OLIVER WYMAN

NAIC
National Association of Insurance Commissioners

VARIABLE ANNUITY STATUTORY RESERVE AND CAPITAL REFORM

QIS II PUBLIC REPORT

FEBRUARY 12, 2018

MARSH & McLENNAN COMPANIES

Confidentiality

Our clients' industries are extremely competitive, and the maintenance of confidentiality with respect to our clients' plans and data is critical. Oliver Wyman rigorously applies internal confidentiality practices to protect the confidentiality of all client information.

Similarly, our industry is very competitive. We view our approaches and insights as proprietary and therefore look to our clients to protect our interests in our proposals, presentations, methodologies and analytical techniques. Under no circumstances should this material be shared with any third party without the prior written consent of Oliver Wyman.

Contents

1. Purpose of this document

1.1. Background

The National Association of Insurance Commissioners ("NAIC") enacted its Risk-Based Capital ("RBC") C3 Phase II initiative in 2006 to prescribe a set of standards for calculating the RBC capital charges for VA products. C3 Phase II was followed in 2009 by Actuarial Guideline XLIII ("AG 43"), which prescribed the reserving standards for VA products.

These standards and their complex interplay challenged VA statutory capital management and in part motivated VA writers to use captive reinsurers. Recognizing the challenges, individual regulators have tried to accommodate these moves while upholding prudent standards of supervision. However, the individual nature of these arrangements, as well as the diversity of practices, led to a perception of an inconsistent application of standards across jurisdictions. The NAIC therefore commissioned an initiative to identify changes to the C3 Phase II and AG 43 frameworks that would mitigate or altogether remove motivations for insurers to use captive reinsurance, while maintaining the regulatory prudence of the standards.

Oliver Wyman was engaged by the NAIC in early 2015 to interview ten specific VA writers to (i) understand motivations for ceding VA liabilities to captive reinsurers, (ii) capture the extent and modes of captive usage, and (iii) identify changes to the statutory framework that might motivate writers to recapture their exposures. On September 10, 2015, Oliver Wyman provided the NAIC with a preliminary report covering a number of ideas for improvements to the current AG 43 and C3 Phase II frameworks, with objectives centered around:

i. Mitigating the asset-liability accounting mismatch between derivative instruments used for hedging purposes and statutory liabilities;

ii. Removing non-economic volatility in statutory capital charges and resultant solvency ratios;

iii. Facilitating greater harmonization across insurers and products for greater comparability.

Subsequently, the NAIC decided to undertake a Quantitative Impact Study ("QIS") in order to assess the efficacy and potential impact of the recommended ideas, identify their strengths and weaknesses, and make amendments accordingly. The QIS involved fifteen participant companies and the American Council of Life Insurers ("ACLI"), spanning from February to July of 2016. Two voluntary Working Groups were also held on a regular basis to discuss in depth issues related to (i) hedge accounting treatments for certain derivative instruments used for hedging purposes and (ii) potential Standard Scenario-related reforms.

Specifically, the QIS aimed to identify the ability of the current statutory framework, as well as potential revisions thereto, to satisfy three core objectives:

i. **Ensure robustness of funding requirements**: reserve and capital requirements should be robust – i.e., adequate to ensure liability defeasance with reasonable confidence – in light of the portfolio risks and risk management activities;

ii. **Promote sound risk management**: additional risk mitigation, in ordinary circumstances, should reduce a portfolio's total funding requirements as well as reduce the pro-cyclicality of surplus and funding requirements net of risk mitigation;

iii. **Promote comparability across insurers and products**: assumptions should be standardized across companies and products – including those subject to other statutory reserving and capital frameworks – where appropriate; in addition, different framework provisions should carry comparable levels of conservatism.

Additionally, Oliver Wyman sought to abide by two additional objectives in designing the potential framework revisions that were tested in the QIS:

i. **Preserve current statutory construct where feasible**: elements that Oliver Wyman sought to preserve included the adherence to principles-based reserving, book value approach to statutory valuation, "time-to-worst" accumulated deficiency approach to determining asset adequacy, use of "real-world" capital markets scenarios, and the use of a Standard Scenario construct to govern assumptions;

ii. **Minimize implementation complexity**: where possible and appropriate, reduce the computational complexity of framework calculations to improve interpretability of results and minimize model risk.

The completion of the initial QIS in 2016 afforded several valuable insights into the current statutory framework as well as Oliver Wyman's initial recommendations – namely:

i. **Economic-based hedging can adversely impact statutory financials under the current statutory framework via multiple mechanisms**, including (i) misalignment of market-sensitivity between economic and statutory funding requirements, especially with respect to interest rates; (ii) lack of statutory funding requirement sensitivity to all risk factors in favorable market conditions, as reserves and Total Asset Requirement ("TAR") become floored at the Standard Scenario Amount or portfolio cash value; and (iii) greater funding requirements when reflecting economic hedging in stochastic calculations due to misalignment between statutory "real-world" scenarios and the risk-adjusted scenarios used in the market consistent no-arbitrage valuation ("fair value") of these derivatives on the statutory balance sheet;

ii. **Under the current statutory framework, full economic hedging of the entire contract is penalized while partial hedging can represent the optimum solution,** because of misalignments between the statutory and economic frameworks, partial instead of full economic hedging minimizes net effective net funding requirement volatility. In fact, full economic hedging often both (i) increases reserves and TAR under recent market conditions and (ii) generates greater surplus volatility compared to partial hedging;

iii. **Oliver Wyman's initial recommendations, presented in September 2015, are appropriate – but additional study is warranted**: solutions studied in the QIS were demonstrated to mitigate the accounting misalignment via both hedge accounting and increased market-sensitivity of the statutory reserves and TAR, while the revision of the C3 calculation framework both (i) stabilizes non-economic volatility in required capital and (ii) shows currently-common hedging strategies can reduce the overall capital requirements;

iv. **Additional opportunities to harmonize practices in the application of AG 43 and C3 Phase II exist**, as the QIS highlighted a number of areas of divergent practices across the industry – arising primarily from a lack of specificity in the current guidelines – with significant impact on funding levels and comparability across insurers.

As a result of these findings, Oliver Wyman judged the QIS to have both reaffirmed the initial ideas presented in September 2015 and motivated the recommendation of several new proposals. On August 23, 2016, Oliver Wyman presented the full set of recommended revisions to AG 43 and C3 Phase II to the Variable Annuity Issues Working Group ("VAIWG") of the NAIC, with a redlined version of both AG 43 and C3 Phase II guidance documents further provided on September 26, 2016. These recommendations are summarized in Section 1.2 of this document.

On August 23, 2016, Oliver Wyman also noted the need to conduct additional testing before the framework revision recommendations are implemented in order to test all recommendations simultaneously as a complete package. Furthermore, while Oliver Wyman's recommendations presented the structural elements of the proposed revisions to AG 43 and C3 Phase II, certain parameters still required finer parameterization.

The recommended framework revisions were exposed on September 15, 2016 for public commentary over a 60-day period. Comment letters collected from industry and other interested parties provided extensive feedback on all of the proposed framework revisions, with alternative proposals for framework revisions suggested in response to several of Oliver Wyman's recommendations. Industry feedback has emphasized the need for adequate additional time for reviewing and testing both (i) recommended framework revisions as a full package of simultaneous changes, and (ii) modifications to the Oliver Wyman recommendations.

In light of Oliver Wyman's recommendations for conducting additional testing and the industry feedback in support of such additional testing, the NAIC resolved to undertake a second QIS ("QIS II") that was conducted in 2017. Specifically, QIS II aimed to serve several purposes:

i. Calibrate specific parameters that remain outstanding in Oliver Wyman's 2016 recommendations;

i. Allow companies to test the recommended framework revisions in greater depth, and simultaneously as a comprehensive package;

ii. Evaluate the impact and efficacy of recommendations that were tested under a different set of parameters during the 2016 QIS and provide an opportunity to identify refinements to these parameters;

iii. Evaluate select potential modifications to Oliver Wyman's recommended framework revisions;

iv. Provide an opportunity for the remainder of the VA industry that did not participate in the 2016 QIS to digest the recommendations and conduct internal financial analysis.

Similar to the first QIS, participants in QIS II were asked to perform a series of statutory reserve and capital calculations with specified scenarios, inputs, and methodologies. These calculations were organized into "Projection Sets", with each Projection Set comprising a set of inputs and calculation methodologies designed to assess a particular aspect of the recommended changes to the current AG 43 and C3 Phase II frameworks. All participating companies were encouraged to complete all Projection Set calculations where feasible. However, participation in QIS II was voluntary; should any Projection Set have proven to be excessively burdensome operationally, participants were permitted to skip the Projection Set.

This document provides a detailed review of all aspects of QIS II, including:

i. The structure and timeline of the QIS, as well as the process undertaken to conduct the QIS;

ii. Technical specifications of the quantitative testing that was performed;

iii. Comprehensiveness of data received for each set of quantitative tests, and procedures for data verification followed prior to discussion of results;

iv. Analytical methodologies followed in assessing participant results;

v. Overview of discussions held with the participants and VAIWG on the implications of testing results, as well as the manner in which the testing results informed Oliver Wyman's ultimate recommendations;

vi. Additional analyses and discussions conducted in separate Working Groups of the QIS – including details of a policyholder behavior experience study conducted as part of a Working Group to inform recommendations on the prescribed behavioral assumptions within the Standard Scenario;

vii. Meeting agendas and conclusions from 21 trilateral meetings held between Oliver Wyman, the VAIWG, and participants throughout the course of QIS II.

Please note that throughout QIS II, a number of public VAIWG calls were also held – culminating in the public VAIWG call held on December 1, 2017 during which Oliver Wyman's final recommendations were exposed. This document is intended to supplement all of the other materials previously released as part of prior public VAIWG calls.

To protect the data confidentiality of participant submissions and – given the highly technical nature of the subject matter evaluated during QIS II – to prevent analyses from being misconstrued, Oliver Wyman agreed with the participants that release of quantitative testing results from QIS II would require approval by participants who have furnished the relevant data. As of the publication date of this document we have not yet obtained the necessary

approval; accordingly, this document does not contain quantitative testing results. However, we are engaged in conversations with the participants on the potential release of select quantitative testing results at a later date, and expect to publish a supplement to this report should we obtain the necessary approvals.

Unless otherwise stated, the interpretations, conclusions, and views presented in this document represent Oliver Wyman's views and are not necessarily endorsed by the QIS II participants, the VAIWG, or the NAIC.

1.2. 2016 Oliver Wyman recommendations

Oliver Wyman presented fourteen recommended proposals on August 23, 2016 at the conclusion of the first QIS. These proposals supported the five general ideas for framework revisions that Oliver Wyman presented to the NAIC on September 10, 2015:

i. Align economically-focused hedge assets with liability valuations;

ii. Reform the Standard Scenario (AG 43 and C3 Phase II);

iii. Align TAR and reserves;

iv. Revise asset admissibility for derivatives and DTAs;

v. Standardize capital markets assumptions.

The specific 2016 recommendations are outlined below in **Table 1**:

Table 1: recommendations presented by Oliver Wyman on August 23, 2016

Ideas		Specific proposals	
1	**Align economically-focused hedge assets with liability valuations**	1A	Endorse hedge accounting for derivatives originated as part of a VA hedge program
		1B	Remove the Working Reserve when calculating scenario GPVAD
		1C	Permit simplified reflection of hedging in liability projections
		1D	Allow higher credit for liability projections with modeled CDHS, but require back-testing
2	**Reform Standard Scenarios (AG 43 and C3 Phase II)**	2A	Align AG 43 Standard Scenario calculations more closely to the stochastic CTE framework
		2B	Remove the C3 Phase II Standard Scenario
		2C	Specify a fuller set of risk factors informed by prevailing conditions and test multiple paths
		2D	Refresh prescribed policyholder behavior assumptions to align with industry experience
3	**Align TAR and reserves**	3A	Require Starting Assets used in liability projections to remain close to the final reserve
		3B	Calculate C3 as the difference between reserves and a tail CTE on the same distribution
4	**Revise admissibility for derivative assets and DTAs**	4A	Increase admissibility limit for designated VA hedges
		4B	Increase admissibility limit for DTAs associated with VA portfolios

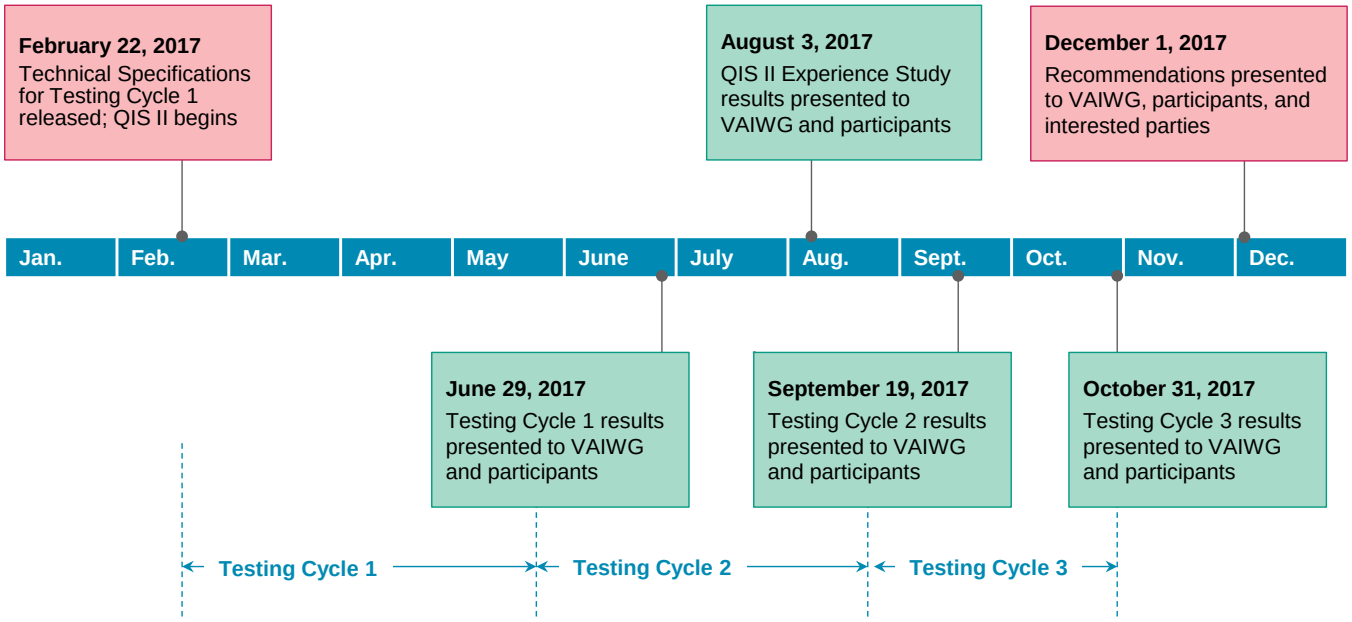
5	**Standardize capital markets assumptions**	5A	Harmonize interest rate and general account net investment income ("NII") assumptions
		5B	Evaluate alternative calibration criteria for equities and other market risk factors

2. QIS structure and Working Groups

QIS II was structured to accommodate variations in portfolio composition, modeling methodology, software capability, and resource availability across participant companies. In particular, QIS II adopted an iterative structure where three cycles of iterative testing was conducted to allow Oliver Wyman and the participants jointly to refine the testing scope and parameterizations in light of emerging insights.

The three Testing Cycles followed the timeline illustrated below in **Figure 1**:

Figure 1: QIS II timeline



Throughout the QIS II, Oliver Wyman also hosted three separate Working Groups that explored certain statutory framework enhancement proposals that either (i) may have required significant additional operational capabilities to complete, or (ii) were relevant for only a select subset of QIS participants. Within each Working Group, Oliver Wyman worked with the participating companies to design additional analyses to conduct in order to assess the efficacy and appropriateness of additional or alternative framework revisions. These Working Groups, in turn, brought any general insights from the additional analyses back to the broader group should such insights have had relevant implications or the potential to inform critically the final framework revision recommendations.

Specifically, the three Working Groups were as follows:

I. **Behavioral Assumptions Working Group**: this Working Group focused on the Standard Scenario policyholder behavioral assumptions that were proposed as part of Proposal 2D in Oliver Wyman's 2016 recommendations (see Section 2 in this document). The Working Group was tasked with two charges:

   A. Evaluate industry-aggregate policyholder behavior experience data to verify and/or determine potential alternative parameterizations of the prescribed Standard Scenario policyholder behavior assumptions included in the 2016 Oliver Wyman recommendations;

   B. Develop a "hybrid governance model" for Standard Scenario policyholder behavior assumptions that specifies prescribed methods to be applied to certain company-specific data – instead of prescribed assumptions – to calibrate behavior assumptions, provided that such data satisfy a set of pre-determined credibility standards. Use of the assumptions output by this model is anticipated to be elective.

II. **Economic Scenario Generation Working Group**: this Working Group assessed the use of proprietary modeling methodologies to generate the stochastic economic scenarios used to calculate the CTE Amount. Specifically, this Working Group sought to:

   i. Survey the range of real-world scenario generation techniques currently employed in industry for interest rates, equity returns, and other asset returns – in particular, those that do not have explicit guidance for either calibration or generation in the current statutory framework;

   ii. Review the strengths and limitations of these techniques;

   iii. Quantify differences in the resultant scenarios generated via the different techniques, as well as the valuation impact of using different techniques given the same set of calibration criteria.

   Insights from this Working Group were used to support the additional investigations suggested by Proposal 5B in Oliver Wyman's 2016 recommendations (see Table 2 in this document) on the governance of, and need for harmonization in, scenario generation for broader risk factors.

III. **Reinsurance Issues Working Group**: this Working Group focused on issues that are specific to VA portfolios that have been reinsured. In particular, this Working Group aimed to ensure that the recommended framework revisions:

   i. Do not inadvertently create unintended and undue consequences – whether financial or operational in nature – for reinsurers of such portfolios;

   ii. Are sufficiently clear and well-specified such that reinsurers of such portfolios may unambiguously apply the intended methodologies therein to their assumed portfolios.

   As all participants in the first QIS were primary VA manufacturers, reinsurer participants in QIS II were encouraged to attend the Reinsurance Issues Working Group such that the final recommendation would be subjected to a sufficiently robust and comprehensive review from a reinsurer's perspective.

3. General instructions and definitions

The general instructions and definitions outlined in this section were applied to all tests that participating companies were requested to complete during QIS II. In general, unless otherwise specified by instructions in sections 4.1, 5.1, and 6.1 of this document, participating companies were requested to use their current interpretations of the statutory provisions and retain the associated existing modeling methodologies.

The remainder of this section is organized into the following sub-sections as detailed below in **Table 2**. For the remainder of this section, the term Valuation Date indicates December 31, 2016.

Table 2: sub-sections of QIS II definitions and general instructions

Sub-section		Topics covered
A	**In-force portfolio**	Scope and state of the in-force portfolio on which the specified calculations should be performed
B	**Market conditions**	Starting market conditions under which the specified calculations should be performed. To assess the effect of the recommended framework changes across different capital markets conditions, QIS II tested both the actual market conditions on the Valuation Date as well as a range of alternative hypothetical market conditions
C	**Output**	Valuation metrics requested for each test, absent additional specifications, and the requested format in which the metrics should be presented

Note that sections 3.1 to 3.3 are replicated from the actual *Technical Specifications* document published for QIS II participants, with select non-applicable sections – e.g., definitions that were not ultimately included in a test – removed. For ease of reference, all paragraphs included in the subsequent sub-sections are numbered.

3.1. In-force portfolio

1. "Baseline Portfolio": unless otherwise defined in a subsequent section of this document, all calculations should be conducted on the Baseline Portfolio.

    a. The purpose of the Baseline Portfolio is to represent, as accurately as possible, the state of each Company's portfolio of in-force variable annuity policies, on a statutory entity-specific basis, within the scope defined in Section 1 as of the Valuation Date.

    b. Consequently, the QIS requests that each company use a Baseline Portfolio that is as representative of its variable annuity in-force on the Valuation Date as is feasible given the QIS II timeline detailed in Section 2. To this end, each Company should strive to include in the Baseline Portfolio as much of the material variable annuity exposure in all of the Company's statutory entities. Participants are encouraged to share a brief rationale for their individual choices for their Baseline Portfolio.

    c. Companies with variable annuity exposure in multiple statutory entities are encouraged to construct separate Baseline Portfolios – one for each statutory entity – and submit separate results. Oliver Wyman will process and aggregate results for up to two entities per full participant. However, if provision of results for standalone entities proves overly burdensome operationally, companies may elect to either (i) use the in-force portfolio for the statutory entity with the largest variable annuity exposure, or (ii) construct the Baseline Portfolio on a consolidated basis.

    d. Participating companies should conduct all calculations reflecting the full extent of affiliated ("captive") and third-party reinsurance in existence – e.g., on a "net" basis – as of the Valuation Date.

e. Should companies use compression techniques to reduce run time requirements, we request that the use of such techniques be indicated and that the degree of compression be provided along with a description of the compression algorithm used

2. "Baseline Portfolio on a Recaptured Basis": the Baseline Portfolio on a Recaptured Basis should be constructed in an identical manner as the Baseline Portfolio, with the sole exception around the treatment of affiliated ("captive") reinsurance transactions.

   a. Affiliated ("captive") reinsurance transactions involving variable annuity portfolios for which the assuming statutory entity does not report statutory reserves and the RBC C3 charge under AG 43 and C3 Phase II, respectively, should be assumed to have never existed.

   b. To this end, participating companies should assume that all ceded business have been held entirely in the primary entity since issue.

   c. The QIS guidance of assuming the simple non-existence – instead of active unwind – of all affiliated reinsurance transactions is designed to not encumber companies with the specific recapture mechanics. Nevertheless, participating companies that believe that the act of unwinding the affiliated reinsurance structure will have a material, non-transient impact on the primary entity are requested to provide a description of the nature of the reinsurance treaty and the expected impact and its cause.

3.2. Market conditions

1. "Baseline Conditions": unless otherwise defined in a subsequent section of this document, all calculations should be conducted assuming that market conditions on the Valuation Date are Baseline Conditions.

   a. The Baseline Conditions should reflect the prevailing capital markets conditions as of market close on the Valuation Date.

   b. For the purpose of determining Baseline Conditions, participating companies may use any data source and apply any market data extrapolation methods – including but not limited to yield curve and implied volatility surface construction – that it deems fit.

   c. The QIS requests each company to provide a description of the data sources and extrapolation methods used in determining the Baseline Conditions.

2. "Stress Conditions": where specified in a subsequent section of the QIS, calculations should be conducted assuming that market conditions on the Valuation Date are Stress Conditions.

   a. The Stress Conditions represent potential future market conditions materially different from baseline conditions. This document may define two types of stresses for each projection set:

      i. *[X]% Equity Stress*: a [X] value that is less than zero indicates that the S&P 500 declines by [X]% relative to the actual S&P 500 on the Valuation Date. An [X] value that is greater than zero indicates a rise in the S&P 500 by [X]% relative to the actual S&P 500 on the Valuation Date. Participating companies may project the changes in all other asset classes – including funds with risk controls – according to the internally-established correlation assumptions or modeling methodologies currently used for statutory calculations. For any future year shock, the convention shall denote total returns including dividends.

      ii. *[Y]% Rate Stress*: a [Y] value that is less than zero indicates that the ISDA mid-market USD par swap curve on the Valuation Date shift downwards in a parallel fashion by [-Y]% points. A [Y] value that is greater than zero indicates that the ISDA mid-market USD par swap curve shifts upwards in a parallel fashion by [Y]% points. Participating companies may project the changes in other yield curves – e.g., sovereign bond yields – according to internally-established correlation assumptions or modeling

methodologies. However, the minimum interest rate at any point in any yield curve should be no lower than 0.01%.

b. Market factors that are not detailed in a given set of Stress Conditions should be kept at Baseline Condition levels.

c. Where a calculation requires that Stress Conditions be used, each company should assume that the specified stress occurs linearly over 21 trading days – or one month, though other suitably short periods of time are also acceptable for reflecting the specified stress as long as companies advise Oliver Wyman of the selected time period.

d. Rebalancing in volatility-control funds should be reflected during the 21-day period (or other selected time period). Participating companies should reflect their own assumptions about rebalancing and volatility control ineffectiveness during this time.

e. Where a calculation stipulates Stress Conditions, these should be reflected in both the attributes of the in-force portfolio and calibration of the stochastic scenarios used in the statutory calculations.

   i. Within the in-force portfolio, market-linked attributes – including but not limited to benefit parameters and annuitization factors – should be updated to reflect their contractual or expected states under the Stress Conditions.

   ii. Within the in-force portfolio, contract account values and asset allocations across funds – as well as within individual funds – should be reflected on a best-efforts basis to reflect their designed or expected states under the Stress Conditions.

   iii. Stochastic scenarios used in the statutory calculations should be calibrated to reflect the Stress Conditions as starting conditions for the scenario generation, as specified, subject to the instructions items (iv) and (v) below.

   iv. For stochastic interest rate scenarios reflecting Stress Conditions, the mean reversion parameter for the 20-year U.S. Treasury bond rate ("MRP") – as described in section A5.2) of the redlined AG 43 document that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II* – should be set equal to the sum of the MRP effective as of the Valuation Date and [Y]% if a calculation appends the phrase "*with MRP effect*" after the description of the *[Y]% Rate Stress*. Otherwise, the MRP should be set to equal that effective as of the Valuation Date.

f. For each set of Stress Conditions, companies are requested to provide detailed disclosures of methodologies and assumptions used to (i) project changes in other asset classes; (ii) project rebalancing in volatility-control funds; and (iii) perform the in-force portfolio alterations.

3.3. Output

1. For all stochastic calculations conducted under the statutory framework, unless otherwise specified in a subsequent section of this document, participating companies should provide:

   a. The Greatest Present Value of Accumulated Deficiency ("GPVAD"), which may be referred to as the "Lowest Present Value of Accumulated Surplus" in the existing C3 Phase II framework, for the "best-efforts" run on a scenario-by-scenario basis across all projection scenarios under the framework(s) tested;

   b. The GPVAD for the "adjusted" run on a scenario-by-scenario basis across all projection scenarios under the framework(s) tested;

   c. The GPVAD for an unhedged run that excludes the impact from hedging – including that from the run-off of currently-held hedge assets – on a scenario-by-scenario basis across all projection scenarios under the framework(s) tested. For the purpose of determining the Starting Assets to use in this projection,

participating companies may replace the currently-held hedge assets with cash or other general account assets provided that the aggregate market value of these replacement assets equals that of the hedge assets.

2. For all stochastic calculations conducted under the existing C3 Phase II framework, companies are requested to provide the C3 Phase II Tax Adjustment and the associated *f* factor calculations, but as a separate output distinct from the GPVADs if feasible.

3. To support a more granular understanding of the quantitative impact, the QIS also requests that companies provide – if readily available – the paths of the following items across all scenarios for all stochastic calculations:

   a. Discount factor at each projection time-step on the applicable tax basis for the test;

   b. Nominal value of accumulated assets at each projection time-step;

   c. Nominal value of accumulated deficiency or surplus at each projection time-step for tests conducted under the current AG 43 or C3 Phase II frameworks, where the Working Reserve is non-zero.

4. For all Standard Scenario calculations conducted under the recommended revised AG 43 and C3 Phase II frameworks, participating companies should provide:

   a. Discount factor at each projection time-step for each market path;

   b. Nominal value of accumulated assets at each projection time-step for each market path;

   c. Nominal value of Accumulated Product Cash Flows, as defined in section A3.2)G) of the redlined version of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, at each projection time-step, for each in-force model point used, and for each market path;

   d. Hedge gains and losses at each projection time-step for each market path.

5. For all Standard Scenario calculations conducted under the current AG 43 and C3 Phase II frameworks, unless otherwise specified in a subsequent section of the QIS, participating companies should provide:

   a. For AG 43, the Greatest Present Value of Accumulated Net Revenue ("GPVANR") results calculated on a seriatim level and aggregated via a simple sum – i.e., not allowing for diversification – to portfolio segment-level.

      Each portfolio segment should be defined by a combination of the living benefit rider and the death benefit rider – though companies may combine into a single segment versions of a particular rider that do not differ materially from each other.

   b. For C3 Phase II, the aggregated GPVANR at the portfolio level, accounting for time diversification of cash flows across policies as allowed by C3 Phase II.

   c. The Basic Adjusted Reserve and Cash Surrender Value on a portfolio-aggregate level.

6. For all fair value calculations, unless otherwise specified in a subsequent section of the QIS, participating companies should provide, as separate items, the market-consistent values of:

   a. Living benefit fees;

   b. Living benefit claims;

   c. Death benefit fees, if not included in the base contract fees;

d. Death benefit fees, if attributed from the base contract fees;

e. Death benefit claims;

f. Total fair value.

Participating companies are requested to provide these results on a portfolio segment-level. Each portfolio segment should be defined by a combination of the living benefit rider and the death benefit rider – though companies may combine into a single segment versions of a particular rider that do not differ materially from each other

4. Cycle I testing

Testing Cycle 1 aimed to evaluate the impact of Oliver Wyman's 2016 recommendations on the stochastic CTE components of the VA reserve and capital framework. In particular, Cycle 1 evaluated the following proposals:

i. 1B: Remove the Working Reserve when calculating scenario GPVAD;

ii. 1C: Permit simplified reflection of hedging in liability projections;

iii. 1D: Allow higher credit for liability projections with modeled CDHS, but require back-testing;

iv. 3A: Require Starting Assets used in liability projections to remain close to the final reserve;

v. 3B: Calculate C3 as the difference between reserves and a tail CTE on the same distribution;

vi. 5A: Harmonize interest rate and general account net investment income assumptions.

All Standard Scenario-related recommendations were deferred until Testing Cycles 2 and 3 for assessment. In addition, while Proposal 5B – evaluate alternative calibration criteria for equities and other market risk factors – was originally planned to be included in Testing Cycle 1, subsequent discussions with the VAIWG deferred its evaluation until Testing Cycle 3.

4.1. Testing specifications

Testing Cycle 1 comprised four primary Projection Sets, as outlined below in **Table 3**.

Ten Projection Sets were originally planned; however, subsequent discussions with the VAIWG and resource constraints among participants eventually required that the scope of evaluated Projection Sets be reduced to four, with the removal of the following Projection Sets:

i. Two Projection Sets related to the linkage between equity scenarios and prevailing interest rates on the valuation date, which the VAIWG decided not to test in QIS II on April 12 – see Section 9.2;

ii. Four Projection Sets on the Standard Scenario, which received relatively few on-time submissions from participants; these Projection Sets were subsequently consolidated with Projection Sets 11-13 in Testing Cycle 2 – see Section 5.1.

Accordingly, the remainder of Section 4 only discusses four Projection Sets, as the other Projection Sets were not used in the eventual presentation of Cycle 1 results to the VAIWG and QIS II participants.

Table 3: Projection Sets requested in Testing Cycle 1

	Projection Set description	Output
1	CTE Amount under current AG 43 and C3 Phase II	Stochastic CTE
2	Standard Scenario Amount under current AG 43 and C3 Phase II	Standard Scenario
3	Fair value of embedded options and guarantees, valued on a market-consistent basis and CTE Amount of all non-guarantee cash flows, valued under the revised AG 43 recommended by Oliver Wyman	Stochastic CTE and fair value of embedded options and guarantees
4	CTE Amount under the revised AG 43 recommended by Oliver Wyman, with select simplifications for testing purposes	Stochastic CTE

Participants were asked to conduct each Projection Set with a specific combination of in-force portfolio composition and assumed starting market conditions, as detailed in **Table 4**. For nomenclature in the remainder of this section, the [X] indicated in **Table 4** should be replaced with the number of the Projection Set – e.g., Projection 2.1 would indicate a calculation conducted using the methodologies prescribed by Projection Set 2, applied to the Baseline Portfolio under Baseline Conditions.

Note that:

i. Eleven combinations of in-force portfolio composition and assumed starting market conditions were originally planned. However, because of resource constraints among participants and to improve interpretability of the results in the Testing Cycle 1 presentation, Oliver Wyman eventually reduced the number of combinations that were ultimately presented down to five;

ii. Projection [X].11 was only requested from companies who engaged in affiliated reinsurance transactions for their variable annuity portfolio. For this calculation, companies were requested to provide the financials for the affiliated reinsurer, under the accounting basis applicable to the affiliated reinsurer as of the Valuation Date, for the portion of liabilities assumed from the Baseline Portfolio;

iii. While Oliver Wyman intended that, except for the circumstances outlined in item (i) above, all Projection Sets be conducted over the full range of starting market conditions outlined in **Table 4**, we were cognizant of the computational intensity of some of the requested Projection Set calculations and the run-time constraints that may be faced by some participants. Accordingly, Oliver Wyman discussed with participating companies throughout the Testing Cycle limitations on computational time and any need to reduce the number of tested starting market conditions outlined in **Table 4** for certain Projection Sets. Oliver Wyman also, separately from this document, issued a one-page prioritization grid that provided guidance on the calculations that should be prioritized for each Projection Set.

Table 4: in-force portfolio compositions and assumed starting market conditions to be used

Projection	Portfolio	Market conditions
[X].1	Baseline Portfolio on a Recaptured Basis	Baseline Conditions
[X].3	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *+3.0% Rate Stress*
[X].5	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-1.0% Rate Stress*
[X].9	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-40% Equity Stress* • *-1.0% Rate Stress*
[X].11	Baseline Portfolio	Baseline Conditions

Note that sections 4.1.1 to 4.1.4 are replicated from the actual *Technical Specifications* document published for QIS II participants, with select non-applicable sections – e.g., Projection Sets that were ultimately not tested – removed. In addition, Oliver Wyman removed from the replicated text in these sections expositions of the purpose of each Projection Set, as the evaluations of the results from each Projection Set – as well as their implications – are covered in greater detail in section 4.2.

For ease of reference, all paragraphs included in the subsequent sub-sections are numbered.

4.1.1. Projection Set 1 – Current Stochastic

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 and C3 Phase II stochastic calculations using the same methodology that they currently use under the existing statutory framework. Participants should also use the same scenario generation techniques and calibration parameters as those that they currently use under the existing statutory framework.

4.1.2. Projection Set 2 – Current Standard Scenario

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 and C3 Phase II Standard Scenario calculations using the same methodology that they currently use under the existing statutory framework.

4.1.3. Projection Set 3 – Fair Value of Options and Guarantees

1. **Definitions**: for this Projection Set, the term Fair Value of Options and Guarantees ("FVOG") is defined as the expected present value of all net cash flows associated with guarantees that have optionality – including both living benefits and death benefits – present in the in-force VA portfolio when viewed under a market-consistent lens.

    i. The FVOG is intended to represents the market cost of a hedge strategy that perfectly replicates the net cash flows – and eliminates all hedgeable market risk – associated with guarantees in the in-force should all assumptions be realized.

    ii. Companies are requested to use the set of market-consistent projection scenarios calibrated by Oliver Wyman for the purpose of the FVOG; these scenarios are provided to each participating company in a separately-attached file.

    iii. Separate account returns should follow the scenarios that Oliver Wyman provides. Where additional indices need to be generated, participants should ensure that the returns assume no risk premium or illiquidity premium above the risk-free rate of return.

    iv. Oliver Wyman will not provide scenarios for general account returns – given different investment policies across participants – and participants should therefore generate such returns using internal methods. However, participants should ensure that the returns are consistent with other fixed income returns – i.e., that they assume no risk premium or illiquidity premium above the risk-free rate of return.

    v. Oliver Wyman understands that some participants, in their existing fair valuations of the VA portfolio, apply an illiquidity premium or spread above the risk-free rate to the discount rate for certain GMWB claims that occur after account depletion. For the purpose of the QIS, participants are requested to use no such illiquidity premium or spread.

    vi. Oliver Wyman understands that many contracts do not charge an explicit "rider fee" for the death benefit, and that fees needed to offset the mortality risk associated with such guarantees are embedded within the base contract fees. Companies should include in the FVOG calculation only the explicit rider fee, and exclude from the calculation any fees that are embedded within the base contract fees.

vii. Participating companies should use best estimate assumptions – including, but not limited to, mortality and policyholder behavior – instead of Prudent Estimate assumptions used under the existing statutory framework.

viii. Companies should not include any margins for non-hedgeable risks or frictional cost of capital in calculating the FVOG.

2. **Calculations**: participating companies are requested to conduct two calculations:

   i. **FVOG**: calculate the FVOG for its in-force portfolio and provide the detailed output as outlined in section 3.3 of this document. Participants that submit statutory calculation results for multiple statutory entities in this QIS II are also requested to submit entity-level FVOG results and associated details for the same statutory entities such that the statutory results and FVOG results may be compared appropriately;

   ii. **Reserves and TAR calculated under the FVOG method**: conduct the full AG 43 stochastic calculation using the methodology outlined in the redlined versions of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, with the following modifications:

      a) In section A1.4)A), remove the requirement that the Starting Asset Amount be set equal to the approximate value of statutory reserves at the start of the projection;

      b) In section A1.4)A), remove the requirement that the Starting Asset Amount over the aggregate Cash Surrender Value of all contracts included in the projection be less than 98% or greater than 102% of the excess of the Conditional Tail Expectation ("CTE") Amount over the aggregate Cash Surrender Value of the same contracts;

      c) In place of items a. and b. above, participants should set the Starting Asset Amount to equal the aggregate Cash Surrender Value of all contracts included in the projection.

      d) Exclude from the projections all cash flows that were incorporated into the FVOG calculation – i.e., those associated with the in-force guarantees that have optionality. Instead, participants should reflect in the first projection time-period a lump-sum expense equal in amount to the FVOG.

      For this calculation, participants should assume that there are no currently-held hedge assets and that the company does not have a Clearly Defined Hedging Strategy ("CDHS"). Because of the removal of CDHS, participants need only conduct the "adjusted" run.

