LATF Fall National Meeting
Supplemental Packet

LIFE ACTUARIAL (A) TASK FORCE
November 29-30, 2023

NAIC FALL NATIONAL MEETING

November 29-30, 2023

TABLE OF CONTENTS

Agenda Item 3

Receive an Update from the
VM-22 Policyholder Behavior Drafting Group

Standard Projection Amount

Base Surrender Rates for Fixed Indexed Annuities

Data Source: 2019/2020 LIMRA Fixed Indexed Annuity Study

Based surrender rates are developed the following types of FIA policies:

- FIA with no GLB
- FIA with exercised GLB
- FIA with not yet exercised GLB

Within each policy type, the base surrender rates are categorized by the following attributes:

- Attained Age Group (0-59, 60-64, 65-69, 70-74, 75-79, 80+)
- Qualified versus Non-Qualified
- Years before the end, at the end, and after the end of the surrender charge period
- In-the-moneyness (0-99%, 100-124%, 125%+)

As there are many possible combinations, we combine or graduate the base surrender rate for some cells with neighboring cells if there are no material differences among them or when the volumes are not credible.

All base surrender rates are rounded to the nearest 0.50%.

FIA, NO GLWB	Under 60		60-69		70-79		80 and over	
	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV
5 or more yrs after expiry	6.5%	362,439,974	7.0%	872,825,158	6.0%	1,559,973,113	5.0%	2,340,007,393
4 yrs after expiry	8.0%	99,004,108	8.5%	246,956,865	6.5%	489,325,310	5.0%	860,771,030
3 yrs after expiry	8.5%	169,071,566	9.5%	411,626,455	7.0%	847,903,742	5.5%	1,457,596,664
2 yrs after expiry	11.0%	265,595,461	12.0%	631,722,030	9.0%	1,313,595,688	7.0%	2,104,219,836
1 yrs after expiry	15.0%	407,046,549	17.5%	978,416,104	13.5%	2,000,926,373	9.0%	2,641,343,425
Yr SC Expires	33.5%	759,132,968	41.5%	2,318,136,057	37.0%	4,162,997,269	23.5%	3,595,604,698
1 yrs to expiry	4.5%	995,033,715	3.5%	3,144,391,095	4.0%	5,080,691,511	4.0%	3,582,846,441
2 yrs to expiry	4.0%	1,265,675,665	3.5%	4,000,947,636	3.0%	5,882,400,721	3.0%	3,693,425,414
3 yrs to expiry	3.0%	1,879,265,259	3.0%	5,677,923,198	3.0%	7,485,708,831	2.5%	4,281,189,408
4 yrs to expiry	2.5%	2,729,096,608	2.5%	7,754,645,406	2.5%	8,934,521,480	2.0%	4,656,360,188
5 yrs or more to expiry	2.0%	23,241,306,391	2.0%	60,089,390,966	2.0%	46,626,860,517	2.0%	14,409,332,402

FIA, with GLWB, under Election, IMF (in the money factor)	Under 60		60-69		70-79		80 and over	
	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV
yrs after SC expiry								
125% and over	1.5%	630,954,157	1.5%	402,890,394	1.0%	1,397,503,656	1.0%	3,392,933
100-124%	1.5%	152,641,122	1.5%	50,580,570	1.5%	284,074,292	1.5%	378,008
Under 100%	3.5%	275,861,301	3.5%	8,296,745	4.5%	208,761,127	4.5%	155,465
Yr SC Expires								
125% and over	1.5%	529,957,319	1.5%	434,899,984	1.0%	1,283,913,770	1.0%	4,702,281
100-124%	2.5%	89,181,339	2.5%	36,759,595	2.5%	199,885,926	2.5%	684,087
Under 100%	10.5%	105,236,356	10.5%	6,123,022	14.0%	98,356,102	14.0%	60,789
prior to expiry								
125% and over	1.0%	4,676,773,692	1.0%	6,149,528,447	1.0%	14,307,818,333	1.0%	61,005,062
100-124%	1.0%	1,730,600,030	1.0%	5,289,677,487	1.0%	8,335,076,617	1.0%	101,603,008
Under 100%	1.0%	501,720,201	1.0%	972,486,678	1.5%	1,417,061,147	1.5%	18,813,144

FIA, with GLWB, No Election, IMF (in the money factor)	Under 60		60-69		70-79		80 and over	
	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV	Surr Rate	XPO_AV
yrs after SC expiry								
125% and over	3.5%	276,875,423	3.5%	1,061,462,156	4.0%	1,367,627,485	4.0%	1,034,837,154
100-124%	8.5%	54,368,043	8.5%	195,923,250	6.0%	446,934,521	6.0%	467,164,213
Under 100%	68.5%	8,159,453	68.5%	34,971,531	50.5%	63,961,566	50.5%	84,575,510
Yr SC Expires								
125% and over	6.0%	403,489,838	6.0%	1,368,282,819	7.5%	1,549,590,971	7.5%	870,131,375
100-124%	16.5%	73,362,908	16.5%	201,431,286	13.5%	420,179,736	13.5%	385,710,160
Under 100%	92.0%	29,567,553	92.0%	108,047,465	86.5%	178,685,007	86.5%	113,832,010
prior to expiry								
125% and over	1.0%	7,472,524,670	1.0%	32,045,809,434	1.5%	25,853,378,194	1.5%	7,232,790,419
100-124%	1.5%	13,044,250,434	1.5%	44,064,092,199	1.5%	26,831,653,607	1.5%	7,470,127,280
Under 100%	1.5%	4,341,759,020	1.5%	11,099,461,190	4.0%	4,781,589,407	4.0%	837,554,667

Withdrawal Rates for Fixed Indexed Annuities

Partial Withdrawal rates are developed the following types of FIA policies:

- FIA without GLB
- FIA with not yet exercised GLB

In addition, we are proposing to replace the Withdrawal Delay Cohort Method with an SPA guardrail around benefit election.

Data Source: 2019/2020 LIMRA Fixed Indexed Annuity Study

Experience data was reviewed based on available granularity, which included:

- Attained Age group (0-59, 60-64, 65-69, 70-74, 75-79, 80+)
- Qualified vs. Non-Qualified Tax Status
- Moneyness levels (for contracts with GLB)
- GLB utilization efficiency (withdrawal amount as %age of GLB limit)
- Calendar year, to make sure inclusion of 2020 data (pandemic era) did not unduly influence assumption

Methodology: Data was grouped for assumption setting when experience was clearly similar. Some rounding was applied based on magnitude of raw experience. (i.e., nearest 5 to 50bps)

Partial Withdrawal Rates for non-GLB FIA contracts, expressed as % of Account Value:

	without GLB; % of AV	
Qualified	ATT_AGE	Qualified
	Under 60	**1.70%**
	60-64	**2.05%**
	65-69	**2.25%**
	70-74	**3.40%**
	75-79	**4.55%**
	80 and over	**6.00%**
Non - Qualified	ATT_AGE	Non - Qualified
	All ages	1.65%

➔ While qualified rates are experience based, they broadly align to RMD rates although not set equal to RMD rates as RMD requirements are set at the taxpayer level, not individual contract level.

Partial Withdrawal Rates for FIA contracts with GLB, but GLB not yet elected (% of AV):

	with GLB (GLB_Wd_Ind = 0)	
Qualified	ATT_AGE	Qualified
	Under 60	**0.95%**
	60-64	**1.15%**
	65-69	**1.40%**
	70-74	**2.70%**
	75-79	**4.30%**
	80 and over	**5.80%**
Non - Qualified	ATT_AGE	Non - Qualified
	Under 70	**1.15%**
	70 and over	**1.65%**

➔ Experience is similar to non-GLB, but slightly lower as presumably partial withdrawal utilization will pick-up upon GLB election.

Partial Withdrawal Rates for FIA contracts with GLB elected:

Similar to VM-21 SPA, for these contracts we assume clients will efficiently utilize the GLB and take 100% of allowable rate for lifetime GLB's (VM-21 used 90%) and 70% for non-lifetime GLB's (VM-21 used 70%) once withdrawals have commenced. We are open to feedback but do believe a high level of efficiency should be assumed.

For contracts not yet withdrawing, we utilized the LIMRA data to benchmark cumulative GLB benefit election rates, to serve as guardrails to the company assumption in place of a Withdrawal Delay Cohort Methodology type approach. The cumulative benefit utilization rates (% of total GLB contracts currently withdrawing) are proposed to be a floor applied to the Company utilization Assumption, and were drafted as follows:

Benefit utilization rate	<= 125% ITM		> 125% ITM	
ATT_AGE	*Qualified*	*Non - Qualified*	*Qualified*	*Non - Qualified*
Under 60	**0.75%**	**1.00%**	**0.75%**	**1.25%**
60-64	**5.00%**	**5.25%**	**8.25%**	**9.25%**
65-69	**14.50%**	**13.25%**	**21.50%**	**20.50%**
70-74	**25.00%**	**20.00%**	**36.75%**	**28.75%**
75 and over	**29.50%**	**22.50%**	**43.50%**	**34.50%**

Dynamic Lapse Rates for Fixed Indexed Annuities

Dynamic Lapse rates are developed the following types of FIA policies:

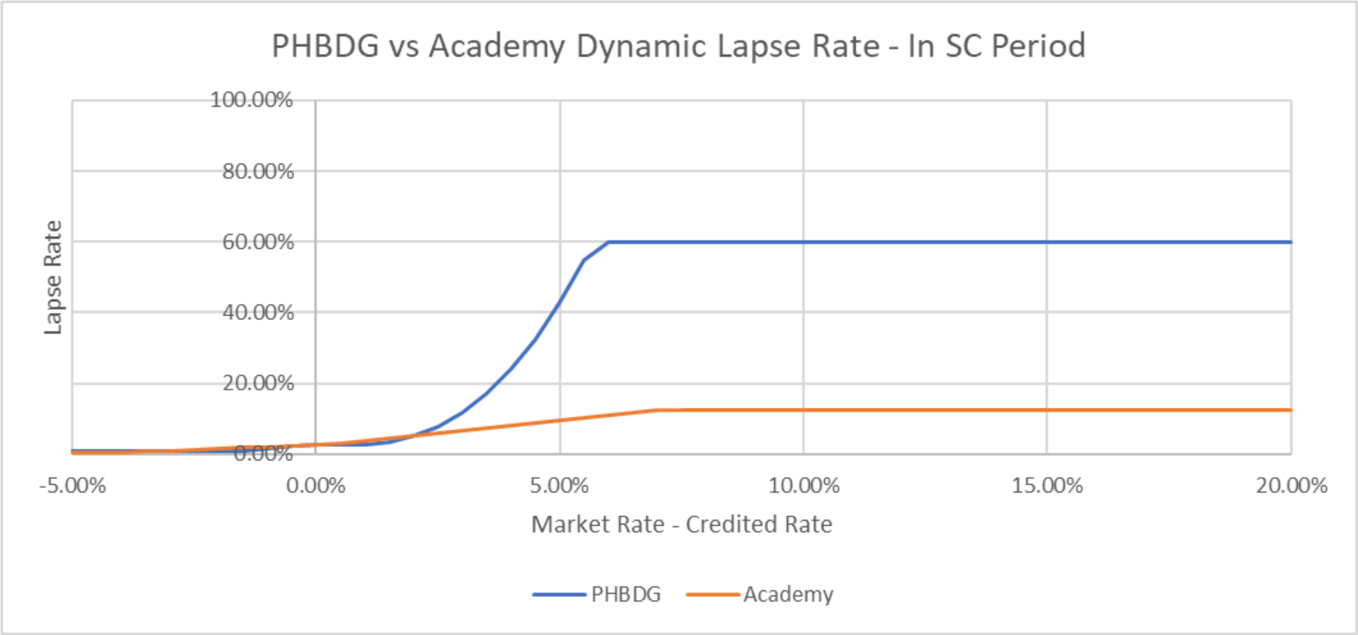
- FIA without GLB
- FIA with GLB

The PHBDG originally proposed the framework in the left column, which featured a multiple of 1.25 and exponent of 2.5 for the Credited Rate – Market Rate adjustment. The Academy responded to the proposal and recommended using a linear formula (exponent 1), with multiples varying by In SC Period/Shock/Post-Shock. The Academy also recommended changes to market rate, minimum/maximum lapse rates, and the buffer factor (the buffer factor defines the minimum difference in rates before dynamic lapse occurs). The Academy also recommended a factor for the ratio of the contract GMIR to the current SNFL rate, but further clarification is needed on how this is intended to function.

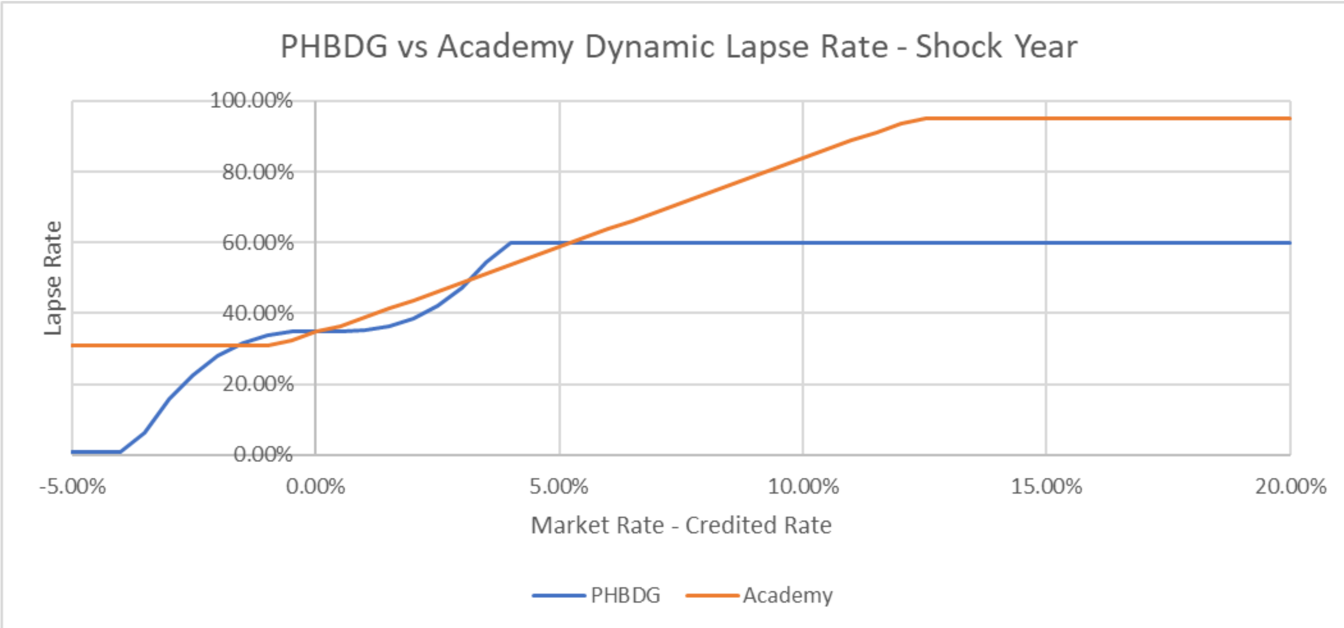
PHBDG	Academy
Total Lapse = (Base Lapse + Rate Factor) * ITM Factor	Total Lapse = (Base Lapse + Rate Factor) * ITM Factor
ITM = PVGMWB/AV	ITM = PVGMWB/AV
ITM Factor = 1 if ITM=<1.25 = $(1.25/ITM)^2$ if ITM>1.25	ITM Factor = 1 if ITM=<1.25 = $(1.25/ITM)^2$ if ITM>1.25 = 0 if AV = 0
Rate Factor = Market Factor × Max [0, 1–5*(SC percentage - MVA)]/100	Rate Factor = Market Factor × Max [0,1–10*(1-CSV/AV)]*GMIR/SNFL Era Factor
Market Factor = -1.25 * $(CR-MR)^{2.5}$ if CR>=MR = 0 if MR>CR>=(MR-BF) = 1.25 * $(MR-BF-CR)^{2.5}$ if CR<= (MR-BF)	Market Factor = X * (CR-MR) if CR>=MR = 0 if MR>CR>=(MR-BF) = Y * (MR-BF-CR) if CR<= (MR-BF)

	X = 1 during the SC period; 5 at shock; 3 thereafter Y = 3 during the SC period; 5 at shock; 6 thereafter
MVA = $[(A/B)^t-1]$ A = [1 + the closing effective yield of the "MVA Index" on the issue date] B = [1 + the closing effective yield of the "MVA Index" two days before the withdrawal, *surrender*, or annuitization] t = the number of days from the date of withdrawal, *surrender*, or annuitization to the next contract anniversary divided by 365, plus the number of whole years from the next anniversary to the end of the surrender charge period.	MVA = Embedded in CSV
CR = crediting rate at the time of projection	CR = crediting rate, or the option budget, at the time of projection
MR = market rate at the time of projection	MR = 10 Year UST + 60%BBB/40% A spread
Min Total Lapse = 1%	Min Rate Factor = -2%; -4% at shock
Max Total Lapse = 60%	Max Rate Factor = 10% during SC period, 60% at shock; 35% thereafter
GMIR/SNFL Era Factor = None; N/A	GMIR/SNFL Era Factor = TBD
Buffer (BF) = 0.50%	Buffer (BF) = 0.25%

Examples:



Assumes 2.5% base lapse rate, 5% surrender charge.



Assumes 35% base lapse rate.

Agenda Item 6

Discuss Comments Received on GOES Acceptance Criteria and Stylized Facts (Additional Material)

November 21, 2023

Ms. Rachel Hemphill, Chair, Life Actuarial (A) Task Force (LATF)
Mr. Philip Barlow, Chair, Life Risk-Based Capital (E) Working Group (Life RBC)
Mr. Mike Yanacheak, Chair, Generator of Economic Scenarios (E/A) Subgroup (GOES Subgroup)
National Association of Insurance Commissioners (NAIC)

Dear Ms. Hemphill, Mr. Barlow, and Mr. Yanacheak,

The American Academy of Actuaries'[1] Economic Scenario Generator Subcommittee (ESGS) appreciates the opportunity to offer our comments on the GOES Stylized Facts and Acceptance Criteria exposed on 10/5/23 (exposure) with the NAIC. The continued open and collaborative dialogue is greatly appreciated, particularly as you move forward with selecting an economic scenario model, stylized facts, and acceptance criteria. While we support the exposed stylized facts, we do have significant concerns with the exposed acceptance criteria and strongly encourage regulators to consider exposing a more comprehensive set of actionable criteria.

Summary

Establishing stylized facts and acceptance criteria are key steps in the traditional economic scenario modeling process. While they are necessary steps in the process, they are not the only factors that should be considered. Rather, they should be combined with a rigorous model selection step that evaluates the strengths and limitations of available models based on the intended purpose of the scenario generator. This is because model forms vary in their ability to reflect key stylized facts and meet acceptance criteria without creating other concerns, such as missing on other factors or requiring excessive overrides, like flooring.

Since no model is perfect, ideally both model selection and the establishment of stylized facts and acceptance criteria are rigorous and comprehensive exercises. This would then result in a model and calibration that is suitable for the intended purpose of the scenario generator and whose limitations are understood. On the other hand, ad hoc model selection paired with heavy use of a floor and a limited set of acceptance criteria risks producing unrealistic and unforeseen results. A more robust set of criteria can help avoid unintended consequences associated with heavy use of a floor, just as a more robust model selection process may avoid model forms that require excessive flooring.

Well-designed model office or field testing can be useful in evaluating aggregate impacts on reserve and

[1] The American Academy of Actuaries is a 19,500-member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues. The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.

capital levels, but do not represent adequate substitutes for rigorous and comprehensive model selection, stylized fact and acceptance criteria processes, and full documentation of the model and its calibration. Just as we would not support a mortality assumption that produced higher mortality in females than in males, even with appropriate levels of reserves and capital produced by model office or field testing, we would support a mortality assumption based on relevant and credible experience data, reasonable future expectations, and a conceptual understanding of theoretical relationships.

The ESGS supports the exposed stylized facts for interest rates, equity returns, and corporate bond fund returns. We note that the exposed stylized facts for equity returns and corporate bond fund returns are identical to those previously proposed by the ESGS. The exposed stylized facts for interest rates have been slightly modified from those previously proposed by the ESGS, but the changes are relatively modest.

The ESGS has significant concerns about the exposed acceptance criteria for interest rates, equity returns, and corporate bond fund returns. A primary concern is the lack of robust criteria around key stylized facts. In the exposure, key stylized facts have no actionable criteria associated with them to ensure they are adequately reflected in the scenarios. This seems especially important given the need to increase volatility to hit low for long targets under the selected model form, as well as the heavy and frequent flooring used to override the extremely negative rates often simulated under such increased volatility.

A stylized fact may state that certain behavior in the scenarios should be consistent with and plausibly more extreme than history, but it would be challenging to ensure such consistency without actionable criteria supporting the stylized fact. Exposed stylized facts for interest rate volatility and slope state that scenarios should generally be consistent with history given the level of interest rates, but there are no actionable criteria in the exposure for ensuring that is the case. This is also true for the distribution of point-in-time interest rates (both initial period and steady state) and median reversion time.

