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