      Companies are requested to use the sets of statutory projection scenarios calibrated by Oliver Wyman and provided to each participating company in a separately-attached file. These scenarios are generated using the American Academy of Actuaries' economic scenario generator – posted on the Society of Actuaries' website – with the mean reversion point for the 20-year U.S. Treasury bond rate calculated via the formula in section A5.2) of the redlined AG 43, but without any other parameter changes. While Oliver Wyman does not anticipate revising the 2016 recommendation that allows companies to generate equity scenarios using proprietary generators as long as the scenarios follow the Gross Wealth Ratios outlined in section A5.3) of the redlined AG 43, we have decided to provide standardized sets of statutory scenarios for testing purposes to allow more consistent and informative analysis of results.

   In the detailed output provided for calculation (ii), participants should not apply any caps or floors to the GPVAD or "Lowest Present Value of Accumulated Surplus"; instead, all values – positive or negative – should be kept as-projected such that the magnitude of the accumulated surplus or deficiency in each scenario may be identified.

4.1.4. Projection Set 4 – Revised Stochastic

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 stochastic calculations using the methodology outlined in the redlined version of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, with the following modifications:

    i. In section A1.4)A), remove the requirement that the Starting Asset Amount be set equal to the approximate value of statutory reserves at the start of the projection;

    ii. In section A1.4)A), remove the requirement that the Starting Asset Amount over the aggregate Cash Surrender Value of all contracts included in the projection be less than 98% or greater than 102% of the excess of the Conditional Tail Expectation ("CTE") Amount over the aggregate Cash Surrender Value of the same contracts;

    iii. In place of items a. and b. above, participants should set the Starting Asset Amount to equal the sum of (i) the aggregate Cash Surrender Value of all contracts included in the projection and (ii) a discretionary non-negative amount in excess of the aggregate Cash Surrender Value, as described below, for three different iterations of the calculation.

    Participants are asked to conduct three iterations of these calculations, each with a different Starting Asset Amount. Oliver Wyman notes that apart from the requested iterations of the Starting Asset Amount, Projection Set 3 is meant to test in full all of the components of stochastic framework outlined in Oliver Wyman's 2016 recommendations. The request for iteration-specific outputs conducted across three different Starting Asset Amounts allows Oliver Wyman to both:

    i. Assess the need for a convergence corridor of a certain width for Proposal 3A in Section 1.2, given participants' estimates of the Net Asset Earned Rate on Additional Assets; and

    ii. Evaluate different parameterizations of the CTE levels involved in Proposal 3B in Section 1.2 by interpolating across the different iterations of the Starting Asset Amount.

    Participants may therefore choose to follow section A1.4)A) of the redlined AG 43 to set the Starting Asset Amount for one of the iterations, but should ensure that the discretionary amount of Starting Assets in excess of the aggregate Cash Surrender Value is meaningfully different across the three different iterations. Participants should subsequently produce the full set of outputs for stochastic calculations as defined in section 3.3 of this document for each iteration.

    Participants are also requested to conduct these calculations both on a pre-tax basis, whereby the effect of Federal Income Tax is excluded from the projection of Accumulated Deficiencies in accordance with the redlined AG 43 guidance document, and after-tax basis, whereby the effect of Federal Income Tax is reflected in the projection of Accumulated Deficiencies. The after-tax calculations will be used to assess the difference between the two methodologies outlined in the redlined RBC LR027 instructions document that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised RBC LR027 (CLEAN) – for QIS II* – i.e.,

    i. Defining the Total Asset Requirement as:

    $$[25\%] \times \big((\text{Pretax CTE}([98]) - \text{Stat. Reserve}) \times 65\% - (\text{Stat. Reserve} - \text{Tax Reserve}) \times 35\%\big)$$

    ii. Defining the Total Asset Requirement as:

    $$[25\%] \times (\text{Aftertax CTE}([98]) - \text{Stat. Reserve})$$

    For the after-tax calculation, participants may use their existing simplifications for modeling tax treatments and cash flows under the current C3 Phase II framework. Participants that do not project tax reserves directly are

requested to provide as output the C3 Phase II Tax Adjustment and the *f* factors computed in accordance with the current C3 Phase II framework.

Companies are requested to use the sets of statutory projection scenarios calibrated by Oliver Wyman and provided to each participating company in a separately-attached file. These scenarios are generated using the American Academy of Actuaries' economic scenario generator – posted on the Society of Actuaries' website – with the mean reversion point for the 20-year U.S. Treasury bond rate calculated via the formula in section A5.2) of the redlined AG 43, but without any other parameter changes. While Oliver Wyman does not anticipate revising the 2016 recommendation that allows companies to generate equity scenarios using proprietary generators as long as the scenarios follow the Gross Wealth Ratios outlined in section A5.3) of the redlined AG 43, we have decided to provide standardized sets of statutory scenarios for testing purposes to allow more consistent and informative analysis of results.

4.2. Results and discussions

4.2.1. Data received

In total, thirteen companies submitted materially error-free testing results that were ultimately used in the Testing Cycle 1 presentation.

Of the thirteen companies, six companies submitted results for Projection Set 3, which enabled Oliver Wyman to conduct additional analyses of the expected impact of framework revisions assuming that the submitting participant implemented a hedge program focused on fully immunizing the guaranteed benefit liability on a fair value basis.

Prior to the presentation of Testing Cycle 1 results, Oliver Wyman shared with each individual participant its own company-specific results in order to:

i. Walk through with individual participants the resultant impact of the tested recommendations and confirm Oliver Wyman's interpretation of the results submitted;

ii. Discuss with individual participants whether the results shown appear to be intuitive or reasonable given the companies' understanding of their own portfolios and risk profiles; and

iii. Where results were unexpected or appeared unreasonable, identify potential sources for data errors, data misinterpretations, or other deviations from expectations.

If a participant did not request revisions to – or retractions of – the company-specific results, Oliver Wyman merged the company's results into the industry-aggregate results to produce the eventual presentation.

4.2.2. Analytical methodology

We assessed the Testing Cycle 1 results via five lenses commonly considered in industry management decisions:

i. **Baseline funding requirement**: defined as the amount of assets required to meet a specific target RBC ratio under existing market conditions as of the Valuation Date – which was set to 2016YE for QIS II. This lens represents the most immediate measure of the balance sheet impact of any changes, but did not allow for the determination of how framework revisions would impact total balance sheet volatility and market sensitivity.

   For the Testing Cycle 1 presentation, Oliver Wyman used a uniform target RBC ratio for all participants at 400% Company Action Level as a proxy for typical target RBC ratios within the US life insurance industry.

ii. **Total surplus volatility**: defined as the change in total statutory surplus across different market conditions, both favorable and unfavorable. This lens allowed for an assessment of changes in the degree of alignment between the accounting standards for hedge instruments – i.e., fair value – and the revised statutory framework. A better alignment meant that a full, fair value-focused hedge program that has extensive coverage of the portfolio's market risk would result in a less volatile surplus across market conditions than under the current framework.

For the Testing Cycle 1 presentation, the market conditions used in quantifying the total surplus volatility are identical to those used for Projections [X].1, [X].3, [X].5, and [X].9, as outlined in **Table 4**.

iii. **Required capital volatility**: defined as the change in the RBC C3 charge across different market conditions, both favorable and unfavorable. The required capital volatility is driven by similar factors as those that drive total surplus volatility, with an additional source of volatility stemming from the mechanics of the C3 calculation. In the current statutory framework, the C3 charge is calculated as the arithmetic difference between two structurally misaligned calculations – i.e., AG 43 reserve and TAR – which produces substantial amounts of non-economic volatility in the C3 charge, thereby driving the motivation for using voluntary reserves and captives to minimize RBC ratio volatility. This lens allowed for a study of the degree to which Oliver Wyman's recommendations would mitigate the non-economic volatility observed in the C3 charge under the current framework.

For the Testing Cycle 1 presentation, the market conditions used in quantifying the required capital volatility are identical to those used for Projections [X].1, [X].3, [X].5, and [X].9, as outlined in **Table 4**.

iv. **Excess surplus volatility**: defined as the volatility of the total statutory surplus in excess of surplus required to meet the target RBC ratio, representing a company's capital distribution capacity or capital accretion needs. This lens combines the results from the prior two lenses – total surplus volatility and required capital volatility – to derive a consolidated measure on the volatility of a company's capital capacity.

For the Testing Cycle 1 presentation, the market conditions used in quantifying the excess surplus volatility are identical to those used for Projections [X].1, [X].3, [X].5, and [X].9, as outlined in **Table 4**.

v. **Effective funding requirement:** defined as the amount of assets required to maintain a specific target RBC ratio, net of hedging gains or losses, across different market condition changes – both favorable and unfavorable. Otherwise said, a company that holds the effective funding requirement at a given Valuation Date should be able to maintain a positive excess surplus across all tested market conditions.

For the Testing Cycle 1 presentation, the market conditions used in quantifying the effective funding requirement were identical to those used for Projections [X].1, [X].3, [X].5, and [X].9, as outlined in **Table 4**. Oliver Wyman defined a uniform target RBC for all participants at 400% Company Action Level for the market conditions in Projections [X].1, [X].3, and [X].5, but allowed the target RBC to be reduced to 300% Company Action Level for Projection [X].9 in light of the severity of the joint equity-interest rate stress.

This lens mimicked the typical solvency risk appetite frameworks in use by many industry participants, where the company would set its hedging and capital management strategies to ensure that the balance sheet has a certain level of resilience vis-à-vis market volatility. Accordingly, a company that has sufficient funding for its target RBC ratio under present market conditions, but that may not have sufficient hedging to remove all statutory balance sheet volatility, may choose to hold an additional "capital buffer" to absorb such residual volatility. The sum of the funding required for the company to reach its target RBC under present market conditions and the capital buffer is correspondingly the effective funding requirement.

Figure 2 to **Figure 6** illustrate the exhibits – filled out with dummy data – that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under each of the five lenses, as well as the target properties of the revised framework under each lens. Note that for results under the revised framework, Oliver Wyman assumed that all participants would use the same hedge "error factor" in weighting the CTE Amount (best-efforts) and CTE Amount (adjusted); to ensure the continued validity of the findings under different reasonable "error factors", multiple "error factors" were tested.

For simplicity, our analysis assumed that each participant was a mono-line VA company with no voluntary reserves or captives usage. In addition, we also assumed that each participant had a relatively small amount of high-quality general account assets with low C1 charges such that there was no meaningful diversification benefit between the C3 charge and other RBC charges in calculating the overall RBC requirement. These simplifications were intended to ensure that the evaluation may:

i. Demonstrate, without confounding factors, whether the recommendations alleviate fundamental framework shortcomings and reduce incentives for captive usage;

ii. Illustrate the expected impact of the recommendations on companies most affected by VA regulations.

However, Oliver Wyman noted that these simplifications need to be accounted for in evaluating the total funding requirement impact of recommended revisions, as:

i. Voluntary reserves and captives are often used to sustain target funding levels; and

ii. Companies with large non-VA businesses may see meaningful diversification benefit in RBC charges.

To quantify the impact of the first simplification, Oliver Wyman requested from all participants captive financials and voluntary reserve usage, both as of the Valuation Date. Based on this information, Oliver Wyman presented to the VAIWG the difference between the amount of assets backing participants' VA portfolios under the first simplification and the actual amount of assets backing participants' VA portfolios as of the Valuation Date. In addition, Oliver Wyman noted that:

i. The aggregate difference was primarily driven by the use of voluntary reserves, as removal of the voluntary reserves would increase the C3 charge and drive a "leverage effect" whereby multiples of the increase in C3 charge needed to be held for the company to meet the target RBC ratio of 400%;

ii. The impact of unwinding participants' captive arrangements as of the Valuation Date was small if voluntary reserves were allowed to be used in place of the captive arrangements – i.e., if the direct insurer immediately posted voluntary reserves equal to the difference between TAR and statutory reserve immediately after recapturing the reinsured business.

We did not quantify the impact of the assumption that there is no meaningful diversification benefit between the C3 charge and other RBC charges. We expect that the impact would be substantial for companies with significant non-VA business and a large general account balance sheet, though we also expect the range of potential impact to be wide given the diversity of business mix across participants. Nevertheless, regardless of the magnitude of the resultant impact, the Testing Cycle 1 presentation would overstate the participant-aggregate level of funding required under the recommended revisions – which may partially or entirely offset the overstatement of funding levels under the current framework, as described in the paragraph above.

Lastly, recall that Proposal 3B of Oliver Wyman's 2016 recommendations included a component that adjusts for the difference between the statutory and the tax reserves on the valuation date in calculating the C3 charge:

$$\begin{aligned}\text{C3} = \text{Scalar} \times \big(&(\text{Tail CTE Amount} - \text{Statutory Reserve}) \times (1 - \text{Federal Income Tax Rate}) \\ &- (\text{Statutory Reserve} - \text{Tax Reserve}) \times \text{Federal Income Tax Rate}\big)\end{aligned}$$

Oliver Wyman recommended the inclusion of this component in order to align tax reflection in the revised C3 calculation with that in the current C3 Phase II framework. The first component of the revised C3 formula captures the tax deductibility of the losses captured in the Tail CTE Amount in excess of the statutory reserve, but does not capture the tax deductibility of the losses captured in the statutory reserve in excess of the tax reserve. In the current C3 Phase II framework, tax deductibility on both types of losses is captured in aggregate through explicit modeling of tax effects in the projection.

Though Oliver Wyman requested tax reserve information for each Projection in each Projection Set, most participants were only able to provide tax reserves as of 2016YE and not under other Valuation Date market conditions that were tested. Accordingly, the Testing Cycle 1 presentation also excluded the impact of differences

between the statutory and tax reserves in calculating C3, implicitly assuming that the tax reserve was equal to the statutory reserve for every participant on the Valuation Date. During the presentation of Testing Cycle 1 results, we noted that the inclusion of differences between statutory and tax reserves in calculating C3 would reduce the C3 charge under the revised framework, the magnitude of which would be assessed in Testing Cycle 3 should more tax reserve information become available.

Figure 2: exhibits used in the Testing Cycle 1 presentation – baseline funding requirement (illustrative data only)

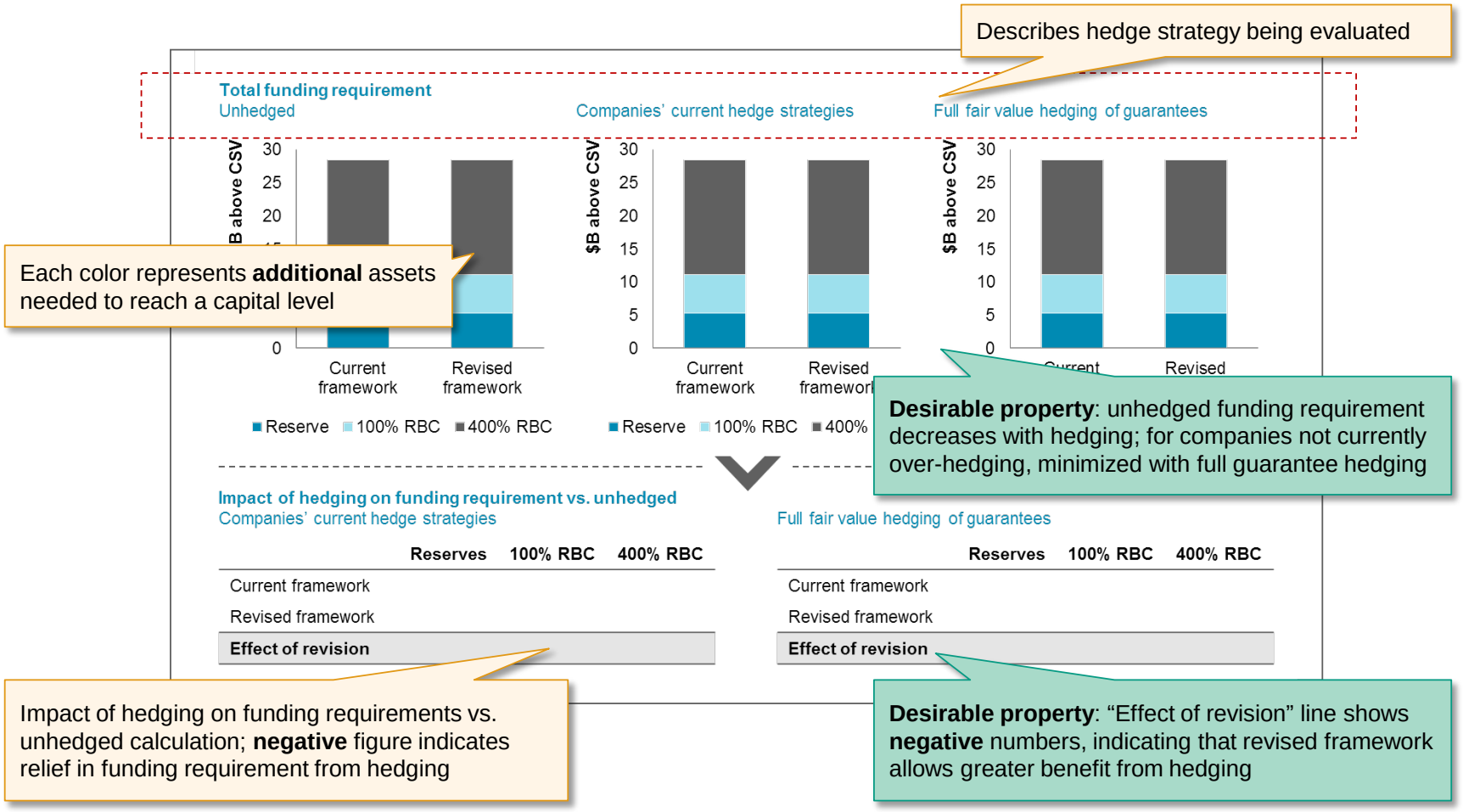


Figure 2 illustrates the exhibits that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under the baseline funding requirement lens. In addition, **Figure 2** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

Figure 3: exhibits used in the Testing Cycle 1 presentation – total surplus volatility (illustrative data only)

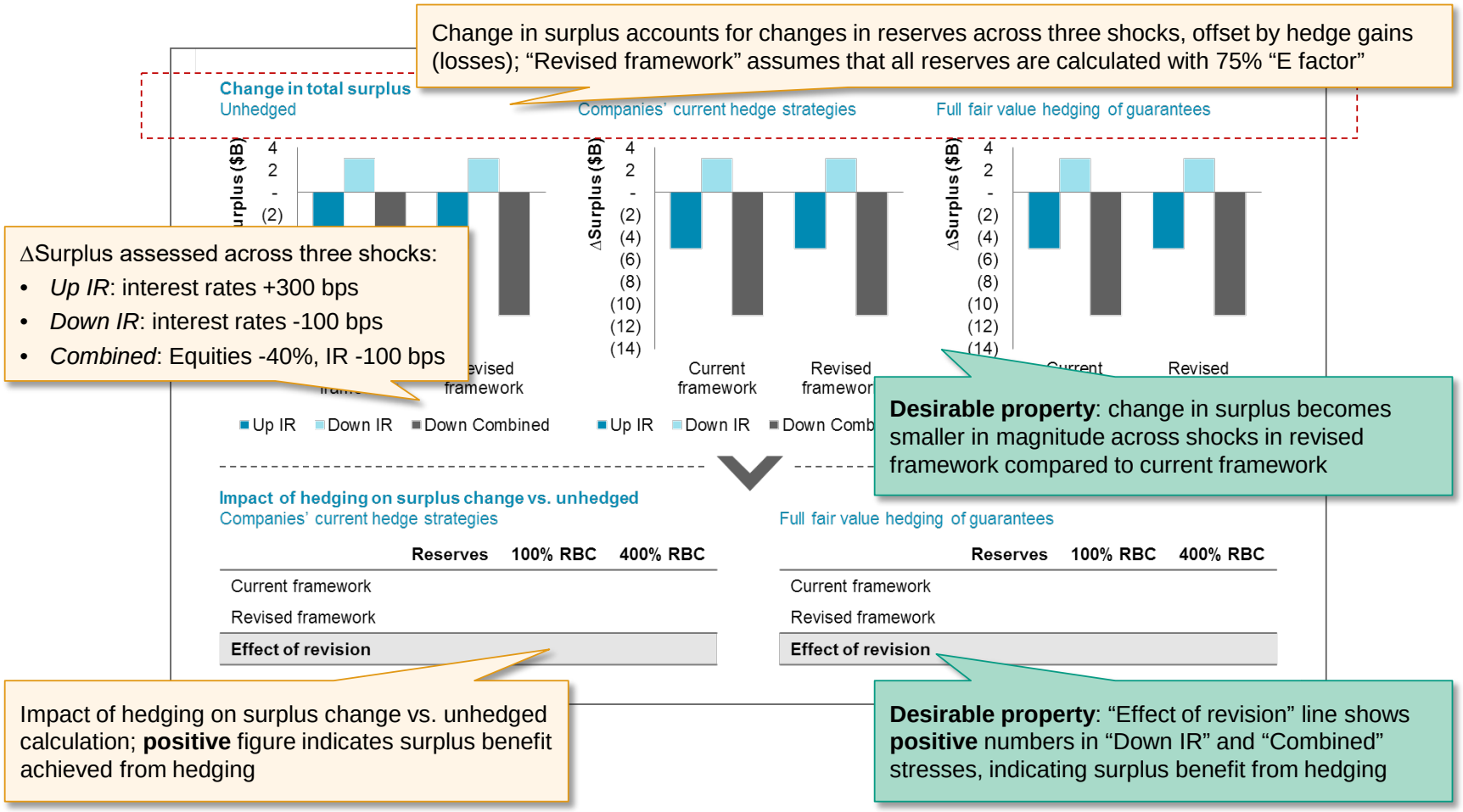


Figure 3 illustrates the exhibits that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under the total surplus volatility lens. In addition, **Figure 3** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

Figure 4: exhibits used in the Testing Cycle 1 presentation – required capital volatility (illustrative data only)

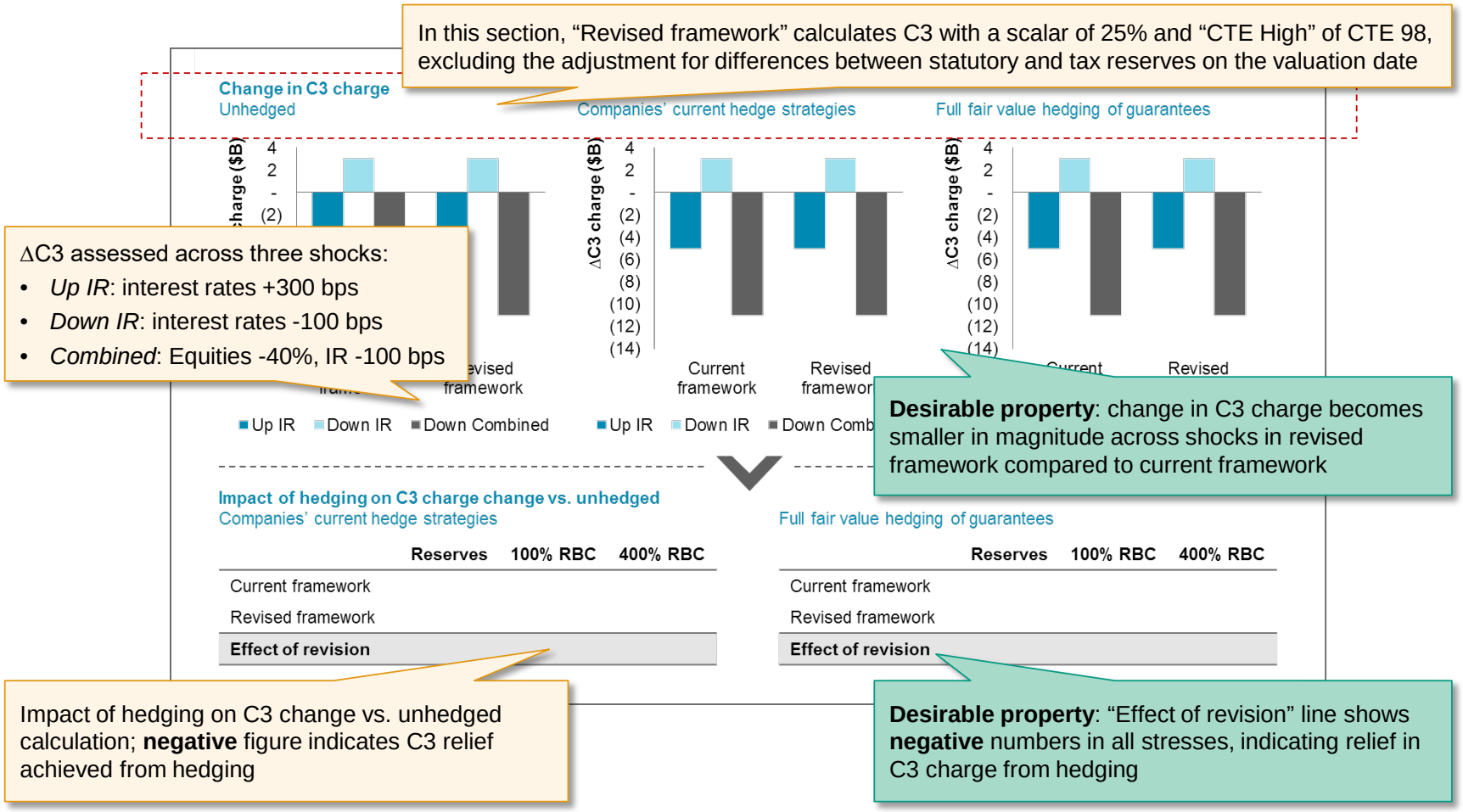


Figure 4 illustrates the exhibits that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under the required capital volatility lens. In addition, **Figure 4** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

Figure 5: exhibits used in the Testing Cycle 1 presentation – excess surplus volatility (illustrative data only)

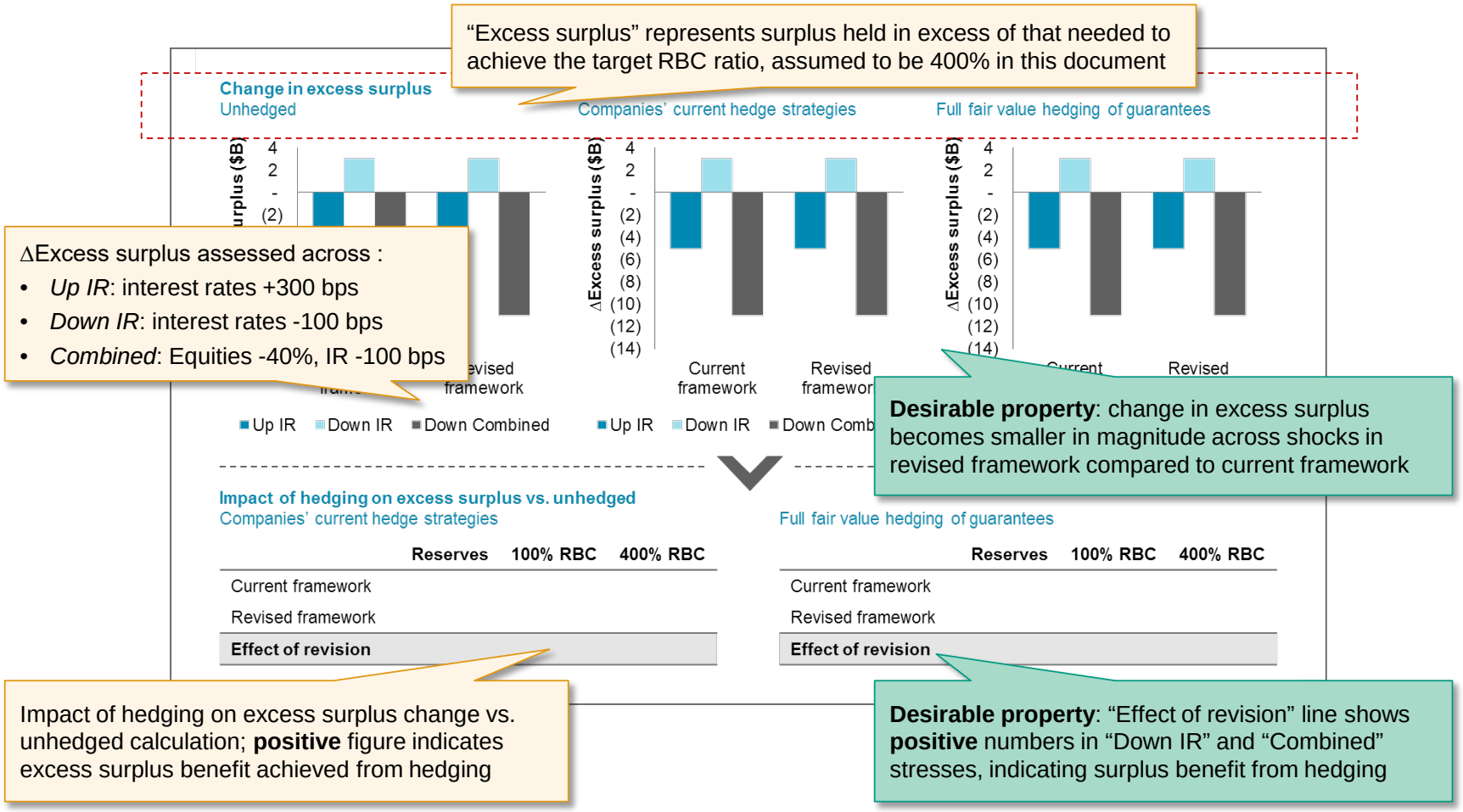


Figure 5 illustrates the exhibits that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under the excess surplus volatility lens. In addition, **Figure 5** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

Figure 6: exhibits used in the Testing Cycle 1 presentation – effective funding requirement (illustrative data only)

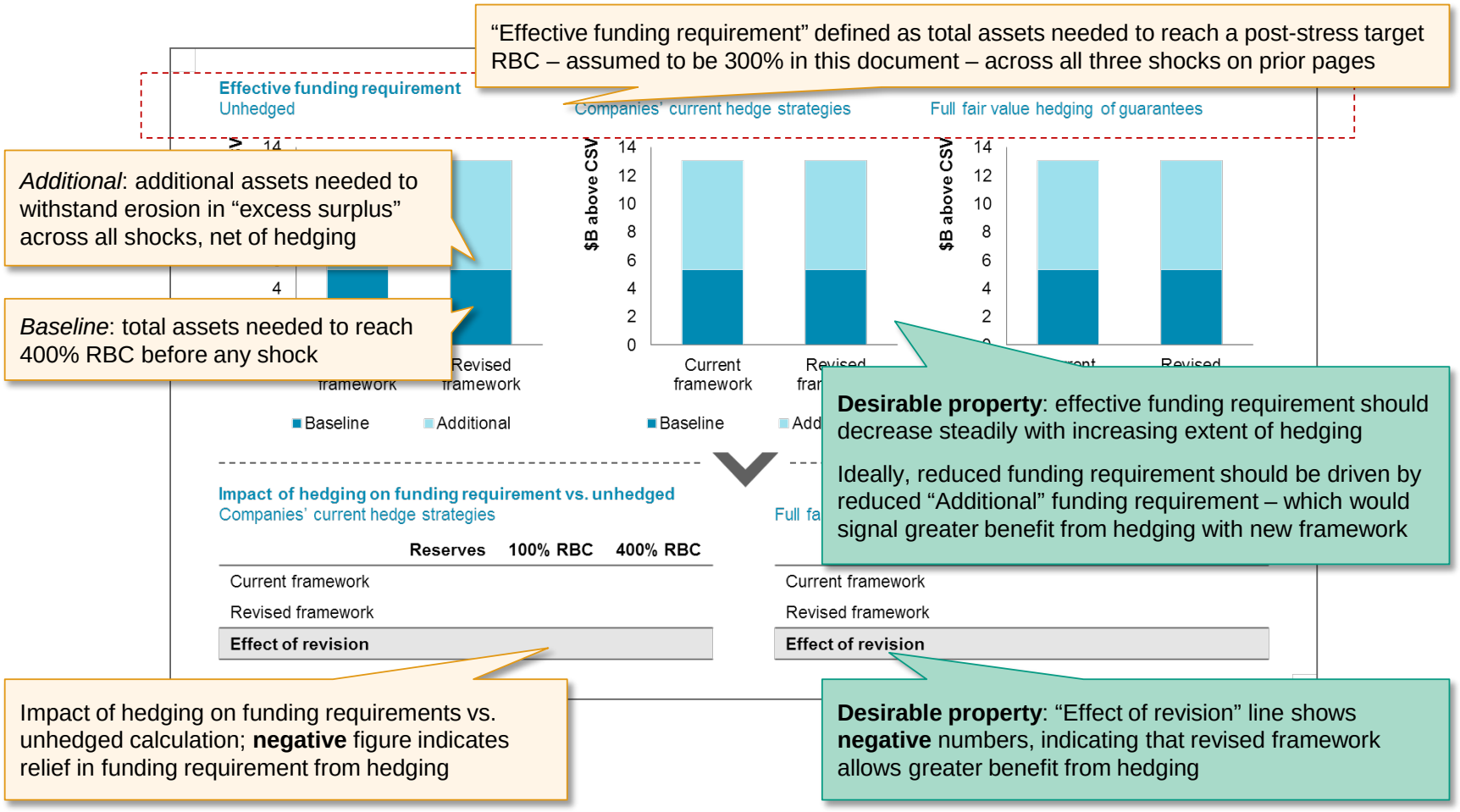


Figure 6 illustrates the exhibits that were used in the Testing Cycle 1 presentation to present participant-aggregate results as viewed under the effective funding requirement lens. In addition, **Figure 6** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

4.2.3. Discussion of results

Overall, testing Cycle 1 results indicate that Oliver Wyman's 2016 recommendations on the stochastic CTE components of the VA reserve and capital framework appear effective in reducing motivations for captive use. We observed that the tested recommendations:

i. Reduce total balance sheet volatility for participants with economically-focused hedge programs;

ii. Reduce volatility in the C3 charge and RBC ratio for most participants;

iii. Mitigate the capital buffer needed to manage multiples of the unstable C3 charge, particularly in stress conditions.

Additionally, the tested recommendations also appeared to align with the set of previously agreed-upon framework enhancement objectives, as we observed that the recommendations:

i. Are effective in improving the risk-sensitivity and overall "signal value" of the RBC ratio by removing non-economic volatility in total surplus and the C3 charge, as well as dis-incentivizing voluntary reserves;

ii. Promote comparability by harmonizing the stochastic capital markets scenarios used across companies and aligning the VA framework with VM-20 in general account modeling.

We also observed that the participant-aggregate "day-zero" TAR impact – calculated under the various assumptions outlined in Section 4.2.2 – was modest under companies' current hedge programs as of the Valuation Date. However, impact of revision varied materially across companies, largely driven by differences in hedge strategies and reflections thereof, as well as scenario generation approaches and portfolio risk profiles.

Specifically, across the five lenses outlined in Section 4.2.2, Oliver Wyman observed:

i. **Baseline funding requirement**: the revised framework maintained a similar TAR as the current framework under current participant hedge strategies. However, there was a moderate reduction in the total funding required to meet the target RBC of 400%, driven primarily by a reduction in C3 charge and partially offset by a reserve increase. Additionally, under the recommendations, a hedge strategy that fully immunizes all guaranteed benefit riders on a fair value basis minimized the C3 charge – a target property given that the hedge strategy would leave the portfolio with little residual VA market risk.

   Separately, while hedging increased reserves under both the current and revised frameworks, the revised framework showed a smaller increase in reserves – mostly as a result of the removal of the Working Reserve and allowance for immediate liquidation of hedges in the AG 43 CTE ("adjusted") calculation;

ii. **Total surplus volatility**: under companies' current hedge strategies, the tested recommendations reduced surplus volatility in interest rate shocks, but increased the volatility in the joint equity-interest rate shock – driven by increased market sensitivity in reserves. However, we note that in developing their current hedge strategies, companies invariably took into account the expected impact of such strategies on the statutory balance sheet. Accordingly, the total surplus volatility under the revised framework may be further reduced if companies decided to optimize their hedge strategies against the statutory balance sheet signature under the revised framework.

   The total surplus volatility analysis also demonstrated that fully hedging the guarantee benefit liability on a fair value basis would erode surplus significantly in a sharply-rising interest rate environment, driven by hedge losses that are offset by changes in reserves after reserves become floored at the portfolio cash surrender value. This issue was not addressed by the revised framework as the revised framework retains the cash surrender value floor for reserves. Accordingly, the analysis indicated a continued need for hedge accounting – i.e., Proposal 1A – to enable interest rate hedging without negative impact to surplus in a rising interest rate environment.

iii. **Required capital volatility**: relative to the current framework, the revised framework sharply reduced the volatility in the C3 charge across different market conditions under all hedge strategies tested – i.e., no hedging, participants' current hedge strategies, and full immunization of the guaranteed rider fair value.

Under the current framework, more extensive hedging typically reduced the C3 charge volatility. This property remained true under the revised framework as well, with C3 charge volatility minimized under the full guarantee immunization hedge strategy. However, the magnitude of the volatility reduction from hedging was lower under the revised framework, as the structural change in the C3 charge calculation had already removed much of the volatility that plagued the current framework.

iv. **Excess surplus volatility**: as the revised framework reduced the volatility in both total surplus and required capital relative to the current framework, it also reduced the volatility in excess surplus.

Under both the current framework and the revised framework, more extensive hedging reduced excess surplus volatility. In addition, the magnitude of the volatility reduction from hedging was also relatively similar between the two frameworks.

v. **Effective funding requirement**: the reduction in excess surplus volatility under the revised framework ultimately translated to a reduction in the "capital buffer" needed to maintain the target RBC ratio. This indicated that under the revised framework, companies would be less burdened with needing to maintain a large capital buffer to offset the non-economic volatility in total balance sheet requirement and RBC ratio present under the current framework, which was cited as the primary reason for captive usage.

