90%	90%	95%	111%
90.0%	97%	88%	90%	101%	127%
95.0%	95%	89%	91%	106%	139%
97.5%	94%	90%	91%	112%	146%
99.0%	93%	90%	92%	114%	168%
Max	94%	95%	117%	84%	194%

Field Test 1B: US Treasury Overview

- Field Test 1B (as of 12/31/21) included a calibration of the Conning GEMS® US Treasury model that was designed to meet regulator acceptance criteria but placed additional emphasis on maintaining realistic term premiums throughout the projection. Towards that end, there was a significantly lower frequency of inversions (e.g.~5% of 1B scenarios had 10 year/2year UST inversions at the end of year 30 compared to ~12% seen in 1A). The average level of inversion was also significantly lower (e.g. in 1B 10 year/2 year UST inversions average ~30 BP at the end of year 30, compared to ~90 BP average inversion level for 1A).
- 1B also included lower and less frequent high interest rates than 1A, but still contained greater and more frequent high interest rates than the AIRG.
- The frequency and severity of negative interest rates were controlled using a shadow floor that preserves the arbitrage free nature of the scenarios. The 1B UST scenario set has a comparable amount of low/negative UST rates to 1A, but significantly more severe and frequent low (and negative) UST rates compared to the AIRG.

1B: 10,000 1-yr UST Scenario Percentiles by Projection Month

Percentile	12	60	120	240	360
Min	-0.59%	-1.08%	-1.24%	-1.18%	-1.19%
1%	-0.10%	-0.51%	-0.61%	-0.59%	-0.58%
10%	0.22%	-0.04%	-0.10%	-0.02%	0.06%
25%	0.42%	0.26%	0.27%	0.37%	0.49%
50%	0.65%	0.65%	0.71%	0.88%	1.28%
75%	0.88%	1.24%	1.67%	2.60%	3.52%
95%	1.76%	3.38%	4.38%	5.99%	7.49%
99%	2.57%	4.89%	6.44%	8.90%	10.64%
Max	4.25%	10.28%	11.63%	17.99%	22.87%

AIRG: 10,000 1-yr UST Scenario Percentiles by Projection Month

Percentile	12	60	120	240	360
Min	0.01%	0.01%	0.01%	0.01%	0.01%
1%	0.01%	0.21%	0.33%	0.32%	0.32%
10%	0.27%	0.66%	0.87%	0.98%	0.99%
25%	0.47%	0.96%	1.22%	1.41%	1.45%
50%	0.69%	1.35%	1.68%	1.99%	2.10%
75%	0.92%	1.78%	2.27%	2.74%	2.90%
95%	1.29%	2.57%	3.40%	4.29%	4.66%
99%	1.59%	3.37%	4.75%	6.17%	6.31%
Max	2.31%	5.82%	10.94%	13.22%	12.76%

Difference	12	60	120	240	360
Min	-0.6%	-1.1%	-1.2%	-1.2%	-1.2%
1%	-0.1%	-0.7%	-0.9%	-0.9%	-0.9%
10%	0.0%	-0.7%	-1.0%	-1.0%	-0.9%
25%	-0.1%	-0.7%	-1.0%	-1.0%	-1.0%
50%	0.0%	-0.7%	-1.0%	-1.1%	-0.8%
75%	0.0%	-0.5%	-0.6%	-0.1%	0.6%
95%	0.5%	0.8%	1.0%	1.7%	2.8%
99%	1.0%	1.5%	1.7%	2.7%	4.3%
Max	1.9%	4.5%	0.7%	4.8%	10.1%

Field Test 1B: Equity Overview

- The 1B equity scenario set used the same calibration as 1A. However, due to the equity-Treasury linkage, the resulting GWFs are different. The largest differences between the 1A and 1B equity GWFs are seen at the upper percentiles at the end of the 30th projection year, with the 1B being substantially lower and more in line with the AIRG.
- The median GWF at the end of the 30th projection year for 1B (7.99) is materially lower than both 1A (8.99) and the AIRG (8.84).
- Finally, the 1st percentile GWF at the end of the 30th projection year for 1b (1.19) was consistent with those of 1A (1.17) and the AIRG (1.22).