Rather than moving forward with a model or scenario generator that engenders such concern, we strongly encourage regulators to expose a more comprehensive set of actionable criteria, which would ensure the model is capable of producing scenarios that adequately reflect the stylized facts under a variety of initial conditions. The ESGS has previously proposed several categories of acceptance criteria to this end, which are included as Appendices to this letter.

Be assured that it is not our intent to suggest that a model or calibration must meet every single possible criterion to be accepted, although passing all criteria would likely increase the probability of that happening. In practice, multiple pieces of criteria may not be met. However, the model or calibration may still be accepted, given satisfactory explanations, prioritizations between criteria, and further expert review. All criteria may be met, especially if the set of criteria is rather limited in scope, and the model form or calibration may not be accepted after a full review by subject matter experts for specific rationales, such as issues related to excessive amounts of flooring. The governance process should include a report on the results of applying the individual criteria to the model or calibration, which would be reviewed by subject matter experts, along with other useful charts, statistics, and holistic judgment prior to accepting or rejecting the model or calibration. The governance process should also periodically review the acceptance criteria themselves, allowing for necessary updates, the removal of criteria that are no longer useful, or the addition of criteria for new areas of concern.

Interest Rates

The ESGS proposed eight categories of acceptance criteria for interest rates, listed below. Of those, three were included in the exposed interest rate acceptance criteria, Low-for-Long and High-for-Long, Min/Max Bounds, and Tail Frequency. No quantitative actionable criteria were exposed for the remaining categories. We strongly recommend adding the latter group of interest rate criteria to the next version of the exposure, which will help ensure the stylized facts are properly reflected in the model and its scenario sets. For example, actionable criteria can help ensure higher interest rates are indeed, on average, more volatile than lower interest rates.[2]

Table in Appendix	Categories of Proposed Criteria for Interest Rates
1.1	Level Criteria – Steady State Period
1.2	Level Criteria – Initial Period
1.3	Low-for-Long and High-for-Long
1.4	Volatility
1.5	Slope
1.6	Min/Max Bounds
1.7	Tail Frequencies
1.8	Median Reversion

Low-for-Long and High-for-Long—Exposed criteria for this category are identical to previously proposed criteria by the ESGS, but only include previously proposed criteria for the initial period. The exposure omitted previously proposed criteria for the steady state period. While such criteria for the initial period is key, the ESGS also supports having criteria to evaluate the steady state low-for-long and high-for-long behavior underlying the model. Note that the exposure also includes criteria specific to starting with 12/31/20 yields at less severe percentiles (5% and 15% instead of 1%), which was not part of but is consistent with the proposal by the ESGS. Since it is critical to understand how the model or calibration performs under a variety of initial conditions,[3] the ESGS supports including criteria that can likewise be applied under a variety of initial conditions. The ESGS believes criteria for the 1st percentile is adequate. At the request of regulators, the ESGS could consider expanding its criteria, which functions under a variety of initial conditions, to include less severe percentiles.

Min/Max Bounds—The ESGS's previously proposed criteria for Min/Max Bounds are not part of the exposure. However, the exposure does include some related criteria. In its current form, rates (all tenors)

[2] This turned out to be a shortcoming with the AIRG. Having explicit volatility criteria in place and monitored over time could have helped identify that sooner. Unrealistic pathwise volatility dynamics in the scenarios can result in distorted cash flow projections that may influence the cost of guarantees modeled, unwarranted hedge breakage, .etc.

[3] Although statutory reserve and capital calculations are based on scenarios / conditions as of the valuation date, actuarial pricing and risk management analyses generally involve forecasting reserves and capital into the future under a variety of economic conditions. Interpreting the results of pricing and risk management analyses requires an understanding of how the distribution of scenarios behaves differently when starting from different sets of initial conditions. A solid understanding (based on a robust set of criteria) of how the model performs under a wide variety of initial conditions can also help with understanding how robust the model's calibration is and under what conditions a recalibration may be warranted.

should generally not be lower than -1.5% and that 99th percentiles of 3M and 10Y rates should not exceed 20% in the first 30 years. The ESGS believes that a minimum bound of -1.5% for all tenors is too extreme, given history, and would recommend min/max bounds that vary by tenor. A 99th percentile of 20% may also be extreme, especially for longer tenors like the 20-year, where the maximum monthly rate in U.S. history is only 15.78%. There are also concerns related to the lack of min/max bounds for slope. The ESGS's previously proposed criteria has separate criteria for the 1Y rate and the 20Y rate, as well as the 20Y-1Y slope (e.g., min/max 1Y bound of -1% to -0.5% / 20% to 24%, min/max 20Y rate of 0% to 0.5% / 17% to 20%).

Tail Frequencies—The ESGS previously proposed Tail Frequencies criteria which are not included in the current exposure. However, the current iteration does include some related criteria, stating that no more than 5% of scenarios should have 3M or 10Y rates that exceed 20% in the first 30 years. The ESGS supports having tail frequency criteria for both low and high rates, with thresholds that vary by tenor. The ESGS's previously proposed tail frequency criteria apply to both low and high rates with thresholds set to historical minimums and maximums that vary by tenor, reflecting the idea that rates which are more extreme than historical rates should be simulated approximately 1 to 3% of the time (0.5% to 1.5% on for each tail, left and right). Similarly, slopes that are more extreme than historical slopes should be simulated approximately 1 to 4% of the time (0.5% to 2% for each tail, left and right). We would note that the ESGS criteria are for individual monthly rates; regulators would be able to request the ESGS develop additional similar criteria for individual scenario paths (i.e., multiple consecutive monthly rates). It is also important to keep in mind that criteria for Min/Max Bounds and other severe rate levels are not that useful without associated frequency criteria, such as the frequency of extremely low/high rates close to the Min/Max Bounds. As exposed, the criteria would not preclude a scenario set where rates are negative half the time.

A note on flooring—There is academic literature on the limitations of the 3-factor affine model structures used to simulate interest rates. Depending on the intended purpose, such model limitations may not be relevant. However, some of those limitations may hamper the model's ability to adequately reflect the stylized facts exposed for this model's intended purpose, determining statutory reserves and capital for long duration life and annuity products. This could result in overly frequent and severe negative rates and distorted volatility and yield curve / term structure relationships, requiring overly excessive post-model overrides such as too much flooring.[4] Some flooring/capping of outlier edge cases due to random noise is reasonable in stochastic models, but any stochastic model should be called into question if it requires overriding a large percentage of rates in a large percentage of scenarios to properly reflect the stylized facts and reasonably satisfy the acceptance criteria developed for the purpose at hand.[5] A rigorous and comprehensive model selection step, based on the stylized facts, can help avoid such a situation. Appendix 4 includes illustrative charts that offer examples of potentially excessive flooring of simulated Treasury yields.

[4] For example, excessive post-model flooring of rates can break some of the desired relationships (e.g., smooth yield curves, minimal arbitrage opportunities) inherent in the model form and the pre-floored rates it produces.

[5] Model forms that may be able to satisfy low for long criteria with significantly less flooring include shadow rate models and certain types of dynamic Nelson Siegel models.

Equity Returns

We will be sharing the results of our updates to and expansion of the 2005 C3P2 Gross Wealth Factors (GWFs) for S&P500 total returns with LATF imminently. Our updated criteria are largely consistent with the 2005 C3P2 GWF criteria when equity reference models have means constrained to 8.75% (as was the case for the 2005 GWFs). The updated GWFs do go further into the tails (include 1st and 99th percentiles), as well as further into the future (30 and 50-year horizons) given the changes to VM-21 (more extreme CTE level of 98%) and expansion of scope to VM-20 (longer duration products). We would strongly recommend regulators use the largely consistent and expanded set of updated GWFs in place of the 2005 GWFs.

We also note that the exposure only contains criteria for the S&P 500 index. This is concerning, as it means the exposure contains no criteria for indices other than the S&P 500, nor does it include criteria for the joint distribution of equity returns and interest rates. Prior NAIC boundary guidance included criteria reflecting the need for other equity indices to include Sharpe ratios (i.e., market price of risk) within 5% of the S&P 500's Sharpe ratio. The ESGS intends to develop criteria for the joint distribution of equity returns and interest rates, such as criteria for quadrants of low interest rates and low equity returns, and low interest rates and high equity returns.

Corporate Bond Fund Returns

The ESGS proposed four categories of acceptance criteria for corporate bond fund returns, listed below. The exposed corporate bond fund return criteria only include Average Excess Return. No quantitative actionable criteria were exposed for the remaining categories. We strongly recommend incorporating the additional three categories of corporate bond fund return criteria into the next exposure, in order to ensure the stylized facts are properly reflected in the model and its scenario sets.

Table in Appendix	Categories of Proposed Criteria for Corporate Bond Fund Returns
3.1	Average Excess Return
3.2	Maximum Excess Return
3.3	Correlations
3.4	Median Reversion

Average Excess Return—The exposed criteria are very similar to the ESGS's previously proposed criteria, with the primary difference related to the use of conservative one-way rather than two-way buffers (average excess returns can only be less than target excess returns). For example, instead of a desired range of 70 to 90 bps (i.e., 80 +/- 10 bps) for the average excess return on 1 to 5-year investment grade corporate bond funds, the exposure uses a desired range 70 to 80 bps. While we understand the desire to be conservative, our recommendation is to leave the scenarios centered economically and apply conservatism via another lever, such as the CTE level.

Closing Remarks

The ESGS appreciates the opportunity to review the exposure. We are confident that the NAIC's collaborative strategy to incorporate public feedback and recommendations will ensure criteria that is beneficial to regulators and industry. We look forward to the discussion at the Fall National Meeting and to continuing to work with you to develop a comprehensive set of acceptance criteria that ensures an economic scenario generator that properly reflects stylized facts and is fit for purpose. Please direct any questions to Amanda Barry-Moilanen, life policy analyst at barrymoilanen@academy.org.

Sincerely,

Jason Kehrberg
Chair, Economic Scenario Generator Subcommittee

Appendix 1—Interest Rate Acceptance Criteria Proposed by AAA

Table 1.1—Level Criteria – Steady State Period

Rate	Statistic (Percentile)	Desired Range
1Y	1st	-0.84% to 0.06%
	5th	-0.70% to 0.10%
	15th	-0.54% to 0.16%
	30th	-0.11% to 0.49%
	Median	1.31% to 3.35%
	70th	4.88% to 6.88%
	85th	6.22% to 8.47%
	95th	9.02% to 11.52%
	99th	13.85% to 16.60%
20Y	1st	0.22% to 1.12%
	5th	0.98% to 1.78%
	15th	1.61% to 2.31%
	30th	2.23% to 2.83%
	Median	3.35% to 4.89%
	70th	5.77% to 7.77%
	85th	7.56% to 9.81%
	95th	9.50% to 12.00%
	99th	13.44% to 16.19%

Notes:

1. Non-Median criteria is based on historical Percentiles Exponentially Weighted (PEWs) using a half-life of 15 years and a data period of 1953.05 to 2021.12, plus or minus a buffer depending on whether the percentile is in the left or right tail respectively.
2. Median criteria are based on historical 40th and 50th PEWs.
3. Steady state statistics can be measured over a single steady state month or multiple consecutive steady state months, e.g., over 240 months (20 years). One option for a 20-year steady state period over which steady state statistics can be measured is months 961-1200, e.g., the last 20 years of a 100-year projection. Another option would be to start the model under steady state conditions and then use the first 20 years.

Table 1.2—Level Criteria – Initial Period

Initial Level of 20Y Rate	Statistic: Percentiles of 20Y Rate					
	End of year 1		End of year 5		End of year 10	
	1%-tile should be less than	99%-tile should be greater than	1%-tile should be less than	99%-tile should be greater than	1%-tile should be less than	99%-tile should be greater than
1%	0.54%	1.92%	0.60%	3.89%	0.72%	6.05%
2%	1.22%	3.30%	0.79%	5.75%	0.81%	8.10%
3%	1.92%	4.66%	1.20%	7.48%	0.95%	9.62%
4%	2.62%	6.01%	1.62%	8.83%	1.23%	10.77%
5%	3.31%	7.22%	2.03%	10.03%	1.50%	11.87%
6%	3.99%	8.38%	2.43%	11.21%	1.75%	12.93%
7%	4.68%	9.52%	2.81%	12.35%	2.00%	13.95%
8%	5.46%	10.64%	3.18%	13.46%	2.23%	14.92%
9%	6.26%	11.76%	3.58%	14.56%	2.45%	15.78%
10%	7.06%	12.86%	4.09%	15.62%	2.66%	16.48%

Notes:

1. Due to the lack of historical data for percentiles of the 20Y rate when starting at a multitude of initial rate levels, criteria were developed by taking the least binding statistic from 3 different reference models (CIR, Black Karasinski, and Brennan Schwartz) calibrated to steady state criteria over 3 different mean reversion speeds (half-lives of 10, 12, and 15 years).
2. These criteria ensure sufficient dispersion in 20Y rate levels at specific points in time during the initial period. The end of years 1, 5, and 10 were selected as round points-in-time to test during the initial period when simulated rates are still materially impacted by starting levels. Other points-in-time could also be considered.
3. 1st and 99th percentiles were selected as the tail severities (reasonably extreme given the purpose). Other percentiles could also be considered.
4. When evaluating an initial calibration of an ESG model, it would be prudent to test the model at a variety of starting 20Y rate levels, e.g., 2%, 5%, and 8%. When evaluating a single candidate scenario set for production, these criteria can be applied by interpolating based on the starting level of the 20Y rate.
5. These criteria were developed for the 20Y rate given its central role in the AIRG and use in other criteria. Similar criteria could also be developed and considered for the 1Y rate.

Table 1.3—Low-for-Long and High-for-Long Criteria

Period	Initial Level of 20Y Rate	Statistic: Percentiles of Geometric Average of 20Y Rate			
		10-year horizon		30-year horizon	
		1%-tile should be less than	99%-tile should be greater than	1%-tile should be less than	99%-tile should be greater than
Initial (from year 0)	1%	0.94%	3.43%	1.50%	6.25%
	2%	1.23%	5.05%	1.68%	7.71%
	3%	1.62%	6.55%	1.86%	8.72%
	4%	2.15%	7.74%	2.06%	9.62%
	5%	2.66%	8.87%	2.26%	10.46%
	6%	3.15%	9.96%	2.50%	11.16%
	7%	3.63%	11.03%	2.78%	11.61%
	8%	4.10%	12.07%	3.06%	11.99%
	9%	4.64%	13.08%	3.34%	12.33%
	10%	5.21%	14.01%	3.65%	12.63%
Steady State (e.g., from year 70)	Any	1.34%	13.57%	1.94%	11.45%

Notes:

1. Due to the lack of historical data for percentiles of the geometric average of the 20Y rate when starting at a multitude of initial rate levels, criteria were developed by taking the least binding statistic from 3 different reference models (CIR, Black Karasinski, and Brennan Schwartz) calibrated to steady state criteria over 3 different mean reversion speeds (half-lives of 10, 12, and 15 years).
2. These criteria ensure sufficient dispersion in geometric average 20Y rate levels over specific horizons during the initial and steady state periods. Horizons of 10 and 30 years are consistent with the NAIC's preliminary low-for-long boundary guidance. Other horizons could also be considered.
3. 1st and 99th percentiles were selected as the tail severities (reasonably extreme given the purpose). Other percentiles could also be considered.
4. When evaluating an initial calibration of an ESG model, it would be prudent to test the model at a variety of starting 20Y rate levels, e.g., 2%, 5%, and 8%. When evaluating a single candidate scenario set for production, these criteria can be applied by interpolating based on the starting level of the 20Y rate.
5. These criteria were developed for the 20Y rate given its central role in the AIRG and use in other criteria. Similar criteria could also be developed and considered for the 1Y rate.

Table 1.4—Volatility Criteria

Statistic	Bucket (beginning of month rate is…)	Desired Range
Annualized standard deviation of monthly changes in the 1Y rate under three different rate level buckets	<= 3%	0.30% to 0.89%
	> 3% to <= 8%	0.58% to 1.73%
	> 8%	1.67% to 5.02%
Annualized standard deviation of monthly changes in the 20Y rate under three different rate level buckets	<= 3%	0.31% to 0.92%
	> 3% to <= 8%	0.37% to 1.12%
	> 8%	0.78% to 2.33%

Notes:

1. Desired range is based on a 50% margin around the historical statistic using a data period of 1953.05 to 2021.12. E.g., the historical annualized standard deviation of monthly changes in the 1Y rate when the beginning of month rate is <= 3% is 0.59%, half of 0.59% is 0.295%, and 0.59% +/- 0.295% results in a desired range of 0.30% to 0.89%.
2. The scenario set statistic can be measured over a single month or multiple consecutive months, e.g., over years 1-10 to evaluate the initial period and years 80-100 to evaluate the steady state period (or could start the model under steady state conditions and then use the first 20 years). Expect more variation for initial period statistics due to the impacts of starting rate and/or volatility levels (e.g., clustering).

Table 1.5—Slope Criteria

Statistic (Percentiles of 20Y-1Y under three different buckets for the 20Y rate level)		Desired Range
1st	<= 3%	-0.32% to 0.18%
	> 3% to <= 8%	-1.73% to -1.23%
	> 8%	-3.43% to -2.93%
5th	<= 3%	-0.23% to 0.27%
	> 3% to <= 8%	-0.97% to -0.47%
	> 8%	-2.06% to -1.56%
10th	<= 3%	-0.11% to 0.39%
	> 3% to <= 8%	-0.71% to -0.21%
	> 8%	-1.79% to -1.29%
15th	<= 3%	-0.01% to 0.49%
	> 3% to <= 8%	-0.56% to -0.06%
	> 8%	-1.46% to -0.96%
85th	<= 3%	2.28% to 2.78%
	> 3% to <= 8%	3.23% to 3.73%
	> 8%	1.94% to 2.44%
90th	<= 3%	2.52% to 3.02%
	> 3% to <= 8%	3.44% to 3.94%
	> 8%	2.05% to 2.55%
95th	<= 3%	2.64% to 3.14%
	> 3% to <= 8%	3.71% to 4.21%
	> 8%	2.41% to 2.91%
99th	<= 3%	2.81% to 3.31%
	> 3% to <= 8%	4.06% to 4.56%
	> 8%	2.76% to 3.26%

Notes:

1. Desired range is based on historical slope percentiles and a data period of 1953.05 to 2021.12, plus or minus a 50 basis point buffer depending on whether the percentile is in the left or right tail respectively. E.g., the historical 1st slope percentile when the 20Y rate is <= 3% is 0.18%, 0.18% less 50 basis points is -0.32%, resulting in a desired range of -0.32% to 0.18%.
2. The scenario set statistic can be measured over a single month or multiple consecutive months, e.g., over years 1-10 to evaluate the initial period and years 80-100 to evaluate the steady state period (or could start the model under steady state conditions and then use the first 20 years). Expect more variation for initial period statistics due to the impacts of starting rate and/or volatility levels (e.g., clustering).

Table 1.6—Min/Max Bounds Criteria

Statistic (over entire projection period)	History (for reference)	Desired Range
1Y Min	0.05%	-1% to -0.5%
1Y Max	16.97%	20% to 24%
20Y Min	0.95%	0% to 0.5%
20Y Max	15.78%	17% to 20%
20Y-1Y Min (when 20Y <= 3%)	0.02%	-1.5% to -0.5%
20Y-1Y Min (when 20Y > 3% to <=8%)	-1.38%	-3.5% to -2%
20Y-1Y Min (when 20Y > 8%)	-3.36%	-5% to -4%
20Y-1Y Max (when 20Y <= 3%)	2.85%	3% to 4%
20Y-1Y Max (when 20Y > 3% to <=8%)	4.15%	4.5% to 6%
20Y-1Y Max (when 20Y > 8%)	2.90%	3.5% to 5.5%

Notes:
1. Historical statistics are based on a data period of 1953.05 to 2021.12.

Table 1.7—Tail Frequencies Criteria

Statistic (Worse-Than-History frequencies during steady state period)	Desired Range
Freq of 1Y < 0.05%	0.5% to 1.5%
Freq of 1Y > 16.97%	0.5% to 1.5%
Freq of 20Y < 0.95%	0.5% to 1.5%
Freq of 20Y > 15.78%	0.5% to 1.5%
Freq of 20Y-1Y (when 20Y <= 3%) < 0.02%	0.5% to 2.0%
Freq of 20Y-1Y (when 20Y > 3% to <=8%) < -1.38%	0.5% to 2.0%
Freq of 20Y-1Y (when 20Y > 8%) < -3.36%	0.5% to 2.0%
Freq of 20Y-1Y (when 20Y <= 3%) > 2.85%	0.5% to 2.0%
Freq of 20Y-1Y (when 20Y > 3% to <=8%) > 4.15%	0.5% to 2.0%
Freq of 20Y-1Y (when 20Y > 8%) > 2.90%	0.5% to 2.0%

Notes:
1. Historical statistics are based on a data period of 1953.05 to 2021.12.
2. Steady state statistics can be measured over a single steady state month or multiple consecutive steady state months, e.g., over 240 months (20 years). One option for a 20-year steady state period over which steady state statistics can be measured is months 961-1200, e.g., the last 20 years of a 100-year projection. Another option would be to start the model under steady state conditions and then use the first 20 years.