However, Oliver Wyman noted that with a hedge strategy that fully immunizes the guaranteed rider fair value, the "capital buffer" required under the revised framework was nevertheless still relatively large. This was driven almost entirely by the surplus erosion in a sharply-rising interest rate environment, which galvanizes our support for hedge accounting – i.e., Proposal 1A – to enable prudent interest rate hedging without negative, non-economic balance sheet impact in a rising interest rate environment.

In addition, Oliver Wyman evaluated the properties of the revised framework across all five lenses with a number of alternative parameterizations of the C3 charge calculation outlined in Proposal 3B of Oliver Wyman's 2016 recommendations – i.e.,

$$\begin{aligned}\text{C3} = \text{Scalar} \times \big(&(\text{Tail CTE Amount} - \text{Statutory Reserve}) \times (1 - \text{Federal Income Tax Rate}) \\ &- (\text{Statutory Reserve} - \text{Tax Reserve}) \times \text{Federal Income Tax Rate}\big)\end{aligned}$$

Oliver Wyman's 2016 recommendations provisionally set the Scalar to 25% and the Tail CTE Amount to CTE 98, but noted that these parameters were subject to calibration in QIS II. While most of the Testing Cycle 1 analyses were conducted with these provisional parameters, Oliver Wyman also tested a number of other combinations of Scalars and Tail CTE Amount confidence levels, with:

i. Scalars ranging from 20% to 50%;

ii. Tail CTE Amounts ranging from CTE 95 to CTE 99.

We observed that while the choice of the Scalar and Tail CTE Amount determined the size of the C3 charge – and therefore the total funding requirement, it did not materially change the desirable properties on the total balance sheet volatility that were attained via the revised framework. Accordingly, we concluded that the parameterization of Proposal 3B may be conducted without needing to re-assess whether the desirable properties on total balance sheet volatility attained by the provisional parameterization of Proposal 3B are still maintained.

Based on these results, Oliver Wyman ultimately decided after Testing Cycle 1 that most of the recommendations related to the stochastic CTE components of the VA statutory framework were effective in achieving their target properties and did not need to be revised during QIS II.

5. Cycle II testing

Testing Cycle 2 aimed to evaluate the impact of Oliver Wyman's 2016 recommendations on the Standard Scenario component of the VA reserve and capital framework. Specifically, Cycle 2 evaluated the following proposals:

i. 2A: Align AG 43 Standard Scenario calculations more closely to the stochastic CTE framework;

ii. 2B: Remove the C3 Phase II Standard Scenario;

iii. 2C: Specify a fuller set of risk factors informed by prevailing conditions and test multiple paths.

In addition, Cycle 2 evaluated several methodologies for reflecting non-guaranteed revenue sharing income within the stochastic CTE calculation that differed from the currently-prescribed methodology in AG 43. The development of these tests was driven by the industry revenue sharing experience collected by the ACLI and presented to the VAIWG during the trilateral VAIWG meeting on May 24, 2017, as discussed in section 9.4.

5.1. Testing specifications

Testing Cycle 2 comprised three primary Projection Sets, as outlined below in **Table 5**.

Originally, Testing Cycle 2 planned for thirteen Projection Sets, with five Projection Sets indicated as low priority. However, the following Projection Sets were subsequently removed from consideration:

i. Five Projection Sets that aimed to evaluate the impact of alternative equity scenarios on the stochastic CTE calculation. At the request of the VAIWG, Oliver Wyman indicated these Projection Sets to be purely voluntary and low priority; because of resource constraints, very few participants submitted results;

ii. Three Projection Sets that were identical to Projection Sets 5-7, but required participants to replace their own Prudent Estimate assumptions with the Standard Scenario assumptions proposed in Oliver Wyman's 2016 recommendations.

   During the development of the Cycle 2 testing specifications, the Behavioral Assumptions Working Group was in the process of reviewing the recommended Standard Scenario assumptions. As such, Oliver Wyman indicated that these Projection Sets would not be used in the Cycle 2 presentation given the Standard Scenario assumptions would be in flux. However, Oliver Wyman noted that any revisions to the previously-recommended assumptions would likely be only parametric in nature, and encouraged participants to conduct these Projection Sets in any case such that they may test their systems readiness and ensure correct structural implementation of the assumptions.

Accordingly, the remainder of Section 5 only discusses five Projection Sets, as the other Projection Sets were not used in the eventual presentation of Cycle 2 results to the VAIWG and QIS II participants.

Table 5: Projection Sets requested in Testing Cycle 2

	Projection Set description	Output
11	Standard Scenario Amount under the revised AG 43 recommended by Oliver Wyman in 2016, but with: − 52 alternative capital market paths provided by Oliver Wyman; − Participants' own Prudent Estimate assumptions; − Reflection of CDHS over the first projection year	Standard Scenario
12	Standard Scenario Amount under the revised AG 43 recommended by Oliver Wyman in 2016, but with:	Standard Scenario

	− 52 alternative capital market paths provided by Oliver Wyman; − Participants' own Prudent Estimate assumptions; − Reflection of CDHS over the full projection timeframe	
13	Standard Scenario Amount under the revised AG 43 recommended by Oliver Wyman in 2016, but with: − 52 alternative capital market paths provided by Oliver Wyman; − Participants' own Prudent Estimate assumptions; − No reflection of CDHS and runoff of currently-held hedges	Standard Scenario
17	CTE Amount under the revised AG 43 recommended by Oliver Wyman, with an alternative reflection of Net Revenue Sharing Income proposed by Oliver Wyman	Stochastic CTE
18	CTE Amount under the revised AG 43 recommended by Oliver Wyman, with an alternative reflection of Net Revenue Sharing Income proposed by ACLI in their 2016 comment letter	Stochastic CTE

Participants were asked to conduct each Projection Set with a specific combination of in-force portfolio composition and assumed starting market conditions, as detailed in **Table 6**. For nomenclature in the remainder of this section, the [X] indicated in **Table 6** should be replaced with the number of the Projection Set – e.g., Projection 11.1 would indicate a calculation conducted using the methodologies prescribed by Projection Set 11, applied to the Baseline Portfolio under Baseline Conditions.

Note that:

i. Eleven combinations of in-force portfolio composition and assumed starting market conditions were originally planned. However, because of resource constraints among participants and to improve interpretability of the results in the Testing Cycle 2 presentation, Oliver Wyman eventually reduced the number of combinations that were ultimately presented down to four;

ii. While Oliver Wyman intended that, except for the circumstances outlined in item (i) above, all Projection Sets be conducted over the full range of starting market conditions outlined in **Table 6**, we were cognizant of the computational intensity of some of the requested Projection Set calculations and the run-time constraints that may be faced by some participants. Accordingly, Oliver Wyman discussed with participating companies throughout the Testing Cycle limitations on computational time and any need to reduce the number of tested starting market conditions outlined in **Table 6** for certain Projection Sets. Oliver Wyman also, separately from this document, issued a one-page prioritization grid that provided guidance on the calculations that should be prioritized for each Projection Set.

Table 6: in-force portfolio compositions and assumed starting market conditions to be used

Projection	Portfolio	Market conditions
[X].1	Baseline Portfolio on a Recaptured Basis	Baseline Conditions
[X].3	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *+3.0% Rate Stress*
[X].5	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-1.0% Rate Stress*
[X].9	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-40% Equity Stress*

- *-1.0% Rate Stress*

Note that sections 5.1.1 to 5.1.5 are replicated from the actual *Technical Specifications* document published for QIS II participants, with select non-applicable sections – e.g., Projection Sets that were ultimately not tested – removed. In addition, Oliver Wyman removed from the replicated text in these sections expositions of the purpose of each Projection Set, as the evaluations of the results from each Projection Set – as well as their implications – are covered in greater detail in section 5.3.

For ease of reference, all paragraphs included in the subsequent sub-sections are numbered.

5.1.1. Projection Set 11 – Revised Standard Scenario #5 (1-Year CDHS)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 Standard Scenario calculation using the methodology outlined in the redlined version of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, with the following modifications:

    i. Remove section A3.1)B); as an acceptable simplification for reducing run-time, participants may use the same in-force for the Standard Scenario calculation as that used in the stochastic calculation;

    ii. Remove sections A3.2)F)3), A3.2)F)4), A3.2)F)5), A3.2)F)6), A3.2)F)7), A3.2)F)8, A3.2)F)10), A3.2)F)11, and A3.2)F)12);

    iii. Remove section A3.2)F)2) on the prescribed maintenance expense assumptions;

    iv. Remove sections A3.2)H) on the Diversification Benefit Adjustment;

    v. For projecting maintenance expenses covered by the section removed in item iii) above, each participant should use the Prudent Estimate assumptions that it uses for conducting the stochastic AG 43 calculations to calculate the Conditional Tail Expectation Amount;

    vi. For projecting policyholder behaviors covered by the sections removed in item ii) above, each participant should use the Prudent Estimate assumptions that it uses for conducting the stochastic AG 43 calculations to calculate the Conditional Tail Expectation Amount. If such assumptions are appropriate only for a stochastic calculation – e.g., random sampling techniques across scenarios, the participant may elect to adopt an alternative modeling technique to implement the assumption provided that the participant informs Oliver Wyman of the chosen alternative modeling technique;

    vii. Replace the first sentence in section A3.2)E) – i.e., "*the Standard Scenario Amount shall be calculated as the greatest of the Scenario Greatest Present Values calculated for each of the three prescribed scenarios*" – with the following text in italics:

    The Standard Scenario Amount for each prescribed scenario shall be calculated as the Scenario Greatest Present Value for that prescribed scenario.

    Subsequently, remove sections A3.2)E)1), A3.2)E)2), and A3.2)E)3). Instead, each participant is requested to use the sets of Standard Scenario market paths provided by Oliver Wyman in a separately-attached file for calculating each Scenario Greatest Present Value. Please note that Oliver Wyman does not anticipate all of the provided market paths to become the prescribed Standard Scenario market paths in the final recommendations; instead, we have decided to provide a wide variety of standardized scenarios for testing purposes to allow comprehensive, consistent, and informative analysis;

    viii. Replace the table in section A3.2)E)4) with the following table:

Returns & indicators	All projection years
Balanced fund returns	Reflect the equity and bond allocations as of the valuation date and any expected asset rebalancing in the projection consistent with fund operations
General account earned rate	Consistent with the manner in which general account assets – including Starting Assets, reinvestment assets, and Additional Invested Assets as defined in section A1.2)D) – is reflected via the methodology outlined in section A1.4)D) of the redlined version of AG 43, including the requirement in A1.4)D)7) for fixed income assets
Fixed account returns	At the option of the actuary, either (i) follow the company's documented crediting practices; or (ii) equal to the larger of the contract's minimum guaranteed crediting rate and the general account earned rate less 200 bps
Implied and realized volatility	Follow the forward volatilities implied by the implied volatility term structure in effect as of the valuation date
Foreign exchange rates	Follow the exchange rates implied by spot exchange rates as of the valuation date and the relevant interest rate term structures

As outlined in section 3.3, participants should submit results for each requested market path separately as instructed in the separately-attached output template.

5.1.2. Projection Set 12 – Revised Standard Scenario #6 (Full CDHS)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to re-conduct all of the calculations in Projection Set 11, with one modification to the redlined AG 43 guidance document: in section A3.2)D), remove the restriction that cash flows associated with a Clearly Defined Hedging Strategy can only be reflected in the first projection year. To this end, the first three paragraphs of A3.2)D) should be replaced with the following text for the purpose of this Projection Set:

   Modeled Hedges. Throughout the projection, cash flows associated with hedging shall be projected in the same manner as that used in the calculation of the CTE Amount (best efforts) as discussed in section A7.3) and shall follow the general guidelines set forth in section A1.1)D). If the company is not following a Clearly Defined Hedging Strategy, or if the hedging strategy does not meet the requirements of Appendix 7, the projections shall not take into account the hedge positions expected to be held in the future through the execution of that strategy. In this case, the costs and benefits of hedging shall be projected in the same manner as that used in the calculation of the CTE Amount (adjusted) as discussed in section A7.3).

   At the end of the first projection year, if the company is not following a Clearly Defined Hedging Strategy, or if the hedging strategy does not meet the requirements of Appendix 7, all projected hedge positions shall be assumed to be liquidated and the projection shall not take into account any new hedge positions. Where applicable, the liquidation value of hedges shall be consistent with the assumed returns in the Standard Scenario from the start of the projection to the date of liquidation, market-consistent valuation of the hedge assets, and other market conditions as of the date of liquidation as defined in section A3.2)D).

5.1.3. Projection Set 13 – Revised Standard Scenario #7 (No CDHS)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to re-conduct all of the calculations in Projection Set 11, with one modification to the redlined AG 43 guidance document: in section A3.2)D), require reflection of the costs and benefits of hedging in the same manner as that conducted to calculate the CTE Amount (adjusted). To this end, the first three paragraphs of A3.2)D) should be replaced with the following text for the purpose of this Projection Set:

   Modeled Hedges. Throughout the projection, the costs and benefits of hedging shall be projected in the same manner as that used in the calculation of the CTE Amount (adjusted) as discussed in section A7.3). However, the projection shall not take into account the hedge positions expected to be held in the future through the execution of the company's hedging strategy.

5.1.4. Projection Set 17 – Revised Stochastic #4 (Alternative Revenue Sharing #1)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 stochastic calculations using the methodology outlined in the redlined version of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, with the following modifications:

   i. Replace section A1.1)E)6)b) in the redlined AG 43 guidance document with the follow text in italics:

      Is the smaller of a) non-contractually guaranteed Net Revenue Sharing Income estimated by the actuary on a Prudent Estimate basis, as described in section III)B)7), and b) the actuary's estimate for non-contractually guaranteed Net Revenue Sharing Income before reflecting any margins for uncertainty multiplied by the following factors:

      (i) 1.0 in the first projection year;

      (ii) 0.9 in the second projection year;

      (iii) 0.8 in the third projection year;

      (iv) 0.7 in the fourth and all subsequent projection years.

   ii. Add the following text in italics after section A1.1)E)6) as A1.1)E)7):

      In each projection year, the actuary may remove from the projected fund management, administration, or advisory fees paid by the separate account to the entity charged with the corresponding fund management, administration, or advisory duties the difference between (i) the actuary's estimate of non-contractually guaranteed Net Revenue Sharing before reflecting any margins for uncertainty and (ii) the amount calculated using the methodology outlined in section A1.1)E)6)b).

   iii. In section A1.4)A), remove the requirement that the Starting Asset Amount over the aggregate Cash Surrender Value of all contracts included in the projection be less than 98% or greater than 102% of the excess of the Conditional Tail Expectation ("CTE") Amount over the aggregate Cash Surrender Value of the same contracts;

   Companies are requested to use the same sets of statutory projection scenarios as those used in Projection Set 4 of Testing Cycle 1.

5.1.5. Projection Set 18 – Revised Stochastic #5 (Alternative Revenue Sharing #2)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.

2. **Calculations**: participating companies are requested to conduct the full AG 43 stochastic calculations using the methodology outlined in the redlined version of AG 43 that Oliver Wyman provided on February 21, 2017 under the file name *20170221 Revised AG 43 (CLEAN) – for QIS II*, with the following modifications:

    i. Remove section A1.1)E)6). Instead, participants should reflect the effect of Net Revenue Sharing Income projected in each scenario using an assumption consistent with what the participant deems to be a Prudent Estimate assumption, but not to exceed 90% of the participant's best estimate assumption;

    ii. In section A1.4)A), remove the requirement that the Starting Asset Amount over the aggregate Cash Surrender Value of all contracts included in the projection be less than 98% or greater than 102% of the excess of the Conditional Tail Expectation ("CTE") Amount over the aggregate Cash Surrender Value of the same contracts.

    Companies are requested to use the same sets of statutory projection scenarios as those used in Projection Set 4 of Testing Cycle 1.

5.2. Alternative Standard Scenario market paths tested

For Projection Sets 11-13, Oliver Wyman provided a number of market paths specifications to be tested. These market paths are summarized in this section; note that the text below was replicated from an actual Technical Memorandum published for QIS II participants.

Oliver Wyman has provided 52 different market paths covering four types of scenarios, as summarized in the following table:

Scenario type	Scenario #
"Intrinsic value" with bond spread	1
"Stressed intrinsic value" with bond spread	2-4
Stress and constant rate of recovery	5-13, 17-25, 29-37, 41-49
Paths corresponding to specific percentiles of stochastic scenarios	14-16, 26-28, 38-40, 50-52

Separate sets of 52 market paths have been provided for the five different starting interest rate conditions requested in Projection Sets 11-16:

i. Baseline: 12/30/2016 conditions;

ii. +3% IR: 12/30/2016 conditions with a parallel 300 bps upward shift in USD swap curve;

iii. +3% IR including MRP: 12/30/2016 conditions with (i) a parallel 300 bps upward shift in the USD swap curve and (ii) a 300 bps upward shift in the mean reversion target of 20-year UST rate;

iv. -1% IR: 12/30/2016 conditions with a parallel 100 bps downward shift in USD swap curve;

v. -1% IR including MRP: 12/30/2016 conditions with (i) a parallel 100 bps downward shift in USD swap curve and (ii) a 100 bps downward shift in the mean reversion target of 20-year UST rate.

"Intrinsic value"-related scenarios

Scenarios 1-4 are analogous to four of the market paths tested in Testing Cycle 1 (S3, S4, S5, and S8), with several modest changes:

i. A uniform 100 bps of earned bond spread has been added to the bond fund returns; in Testing Cycle 1, participants were requested to include their own estimate of the earned spread based on general guidance provided in the redlined AG 43;

ii. Application of the interest rate shock in the first year has been revised such that the projected equity returns track the forward rates in the same scenario more closely; and

iii. For better comparability with the other market paths tested in Cycle 2, swap rates provided in the scenario files have been redefined to represent par rates instead of spot rates.

Details of Scenarios 1-4 are provided in the following table:

#	*Eq. stress*	*Eq. recovery*	*IR stress*[1]	*IR recovery*	*Bond stress*	*Bond recovery*
1	0.0%	Follow forwards	0 bps	Follow forwards	0.0%	Follow forwards + 100 bps p.a.
2	-13.5%		0 bps		-2.7%	
3	0.0%		-25% of 10-year swap		+150% of 10-year swap	
4	-9.5%		-18% of 10-year swap		+106% of 10-year swap	

Scenarios with initial stress and constant rate of recovery

These scenarios test a wide range of market paths of the following structure:

i. Initial equity and/or interest rate stress occurring smoothly over the first projection year;

ii. Constant per annum equity returns for all projection years thereafter.

A wide range of market paths is tested in order to identify the market path that most closely reflects the CTE70 level of funding requirement for each participant.

Details of these scenarios are provided in the following table:

#	*Equity stress*	*Eq. recovery*	*IR stress*	*IR recovery*	*Bond stress*	*Bond recovery*
5	0.0%	0% p.a.	0 bps	Mean path in Academy IRG	0.0%	Trailing avg. of UST rate + 100 bps
6		2% p.a.				
7		4% p.a.				
8	-6.75%	0% p.a.			-1.35%	
9		2% p.a.				
10		4% p.a.				
11	-13.5%	0% p.a.			-2.7%	
12		2% p.a.				

[1] All stresses applied as parallel shifts in the yield curve. Stresses are floored at -50 bps.

#	Equity stress	Eq. recovery	IR stress	IR recovery	Bond stress	Bond recovery
13		4% p.a.				
17	0.0%	0% p.a.	0 bps	Mean path in Academy IRG with -100 bps MRP shock	0.0%	Trailing avg. of UST rate + 100 bps
18		2% p.a.				
19		4% p.a.				
20	-6.75%	0% p.a.			-1.35%	
21		2% p.a.				
22		4% p.a.				
23	-13.5%	0% p.a.			-2.7%	
24		2% p.a.				
25		4% p.a.				
29	0.0%	0% p.a.	-50 bps	Mean path in Academy IRG	+3.0%	Trailing avg. of UST rate + 100 bps
30		2% p.a.				
31		4% p.a.				
32	-6.75%	0% p.a.			+1.65%	
33		2% p.a.				
34		4% p.a.				
35	-13.5%	0% p.a.			+0.3%	
36		2% p.a.				
37		4% p.a.				
41	0.0%	0% p.a.	-50 bps	Mean path in Academy IRG with -100 bps MRP shock	+3.0%	Trailing avg. of UST rate + 100 bps
42		2% p.a.				
43		4% p.a.				
44	-6.75%	0% p.a.			+1.65%	
45		2% p.a.				
46		4% p.a.				
47	-13.5%	0% p.a.			+0.3%	
48		2% p.a.				
49		4% p.a.				

Scenarios corresponding to specific percentiles of stochastic scenarios

These scenarios are based on specific percentiles of cumulative returns from the sets of 1,000 stochastic scenarios used in Projection Set 4 of Testing Cycle 1.

Similar to the prior set of scenarios, these scenarios aim to determine market paths that produce funding requirements equivalent to CTE 70 in the stochastic calculation. A range of percentiles and interest rate paths are tested to help determine the market paths that most closely align with the CTE 70 level of funding requirement for each participant.

Details of these scenarios are provided in the following table:

#	*Eq. stress & recovery*	*IR stress*	*IR recovery*	*Bond stress & recovery*
14	5th percentile cumulative returns from stochastic	0 bps	Mean path in Academy IRG	5th percentile cumulative returns from stochastic
15	10th percentile cumulative returns from stochastic			10th percentile cumulative returns from stochastic
16	15th percentile cumulative returns from stochastic			15th percentile cumulative returns from stochastic
26	5th percentile cumulative returns from stochastic	0 bps	Mean path in Academy IRG with -100 bps MRP shock	5th percentile cumulative returns from stochastic
27	10th percentile cumulative returns from stochastic			10th percentile cumulative returns from stochastic
28	15th percentile cumulative returns from stochastic			15th percentile cumulative returns from stochastic
38	5th percentile cumulative returns from stochastic	-50 bps	Mean path in Academy IRG	5th percentile cumulative returns from stochastic
39	10th percentile cumulative returns from stochastic			10th percentile cumulative returns from stochastic
40	15th percentile cumulative returns from stochastic			15th percentile cumulative returns from stochastic
50	5th percentile cumulative returns from stochastic	-50 bps	Mean path in Academy IRG with -100 bps MRP shock	5th percentile cumulative returns from stochastic
51	10th percentile cumulative returns from stochastic			10th percentile cumulative returns from stochastic
52	15th percentile cumulative returns from stochastic			15th percentile cumulative returns from stochastic

Scenario file format

Each scenario file provided has the following fields:

Field	Description
SWAP	USD swap rates, quoted as par rates
T_SWAP_SPREAD	Spread between par USD swap rates and corresponding par US Treasury rates
US_EQ	Gross Wealth Ratio of equity funds
BOND	Gross Wealth Ratio of bond funds
MM	Gross Wealth Ratio of money market funds
VOL	At-the-money-forward implied volatility on the SPX index

As in Testing Cycle 1, for calculating returns of balanced and volatility-control funds, participants should reflect the equity and bond allocations as of the valuation date and any expected asset rebalancing in the projection consistent with fund operations.

5.3. Results and recommendations

5.3.1. Data received

In total, twelve companies submitted testing materially error-free testing results for Projection Sets 11-13 that were ultimately used in the Testing Cycle 2 presentation, while seven companies submitted testing results for Projection Sets 17-18.

Prior to the presentation of Testing Cycle 2 results, Oliver Wyman shared with each individual participant its own company-specific results and went through the same data validation procedures as those from Testing Cycle 1, as outlined in Section 4.2.1. If a participant did not request revisions to – or retractions of – the company-specific results, Oliver Wyman merged the company's results into the industry-aggregate results to produce the eventual presentation.

5.3.2. Analytical methodology

Overall, the specifications and analytical methodology used by Oliver Wyman Projection Sets 11-13 in Testing Cycle 2 aimed to ensure that the existing Standard Scenario construct within AG 43, after the recommended revisions, would be able to meet the VAIWG's stated purpose for the Standard Scenario. In April 2017, the VAIWG articulated that the purpose of the Standard Scenario is to govern actuarial assumptions, modeling choices, and other components where the CTE calculation incorporates material company judgment – but not the stringency of the capital markets scenarios used in the CTE calculation, which was to be governed separately within the CTE calculation. Additional details of the VAIWG discussions around the purpose and construct of the Standard Scenario are outlined in Section 9.6 of this document.

Oliver Wyman interpreted the VAIWG's guidance on the purpose of the Standard Scenario to indicate that for effective governance of companies' actuarial assumptions and modeling choices, the Standard Scenario Amount should exceed the CTE Amount if and only if:

i. A company uses assumptions or practices that substantially deviate from industry experience or accepted practices (unless such deviations can be justified by the company); and

ii. Such deviations result in materially-lower CTE-based reserve.

These objectives imply that if a company uses the same actuarial assumptions and modeling choices between the Standard Scenario calculation and the CTE calculation, then the company should see a Standard Scenario Amount similar in magnitude to the CTE Amount under all market conditions and portfolio compositions, as:

i. If the Standard Scenario Amount is materially below the CTE Amount, it may not necessitate reserve strengthening even if a company's assumptions in the CTE calculation were materially deviant;

ii. If the Standard Scenario Amount materially exceeds the CTE Amount, it would needlessly force a company with reasonable assumptions in the CTE calculation to hold an additional "redundant" reserve.

As such, results of Projection Sets 11-13 were assessed primarily to determine whether this particular target property was achieved under the revised Standard Scenario construct and any of the 52 alternative market paths provided by Oliver Wyman.

For each participant, the result for each of the 52 alternative market paths was evaluated individually and not aggregated in any fashion with the results from the other 51 market paths. In the evaluation of each market path, the Standard Scenario Amount was compared against the participant's CTE Amount under the same Valuation Date market conditions from Projection Set 4 of Testing Cycle 1 to determine an "equivalent CTE level" – defined as the CTE level at which the CTE Amount from Testing Cycle 1 would be equal to the Standard Scenario Amount. Subsequently, we plotted the distribution of "equivalent CTE levels" for each alternative market path across all participants on histograms, with the expectation that if the target property were met:

i. Most, if not all, participants would have the equivalent CTE level close to CTE 70; and

ii. For each participant, the equivalent CTE level would remain relatively stable – i.e., close to CTE 70 – across all tested Valuation Date market conditions.

The distributions of "equivalent CTE levels" were plotted for each of the three different methodologies for reflecting CDHS that were tested in Projection Sets 11-13. **Figure 7** and **Figure 8** illustrate the exhibits – filled out with dummy data – that were used in the Testing Cycle 2 presentation to present these results.

Results of Projection Sets 17-18 were used to assess the impact of alternative methods for reflecting non-guaranteed revenue sharing income. Specifically, Oliver Wyman evaluated the aggregate baseline funding requirement – defined in the same manner as that used in Testing Cycle 1, as outlined in Section 4.2.2 – across participants under both (i) the methodology under the existing AG 43 guidance and (ii) the alternative methodologies proposed via Projection Sets 17-18. In doing so, we also made the same simplifications as those that we made in Testing Cycle 1 – i.e., assuming that each participant was a mono-line VA company with no voluntary reserve or captive usage.

Figure 7: exhibits used in the Testing Cycle 2 presentation – distribution of equivalent CTE levels (illustrative data only)

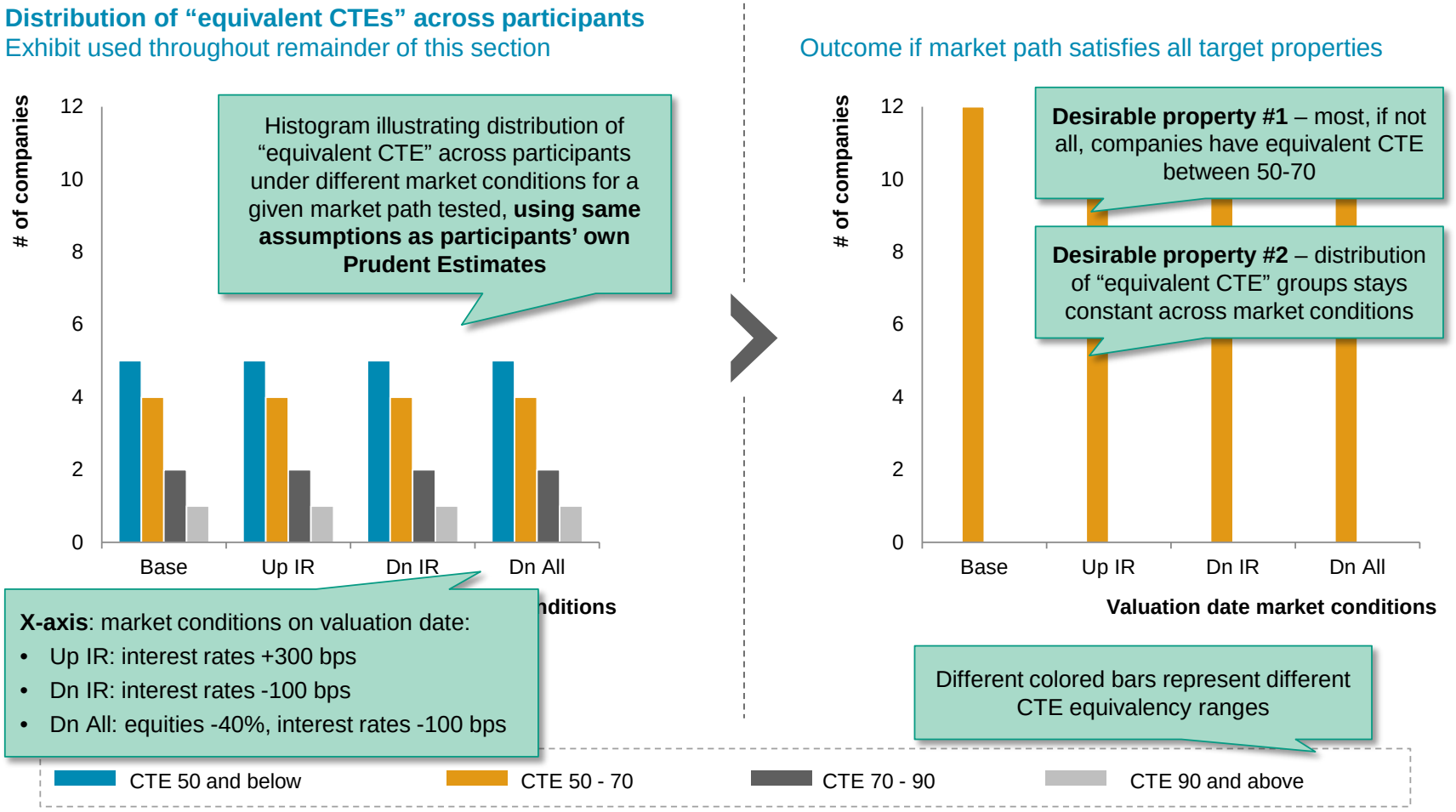


Figure 7 illustrates the exhibits that were used in the Testing Cycle 2 presentation to evaluate the distribution of equivalent CTE levels across the participant group and different Valuation Date market conditions for each of the 52 alternative market paths tested. In addition, **Figure 7** also indicates the desirable properties of the tested recommendations, representing the criteria by which Oliver Wyman evaluated the effectiveness of the tested recommendations.

Figure 8: exhibits used in the Testing Cycle 2 presentation – impact of alternative CDHS reflection methodologies (illustrative data only)

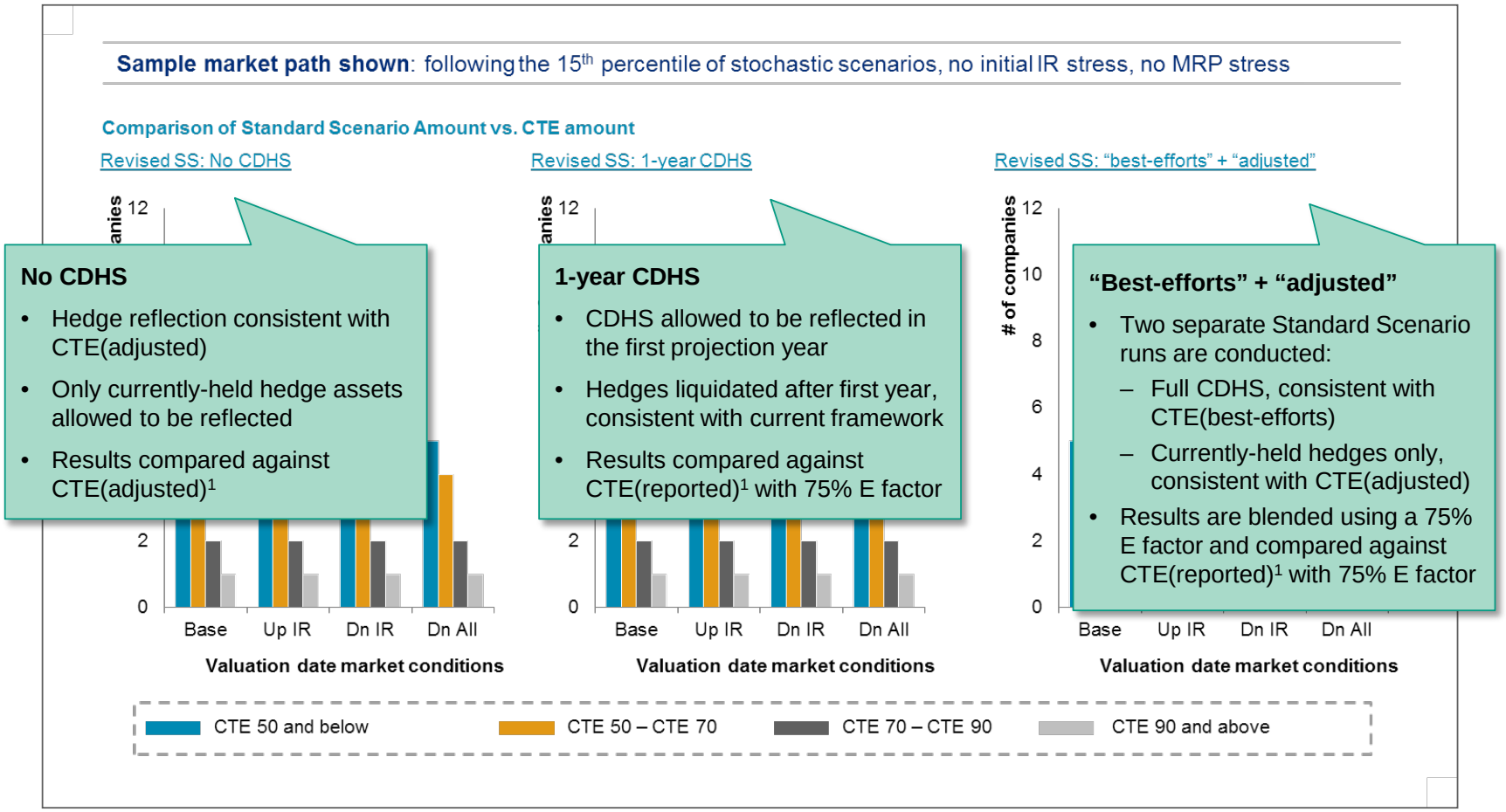


Figure 8 illustrates the exhibits that were used in the Testing Cycle 2 presentation to evaluate the impact of different methodologies for reflecting CDHS within the Standard Scenario. For each CDHS reflection methodology, Oliver Wyman plotted the distribution of equivalent CTE levels across the participant group and different Valuation Date market conditions in the same manner as that illustrated in **Figure 7**.

5.3.3. Discussion of results

Results from all 52 alternative market paths tested in Projection Sets 11-13 demonstrated that none of the tested combinations of hedging reflection methodology and alternative market path was able to satisfy the target property being assessed – i.e.,

> *If a company uses the same actuarial assumptions and modeling choices between the Standard Scenario calculation and the CTE calculation, then the company should see a Standard Scenario Amount similar in magnitude to the CTE Amount under all market conditions and portfolio compositions.*

Specifically, the results indicated that no standardized market path could reliably produce a Standard Scenario Amount that is in-line with, and stable with respect to, the CTE Amount for the same portfolio and under the same actuarial assumptions. Notably:

i. For each standardized market path and Valuation Date market condition tested, we observed significant dispersion in the "equivalent CTE level" across participants – with some participants seeing equivalent CTE levels below CTE 50 while others seeing equivalent CTE levels exceeding CTE 90;

ii. For each standardized market path tested and each participant, we observed significant instability in the "equivalent CTE level" across different Valuation Date market conditions.

These results demonstrated that the VAIWG's stated purpose for the Standard Scenario could not be achieved by using a single, standardized market path given the diversity of portfolio compositions and risk profiles, both across the industry and across different market conditions, a critical flaw in the current framework. Indeed, theoretically and in the extreme, a CTE 70-equivalent market path for traditional VA contracts with heavy separate account equity investments may not be at all adverse for VA contracts with heavy separate account fixed income investments, or some modern VA contracts with "up-equity risk" – i.e., call option-like guarantees.

Based on these results, Oliver Wyman noted to the VAIWG in the Testing Cycle 2 presentation that the Standard Scenario framework needs to be more flexible and allow specific calibration to each company's portfolio for it to achieve the VAIWG's stated purpose. We subsequently presented two alternative proposals for a revised Standard Scenario framework that had the following features:

Proposal A:

i. The Standard Scenario Amount would be calculated as the GPVAD along a single market path under prescribed actuarial and expense assumptions, with the GPVAD calculation following Proposal 2A from Oliver Wyman's 2016 recommendations and incorporating all other relevant additional recommendations concerning the GPVAD at the conclusion of QIS II;

ii. The reflection of hedging along the single market path should be identical to that used in each company's calculation of CTE ("adjusted") – i.e., either reflecting only expiration of currently-held hedge assets, or permitting all currently-held hedge assets to be replaced by other general account assets prior to the commencement of projection;

iii. The single market path would have two components – an initial stress that can be calibrated individually by each company, and a recovery path that is standardized across the industry;

iv. The calibration of the initial stress in the single market path would be conducted in such a manner that the resultant Standard Scenario Amount, if calculated using the same actuarial assumptions and modeling choices as the CTE calculation, would have an "equivalent CTE level" of CTE 65.