1B: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.51	0.30	0.26	0.34	0.27
1.0%	0.71	0.61	0.61	0.82	1.19
2.5%	0.78	0.70	0.76	1.05	1.59
5.0%	0.83	0.80	0.90	1.33	2.07
10.0%	0.88	0.92	1.08	1.68	2.72
25.0%	0.98	1.12	1.42	2.47	4.57
50.0%	1.08	1.38	1.90	3.78	7.99
75.0%	1.17	1.68	2.56	5.85	13.71
90.0%	1.26	2.00	3.32	8.61	23.14
95.0%	1.32	2.24	3.94	10.91	32.00
97.5%	1.36	2.50	4.53	13.70	43.02
99.0%	1.42	2.80	5.44	17.25	61.86
Max	1.83	4.67	14.21	76.72	258.35

AIRG: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.41	0.32	0.26	0.35	0.38
1.0%	0.70	0.62	0.66	0.83	1.22
2.5%	0.76	0.72	0.77	1.10	1.69
5.0%	0.82	0.81	0.92	1.41	2.25
10.0%	0.89	0.93	1.12	1.83	3.09
25.0%	0.98	1.16	1.51	2.74	5.11
50.0%	1.09	1.45	2.09	4.27	8.84
75.0%	1.19	1.81	2.88	6.80	15.35
90.0%	1.30	2.22	3.81	10.15	24.98
95.0%	1.37	2.48	4.44	12.92	34.25
97.5%	1.44	2.72	5.17	15.65	45.88
99.0%	1.52	3.06	6.18	20.49	60.45
Max	1.92	4.77	11.86	66.94	235.95

1B/AIRG: GWF Ratios by Projection Month

SP500	12	60	120	240	360
Min	124%	94%	102%	98%	71%
1.0%	102%	98%	93%	99%	98%
2.5%	103%	98%	99%	95%	94%
5.0%	101%	99%	97%	94%	92%
10.0%	100%	99%	96%	92%	88%
25.0%	99%	97%	94%	90%	89%
50.0%	99%	95%	91%	88%	90%
75.0%	99%	93%	89%	86%	89%
90.0%	97%	90%	87%	85%	93%
95.0%	96%	90%	89%	84%	93%
97.5%	95%	92%	88%	88%	94%
99.0%	94%	91%	88%	84%	102%
Max	95%	98%	120%	115%	109%

Field Test 2A: US Treasury Overview

- Field Test 2A (as of 12/31/19 + 200 BP) used the same calibration as 1A (Conning Calibration with a Generalized Fractional Floor) but with a 12/31/19 starting yield curve modified using a 200 BP increase across all maturities.
- The higher starting interest environment leads to greater and more frequent high interest rates and less severe and less frequent low interest rates in 2A compared to 1A.
- Compared to the AIRG with a 12/31/19 + 200 BP starting interest environment, the 2A scenario set has a greater frequency and severity of high UST rates and more prevalent and severe low (and negative) UST rates.

2A (12/31/19 + 200 BP): 10,000 1-yr UST Scenario Percentiles by Projection Month

Percentile	12	60	120	240	360
Min	-0.13%	-0.78%	-0.82%	-0.89%	-0.92%
1%	0.29%	-0.27%	-0.42%	-0.49%	-0.53%
10%	1.34%	0.19%	0.02%	-0.04%	-0.06%
25%	2.26%	0.87%	0.39%	0.31%	0.32%
50%	3.34%	2.89%	2.69%	2.43%	2.54%
75%	4.49%	5.15%	5.38%	5.47%	5.53%
95%	6.19%	8.80%	10.06%	10.86%	11.30%
99%	7.44%	11.88%	13.61%	15.32%	15.70%
Max	11.48%	17.62%	22.91%	27.07%	28.97%

AIRG (12/31/19 + 200 BP): 10,000 1-yr UST Scenario Percentiles by Projection Month

Percentile	12	60	120	240	360
Min	0.31%	0.01%	0.01%	0.01%	0.01%
1%	1.25%	0.47%	0.34%	0.29%	0.31%
10%	1.82%	1.22%	1.06%	1.04%	1.00%
25%	2.16%	1.72%	1.58%	1.53%	1.50%
50%	2.53%	2.35%	2.24%	2.21%	2.18%
75%	2.92%	3.06%	3.08%	3.10%	3.05%
95%	3.55%	4.39%	4.77%	4.96%	4.94%
99%	4.06%	5.66%	6.73%	7.29%	6.73%
Max	5.24%	9.85%	16.66%	15.13%	13.59%

2A-AIRG: 10,000 1-yr UST Scenario Percentiles by Projection Month

Difference	12	60	120	240	360
Min	-0.4%	-0.8%	-0.8%	-0.9%	-0.9%
1%	-1.0%	-0.7%	-0.8%	-0.8%	-0.8%
10%	-0.5%	-1.0%	-1.0%	-1.1%	-1.1%
25%	0.1%	-0.8%	-1.2%	-1.2%	-1.2%
50%	0.8%	0.5%	0.4%	0.2%	0.4%
75%	1.6%	2.1%	2.3%	2.4%	2.5%
95%	2.6%	4.4%	5.3%	5.9%	6.4%
99%	3.4%	6.2%	6.9%	8.0%	9.0%
Max	6.2%	7.8%	6.3%	11.9%	15.4%