Table 1.8—Median Reversion Criteria

Statistic (Year median rate/slope reaches midpoint between initial and ultimate levels)	Desired Range
1Y rate	10 to 20 years
20Y rate	10 to 20 years
20Y-1Y slope	2 to 8 years

Notes:
1. The midpoint can be determined as the average of the starting (beginning of year 0) level and the median ultimate (e.g., end of year 100) level.
2. Criteria may not perform well if the median path is materially nonmonotonic.

Appendix 2—Equity Return Acceptance Criteria Proposed by AAA

Table 2.1—Gross Wealth Factor (GWF) Criteria

GWF Percentiles	Horizon 1 year	5 years	10 years	20 years	30 years	50 years
Min	0.48	0.28	0.33	0.32	0.56	0.85
1st	0.71	0.64	0.71	0.99	1.55	4.15
5th	0.83	0.84	1.02	1.62	2.73	8.63
10th	0.89	0.98	1.22	2.10	3.74	12.78
15th	0.93	1.07	1.38	2.46	4.55	16.49
30th	1.02	1.28	1.76	3.41	6.84	27.56
70th	1.17	1.73	2.70	6.14	13.50	62.71
85th	1.24	1.97	3.27	8.41	20.39	112.78
90th	1.28	2.09	3.58	9.59	23.93	142.63
95th	1.33	2.28	4.08	11.43	30.68	195.72
99th	1.42	2.67	5.10	15.83	45.17	333.02
Max	1.67	3.75	8.01	29.20	99.48	1019.62

Notes:
1. These criteria are based on reference models fit to S&P 500 total returns from 1957.03 through 2022.12 with the mean total return constrained to be 8.75%. Unconstrained mean total returns ranged from 11.37% to 11.94% across the reference models.
2. To ensure sufficient dispersion in the distribution, left tail percentiles should be less than their respective criteria, and right tail percentiles should be greater than their respective criteria.

Table 2.2—Relationship to Interest Rates (Joint/Quadrant) Criteria

TBD

Table 2.3—Relationship to S&P 500 Criteria

TBD

As a placeholder, we suggest using as criteria that Sharpe ratios for total returns on other indices be within 5% of the Sharpe Ratio for the S&P 500.

Appendix 3—Corporate Bond Fund Return Acceptance Criteria Proposed by AAA

Table 3.1—Average Excess Return Criteria

Corporate Bond Fund	Desired Range for Average Steady State Excess Return
IG 1-5	70 to 90 bps
IG 5-10	69 to 89 bps
IG Long	56 to 76 bps
High Yield	220 to 260 bps

Notes:

1. Excess return equals total return on corporate bond fund less total return on government bond fund of similar duration.
2. Criteria is based on prescribed VM-20 ultimate spreads as of 12/31/21 and Bloomberg bond fund data from 1991 to 2021.
3. Suggested period for determining average steady state excess returns is years 20-30 (months 241-360). Alternatively, the first 10 years of the projection can be used if the model is started with initial conditions equal to steady state.

Table 3.2—Maximum Excess Return Criteria

Corporate Bond Fund	Maximum excess return should be less than
IG 1-5	157 bps
IG 5-10	241 bps
IG Long	263 bps
High Yield	548 bps

Notes:

1. Excess return equals total return on corporate bond fund less total return on government bond fund of similar duration.
2. Criteria determined by adding 50 bps to average prescribed VM-20 ultimate spreads as of 12/31/21.
3. Criteria can be applied over the entire projection (i.e., applies to both initial and steady state periods).

Table 3.3—Correlations Criteria

	Corp Bond Fund	SPX Variance	SPX Return	Spread			Excess Return		
				IG 1-5	IG 5-10	IG Long	IG 1-5	IG 5-10	IG Long
Spread	IG 1-5	0.5 to 0.7	-0.5 to -0.7						
	IG 5-10	0.5 to 0.7	-0.5 to -0.7	>0.8					
	IG Long	0.5 to 0.7	-0.5 to -0.7	>0.8	>0.8				
	High Yield	0.5 to 0.7	-0.5 to -0.7	>0.8	>0.8	>0.8			
Excess Return	IG 1-5	-0.5 to -0.7	0.5 to 0.7						
	IG 5-10	-0.5 to -0.7	0.5 to 0.7				>0.8		
	IG Long	-0.5 to -0.7	0.5 to 0.7				>0.8	>0.8	
	High Yield	-0.5 to -0.7	0.5 to 0.7				>0.8	>0.8	>0.8

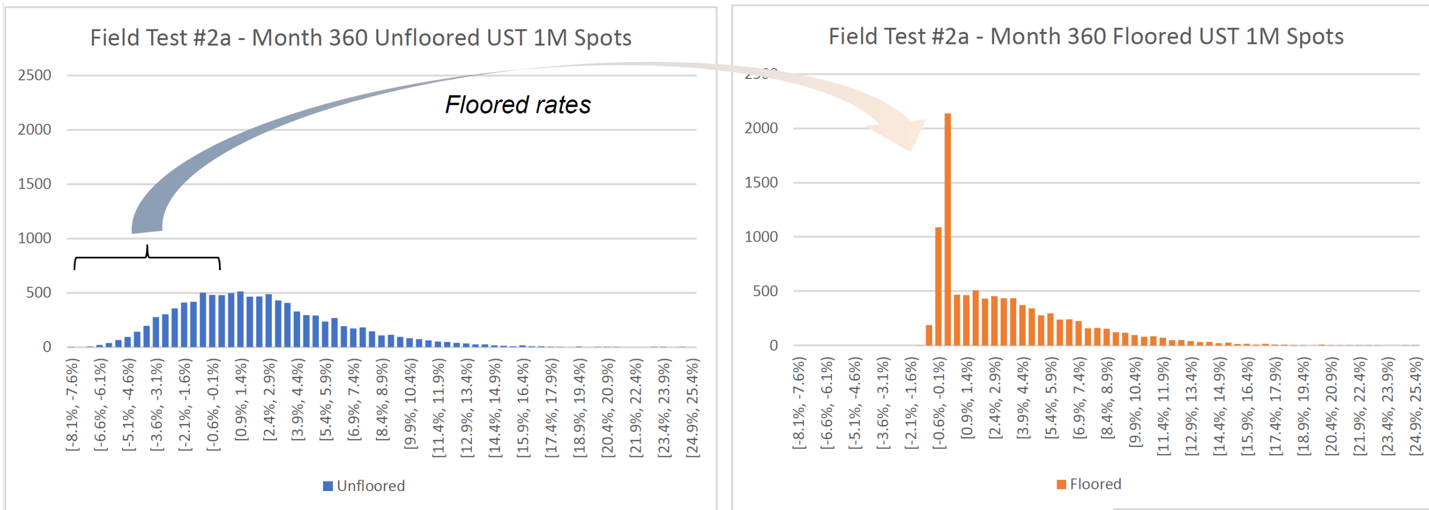
Notes:
1. Criteria based on Bloomberg bond fund data from 1991 to 2021.
2. Criteria can be applied over the entire projection (i.e., applies to both initial and steady state periods).

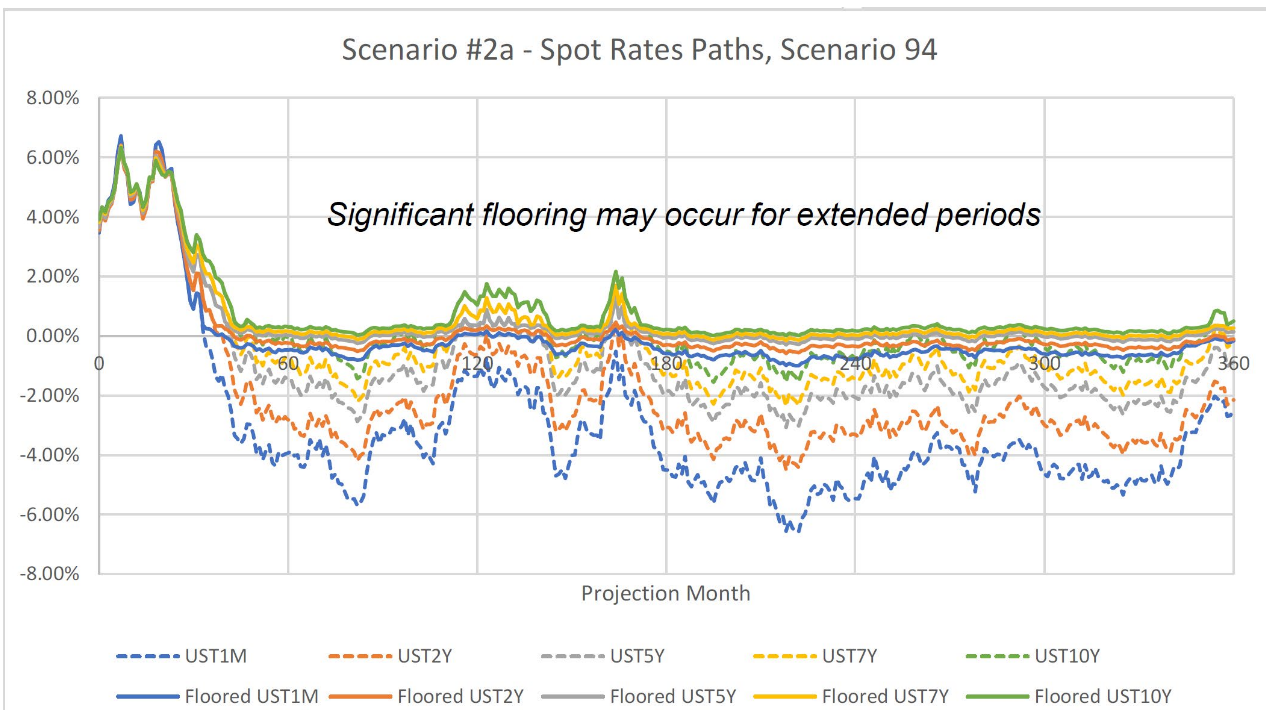
Table 3.4—Median Reversion Criteria

Statistic (Year median spread reaches midpoint between initial and ultimate levels)	Desired Range
IG 1-5	22 to 26 months
IG 5-10	22 to 26 months
IG Long	22 to 26 months
High Yield	22 to 26 months

Notes:
1. The midpoint can be determined as the average of the starting (beginning of year 0) level and the median ultimate (e.g., end of year 100) level.
2. Criteria based on VM-20, which prescribes a 4-year grading period for general account fixed income credit spreads (i.e., midpoint at 24 months).
3. Criteria may not perform well if the median path is materially nonmonotonic.

Appendix 4—Charts illustrating examples of potentially excessive flooring of simulated Treasury yields





Agenda Item 8

Review Results of New Calibration of the GOES

Agenda

1. Review of Treasury Scenarios vs Acceptance Criteria
2. Review of Equity Scenarios vs Acceptance Criteria

Review of New Treasury Calibration Against Acceptance Criteria

Treasury Calibration Approach

Field Test 1A

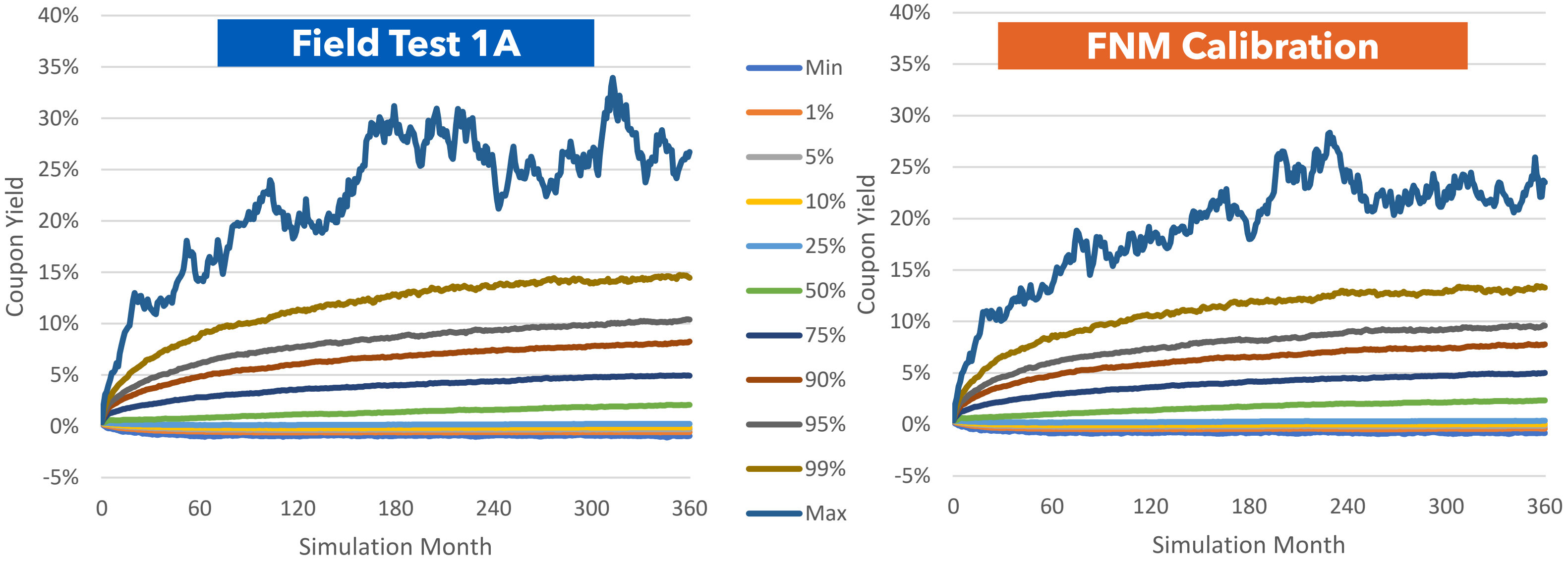
- In the NAIC's 2022 GOES Field Test, the 1A scenario set utilized the Conning calibration with a generalized fractional floor and a starting date of 12/31/21.
- 1A was designed to meet the regulator acceptance criteria, including the constraining low for long criteria based on the 10- and 30-year geometric averages and level of interest rates at 12/31/20.
- Field test participants and other commenters noted several issues with scenario set 1A, including:
  - UST rates in excess of 25%
  - High frequency, severity, and duration of inversions
  - High frequency and severity of negative interest rates

Fall National Meeting (FNM) Calibration

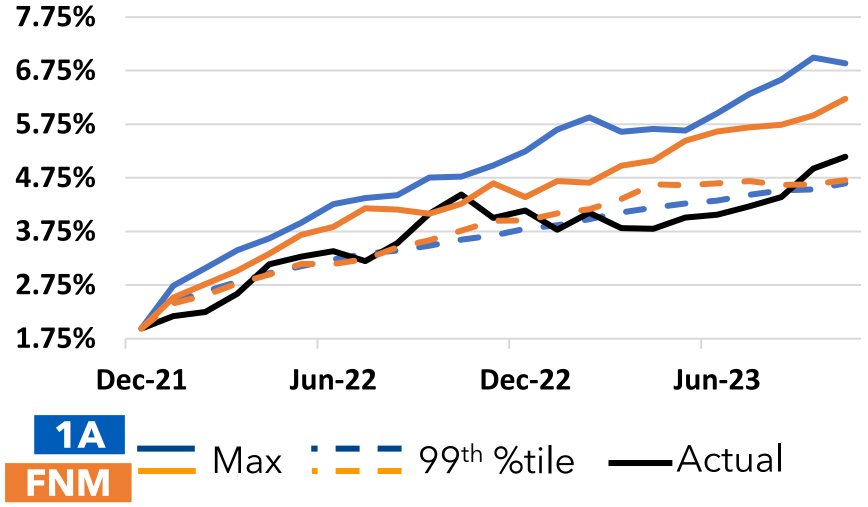
- For the 10/5/23 exposure of acceptance criteria, regulators relaxed the NAIC 30-year low for long criteria. Additional low for long and high for long criteria recommended by the Academy were also included.
- With the relaxed NAIC low for long criteria, Conning performed testing using the parameters and flooring from the 1A calibration, then varied the mean reversion speed parameter to allow the model to revert more quickly and address some of the issues noted with 1A.

Item	Category	Criteria
T1.	Prevalence of High Rates, Upper Bound on Treasury Rates	**a) The scenario set should reasonably reflect history, with some allowance for more extreme high and low interest rate environments** b) Upper Bound: i. [20%] is >= [99%]-tile on the 3M yield fan chart, and no more than [5%] of scenarios have 3M yields that go above [20%] in the first 30 years ii. [20%] is >= [99%]-tile on the 10Y yield fan chart, and no more than [5%] of scenarios have 10Y yields that go above [20%] in the first 30 years

10,000 1Y UST Scenarios as of 12/31/21 Fan Charts



10,000 20Y UST Scenarios as of 12/31/21 : Max and 99th Percentiles for 1A and FNM vs Actual

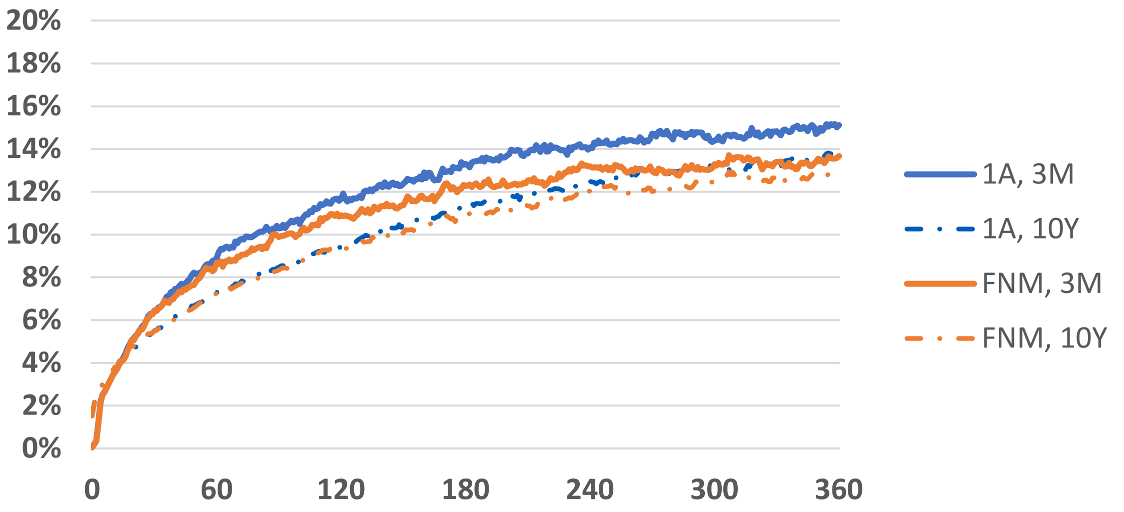


There is some subjectivity in the T1a criterion, and a broader discussion is needed to evaluate fully. Looking at maximum rates over the first 30 years, the 1Y UST is moderated in the FNM Calibration with a high of ~28% compared to ~34% for 1A. The historical max 1Y UST was ~17%. However, there was a slight miss for the FNM Calibration with a 12/31/21 start date for the 20Y UST compared to the actual in October of 2022. 1A was able to produce 20Y rates as high as the actual with a 12/31/21 start date.