Proposal B:

i. The Standard Scenario Amount would be calculated as the CTE Amount under prescribed actuarial and expense assumptions, with the calculation incorporating all other relevant additional recommendations concerning the CTE Amount at the conclusion of QIS II;

ii. The reflection of hedging in the Standard Scenario calculation should be identical to that used in each company's calculation of CTE ("adjusted") – i.e., either reflecting only expiration of currently-held hedge assets, or permitting all currently-held hedge assets to be replaced by other general account assets prior to the commencement of projection;

iii. The "Standard Scenario CTE Amount" would be calculated at a lower confidence level such that if it were calculated using the same actuarial assumptions and modeling choices as the CTE Amount, it would have an "equivalent CTE level" of CTE 65.

The establishment of CTE 65 as Oliver Wyman's target "equivalent CTE level" for both proposals reflected the VAIWG's articulation of the desired stringency of the prescribed policyholder behavioral assumptions within the Standard Scenario. Specifically, as discussed in Section 9.6, the VAIWG expressed a desire only to "catch outliers" in the establishment of company-defined model assumptions and choices via the prescribed Standard Scenario assumptions. Oliver Wyman interpreted this statement to indicate that regulators are willing to give "the benefit of the doubt" to companies when it pertains to company-defined assumptions and model choices that deviate modestly, but not meaningfully, from the prescribed assumptions.

Because the prescribed behavioral assumptions are numerous and have complex interaction effects with each other, Oliver Wyman decided to calibrate the prescribed assumptions to observed industry experience without significant margins in either direction – as outlined in Section 7. Instead, Oliver Wyman decided to capture the VAIWG's articulation by adjusting the implied target CTE confidence level for the Standard Scenario – i.e., within a CTE-based solvency framework, the VAIWG's articulation indicates that the Standard Scenario's confidence level should be below CTE 70 when calculated using the prescribed assumptions. This would effectively establish a reserve "buffer" equal to the difference in stochastic reserves between the chosen CTE level and CTE 70.

Oliver Wyman selected CTE 65 as the target confidence level for the Standard Scenario Amount, as it had value as a regulatory precedent given that CTE 65 was an initial proposal for the AG 43 CTE confidence level before being later revised to CTE 70. Subsequent testing in Testing Cycle 3 – as discussed in Section 6 – also indicated that this buffer was prudent in size, never exceeding for any QIS II participant 2% of its cash surrender value under any tested market conditions.

At the conclusion of the Testing Cycle 2 presentation, the VAIWG expressed openness to further development and testing of both proposals; accordingly, both proposals were tested in Testing Cycle 3 via Projection Sets 26-27 and 34, respectively.

Results from Projection Sets 17-18 on alternative methodologies for reflecting non-guaranteed revenue sharing income were as expected. Both alternative methodologies for reflecting revenue sharing resulted in modestly lower statutory reserves, C3 charges, and total baseline funding requirements, with the reductions being similar between the two methodologies. At the same time, it was also clear that the magnitude of the aggregate impact was substantially lower than those observed for other recommendations in Testing Cycle 1.

6. Cycle III testing

Testing Cycle 3 aimed to provide insight into the expected impact of a consolidated set of framework revisions that Oliver Wyman expected to be potential candidates to be included in the final recommendations, taking into account findings and VAIWG guidance obtained through Testing Cycles 1 and 2. Note that certain framework revision candidates, such as alternative reflections of revenue sharing, were not tested in Testing Cycle 3, as Oliver Wyman deemed these revisions to have been sufficiently tested in the prior Testing Cycles in light of their intuitiveness or linearity in valuation impact.

More specifically, Cycle 3 evaluated three items listed below:

i. **CTE Amount under alternative equity scenarios**: impact of alternative equity scenarios, calibrated using a longer span of US equity market history, on the CTE Amount for both statutory reserve and C3 calculations. This assessment incorporated all of the previously-tested Oliver Wyman recommendations from Testing Cycle 1;

ii. **Proposal A of the alternative Standard Scenario construct**: appropriateness and impact of "Proposal A" of the alternative Standard Scenario construct that was discussed during the Testing Cycle 2 presentation, as outlined in Section 5.3.3;

iii. **Proposal B of the alternative Standard Scenario construct**: appropriateness and impact of "Proposal B" of the alternative Standard Scenario construct that was discussed during the Testing Cycle 2 presentation, as outlined in Section 5.3.3;

6.1. Testing specifications

Testing Cycle 3 comprised five primary Projection Sets, as outlined below in **Table 7**.

Originally, Testing Cycle 3 planned for eleven Projection Sets. However, six Projection Sets were subsequently removed from consideration, as these Projection Sets:

i. Represented additional tests of the same Standard Scenario construct as that which was tested extensively in Testing Cycle 2 and have been demonstrated to have low probability of achieving the VAIWG's stated objectives for the Standard Scenario;

ii. Received, as a result, few submissions of results from participants.

Accordingly, the remainder of Section 6 only discusses five Projection Sets, as the other Projection Sets were not used in the eventual presentation of Cycle 3 results to the VAIWG and QIS II participants.

Table 7: Projection Sets requested in Testing Cycle 3

Projection Set	description	Output
24	CTE Amount under the revised AG 43 recommended by Oliver Wyman and tested in Projection Set 4, with an alternative set of equity calibration criteria – but no other changes from Projection Set 4	Stochastic CTE
25	Fair value of options and guarantees valued on a market-consistent basis and CTE Amount of all non-guarantee cash flows, valued under the revised AG 43 recommended by Oliver Wyman in Projection Set 24	Stochastic CTE and fair value of options and guarantees
26	Standard Scenario Amount calculated with company-specific market paths; this Projection Set is meant to identify the company-specific market path for each participant and therefore uses participants' own Prudent Estimate assumptions	Standard Scenario

27	Standard Scenario Amount calculated with company-specific market paths; this Projection Set uses Oliver Wyman's recommended prescribed assumptions	Standard Scenario
34	Standard Scenario Amount determined as the stochastic CTE Amount, but calculated using Oliver Wyman's recommended prescribed assumptions	Stochastic CTE

Participants were asked to conduct each Projection Set with a specific combination of in-force portfolio composition and assumed starting market conditions, as detailed in **Table 8**. For nomenclature in the remainder of this section, the [X] indicated in **Table 8** should be replaced with the number of the Projection Set – e.g., Projection 24.1 would indicate a calculation conducted using the methodologies prescribed by Projection Set 24, applied to the Baseline Portfolio under Baseline Conditions.

Table 8: in-force portfolio compositions and assumed starting market conditions to be used

Projection	Portfolio	Market conditions
[X].1	Baseline Portfolio on a Recaptured Basis	Baseline Conditions
[X].2	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *+3.0% Rate Stress*
[X].3	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-1.0% Rate Stress*
[X].4	Baseline Portfolio on a Recaptured Basis	Stress Conditions • *-40% Equity Stress* • *-1.0% Rate Stress*

Note that sections 6.1.1 to 6.1.5 are replicated from the actual *Technical Specifications* document published for QIS II participants, with select non-applicable sections – e.g., Projection Sets that were ultimately not tested – removed. In addition, Oliver Wyman removed from the replicated text in these sections expositions of the purpose of each Projection Set, as the evaluations of the results from each Projection Set – as well as their implications – are covered in greater detail in section 6.2.

For ease of reference, all paragraphs included in the subsequent sub-sections are numbered.

6.1.1. Projection Set 24 – Revised Stochastic (Alternative equity calibration criteria)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.
2. **Calculations**: participating companies are requested to re-conduct all of the calculations in Projection Set 4 in Testing Cycle 1, with one modification to the redlined AG 43: in section A5.3), replace all calibration Gross Wealth Ratios for S&P 500 Total Returns with the following values:

Calibration Point	One Year	Five Year	Ten Year	Twenty Year
2.5%	72%	62%	57%	
5.0%	79%	71%	69%	114%
10.0%	87%	84%	94%	144%
90.0%	133%	229%	420%	1103%
95.0%	141%	263%	515%	1475%
97.5%	145%	304%	599%	

These equity calibration criteria were derived by recalibrating the American Academy of Actuaries' economic scenario generator to incorporate historical data from 1928 to 2016, as the current generator uses only historical data from 1955 to 2003. Specifically, the following steps were undertaken to recalibrate the economic scenario generator:

i. Daily S&P 500 Total Return ("SPTR") index data was obtained from Bloomberg from 12/30/1927 to 12/30/2016 via TOT_RETURN_INDEX_GROSS_DVDS; note that reliable daily SPTR data before 1927 were unable to be sourced;

ii. Daily SPTR log-returns were calculated and used to construct a monthly time series of realized volatility – i.e., for month t, the realized volatility was calculated as the annualized sample standard deviation of all log-returns in month t;

iii. Daily STPR log-returns were also used to construct a monthly time series of realized monthly log-returns – i.e., for month t, the monthly realized log-return was calculated as the sum of the daily log-returns;

iv. The three volatility parameters – i.e., (ϕ, τ, v) – in the Academy generator were re-estimated by applying Maximum Likelihood Estimation ("MLE") to the monthly time series of realized volatility;

v. The drift parameter A, representing the long-term risk-free return, was kept at 0.055;

vi. The correlation parameter ρ between the Standard Normal driving the stochasticity of the log-volatility process and that driving the stochasticity of the log-return process was re-estimated by applying MLE to the monthly time series of realized log-returns, without applying any constraints on the overall mean annual returns;

vii. The drift parameter C was kept at -0.9, while the drift parameter B was modified slightly such that the overall mean annual return remained at ~8.75%. Note that no re-estimation of these parameters were attempted; accordingly, the drift parameters A, B, and C for this Projection Set are very similar to those in the current Academy generator.

The following recalibrated parameters were subsequently used in the 2015.03 version of the Academy generator to produce the set of statutory projection scenarios to be used for this Projection Set:

τ	ϕ	σ_v	ρ	A	B	C
0.1332	0.2621	0.3538	-0.2670	0.0550	0.5800	-0.9000

Companies are requested to use the sets of statutory projection scenarios calibrated by Oliver Wyman and provided to each participating company in a separately-attached file. While Oliver Wyman does not anticipate revising the 2016 recommendation that allows companies to generate equity scenarios using proprietary generators as long as the scenarios follow the Gross Wealth Ratios outlined in section A5.3) of the redlined

AG 43, we have decided to provide standardized sets of statutory scenarios for testing purposes to allow more consistent and informative analysis of results.

Unlike Projection Set 4 in Testing Cycle 1, this Projection Set does not require participants to perform three iterations of the calculation with different Starting Asset Amounts; instead, participants should follow section A1.4)A) of the redlined AG 43 to set the Starting Asset Amount. In addition, in section A1.4)A) of the redlined AG 43, companies should also remove the requirement that the Starting Asset Amount over the aggregate Cash Surrender Value of all contracts included in the projection be less than 98% or greater than 102% of the excess of the Conditional Tail Expectation ("CTE") Amount over the aggregate Cash Surrender Value of the same contracts.

6.1.2. Projection Set 25 – FVOG (Alternative equity calibration criteria)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.
2. **Calculations**: participating companies are requested to re-conduct all of the calculations in Projection Set 3 of Testing Cycle 1, but using the same calibration Gross Wealth Ratios for S&P 500 Total Returns as those used in Projection Set 24 of Testing Cycle 3.

   To this end, participants should use the same set of statutory projection scenarios – calibrated and provided by Oliver Wyman in a separately-attached file – as those used in Projection Set 24. While Oliver Wyman does not anticipate revising the 2016 recommendation that allows companies to generate equity scenarios using proprietary generators as long as the scenarios follow the Gross Wealth Ratios outlined in section A5.3) of the redlined AG 43, we have decided to provide standardized sets of statutory scenarios for testing purposes to allow more consistent and informative analysis of results.

6.1.3. Projection Set 26 – CTE 65-equivalent Standard Scenario (own assumptions)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.
2. **Calculations**: participating companies are requested to conduct the full AG 43 Standard Scenario calculation using the methodology outlined in the redlined version of AG 43 that Oliver Wyman provided on August 7, 2017 under the file name *20170807 Revised AG 43 SS – for QIS II Cycle 3.pdf*, with the following modifications:

   i. Replace section A3.2)E)1) with the following text in italics:

   To determine the Standard Scenario Market Path, the company should calculate the Scenario Greatest Present Value for nine scenarios with the following separate account equity fund returns, gross of any fees chargeable to the separate account value:

Scenario	First year	Subsequent years
Scenario 1	+10.0%	+3.0%
Scenario 2	+5.0%	+3.0%
Scenario 3	+0.0%	+3.0%
Scenario 4	-5.0%	+3.0%
Scenario 5	-10.0%	+3.0%
Scenario 6	-15.0%	+3.0%
Scenario 7	-20.0%	+3.0%
Scenario 8	-25.0%	+3.0%

Scenario 9	-30.0%	+3.0%

Other market path specifications are identical across the nine scenarios and are as follows:

(i) US Constant Maturity Treasury rates shall follow the mean interest rate path from the prescribed interest rate scenario generator with prescribed parameters described in section A5.2). The mean interest rate path is the interest rate path that would be generated from the prescribed interest rate scenario generator should the realized interest rate volatility be set to zero in all projection time periods;

(ii) All separate account bond fund returns, gross of any fees chargeable to the separate account value, in the first projection year shall equal 20% of the equity fund return in the first projection year;

(iii) All separate account bond fund returns, gross of any fees chargeable to the separate account value, shall follow the trailing 6-year average of the projected 6-year US Constant Maturity Treasury rate plus the effect of 100 basis points of earned spread – i.e., net of default costs – per annum, from the second projection year until the end of the projection;

(iv) All other indices and returns shall follow the instructions outlined in section A3.2)E)2) for all projection periods, including the first projection year.

In calculating Scenario Greatest Present Values across the nine scenarios for the for the purpose of determining the Standard Scenario Market Path, the actuary shall replace the prescriptions in section A3.2)F) with the company's own Prudent Estimate assumptions.

The Standard Scenario Market Path shall be the scenario for which the Scenario Greatest Present Value is the closest to, but not in excess of, the 65 CTE Amount(adjusted). The 65 CTE Amount(adjusted) shall be calculated in the same manner as the 70 CTE Amount(adjusted) as defined in section A7.3), with the sole exception that the 65 CTE Amount shall be equal to the numerical average of the 35 percent – instead of 30 percent – largest values of the Scenario Greatest Present Values.

ii. Replace the first three paragraphs of A3.2)D) with the following text:

Modeled Hedges. Throughout the projection, the costs and benefits of hedging shall be projected in the same manner as that used in the calculation of the CTE Amount (adjusted) as discussed in section A7.3). However, the projection shall not take into account the hedge positions expected to be held in the future through the execution of the company's hedging strategy.

iii. Remove section A3.2)F). Instead, for projecting maintenance expenses and policyholder behavior covered by section A3.2)F), each participant should use the Prudent Estimate assumptions that it uses for conducting the stochastic AG 43 calculations to calculate the Conditional Tail Expectation Amount. If such assumptions are appropriate only for a stochastic calculation – e.g., random sampling techniques across scenarios, the participant may elect to adopt an alternative modeling technique to implement the assumption provided that the participant informs Oliver Wyman of the chosen alternative modeling technique.

Please note that:

i. The modifications outlined above to section A3.2)E)1) represents the potential Standard Scenario market path construct that Oliver Wyman would like to test. For greater consistency in testing, participants are requested to use a uniform set of scenario files provided by Oliver Wyman;

ii. For each participant, Oliver Wyman will subsequently compare the Scenario Greatest Present Value for each provided market path against the participant's CTE Amount results from Projection Set 4 – i.e., under the current equity calibration criteria – and Projection Set 24. Accordingly, Oliver Wyman

would be able to identify the market path that satisfies the desired property of being close to the 65 CTE Amount(adjusted) under both the current equity calibration criteria and the alternative equity calibration criteria tested in Projection Set 24.

As outlined in section 3.3, participants should submit results for each requested market path separately as instructed in the separately-attached output template.

6.1.4. Projection Set 27 – CTE 65-equivalent Standard Scenario (prescribed assumptions)

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.
2. **Calculations**: participating companies are requested to re-conduct all of the calculations in Projection Set 26, but retain section A3.2)F) in the file *20170807 Revised AG 43 SS – for QIS II Cycle 3.pdf*. As such, for projecting maintenance expenses and policyholder behavior, participants should use the prescribed assumptions outlined in the file *20170807 Revised AG 43 SS – for QIS II Cycle 3.pdf*.

   As outlined in section 3.3, participants should submit results for each requested market path separately as instructed in the separately-attached output template.

6.1.5. Projection Set 34 – Stochastic under Prescribed Assumptions

1. **Definitions**: this Projection Set does not require any additional definitions beyond those outlined in Section 3.
2. **Calculations**: participating companies are requested to re-conduct the "adjusted" run and the "unhedged" run that excludes the impact from hedging – including that from the run-off of currently-held hedge assets – in Projection Set 4 in Testing Cycle 1, with one modification to the redlined AG 43: in section III)B)7), the definition of "Prudent Estimate" should be changed to refer to section A3.2)F) in the file *20170918 Revised AG 43 SS – For QIS II Cycle 3.pdf*. As such, for projecting maintenance expenses and policyholder behavior, participants should use the prescribed assumptions outlined in the file *20170918 Revised AG 43 SS – For QIS II Cycle 3.pdf*.

   To maintain a tractable run-time, participants may adopt random sampling techniques – either across scenarios or across policies – to implement the Withdrawal Delay Cohort Method as outlined in section A3.2)F)6) of the file *20170918 Revised AG 43 SS – For QIS II Cycle 3.pdf*. As output, participants should submit scenario-specific GPVADs for each run in the same template as that used for Projection Sets 4 and 24.

   We request that this Projection Set be conducted across all four starting market conditions as outlined in Table 10 of the Technical Specifications, though we note that the Projection with Baseline Conditions is the highest priority. Please also note that submission of this Projection Set is entirely voluntary.

6.2. Results and recommendations

6.2.1. Data received

In total, eleven companies submitted Projection Set 24 testing results that were ultimately used in the Testing Cycle 3 presentation. Among the eleven companies, five submitted results for Projection Set 25. Similar to in Testing Cycle 1, results for Projection Set 25 enabled Oliver Wyman to conduct additional analyses of the expected impact of framework revisions assuming that the submitting participant implemented a hedge program focused on immunizing the guaranteed benefit liability on a fair value basis.

Nine companies submitted testing results for Projection Sets 26 and 27, which evaluated the alternative Standard Scenario construct that was discussed as "Proposal A" during the Testing Cycle 2 presentation as outlined in Section 5.3.3. Of the nine companies, six submitted results for Projection Set 34, which evaluated the alternative Standard Scenario construct that was discussed as "Proposal B", as outlined in Section 5.3.3.

Prior to the presentation of Testing Cycle 3 results, Oliver Wyman shared with each individual participant its own company-specific results and went through the same data validation procedures as those from Testing Cycle 1, as outlined in Section 4.2.1. If a participant did not request revisions to – or retractions of – the company-specific results, Oliver Wyman merged the company's results into the industry-aggregate results to produce the eventual presentation.

6.2.2. Analytical methodology

Testing Cycle 3 conducted the following analyses on the testing results collected:

i. **CTE Amount properties**: results from Projection Sets 24-25 were evaluated in conjunction with results from Testing Cycle 1 via the same five lenses as outlined in Section 4.2.2 in order to assess whether the recalibration of the equity scenarios materially changed the desirable properties of the revised framework that were observed in Testing Cycle 1;

ii. **CTE Amount confidence levels**: with the results from Projection Sets 24-25 and Testing Cycle 1, a number of different CTE confidence levels were evaluated in calculating the tail CTE Amount for determining the C3 charge under Proposal 3B of Oliver Wyman's 2016 recommendations. Specifically, we assessed the impact of hedging on companies' total funding requirements under different tail CTE confidence levels, as well as the impact of the equity scenario recalibration on the CTE Amount at different confidence levels;

iii. **Standard Projection Amount – Proposal A**: results from Projection Sets 26-27 were evaluated to identify the number of participants for whom the Standard Scenario construct being tested – i.e., Proposal A outlined in Section 5.3.3 – would lead to an increase in reserves, the aggregate magnitude of such an increase, and the stability of the additional reserves across different Valuation Date market conditions;

iv. **Standard Projection Amount – Proposal B**: results from Projection Sets 34 were compared against results from Projection Set 26-27 to assess whether the Standard Scenario construct described as Proposal B in Section 5.3.3 would lead to materially different reserves from those obtained via Proposal A across different market conditions;

v. **Aggregate framework impact**: using results from all Projection Sets in Testing Cycles 1 and 3, all of Oliver Wyman's recommendations that were tested were evaluated simultaneously to determine both the aggregate impact on overall funding requirements, as well as whether the desirable properties of the revised framework that were observed in Testing Cycle 1 were maintained.

Figure 9 and **Figure 10** illustrate the exhibits – filled out with dummy data – that were used in the Testing Cycle 3 presentation to present these results. Note that in the remainder of this section, we have elected to use the term Standard Projection Method to denote both alternative Standard Scenario constructs tested in order to remain consistent with the terminology of Oliver Wyman's ultimate redlined AG 43 and C3 Phase II guidance documents.

Consistent with Testing Cycles 1 and 2, for simplicity, Oliver Wyman assumed in these analyses that each participant was a mono-line VA company with no captive usage. However, unlike Testing Cycles 1 and 2, Oliver Wyman evaluated Testing Cycle 3 results under two separate assumptions regarding voluntary reserve usage:

i. Consistent with Testing Cycles 1 and 2, assume that no participant uses voluntary reserves under both the current statutory framework and the revised statutory framework being evaluated;

ii. Assume that no participant uses voluntary reserves under the revised statutory framework being evaluated, but that under the current statutory framework, all participants elect to post voluntary reserves in an amount sufficient to make the C3 charge zero.

Under both assumptions, we also assumed that each participant had a relatively small amount of high-quality general account assets with low C1 charges such that there was no meaningful diversification benefit between the C3 charge and other RBC charges in calculating the overall RBC requirement.

Consistent with Testing Cycle 1, for CTE-related results under the revised framework, Oliver Wyman assumed that all participants would use the same hedge "error factor" in weighting the CTE Amount (best-efforts) and CTE Amount (adjusted). However, we did not evaluate multiple "error factors" in Testing Cycle 3, as multiple "error factors" were already evaluated in Testing Cycle 1 to ensure the continued validity of the findings under different reasonable "error factors".

Separately, consistent with Testing Cycle 1, Oliver Wyman excluded the component that adjusts for the difference between the statutory and the tax reserves on the valuation date in the revised C3 charge calculation for most of the Testing Cycle 3 exhibits. However, in the analysis of the aggregate framework impact, Oliver Wyman included this component for as many companies as possible given the tax reserve information that we were able to collect in Testing Cycle 3. Similar to Testing Cycle 1, though Oliver Wyman requested tax reserve information for each Projection in each Projection Set, most participants were only able to provide tax reserves as of 2016YE and not under other Valuation Date market conditions that were tested.

We note that across all of these analyses, apart from the exceptions noted above and in Section 6.1, the specific calculations that were used to determine various quantities – e.g., the CTE Amount, the Additional Standard Projection Amount, and the C3 charge – were identical to those that were ultimately recommended in the redlined AG 43 guidance exposed on December 1, 2017.

Figure 9: exhibits used in the Testing Cycle 3 presentation – CTE properties and aggregate framework impact (illustrative data only)

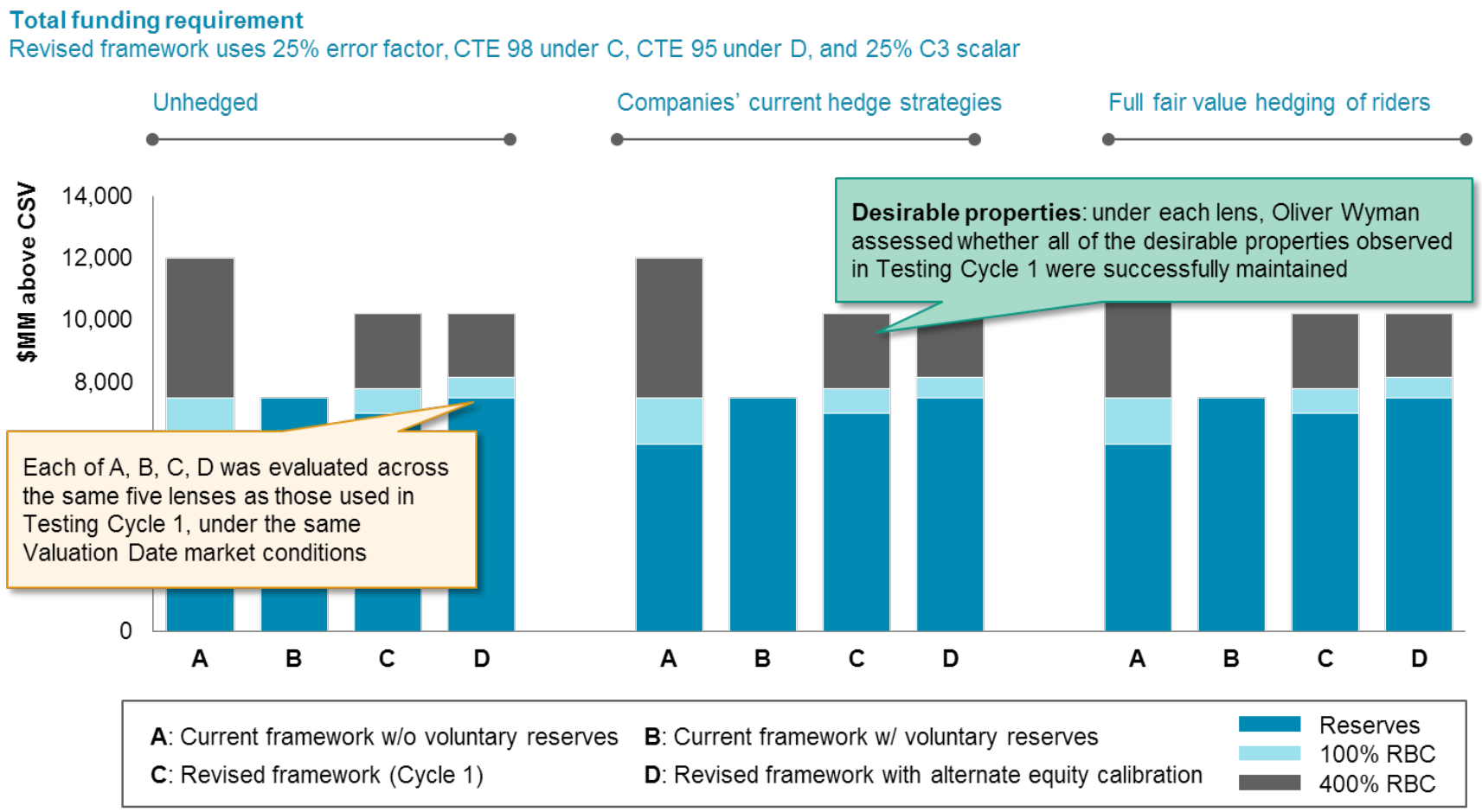


Figure 9 illustrates the exhibits that were used in the Testing Cycle 3 presentation to evaluate (i) properties of the CTE Amount and (ii) aggregate impact of all recommended framework revisions. Note that these exhibits were produced for each of the five evaluation lenses introduced in Testing Cycle 1 and outlined in Section 4.2.2.

The structures of these exhibits were intended to be identical to those illustrated by **Figure 2** to **Figure 6** and used in Testing Cycle 1, with the exception that under each hedge strategy grouping, two more sets of results were included – (i) the current framework with voluntary reserve usage sufficient to eliminate the C3 charge, and (ii) the revised framework calculated using the alternative equity scenarios in Projection Set 24.

Figure 10: exhibits used in the Testing Cycle 3 presentation – evaluation of the Standard Projection Method (illustrative data only)

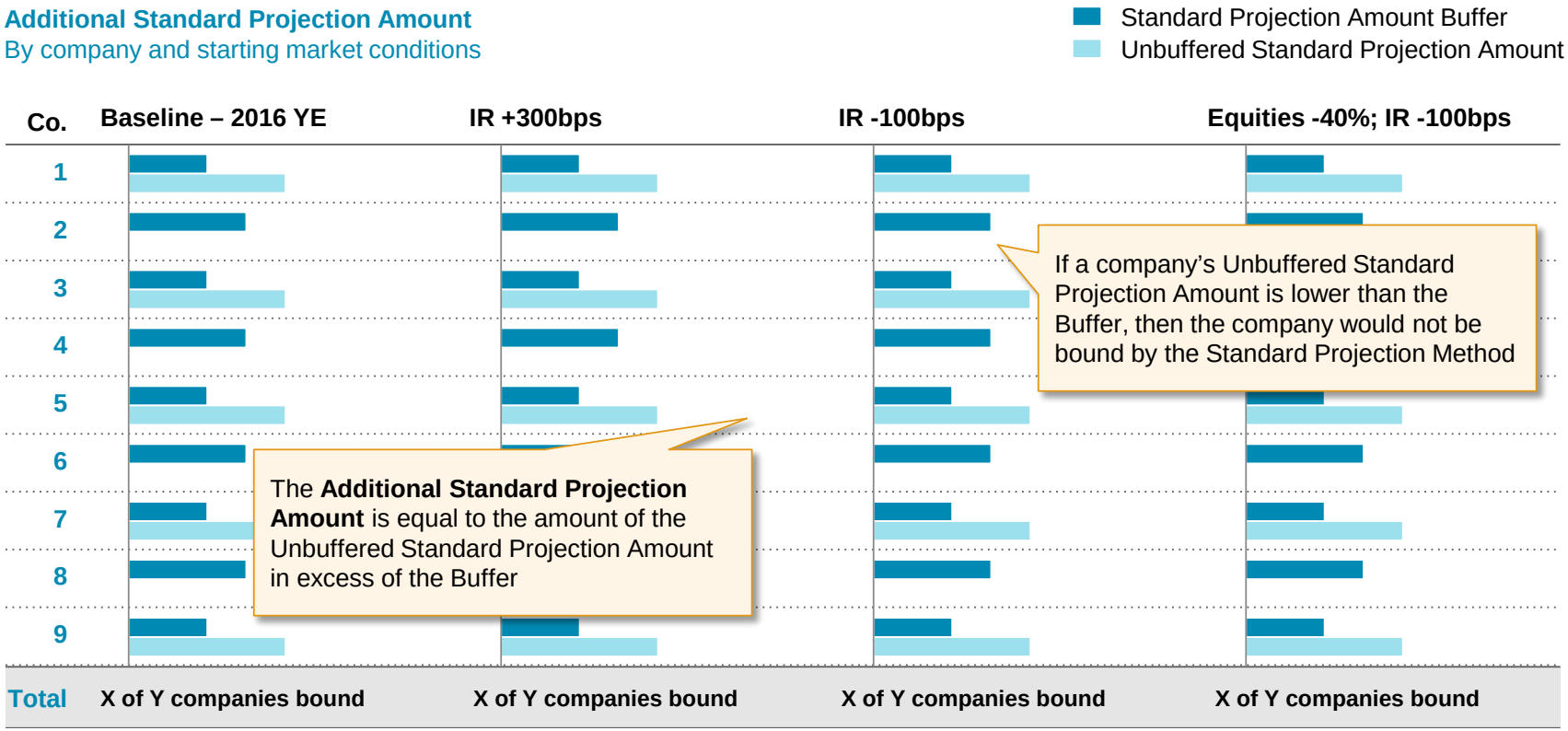


Figure 10 illustrates the exhibits that were used in the Testing Cycle 3 presentation to evaluate the results of each of the two proposals for the Standard Projection Method. Anonymized company-specific results were displayed in these exhibits to identify the number of companies "bound" by the Standard Projection Method – i.e., where the Additional Standard Projection Amount exceeded zero – under each Valuation Date market condition.

For each company and each Valuation Date market condition, two quantities were plotted – the Unbuffered Standard Projection Amount, and the Standard Projection Amount Buffer, both expressed as percentages of the company's cash surrender value. The difference between the two quantities – after being floored at zero – would therefore be the Additional Standard Projection Amount for that company.

6.2.3. Discussion of results

The following results were observed across the five different analyses conducted in Testing Cycle 3.

6.2.3.1. CTE Amount properties

Holding constant all other framework parameters, using the alternative equity scenarios that were tested in Projection Sets 24-25 increased participant-aggregate funding requirements. Of note, the funding requirement increase was driven by increases in both reserves and the C3 charge.

However, the market sensitivity of total surplus, the C3 charge, and excess surplus did not change materially from that which was observed in Testing Cycle 1. This indicated that changing the equity scenarios that are used in calculating the CTE Amount would still maintain the desirable properties of the revised framework that were observed in Testing Cycle 1. As such, Oliver Wyman concluded that the impact from altering the equity calibration criteria under the current AG 43 framework would only be on total funding requirements – not on total balance sheet volatility.

6.2.3.2. CTE Amount confidence levels

Based on the conclusions in Section 6.2.3.1, Oliver Wyman judged that the total funding requirement impact from altering the current equity calibration criteria may be adjusted by calibrating the C3 charge calculation under Proposal 3B of Oliver Wyman's 2016 recommendations – i.e.,

$$\begin{aligned}\text{C3} = \text{Scalar} \times \big(&(\text{Tail CTE Amount} - \text{Statutory Reserve}) \times (1 - \text{Federal Income Tax Rate}) \\ &- (\text{Statutory Reserve} - \text{Tax Reserve}) \times \text{Federal Income Tax Rate}\big)\end{aligned}$$

Note that for simplicity, the formula above assumes that the statutory reserve is equivalent to the AG 43 CTE Amount and does not have any Additional Standard Projection Amount.

Specifically, there are two parameters under Proposal 3B that may be calibrated:

i. The confidence level used to calculate the Tail CTE Amount;

ii. The scalar applied to the full C3 calculation.

Oliver Wyman ultimately decided to set the scalar applied to the full C3 calculation to 25%, as:

i. The scalar determines the level of voluntary reserves that can fully eliminate the required capital for market risk – i.e., the C3 charge;

ii. We believe that, to preserve the meaningfulness of the RBC ratio in judging the financial health of a company and to identify weakly-capitalized insurers, regulatory prudence should require the C3 charge to remain non-zero until a company's funding level for its VA portfolio is equivalent to a typical capitalization target for a healthy US life insurer;

iii. As a 400% Company Action Level RBC ratio is the typical target capitalization level for US life insurers today, setting the scalar at 25% would ensure that voluntary reserves would not eliminate the C3 charge until the funding level for the VA portfolio effectively matches a 400% RBC ratio.

The establishment of 25% as the recommended scalar for the C3 charge calculation left just one parameter to calibrate – the confidence level used to calculate the Tail CTE Amount. As discussed in Section 4.2.3, results from Testing Cycle 1 established that the parameterization of the Tail CTE Amount does not materially change the desirable properties on total balance sheet volatility attained via the revised framework. Hence, Oliver Wyman decided on the recommended Tail CTE Amount in the following manner:

i. **Primary decision criterion**: we judged that the Tail CTE Amount should be selected in a manner such that hedging reduces the total funding requirement of the portfolio. This particular decision criterion was proposed by Oliver Wyman during the VAIWG meeting on May 10 – as outlined in Section 9.1 – with no regulator objection and explicit support from select regulators.

   Accordingly, we used this decision criterion as the critical determinant of the Tail CTE Amount parameterization, though we noted in the Testing Cycle 3 presentation that the property may not be attainable in all market conditions – e.g., in conditions with particularly low interest rates.

ii. **Additional consideration**: as an additional consideration, we assessed the impact of different Tail CTE Amount confidence levels on the overall total funding requirements relative to both the current framework without usage of voluntary reserves and the current framework with voluntary reserve usage.

   We noted that if the VAIWG wished to keep the same level of conservatism in the revised VA reserve and capital framework as the current framework, the total funding requirement should be between (a) that under the current framework without any voluntary reserves; and (b) that under the current framework with voluntary reserves used in an amount that reduces the C3 charge to zero. Choosing a Tail CTE Amount that causes total funding requirements to stray outside these benchmarks would suggest a material shift in the level of conservatism embedded in the VA statutory framework.

In assessing confidence levels for the Tail CTE Amount against the primary decision criterion, we observed that:

i. The Tail CTE Amount confidence level that is required for hedging to reduce the total funding requirement at the 400% RBC level differed by company, driven by different portfolio compositions, risk profiles, actuarial assumptions, and modeling approaches for reflecting hedging;

ii. The Tail CTE Amount confidence level that is required for hedging to reduce the total funding requirement at the 400% RBC level differed by market conditions. A higher confidence level was needed in low interest rate conditions; indeed, in an interest rate environment represented by a 100 basis point downward parallel shift of the 2016YE market conditions, a confidence level in excess of CTE 99 was required for many participants;

   Thus, Oliver Wyman noted to the VAIWG that choosing the Tail CTE Amount confidence level requires a regulatory view of the "worst market condition in which hedging should reduce total funding requirement at the 400% RBC level". In addition, we expressed our view that setting this "threshold market condition" similar to 2016YE levels is reasonable, as it anchors to an interest rate environment similar to that experienced in recent years, but not at the most extreme end – e.g., June 2016;

iii. With a 25% Scalar, the Tail CTE Amount confidence level that is required for hedging to reduce the funding requirement at the 100% RBC level – i.e., the Total Asset Requirement – exceeds CTE 99 for most participants under all tested market conditions.

With these observations in mind, Oliver Wyman ultimately decided that under the primary decision criterion, the following confidence levels represented the most suitable choices for the Tail CTE Amount:

i. CTE 98 calculated with the equity calibration criteria under the current AG 43 guidance;

ii. CTE 95 calculated with the alternative equity scenarios tested in Projection Set 24.

At these confidence levels, we found that under any interest rate condition similar to or higher than 2016YE levels, hedging would reduce the total funding requirement at the 400% RBC level for most participants that submitted Projection Set 24-25 results. These confidence levels also represent points of diminishing marginal impact, where substantial additional increases in the confidence level would have been required for the remaining participants to observe total funding requirement reductions from hedging under 2016YE market conditions.

