

William H. Wilton  
[Wilton579@sbcglobal.net](mailto:Wilton579@sbcglobal.net)

April 20, 2022

Reggie Mazyck  
National Association of Insurance Commissioners  
1100 Walnut Street – Suite 1500  
Kansas City, MO 64106-2197

Re: ESG Field Test Specifications

I appreciate the opportunity to provide comments and some additional thoughts on the Economic Scenario Generator (ESG) Reserves and Capital Field Test Specifications. The implementation of a new generator potentially has a significant impact on consumers, companies and the industry and we should make sure we do all the analysis necessary to fully understand its impact on companies and the products they will be able to offer to consumers.

### **Section 1: A. 3 Metrics**

<b>3. Metrics</b>	<ul style="list-style-type: none"><li>• Which particular interest rate and equity scenarios cause the greatest stress?</li><li>• How do results compare for CTE70 vs. CTE98? Calculate different CTE levels (e.g., CTE70, CTE98, CTE90) to compare to existing requirements.</li><li>• <b>How do the metrics perform with different scenario set sizes?</b></li></ul>
-------------------	---

What is envisioned for “different scenario set sizes”? Is this to test the number of scenarios for convergence or is it to test the adequacy / inadequacy of the scenario picking tool?

### **Section II: D. ESG Models and Scenarios, 3<sup>rd</sup> bullet**

- 10,000 scenarios will be provided along with 1,000, 500, 200, and 40 scenario subsets

### **Section II: E. ESG Models and Scenarios, 7<sup>th</sup> bullet**

- Field test participants may choose the number of scenarios included in their calculation of reserves or capital for each line of business, with the exception of C3 Phase I where runs will be subject to a minimum of 200 scenarios.

Are the scenario subsets based on the scenario-picker tool? Is the scenario picker tool still based on the 20-year Treasury? If so, are we getting what we want? If the objective of the field test is to validate the generator, are we diluting the analysis with artificial variability and precision due to the scenario picker? It would seem that for a field test you would want consistency among companies using different scenario set sizes. The companies that are running 10,000 scenarios could easily report results if the 200, 500 or 1000 scenario subsets were sequentially the first x

scenarios of the 10,000. I see minimal value with a 40-scenario subset in meeting the field test objectives.

For VM-21, are the scenario subsets appropriate? Will the scenario subsets for VM-21 be based on the 20-year Treasury and not equity markets?

In my opinion, the use of any scenario reduction techniques should be left to the company to validate and not be part of this ESG field test. Once the parameterization of the scenario generator is established, we can test if the scenario picker is still acceptable for use as permitted in Section 7.G.2.c. of VM-20.

<b>3. Metrics</b>	<ul style="list-style-type: none"><li>• Which particular interest rate and equity scenarios cause the greatest stress?</li><li>• How do results compare for CTE70 vs. CTE98? Calculate different CTE levels (e.g., CTE70, CTE98, CTE90) to compare to existing requirements.</li><li>• How do the metrics perform with different scenario set sizes?</li></ul>
-------------------	--

Do we want to capture more than just the CTE metrics? Do we want to have a “picture” of the distribution in the tail? Would percentile metrics for some of the key components add additional value? For example, percentiles 65 – 95 at 5% increments with 1% increments from 95% to 100%? One of the underlying questions may be is CTE still the best metric to quantify reserves and capital? Are we concerned about the average loss in the tail or should we be focused on the level of reserves or capital required to be x% confident that they are sufficient? Should reserves be determined based on negative interest rate scenarios? Is an 80<sup>th</sup> percentile value a better representative of an appropriate reserve for moderately adverse conditions? Collecting more data than just the CTE may provide insight into the numerous questions about the tail that will likely arise.

For example, sample metrics to collect may include:

	Component (e.g. Scenario Reserve)			
	Scenario Set			
Metric	1-200	1-500	1-1000	1-10000
Percentile				
65				
70				
75				
80				
85				
90				
95				
96				
97				
98				
99				
100				
CTE 70				
CTE 90				
CTE 98				

## Section II: E. Metrics/Output 1<sup>st</sup> bullet, 4<sup>th</sup> sub-bullet:

- C3 Phase I
  - Reserves that were cash flow tested for asset adequacy
  - The C3 Phase I results should be summarized by applying the weights in the table below to the respective percentiles.

----- Percentile -----	92	92.5	93	93.5	94	94.5	95	95.5	96	96.5	97	97.5	98
Weighting	.02	.04	.06	.08	.10	.12	.16	.12	.10	.08	.06	.04	.02

For C3 Phase I, it would seem that the current scenario subset (generally 50) and the full 200 scenario set should be run. This would facilitate the attribution analysis that may be ultimately desired.

1. 50 Scenario – weighted 95<sup>th</sup> percentile (C3P1 Generator)
2. 200 Scenarios – weighted 95<sup>th</sup> percentile, 95<sup>th</sup> %-ile, & CTE90 (C3P1 Generator)
3. 200 Scenarios – 95<sup>th</sup> %-ile, & CTE90 with VM prescribed scenario generator
4. 200 Scenarios – 95<sup>th</sup> %-ile, & CTE90 with Treasury model - “Non-shadow”
5. 200 Scenarios – 95<sup>th</sup> %-ile, & CTE90 with Treasury model - “Shadow”
6. >200 Scenarios – 95<sup>th</sup> %-ile, & CTE90 with Treasury model - “Non-shadow”
7. >200 Scenarios – 95<sup>th</sup> %-ile, & CTE90 with Treasury model - “Shadow”

The 50 scenario set is currently used by many companies for their C3 Phase I calculation. The 200 scenario set will highlight the use of the full distribution, as opposed to a scenario subset designed to capture the tail, and also highlight the potential impact of a movement to a CTE metric. The VM prescribed economic scenario generator will illustrate the impact the current AIRG would have on the current methodology. The last 4 scenario sets would highlight the impact a new scenario generator would have on the results that would otherwise be produced, including the impact of running more than 200 scenarios.

I thank you for the opportunity to provide these comments and thoughts on this exposure.

Sincerely,

A handwritten signature in black ink, appearing to read "William H. Wilton".

William H. Wilton, FSA, MAAA