Field Test 2A: Equity Overview

- The targets of the 2A equity scenarios is designed to align the GWF at the end of the 30th projection year (8.97) with those produced by the AIRG (8.84) no matter the starting interest rate environment. However, there is still an impact to the 2A equity scenarios due to the increased starting interest rate environment and the equity-Treasury linkage compared to the 1A equity scenarios.
- The largest differences between the 2A and 1A equity GWFs are seen at the upper percentiles at the end of the 30th projection year, for example the 99th percentile GWF for 1b is 127.28 at the end of the 30th year compared to 101.58 for the 1A scenario set.
- The same considerations apply when comparing 2A to the AIRG with a 12/31/19 + 200 BP starting interest rate environment, with the largest differences between the GWFs of 2A and the AIRG occurring in the higher percentiles and later projection years.

2A: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.51	0.30	0.26	0.40	0.36
1.0%	0.73	0.65	0.64	0.83	1.07
2.5%	0.79	0.75	0.80	1.08	1.46
5.0%	0.84	0.85	0.95	1.34	1.93
10.0%	0.90	0.97	1.15	1.73	2.63
25.0%	1.00	1.20	1.54	2.64	4.71
50.0%	1.10	1.48	2.11	4.38	8.97
75.0%	1.20	1.82	2.96	7.42	18.20
90.0%	1.29	2.19	4.01	12.10	35.66
95.0%	1.35	2.46	4.74	16.60	54.53
97.5%	1.39	2.73	5.63	22.33	83.32
99.0%	1.45	3.10	7.00	30.39	127.28
Max	1.87	5.11	15.80	86.26	817.22

AIRG: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.41	0.32	0.26	0.35	0.38
1.0%	0.70	0.62	0.66	0.83	1.22
2.5%	0.76	0.72	0.77	1.10	1.69
5.0%	0.82	0.81	0.92	1.41	2.25
10.0%	0.89	0.93	1.12	1.83	3.09
25.0%	0.98	1.16	1.51	2.74	5.11
50.0%	1.09	1.45	2.09	4.27	8.84
75.0%	1.19	1.81	2.88	6.80	15.35
90.0%	1.30	2.22	3.81	10.15	24.98
95.0%	1.37	2.48	4.44	12.92	34.25
97.5%	1.44	2.72	5.17	15.65	45.88
99.0%	1.52	3.06	6.18	20.49	60.45
Max	1.92	4.77	11.86	66.94	235.95

2A/AIRG: GWF Ratios by Projection Month

SP500	12	60	120	240	360
Min	124%	95%	99%	116%	94%
1.0%	104%	104%	97%	100%	88%
2.5%	105%	103%	103%	98%	87%
5.0%	103%	105%	103%	95%	86%
10.0%	101%	104%	102%	94%	85%
25.0%	101%	103%	102%	96%	92%
50.0%	101%	102%	101%	103%	101%
75.0%	101%	100%	103%	109%	119%
90.0%	100%	99%	105%	119%	143%
95.0%	98%	99%	107%	129%	159%
97.5%	97%	100%	109%	143%	182%
99.0%	96%	101%	113%	148%	211%
Max	98%	107%	133%	129%	346%

Field Test 5A: Treasury and Equity Overview

- The 5A scenario set uses the exact same UST scenarios as 1A.
- For the 5A equity scenario set, the Conning's original equity model calibration is used that includes the full impact of the equity-Treasury linkage. With 5A's lower overall UST rates, the equity GWFs at the lower percentiles are much more severe than the AIRG and other field test scenario sets. For example, the 1st percentile of equity GWFs for 5A is .39, compared to 1.22 for the AIRG and 1.19 for 1A.
- The median GWF at the end of the 30th projection year for 5A (5.88) is significantly lower than with both 1A (8.99) and the AIRG (8.84).

5A: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.47	0.13	0.06	0.04	0.05
1.0%	0.71	0.45	0.36	0.38	0.39
2.5%	0.76	0.57	0.48	0.54	0.65
5.0%	0.82	0.67	0.63	0.73	0.95
10.0%	0.87	0.80	0.82	1.04	1.48
25.0%	0.96	1.02	1.20	1.79	2.93
50.0%	1.05	1.28	1.69	3.09	5.88
75.0%	1.14	1.56	2.31	5.11	11.43
90.0%	1.21	1.85	3.02	8.11	21.44
95.0%	1.26	2.04	3.59	10.76	32.94
97.5%	1.30	2.23	4.11	13.83	47.77
99.0%	1.35	2.50	4.83	18.95	71.23
Max	1.68	3.79	10.89	64.69	494.22