Subsequently, Oliver Wyman evaluated the choices of CTE 98 and CTE 95 against the additional consideration of the impact on total funding requirement relative to that under the current VA statutory framework – and specifically, relative to the two benchmarks of (a) the current framework without voluntary reserves and (b) the

current framework with voluntary reserves sufficient to reduce the C3 charge to zero. The major observations from this evaluation were that:

i. CTE 98 under the current equity calibration criteria and CTE 95 under the alternative equity calibration criteria both resulted in participant-aggregate funding requirements that resided between the two benchmarks across all tested market conditions;

ii. These observations were maintained both when assuming no hedging from any participant and under companies' current hedge programs.

We then turned our attention to the impact of the selected confidence levels on the Total Asset Requirement – i.e., 100% RBC level, which yielded the observations outlined in **Table 9**.

Table 9: impact of CTE Amount confidence levels on Total Asset Requirement – i.e., 100% RBC level

Hedging	CTE 98 with current equity scenarios	CTE 95 with alternative equity scenarios
No hedging	• Moderately lower than current framework across all tested market conditions	• Moderately higher than current framework under 2016YE market conditions • Moderately lower than current framework in the joint equity-interest rate stress tested
Companies' current hedge strategies	• Similar to current framework under 2016YE market conditions • Substantially lower than current framework in the joint equity-interest rate stress tested	• Higher than the current framework under 2016YE market conditions • Similar to current framework in the joint equity-interest rate stress tested

Based on the set of observations discussed in this section, Oliver Wyman judged that the chosen confidence levels for the Tail CTE Amount under both the current equity calibration criteria and the alternative equity scenarios tested in Projection Set 24 were appropriate for their respective equity scenarios. However, Oliver Wyman also noted to the VAIWG in the Testing Cycle 3 presentation that selecting CTE 95 under the alternative equity scenarios had several significant differences from CTE 98 under the current equity calibration criteria:

i. Hedging would be more likely to reduce a company's Total Asset Requirement – i.e., 100% RBC – under the alternative equity scenarios than under the current equity calibration criteria; in other words, removal of hedging would be more likely to reduce a company's Total Asset Requirement under the current equity calibration criteria than under the alternative equity scenarios;

ii. As outlined in **Table 9**, with a Scalar of 25%, selecting CTE 98 under the current equity calibration criteria caused the Total Asset Requirement to be below that under the current framework – and sometimes substantially so. With CTE 95 under alternative equity scenarios, the Total Asset Requirement was similar to or higher than that under the current framework across all tested market conditions.

Maintaining the framework property that reflection of more extensive hedging be more likely to reduce the Total Asset Requirement was a significant factor behind our recommendation that CTE 95 under the alternative equity scenarios be used in place of CTE 98 under the existing equity calibration criteria. Oliver Wyman re-emphasized this point in presenting our recommendations to the VAIWG on December 1, 2017, stating that:

i. Though the recommended framework revisions reduce disincentives to hedge relative to the current framework, hedging still increases the Total Asset Requirement under 2016YE interest rate conditions and in more adverse interest rate conditions;

ii. Hence, for a hypothetical weakly-capitalized company with heavy VA exposure, the revised framework may incentivize reductions in the extensiveness of hedging to increase the company's RBC ratio, which we view as the key remaining weakness in the framework;

iii. Electing to use CTE 95 with the alternative equity scenarios in place of CTE 98 with the existing equity calibration criteria would mitigate this issue.

A more detailed illustration of the issue is provided in **Figure 11** – an exhibit replicated from our presentation on December 1, 2017. Fundamentally, the issue is that the statutory framework, even after adopting all of Oliver Wyman's recommended revisions, would still allow – and therefore incentivize – a weakly-capitalized company to improve its capital position by reducing the extensiveness of its hedge program and thereby assuming a greater quantum of market risk in low interest rate environments.

We believe that this property is inappropriate conceptually and imprudent from a regulatory point of view:

i. Increasing the amount of market risk retained by a company should ordinarily lead to deterioration in the metric that is used to measure its capital strength;

ii. Allowing this property creates a substantial regulator burden to monitor the actions of a hypothetical weakly-capitalized company, as the RBC ratio would counterintuitively signal an improvement in capital strength should the hypothetical company reduce hedging and retain more risk;

iii. If a regulator deems that the amount of additional market risk assumed via reduced hedging is excessive, the regulator would be mandating that the company take an action that further reduces its RBC ratio. Should the company's capital position further deteriorate in the interim – which is highly possible given the speed at which equity markets move, such an action may push the company into insolvency.

Separately, Oliver Wyman believes that the alternative equity scenarios better mitigate the "scenario generation" risk embedded in the statutory framework and described in Section 9.2 – i.e., that:

i. The equity scenarios within the framework fundamentally reflect a subjective view of the distribution of potential future market outcomes;

ii. Even if such scenarios are calibrated against historical data, there is material parameter uncertainty in the calibration, as changes to the calibration window used or inclusion of additional data – e.g., equity returns from other geographies – can often cause large changes in the parameters.

The latter point is evidenced by the material change observed in the equity scenarios from the simple recalibration conducted in QIS II for Projection Set 24 – i.e., where the calibration window was widened to include a greater amount of data. As discussed in Section 6.1.1, the current equity calibration criteria were determined based on US equity returns from 1955 to 2003, while the recalibration that Oliver Wyman conducted in QIS II lengthened the calibration window to span from 1926 to 2016. Within the additional data included, returns from 1929 to 1933 drove the vast majority of the differences between the alternative equity scenarios and the current equity calibration criteria, illustrating the sharp sensitivity of the scenario parameters to a few years' worth of data.

Select QIS participants contended that it may not be appropriate to include historical US equity return data in the earlier part of Oliver Wyman's lengthened calibration window – i.e., the period encompassing the Great Depression – in recalibrating the equity scenarios, as:

i. The current legal and financial framework of the US economy reflects the lessons learned from the Great Depression. The enactment of the Securities Act of 1933, the Securities Exchange act of 1934, the Investment Company Act of 1940 and the Banking Acts of 1933 (Glass-Steagall) and 1934 established the modern regulatory framework for economic activity, including formation of the SEC, increased powers of the Federal Reserve, establishment of the FDIC, and greater regulation of commercial and investment banking – enhanced recently by the Dodd-Frank Act. As a consequence, lengthy and severe market downturns are less likely than before;

ii. Market indices prior to 1957 were inherently more volatile than modern broad market indices due to the smaller number of companies used in the index. While the S&P equity index tracking began in 1923, experience from 1923 to 1926 was measured using a small number of companies. From 1926 to 1956, the index comprised 90 companies, and the index was only expanded to encompass 500 companies in

1957. When the index had 90 stocks, there were only three industry sectors – whereas the S&P 500 today contains ten industry sectors.

In response to these comments, Oliver Wyman noted that:

i. As discussed in greater detail in Section 9.2, historical US equity return data from 1871 to 1955 were presented by participants and the ACLI as counterarguments against Oliver Wyman's recommendation from 2016 that the equity calibration criteria should be linked with prevailing interest rate conditions. This data was given strong consideration by the VAIWG, signaling to Oliver Wyman that the calibration of alternative equity scenarios should also contemplate data from the early part of the 20th century, given that the VAIWG had found it valuable to consider historical evidence this time period in forming decisions on the structure of the equity scenarios;

ii. The equity calibration criteria represent nominal equity returns. Interest rate levels in 1955-2003 were substantially higher than today and in most previous historical periods. If interest rate levels indeed affected nominal equity returns in this time period, such an effect would not be captured if historical data from outside this period were not considered given the VAIWG's decision not to link equity calibration criteria with prevailing interest rates;

iii. Substantial differences in calibration criteria would be observed if foreign equity return data in recent years were included in the calibration data;

iv. The current calibration criteria were determined assuming a mean equity return of 8.75%; given present interest rate levels, this would imply a large equity risk premium. In the American Academy of Actuaries' (the "Academy") report on C3 Phase II presented to the NAIC's Capital Adequacy Task Force in June 2005 supporting the current equity calibration criteria, the Academy noted that based on average 3-month US Treasury bill returns, that that 3.5-3.75% per annum does not seem excessive. However, this was based on the Academy's determination that "over the last 50 and 20 years, respectively, the average returns on 3-month Treasury bills were approximately 5.30% and 5.15% … hence, taking a long-term perspective … the range 5-5.25% seems sensible for future risk-free rates";[2]

v. There is substantial uncertainty present in the model specification and parameterization of any "real world" scenario generator, particularly in the long-term tail scenario outcomes that drive reserves and capital in the CTE calculation. In addition, using any calibration window within a historical data-based calibration approach contains an element of arbitrariness.

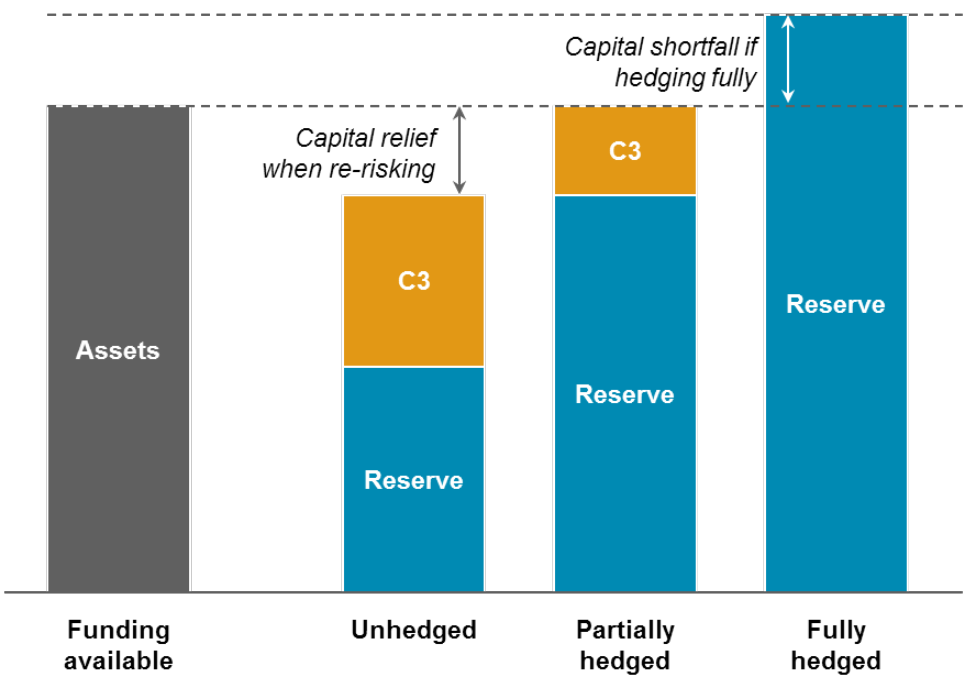
The calibration of the existing VM-20 economic scenario generator relied on daily time series of US equity total return data to calibrate the stochastic volatility process. Oliver Wyman thus sought to obtain the longest-possible daily US equity total return time series between 1871 and 2016 for recalibrating the VM-20 generator and producing the alternative equity scenarios. At the time when Projection Set 24 was being developed, Oliver Wyman was only able to obtain a daily time series of S&P 500 Total Return dating as far back as 1926 for recalibrating the equity scenarios. However, we would support recalibration of the equity scenarios using additional historical data from before 1926, should such data become available in daily time series format.

[2] "Recommended Approach for Setting Regulatory Risk-Based Capital Requirements for Variable Annuities and Similar Products – Presented by the American Academy of Actuaries' Life Capital Adequacy Subcommittee to the National Association of Insurance Commissioners' Capital Adequacy Task Force", June 2005.

Figure 11: exhibit used in the recommendations presentation on December 1, 2017 on incentives to reduce hedging near insolvency

Though revisions reduce disincentives to hedge, hedging still increases TAR with '16YE IR – which may incentivize reductions in hedging near insolvency

Assets vs. TAR for a sample company under different hedge strategies
Under today's interest rate environment



- Revised framework enables hedging to reduce assets required to reach a typical target RBC ratio in most market conditions
- This criterion was agreed in a VAIWG call and formed the basis behind the calibration of the CTE 95 parameter for calculating C3
- However, we were unable to ensure that hedging reduces TAR in low interest rate conditions – including current conditions
  - Cost of hedging in low IR conditions exceeds CTE 70 unless underlying CTE scenarios are overhauled
  - Required changes to scenarios would increase reserves and TAR substantially (e.g., link equity returns with interest rates)
- Thus, as a company's assets approach TAR, it would be incentivized to release hedges to lower TAR and increase RBC ratio
- The company also cannot hedge more if it hasn't already, as hedging would "lock in" its loss and lead to immediate breach of TAR

We view this property as the key remaining weakness in the framework, as it increases reliance on regulator intervention for companies experiencing financial distress

6.2.3.3. Standard Projection Amount – Proposal A

As expected, Oliver Wyman observed significant variation in the specifications of the scenario along which the Scenario GPVAD was closest to that of CTE 70 (the "CTE 70-equivalent scenario") when both calculations used the same set of actuarial assumptions – i.e., the participants' own Prudent Estimate assumptions. Consistent with findings from Testing Cycle 2, the dispersion in the CTE 70-equivalent scenarios suggested a need for the Standard Scenario construct to incorporate some company-specific elements to:

i. Account for different joint market-actuarial risk profiles across companies;

ii. Identify key vulnerabilities and drivers of funding requirements within each company's portfolio.

After participants replaced their Prudent Estimate assumptions with the prescribed assumptions recommended by Oliver Wyman in Projection Set 27, we found that:

i. The Total Standard Projection Amount exceeded the 70 CTE Amount ("adjusted") for five out of the nine participants who submitted results under 2016YE market conditions. This suggested that the prescribed assumptions recommended by Oliver Wyman are "middle-of-the-pack" in their level of conservatism relative to submitting participants' own Prudent Estimate assumptions;

   Across the other Valuation Date market conditions tested, the number of participants for whom the Total Standard Projection Amount exceeded the 70 CTE Amount ("adjusted") rose to a maximum of six after a 300 basis point interest rate increase, and fell to a minimum of three after a 100 basis point interest rate decrease. In the joint equity-interest rate market stress, four participants saw their Total Standard Projection Amount exceed the 70 CTE Amount ("adjusted").

ii. The Additional Standard Projection Amount was non-zero for two out of the nine participants who submitted results under 2016YE market conditions.

   Across the other Valuation Date market conditions tested, the number of participants for whom the Additional Standard Projection Amount was non-zero remained at two for both interest rate shocks. In the joint equity-interest rate market stress, four participants saw their Additional Standard Projection Amount become non-zero.

For several participants for whom the Additional Standard Projection Amount was non-zero, we additionally evaluated the equivalent CTE level – based on CTE Amounts calculated under their own Prudent Estimate assumptions – of their Total Standard Projection Amount. In select instances, the Total Standard Projection Amount exceeded CTE 80 materially, suggesting that the Additional Standard Projection Amount would be large enough to reduce the C3 charge materially. In order to avoid the situation where the C3 charge is reduced to zero because of the Additional Standard Projection Amount – and thereby eliminating the meaningfulness of the resultant RBC ratio, Oliver Wyman ultimately recommended that the Additional Standard Projection Amount be added to the Tail CTE Amount when calculating the C3 charge such that it would not be able to reduce the C3 charge – i.e.,

$$\begin{aligned} \text{C3} = \text{Scalar} \times \big(&(\text{Tail CTE Amount} + \textbf{Add'l Standard Projection Amount} - \text{Statutory Reserve}) \\ &\times (1 - \text{Federal Income Tax Rate}) - (\text{Statutory Reserve} - \text{Tax Reserve}) \times \text{Federal Income Tax Rate} \big) \end{aligned}$$

The addition of the Additional Standard Projection Amount to the Tail CTE Amount is analogous to calculating the Tail CTE Amount with the prescribed Standard Projection Method assumptions recommended by Oliver Wyman (but with the subsequent application of the same buffer that is applied on the Unbuffered Standard Projection Amount). We believe that this is a conceptually appropriate calculation to the extent that the C3 charge is meant to capture the required capital for market risk, as there is no reason to believe that the market risk of a portfolio would be reduced by differences between the company's own Prudent Estimate assumptions and the prescribed assumptions recommended by Oliver Wyman. Further, we note that the addition removes a source of C3 volatility – i.e., the impact of different market conditions on the valuation difference between a company's own Prudent Estimate assumptions and the prescribed assumptions.

6.2.3.4. Standard Projection Amount – Proposal B

For the seven participants that submitted data for Projection Set 34, we observed that the Additional Standard Projection Amount as calculated via Proposal B of the alternative Standard Scenario construct was similar to that calculated under Proposal A across all tested market conditions for six of the participants, but materially different from that calculated under Proposal A for one participant.

Based on these results, we anticipate that for select companies, the specific construct of the prescribed market paths under Proposal A – i.e., an initial stress, followed by recovery – may be dissimilar from those market paths that underlie the calculation of their CTE Amount. As a result, while we anticipate that the majority of companies would see similar Standard Projection Amounts under both Proposals A and B, these select companies may see meaningfully different quantities under the two Proposals. Oliver Wyman believes that from the perspective of governing actuarial assumptions and model choices within the CTE Amount calculation itself, Proposal B represents a more direct approach for achieving such an objective – and may therefore be conceptually preferable. Nevertheless, there may also be regulatory and practical disadvantages to Proposal B that can be addressed by Proposal A. During the Testing Cycle 3 presentation, Oliver Wyman outlined the following advantages and disadvantages of Proposal A vs. Proposal B:

i. **Proposal A**: allows regulators to retain full control over the panel of market paths used in the Standard Projection Method and reduces computational burden for point-in-time balance sheet projections, as only a small set of scenarios are run. However, for balance sheet roll-forward analyses – e.g., for capital planning, this approach is more difficult than Proposal B given the larger number of steps required in calculating the Total Standard Projection Amount;

ii. **Proposal B**: does not allow regulators to retain control over the market paths used in the Standard Projection Method unless the stochastic economic scenario generator is fully prescribed. In addition, point-in-time balance sheet projections are computationally-intensive to conduct, as more scenarios are needed than in Proposal A. However, balance sheet roll-forward analyses would be simpler under Proposal B than under Proposal A, as Proposal B has fewer steps in calculating the Total Standard Projection Amount.

Accordingly, Oliver Wyman ultimately recommended that both calculation methodologies be permitted in the redlined AG 43 document.

6.2.3.5. Aggregate framework impact

As only several participants saw Additional Standard Projection Amounts in excess of zero, incorporation of the Standard Projection Method into the CTE-only analysis observed in Section 6.2.3.1 had a relatively minor impact on participant-aggregate results. In particular, we observed that all of the desirable properties on total balance sheet volatility were maintained after inclusion of the Standard Projection Method.

Four participants provided the full set of requested tax reserve information – including tax reserves under market conditions other than 2016YE conditions – to enable our analysis on the impact from excluding the component that adjusts for the difference between statutory and tax reserves on the Valuation Date. Results from the analyses indicated that for these four participants in aggregate, exclusion of the tax reserve component from the C3 charge calculation led to:

i. An overstatement of the C3 charge under the revised framework by approximately 15-25% as of 2016YE, depending on the set of equity scenarios used and hedge strategy assumed – indicating substantial differences between the statutory and tax reserves;

ii. An overstatement of the “capital buffer” component of the effective funding requirement – i.e., the amount of effective funding requirement in excess of the baseline funding requirement – under the revised framework by approximately 5%.

We also note that for select participants in select market conditions, inclusion of the tax reserve adjustment reduced the C3 charge to zero. However, the incidence of this phenomenon was low, as the difference between

statutory and tax reserves needed to be sufficiently large such that the “tax credit” it offers equals or exceeds the tax-effected difference between the Tail CTE Amount and the statutory reserve, without application of the Scalar.

Based on these results, Oliver Wyman decided after Testing Cycle 3 that the framework revisions tested were effective in achieving their target properties, and subsequently formalized them as recommendations for the revision of the VA statutory reserve and capital framework.

7. Behavioral Assumptions Working Group

The prescribed policyholder behavior assumptions for the Standard Scenario Method within the existing AG 43 framework were developed largely in the absence of relevant experience. As a result, emerging experience data have shown substantial deviations from these prescriptions over the recent years. In light of this, Oliver Wyman recommended in 2016 that the prescribed behavioral assumptions be refreshed to align with emerging data, and provided a set of preliminary revised assumptions, constructed based on Oliver Wyman's prior analyses of the industry's policyholder behavioral experience.

As outlined in Section 2 of this document, one of the primary charges of the Behavioral Assumptions Working Group in QIS II was to evaluate industry-aggregate policyholder behavior experience data to validate or, where required, define the parameterizations of the prescribed Standard Scenario policyholder behavior assumptions included in the 2016 Oliver Wyman recommendations.

To this end, the Behavioral Assumptions Working Group undertook the following tasks:

i. Develop a list of material product features that may influence policyholder behavior;

ii. Agree on a list of material product features and policyholder behavioral actions to include in the prescribed Standard Scenario assumptions;

iii. Gather and aggregate anonymized policyholder behavior experience data collected by as many participating companies as possible in recent years;

iv. After accounting for differences in company data definitions, construct an industry-aggregate policyholder behavior experience dataset. Existing experience analyses and reports from individual participants were evaluated in full but were not considered valid substitutes for company data;

v. Conduct specific behavioral analyses on the industry policyholder behavior experience dataset with industry participation – primarily for data clarification and confirmation;

vi. Review results from Oliver Wyman's analyses to identify patterns, trends, drivers, and discriminants relevant for the behavioral assumption prescriptions;

vii. Compare the results from Oliver Wyman's analyses against the Standard Scenario behavioral assumptions that were recommended by Oliver Wyman in Proposal 2D in 2016 to identify areas where revisions to the recommendations may be warranted.

The remainder of this section details Oliver Wyman's methodologies in conducting the tasks listed above (collectively, the "QIS II Experience Study"), outlines the primary findings, and discusses the implications of such findings on Oliver Wyman's eventual recommended behavioral assumption prescriptions for the Standard Projection Method.

7.1. Process and scope of the experience study

7.1.1. Process and timeline

Weekly meetings of the Behavioral Assumptions Working Group commenced on March 13, 2017 and were held, for the majority of QIS II, on a weekly basis; the Working Group officially concluded on September 25, 2017. All QIS II participants elected to attend the Behavioral Assumptions Working Group.

In March, Oliver Wyman issued a set of specific data requests for individual companies that described in detail the format and data fields required for the study, as outlined in Section 7.2.1 of this document. Each participating company then individually submitted experience data to Oliver Wyman. The data collected was subsequently reviewed by Oliver Wyman to ensure data integrity. Follow-up data requests and clarifications were common as

Oliver Wyman worked to clean company data and ensure standardized data definitions across company datasets. In many cases, companies were not able to submit the full set of necessary data fields for all policyholder experience studied. For this reason, most analyses conducted ultimately included data from only a subset of the participating companies. A summary of data used to study various behavioral assumptions is detailed in later sections of this document where specific experience study results are discussed.

After ensuring the integrity of the experience data collected, Oliver Wyman performed a number of transformations – as detailed in Section 7.2.2 of this document – to normalize each individual participant's dataset into a common format that could be subsequently aggregated across all submissions. However, prior to aggregating company datasets into an industry-level dataset, Oliver Wyman conducted the full set of all experience study analyses on a company-by-company basis. The company-specific results were then shared with the individual participants in order to:

i. Confirm with individual participants that the company-level results were consistent with the companies' understanding of their own portfolios and behavioral experience; and

ii. Where Oliver Wyman results were inconsistent with the companies' own understanding, identify potential sources for data errors, data misinterpretations, or other deviations from expectation.

Where participants observed deviations in Oliver Wyman's results from their expectations, Oliver Wyman set up discussions on an as-needed basis to resolve potential issues.

Only after a participant confirmed that Oliver Wyman's results were in-line with its expectations did Oliver Wyman include the participants' data in industry-level exhibits. Company-specific results that Oliver Wyman was unable to validate with the respective participant were not included in any industry-level exhibits presented either in the Behavioral Assumptions Working Group during QIS II or in the exhibits of this document and were not considered in the parameterization of the recommended Standard Scenario policyholder behavior assumption prescriptions.

Beginning in June 2017, industry-level results for each experience study analysis within the scope outlined in Section 7.1.2 of this document were shared with the Behavioral Assumptions Working Group and discussed during the weekly meetings. In these sessions, Oliver Wyman:

i. Discussed the dataset and analytical methods underlying the experience study results;

ii. Compared the industry-level results with the recommendations for Proposal 2D presented in August 2016;

iii. Proposed revisions to Proposal 2D, where necessary; and

iv. Gathered feedback from participating industry representatives regarding the interpretation of industry results as well as implications on revisions to the 2016 recommendations.

Concurrently, participating industry representatives provided feedback to Oliver Wyman regarding the interpretation of industry results as well as implications on revisions to the 2016 recommendations.

On August 3, 2017, Oliver Wyman presented the full set of finalized QIS II Experience Study results, as well as the revised recommendations for the prescribed behavioral assumptions for the Standard Projection Method, to the VAIWG and QIS II participants in an in-person meeting in Philadelphia, PA. During the presentation, regulators of the VAIWG reviewed the experience study results against the revised assumptions and provided feedback on select revisions. After the completion of the meeting, Oliver Wyman incorporated the VAIWG's feedback before finalizing the list of recommended prescribed assumptions for the Standard Projection Method that were presented on December 1, 2017.

7.1.2. Products and assumptions reviewed

The QIS II Experience Study focused on the most critical combinations of policyholder behavioral options and product types. The determination of these critical combinations took into consideration:

i. Survey responses from Behavioral Assumptions Working Group participants toward the beginning of QIS II, when asked about the most critical combinations of policyholder behaviors and products;

ii. Prevalence of various product classes in the industry's in-force portfolios, in aggregate, and the continued pace of sales;

iii. Complexity of product features and their interactions with policyholder behavioral incentives; and

iv. Availability of relevant industry experience data collected.

Ultimately, Oliver Wyman decided to focus the QIS II Experience Study on the following combinations of policyholder behavioral options and product classes:

i. Withdrawal delay – i.e., the timing of the policyholder's first withdrawal from the account value – for GMWBs;

ii. Withdrawal amount – i.e., the amount withdrawn in each year once the policyholder has taken his or her first withdrawal – for GMWBs;

iii. Full surrender ("lapse") – i.e., a policyholder's termination of the contract and associated guarantees in exchange for the full repayment – for GMWBs;

iv. Withdrawal amount for standalone GMDBs – i.e., contracts that do not have living benefit guarantees;

v. Full surrender for standalone GMDBs;

vi. Withdrawal delay and withdrawal amount for GMIBs, including traditional GMIBs and hybrid GMIBs; and

vii. Full surrender for GMIBs, including traditional GMIBs and hybrid GMIBs.

For all GMIB-related studies, Oliver Wyman received experience data from only a small number of Behavioral Assumptions Working Group participants. The QIS II Experience Study included analyses of these datasets to inform revisions to GMIB-related prescribed assumptions for the Standard Projection Method. However, to ensure confidentiality and anonymity of experience data, discussions of the results from these studies were held only with the subset of participants from whom relevant data was collected, and not with the broader Working Group or any other parties in the QIS. For the same reason, we have also excluded discussions of these studies from the remainder of this document.

7.1.3. ACLI engagement with Ruark Consulting

In parallel with the QIS II Experience Study, the ACLI engaged Ruark Consulting, LLC ("Ruark") to conduct an independent, industry-aggregate experience study (the "Ruark Experience Study").

The Ruark Experience Study was conducted using the industry-aggregate dataset that Ruark has collected and maintained via its prior engagements with individual companies and, unlike the QIS II Experience Study dataset, was not developed for the explicit purpose of QIS II or the Behavioral Assumptions Working Group. Accordingly, Oliver Wyman did not have perfect visibility into the data processing, normalization, and analytical techniques employed by Ruark throughout the experience study, and therefore have not included any formal documentation of their process in this document. Nevertheless, we note that Oliver Wyman and Ruark held a number of discussions – with participation from the ACLI and select Behavioral Assumptions Working Group participants – to

discuss the data processing and analytical approaches used in the QIS II Experience Study and the Ruark Experience Study, so that the respective studies may, on a best-efforts basis, use a consistent set of approaches.

We noted the following differences in data and approaches between the Ruark Experience Study and the QIS II Experience Study:

i. The Ruark Experience Study dataset contains a larger amount of experience data, as Ruark collected data from a larger number of companies – including those that are not participants of the QIS II – and over a longer period of time;

ii. Due to its larger dataset, the Ruark Experience Study evaluated withdrawal delay behavior for GMWBs by examining the cumulative withdrawal rates based on rates of annual income initiation by issue age and contract year, rather than solely by attained age as in the approach outlined in Section 7.3.1 of this document. This did not produce materially different conclusions on the relevant policyholder behavior for these contracts; and

iii. In calculating the guarantee in-the-moneyness for standalone GMDBs, the Ruark Experience Study did not reflect future guaranteed roll-up features in the benefit base when calculating the guarantee's actuarial present value, rather than the approach outlined in **Table 10** of this document. We understand that Ruark's motivation for this simplification was to expedite the work for this relatively small segment in context of the required timelines, and mirror their observations of common industry practice. This had the effect of producing a different distribution of exposure by numerical moneyness categories, but not a materially different conclusion on the relevant policyholder behavior for these contracts from that described in Section 7.7.2 of this document.

Other differences may exist between the two experience studies; however, they were not explicitly identified in the conversations between Oliver Wyman and Ruark.

Results from the Ruark Experience Study were shared with Oliver Wyman and the Behavioral Assumptions Working Group from July to September. These results were compared against:

i. Results from the QIS II Experience Study; and

ii. Revised recommendations for prescribed assumptions for the Standard Projection Method.

Where material differences were observed between the Ruark Experience Study results and the QIS II Experience Study results, Oliver Wyman conducted additional analyses to investigate and determine whether further revisions to the recommended assumptions were needed. Overall, where comparable studies were provided, Oliver Wyman judged that the findings of the Ruark Experience Study largely corroborated the findings of the QIS II Experience Study; accordingly, no changes were made to the recommendations in light of the Ruark Experience Study.

7.2. Experience data collection and processing

7.2.1. Data request and collection

On March 14, 2017, Oliver Wyman issued the following Request for Information ("RFI") to all participants of the Behavioral Assumptions Working Group:

REQUEST FOR INFORMATION
March 14, 2017
RE: Behavioral Assumptions Working Group Data Request

Dear Behavioral Assumptions Working Group participants,

To evaluate the suitability of the recommended policyholder behavior assumption prescriptions in Oliver Wyman's 2016 recommendations, the QIS team would like to request policyholder behavior experience data from participants of the Behavioral Assumptions Working Group. In particular, we would like to request that three types of data be provided by willing participants:

i. **Policy census table**: a single, seriatim file containing the static policyholder "census" data – i.e., policyholder and contract attributes;
ii. **Period-end snapshots of the in-force**: files containing snapshots of the in-force VA portfolio at regular intervals, preferably monthly – though quarterly or annual are also acceptable;
iii. **Policy year-level data**: files containing in-force data aggregated on a policy-year level – i.e., each record would indicate the state of a contract in a given policy year, with fields such as "withdrawals" aggregated to represent the total withdrawals taken during the policy year.

Participants may submit data in CSV, Access, or SAS formats. To facilitate data processing, we have provided a set of requested field headers and data types outlined in the remainder of this document, though we would note that:

i. For fields or tables requiring information that a participant does not currently collect or cannot easily provide, the participant may skip them or simply leave the values blank;
ii. Participant that cannot easily re-format experience data into the requested format may provide data in a more readily-available format; however, we would request that these participants also furnish a table mapping its own field headers to the headers requested in this document;
iii. Similarly, participants may use internally-used product or rider codes; however, we would request that these participants also furnish a mapping table with the marketing names.

We ask that participants provide a sample dataset – e.g., 100 records – by March 22 to allow the QIS team to become familiar with the data format, with the full dataset to follow by March 31, 2017. Should you anticipate difficulties in meeting these requested return dates, or if you have any further inquiries, please don't hesitate to let us know via **naic.qis@oliverwyman.com.**

Table A: requested format of the policy census table

Note: for joint policies, please provide the following information for the younger annuitant.

If you have one or more products that have important product features not captured by the fields below, please feel free to include additional fields in your submitted data containing such information. Additionally, you may simply provide a mapping table of rider name or ID to product features and omit the product feature-oriented fields – e.g., roll-up rates and terms – near the bottom of this table.

Field Header	Description of Field	Data type	Example
Company_Name	Name of issuing company	Text	ABC Ins. Co.
Policy_Number	Unique policy number	Text	10119240
Product_ID	Base contract product name or ID	Text	AA2841H
DB_Rider	GMDB rider name or ID	Text	ROP
DB_D4D_Ind	If GMDB has dollar-for-dollar withdrawals	Binary, 1=Yes	1
AB_Rider	GMAB rider name or ID	Text	ROP
IB_Rider	GMIB rider name or ID	Text	6% rollup
WB_Rider	GMWB rider name or ID	Text	5% lifetime
Issue_Date	Base contract issue date	Date, mm/dd/yyyy	01/01/2013
Issue_Age	Annuitant age at issue	Integer	56

Sex	Annuitant sex	Text = M or F	M
DOB	Annuitant date of birth	Date	05/07/1957
Joint_Status	Indicates whether the policy is a joint policy	Text = Y or N	N
AB_Eff_Date	Accumulation benefit effective date	Date, mm/dd/yyyy	01/01/2013
IB_Eff_Date	Income benefit effective date	Date, mm/dd/yyyy	01/01/2013
WB_Eff_Date	Withdrawal benefit effective date	Date, mm/dd/yyyy	01/01/2013
First_WD_Date	Date of the contract's first withdrawal	Date, mm/dd/yyyy	01/01/2013
SC_Period	Surrender charge period in years	Integer	5
Tax_Status	Tax status, qualified or non-qualified	Text = Q or N	Q
Term_Date	Policy termination date	Date, mm/dd/yyyy	01/02/2015
Term_Reason	Policy termination reason	Text, one of: D – Death A – Annuitization F – Free-look S – Full surrender E – Exercise of another benefit O – Other	D
Total_Premium	Cumulative gross premiums	Floating point	39350.91
DB_Rollup_Rate	GMDB rollup rate	Floating point	*0.05*
DB_Rollup_Term	Policy year when GMDB rollup terminates	Floating point	*999*
AB_Rollup_Rate	GMAB rollup rate	Floating point	*0.05*
AB_Rollup_Term	Policy year when GMAB rollup terminates	Floating point	*10*
IB_Rollup_Rate	GMIB rollup rate	Floating point	*0.05*
IB_Rollup_Term	Policy year when GMIB rollup terminates	Floating point	*999*
WB_Rollup_Rate	GMWB rollup rate	Floating point	*0.05*
WB_Rollup_Term	Policy year when GMWB rollup terminates	Floating point	*10*
AB_Yrs_To_Maturity	Years until GMAB matures	Floating point	*3.2*
IB_Yrs_To_Exercise	Years until GMIB becomes exercisable	Floating point	*7.6*

Table B: requested format of period-end snapshots of the in-force

Note: if you have one or more products where certain important product features are dynamic – or time-varying – and cannot be captured fully in the census table, please feel free to include additional fields in the period-end snapshots of the in-force illustrating the product feature. For instance, for GMWBs where the roll-up period resets upon a step-up, you may include a field that has the remaining roll-up term.

Field Header	Description of Field	Data type	Example
Company_Name	Name of issuing company	Text	*ABC Ins. Co.*
Policy_Number	Unique policy number	Text	*10119240*
Val_Date	Valuation Date	Date, mm/dd/yyyy	*03/01/2017*
Attained_Age	Attained age on the valuation date	Integer	*60*
Policy_Yr	Policy year; first year after policy issue = 1	Integer	*1*
Acct_Val	Account value on the valuation date	Floating point	*24566.34*

DB_Base	GMDB base on the valuation date	Floating point	*90000.00*
AB_Base	GMAB base on the valuation date	Floating point	*89833.03*
IB_Unred_Base	Non-declining GMIB base on valuation date	Floating point	*100000.00*
IB_Rem_Base	Declining GMIB base on the valuation date	Floating point	*85378.23*
WB_Unred_Base	Non-declining GMWB base on valuation date	Floating point	*100000.00*
WB_Rem_Base	Declining GMWB base on valuation date	Floating point	*85378.23*
WD_Ind	Whether a withdrawal was taken in the period	Binary, 1= Yes	*1*
Sys_WD_Amt	Amount withdrawn as part of a systematic withdrawal program within the period	Floating point	*2500.00*
Total_WD_Amt	Total amount withdrawn within the period	Floating point	*2500.00*
Total_WD_To_Date	Total amount withdrawn to date	Floating point	*2500.00*
IB_D4D_Max _WD	GMIB maximum dollar-for-dollar withdrawal amount per policy year	Floating point	*4000.00*
WB_Max_WD	GMWB maximum withdrawal amount per policy year	Floating point	*4000.00*
IB_Guaranteed_AF	GMIB guaranteed annuity payout rate, quoted as % of GMIB base received p.a.	Floating point, 1 = 100%	*0.047*
Current_AF	Current annuity payout rate on valuation date, quoted as % of premium received p.a.	Floating point, 1 = 100%	*0.053*
CSV	Cash surrender value on the valuation date	Floating point	*67090.45*
Total_Premium	Cumulative gross premiums on valuation date	Floating point	*90928.02*
Fixed_Acct_Pct	Percent of account value allocated to the fixed account on the valuation date	Floating point, 1 = 100%	*0.17*

Table C: requested format of policy year-level data

Note: for many fields, the requested format below includes both beginning-of-policy year ("BOPY") and end-of-policy-year ("EOPY") values. However, should this become burdensome to produce, participants may simply provide one of (i) BOPY, (ii) EOPY, or (iii) policy year average values.
Additionally, similar to **Table B**, if you have one or more products where certain important product features are dynamic – or time-varying – and cannot be captured fully in the census table, please feel free to include additional fields in the policy year-level data illustrating the product feature.