Item	Category	Criteria
T1.	Prevalence of High Rates, Upper Bound on Treasury Rates	a) The scenario set should reasonably reflect history, with some allowance for more extreme high and low interest rate environments **b) Upper Bound:** **i. [20%] is >= [99%]-tile on the 3M yield fan chart, and no more than [5%] of scenarios have 3M yields that go above [20%] in the first 30 years** **ii. [20%] is >= [99%]-tile on the 10Y yield fan chart, and no more than [5%] of scenarios have 10Y yields that go above [20%] in the first 30 years**

	2022 Field Test 1A			**NAIC Fall National Meeting Calibration**		
	99th Percentile	**% Scenarios >20%**	**Pass/Fail Criteria**	**99th Percentile**	**% Scenarios >20%**	**Pass/Fail Criteria**
b.i	15.12%	0.1% <5%	✓Pass	13.98%	0.0% <5%	✓Pass
b.ii	13.87%	0.0% <5%	✓Pass	13.56%	0.0% <5%	✓Pass

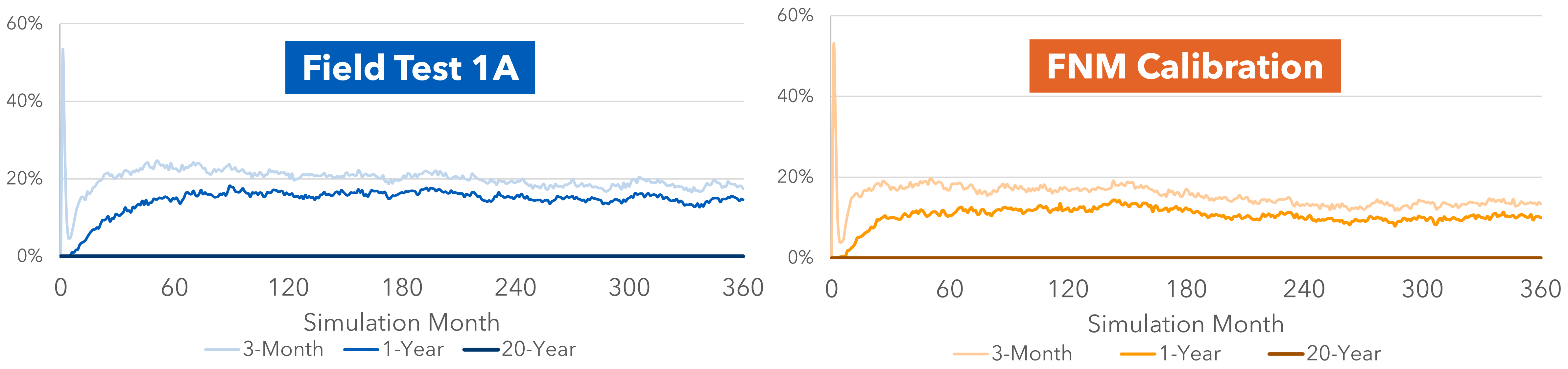
10,000 10Y and 3M UST Scenarios as of 12/31/21 : 99th Percentiles for 1A and FNM



Both the 1A and FNM Calibrations pass the objective T1b criteria for both the 3M UST and the 10Y UST.

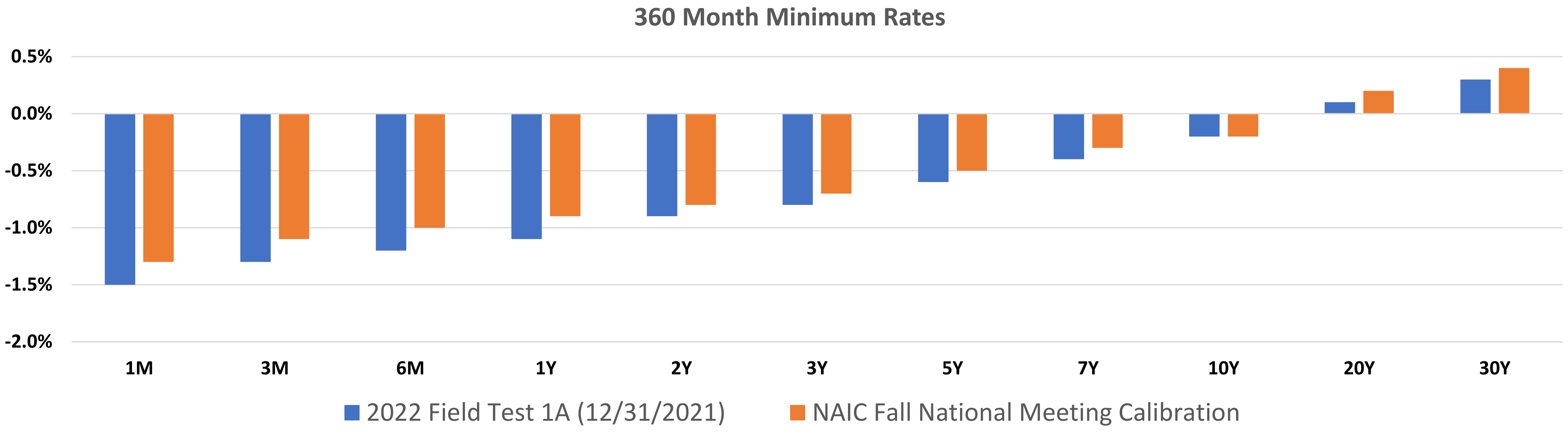
Item	Category	Criteria
T2.	Lower Bound on Negative Interest Rates, Arbitrage Free Considerations	Apply the following guidance for negative rates: **a) All maturities could experience negative interest rates** b) Interest rates may remain negative for multi-year time periods c) Rates should generally not be lower than -1.5%

Frequency of Negative Rates, 12/31/21 Scenario Sets



While the T2a criterion allows for all maturities to experience negative interest rates, there are no negative rates for the 20Y UST in the first thirty years in either 1A or the FNM Calibration. The FNM calibration has materially fewer negative interest rates than 1A (e.g. 9.2% of negative 1Y UST rates at year 30, vs 14% for 1A)

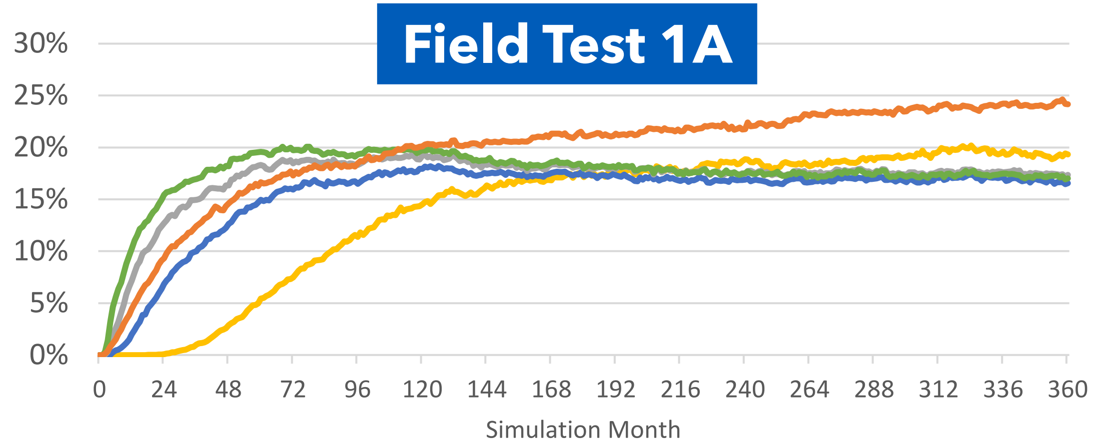
Item	Category	Criteria
T2.	Lower Bound on Negative Interest Rates, Arbitrage Free Considerations	Apply the following guidance for negative rates: a) All maturities could experience negative interest rates b) Interest rates may remain negative for multi-year time periods **c) Rates should generally not be lower than -1.5%**

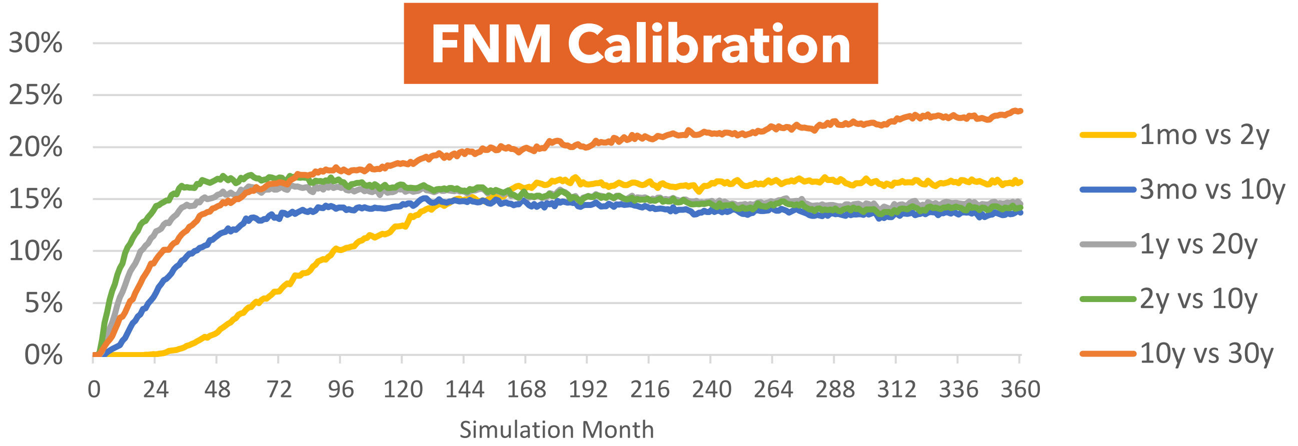


Both the 1A and Fall National Meeting Calibrations pass the objective T2c criteria, with minimums for the first 30 years of the projection never going below -1.5% for all maturities. The FNM Calibration has minimums that are less negative (or more positive) compared to 1A.

Item	Category	Criteria
T3.	Initial Yield Curve Fit, Yield Curve Shapes in Projection, and Steady State Yield Curve Shape	a) Review initial actual vs. fitted spot curve differences for a sampling of 5 dates representing different shapes and rate levels for the entire curve and review fitted curves qualitatively to confirm they stylistically mimic the different actual yield curve shapes **b) The frequency of different yield curve shapes in early durations should be reasonable considering the shape of the starting yield curve (e.g. a flatter yield curve leads to more inversions).** c) The steady state curve has normal shape (not inverted for short maturities, longer vs shorter maturities, or between long maturities)

Inversion Frequencies, 12/31/21 Scenario Sets





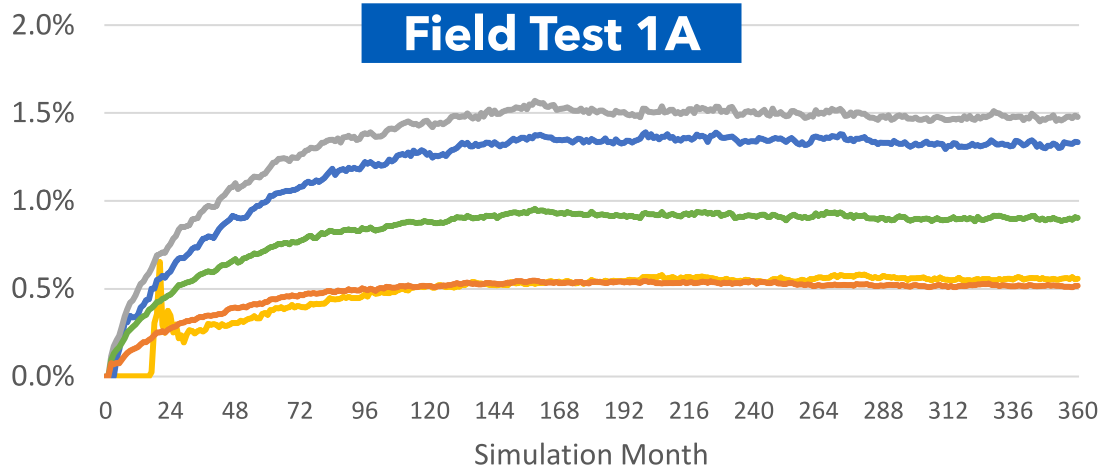
Historical Inversion Data	1m > 2y	3m > 10y	2y > 10y	10y > 30y	1y > 20y
% Inversions, 4/1953 to 3/2021*	10%	10%	19%	22%	16%
% Inversions, 12/31/21 to 11/21/23**	42%	57%	74%	6%	63%

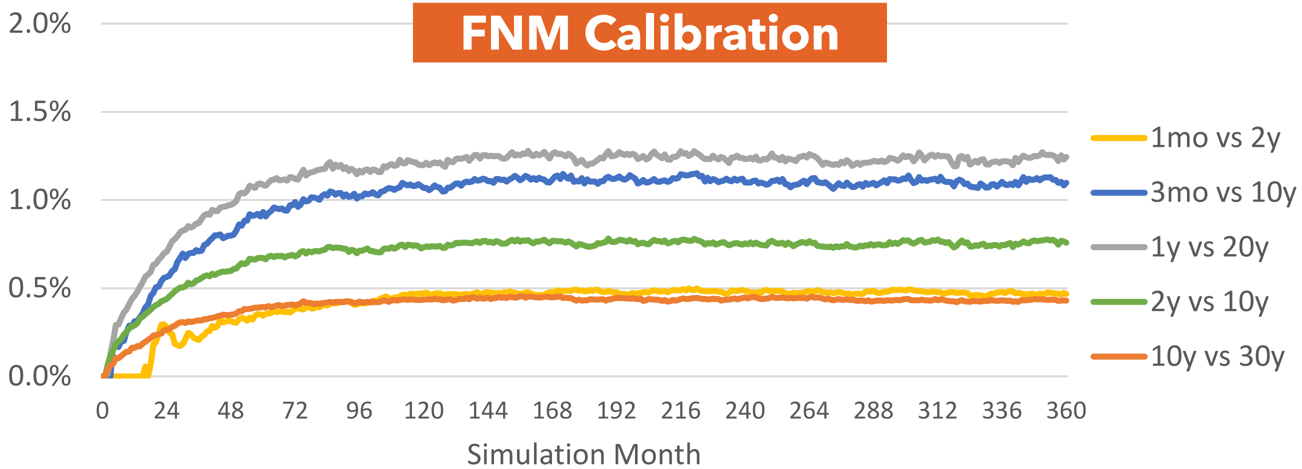
12/31/21 had a typical normal yield curve shape, so starting from zero inversions and moving to higher ultimate levels fits the T3b criteria. The frequency of inversions at the end of 30 years is less in the FNM calibration compared to 1A.

***Based on month-end data** ****Based on daily data**

Item	Category	Criteria
T3.	Initial Yield Curve Fit, Yield Curve Shapes in Projection, and Steady State Yield Curve Shape	a) Review initial actual vs. fitted spot curve differences for a sampling of 5 dates representing different shapes and rate levels for the entire curve and review fitted curves qualitatively to confirm they stylistically mimic the different actual yield curve shapes **b) The frequency of different yield curve shapes in early durations should be reasonable considering the shape of the starting yield curve (e.g. a flatter yield curve leads to more inversions).** c) The steady state curve has normal shape (not inverted for short maturities, longer vs shorter maturities, or between long maturities)

Average Level of Inversion, 12/31/21 Scenario Sets





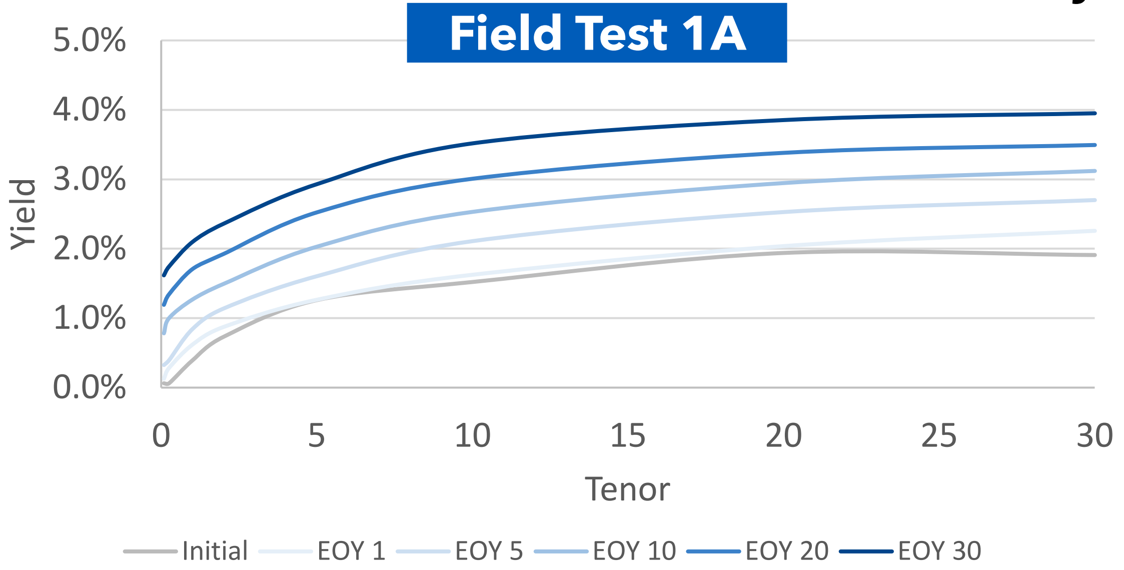
Historical Inversion Data	1m > 2y	3m > 10y	2y > 10y	10y > 30y	1y > 20y
Average Inversion, 4/1953 to 3/2021*	0.33%	0.54%	0.38%	0.22%	0.63%
Average Inversion, 12/31/21 to 11/21/23**	0.59%	1.18%	0.56%	0.05%	0.74%

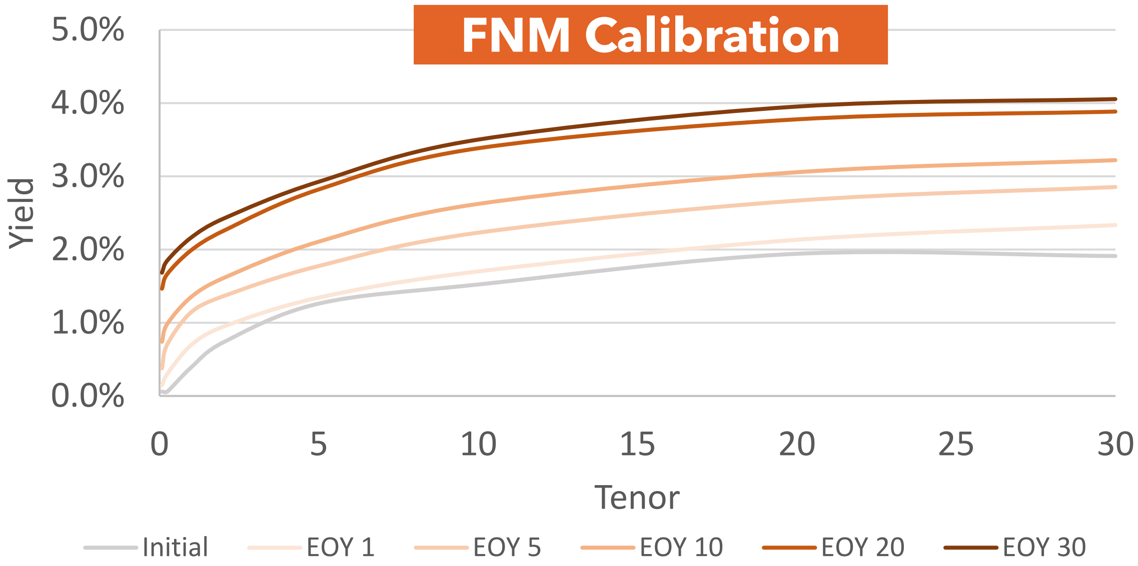
The average level of inversion (when looking only at scenarios that are inverted) at the end of 30 years is higher for the FNM and 1A calibrations compared to the average historical level from 1953 to 2021. The FNM calibration has lower levels of inversion than 1A.

*Based on month-end data **Based on daily data

Item	Category	Criteria
T3.	Initial Yield Curve Fit, Yield Curve Shapes in Projection, and Steady State Yield Curve Shape	a) Review initial actual vs. fitted spot curve differences for a sampling of 5 dates representing different shapes and rate levels for the entire curve and review fitted curves qualitatively to confirm they stylistically mimic the different actual yield curve shapes b) The frequency of different yield curve shapes in early durations should be reasonable considering the shape of the starting yield curve (e.g. a flatter yield curve leads to more inversions). **c) The steady state curve has normal shape (not inverted for short maturities, longer vs shorter maturities, or between long maturities)**

Median Yields at Selected Projection Months, 12/31/21 Scenario Sets





For both the 1A and FNM calibrations with a start date of 12/31/21, the median yield curve at the end of 30 years was normal, meeting the T3c criteria.