AIRG: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.41	0.32	0.26	0.35	0.38
1.0%	0.70	0.62	0.66	0.83	1.22
2.5%	0.76	0.72	0.77	1.10	1.69
5.0%	0.82	0.81	0.92	1.41	2.25
10.0%	0.89	0.93	1.12	1.83	3.09
25.0%	0.98	1.16	1.51	2.74	5.11
50.0%	1.09	1.45	2.09	4.27	8.84
75.0%	1.19	1.81	2.88	6.80	15.35
90.0%	1.30	2.22	3.81	10.15	24.98
95.0%	1.37	2.48	4.44	12.92	34.25
97.5%	1.44	2.72	5.17	15.65	45.88
99.0%	1.52	3.06	6.18	20.49	60.45
Max	1.92	4.77	11.86	66.94	235.95

5A/AIRG: GWF Ratios by Projection Month

SP500	12	60	120	240	360
Min	114%	40%	24%	11%	13%
1.0%	101%	73%	54%	46%	32%
2.5%	100%	79%	62%	49%	39%
5.0%	100%	83%	68%	51%	42%
10.0%	99%	86%	73%	57%	48%
25.0%	98%	88%	79%	65%	57%
50.0%	97%	88%	81%	72%	66%
75.0%	96%	86%	80%	75%	74%
90.0%	94%	83%	79%	80%	86%
95.0%	92%	82%	81%	83%	96%
97.5%	91%	82%	80%	88%	104%
99.0%	89%	82%	78%	92%	118%
Max	87%	80%	92%	97%	209%

Field Test 6: Treasury and Equity Overview

- The field test 6 scenario set uses the exact same UST scenarios as 1A.
- The equity calibration for scenario set 6 assumes a constant mean equity return independent of rates and increases alignment with AIRG equity model GWFs.
- The median GWF at the end of the 30th projection year for 6 is 9.49, which is close but somewhat higher than the the corresponding GWFs for both 1A (8.99) and the AIRG (8.84).
- While there are differences (somewhat lower GWFs in low percentiles, lower GWFs at higher percentiles), the equity scenarios from 6 overall are more consistent with those produced by the AIRG than other field test scenario sets.

6: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.43	0.14	0.13	0.31	0.23
1.0%	0.71	0.57	0.59	0.79	1.20
2.5%	0.77	0.68	0.76	1.08	1.73
5.0%	0.83	0.80	0.92	1.41	2.32
10.0%	0.89	0.94	1.14	1.85	3.20
25.0%	0.99	1.19	1.58	2.90	5.41
50.0%	1.09	1.50	2.17	4.55	9.49
75.0%	1.19	1.82	2.90	6.83	15.89
90.0%	1.28	2.15	3.66	9.85	24.35
95.0%	1.33	2.34	4.22	12.01	31.70
97.5%	1.38	2.52	4.76	14.36	39.68
99.0%	1.43	2.75	5.37	17.19	52.06
Max	1.79	3.97	9.38	33.26	135.23

AIRG: 10,000 SP500 GWF %-tiles by Projection Month

	12	60	120	240	360
Min	0.41	0.32	0.26	0.35	0.38
1.0%	0.70	0.62	0.66	0.83	1.22
2.5%	0.76	0.72	0.77	1.10	1.69
5.0%	0.82	0.81	0.92	1.41	2.25
10.0%	0.89	0.93	1.12	1.83	3.09
25.0%	0.98	1.16	1.51	2.74	5.11
50.0%	1.09	1.45	2.09	4.27	8.84
75.0%	1.19	1.81	2.88	6.80	15.35
90.0%	1.30	2.22	3.81	10.15	24.98
95.0%	1.37	2.48	4.44	12.92	34.25
97.5%	1.44	2.72	5.17	15.65	45.88
99.0%	1.52	3.06	6.18	20.49	60.45
Max	1.92	4.77	11.86	66.94	235.95

6/AIRG: GWF Ratios by Projection Month

SP500	12	60	120	240	360
Min	106%	44%	50%	88%	60%
1.0%	101%	92%	89%	95%	98%
2.5%	102%	95%	99%	98%	102%
5.0%	101%	100%	100%	100%	103%
10.0%	100%	101%	102%	101%	103%
25.0%	101%	103%	104%	106%	106%
50.0%	100%	103%	104%	107%	107%
75.0%	100%	100%	101%	101%	104%
90.0%	99%	97%	96%	97%	97%
95.0%	97%	95%	95%	93%	93%
97.5%	96%	92%	92%	92%	86%
99.0%	94%	90%	87%	84%	86%
Max	93%	83%	79%	50%	57%