Field Header	Description of Field	Data type	Example
Company_Name	Name of issuing company	Text	ABC Ins. Co.
Policy_Number	Unique policy number	Text	10119240
Policy_Yr	Policy year; first year after policy issue = 1	Integer	1
Attained_Age	Attained age at the beginning of policy year	Integer	60
Acct_Val_BOPY	Account value at beginning of policy year	Floating point	24566.34
Acct_Val_EOPY	Account value at end of policy year	Floating point	24566.34
DB_Base_BOPY	GMDB base at beginning of policy year	Floating point	90000.00
DB_Base_EOPY	GMDB base at end of policy year	Floating point	90000.00
AB_Base_BOPY	GMAB base at beginning of policy year	Floating point	89833.03

AB_Base_EOPY	GMAB base at end of policy year	Floating point	89833.03
IB_Unred_Base_BOPY	Non-declining GMIB base at beginning of policy year	Floating point	100000.00
IB_Unred_Base_EOPY	Non-declining GMIB base at end of policy year	Floating point	100000.00
IB_Rem_Base_BOPY	Declining GMIB base at beginning of policy year	Floating point	85378.23
IB_Rem_Base_EOPY	Declining GMIB base at end of policy year	Floating point	85378.23
WB_Unred_Base_BOPY	Non-declining GMWB base at beginning of policy year	Floating point	100000.00
WB_Unred_Base_EOPY	Non-declining GMWB base at end of policy year	Floating point	
WB_Rem_Base_BOPY	Declining GMWB base at beginning of policy year	Floating point	85378.23
WB_Rem_Base_EOPY	Declining GMWB base at end of policy year	Floating point	85378.23
WD_Ind	Whether a withdrawal was taken in the period	Binary, 1= Yes	1
Sys_WD_Amt	Amount withdrawn as part of a systematic withdrawal program within the period	Floating point	*2500.00*
Total_WD_Amt	Total amount withdrawn within policy year	Floating point	*2500.00*
Total_WD_To_Date	Total amount withdrawn to date	Floating point	*2500.00*
IB_D4D_Max _WD	GMIB maximum dollar-for-dollar withdrawal amount in the policy year	Floating point	*4000.00*
WB_Max_WD	GMWB maximum withdrawal amount in the policy year	Floating point	*4000.00*
IB_Guaranteed_AF	GMIB guaranteed annuity payout rate, quoted as % of GMIB base received p.a., at the beginning of the policy year	Floating point, 1 = 100%	*0.047*
Current_AF	Current annuity payout rate, quoted as % of premium received p.a., at the beginning of the policy year	Floating point, 1 = 100%	*0.053*
CSV_BOPY	Cash surrender value at beginning of policy year	Floating point	*67090.45*
CSV_EOPY	Cash surrender value at end of policy year	Floating point	*67090.45*
Total_Premium_EOPY	Cumulative gross premiums at beginning of policy year	Floating point	*90928.02*
Total_Premium_EOPY	Cumulative gross premiums at end of policy year	Floating point	*90928.02*
Fixed_Acct_Pct_BOPY	Percent of account value allocated to the fixed account at beginning of policy year	Floating point, 1 = 100%	*0.17*
Fixed_Acct_Pct_EOPY	Percent of account value allocated to the fixed account at end of policy year	Floating point, 1 = 100%	*0.17*

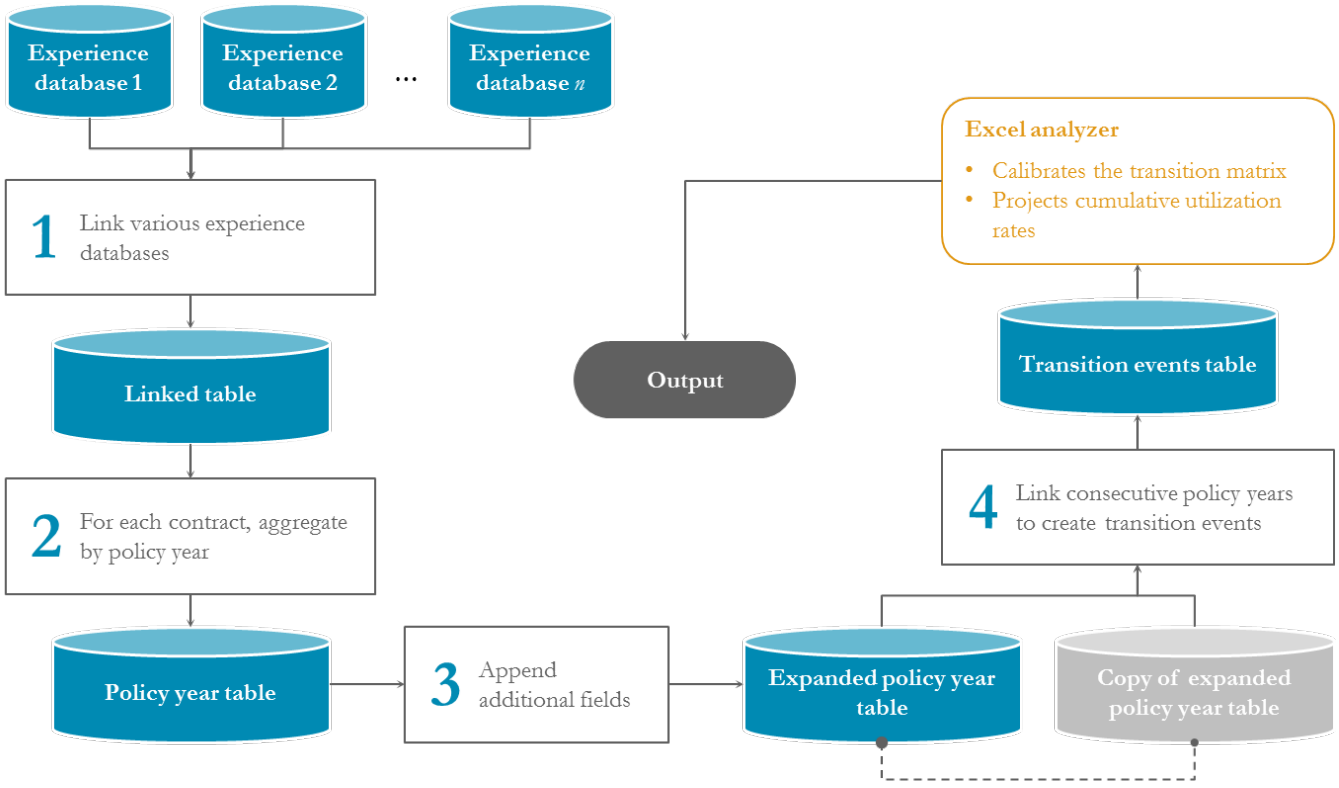
In total, nine participants provided experience data, though such data varied significantly in their coverage of product types, behavioral options, and calendar years. Specifics of the experience data provided and used for each analysis in the QIS II Experience Study are detailed in the discussions of the individual analyses throughout Sections 7.3 to 7.7 of this document.

7.2.2. Data processing and normalization

In order to normalize the different experience data formats provided by different participants, Oliver Wyman strived to develop the industry-aggregate dataset to be consistent in format with that described in Table C of the experience data RFI provided in Section 7.2.1 above. This aspirational dataset format would have been policy year-level – i.e., each record would have indicated the state of a contract in a given policy year, with fields such as "withdrawals" aggregated to represent the total withdrawals taken during the policy year.

Though Oliver Wyman requested experience data in this format via **Table 3** of the experience data RFI, very few participants were able to provide data of a similar format readily during the course of the QIS. Instead, the majority of participants elected to provide experience data only in the format outlined by Table1 and Table 2 of the RFI – i.e., a policy "census" table combined with monthly tables that represent month-end snapshots of the in-force. In order to develop the industry-aggregate dataset from company-specific data of this format, Oliver Wyman undertook the following data processing steps outlined in **Figure 12**:

Figure 12: Data processing flow chart



1. **Linkage**: Companies typically supplied experience databases in multiple database files – e.g., census table and period-end in-force snapshot tables – that needed to be linked together before further processing is possible;
2. **Aggregation**: Data in the format of month-end in-force snapshots were compressed into a single-year resolution by policy year. The policy year for each period-end snapshot record was assigned based on the policy year on the valuation date of the record, and each subsequent policy year-level record was developed from combining twelve monthly records.

In developing the policy year-level dataset, data fields that represent the contract status – such as account value and benefit base – were taken on the valuation date of the last month-end snapshot record included for each policy year. Transaction-based data fields, such as withdrawals, were aggregated across all twelve month-end snapshot records for each policy year.

3. **Expansion**: Additional data fields, most of which were calculated from pre-existing fields, were appended to the policy year table. Depending on the specific policyholder behavior analyzed, different additional data fields were required.

   The key data fields added to the aggregated, policy year-level dataset are detailed below in **Table 10**:

Table 10: List of data fields compiled or created for the QIS behavior study

Field	Description of Field
Withdrawal amount	For all contracts, the total amount withdrawn by the policyholder in the policy year recorded, expressed as a percentage of the contract's account value
Withdrawal efficiency	For GMWB contracts only, the total amount withdrawn by the policyholder in the policy year recorded, expressed as a percentage of the contract's guaranteed maximum annual withdrawal amount
Withdrawal state	For GMWB contracts only, classification of each contract's withdrawal behavior during a policy year. The withdrawal state captures the contract's withdrawal delay, as well as the withdrawal amount in that policy year. The following withdrawal states were assigned: i. **Not withdrawing** – the contract has never taken a withdrawal up to and including the policy year recorded; ii. **Pause** – the contract has a withdrawal efficiency of 0% in the policy year recorded, but had previously taken a withdrawal; iii. **Partial withdrawal** – the contract has a withdrawal efficiency of >0% but ≤90% in the policy year recorded; iv. **Full withdrawal** – the contract has a withdrawal efficiency between 90% and 110% in the policy year recorded; v. **Excess withdrawal** – the contract has a withdrawal efficiency ≥110% in the policy year recorded.
Termination state	Classification of whether the contract has experienced a termination event – e.g., full surrender, annuitization, or death – in the policy year
Conforming withdrawer indicator	For GMWB contracts only, classification of whether a contract is a "conforming withdrawer", defined in the QIS II Experience Study as a contract that has taken a withdrawal in or before the policy year recorded but that had never taken an excess withdrawal up to and including that policy year
Immediate withdrawal indicator	For GMWB contracts only, indicates whether a contract elected to begin withdrawing in the first policy year after the inception of the contract
Incomplete record indicator	Indicates whether a record represents a complete policy year – i.e., twelve full policy months' worth of data for surviving policies. Because some critical contract features – e.g., guaranteed maximum annual withdrawal amount for GMWBs – are defined on a policy year basis, records that do not capture a full policy year's information cannot be used to identify the contract state reliably and thus may need to be discarded in some analyses, such as those for GMWB withdrawal delay and withdrawal amount
In-the-moneyness	The ratio between the Guarantee Actuarial Present Value ("GAPV") of a contract's guaranteed benefit, as defined in Section A3.2)F)3) of the redlined AG 43 document, and its account value.

Field	Description of Field
	Note that Section A3.2)F)3) requires that the GAPV be calculated using the 10-year U.S. Treasury rate on the valuation date. However, to expedite data processing in the QIS II Experience Study, Oliver Wyman rounded all 10-year U.S. Treasury rates to the nearest 100 basis point. Note also that in Oliver Wyman's 2016 recommendations, calculation of in-the-moneyness for certain types of guarantees required the computation of a quantity referred to as the Maximum GAPV – i.e., the maximum GAPV, in present value terms, that is attainable via the range of possible exercise options within a guaranteed benefit. Oliver Wyman continues to believe that the Maximum GAPV metric is a superior reflection of guarantee in-the-moneyness than the simple GAPV metric, but – based on participant feedback – decided during the QIS II Experience Study to remove the concept of Maximum GAPV from our recommendations to reduce implementation complexity for companies.
Premium size	Classification of the premium contribution of a contract into three bands: i. Small: ≤ $100,000; ii. Medium: > $100,000, but ≤ $250,000; and iii. Large: > $250,000

Note that the expanded data fields were generated using the policy year-level dataset. Accordingly, all values used in the calculations outlined in **Table 10** above are consistent in definitions with those for the policy year-level dataset itself. For instance, where account value was used to create an expanded data field, the account value would correspond to that recorded on the valuation date of the final monthly snapshot included in the aggregation for that particular policy year record.

4. **Transition**: for each contract, records representing consecutive policy years were linked together to derive transition events, which indicate the contract's movement – or lack thereof – between withdrawal and termination states as the policyholder traverses policy years.

   Because two complete policy years are required to form a transition event, a contract with *n* complete policy years of data would contribute *n-1* transition events to the experience database. As such, much like all its precursor databases, the ultimate transition event table constitutes a longitudinal database where the status and actions of each contract is tracked through time. The transition table is ultimately used for all analyses in the QIS II Experience Study except studies on withdrawal amounts, which were conducted directly on the expanded policy year-level dataset.

Finally after all data processing and normalization steps outlined in this section had been completed, the relevant datasets were exported into Excel for segmentation, analysis, and interpretation.

7.2.3. Credibility analysis

For several of the termination-related analyses in the QIS II Experience Study – in particular, the full surrender studies outlined in Sections 7.5 and 7.7 of this document, relevant experience data were limited. Accordingly, Oliver Wyman included credibility statistics in disclosure of these results as an additional item.

Traditionally and typically, companies use Limited Fluctuation theory to study credibility for binary events, including the types of termination-related analyses conducted in the QIS II Experience Study. However, for rare events – such as full surrender rates for deeply in-the-money contracts – Limited Fluctuation theory often requires an impractically-large number of observations to be recorded to achieve high credibility scores.

Consequently, for the QIS II Experience Study, Oliver Wyman relied on the Clopper-Pearson binomial proportion confidence interval to judge credibility and provide a visual representation of confidence alongside each

termination analysis. This test calculates the highest and lowest values of a hypothetical "true" incidence rate for which the observed incidence rate is a possible sampling outcome, assuming a binomial distribution of outcomes. As such, the width of this interval will depend on the number of observations used to calculate the lapse rate for a given policy cell as well as the observed lapse rate. Intervals presented throughout this document represent the 95th percentile confidence interval around the observed termination rate, indicating that with 95% certainty, the true population termination rate lies within this interval should the underlying distribution assumptions remain true.

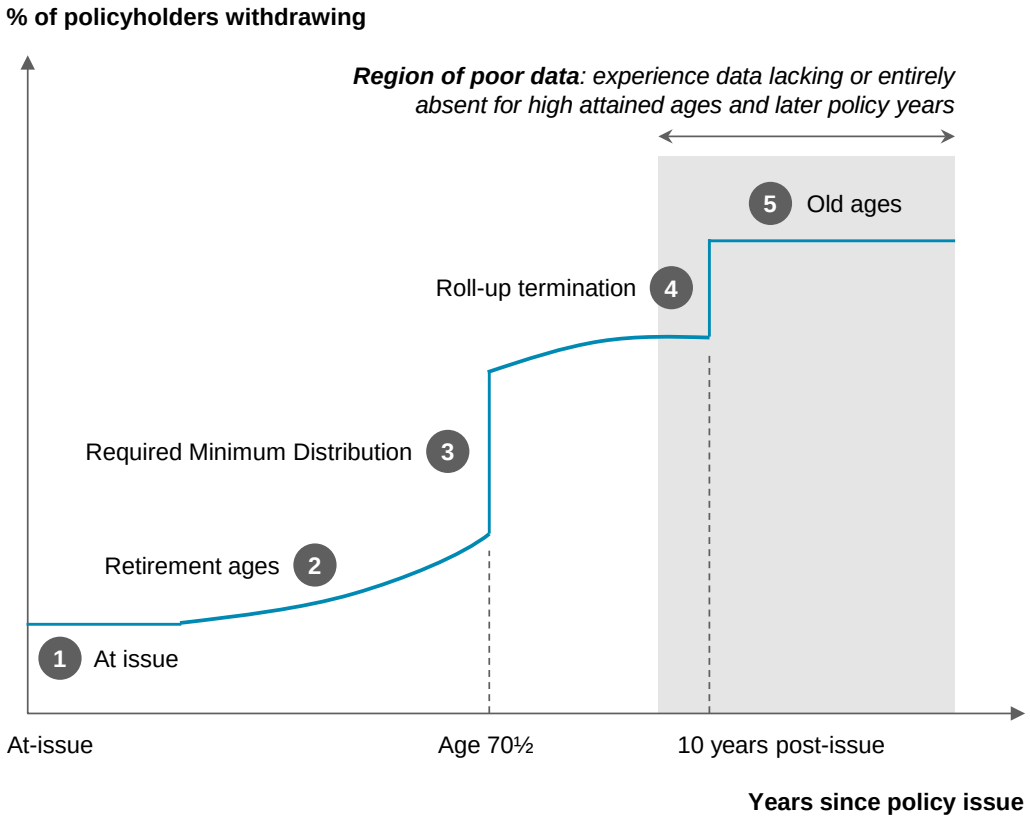
It is important to note that, due to its use of the binomial distribution, the Clopper-Pearson interval assumes that observations sampled to calculate an observed termination rate are entirely independent. In practice, multiple observations are gathered for single policies across study years, likely introducing some level of autocorrelation between observations. Additionally, termination rates are calculated by account value in the QIS II Experience Study, while the confidence intervals shown represent a view calculated by policy count rather than account value. However, for purposes of this report, Oliver Wyman deemed these simplifications acceptable due to the illustrative intent of the calculated interval.

7.3. Study details – withdrawal delay for GMWBs

The GMWB withdrawal delay study aimed to evaluate the accuracy of the cumulative withdrawal rates projected by the Withdrawal Delay Cohort Method outlined in section A3.2)F)6) of the redlined AG 43 document vis-à-vis industry experience for GMWB products. For a specific vintage of a GMWB product, the cumulative withdrawal rate at each point in time represents the proportion of the portfolio of contracts belonging to that vintage that have begun withdrawing by that point in time, excluding contracts that have been terminated.

Oliver Wyman has typically observed that the profile of cumulative withdrawal rates for a single vintage of GMWB contracts through time follows the pattern illustrated in **Figure 13**.

Figure 13: Cumulative withdrawal rate profile for a single vintage of contracts through time



Specifically, there are five phases of distinct withdrawal behaviors:

i. **At-issue**: a portion of policyholders begin to withdraw in the first policy year after contract issue, often at the expense of forgoing future guaranteed benefit growth ("roll-up") or other deferral bonuses;

ii. **Retirement ages**: between the attained ages of 60-70, non-withdrawing policyholders begin to initiate withdrawals at a steady rate each year, leading to a rise in the cumulative withdrawal rates;

iii. **Required Minimum Distribution ("RMD")**: policyholders reaching age 70½ are required by the IRS to withdraw a certain portion of tax-qualified assets each year. Many such policyholders elect to withdraw a portion of the required amount from their VA contracts within qualified accounts, thereby leading to a sharp rise in the cumulative withdrawal rate for qualified policies;

iv. **Roll-up termination**: many GMWB products only provide guaranteed benefit growth ("roll-up") or other deferral bonuses for a finite number of years after issue – e.g., 10 years. The termination of these deferral bonuses materially reduces incentives to defer withdrawals further, thereby triggering a sharp increase in cumulative withdrawal rates;

v. **Old ages**: few policyholders initiate withdrawals in later attained ages, leading to expectations that there may be a "never withdraw" cohort of policyholders that do not end up taking a single withdrawal within the policy lifespan.

The remainder of this section outlines further details of the GMWB withdrawal delay study – including the general approach, summary of findings, and Oliver Wyman's recommended assumptions.

7.3.1. General approach

As discussed in Section 7.2.2 of this document, for a given GMWB contracts within the dataset, Oliver Wyman characterized its withdrawal status in a specific policy year by one of several "withdrawal states" that define:

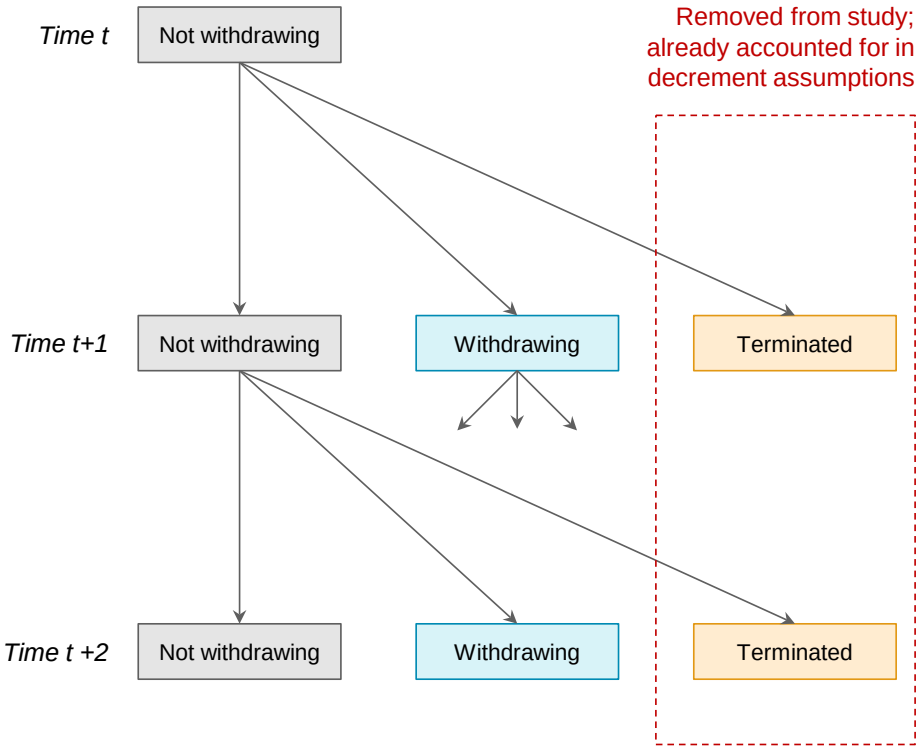
i. If the contract has begun taking income; and

ii. If the contract has terminated

Accordingly, the withdrawal states can be seen as evolving according to a discrete-time process with annual time-steps that coincide with the contract's policy year.

The GMWB withdrawal delay analysis models withdrawal behavior as a time-inhomogeneous Markov chain whereby a contract may transition from one state to another at the conclusion of each policy year, as illustrated in **Figure 14**.

Figure 14: Illustration of Markov process used to model transitions between withdrawal states



Note that:

i. All withdrawal states other than "not withdrawing" are considered "withdrawing", unless the record also has a termination state that represents a termination in that policy year; and

ii. Records that represent policy years in which the contract has terminated were removed from the withdrawal delay study, as these terminations would be accounted for in separate, dedicated termination studies – as outlined in Section 7.5 of this document.

The withdrawal delay study uses the Markov framework – instead of examining overall withdrawal rates exhibited by a snapshot of the historical in-force – in order to provide a clearer view of the withdrawal dynamics. In particular, the Markov framework offers three distinct advantages:

i. Reflects active decisions made in a period instead of the cumulative effect of prior decisions;

ii. Eliminates survival bias from portfolio aging; and

iii. Allows for forward projection of future cumulative withdrawal rates for policies later policy years if the transition rates can be characterized and do not vary significantly across multiple variables.

As discussed in greater detail in Section 7.3.1.1, the Markov framework defines the state space as the withdrawal states that account for a single year of withdrawal history – i.e., a contract's withdrawal status in its current policy year. While a wider state space may be defined by states that cover multiple years of withdrawal history, the short histories of most participants' GMWB experience data preclude extensive analyses from being meaningfully undertaken on a wider state space.

The transition rates obtained from the study can be used to compare directly the propensity for initiating withdrawals across various policyholder and contract segments in order to identify skews in withdrawal behaviour. In addition, the transition rates also provide insight into the interactions between withdrawal efficiency behaviours

– that is, whether policyholders of a particular withdrawal efficiency behaviour would be more likely to exhibit certain other efficiency behaviours in the future.

Furthermore, through *n*-step application of the transition rates to an initial distribution of withdrawal states, the evolution of distributions across withdrawal states in the future can be determined. In this manner, the future cumulative withdrawal rates may be projected with a detailed breakdown of the expected composition of withdrawals across various efficiency levels.

Calibration of the transition rates is detailed in Section 7.3.1.1, while application of the transition rates to project the *n*-step evolution of the distribution across withdrawal states is presented in Section 7.3.1.2.

7.3.1.1. Transition matrix calculation

For a given contract, the transition matrix $\mathbf{P}_t$ has elements $\mathbf{P}_t(i,j)$ representing the probability of the contract's being in withdrawal state j in policy year $t+1$ given that it resides in state i in policy year t. To estimate the unknown matrix $\mathbf{P}_t$, the empirical transition matrix $\widehat{\mathbf{P}}_\Omega$ is calibrated from the transition event table for all contracts with the attribute set Ω. Each entry $\mathbf{P}_\Omega(i,j)$ is calculated as the proportion of policies that have withdrawal state i in any given policy year that transitioned to state j in the subsequent policy year, weighed by the living benefit base ("*LBB*"):

$$\mathbf{P}_\Omega(i,j) = \frac{\sum_\Omega LBB_{i\to j}}{\sum_\Omega LBB_i}$$

The experience dataset used in computing the transition matrices was a filtered version of the transition table described in Section 7.2.2. The filter excluded the following types of records:

i. Records that represent contracts that are immediate withdrawers, as defined in **Table 10**; and

ii. Records that represent policy years in which the policy terminates, as termination events were studied separately in the QIS II Experience Study.

It should be noted that as the attribute set Ω may include any combination of contract characteristics, potential time-inhomogeneity of the transition probabilities may be captured by including in Ω a discretized chronological variable, such as policy year or attained age. Possible dynamic behavior – such as sensitivity to the living benefit in-the-moneyness – can also be accounted for in this manner by specifying different transition matrices for different levels of the dynamic variable.

7.3.1.2. Cumulative withdrawal rate projection

Given a population of contracts, the distribution of withdrawal states in state space $\boldsymbol{S}$ during policy year t can be denoted by the row vector x_t, where each element $x_t(i)$ represents the probability of the contract's being in withdrawal state i. As described in Section 7.3.1.1, for a given contract, the transition matrix in its policy year t, $\mathbf{P}_t$, has each element $\mathbf{P}_t(i,j)$ representing the conditional probability of transitioning from state i to state j by policy year $t+1$, contingent on the contract's not surrendering or otherwise terminating in any manner. Then, the distribution of withdrawal states x_{t+1} in policy year $t+1$, conditional on the contract's survival, is equivalent to:

$$x_{t+1} = x_t\mathbf{P}_t \to x_{t+1}(j) = \sum_{i\epsilon\boldsymbol{S}} x_t(i)\mathbf{P}_t(i,j)$$

Prior Oliver Wyman studies on GMWB withdrawal have identified that a policyholder's attained age is a material driver of withdrawal behavior – and consequently of transition rates among states – while policy year, controlling for attained age, offers little additional explanatory power. As a result, for projection purposes, the transition matrix $\mathbf{P}_t$ in policy year t of a given contract is taken to be equivalent to the empirical transition matrix $\widehat{\mathbf{P}}_a$ that is

calibrated from all policies in a particular segment with attained age a in the starting state, regardless of the policy year. The projected distribution of withdrawal states $\hat{x}_{a+1}$ in the following year, where the policyholder would reach attained age $a+1$, can then be calculated as:

$$\hat{x}_{a+1} = \hat{x}_a \widehat{\mathbf{P}}_a \rightarrow \hat{x}_{a+1}(j) = \sum_{i \epsilon \boldsymbol{S}} \hat{x}_a(i) \widehat{\mathbf{P}}_a(i,j)$$

Because the prescribed policyholder behavioral assumptions for the Standard Projection Method are meant to apply only to those GMWB contracts that have not yet initiated a steady stream of withdrawals, as outlined in Section A3.2)F)4) of the redlined AG 43 document, the starting distribution of states $\tilde{x}_{a_0}$ was set to 100% "Not withdrawing" in the QIS II Experience Study.

To capture the differentiation and skews in withdrawal behaviour across different policyholder segments, withdrawal state distributions are computed via a bottom-up approach: each segment that exhibits distinctive withdrawal behaviour is projected separately, and the projected withdrawal states are subsequently aggregated based on the in-force composition. Based on prior Oliver Wyman GMWB withdrawal studies, tax status and premium size bands are typically used as the major segmentation dimensions for this purpose. With the expanded set of data fields outlined in Section 7.2.2, this yields six individually-projected segments.

For products that have a finite deferral bonus period, Oliver Wyman expects to observe a sharp increase in the cumulative withdrawal rates in the policy year immediately following the termination of the deferral bonus. To model this phenomenon in the cumulative withdrawal rate projection, an external adjustment needs to be made to the Markov chain-based cumulative withdrawal rate projection. Because very few participants possess relevant experience data to calibrate the degree of sharp increase in cumulative withdrawal rates, the QIS II Experience Study did not examine the external adjustment required for all participants. Instead, for products that have finite deferral bonus periods but insufficient data to model the end of the deferral bonus period, the cumulative withdrawal rate projection was conducted only for the duration of the deferral bonus period. As discussed further in Section 7.3.2.2, Oliver Wyman was able to conduct a study on the external adjustment required for the end of the deferral bonus period based on a small subset of participants with sufficient data.

7.3.1.3. Aggregation of industry experience

The GMWB withdrawal delay analysis as described in this section was first conducted for individual products. Subsequently, the individual product-level results were aggregated to the company level – and then the industry level – using a simple average, calculated at each attained age.

Oliver Wyman elected to use a simple average for aggregating across products for the following reasons:

i. **Representation**: Oliver Wyman has observed that GMWB withdrawal delay behavior is sensitive to product features – which often provide differing incentives for withdrawal vs. deferral across the industry. Accordingly, using weighted average-based methods for aggregation risks marginalizing the results for products with smaller exposures in the experience dataset, which may mask product-specific weaknesses in Oliver Wyman's recommended behavioral assumption prescriptions even if the prescriptions appear to be appropriate for dominant products in the dataset;

ii. **Confidentiality**: similarly, weighted average-based aggregations would disproportionately reflect the results of the largest products in the industry, which may risk the data confidentiality of specific carriers with large and seasoned in-force GMWB portfolios.

7.3.2. Results and discussion for lifetime WBs

7.3.2.1. Data description and credibility

The underlying experience data for the study is characterized below in **Table 11**:

Table 11: Overview of experience data included in the study

Data attribute	Products with roll-ups	Products without roll-ups
Companies included	6	4
Products included	27	7
Range of calendar years	2006 – 2016	2006 - 2016

There are a small number of GMWB products in industry that have both lifetime and non-lifetime withdrawal options. These products allow the policyholder to use the policy as a lifetime and a non-lifetime contract (i) without the need to make a pre-determination upon issue and (ii) with the flexibility of switching between the two in the middle of the contract lifespan. In our withdrawal timing analyses, we have included such products in both the lifetime GMWB and non-lifetime GMWB datasets. We note, however, that these products represent a minor proportion of the total population studied.

Figure 15 and **Figure 16** below illustrate the total annual exposure in the GMWB delay study dataset for lifetime GMWBs across the key segmentation dimensions in the study. In particular, we note that:

i. As shown in **Figure 15**, experience data becomes substantially less credible for attained ages in excess of 80 years old – particularly for qualified policies;

ii. As shown in **Figure 16**, experience data becomes very scarce after the 10th policy year.

The second observation is of particular importance, as many of the industry's lifetime GMWB products possess a 10-year deferral bonus period. As noted in Section 7.3.1.2, Oliver Wyman expects that, with the termination of the deferral bonus after 10 years, a sharp increase in withdrawal rates would be observed. The lack of robust experience study supporting or refuting this phenomenon, therefore, represents a key limitation of the QIS II Experience Study for GMWB withdrawal delay.

Figure 15: total annual exposure in GMWB delay study dataset, by tax status and attained age

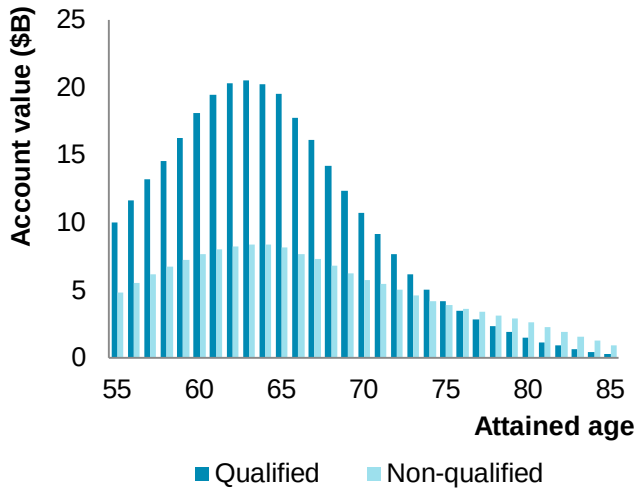
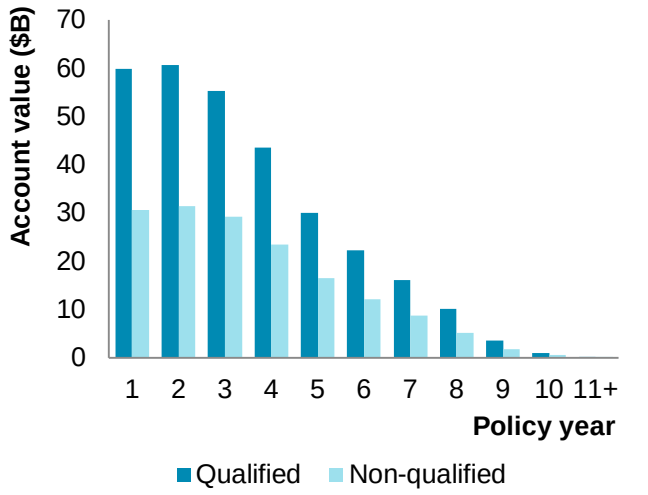


Figure 16: total annual exposure in GMWB delay study dataset, by tax status and policy year



7.3.2.2. Study results

Overall, results from the study demonstrated that:

i. Withdrawal rates are low in the earlier ages – i.e., below attained age 60;

ii. Withdrawal rates steadily increase in the retirement ages – i.e., after attained age 60;

iii. For qualified policies, withdrawal rates increase sharply after attained age 70½, driven by Required Minimum Distributions ("RMD"); and

iv. For non-qualified policies, withdrawal rates continue to rise at a steady pace after attained age 70½ as there is no RMD effect.

Note that in the study, Oliver Wyman only projected cumulative withdrawal rates for the first ten policy years in order to account for the lack of industry experience data and product-specific effects past the first ten years. As discussed in Section 7.3.1.2:

i. Many lifetime GWMB products with guaranteed roll-up features in the benefit base terminate the roll-up after the tenth policy anniversary;

ii. We expect that the end of the rollup period would cause a significant increase in cumulative withdrawal rates, as the policyholder loses a significant incentive to defer withdrawals further; but

iii. As shown in **Figure 16**, most of the industry's experience has not reached the tenth policy anniversary yet, thus precluding direct observation of the magnitude of this expected "shock withdrawal" phenomenon.

Nevertheless, in the industry experience collected for this study, there were four rollup lifetime GMWB products for which the experience dataset does encompass the end of the deferral bonus period. Oliver Wyman studied these products individually to ascertain the existence and assess the magnitude of their "shock withdrawal" event at the end of the deferral bonus period.

Specifically, we studied the rate at which a non-withdrawing contract transitions to withdrawing in the subsequent policy year – i.e., the withdrawal transition rate in each policy year. We observed that:

i. The transition rate is steady in the years immediately before the end of the deferral bonus period;

ii. In the year immediately after the end of the rollup period, the transition rate increases sharply, indicating that the expected "shock withdrawal" phenomenon exists;

iii. In the second policy year after the end of the deferral bonus period, the transition rate decreases quickly – but remains moderately elevated relative to transition rates before the end of the deferral bonus period;

iv. In later policy years, while the transition rates appear to continue trending downward, they remain either equal to or higher than the transition rates within the deferral bonus period.

We note that one of these four products studied has a ten-year deferral bonus period, while the remaining three products have deferral bonus periods that are significantly shorter than ten years. However, there does not appear to be a substantial difference in the magnitude of the "shock withdrawal" event, particularly between the product with the ten-year deferral bonus period and products with significantly shorter deferral bonus periods.

7.3.2.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.3.2.2 against our recommended prescribed assumptions for the Standard Projection Method. These comparison exhibits were presented to the VAIWG and QIS participants and were constructed in the following manner:

i. Projected cumulative withdrawal rates generated from the experience study dataset are identical to those in Section 7.3.2.2;

ii. Cumulative withdrawal rates representing Oliver Wyman's recommendation were generated via the Withdrawal Delay Cohort Method outlined in section A3.2)F)6) of the redlined AG 43 document – first individually for each product included in the experience study dataset, then aggregated to the industry-level via simple average;

iii. For each issue age, attained age, and tax status, the percentage point difference between the projected cumulative withdrawal rate and the cumulative withdrawal rate generated via Oliver Wyman's recommendation was determined – first individually for each product included in the experience dataset, then aggregated to the industry-level via a root mean squared error ("RMSE") calculation; and

iv. Cumulative withdrawal rates representing the current AG 43 Standard Scenario prescription – i.e., all contracts commence withdrawals at the earliest possible time – were also plotted.

Based on these comparison exhibits, Oliver Wyman believes that the recommended Withdrawal Delay Cohort Method follows industry experience for lifetime GMWBs closely, both in aggregate and on a product-specific level – as evidenced by relatively low RMSEs that are typically below 10 percentage points. The Withdrawal Delay Cohort Method captures differences in withdrawal behavior between qualified and non-qualified contracts, as well as between products with higher benefit growth potential – and therefore greater withdrawal deferral incentive – and products with lower benefit growth potential. Additionally, the recommended Withdrawal Delay Cohort Method represents a substantial improvement over the accuracy of the current AG 43 prescribed assumption, which sets the cumulative withdrawal rate at 100% at the earliest possible age of eligibility for guaranteed withdrawals for all products, regardless of withdrawal incentive.