Item	Category	Criteria
T4.	Low For Long: 12/31/20 Starting Conditions	a) At least 10% of scenarios need a 10-year geometric average of the 20-year UST below **1.45%** b) At least 5% of scenarios need a 30-year geometric average of the 20-year UST below **1.95%** **Note: As part of the model acceptance process, a given calibration of the GOES will be tested at multiple starting dates. This criteria is relevant for the 12/31/20 starting yield curve.**

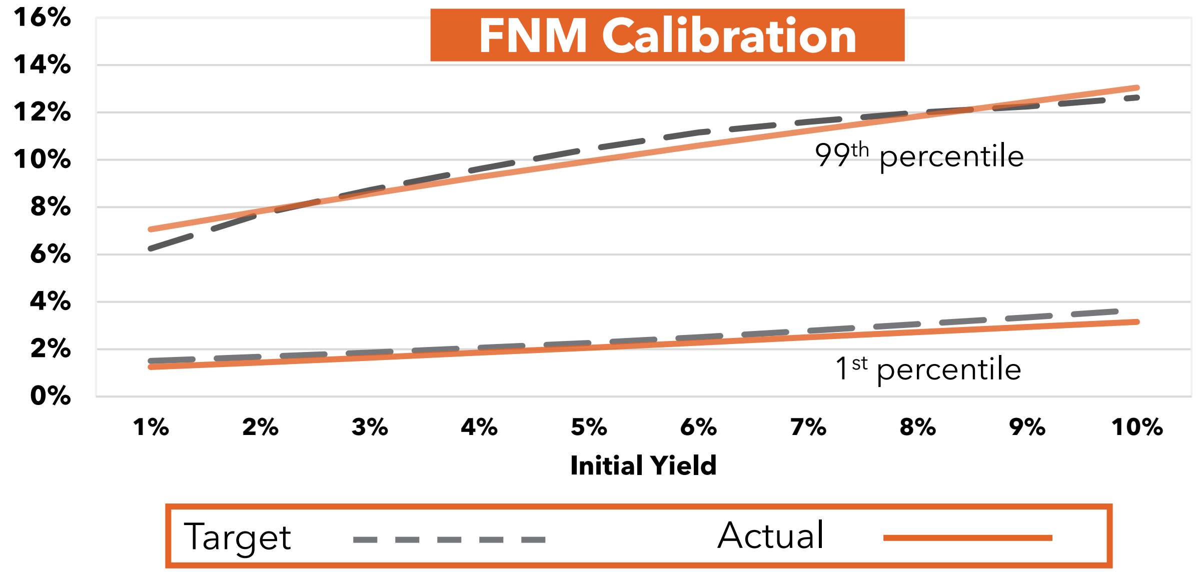
	A			B		
	90th Percentile of 10Y Geometric Average	**Criteria**	**Pass / Fail**	**95th Percentile of 30Y Geometric Average**	**Criteria**	**Pass / Fail**
Field Test 1A	1.37%*	1.45%	✓Pass	1.51%*	1.95%	✓Pass
NAIC Fall National Meeting Calibration	1.35%	1.45%	✓Pass	1.75%	1.95%	✓Pass

Both the 1A and Fall National Meeting Calibrations pass the objective T4 criteria calculated as of 12/31/20. Note that the 1A calibration was determined using more stringent 30-year geometric average targets.

*as of 12/31/21

Item	Category	Criteria
T5.	Low- and High-For-Long at Varying Starting Conditions	a) For each scenario, calculate the geometric average of the [20-year] UST yield over the first [10] and [30] years of the projection. b) Calculate the [1st] and [99th] percentiles of the distribution of geometric average rates (for both the 10 and 30-year horizons). c) Look up criteria based on the starting level of the 20-year UST yield (interpolate if necessary).

Geometric Average of 20Y UST over 30 years



	10-Year				30-Year			
	1st Percentile		99th Percentile		1st Percentile		99th Percentile	
Starting Yield of 20Y UST	*Target*	Actual	*Target*	Actual	*Target*	Actual	*Target*	Actual
1%	*0.94%*	0.68%	*3.43%*	4.65%	*1.50%*	1.24%	*6.25%*	7.06%
2%	*1.23%*	1.01%	*5.05%*	5.87%	*1.68%*	1.44%	*7.71%*	7.83%
3%	*1.62%*	1.37%	*6.55%*	7.02%	*1.86%*	1.64%	*8.72%*	**8.57%**
4%	*2.15%*	1.76%	*7.74%*	8.09%	*2.06%*	1.85%	*9.62%*	**9.28%**
5%	*2.66%*	2.18%	*8.87%*	9.12%	*2.26%*	2.06%	*10.46%*	**9.94%**
6%	*3.15%*	2.59%	*9.96%*	10.14%	*2.50%*	2.27%	*11.16%*	**10.60%**
7%	*3.63%*	3.03%	*11.03%*	11.14%	*2.78%*	2.50%	*11.61%*	**11.22%**
8%	*4.10%*	3.46%	*12.07%*	12.12%	*3.06%*	2.72%	*11.99%*	**11.83%**
9%	*4.64%*	3.92%	*13.08%*	13.08%	*3.34%*	2.94%	*12.25%*	12.45%
10%	*5.21%*	4.36%	*14.01%*	14.03%	*3.65%*	3.16%	*12.63%*	13.05%

The FNM calibration meets all of the 10-year geometric average low for long and high for long criteria for varying starting levels. However, there are some misses for the 3% to 8% starting environments on the high for long criteria. In order to meet all of these criteria, Conning could slow down the mean reversion speed or make other potential changes - leading to other tradeoffs.

Equity Model Acceptance Criteria

Equity Model Acceptance Criteria

Item	Category	Criteria
E1.	Low and High Accumulated Equity Returns	Use the former C3 Phase II equity model Calibration Criteria as a rough placeholder benchmark when evaluating equity scenarios.

Large Cap (S&P 500) C3 Phase II Calibration Criteria

Percentile	1 year	5 years	10 years	20 years
2.5%	0.78	0.72	0.79	
5.0%	0.84	0.81	0.94	1.51
10.0%	0.9	0.94	1.16	2.1
90.0%	1.28	2.17	3.63	9.02
95.0%	1.35	2.45	4.36	11.7
97.5%	1.42	2.72	5.12	

Equity Model Acceptance Criteria – Large Cap

Criteria are incomplete for Large Cap

a) GEMS links to Treasury
- Need to know starting level
- Need to know which calibration

b) GEMS model has stochastic volatility
- Need to know starting volatility level

Conning's approach

a) Set initial level to long-term values
- Treasury Yields based on long-term State Values
- Equity Volatility set to long-term value

b) Use proposed Treasury calibration: 25% Speed Reduction

Equity Model Acceptance Criteria – Large Cap

Targets

Percentiles	1-year	5-year	10-year	20-year
2.5%	0.78	0.72	0.79	
5%	0.84	0.81	0.94	1.51
10%	0.90	0.94	1.16	2.10
90%	1.28	2.17	3.63	9.02
95%	1.35	2.45	4.36	11.70
97.5%	1.42	2.72	5.12	

<= cumulative returns should be < this

<= cumulative returns should be > this

S&P – returns in USD: FNM Calibration as of 12/31/21

Percentiles	1-year	5-year	10-year	20-year
2.5%	**0.78**	0.72	0.74	
5%	0.83	0.80	0.89	1.19
10%	0.89	0.93	1.07	1.55
90%	**1.27**	**2.08**	**3.58**	10.01
95%	**1.32**	**2.31**	**4.24**	13.52
97.5%	**1.37**	**2.55**	**5.03**	

Red bold cells are those where the S&P cumulative return is above (below) the applicable target

Proposed Parameters	
Fixed Return	0.005137
Fixed Jump Intensity	3.595725
Initial Variance	0.0125
Alpha	0.005774
Beta	0.462695
Sigma	0.07468

Comparison to targets

- Slight miss on 1-Year 2.5%: 6 bps too high
- Bigger miss on the upside: GEMS Model produces negative skew in-line with history; Targets are based on model with positive skew

Equity Model Acceptance Criteria – Other Equity Indices

Proposed Targets are only for Large Cap

a) Need targets for the other 5 indices
b) Used 95%/105% Sharpe Ratio range that was utilized in original Field Test
   - Started with 3 native indices: Mid Cap, Small Cap and Aggressive US with same 70/30 weighting
   - Adjusted Means to get within the range

	Actual		Excess			
Investment	Mean	St Dev	Mean	St Dev	Sharpe Ratio	**Pass?**
1m Treasury	3.22%	3.46%	0.00%	0.00%		
Large Cap	8.46%	16.91%	5.09%	16.00%	31.82%	
Mid Cap	8.70%	18.18%	5.32%	17.27%	30.79%	TRUE
Small Cap	9.35%	19.94%	5.95%	19.00%	31.33%	TRUE
Aggressive US Equity	11.18%	24.88%	7.73%	23.85%	32.43%	TRUE
International Equity	8.87%	19.00%	5.51%	18.21%	30.24%	TRUE
Aggressive Foreign Equity	12.29%	29.13%	8.77%	27.81%	31.52%	TRUE

Attachment Eight
Life Actuarial (A) Task Force
11/29-11/30

Equity Model Acceptance Criteria – GWF Distribution

Original Field Test used GWF Targets

a) Included here for completeness even though they weren't targeted

	AIRG					Fall National Meeting Calibration			
	US	INT	SMALL	AGGR		Large	International	Small	Aggr US Equity
Mean	1258.66%	1441.19%	1823.15%	2572.98%		1287.33%	1416.16%	1527.09%	2397.33%
St Dev	1204.05%	1591.29%	2844.95%	5305.59%		2044.07%	2001.52%	2748.97%	5258.72%
Wealth CV	95.66%	110.42%	156.05%	206.20%		158.78%	141.33%	180.01%	219.36%
Percentile									
1%	122.79%	101.89%	62.77%	40.14%		89.45%	87.16%	66.44%	52.36%
5%	229.86%	210.67%	154.37%	104.14%		167.89%	159.67%	145.67%	126.83%
25%	523.88%	508.42%	467.63%	426.47%		400.92%	420.05%	407.99%	440.54%
50%	896.98%	952.52%	992.93%	1066.49%		741.18%	812.91%	809.11%	1051.94%
75%	1570.65%	1763.08%	2120.32%	2662.99%		1438.25%	1617.85%	1671.67%	2461.83%
95%	3480.18%	4304.80%	6124.85%	9576.36%		3923.31%	4533.45%	4953.90%	8409.61%
99%	5852.07%	7702.80%	11770.24%	24555.24%		8419.26%	9831.84%	11462.97%	20224.77%

Agenda Item 10

Hear an Update from the American Academy of Actuaries Economic Scenario Generator Subcommittee on Equity Acceptance Criteria

AMERICAN ACADEMY of ACTUARIES

November 22, 2023

Ms. Rachel Hemphill, Chair, Life Actuarial (A) Task Force (LATF)
Mr. Philip Barlow, Chair, Life Risk-Based Capital (E) Working Group (Life RBC)
Mr. Mike Yanacheak, Chair, Generator of Economic Scenarios (E/A) Subgroup (GOES Subgroup)
National Association of Insurance Commissioners (NAIC)

Dear Ms. Hemphill, Mr. Barlow, and Mr. Yanacheak,

The Economic Scenario Generator Subcommittee (ESGS) of the American Academy of Actuaries'[1] (the Academy) appreciates the ongoing opportunity to present proposals for stylized facts and acceptance criteria related to LATF's Generator of Economic Scenarios (GOES) project. These comments offer feedback and potential criteria related to Gross Wealth Factors (GWF) for equity index returns.

Executive Summary

In response to LATF's request, the ESGS has analyzed the forward-looking GWFs for S&P 500 total returns were previously developed by the Academy and proposed to LATF in 2005 and, using more recent historical data on S&P 500 total returns, developed the following updates.

In the course of this work, the ESGS found that equity returns can be effectively modeled using a constant mean, with the excess of that constant mean over the long-term target for interest rates representing a premium in exchange for the risk of holding an equity index over Treasuries. This is an equity risk premium (ERP. It is important to understand the differences and advantages of using a constant mean equity return rather than a constant mean equity risk premium to model returns on equity indices. While using a constant mean ERP is suitable for some purposes, we did not find it suitable for purposes with longer horizons that rely on realistic tail distributions, such as statutory reserves and capital for long duration, market-sensitive life and annuity products.

The methodology used to update GWFs and the updated GWFs themselves is addressed below. While the ESGS did not achieve consensus around a single target level for long-term equity returns, we offer three different sets of updated GWFs for consideration: one where the mean equity return was left unconstrained (straight fit to historical data resulting in a mean of 11.64%); one where the mean equity returns was constrained to 8.75% (the level used for the 2005 GWFs); and one where the mean equity return was constrained to 10.00% (roughly in the middle of the other two).

[1] The American Academy of Actuaries is a 19,500-member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues. The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.

The relationship between equity returns and interest rates

After reviewing historical data and economic theory, the ESGS believes a suitable method for modeling equity returns is to use a constant mean. Such a method produces a moderately inverse relationship between equity returns and interest rates, which is supported by history. It results in a higher ERP when simulated interest rates are low and vice versa, as well as an average ERP over the long term that is the positive excess of the constant mean for equity returns over the long-term target for interest rates. We have included theoretical and empirical rationale for the above statement, discussing the differences and advantages of using a method that employs a constant mean equity return rather than a constant mean ERP.

In 2005, the Academy proposed a set of GWF for validating equity scenarios used to determine capital for variable annuity products under C-3 Phase II. The ESGS has reviewed and updated those factors considering subsequent equity market performance and our previously proposed stylized facts for equity returns (see Appendix 1), which feature the concepts that:

- Cumulative equity returns tend to exceed the compounded risk-free rate (positive observed equity risk premium) over long time horizons, but over short time horizons the observed equity risk premium fluctuates due to several factors and can be negative." (Stylized Fact #2)[2]
- "Cumulative equity returns over long time horizons are not materially impacted by initial market conditions." (Stylized Fact #4)[3]

As discussed in academic, investment industry, and other papers, economic theory and empirical data suggest an ERP which is not constant over time[4] but instead one which varies over time in a countercyclical (inverse) manner,[5] expanding and contracting with business cycles. It may be higher in a recession[6] and when interest rates are low[7], and it may depend on other factors such as investor risk aversion and consumption preferences, inflation, quality and availability of earnings information, and government and monetary policy[8] (see figure 1).

Note that while the capital asset pricing model (CAPM) assumes a constant ERP, leading ERP research[5,6,7] indicates an ERP that changes over time. As a practical matter, market users of CAPM regularly update

---

[2] For example, the Dividend Discount Model implies that equity valuations (i.e., present value of projected earnings) decrease when risk-free interest rates increase (i.e., an inverse relationship due to increased discounting when rates are high and vice versa). However, this inverse relationship can be magnified if the increase in interest rates is due to Fed policy to slow the economy, which will presumably also have a negative effect on projected earnings).

[3] This stylized fact is implicitly in the current AIRG as well as the 2005 C-3 Phase II GWF calibration standard.

[4] Expansionism: The Impact of the Fed's Monetary Regime on the Equity Risk Premium (Global Financial Data White Paper)

[5] The Equity Risk Premium: A Contextual Literature Review (CFA Institute Research Institute)

[6] Kroll Recommended U.S. Equity Risk Premium and Corresponding Risk-Free Rates to be Used in Computing Cost of Capital: January 2008 – Present

[7] The Equity Risk Premium: A Review of Models (Federal Reserve Bank of New York Staff Report No. 714)

[8] Equity Risk Premiums (ERP): Determinants, Estimation, and Implications—The 2021 Edition, with annual updates also available at Professor Damodaran's website

ERPs as a key input to the model[3], effectively treating the ERP as constant for the time horizon of interest. As such, the ERP is kept constant for the duration of the current period's cash flow projection and updated from one period to the next to reflect changes in the ERP over time.

Figure 1: Factors considered in Kroll's December 9, 2020, U.S. ERP Recommendation[9]
(relative change from March to November 2020)

FACTOR	CHANGE	EFFECT ON ERP
U.S. Equity Markets	▲	▼
Implied Equity Volatility	▼	▼
Corporate Spreads	▼	▼
Economic Policy Uncertainty (EPU) and Equity Uncertainty Indices	▼	▼
Historical Real GDP Growth and Forecasts	▼	▲
Unemployment Environment	▼	▲
Consumer Confidence	▼	▲
Business Confidence	▲	▼
Sovereign Credit Ratings	◀▶	◀▶
Damodaran Implied ERP Model	▼	▼
Default Spread Model	▼	▼

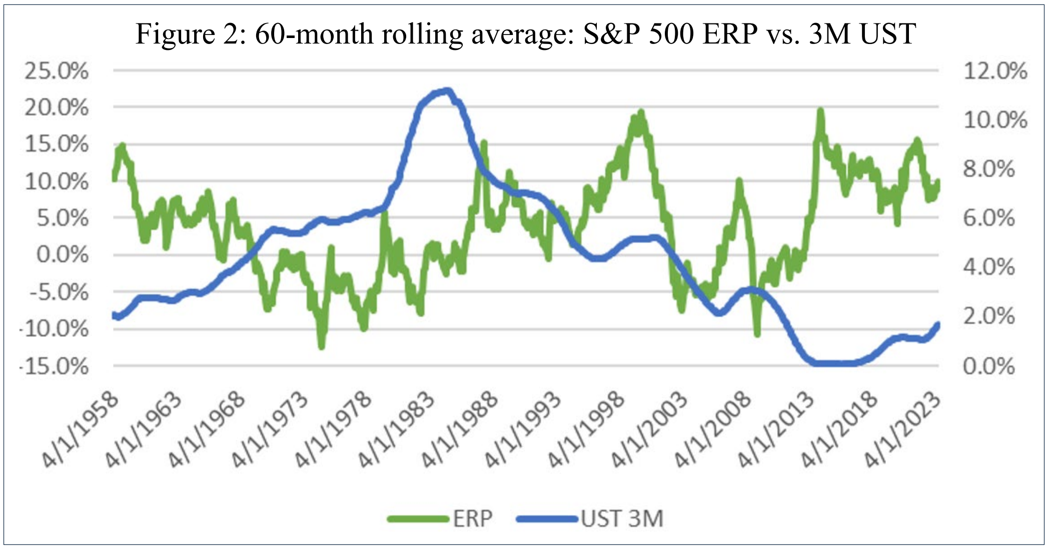
Under quantitative equity valuation models such as the discounted cash flow (DCF) or dividend discount model (DDM), the price of a stock is the present value of future dividends. All else being equal, a stock's price is expected to fall as interest rates increase. In addition, Federal Reserve monetary policy, which increases rates to slow the economy in the near-term while maintaining stability in the longer-term, may also reduce projected earnings and dividend growth (e.g., lower revenues, higher capital / borrowing costs), further adversely affecting stock prices in the near-term.

Changes in investor risk appetites as interest rates increase may also decrease the relative attractiveness of equities, such as the There Is No Alternative (TINA) effect widely cited in industry publications. TINA describes the preference for stocks and other risky assets during periods of near-zero interest rates, compared to a preference for Treasuries when risk-free rates on Treasuries are high and close to the risky returns on stocks.[8]

In addition to economic theory, it's also useful to consider empirical data and see how well a single-regime[9] or regime-switching model can be fit to empirical data. Exploratory graphical analysis of empirical data (see Figure 2) suggests an ERP that tends to be higher when interest rates are low and lower when interest rates

[9] Duff & Phelps Recommended U.S. Equity Risk Premium Decreased as COVID-19 Impact Recedes

are high (vs. randomly dispersed around a constant level).



Statistical tests based on simple regression (i.e., a single regime) models, such as the 2022 Blitz paper[10], also tend to reject the hypothesis that a higher risk-free return implies higher total average stock returns. Rather, they show expected stock returns appear to be inversely related to the level of the risk-free return.

In addition to using single regime models, historical data may also be analyzed using more complex models that allow for the possibility of three types of regimes:

- Expected equity returns increase 1-for-1 as interest rates increase (expected ERP is constant)
- Expected equity returns remain constant as interest rates increase (expected ERP varies inversely with rates)
- Expected equity returns decrease 1-for-1 as interest rates increase (expected ERP has a strongly inverse relationship with rates).

In all three regimes, random variations result in a distribution of ERPs and returns around the expected values.

The ESGS's maximum likelihood estimation of such a model using historical S&P 500 total returns suggests the S&P 500 tends to move in the same direction as interest rates (a constant ERP relationship where equity returns increase as interest rates increase) about 10% of the time and in the opposite direction as interest rates (a strongly inverse ERP relationship where equity returns decrease as interest rates increase) about 30% of the time. However, we also found that 60% of the time equity returns tend not to move with interest rates but stay centered around a *constant* mean equity return, albeit a mean that is higher than average interest rate levels. This reflects an ERP that varies moderately inversely with interest rates, especially when interest rates are less volatile (see Figures 2 and 3). As with the 2022 Blitz paper, these findings suggest rejecting a constant mean ERP relationship where equity returns increase as interest rates increase and vice versa.

Hypothesis testing of different relationships between interest rates and equity returns based on monthly

[10] Expected Stock Returns When Interest Rates Are Low (Blitz, 2022)

historical data from April 1953 to December 2020 indicates neither a constant mean ERP relationship nor a constant mean equity return relationship can be rejected in the middle eight deciles of the distribution of risk-free interest rates (see Table 1). However, at the top and bottom deciles, i.e., 3M UST yield below 15bps and above 8.33%, the constant mean ERP relationship ought to be rejected with p-values at a 0.3% and 1.6% significance level respectively. Note that this analysis accounts for the underlying volatility of the equity returns. Alternatively, unconstrained regression of the ERP relationship in the data produces a much higher significance level of 8.2% and 35.7% in the bottom and top deciles, pointing to an inverse relationship between risk free rates and equity returns, a similar conclusion in the 2022 Blitz paper. Given the purpose at hand, modeling cash flows for the determination of statutory capital, it is important to reflect a plausible relationship between interest rates and equity returns in low and high tails which are likely to drive the total asset requirement (TAR) upon which capital is based (see Appendix 4 for additional details on this analysis).

Figure 3: Historical frequency of degree of relationship between monthly S&P price index returns and short-term interest rates observed from 1953.04 to 2020.12

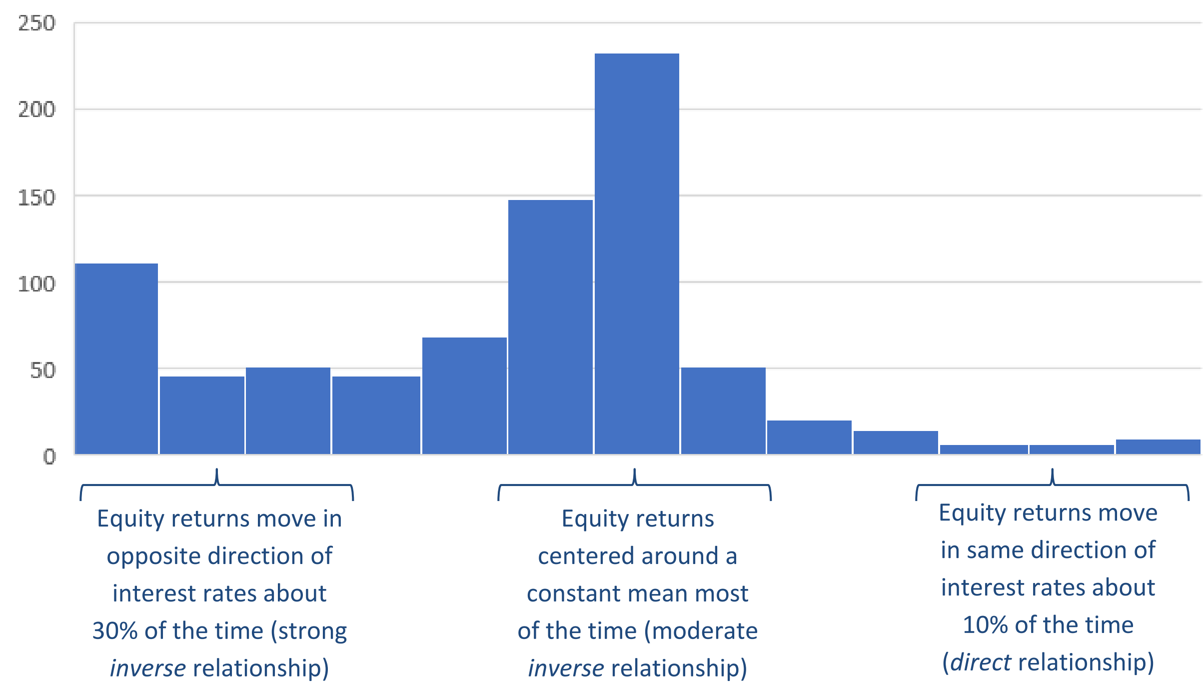


Table 1: p-values from constrained and unconstrained regression fits to monthly historical S&P 500 data from 1953.04 to 2020.12.				
Deciles of 3M UST Yield	**Average 3M UST Yield**	**Regression constrained to constant mean ERP**	**Regression constrained to constant mean Equity Return**	**Regression unconstrained**
0-10	0.1%	0.3%	2.4%	8.2%
10-20	0.8%	29.5%	50.0%	34.6%
20-30	1.9%	40.1%	42.5%	30.5%
30-40	2.8%	29.6%	20.8%	15.5%
40-50	3.7%	20.5%	16.1%	13.5%
50-60	4.6%	43.9%	46.2%	48.1%
60-70	5.2%	21.6%	17.1%	14.1%
70-80	6.0%	35.9%	45.6%	46.8%
80-90	7.5%	39.9%	41.7%	28.8%
90-100	10.7%	1.6%	12.6%	35.7%
Mean ERP =		***Constant***	***Constant - (1.00 × Risk Free Rate)***	***Constant - (1.74 × Risk Free Rate)***
Mean Equity Return =		***Risk Free Rate + Constant***	***Constant***	***Constant - (0.74 × Risk Free Rate)***

After reviewing historical data and economic theory, a suitable way to model an ERP that exhibits a moderately inverse relationship with interest rates uses a constant mean for simulated equity returns, resulting in a higher ERP when simulated interest rates are low and vice versa. It also uses an average ERP over the long term that is the positive excess of the long-term target for equity returns over the long-term target for interest rates. Under such an approach, the mean equity return not only stays constant throughout the projection, but also generally stays constant from one period to the next.[11] However, initial ERP levels would change from one period to the next as interest rates move. Under such an approach, steady state equity return targets should align with steady state interest rate levels, reflecting an appropriate long-term equity risk premium and sufficient dispersion of equity returns to allow for appropriate joint probabilities of low/high equity returns and low/high interest rates.

While some applications model the ERP as varying around a constant expected mean, such as GEMS where equity returns increase with interest rates and vice versa, this modeling simplification is less aligned with economic theory and historical data. It is more commonly and appropriately used for shorter-term projections of real-world scenarios, particularly when the computationally efficient generation of "nested" risk neutral scenarios is a priority (Solvency II), strategic asset allocation where the focus is on the middle of the distribution, or non-equity sensitive liabilities (short-duration property and casualty liabilities). It is not suitable for purposes with longer horizons that rely on realistic tail distributions, such as statutory reserves and capital for long duration or market-sensitive life and annuity products. An increasing return / constant expected ERP approach is more likely to distort reserve and capital levels by inadequately reflecting historically observed tail dynamics. It will also cause a distribution of equity returns that shifts from one period to the next in an unintuitive manner as interest rates move, potentially resulting in artificial and unintuitive volatility and/or distorting hedging and risk management decisions and costs. Workarounds that maintain this relationship in the ESG

[11] For example, major and long-lasting changes in fed policy or market dynamics may warrant a review and potential change to long term targets for equity returns.

model, but adjust the model's parameters whenever initial conditions change, may partially mitigate some of these impacts. It may also introduce other unintuitive dynamics, adding additional unnecessary operational and governance-type complexities.

In summary, the 4th stylized fact proposed for equity returns supports continued use of a static set of GWFs as criteria for equity returns, such as the set of GWFs proposed by the Academy in 2005 for validating equity scenarios used to determine capital for variable annuity products under C-3 Phase II. Although such criteria generally would not be expected to change over time, it is prudent to review them periodically, consider interim equity market experience, and determine if any deviations are material enough to warrant a change to the criteria. Under a constant mean equity return approach, there is still a relationship between interest rates and equity returns, but it is implicit in the long-term targets set for equity returns and interest rates. It is important to consider the reasonableness of the implied long-term ERP when reviewing and resetting those long-term targets.

Methodology and updates to criteria for equity GWFs

As in 2005, our updated analysis of equity GWFs was based on taking several model forms, consistent with our equity stylized facts and commonly used to model equity returns, and fitting them to monthly S&P 500 total returns from March 1957 through December 2022, including twenty additional years of relevant and credible data than the original 2005 analysis. This parallels the 2005 analysis done by the Academy to develop the first C-3 Phase II GWF criteria for equity returns. In that work, several similar equity model forms were used to inform what realistic equity return behavior could look like given the lack of historical data and limited availability of non-overlapping multi-year periods in the historical data set[12]. Appendix 2 details the model forms considered, fitted, and used to inform realistic equity market behavior and develop criteria for GWFs, such as Heston, SLV2, RSLN2, and Heston with Jump[13]. Tables at the end of Appendix 2 show annualized equity returns across the reference models for 1-year and 20-year horizons, compared to similar statistics for the AIRG and the GEMS Field Test (FT) #1a scenario sets, as well as to the distribution of rolling 1-year returns observed in history.

We initially fit reference models to history without constraining the mean equity return. Table 2 below shows the resulting unconstrained means for the various reference models are about 11.50%, or 2.75% higher than the constrained mean of 8.75% used in the Academy's 2005 analysis.

---

[12] As in 2005, the data period begins in the mid-1950s, when the Fed, securities regulation, and the S&P 500 began to resemble their modern-day counterparts (before the mid-1950s the S&P 500 was less of a broad large cap index, and securities regulation and fed/monetary policy functioned in a markedly different manner, e.g., lack of trade limits, significantly less national debt as a percentage of GDP). Note that reference models appropriately calibrated to this data period, which doesn't include the 1930s great depression, are still able to simulate large drops and sustained losses that are even worse than those experienced during the great depression.

[13] Heston with Jump is similar to GEMS, but without the constant expected ERP (increasing return) relationship.

Table 2:	30-year mean and median from unconstrained fits to 1957.03 to 2022.12	
Model	*Unconstrained mean*	*Unconstrained median*
*Heston**	11.47%	10.39%
SLV1	10.34%	9.13%
*SLV2**	11.37%	10.53%
SLV3	11.61%	10.81%
*RSLN2**	11.94%	10.48%
RSDD2	11.22%	10.34%
LN	11.68%	10.43%
*Jump**	11.80%	10.79%
AIRG†	8.81%	7.59%

** GWF tables below are based on the least binding percentile across these four selected reference models. The average mean across those four selected reference models is 11.64%.*

† The AIRG's constrained mean and median are provided for reference.

While the ESGS did not achieve consensus around a single target level for long-term equity returns, the group recommends the use of a *best-estimate* target, with appropriately disperse tails, rather than a target that is purposely set low or high. Since some products and risk management strategies perform worse when equity returns are low while others perform better, our recommendation is a distribution for stochastically modeled risk factors that is best estimate, with the CTE level used as the source of prescribed conservatism in statutory reserve and capital calculations. Although a long-term, best-estimate mean equity return target of 11.64% based on a pure fit to history may be too high, some felt a target of 8.75% (a 2.89% haircut to the historical fit of 11.64%) may be too low, indicating something in the middle may be more appropriate[14]. Regardless of the long-term target used for equity returns, it is important to periodically review the long term mean ERP implied by the long-term targets used for equity returns and interest rates for reasonableness. For purposes of this illustration, Tables 3.1, 3.2, and 3.3 below show the least binding GWF[15] across the four reference models using a constrained mean of 8.75%, a constrained mean of 10.00%, and an unconstrained mean of 11.64% (see Appendix 3 for a graphical representation of Table 3.2). We found that the least binding reference model was largely the same regardless of whether the mean was left unconstrained or constrained to 8.75% or 10%.

[14] Best-estimate forward-looking views on equity returns are often based on more than just historically observed equity returns, and also incorporate things such as expectations around real GDP growth, inflation, fed/monetary policy, industry surveys, etc.

[15] The least binding GWF across the four selected reference models is the maximum for low-return (left) tail percentiles and the minimum for high-return (right) tail percentiles (means and medians are shown for informational purposes and are *averages* across reference models).

Table 3.1:	Updated GWFs using a constrained mean return of 8.75%					Least binding reference model					
	Horizon (years)					*Horizon (years)*					
Percentile	*1*	*5*	*10*	*20*	*30*	*50*	*1*	*5*	*10*	*20*	*30*	*50*
Min	0.48	0.28	0.33	0.32	0.56	0.85	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
1%	0.71	0.64	0.71	0.99	1.55	4.15	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
5%	0.83	0.84	1.02	1.62	2.73	8.63	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
10%	0.89	0.98	1.22	2.10	3.74	12.78	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
15%	0.93	1.07	1.38	2.46	4.55	16.49	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
30%	1.02	1.28	1.76	3.41	6.84	27.56	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
Median	1.09	1.48	2.15	4.47	9.23	39.98	n/a	n/a	n/a	n/a	n/a	n/a
70%	1.17	1.73	2.70	6.14	13.50	62.71	Heston	SLV2	RSLN2	RSLN2	RSLN2	RSLN2
85%	1.24	1.97	3.27	8.41	20.39	112.78	SLV2	SLV2	SLV2	SLV2	SLV2	RSLN2
90%	1.28	2.09	3.58	9.59	23.93	142.63	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
95%	1.33	2.28	4.08	11.43	30.68	195.72	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
99%	1.42	2.67	5.10	15.83	45.17	333.02	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
Max	1.67	3.75	8.01	29.20	99.48	1019.62	SLV2	Jump	SLV2	SLV2	SLV2	SLV2
Mean	1.09	1.52	2.31	5.38	12.38	65.77	n/a	n/a	n/a	n/a	n/a	n/a

Table 3.2:	Updated GWFs using a constrained mean return of 10.00%					Least binding reference model					
	Horizon (years)					*Horizon (years)*					
Percentile	*1*	*5*	*10*	*20*	*30*	*50*	*1*	*5*	*10*	*20*	*30*	*50*
Min	0.49	0.29	0.36	0.40	0.80	1.51	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
1%	0.72	0.68	0.79	1.25	2.18	7.36	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
5%	0.84	0.89	1.15	2.03	3.84	15.27	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
10%	0.90	1.04	1.37	2.64	5.27	22.62	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
15%	0.94	1.14	1.55	3.09	6.41	29.20	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
30%	1.03	1.36	1.97	4.29	9.64	48.80	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
Median	1.11	1.57	2.41	5.62	13.00	70.81	n/a	n/a	n/a	n/a	n/a	n/a
70%	1.18	1.83	3.03	7.72	19.03	111.04	Heston	SLV2	RSLN2	RSLN2	RSLN2	RSLN2
85%	1.26	2.08	3.67	10.57	28.73	199.71	SLV2	SLV2	SLV2	SLV2	SLV2	RSLN2
90%	1.29	2.21	4.02	12.05	33.72	252.57	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
95%	1.34	2.42	4.57	14.37	43.23	346.58	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
99%	1.44	2.83	5.71	19.90	63.64	589.72	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
Max	1.69	3.97	8.98	36.70	140.17	1805.56	SLV2	Jump	SLV2	SLV2	SLV2	SLV2
Mean	1.10	1.61	2.59	6.76	17.45	116.46	n/a	n/a	n/a	n/a	n/a	n/a

Table 3.3:	Update GWFs using an unconstrained mean return of 11.64%					Least binding reference model					
	Horizon (years)					*Horizon (years)*					
Percentile	*1*	*5*	*10*	*20*	*30*	*50*	*1*	*5*	*10*	*20*	*30*	*50*
Min	0.49	0.31	0.41	0.51	1.15	2.80	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
1%	0.73	0.72	0.90	1.60	3.15	13.63	Heston	SLV2	SLV2	SLV2	SLV2	SLV2
5%	0.85	0.95	1.30	2.60	5.56	28.30	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
10%	0.92	1.11	1.55	3.37	7.63	41.92	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
15%	0.95	1.21	1.75	3.96	9.28	54.11	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
30%	1.04	1.44	2.23	5.52	13.96	90.53	SLV2	SLV2	SLV2	Jump	SLV2	Jump
Median	1.12	1.69	2.79	7.56	20.27	148.11	n/a	n/a	n/a	n/a	n/a	n/a
70%	1.20	1.95	3.43	10.18	29.42	238.65	Heston	SLV2	SLV2	SLV2	SLV2	Heston
85%	1.27	2.22	4.15	13.53	41.60	377.39	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
90%	1.31	2.35	4.55	15.42	48.82	468.01	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
95%	1.36	2.57	5.17	18.39	62.60	642.20	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
99%	1.46	3.01	6.46	25.47	92.14	1092.72	SLV2	SLV2	SLV2	SLV2	SLV2	SLV2
Max	1.71	4.31	10.16	46.96	202.94	3345.63	SLV2	Jump	SLV2	SLV2	SLV2	SLV2
Mean	1.12	1.73	3.01	9.10	27.28	245.63	n/a	n/a	n/a	n/a	n/a	n/a

Tables 4.0, 4.1, 4.2, and 4.3 compare the updated GWFs to the prior set of C-3 Phase II GWFs[16]. Table 4.1 shows that when using the same constrained mean equity return of 8.75%, considering 20 years of additional data does not result in significantly different GWFs[17]. However, the updated GWFs in the above tables are now available for longer horizons and additional percentiles than the prior set of C-3 Phase II GWFs, given the expanded use of economic scenarios to for reserves, longer duration life insurance liabilities (e.g., VM-20 liabilities tend to be of longer duration than VM-21 liabilities), and the adoption of more extreme tail CTE levels in capital requirements (e.g., CTE98 vs. CTE90).

[16] The 2005 C-3 Phase II GWF calibration standard was based on an SLV model fit to S&P 500 total returns from 1955.12 to 2003.12, with the mean constrained to 8.75%, and further adjustments made to accommodate a wider range of reasonably fit ESG model forms (i.e., LN, RSLN). Note that the 2005 standard had criteria for percentiles of 2.5% and 97.5%, whereas updated GWFs have criteria for percentiles further in the tail at 1% and 99%.

[17] This is consistent with the last 20 years having both the 2008-2009 financial crisis as well as periods of very favorable returns. Note that the Academy's 2012 analysis suggested that the 2005 set of C-3 Phase II GWF criteria allowed the volatility and returns seen during the 2008-2009 financial crisis and did not warrant an update at that time.

Tables 4.0, 4.1, 4.2, and 4.3: Updated gross wealth factors compared to prior C3P2 calibration standard

Table 4.0:	Prior C3P2 GWF calibration standard (mean return constrained to 8.75%)			
	Horizon (years)			
Percentile	*1*	*5*	*10*	*20*
2.5%	0.78	0.72	0.79	n/a
5%	0.84	0.81	0.94	1.51
10%	0.90	0.94	1.16	2.10
90%	1.28	2.17	3.63	9.02
95%	1.35	2.45	4.36	11.70
97.5%	1.42	2.72	5.12	n/a

Note: The tables below only show the subset of updated GWFs that can be compared to the prior C3P2 calibration standard.

Table 4.1:	Proposed update to GWFs using constrained mean return of 8.75%				As a percentage of prior C3P2 calibration standard			
	Horizon (years)				*Horizon (years)*			
Percentile	*1*	*5*	*10*	*20*	*1*	*5*	*10*	*20*
5%	0.83	0.84	1.02	1.62	99%	104%	109%	107%
10%	0.89	0.98	1.22	2.10	99%	105%	105%	100%
90%	1.28	2.09	3.58	9.59	100%	96%	99%	106%
95%	1.33	2.28	4.08	11.43	98%	93%	94%	98%

Table 4.2:	Proposed update to GWFs using constrained mean return of 10.00%				As a percentage of prior C3P2 calibration standard			
	Horizon (years)				*Horizon (years)*			
Percentile	*1*	*5*	*10*	*20*	*1*	*5*	*10*	*20*
5%	0.84	0.89	1.15	2.03	100%	110%	122%	135%
10%	0.90	1.04	1.37	2.64	100%	111%	118%	126%
90%	1.29	2.21	4.02	12.05	101%	102%	111%	134%
95%	1.34	2.42	4.57	14.37	100%	99%	105%	123%

Table 4.3:	Proposed update to GWFs using unconstrained mean return (11.64%)				As a percentage of prior C3P2 calibration standard			
	Horizon (years)				*Horizon (years)*			
Percentile	*1*	*5*	*10*	*20*	*1*	*5*	*10*	*20*
5%	0.85	0.95	1.30	2.60	101%	117%	138%	172%
10%	0.92	1.11	1.55	3.37	102%	118%	134%	161%
90%	1.31	2.35	4.55	15.42	102%	108%	125%	171%
95%	1.36	2.57	5.17	18.39	101%	105%	119%	157%

Depending on the desired long-term mean equity return target, tables 3.1, 3.2, or 3.3 can be considered as a potential update to the 2005 set of C-3 Phase II GWF criteria[18]. Such criteria are applied to a scenario set by checking if the corresponding percentiles from the scenario set are more extreme than the criteria. Although such a one-way check helps ensure the distribution of GWFs is plausibly more extreme than history, when applying the criteria, it may also be useful to note the magnitude of the differences between criteria and the corresponding percentiles from the scenarios set. Differences that are too large may indicate a distribution of GWFs that is unreasonably more extreme than history.

In summary, after reviewing economic theory and a data period that includes 20 additional years of experience, the Academy supports a modeled ERP that is implicit in the long-term best-estimate targets (i.e., means) used for interest rates and long-term equity returns, resulting in a distribution of equity returns that does not change from period to period, instead changing returns only when long-term targets are revised. GWFs reflecting the additional 20 years of experience (more extreme percentiles of 1% and 99%, and longer horizons of 30 and 50 years) have been illustrated using a constrained mean of 8.75%, a constrained mean of 10.00%, and an unconstrained mean of 11.64%.

The Academy looks forward to our continued collaboration with regulators on the GWFs illustrated, including reasonable "best-estimate" ranges for long-term targets on interest rates, equity returns, and ERP. We look forward to providing additional comments on the overall application of a complete set of stylized facts and acceptance criteria, as well as proposal for "quadrant" criteria related to the tails of the joint distribution of equity returns and interest rates, such as the prevalence of low interest rates and low equity returns, and high interest rates and low equity returns[19]. Please direct any questions to Amanda Barry-Moilanen, life policy analyst at barrymoilanen@academy.org.