Section A3.2)F)6)g) of the redlined AG 43 document describes Oliver Wyman's recommendation for the magnitude of the "shock withdrawal" phenomenon at the end of the deferral bonus period. Specifically, the recommendation sizes the "shock withdrawal" to be equal to 35% of the proportion of policyholders that:

i. Do not withdraw inside the deferral bonus period; but

ii. Are projected to withdraw eventually – i.e., that are not within the "never withdrawal cohort" described in section A3.2)F)6)m) of the redlined AG 43 document.

The 35% figure was selected to target an approximately 25% transition rate for a non-withdrawing contract at the end of the deferral bonus period to begin withdrawing in the first policy year after the deferral bonus period, consistent with the upper end of industry experience observed. Note that the denominator used to calculate the transition rates that we referred to in Section 7.3.2.2 include contracts that may eventually never withdraw, whereas the 35% recommendation excludes the "never withdrawal cohort".

In developing the recommendations for the lifetime GMWB withdrawal delay assumptions, Oliver Wyman aimed to arrive at a set of prescribed assumptions that:

i. Are sufficiently simple and parsimonious in parameters to allow facile interpretation and application;

ii. Reflect prevailing industry experience data accurately; and

iii. Are sufficiently robust to capture the wide range of industry products that are currently in-force – as well as those that may be developed in the future – without creating undue biases.

Importantly, Oliver Wyman did not attempt to apply explicit conservatism margins on top of industry experience data. For many behavioral assumptions, margins that increase conservatism for one product design may in fact reduce conservatism for another product design; accordingly, to avoid unintended consequences, Oliver Wyman strived to develop accurate and unbiased assumptions that represent reasonable "best-estimates".

The sole exception within the withdrawal delay assumption to this view of conservatism margins lies in the "shock withdrawal" assumption. In targeting the equivalent transition rate, Oliver Wyman deliberately selected the upper end of industry experience – instead of the mean or median – to reflect the high degree of uncertainty in this particular assumption given the dearth of relevant industry experience data at this point in time. Oliver Wyman believes that this choice is appropriate given that section III)8) of the current AG 43 guidelines require that a Prudent Estimate assumption incorporate larger conservatism margins for assumptions with greater uncertainty, and have noted this as both a limitation and an area of focus in future experience studies.

7.3.3. Study results for non-lifetime WBs

7.3.3.1. Data description and credibility

The underlying experience data for the study is characterized below in **Table 12**:

Table 12: Overview of experience data included in the study

Data attribute	All products
Companies included	6
Products included	9
Range of calendar years	2006 - 2016

There are a small number of GMWB products in industry that have both lifetime and non-lifetime withdrawal options. These products allow the policyholder to use the policy as a lifetime and a non-lifetime contract (i) without the need to make a pre-determination upon issue and (ii) with the flexibility of switching between the two in the middle of the contract lifespan. In our withdrawal timing analyses, we have included such products in both the lifetime GMWB and non-lifetime GMWB datasets. We note, however, that these products represent a minor proportion of the total population studied.

Figure 17 and **Figure 18** below illustrate the total annual exposure in the GMWB delay study dataset for non-lifetime GMWBs across the key segmentation dimensions in the study. In particular, we note that:

i. As shown in **Figure 17**, experience data becomes substantially less credible for higher attained ages – particularly for qualified policies;

ii. As shown in **Figure 18**, experience data becomes scarce after the 10th policy year.

Similar to the lifetime GMWB study, the deterioration of experience data quality for contracts with high attained ages and policy years limits the robustness of the withdrawal delay study for non-lifetime GMWBs. However, we expect that the limitation caused by the second observation above would be less severe for non-lifetime GMWBs compared to lifetime GMWBs, as few non-lifetime GMWBs in the study contained substantial deferral bonuses – e.g., guaranteed benefit roll-up.

Figure 17: total annual exposure in GMWB delay study dataset, by tax status and attained age

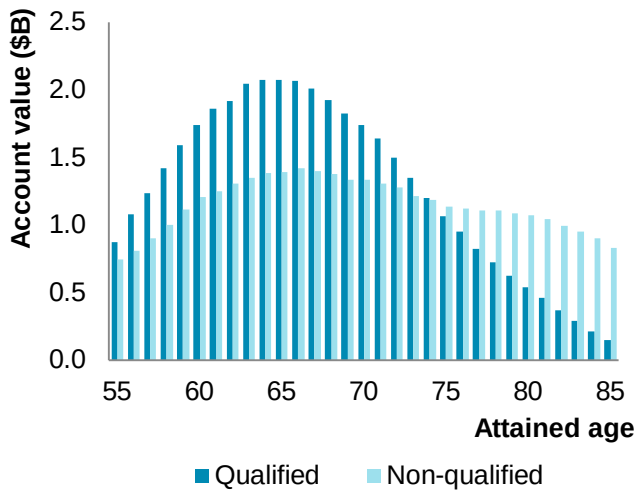
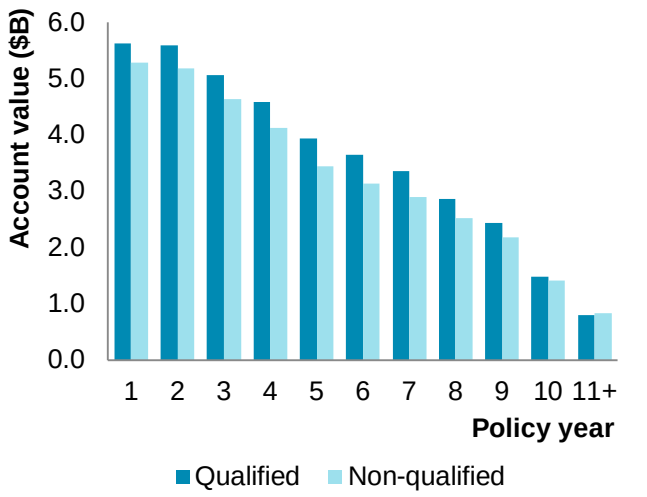


Figure 18: total annual exposure in GMWB delay study dataset, by tax status and policy year



7.3.3.2. Study results

Overall, results from the study demonstrated that:

i. Withdrawal rates are low in the earlier ages – i.e., below attained age 60;

ii. Withdrawal rates steadily increase in the retirement ages – i.e., after attained age 60;

iii. For qualified policies, withdrawal rates increase sharply after attained age 70½, driven by Required Minimum Distributions ("RMD"); and

iv. For non-qualified policies, withdrawal rates continue to rise at a steady pace after attained age 70½ as there is no RMD effect.

These findings are consistent with those identified for lifetime GMWBs, as outlined in Section 7.3.2.2.

7.3.3.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.3.3.2 against our recommended prescribed assumptions for the Standard Projection Method. These comparison exhibits were presented to the VAIWG and QIS participants and were constructed in the same manner as those in Section 7.3.2.3.

As is the case with lifetime GMWBs, the Withdrawal Delay Cohort Method for non-lifetime GMWBs also captures differences in withdrawal behavior between qualified and non-qualified contracts, as well as between products with higher benefit growth potential and those with lower benefit growth potential. Additionally, the recommended Withdrawal Delay Cohort Method represents a substantial improvement over the accuracy of the current AG 43 prescribed assumption, which sets the cumulative withdrawal rate at 100% at the earliest possible age of eligibility for guaranteed withdrawals for all products, regardless of withdrawal incentive.

Oliver Wyman notes that the Withdrawal Delay Cohort Method has lower predictive power for non-lifetime GMWB withdrawal behavior than for predicting lifetime GMWB behavior. Nevertheless, we believe that the Withdrawal Delay Cohort Method still follows industry experience for non-lifetime GMWBs sufficiently closely – both in aggregate and on a product-specific level – such that modifications to the recommendations are not warranted. In

particular, Oliver Wyman judged that creating a differentiated Withdrawal Delay Cohort Method for non-lifetime GMWB would introduce excessive complexity inappropriate for the marginal improvement in predictive accuracy.

7.3.4. Limitations

While Oliver Wyman believes that the analytical methodology for the study was appropriate and that the results of the study support the recommendations for behavioral assumption prescriptions within the Standard Projection Method, several key limitations remain and are summarized below in Table 13.

Table 13: Limitations of the GMWB withdrawal delay study

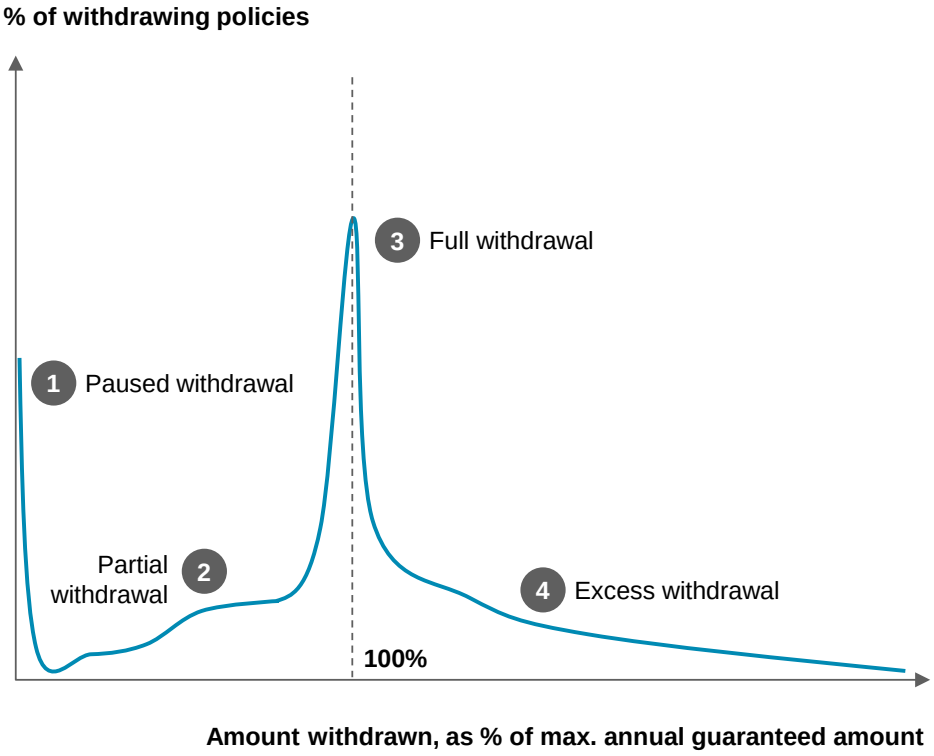
Limitation	Description
Lack of experience data for policy durations beyond 10 years	Experience data is limited for contracts beyond the tenth policy year. This data limitation precludes robust analysis of the "shock withdrawal" phenomenon that is expected at the end of the deferral bonus period, as most in-force products have 10-year deferral bonus periods
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns such that non-withdrawing policies are increasingly less likely to withdraw; ii. Possible improvement in the economic efficiency of policyholders' behavior patterns through time
Ad hoc withdrawal patterns not differentiated and studied in depth	The study – as well as Oliver Wyman's recommended behavioral assumption prescriptions – focuses on contracts that perform persistent withdrawals to obtain a retirement income stream from the contract. While such withdrawal pattern is dominant in the experience data collected, withdrawals of more ad hoc natures also exist. These were not studied separately, but rather were aggregated with the persistent, income-oriented withdrawals
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential additional data skews or behavioral drivers

7.4. Study details – withdrawal amount for GMWBs

The GMWB withdrawal amount study aimed to evaluate the withdrawal efficiency ratio, calculated in a manner consistent with **Table 10** of this document – i.e., the amount withdrawn in each policy year, expressed as a percentage of the contract's guaranteed maximum annual withdrawal amount for that policy year – for contracts that have already taken a first withdrawal.

Oliver Wyman has typically observed that the distribution of withdrawal efficiency ratios for modern GMWB contracts follows the pattern illustrated in **Figure 19**.

Figure 19: Distribution and classes of withdrawal efficiency ratios



Specifically, we observe four classes of withdrawal efficiency ratios:

i. **Paused withdrawal**: the policyholder does not withdraw during the policy year, but has previously taken a withdrawal in a prior policy year;

ii. **Partial withdrawal**: the policyholder withdraws a non-zero amount less than the guaranteed maximum annual withdrawal amount during the policy year. A large portion of such withdrawals are taken by policyholders with qualified contracts who withdraw just enough to satisfy RMD requirements;

iii. **Full withdrawal**: the policyholder withdraws an amount over the policy year equal to the guaranteed maximum annual withdrawal amount;

iv. **Excess withdrawal**: the policyholder withdraws an amount over the policy year in excess of the guaranteed maximum annual withdrawal amount. For most modern GMWBs, excess withdrawals reduce the benefit base by the same proportion as that by which the account value is reduced, though some older products have harsher penalties.

We note that the definition of withdrawal efficiency classes outlined in **Table 10** incorporates small tolerance bands around the withdrawal efficiency ratios to accommodate potential data imprecisions in the experience dataset. Such imprecision include, for instance, the use of month-end snapshots to construct policy year-level records, even though the policy anniversary date may not coincide with the end of the month.

The remainder of this section outlines further details of the GMWB withdrawal amount study – including the general approach, summary of findings, and Oliver Wyman's recommended assumptions.

7.4.1. General approach

The experience dataset used in this study was a filtered version of the aggregated, policy year-level dataset described in Section 7.2.2. The filter excluded the following types of records:

i. Records in which the contract has not yet taken a first withdrawal;

ii. Records representing the first policy year in which a contract has taken a withdrawal, as a substantial number of such records represent policyholders who have elected to take future withdrawals equal to 100% of the guaranteed maximum annual withdrawal amount on a systematic withdrawal program, but have made such election in the middle of the policy year. Accordingly, the withdrawal amount in this particular policy year may appear to be below the guaranteed maximum, even though the policyholder intends to withdraw the maximum in all future years;

iii. Records that represent policy years in which the policy terminates; and

iv. Records that represents policy years in which the policy takes an excess withdrawals – i.e., where the amount withdrawn is greater than 110% of the guaranteed maximum annual withdrawal amount.

Note that excess withdrawals, while excluded from the GMWB withdrawal amount study, were included in the GMWB full surrender and excess withdrawal study outlined in Section 7.5. Oliver Wyman decided to treat excess withdrawals in this manner for the following reasons:

i. The typical contractual benefit adjustments for excess withdrawals differ substantially from those for non-excess withdrawals, which precludes excess withdrawals from being modeled using the same modeling approach as non-excess withdrawals; and

ii. For most of the industry's GMWB products, excess withdrawals are effectively a "fractional" full surrender, as they affect the portfolio in a similar manner as full surrenders – i.e., the benefit is reduced by the same proportion as the account value.

The experience dataset was then segmented along the same dimensions and in the same manner as the recommended GMWB withdrawal amount assumption outlined in Section A3.2)F)4) of the redlined AG 43 document. In each segment, the aggregate withdrawal efficiency was calculated as:

$$\text{Withdrawal Efficiency} = \frac{\sum \text{Total amount withdrawn}}{\sum \text{Guaranteed maximum annual withdrawal amount}}$$

In this calculation, the following treatment was applied for records that represent "pauses" – i.e., policy years in which a contract did not take a withdrawal, even though the contract had previously taken a withdrawal, as defined in **Table 10**:

i. For GMWBs that recommence the guarantee roll-up feature in "paused" policy years, the pause was treated as a fully-efficient withdrawal – i.e., the amount withdrawn equals the guaranteed maximum; and

ii. For GMWBs that do not recommence the guarantee roll-up feature in "paused" policy years, the pause was treated as a 0% efficient withdrawal.

Such a bifurcated treatment was applied in the experience study as, from a valuation perspective, "pauses" in products that do not recommence the roll-up feature are the maximally-inefficient form of non-excess withdrawals. However, "pauses" in products that recommence the roll-up feature may not be inefficient at all from a valuation perspective, as the benefit richness increases to compensate for the forgone withdrawal opportunity; in other words, the unwithdrawn portion of the guaranteed maximum is effectively "credited" as a benefit increase.

The GMWB withdrawal amount analysis as described in this section was first conducted for individual products. Subsequently, the distributions of individual product-level results across various segmentation dimensions were plotted in box-and-whisker plots to facilitate interpretation, as discussed in Sections 7.4.2.2 and 7.4.3.2.

7.4.2. Study results for lifetime WBs

7.4.2.1. Data description and credibility

The underlying experience data for the study is characterized below in **Table 14**.

Table 14: Overview of experience data included in the study

Data attribute	All products
Companies included	5
Products included	23
Range of calendar years	2006 - 2016

There are a small number of GMWB products in industry that have both lifetime and non-lifetime withdrawal options. These products allow the policyholder to use the policy as a lifetime and a non-lifetime contract (i) without the need to make a pre-determination upon issue and (ii) with the flexibility of switching between the two in the middle of the contract lifespan. In our withdrawal amount analyses, we have excluded such products from both the lifetime GMWB and non-lifetime GMWB datasets.

Overall, the experience dataset for the lifetime GMWB withdrawal amount study is very robust, with over $50 billion in total annual exposure across the companies and products outlined above in **Table 14**. Qualified contracts constitute $37 billion in exposure, while non-qualified contracts make up $14 billion in exposure.

7.4.2.2. Study results

Oliver Wyman presented results of the study to the VAIWG and QIS participants via box-and-whiskers plot of the distribution of product-level withdrawal efficiency ratios across all studied lifetime GMWB products, segmented by tax status. While outliers exist, the inner quartiles of the box-and-whiskers plot show a tight range of withdrawal efficiency ratios across most lifetime GMWB products – i.e., annual withdrawals are typically between 80-90% of the guaranteed maximum annual withdrawal amount.

7.4.2.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.4.2.2 against Oliver Wyman's recommended prescribed assumptions for the Standard Projection Method. These comparison exhibits were presented to the VAIWG and QIS participants and were constructed in the following manner:

i. Withdrawal efficiency ratios from the experience study dataset are identical to those in Section 7.4.2.2;

ii. The withdrawal efficiency ratio recommended by Oliver Wyman to be used for all lifetime GMWB withdrawals projected in the Standard Projection Method was plotted in orange; and

iii. The withdrawal efficiency ratio in the current AG 43 Standard Scenario prescription for all lifetime GMWB withdrawals taken after attained age 60 was plotted in grey.

In developing the recommended withdrawal efficiency ratio for the Standard Projection Method, Oliver Wyman deliberately targeted the upper quartile of withdrawal efficiency ratios observed across products studied in the QIS II Experience Study to account for the highly simplified nature of the prescribed withdrawal modeling approach. As discussed in Section 7.4.4, for simplicity and computational tractability, Oliver Wyman recommended that all withdrawals for GMWBs be depicted as a fixed, prescribed percentage of the guaranteed maximum annual

withdrawal amount. However, in reality, countless patterns of withdrawal amounts may exist, which may have substantially different valuation results than those obtained under the simplified modeling approach.

We believe that the conservatism in targeting the upper quartile of experience study results is appropriate given that section III)8) of the current AG 43 guidelines require that a Prudent Estimate assumption incorporate larger conservatism margins for assumptions with greater uncertainty. At the same time, we have noted this as both a limitation and an area of focus in future experience studies.

7.4.3. Results and discussion for non-lifetime WBs

7.4.3.1. Data description and credibility

The underlying experience data for the study is characterized below in Table 15.

Table 15: Overview of experience data included in the study

Data attribute	All products
Companies included	6
Products included	9
Range of calendar years	2006 - 2016

There are a small number of GMWB products in industry that have both lifetime and non-lifetime withdrawal options. These products allow the policyholder to use the policy as a lifetime and a non-lifetime contract (i) without the need to make a pre-determination upon issue and (ii) with the flexibility of switching between the two in the middle of the contract lifespan. In our withdrawal amount analyses, we have excluded such products from both the lifetime GMWB and non-lifetime GMWB datasets.

Overall, the experience dataset for the lifetime GMWB withdrawal amount study is very robust, with $38 billion in total annual exposure across the companies and products outlined above in **Table 15**. Qualified contracts constitute $25 billion in exposure, while non-qualified contracts make up $13 billion in exposure.

7.4.3.2. Study results

Oliver Wyman presented results of the study to the VAIWG and QIS participants via box-and-whiskers plot of the distribution of product-level withdrawal efficiency ratios across all studied lifetime GMWB products, segmented by tax status.

Compared to the results for lifetime GMWBs discussed in Section 7.4.2.2, we observe that:

i. Non-lifetime GMWB products typically have lower withdrawal efficiency ratios, with most residing between 50-70% of the guaranteed maximum annual withdrawal amount;

ii. The dispersion of withdrawal efficiency ratios across non-lifetime GMWB products is substantially wider.

These observations align with our general expectations, as non-lifetime GMWB policyholders should have lower incentive to withdrawal the full guaranteed maximum, given there is only a finite cumulative amount that can be withdrawn from the benefit regardless of the policyholder's longevity.

7.4.3.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.4.3.2 against our recommended prescribed assumptions for the Standard Projection Method. The comparison exhibit was constructed in the same manner as that in Section 7.4.2.3.

Similar to the recommended assumption for lifetime GMWBs, in developing the recommended withdrawal efficiency ratio for the Standard Projection Method, Oliver Wyman deliberately targeted the upper quartile of withdrawal efficiency ratios observed across products studied in the QIS II Experience Study to account for the highly simplified nature of the prescribed withdrawal modeling approach.

7.4.4. Limitations

While Oliver Wyman believes that the analytical methodology for the study was appropriate and that the results of the study support the recommendations for behavioral assumption prescriptions within the Standard Projection Method, several key limitations remain and are summarized in Table 16.

Table 16: Limitations of the GMWB withdrawal amount study

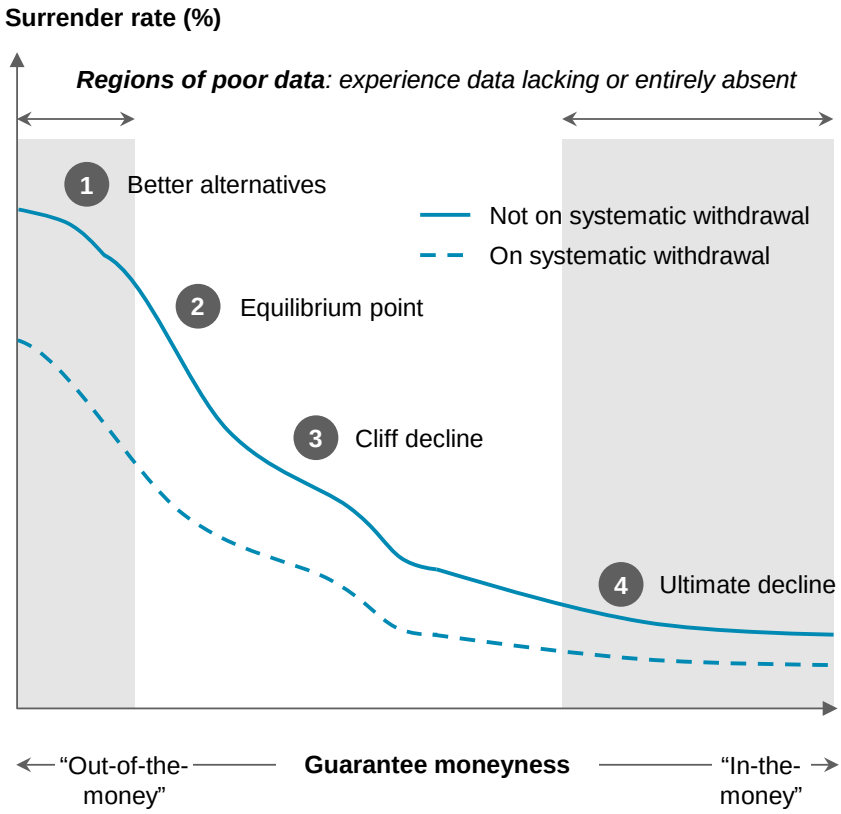
Limitation	Description
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns; ii. Possible improvement in the economic efficiency of policyholders' withdrawal patterns through time
Ad hoc withdrawal patterns not differentiated and studied in depth	The study – as well as Oliver Wyman's recommended behavioral assumption prescriptions – focuses on contracts that perform persistent withdrawals to obtain a retirement income stream from the contract. While such withdrawal pattern is dominant in the experience data collected, withdrawals of more ad hoc natures also exist. These were not studied separately, but rather were aggregated with the persistent, income-oriented withdrawals
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential additional data skews or behavioral drivers
Recommended modeling approach in Standard Projection Method simplifies actual withdrawal dynamics	For simplicity and computational tractability, Oliver Wyman recommended that all withdrawals for GMWBs be depicted as a fixed, prescribed percentage of the guaranteed maximum annual withdrawal amount. However, in reality, countless patterns of withdrawal amounts may exist, which may have substantially different valuation results than those obtained under the prescribed simplification that all withdrawals are a fixed percentage of the guaranteed maximum annual withdrawal amount

7.5. Study details – full surrender and excess withdrawal for GMWBs

The GMWB full surrender and excess withdrawal study aimed to evaluate the total annual lapse rate for GMWB portfolios from full surrenders and excess withdrawals. As discussed in Section 7.4.1 of this document, Oliver Wyman decided to combine the study of excess withdrawals and full surrenders given the similarities in benefit adjustment – and consequently valuation impact – between the two events for the majority of the industry's in-force products.

Oliver Wyman has typically observed that full surrender rates for GMWB contracts follow the pattern illustrated in **Figure 20** – particularly for contracts outside the surrender charge period.

Figure 20: Full surrender rate profile across guarantee in-the-moneyness levels



Specifically, we observe four general regions in the chart:

i. **Better alternatives**: the cash surrender value of the contract exceeds the value of the guarantee and good alternative investments exist. Accordingly, the policyholder has an economic incentive to surrender and reinvest the cash surrender value proceeds in a different product;

ii. **Equilibrium point**: the policyholder can obtain a similar guaranteed income stream by surrendering and investing the cash surrender value proceeds into an alternative product;

iii. **Cliff decline**: the guarantee becomes more in-the-money and the financial advisor finds it increasingly harder to make the case that surrendering and reinvesting the proceeds in an alternative product is economically favorable for the policyholder;

iv. **Ultimate decline**: when the guarantee is deeply in-the-money, surrenders are almost exclusively driven by the policyholder's need for immediate liquidity.

We also note that, as discussed in Section 7.5.3, there are two regions of in-the-moneyness levels where relevant industry experience for full surrender rates – particularly out of the surrender charge period – is sparse:

i. Contracts with deeply in-the-money guarantees, given the lack of a sustained market downturn in recent years;

ii. Contracts with far out-of-the-money guarantees, given most modern GMWB products have automatic step-up features in their benefit base and have not been exposed to a rising interest rate environment.

The remainder of this section outlines further details of the GMWB full surrender and excess withdrawal study – including the general approach, summary of findings, and Oliver Wyman's recommended assumptions.

7.5.1. General approach

The experience dataset used in this study to evaluate full surrender behavior was the aggregated transition table described in Section 7.2.2, similar to that used for the GMWB withdrawal delay study – but focusing primarily on termination states instead of withdrawal states. The experience dataset was segmented along the same dimensions and in the same manner as the recommended GMWB full surrender assumption outlined in Section A3.2)F)7) of the redlined AG 43 document. In each segment, the aggregate annual full surrender rate was calculated as:

$$\text{Full Surrender Rate} = \frac{\sum \text{Account Value}_{\text{Persisting}\rightarrow\text{Full Surrender}}}{\sum \text{Account Value}_{\text{Persisting}}}$$

The notation for this formula is consistent with that used for the withdrawal state transition rate calculation in Section 7.3.1.1. Accordingly, $\text{Account Value}_{\text{Persisting}\rightarrow\text{Full Surrender}}$ represents the aggregate account value of records that transition from a non-terminated state in policy year t to a full surrender termination state in policy year $t+1$, while $\text{Account Value}_{\text{Persisting}}$ represents the aggregate account value of all records that are in the non-terminated state in policy year t, including those that do not transition to a full surrender termination state in policy year $t+1$.

Note that in the annual full surrender rate calculation, Oliver Wyman has elected to use as the denominator the aggregate account value in the policy year prior to the policy year in which the surrender occurred because of data constraints. In many of the participants' datasets, the account value of the policy year in which the contract surrendered was not available. Nonetheless, Oliver Wyman believes that the full surrender rate calculated in this manner should not lead to material, systematic distortions in the experience study results.

The experience dataset used in this study to evaluate excess withdrawal behavior was the aggregated policy year-level dataset described in Section 7.2.2, similar to that used for the GMWB withdrawal amount study – but filtering out only records that represent policy years in which the policy terminates. Consistent with the full surrender study, the experience dataset was segmented along the same dimensions and in the same manner as the recommended GMWB full surrender assumption outlined in Section A3.2)F)7) of the redlined AG 43 document. In each segment, the aggregate excess withdrawal rate was calculated as:

$$\text{Excess Withdrawal Rate} = \frac{\sum \text{Total amount withdrawn via excess withdrawals}}{\sum \text{Account Value in prior policy year}}$$

Note that to remain consistent with the full surrender study, the denominator used in calculating the excess withdrawal rate was the aggregate account value recorded at the end of the policy year prior to the policy year in which each excess withdrawal was taken. In this manner, the total lapse rate – which includes reductions in the GMWB portfolio from both full surrenders and excess withdrawals – may be calculated as:

$$\text{Lapse Rate} = \text{Full Surrender Rate} + \text{Excess Withdrawal Rate}$$

Note that because of the scarcity of experience data for key segments – e.g., contracts with deeply in-the-money guarantees, Oliver Wyman did not present aggregated results via distributions of product-level results or simple averages across products. Instead, Oliver Wyman presented only aggregated results of the study where the aggregate full surrender and excess withdrawal rates effectively represented exposure-weighted participant averages, as shown in Section 7.5.2.2.

Accordingly, unlike in Sections 7.3 and 7.4, we also did not differentiate between lifetime GMWBs and non-lifetime GMWB products in presenting the results of the GMWB full surrender and excess withdrawal study.

7.5.2. Results and discussion

7.5.2.1. Data description

The underlying experience data for the study is characterized below in Table 17.

Table 17: Overview of experience data included in the study

Data attribute	Lifetime GMWBs	Non-lifetime GMWBs
Companies included	4	2
Products included	46	7
Range of calendar years	2006 – 2016	2006 - 2016

There are a small number of GMWB products in industry that have both lifetime and non-lifetime withdrawal options. These products allow the policyholder to use the policy as a lifetime and a non-lifetime contract (i) without the need to make a pre-determination upon issue and (ii) with the flexibility of switching between the two in the middle of the contract lifespan. In our full surrender and excess withdrawal analyses, we have excluded such products from the experience dataset.

Figure 21 to **Figure 26** below illustrate the total annual exposure in the GMWB full surrender study dataset, for contracts with non-zero surrender charge periods, across the key segmentation dimensions in the study – i.e.,

i. Whether the policy is in the surrender charge period, in its first policy year after the surrender charge period, or in subsequent policy years;

ii. Whether the policy is a conforming withdrawer, as defined in **Table 10**; and

iii. Guarantee in-the-moneyness bands, as calculated via the GAPV-based approach outlined in **Table 10**.

We note that, as is evident in these figures, relevant experience data for contracts with deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years. Accordingly, we have noted this as a major limitation in Section 7.5.3 of this document and identified it as a focal area for future experience studies conducted for the purpose of verifying or updating the prescribed behavioral assumptions within the Standard Projection Method.

Figure 21: total annual exposure, in CDSC period, conforming withdrawers

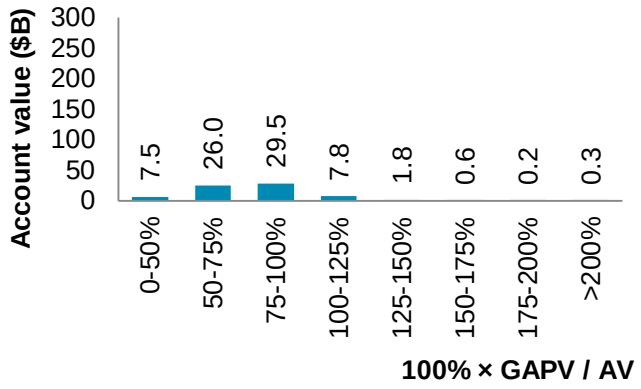


Figure 22: total annual exposure, in CDSC period, other policies

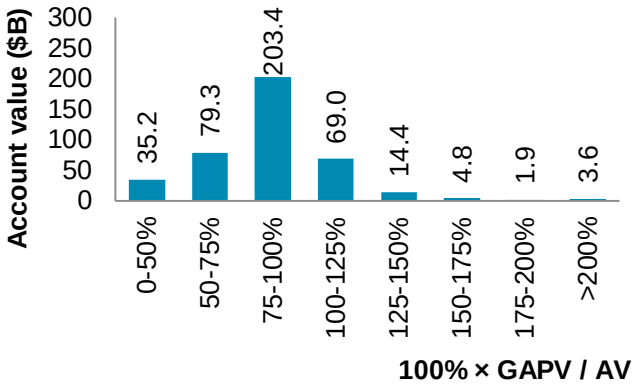


Figure 23: total annual exposure, first year after CDSC period, conforming withdrawers

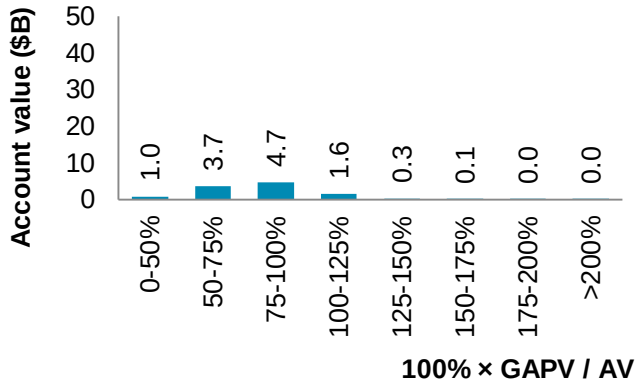


Figure 24: total annual exposure, first year after CDSC period, other policies

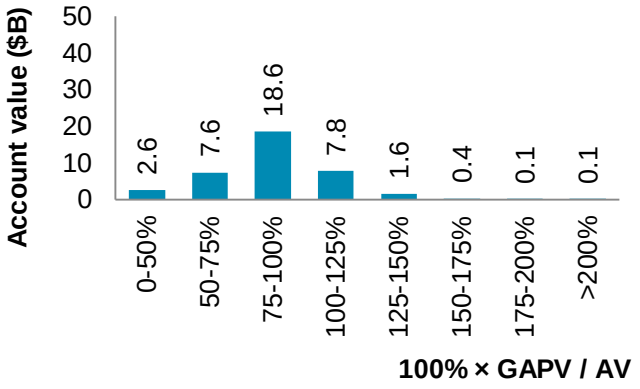


Figure 25: total annual exposure, subsequent post-CDSC years, conforming withdrawers

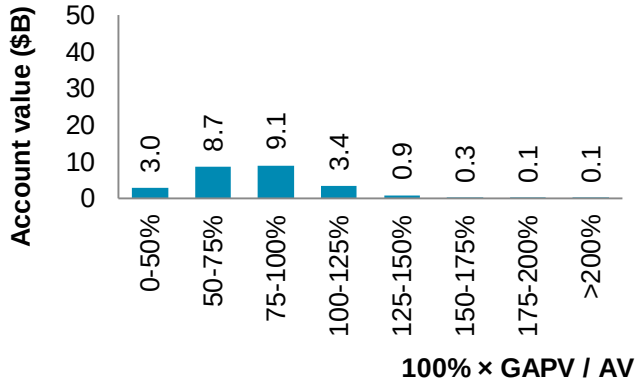


Figure 26: total annual exposure, subsequent post-CDSC years, other policies

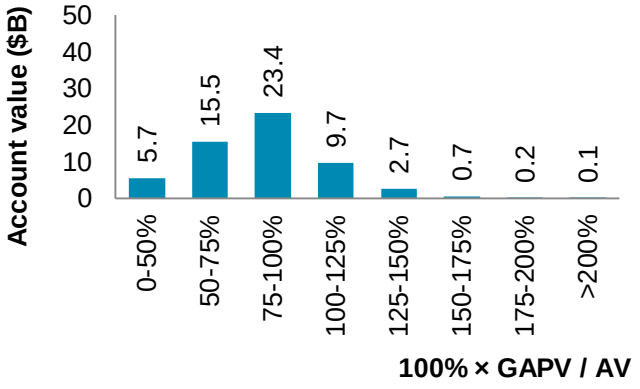


Figure 27 to **Figure 30** below illustrate the total annual exposure in the GMWB full surrender study dataset, for contracts with no surrender charge periods, across the key segmentation dimensions in the study – i.e.,

i. Whether the policy is in the first three policy years or subsequent policy years;

ii. Whether the policy is a conforming withdrawer, as defined in **Table 10**; and

iii. Guarantee in-the-moneyness bands, as calculated via the GAPV-based approach outlined in **Table 10**.

Similar to the dataset for contracts with non-zero surrender charge periods, relevant experience data for contracts with deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years.