Sincerely,

Jason Kehrberg
Chair, Economic Scenario Generator Subcommittee

[18] The Min, Max, Median and Mean, as well as the Least binding reference model, are shown for reference and not specifically for inclusion in an updated set of GWF criteria for equity returns. If a smaller set of GWF criteria is desired, the Academy would recommend keeping the more extreme percentiles of 1%, 5%, 95%, and 99%, along with the somewhat less extreme percentiles of 15% and 85% (i.e., dropping percentiles of 10%, 30%, 70%, and 90%). Note that these updated GWF criteria were developed for use on sets of 10,000 scenarios but could be considered for use on sets with fewer scenarios. Also note that these updated GWF criteria have been expressed in terms of gross wealth factors ("GWFs") but can also be expressed in terms of geometric average returns ("GAVGs") by using the formula: *GAVG = GWF^(1/horizon) - 1*, where *horizon* is in years.

[19] LATF has previously requested the ESGS develop a proposal for such "quadrant" criteria.

Appendix 1—Stylized facts for Equity Returns (presented by Academy to LATF on 9/29/22)

1. Equity indices (indeed, all asset classes) tend to exhibit **consistent risk/reward relationships** over **long** time horizons.
2. Cumulative equity returns tend to exceed the compounded risk-free rate (positive observed **equity risk premium**) over long time horizons, but over short time horizons the observed equity risk premium fluctuates due to several factors and can be negative.
3. Equities **fluctuate between bull and bear markets** (bubbles tend to burst)—markets can experience significant losses but eventually tend to **move back into positive territory** (negative cumulative equity returns become less likely over longer time horizons).
4. Cumulative equity returns ***over long time horizons*** **are not materially impacted by initial market conditions**.
5. The **volatility of equity returns varies over time but has a strong tendency to revert to normative levels**. This allows for both extreme gains and extreme losses over short time periods (i.e., the distribution has fat tails, or **positive kurtosis**). Furthermore, the **volatility of equity returns is higher in bear markets**. This increases the probability of extreme losses relative to extreme gains (i.e., the distribution has a longer left tail, or **negative skewness**).
6. Equity markets contain **pathwise dynamics** over long time horizons that aren't present in the distribution of single-period returns. Future equity scenarios should have reasonable distributions of cumulative equity returns over long time horizons (e.g., 10, 20, 30 years), especially since these distributions are key to the performance of long-duration life and annuity products.
7. Future equity scenarios should include events that are **plausibly more extreme than history**.
8. Equity returns have both a **price and dividend component**, and they behave differently—dividend returns tend to be more stable than price returns.
9. Returns between different equity indices **are generally positively correlated** over **long** time horizons. This correlation may increase sharply in bear markets, but it tends to revert to normative levels in a short period of time.

Appendix 2—Detail on reference models considered for equity returns

A. Heston with Jumps (“Jump”)[20]

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$
- Stochastic equity variance follows Heston/CIR:

(a) $$v_t = \max\left[\tau^2(1-\zeta) + v_{t-1}\zeta + \sigma\sqrt{\tau^2/_{2\varphi}(1-\zeta)^2 + {}^{v_{t-1}}/_{\varphi}(\zeta-\zeta^2)}\, Z_t^v,\ v_{min}\right]$$

where:

- τ^2, φ *and* σ are the steady state target, reversion speed, and monthly diffusion coefficient of the monthly variance process
- $\zeta = e^{-\varphi}$
- $Z_t^v \sim N(0,1)$
- v_{min} is a suitably small floor to ensure stability of the variance process
- v_0 is set to $v_0 = initial\ volatility^2$

- Periodic log-return follows:

(b) $$lr_t = (A + (C - .5)v_{t-1})\,\Delta_t + \sqrt{v_{t-1}\,\Delta_t}\, Z_t^{lr} - \lambda_t m + N_t\mu_j + \sigma_j\sqrt{N_t} Z_t^j$$

where:

- v_{t-1} is the beginning-of-the-period Heston variance defined above
- $Z_t^{lr} \sim N(0,1)$, where $\langle Z_t^{lr}, Z_t^v\rangle = \rho$ is the correlation parameter between variance and log-return process
- $\lambda_t = v_{t-1}\,\lambda_1\,\Delta_t$ is the intensity of Poisson counting process tied to the beginning-of-period Heston variance v_{t-1}
- $N_t \sim Poisson(\lambda_t)$ is a Poisson random variable that is un-correlated with Heston variance and the core return process
- $m = e^{\mu_j + .5\sigma_j^2} - 1$, with μ_j and σ_j as mean and volatility of the variable jump γ
- Z_t^j is a standard normal variate that drives the log-normal jump diffusion of γ

[20] Note that GEMS employs a Heston with Jumps model for equity returns.

B. Heston

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$
- Stochastic equity variance follows Heston/CIR:

(a) $$v_t = \max\left[\, \tau^2(1-\zeta) + v_{t-1}\zeta + \sigma\sqrt{\tau^2/_{2\varphi}\,(1-\zeta)^2 + {}^{v_{t-1}}/_{\varphi}\,(\zeta-\zeta^2)}\; Z_t^v,\ v_{min}\right]$$

where:

- τ^2, φ *and* σ are the steady-state target, monthly reversion speed, and diffusion coefficient of the monthly variance process
- $\zeta = e^{-\varphi}$
- $Z_t^v \sim N(0,1)$
- v_{min} is a suitably small floor to ensure stability of the variance process
- v_0 is set to $v_0 = initial\ volatility^2$

- Periodic log-return follows:

(b) $$lr_t = (\mu_0 - .5v_{t-1})\,\Delta_t + \sqrt{v_{t-1}\,\Delta_t}\,Z_t^{lr}$$

where:

- v_{t-1} is the beginning-of-the-period Heston variance defined above
- $Z_t^{lr} \sim N(0,1)$, where $\langle Z_t^{lr}, Z_t^v \rangle = \rho$ is the correlation parameter between variance and log-return process

C. Stochastic Log Volatility ("SLV")[21]

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$
- Stochastic equity log-volatility follows OU process:

  (a) $lv_t = \min\left[\varphi\tau + (1-\varphi)lv_{t-1},\ log(SoftMaxVol)\right] + \sigma\, Z_t^{lv}$

  where:

  - $\tau,\ \varphi\ and\ \sigma$ are the steady-state target, reversion speed, and monthly diffusion coefficient of the monthly log-volatility process
  - $Z_t^v \sim N(0,1)$
  - $vol_t = \max\left[\min\left(e^{lv_t}, MaxVol\right),\ MinVol\right]$
  - lv_0 is set to $lv_0 = log(initial\ volatility)$

- Periodic log-return follows:

  (b) $lr_t = (A + B\ vol_t + C\ vol_t^2)\ \Delta_t + vol_t\sqrt{\Delta_t}\ Z_t^{lr}$

  where:

  v_{t-1} is the beginning-of-the-period Heston variance defined above

  $Z_t^{lr} \sim N(0,1)$, where $\langle\ Z_t^{lr},\ Z_t^v \rangle = \rho$ is the correlation parameter between log-volatility and log-return process

[21] Note that the AIRG employs an SLV model for equity returns.

D. Lognormal ("LN")[22]

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$

  (a) $lr_{1,t} = \mu\Delta_t + \sigma\sqrt{\Delta_t}Z_t$

  where:

  - $Z_t \sim N(0,1)$

[22] The LN model was included in our analysis due to its simple nature and historical significance, however over longer time horizons it does not meet all our stylized facts for equity returns and so was not used to derive the updated tables of least binding GWFs. In particular, the LN model lacks negative skew, and doesn't exhibit excess kurtosis over longer time horizons.

E. Regime-Switching Lognormal Model for 2 regimes ("RSLN2")

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$
- Periodic log-return for two regimes follows:

(a) $lr_{1,t} = \mu_1 \Delta_t + \sigma_1 \sqrt{\Delta_t} Z_t$

(b) $lr_{2,t} = \mu_2 \Delta_t + \sigma_2 \sqrt{\Delta_t} Z_t$

where:

- $Z_t \sim N(0,1)$
- $p11, and\ p21$ are monthly transition probabilities indicating continuing in state 1, and migrating from state 2 to state 1 respectively.
- For each scenario, the initial state is initialized to 1 if the scenario-specific $U(0,1) < p_1 = {}^{p_{21}}/_{p_{21} + p_{12}}$, and to state 2 otherwise.
- Transition states are evolved using independent and identically distributed uniform variates compared to transition probabilities in the subsequent projection steps.

F. Regime-Switching Lognormal Model for 2 regimes with Draw Down (“RSDD2”)[23]

- Process specified for discrete monthly timestep, $\Delta_t = 1/12$
- Periodic log-return for two regimes follows:

(a) $lr_{1,t} = \mu_1 \Delta_t + \varphi_1 DD_t + \sigma_1 \sqrt{\Delta_t} Z_t$

(b) $lr_{2,t} = \mu_2 \Delta_t + \varphi_2 DD_t + \sigma_2 \sqrt{\Delta_t} Z_t$

where:

- $Z_t \sim N(0,1)$
- Draw Down at each projection point, t, is defined as $DD_t = min(0,\ DD_{t-1} + lr_{t-1})$ and is initialized with $DD_1 = 0$.
- $p11, and\ p21$ are monthly transition probabilities indicating continuing in state 1 and migrating from state 2 to state 1 respectively.
- For each scenario, the initial state is initialized to 1 if the scenario-specific $U(0,1) < p_1 = {p_{21}}/{p_{21} + p_{12}}$, and to state 2 otherwise.
- Transition states are evolved using independent and identically distributed uniform variates compared to transition probabilities in the subsequent projection steps.

[23] The RSDD2 model was included in our analysis because it met our stylized facts for equity returns, but it was not used to derive the updated tables of least binding GWFs. RSDD2 has theoretical and empirical support but it is sensitive to the data period used and risks understating extreme events like the great depression. Had RSDD2 been used to derive the updated tables of least binding GWFs, the distribution of GWFs would have been narrower, i.e., *less* constraining.

G. Parameters and sum of log likelihood for reference models fit using unconstrained Maximum Likelihood Estimation (MLE) based on S&P 500 monthly returns from 1957.03 to 2022.12

	Heston	Jump	SLV	AIRG
tau	0.14694	0.14242	0.13076	0.12515
phi	0.09317	0.08436	0.09871	0.35229
sigma	0.04130	0.03805	0.16559	0.32645
A	0.10844	0.10886	0.09904	0.05500
B	0.00000	0.00000	0.00000	0.56000
C	0.00000	0.13580	2.45530	-0.90000
correlation (skew)	-0.54794	-0.58593	-0.68936	-0.24880
initial vol	0.14467	0.14242	0.15010	0.14760
min vol	0.03000	0.03000	0.03000	0.03050
soft max vol	0.30000	0.30000	0.30000	0.30000
max vol	0.35000	0.35000	0.35000	0.79880
mu_jump	0	-0.14740	0	0
sigma_jump	0	0.07000	0	0
lambda_jump	0	2.51937	0	0
MLE: Sum of LL	***1,430***	***1,435***	***1,447***	***1,418***

	RSLN2	RSDD2	LN
p11	0.93540	0.94077	1.00000
p21	0.10313	0.17652	0.00000
mu1	0.16570	0.13209	0.09910
mu2	-0.00720	-0.15209	0.00000
sigma1	0.09901	0.10749	0.14835
sigma2	0.20042	0.21292	0.10000
phi1	0	-0.06935	0
phi2	0	-0.00317	0
MLE: Sum of LL	***1,413***	***1,421***	***1,418***

H. Log return statistics for reference models fit using unconstrained Maximum Likelihood Estimation (MLE) based on S&P 500 monthly returns from 1957.03 to 2022.12

Monthly Log Return Statistics (first 600 months, 10k scenarios)								
	Heston	**Jump**	**SLV**	**RSLN2**	**RSDD2**	**LN**	**AIRG**	**History**
mean	0.81%	0.84%	0.83%	0.83%	0.83%	0.83%	0.61%	0.83%
st. dev.	4.25%	4.26%	4.28%	4.29%	4.29%	4.28%	4.36%	4.28%
skew	-0.03	-0.32	-0.69	-0.33	-0.56	0.00	-0.67	-0.67
kurt	4.25	5.79	5.47	4.39	4.63	3.00	7.00	5.32

Annual Log Return Statistics (first 50 years, 10k scenarios)								
	Heston	**Jump**	**SLV**	**RSLN2**	**RSDD2**	**LN**	**AIRG**	**History**
mean	9.78%	10.07%	9.93%	9.79%	9.90%	9.90%	7.31%	10.16%
st. dev.	14.92%	14.88%	14.14%	15.96%	15.72%	14.84%	15.13%	15.41%
skew	-0.58	-0.66	-0.74	-0.53	-0.66	0.00	-0.41	-0.90
kurt	4.20	4.41	4.77	3.73	4.36	3.01	4.12	4.54
Average Annual Return (30 yrs)	11.47%	11.80%	11.37%	11.94%	11.22%	11.68%	8.81%	n/a
Amount over AIRG	2.66%	2.99%	2.56%	3.13%	2.42%	2.87%	0.00%	n/a

I. Distribution of 1-yr and 20-yr returns for reference models fit using Maximum Likelihood Estimation (MLE) based on S&P 500 monthly returns from 1957.03 to 2022.12

The tables below show annualized equity returns across the reference models for 1-year and 20-year horizons compared to similar statistics for the AIRG and the GEMS Field Test (FT) #1a[24] scenario sets as well as to the distribution of rolling 1-year returns observed in history.

All reference models except for Lognormal (LN) allow for explicit return/volatility (negative skew) and produce scenarios with 1-year losses and gains that exceed the worst seen in history since 1957.03. Furthermore, when mean equity returns are constrained to 8.75% all reference models produce scenarios with negative returns over the first 20 years, an event not seen in history even including the great depression, and approximately 1% of scenarios experience negative returns over the first 20 years.

GEMS FT #1a shows a much lower median of 6.7% over the first year due to equity returns in that scenario set keying off the short interest rate which starts near 0% in that scenario set, while its first percentile is like the reference models (other than LN), i.e., reference models that model equity returns using a constant mean are able to achieve rates as low as GEMS FT #1a, even when starting short rates are near 0%. Over the first 20 years, GEMS FT #1a (which was calibrated to align with the AIRG 30-year GWF over the long-term) shows a median of 7.2% that is similar to the reference models when the mean equity return is constrained to 8.75% (first percentiles are also similar, as was the case over the first year).

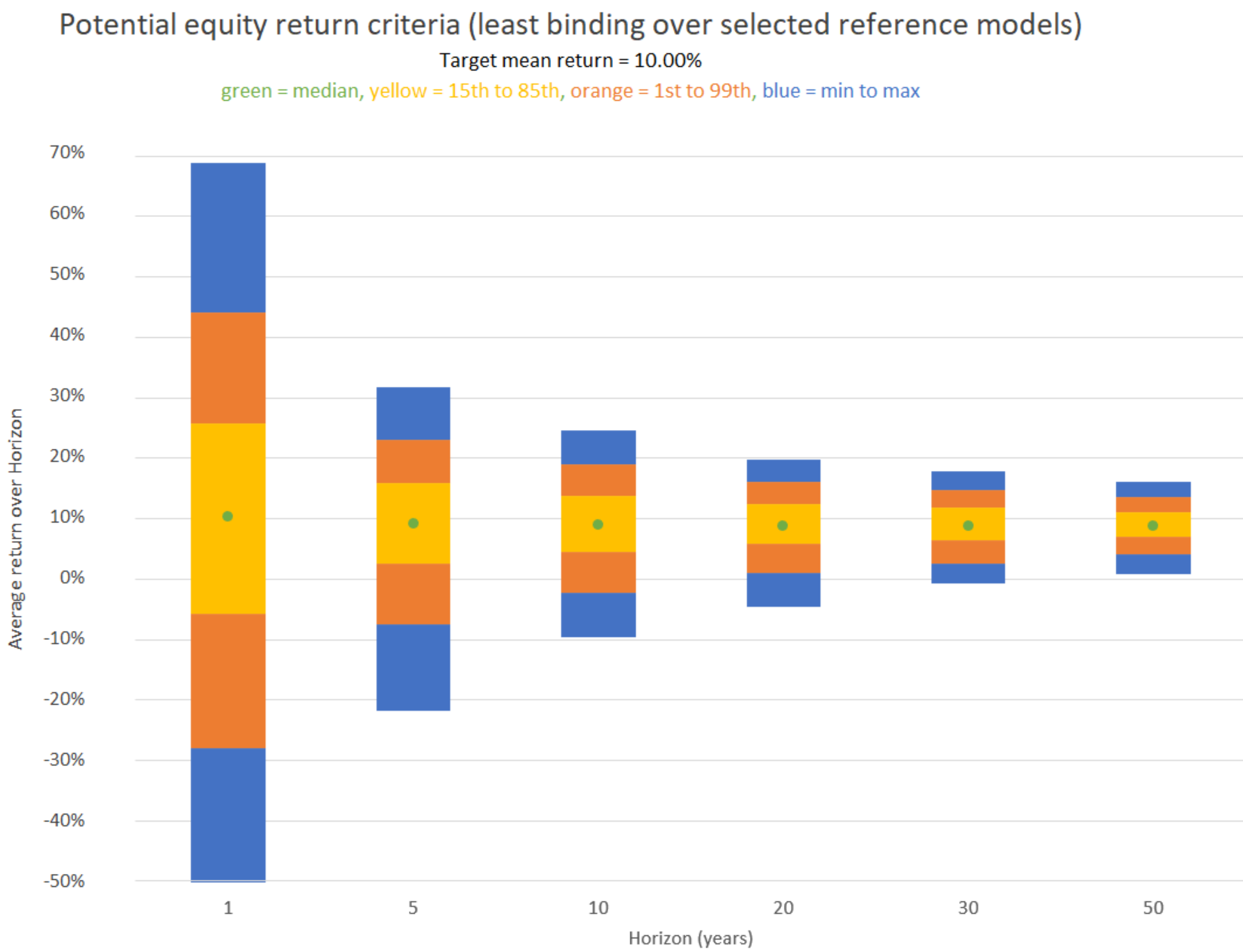
Distribution of 1-yr Return (Historical MLE, unconstrained mean)								**GEMS**	
	Heston	**Jump**	**SLV**	**RSLN2**	**RSDD2**	**LN**	**AIRG**	**FT #1a**	**History**
min	-50.7%	-51.5%	-52.3%	-58.3%	-51.8%	-41.9%	-59.5%	-49.7%	-43.3%
1	-26.9%	-27.7%	-28.0%	-29.2%	-31.1%	-22.1%	-28.5%	-29.2%	-33.2%
5	-15.1%	-15.0%	-14.8%	-17.6%	-18.2%	-13.6%	-17.3%	-18.2%	-15.2%
15	-5.5%	-4.8%	-4.5%	-5.7%	-6.4%	-5.4%	-7.7%	-8.7%	-5.0%
30	3.2%	3.9%	4.3%	3.8%	2.1%	2.3%	0.5%	-0.9%	5.3%
50	11.7%	12.5%	12.3%	12.4%	9.7%	10.8%	8.2%	6.7%	13.3%
70	19.7%	20.5%	20.2%	20.7%	17.6%	19.6%	16.4%	14.1%	19.9%
85	28.2%	28.2%	27.3%	28.8%	25.2%	28.8%	25.3%	21.7%	27.9%
95	38.1%	37.3%	36.1%	39.6%	34.8%	40.8%	36.5%	30.6%	37.8%
99	49.5%	48.5%	45.9%	53.0%	47.0%	56.1%	51.6%	41.0%	48.2%
max	82.9%	72.1%	71.0%	96.1%	78.9%	89.3%	85.2%	80.9%	61.1%

[24] Note that GEMS Field Test scenario (FT) set #2a would produce higher GWFs than GEMS FT #1a. This is because those two scenario sets model equity returns as a constant mean equity risk premium over the short rate, and starting short rates were significantly higher in GEMS FT #2a than GEMS FT #1a, especially for the one-year horizon.

Distribution of 20-yr Return (Historical MLE, unconstrained mean)								**GEMS**	
	Heston	**Jump**	**SLV**	**RSLN2**	**RSDD2**	**LN**	**AIRG**	**FT #1a**	**History**
min	-3.7%	-3.6%	-3.3%	-6.7%	-0.5%	-3.0%	-4.4%	-4.5%	n/a
1	1.2%	1.2%	2.4%	0.9%	3.7%	2.2%	-0.6%	-0.9%	n/a
5	4.0%	4.3%	4.9%	3.7%	5.5%	4.6%	1.7%	1.5%	n/a
15	6.5%	7.0%	7.1%	6.2%	7.3%	6.7%	3.9%	3.4%	n/a
30	8.5%	8.9%	8.9%	8.4%	8.7%	8.5%	5.8%	5.3%	n/a
50	10.5%	10.9%	10.6%	10.6%	10.3%	10.4%	7.7%	7.2%	n/a
70	12.4%	12.8%	12.3%	12.7%	11.9%	12.4%	9.5%	9.3%	n/a
85	14.2%	14.5%	13.9%	14.8%	13.6%	14.4%	11.4%	11.3%	n/a
95	16.3%	16.4%	15.7%	17.2%	15.6%	16.7%	13.7%	14.0%	n/a
99	18.4%	18.7%	17.6%	19.9%	17.8%	19.3%	16.2%	17.1%	n/a
max	22.1%	22.2%	21.2%	27.6%	23.1%	28.4%	20.3%	22.3%	n/a

Distribution of 20-yr Return (Historical MLE, mean constrained to 8.75%)								**GEMS**	
	Heston	**Jump**	**SLV**	**RSLN2**	**RSDD2**	**LN**	**AIRG**	**FT #1a**	**History**
min	-6.1%	-6.2%	-5.6%	-9.4%	-2.7%	-5.5%	-4.4%	-4.5%	n/a
1	-1.2%	-1.5%	0.0%	-2.0%	1.4%	-0.4%	-0.6%	-0.9%	n/a
5	1.5%	1.5%	2.4%	0.7%	3.1%	1.8%	1.7%	1.5%	n/a
15	3.9%	4.1%	4.6%	3.2%	4.9%	3.9%	3.9%	3.4%	n/a
30	5.8%	5.9%	6.3%	5.3%	6.3%	5.7%	5.8%	5.3%	n/a
50	7.8%	7.8%	8.0%	7.4%	7.9%	7.5%	7.7%	7.2%	n/a
70	9.7%	9.7%	9.7%	9.5%	9.4%	9.4%	9.5%	9.3%	n/a
85	11.4%	11.4%	11.2%	11.5%	11.0%	11.4%	11.4%	11.3%	n/a
95	13.4%	13.2%	13.0%	13.9%	13.0%	13.7%	13.7%	14.0%	n/a
99	15.6%	15.5%	14.8%	16.5%	15.1%	16.2%	16.2%	17.1%	n/a
max	19.2%	18.9%	18.4%	24.0%	20.3%	25.0%	20.3%	22.3%	n/a

Appendix 3—Graphical view of table 3.2 showing proposed criteria for selected percentiles based on a constrained mean equity return of 10.00%



Appendix 4—Additional analysis on the relationship between equity returns and interest rates

A. Summary of additional analysis

Historical Observations:

- Realized equity return (and ERP) are *inversely related* to the 3M UST rate in the top and bottom deciles, i.e., 17.6% equity return (17.6% ERP) when 3M UST rates are below 15 basis points, and 0.6% equity return (-9.9% ERP) when UST 3M rates are above 8.33%.
- The Fed's mandate includes balancing moderate inflation against employment and economic growth. Monetary tightening (e.g., when inflation is high) and easing (e.g., to stimulate the economy) rely on fundamental economic principles and contribute to the observed equity/rate dynamics of equity returns being *inversely* related to the level of 3M UST rates at tail levels.

GEMS assumed relationship:

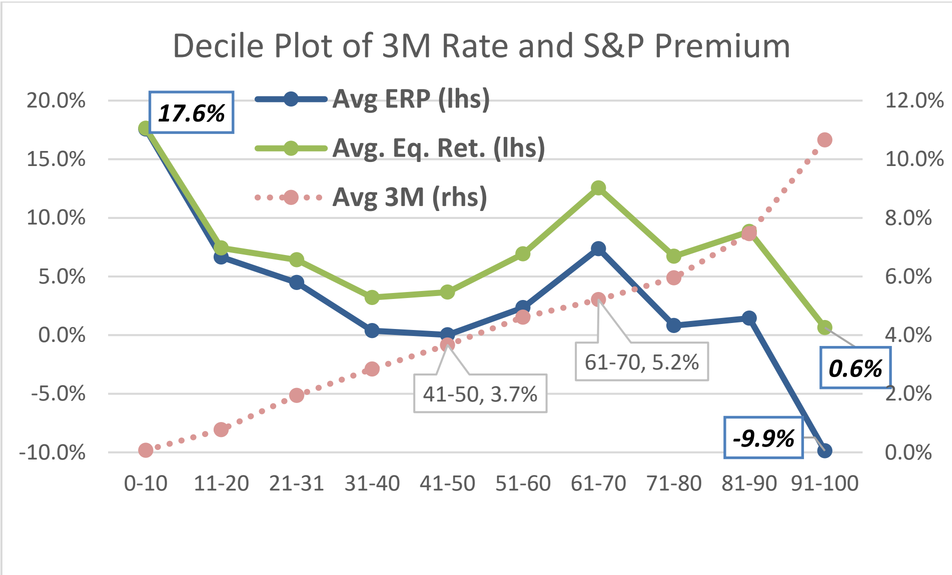
- GEMS assumes a constant expected ERP regardless of the level of rates, i.e., assumes expected equity return is an *increasing* function of the short (e.g., 3M) UST rate, which is opposite the historical relationship observed in the tails.
- Historical data suggests a constant ERP should be rejected at p-values of **0.3%** and **1.6%** in low- and high-rate deciles respectively. Low p-values indicate the model is not able to explain the realized equity return, even accounting for historical volatility.

AIRG assumed relationship:

- The AIRG model assumes a *constant* expected equity return, i.e., an ERP that decreases with the level of short (e.g., 3M) UST rate.
- While stylized, this assumption represents a middle ground between an inverse relationship seen in history and an increasing relationship embedded in GEMS.
- The AIRG model would be rejected at a p-value of 2.4% in the low-rate decile but shows p-values of 10%+ across the rest of the distribution.

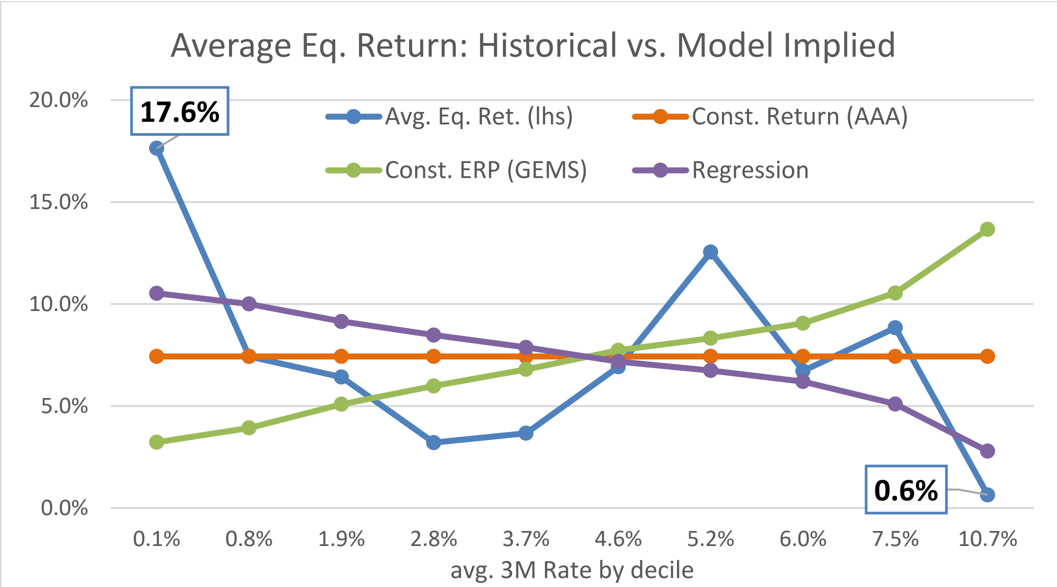
> Both the AIRG (constant mean equity return) and GEMS (constant mean equity risk premium) assume ERP dynamics which are stylized simplifications of the complete historical record. Both may adequately account for historical equity returns in the middle 80% of the distribution and on average over the long term, but the constant mean equity return approach better accounts for the historical dynamics (and underlying Fed policy) observed in the tails (i.e., top and bottom deciles).

B. Realized Excess Equity Returns are Inversely Tied to Tail Rate Levels



- Based on monthly 3m treasury and S&P price index data from 4/1953 to 12/2020, covering 814 months, or 67.8yrs.
- Monthly ERP calculated as return on S&P less average 3M yield through the month, expressed on continuous/log basis.
- Each decile represents 81-82 monthly points, or 6.8yrs, where data was grouped by 3M rate.
- Average Eq. Return and ERP were then calculated for each decile, and annualized.
- Realized equity return and ERP are inversely related to 3M rate in the top and bottom deciles: 17.6% equity return (17.6% ERP) where rates are below 15bp, and 0.6% return (-9.9% ERP) when rates are above 8.33%.
- Historically, the 3M treasury rate is strongly tied to Fed Funds rate, which is typically targeted by the Fed to achieve its objectives under different market environments: Easing post 2008 Financial Crisis intended to stimulate economic growth and employment, resulted in near 0% short rates for most of the last 11 years. The policy, facilitating borrowing and spending, had contributed to growth in equities and outperformance over the money market.
- Fed tightening intended to control high inflation in 70s and 80s lead to double-digit short rates, especially in late 70s through early 80s. As intended, the policy of flooring borrowing rates at historically high levels stimulated savings while stifling inflation and economic growth, contributing to money-market outperformance vs. equities over an extended period.

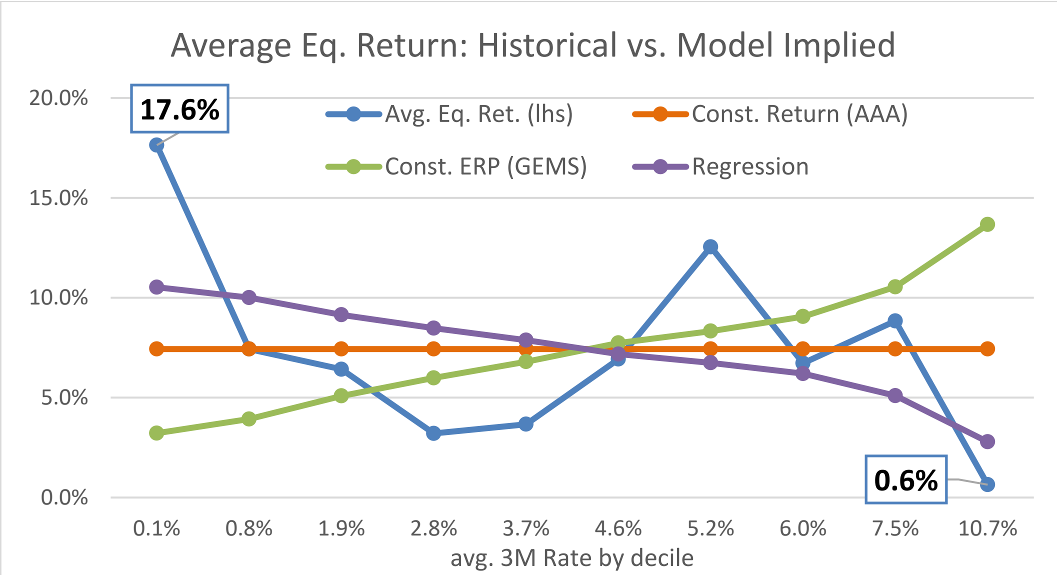
C. Simplified Modeling of Equity Return/ERP



	Avg. ERP (RF)	Avg. Return (RF)	Avg. Implied ERP	Avg. Implied Return
Constant Return (AIRG)	*7.43% - RF*	*7.43%*	3.15%	7.43%
Constant ERP (GEMS)	*3.15%*	*RF + 3.15%*	3.15%	7.43%
Regression Line	*10.58% - 1.74 x RF*	*10.58% - .74 x RF*	3.13%	7.41%

- Above assumes: **Equity Return = RF + ERP**, where RF is the return associated with risk free rate, and ERP is the Equity Risk Premium or excess earned over RF.
- Analysis considers constant ERP, constant Return, and ERP as a linear function of risk-free rate.
- While no considered approach perfectly captures historical data, a regression line showing inversely related equity return and risk-free rate, best aligns with the decile distribution.
- Constant (expected) Return (used in AIRG) implies ERP that is inversely related to risk-free rate.
- Constant (expected) ERP assumption, such as used in GEMS, results in increasing equity returns as a function of short rate – *directionally opposite* to what has been observed. Economic scenarios based on constant ERP would tend to:
    - Produce overly punitive equity returns in low-rate tail scenarios.
    - Understate the risk of adverse equity performance in high-rate tail scenarios.

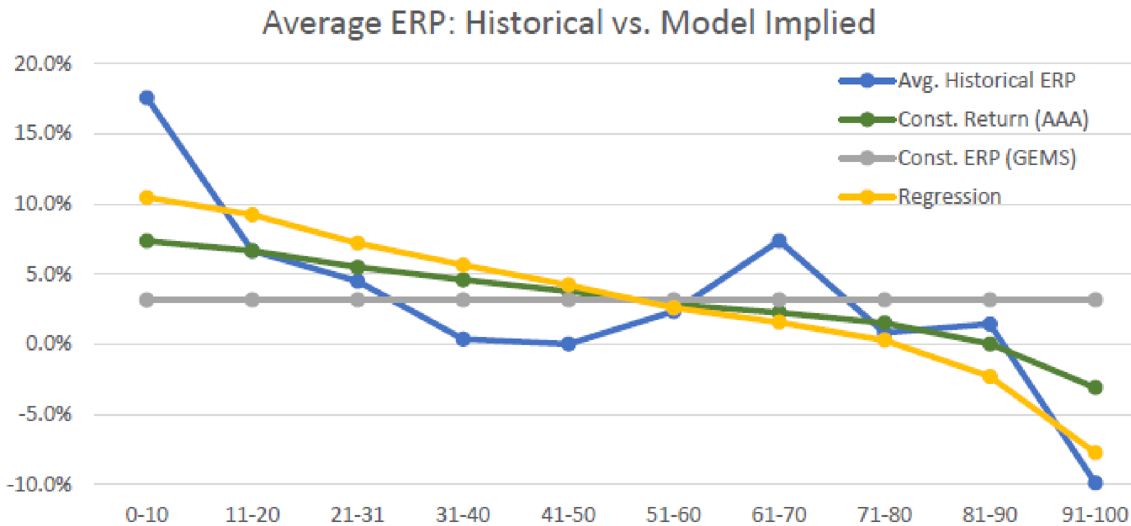
D. Historical Example: Fitting to January 2011 Treasury Curve



Deciles of 3M UST Yield	Average 3M UST Yield	Regression constrained to constant mean ERP	Regression constrained to constant mean Equity Return	Regression unconstrained
0-10	0.1%	0.3%	2.4%	8.2%
10-20	0.8%	29.5%	50.0%	34.6%
20-30	1.9%	40.1%	42.5%	30.5%
30-40	2.8%	29.6%	20.8%	15.5%
40-50	3.7%	20.5%	16.1%	13.5%
50-60	4.6%	43.9%	46.2%	48.1%
60-70	5.2%	21.6%	17.1%	14.1%
70-80	6.0%	35.9%	45.6%	46.8%
80-90	7.5%	39.9%	41.7%	28.8%
90-100	10.7%	1.6%	12.6%	35.7%
Mean ERP =		***Constant***	***Constant - (1.00 × Risk Free Rate)***	***Constant - (1.74 × Risk Free Rate)***
Mean Equity Return =		***Risk Free Rate + Constant***	***Constant***	***Constant - (0.74 × Risk Free Rate)***

- P-values were calculated to test the null hypothesis where average observed Equity Return within each decile was generated by each of the simplified ERP models.
- A small value of p, below a significance level (a popular choice is 5%) implies that the null hypothesis can be rejected with high confidence.
- Constant ERP produces very small p-values in the tails, especially in the first decile, where p = 0.3% implies that the model would produce average ERP/equity return > 17.6% over 82 months with only 0.3% probability.
- Constant Return assumption (used in AIRG) improves upon Constant ERP in the tails and is better capable of generating scenarios that reflect historical macro-economic/fed policy interactions.

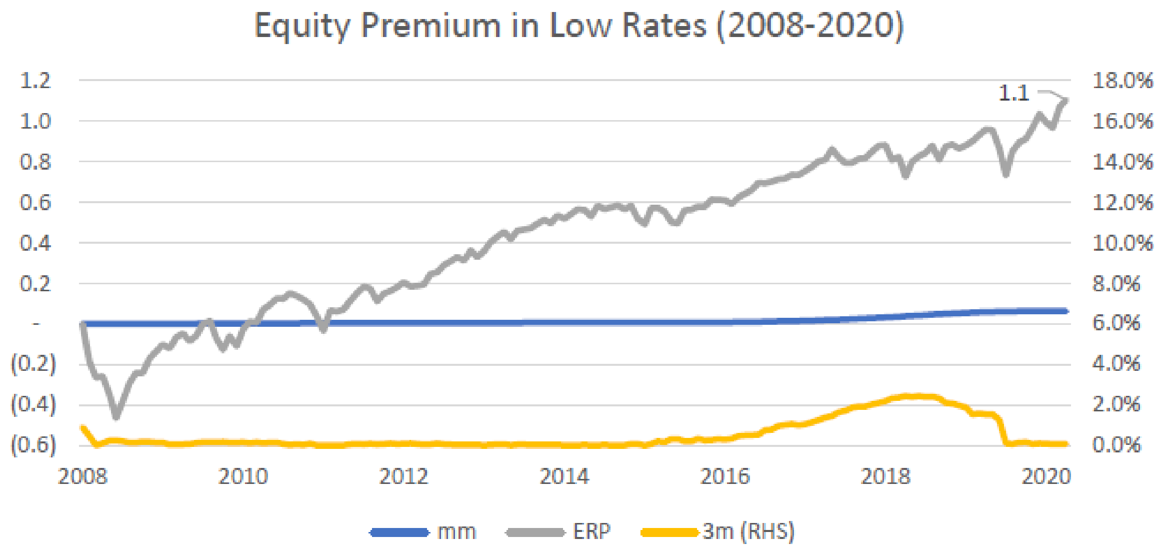
E. Simplified Modeling of ERP

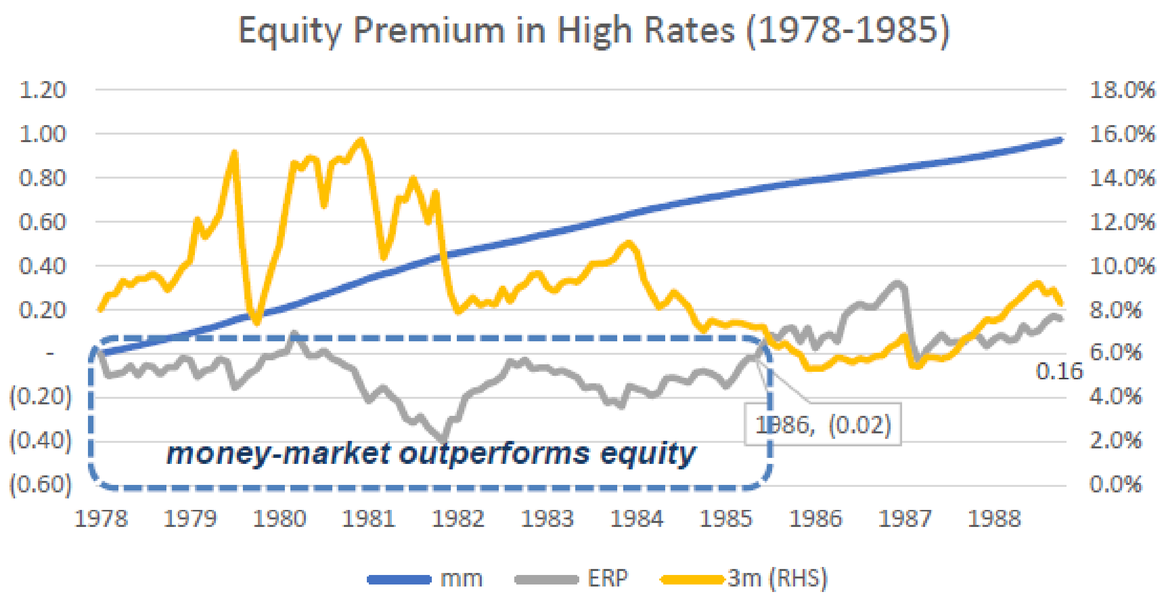


	Avg. ERP (RF)	Avg. Return (RF)	Avg. Implied ERP	Avg. Implied Return
Constant Return (AIRG)	*7.43% - RF*	*7.43%*	3.15%	7.43%
Constant ERP (GEMS)	*3.15%*	*RF + 3.15%*	3.15%	7.43%
Regression Line	*10.58% - 1.74 x RF*	*10.58% - .74 x RF*	3.13%	7.41%

- Above assumes: ***Equity Return = RF + ERP***, where RF is the return associated with risk-free rate, and ERP is the Equity Risk Premium or excess earned over RF.
- Analysis considers constant ERP, constant Return, and ERP as a linear function of risk-free rate.
- Rather than focusing on Equity Return by decile, this view of the data focuses on averaged realized ERP by decile.
- The chart suggests Constant Return assumption is better aligned with historical experience in tail deciles.

F. Historical Tails





- Cumulative Money Market return and ERP are presented on a continuous (log-return) basis, scaled on the left-hand side.
  - Example: wealth ratio (year 10) = exp(cumulative log return (year 10)).