Figure 27: total annual exposure, first three policy years, conforming withdrawers

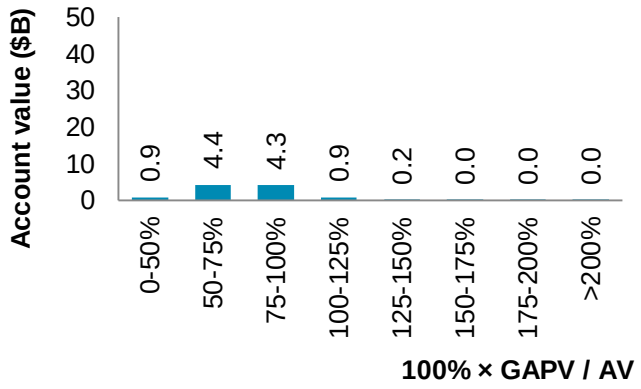


Figure 28: total annual exposure, first three policy years, other policies

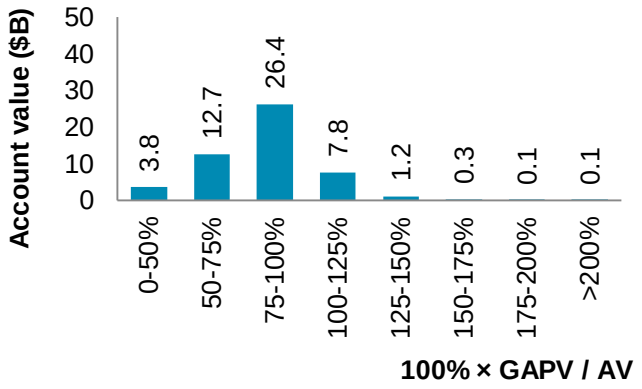


Figure 29: total annual exposure, subsequent policy years, conforming withdrawers

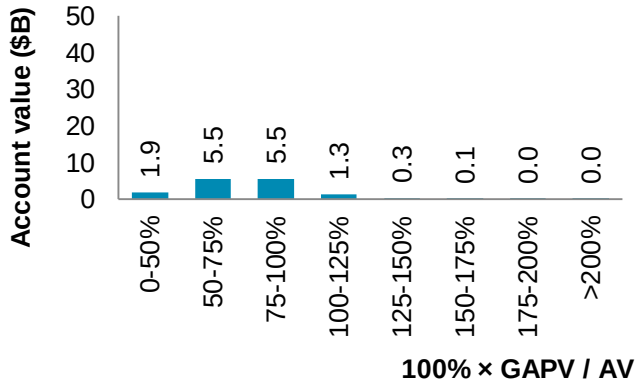
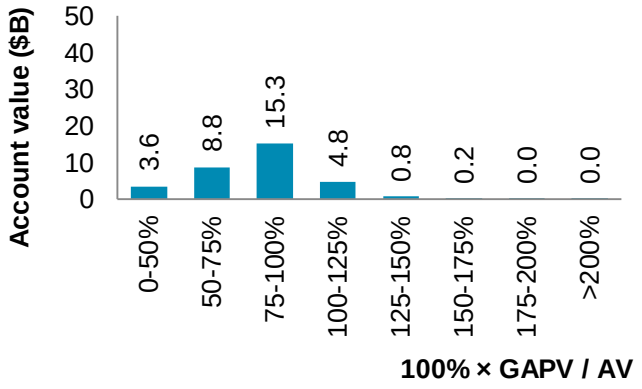


Figure 30: total annual exposure, subsequent policy years, other policies



7.5.2.2. Study results

Results of the study were presented as the aggregated total lapse rate – i.e., including both full surrenders and excess withdrawals – segmented by policy year relative to the surrender charge period, guarantee in-the-moneyness, and whether the contract is a conforming withdrawer. For each key segment, Oliver Wyman also presented the 95% Clopper-Pearson binomial proportion confidence interval for the full surrender rate, calculated in a manner consistent with Section 7.2.3. Note that these confidence intervals were meant only to illustrate the

uncertainty in the calculated full surrender rates and do not capture potential additional variation in the calculated excess withdrawal rates.

Oliver Wyman observed the following total lapse rate behaviors within the surrender charge period for contracts with non-zero surrender charge periods:

i. The observed lapse rate is lower than 5.0% for most guarantee in-the-moneyness segments, with slight sensitivity to in-the-moneyness levels – though close to no sensitivity to whether the policyholder is a conforming withdrawer;

ii. Excess withdrawals constitute approximately 0.5% of the overall lapse rate across most in-the-moneyness levels, except for contracts that are far out-of-the-money where the excess withdrawal rate was higher than 0.5%; and

iii. Credibility is low for deeply in-the-money contracts that are conforming withdrawers, as indicated by the wide confidence intervals in the chart.

As discussed in Section 7.5.3, in our study, we have chosen to group all contracts with guarantee in-the-moneyness in excess of 200% together as one data point in light of data limitations for contracts with deeply in-the-money guarantees. However, we note that most of the data points underlying this in-the-moneyness group are concentrated near 200%, and that data for contracts with deeper in-the-moneyness are very sparse. We expect lapse rates to continue declining gradually as guarantee in-the-moneyness increases further beyond 200%, and note that future studies – if conducted with more experience data – should break out this in-the-moneyness segment further.

Oliver Wyman observed the following total lapse rate behaviors in the first policy year after the surrender charge period for contracts with non-zero surrender charge periods:

i. The observed lapse rate can reach 15-20% for contracts that are far out-of-the-money on a GAPV basis;

ii. The observed lapse rate exhibits substantial sensitivity to guarantee in-the-moneyness and whether the contract is a conforming withdrawer, with high in-the-moneyness and conforming withdrawer both associated with decreasing lapse rates;

iii. Deeply in-the-money conforming withdrawers constitute the segment with the lowest lapse rates, with lapse rates close to 1.0%;

iv. Excess withdrawals constitute approximately 0.5% of the overall lapse rate across most in-the-moneyness levels, except for contracts that are far out-of-the-money where the excess withdrawal rate was higher than 0.5%; and

v. Credibility is low for deeply in-the-money contracts – particularly those that are conforming withdrawers, as indicated by the wide confidence intervals in the chart, though the results appear intuitive in the context of the functional dependency observed with respect to moneyness.

Oliver Wyman observed the following total lapse rate behaviors in all subsequent policy years after the surrender charge period for contracts with non-zero surrender charge periods:

i. The observed lapse rate reaches approximately 12% for contracts that are far out-of-the-money on a GAPV basis;

ii. The observed lapse rate exhibits substantial sensitivity to guarantee in-the-moneyness and whether the contract is a conforming withdrawer, with high in-the-moneyness and conforming withdrawer both associated with decreasing lapse rates;

iii. Deeply in-the-money conforming withdrawers constitute the segment with the lowest lapse rates, with lapse rates close to 1.0%; for contracts that are not conforming withdrawers, lapse rates for deeply in-the-money contracts reach approximately 2.5%;

iv. Excess withdrawals constitute approximately 0.5% of the overall lapse rate across most in-the-moneyness levels, except for contracts that are far out-of-the-money where the excess withdrawal rate was higher than 0.5%; and

v. Credibility is low for deeply in-the-money contracts – particularly those that are conforming withdrawers, as indicated by the wide confidence intervals in the chart.

Most of these observations noted above are consistent with the lapse behaviors observed for contracts that have non-zero surrender charge periods, though we note additionally that the experience dataset for contracts with no surrender charge periods was substantially smaller than that for contracts with non-zero surrender charge periods. Accordingly, data credibility issues were exacerbated for contracts with zero surrender charges, with very wide confidence intervals shown for deeply in-the-money contracts. Nevertheless, the general magnitude of lapse rates across the various in-the-moneyness bands is consistent with those observed for contracts with non-zero surrender charge periods, with slightly higher lapse rates for deeply in-the-money contracts.

7.5.2.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.5.2.2 against our recommended prescribed assumptions for the Standard Projection Method.

The comparison exhibit was constructed in the following manner:

i. Full surrender rates from the experience study dataset are identical to those in Section 7.5.2.2;

ii. Full surrender rates recommended by Oliver Wyman were plotted in orange; and

iii. Full surrender rates in the current AG 43 Standard Scenario were plotted in grey.

In the first policy year after the surrender charge period for contracts with non-zero surrender charge periods, the prescribed lapse rates in the current AG 43 Standard Scenario exceed actual experience across most guarantee in-the-money levels. Oliver Wyman's recommended assumptions therefore reduce the prescribed lapse rates to become more in-line with actual experience observed.

As discussed in Section 7.5.2.2, Oliver Wyman did not observe significant differentiation in lapse rates within the surrender charge period between conforming withdrawers and other policies after controlling for guarantee in-the-moneyness. Nevertheless, we decided to maintain such differentiation in the recommendation for simplicity – as substantial differentiation was observed for lapse rates after the surrender charge period, and judged the resultant deviation from actual experience to be acceptable, particularly as the assumption would only apply for the duration of the surrender charge period.

In the first year after the surrender charge period for contracts with non-zero surrender charge periods:

i. The recommendation represents a significant improvement over the prescribed assumption in the current AG 43 Standard Scenario in capturing the existence of the "shock lapse" effect in this policy year;

ii. The recommendation aligns relatively closely to the experience for in-the-money contracts, but appears to deviate for out-of-the-money contracts;

iii. In particular, even though we expect conforming withdrawers to exhibit lower lapse rates than other policies at all guarantee in-the-moneyness levels, results of the experience study seem to indicate otherwise for out-of-the-money contracts – a phenomenon not captured by our recommendations;

iv. However, relevant experience data for out-of-the-money contracts is relatively sparse, as most GMWB products in industry offer benefit step-ups and have seldom experienced a high interest rate environment. We have noted the latter limitation in Section 7.5.3.

While the fit of the recommendation to experience for the first policy year after the surrender charge period is not ideal, given the assumption applies only for one policy year, we judged the recommendation to be sufficiently in-line with the results of the experience study to be appropriate for use in the Standard Projection Method.

In subsequent years after the surrender charge period for contracts with non-zero surrender charge periods:

i. The recommendation represents a significant improvement over the prescribed assumption in the current AG 43 Standard Scenario in capturing a more gradual sensitivity to guarantee in-the-moneyness, instead of the "cliff-like" sensitivity embedded in the current AG 43 prescription;

ii. The recommendation aligns relatively closely to the experience across most in-the-moneyness bands, except contracts with far out-of-the-money guarantees held by conforming withdrawers; and

iii. While relevant experience data for deeply in-the-money contracts is relatively sparse – as noted in Section 7.5.3, the recommendation appears to be in-line with the aggregate results from the experience study for both conforming withdrawers and other policies, with close to no observable conservatism margin implicitly embedded.

In addition, Oliver Wyman notes that the select members of the VAIWG provided explicitly commentary during the in-person meeting in Philadelphia, PA on August 3 that the prescribed lapse rates for in-the-money contracts should not be meaningfully higher than the aggregate results from the QIS II Experience Study.

Overall, Oliver Wyman believes that the recommended lapse assumption for GMWB contracts balances simplicity of application and accuracy in reflecting prevailing participant experience. Importantly as with most other recommended assumptions discussed in this document, Oliver Wyman did not attempt to apply explicit conservatism margins on top of industry experience data – despite the existence of several sources of data limitations as outlined in Section 7.5.3. We elected not to apply additional conservatism margins in recognition that explicit conservatism margins have been applied elsewhere in the full set of recommended assumptions for more at-risk assumptions – as discussed in Sections 7.3.2.3 and 7.4.2.3 – such that the set of recommended assumptions, collectively, should represent a degree of conservatism consistent with the interpretation of "Prudent Estimate".

Lastly, for policies with no surrender charge period, the recommended assumption for the first three policy years is identical to that for contracts within the surrender charge period, while the recommended assumption for subsequent policy years is identical to that for contracts with non-zero surrender charge periods outside the surrender charge period. Although results of the experience study indicate that these contracts may have a moderately different level of sensitivity to guarantee in-the-moneyness than contracts with non-zero surrender charge periods, Oliver Wyman deemed that the difference is not sufficiently material to warrant the additional complexity that would arise by prescribing a separate assumption for contracts with no surrender charge periods.

7.5.3. Limitations

While Oliver Wyman believes that the analytical methodology for the study was appropriate and that the results of the study support the recommendations for behavioral assumption prescriptions within the Standard Projection Method, several key limitations remain and are summarized in Table 18.

Table 18: Limitations of the GMWB full surrender and excess withdrawal study

Limitation	Description
Lack of experience data for deeply in-the-money contracts	Experience for contracts that have deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years In our study, we have chosen to group all contracts with guarantee in-the-moneyness in excess of 200% together as one data point. However, we note that most of the data underlying this in-the-

Limitation	Description
	moneyness group are concentrated near 200%, and that data for contracts with deeper in-the-moneyness are very sparse
Lack of experience data in high interest rate environments	Relatively little to no experience data exist for modern GMWBs in a high interest rate environment, which precludes analysis of whether the GAPV-based in-the-moneyness definition adequately captures the potential interest rate sensitivity of surrender behavior
Lack of variation by GMDB type	The study does not differentiate by GMDB type or richness, as most GMWB contracts in the experience dataset have GMDBs that do not have roll-up features, but are instead return-of-premium or ratchet-only
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns such that persisting policies are increasingly less likely to surrender; ii. Possible improvement in the economic efficiency of policyholders' behavior patterns through time
Potential mortality bias not studied	The study did not consider the potential existence of a mortality bias in surrender behavior, where surrendering policies – particularly those that have in-the-money guarantees – may have higher mortality and are thus not as economically valuable as that which would be expected should the policy have a more typical mortality profile
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential additional data skews or behavioral drivers

7.6. Study details – withdrawal amount for standalone GMDBs

7.6.1. General approach

The standalone GMDB withdrawal amount study aimed to evaluate the withdrawal amount, calculated in a manner consistent with **Table 10** of this document – i.e., the amount withdrawn in each policy year, expressed as a percentage of the contract's account value.

The experience dataset used in this study was a filtered version of the aggregated, policy year-level dataset described in Section 7.2.2, excluding only records that represent policy years in which the policy terminates. The experience dataset was segmented along the same dimensions and in the same manner as the recommended standalone GMDB withdrawal amount assumption outlined in Section A3.2)F)4) of the redlined AG 43 document. In each segment, the aggregate withdrawal amount was calculated as:

$$\text{Withdrawal Amount} = \frac{\sum \text{Total amount withdrawn}}{\sum \text{Contract account value}}$$

The standalone GMDB withdrawal amount analysis as described in this section was first conducted for individual products. Subsequently, the medians of individual product-level results across various segmentation dimensions were presented to facilitate interpretation, as discussed in Section 7.6.2.2, for the same reasons as those underlying the use of simple averages for the GMWB withdrawal delay study, outlined in Section 7.3.1.3.

7.6.2. Results and discussion

7.6.2.1. Data description and credibility

The underlying experience data for the study is characterized below in Table 19.

Table 19: Overview of experience data included in the study

Data attribute	Products with roll-ups	Products without roll-ups
Companies included	4	4
Products included	10	17
Range of calendar years	2006 – 2016	2006 - 2016

Figure 31 and **Figure 32** below illustrate the total annual exposure in the standalone GMDB withdrawal study dataset for rollup and non-rollup GMDBs, respectively, across different policy years.

Figure 31: total annual exposure in rollup GMDB withdrawal study dataset, by policy year

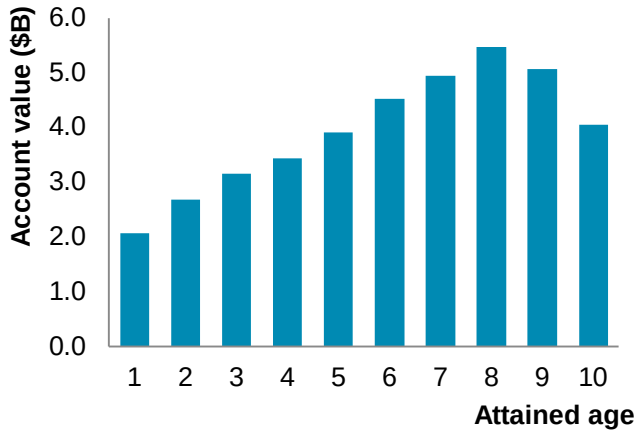
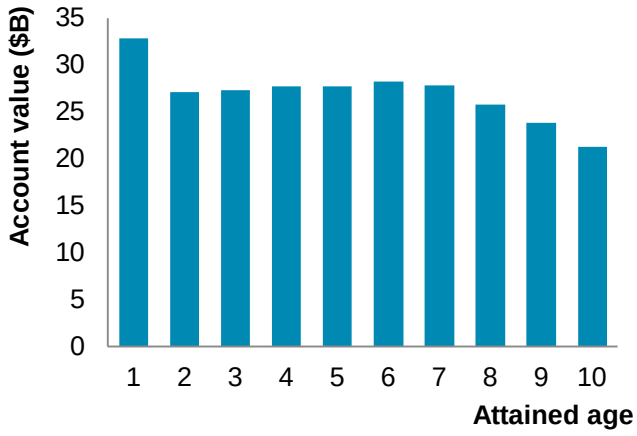


Figure 32: total annual exposure in non-rollup GMDB withdrawal study dataset, by policy year



7.6.2.2. Study results

Results of the study were expressed as aggregate withdrawal rates – i.e., weighted by exposure – across all studied standalone GMDB products, segmented by policy year and the presence of a roll-up feature.

The results demonstrated a rising trend in withdrawal rates – expressed as the ratio between the amount withdrawn in each year and the contract account value – across early policy years, though in later policy years the withdrawal rates appear to plateau. In addition, rollup GMDBs appear to have lower withdrawal rates than non-rollup GMDBs, consistent with expectations as the benefit rollup would act as an incentive to defer withdrawals.

7.6.2.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.6.2.2 against our recommended prescribed assumptions for the Standard Projection Method.

These comparison exhibits were constructed in the following manner:

i. Withdrawal rates from the experience study dataset are identical to those in Section 7.6.2.2;

ii. Withdrawal rates recommended by Oliver Wyman to be used for all standalone GMDB withdrawals projected in the Standard Projection Method were plotted in orange; and

iii. Withdrawal rates in the current AG 43 Standard Scenario prescription for all standalone GMDB withdrawals were indicated to be 0% at all times.

For simplicity and parsimony of parameters, Oliver Wyman decided to recommend that a single withdrawal rate be used for each class of standalone GMDBs across all policy years, instead of prescribing a schedule of withdrawal rates that differ by policy year.

In summary, Oliver Wyman noted that for rollup GMDBs, the recommended assumption appeared to be higher than the aggregate withdrawal rates in the first four policy years, but lower in the subsequent policy years. For non-rollup GMDBs, the recommended assumption was higher than the aggregate withdrawal rates in the first six policy years, but lower in the subsequent policy years.

7.6.3. Limitations

While Oliver Wyman believes that the analytical methodology for the study was appropriate and that the results of the study support the recommendations for behavioral assumption prescriptions within the Standard Projection Method, several key limitations remain and are summarized in **Table 20**.

Table 20: Limitations of the standalone GMDB withdrawal amount study

Limitation	Description
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns; ii. Possible improvement in the economic efficiency of policyholders' withdrawal patterns through time
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential additional data skews or behavioral drivers
Recommended modeling approach in Standard Projection Method simplifies actual withdrawal dynamics	For simplicity and computational tractability, Oliver Wyman recommended that all withdrawals be depicted as a fixed, prescribed percentage of the account value. However, in reality, countless patterns of withdrawal amounts may exist, which may – for GMDB contracts that reduce the benefit base on a dollar-for-dollar basis for withdrawals – have substantially different valuation results than those obtained under the prescribed simplification that all withdrawals are a fixed percentage of the account value

7.7. Study details – full surrender for standalone GMDBs

7.7.1. General approach

The standalone GMDB full surrender study aimed to evaluate the annual lapse rate for standalone GMDB portfolios from full surrenders.

The experience dataset used in this study to evaluate full surrender behavior was the aggregated transition table described in Section 7.2.2, similar to that used for the GMWB withdrawal delay study – but focusing primarily on termination states instead of withdrawal states. The experience dataset was segmented along the same dimensions and in the same manner as the recommended standalone GMDB full surrender assumption outlined in Section A3.2)F)7) of the redlined AG 43 document. In each segment, the aggregate annual full surrender rate was calculated as:

$$\text{Full Surrender Rate} = \frac{\sum \text{Account Value}_{\text{Persisting}\rightarrow\text{Full Surrender}}}{\sum \text{Account Value}_{\text{Persisting}}}$$

The notation for this formula is consistent with that used for the withdrawal state transition rate calculation in Section 7.3.1.1. Accordingly, $\text{Account Value}_{\text{Persisting}\rightarrow\text{Full Surrender}}$ represents the aggregate account value of records that transition from a non-terminated state in policy year t to a full surrender termination state in policy year $t+1$, while $\text{Account Value}_{\text{Persisting}}$ represents the aggregate account value of all records that are in the non-terminated state in policy year t, including those that do not transition to a full surrender termination state in policy year $t+1$.

Note that in the annual full surrender rate calculation, Oliver Wyman has elected to use as the denominator the aggregate account value in the policy year prior to the policy year in which the surrender occurred because of data constraints. In many of the participants' datasets, the account value of the policy year in which the contract surrendered was not available. Nonetheless, Oliver Wyman believes that the full surrender rate calculated in this manner should not lead to material, systematic distortions in the experience study results.

In addition, note that because of the scarcity of experience data for key segments – e.g., contracts with deeply in-the-money guarantees, Oliver Wyman did not present aggregated results via distributions of product-level results or simple averages across products. Instead, Oliver Wyman presented only aggregated results of the study where the aggregate full surrender and excess withdrawal rates effectively represented exposure-weighted participant averages, as shown in Section 7.7.2.2.

7.7.2. Results and discussion

7.7.2.1. Data description

The underlying experience data for the study is characterized below in Table 21.

Table 21: Overview of experience data included in the study

Data attribute	All products
Companies included	3
Products included	10
Range of calendar years	2006 - 2016

Figure 33, **Figure 34**, and **Figure 35** below illustrate the total annual exposure in the standalone GMDB full surrender study dataset across the key segmentation dimensions in the study – i.e.,

i. Whether the policy is in the surrender charge period, in its first policy year after the surrender charge period, or in subsequent policy years; and

ii. Guarantee in-the-moneyness bands, as calculated via the GAPV-based approach outlined in **Table 10**.

We note that, as is evident in these figures, relevant experience data for contracts with deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years. Accordingly, we have noted

this as a major limitation in Section 7.7.3 of this document and identified it as a focal area for future experience studies conducted for the purpose of verifying or updating the prescribed behavioral assumptions within the Standard Projection Method.

Figure 33: total annual exposure, in CDSC period

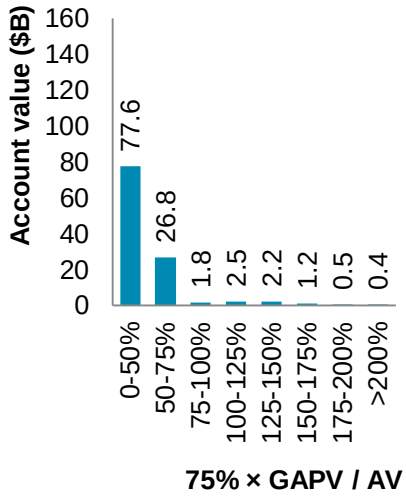


Figure 34: total annual exposure, first year after CDSC period

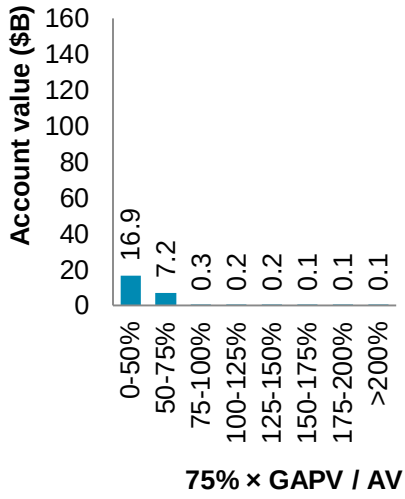
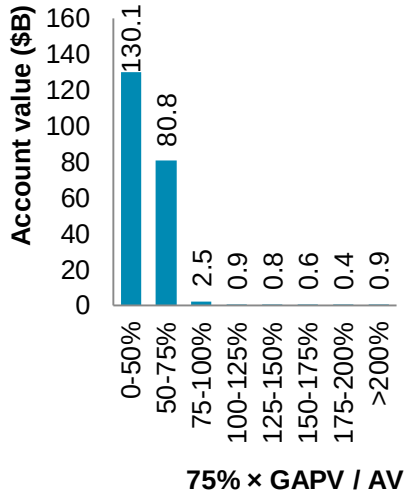


Figure 35: total annual exposure, subsequent post-CDSC years



7.7.2.2. Study results

Results of the study were presented as the aggregated full surrender rate segmented by policy year relative to the surrender charge period and guarantee in-the-moneyness. For each key segment, Oliver Wyman also presented the 95% Clopper-Pearson binomial proportion confidence interval for the full surrender rate, calculated in a manner consistent with Section 7.2.3. Note that unlike the full surrender study for GMWBs, Oliver Wyman only included contracts that have non-zero surrender charge periods.

We noted the following observations:

i. The observed lapse rate in the surrender charge period is lower than 5.0% for most guarantee in-the-moneyness segments, with slight sensitivity to in-the-moneyness levels;

ii. The observed lapse rate can reach 25% for contracts that are far out-of-the-money on a GAPV basis in the first year after the surrender charge period, and approximately 12% in subsequent years after the surrender charge period; and

iii. After the surrender charge period, the observed lapse rate exhibits substantial sensitivity to guarantee in-the-moneyness, with deeply in-the-money contracts exhibiting lapse rates as low as 3.0% after the first year out of the surrender charge period.

As discussed in Section 7.7.3, in our study, we have chosen to group all contracts with guarantee in-the-moneyness in excess of 200% together as one data point in light of data limitations for contracts with deeply in-the-money guarantees. However, we note that most of the data points underlying this in-the-moneyness group are concentrated near 200%, and that data for contracts with deeper in-the-moneyness are very sparse. We expect lapse rates to continue declining gradually as guarantee in-the-moneyness increases further beyond 200%, and note that future studies – if conducted with more experience data – should break out this in-the-moneyness segment further.

7.7.2.3. Proposed assumptions

Oliver Wyman subsequently compared the study results presented in Section 7.7.2.2 against our recommended prescribed assumptions for the Standard Projection Method.

The comparison exhibit was constructed in the following manner:

i. Full surrender rates from the experience study dataset are identical to those in Section 7.7.2.2;

ii. Full surrender rates recommended by Oliver Wyman were plotted in orange; and

iii. Full surrender rates in the current AG 43 Standard Scenario were plotted in grey.

Based on these comparisons, Oliver Wyman judged that results from the standalone GMDB study – when defining guarantee in-the-moneyness as 75% of the ratio between the GAPV and the contract account value – are sufficiently similar to the recommended schedule of lapse rates for GMWB contracts that are not conforming withdrawers. Accordingly, to promote simplicity and parsimony of parameters, Oliver Wyman elected to use the same schedule of GMWB lapse rates – i.e., the Standard Table for Full Surrenders, as detailed in section A3.2)F)7) of the redlined AG 43 document – for projecting standalone GMDB full surrenders in the Standard Projection Method.

7.7.3. Limitations

While Oliver Wyman believes that the analytical methodology for the study was appropriate and that the results of the study support the recommendations for behavioral assumption prescriptions within the Standard Projection Method, several key limitations remain and are summarized in **Table 22**.

Table 22: Limitations of the standalone GMDB full surrender study

Limitation	Description
Lack of experience data for deeply in-the-money contracts	Experience for contracts that have deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years In our study, we have chosen to group all contracts with guarantee in-the-moneyness in excess of 200% together as one data point. However, we note that most of the data underlying this in-the-moneyness group are concentrated near 200%, and that data for contracts with deeper in-the-moneyness are very sparse
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns such that persisting policies are increasingly less likely to surrender; ii. Possible improvement in the economic efficiency of policyholders' behavior patterns through time
Potential mortality bias not studied	The study did not consider the potential existence of a mortality bias in surrender behavior, where surrendering policies – particularly those that have in-the-money guarantees – may have lower mortality and are thus not as economically valuable as that which would be expected should the policy have a more typical mortality profile
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential

Limitation	Description
	additional data skews or behavioral drivers

7.8. Suggestions for ongoing updates

7.8.1. Assumption update process

Oliver Wyman has recommended that a regular review and update process for the prescribed behavioral assumptions within the Standard Projection Method be established to ensure that the prescribed assumptions remain up-to-date in accuracy and robustness to face (i) emerging industry experience data and (ii) ongoing product design innovations.

As limitations driven by experience data shortages will inevitably arise in the assumption review and update process – particularly if there is a high pace in product design innovations, Oliver Wyman anticipates that the update process cannot feasibly be conducted in an algorithmic manner and instead will necessarily involve a degree of judgment. Accordingly, Oliver Wyman would suggest establishing the process such that:

i. The default prescribed assumptions for the Standard Projection Method shall be those in effect immediately prior to the commencement of the assumption review and update process;

ii. The default prescribed assumptions shall remain in effect until the regulators approves promulgation of a new set of prescribed assumptions and such new set of prescribed assumptions enters into effect;

iii. The assumption review shall be executed by an independent party selected by the NAIC;

iv. The assumption review may include other parties – such as representatives from the industry, or other third-party institutions – as the Reviewer judges to be reasonable and appropriate to support the assumption review and any related activities, such as experience studies;

v. Upon completion of the assumption review, the Reviewer shall recommend either that (i) the default prescribed assumptions be maintained or (ii) a new set of prescribed assumptions be implemented;

vi. If the Reviewer recommends that a new set of prescribed assumptions be implemented, the Reviewer shall provide (i) the recommended new set of prescribed assumptions and (ii) supporting evidence to the regulators for decision-making.

In addition, to ensure the integrity of the prescribed assumptions, Oliver Wyman believes that recommendations provided by the Reviewer should be informed by an independent experience study of a similar scale as the QIS II Experience Study, conducted by the Reviewer.

7.8.2. Suggested focus areas

Future experience studies conducted for the purpose of verifying or updating the prescribed behavioral assumptions within the Standard Projection Method should strive to address all limitations outlined in this document. Accordingly, these limitations – outlined in Sections 7.3.4, 7.4.4, 7.5.3, 7.6.3, and 7.7.3 – constitute the suggested focal areas for future experience studies.

Recognizing that these limitations differ in their nature – e.g., some are caused by experience data shortage while others are caused by simplicity or modeling tractability concerns, Oliver Wyman has divided these limitations into several categories, as outlined in **Table 23** to **Table 25** below.

Table 23: experience data limitations; to be addressed when more experience data become available

Limitation	Description	Assumptions affected
Lack of long-dated experience data	Experience data is limited for contracts beyond the tenth policy year. This data limitation precludes robust analysis of the "shock withdrawal" phenomenon that is expected at the end of the deferral bonus period, as most in-force products have 10-year deferral bonus periods	Withdrawal delay for GMWBs
Lack of experience data for deeply in-the-money contracts	Experience for contracts that have deeply in-the-money guarantees is sparse given the lack of a sustained market downturn in recent years In our study, we have chosen to group all contracts with guarantee in-the-moneyness in excess of 200% together as one data point. However, we note that most of the data underlying this in-the-moneyness group are concentrated near 200%, and that data for contracts with deeper in-the-moneyness are very sparse. Future studies – if conducted with more experience data – should break out this in-the-moneyness segment further	Lapse for GMWBs and standalone GMDBs
Lack of experience data in high interest rate environments	Relatively little to no experience data exist for modern GMWBs in a high interest rate environment, which precludes analysis of whether the GAPV-based in-the-moneyness definition adequately captures the potential interest rate sensitivity of surrender behavior	Lapse for GMWBs
Lack of variation by GMDB type	The study does not differentiate by GMDB type or richness, as most GMWB contracts in the experience dataset have GMDBs that do not have roll-up features, but are instead return-of-premium or ratchet-only	Lapse for GMWBs

Table 24: limitations driven by need for simplicity in regulation; to be monitored on a regular basis and addressed if simplifications begin causing substantial inaccuracies relative to experience data

Limitation	Description	Assumptions affected
Potential survival bias or improvement effects not studied	The study did not account for potential changes in policyholder behavior through time, including but not limited to: i. Possible survival bias, or "burn off" effects, where policies exhibit inertia in behavioral patterns; ii. Possible improvement in the economic efficiency of policyholders' behavior patterns through time	All assumptions
Ad hoc withdrawal patterns not differentiated and studied in depth	The study – as well as Oliver Wyman's recommended behavioral assumption prescriptions – focuses on contracts that perform persistent withdrawals to obtain a retirement income stream from the contract. While such withdrawal pattern is dominant in the experience data collected, withdrawals of more ad hoc natures also exist. These were not studied separately, but rather were aggregated with the persistent, income-oriented withdrawals	Withdrawal delay and amount for GMWBs
Potential mortality bias not studied	The study did not consider the potential existence of a mortality bias in surrender behavior, where surrendering	Lapse for GMWBs and standalone GMDBs

Limitation	Description	Assumptions affected
	policies may have higher or lower mortality than the rest of the in-force and are thus not as economically valuable as that which would be expected should the policy have a more typical mortality profile	
Limited additional analyses for skew	The study followed the segmentation dimensions outlined in Proposal 2D of Oliver Wyman's 2016 recommendations. Other segmentation dimensions were not systematically assessed to identify potential additional data skews or behavioral drivers	All assumptions

Table 25: modeling tractability-related limitations; to be addressed if the industry's collective modeling capabilities improve to a level that may support more sophisticated modeling approaches

Limitation	Description	Assumptions affected
Recommended modeling approach in Standard Projection Method simplifies actual withdrawal dynamics	For simplicity and computational tractability, Oliver Wyman recommended that: i. All withdrawals for GMWBs be depicted as a fixed, prescribed percentage of the guaranteed maximum annual withdrawal amount; ii. All withdrawals for standalone GMDBs be depicted as a fixed, prescribed percentage of the account value. However, in reality, countless patterns of withdrawal amounts may exist, which may have substantially different valuation results than those obtained under the prescribed simplification.	Withdrawal for GMWBs and standalone GMDBs

7.9. The hybrid governance model

As outlined in Section 2 of this document, a secondary objective of the Behavioral Assumptions Working Group was to develop an elective "hybrid governance model" for Standard Scenario policyholder behavior assumptions. Upon election by a company, the hybrid governance model would specify prescribed methods to be applied to certain company-specific data to calibrate behavior assumptions to be used within the Standard Scenario in place of the prescribed assumptions.

The Behavioral Assumptions Working Group began developing the hybrid governance model based on the prescribed approach within VM-20 for setting Prudent Estimate mortality assumptions for the Deterministic Reserve and Stochastic Reserve calculations. Specifically, the initial prototype of the hybrid governance model contained the following steps:

i. Determine the segmentation dimensions for each policyholder behavior assumption that is prescribed in the revised Standard Scenario; these segmentation dimensions must contain at least those used in the prescribed assumptions, but may be more granular if the company elects to do so;

ii. Calculate, using the company's own experience data, the relevant policyholder behavior rates for each segment; all aspects of each calculation would be fully specified – e.g., weights applied in calculating each rate, the time period over which each rate is measured;

iii. Apply a set of prescribed margins or adjustments to the rates derived from the company's own experience data. These margins or adjustments are meant to apply the same level of conservatism as that used in determining the prescribed assumptions; correspondingly, certain assumptions – e.g., the

withdrawal efficiency ratio for lifetime GMWBs, as discussed in Section 7.4.2.3 – may have larger margins or adjustments than others;

iv. Calculate the credibility of the company experience data using a prescribed method – e.g., Limited Fluctuation – for each segment used in the company's analysis; accordingly, in electing to increase the granularity of its data segmentation, a company would need to consider the tradeoff between greater analytical sophistication and lower data credibility per segment;

v. Within each segment in the company's own experience analysis, blend the company's results with the corresponding rates from the prescribed assumption based on the credibility of the company's data;

vi. Apply a set of programmatic adjustments, if necessary, such that the resultant blended assumption does not have any unintuitive relationships.

This basic structure of the hybrid governance model received favorable feedback from most participants within the Working Group, though there were a number of areas where additional work was deemed necessary. These areas included, but were not limited to:

i. The parameterization of the prescribed margins or adjustments in step (iii);

ii. The credibility study method prescribed in step (iv), as well as the granularity of segmentation at which the credibility study should be conducted – e.g., at the individual segment-level as defined by the company, at the individual segment-level as defined by the prescribed assumptions, or in aggregate across all segments;

iii. The specifications of the programmatic adjustments, if any, in step (vi).

However, because of resource constraints and requested prioritization of the QIS II Experience Study by the Working Group participants, the hybrid governance model was not able to be fully developed within the timeframe of QIS II. In particular, several participants expressed the view that in light of the revised prescribed assumptions, the importance of the hybrid governance model did not warrant significant resource allocation from QIS II at the potential expense of other aspects of the QIS. Nevertheless, Oliver Wyman encourages ongoing development of the hybrid governance model by industry or another third-party organization, and continues to support the hybrid governance model as an elective alternative standard for the prescribed policyholder behavior assumptions in the Standard Projection Method.

8. Other QIS II Working Groups

8.1. Economic Scenario Generator Working Group

As outlined in Section 2, the Economic Scenario Generator Working Group aimed to:

i. Assess the prevalence of using proprietary economic scenario generators in calculating the CTE Amount;

ii. Understand the rationale for such use; and

iii. Develop a governance structure around proprietary scenario generation – insofar as strong rationale exists for continued permission of companies to use proprietary generators.

The Working Group was attended by a subset of QIS II participants.

Initial discussions within the Working Group identified three major reasons for using proprietary economic scenario generators in calculating the CTE Amount – and for continued allowance of such use:

i. **Systemic risk**: given risks involved in scenario generation, undue reliance on one generator increases the impact of any potential errors with the "designated generator";

ii. **Breadth of asset class coverage needed**: separate account funds for VA products may be invested in a wide range of asset classes, and certain funds have complex rebalancing dynamics – e.g., target-volatility funds. These factors make it impractical to provide prescribed scenarios or define a prescribed generator;

   While prescribing scenarios or a generator for a subset of these asset classes and fund types may be a partial solution, it suffers from the weakness that it is difficult to ensure consistency across asset classes and fund types if they are not all developed in an integrated fashion from the same generator;

iii. **Additional, company-specific needs**: for risk management purposes, companies may need to generate a larger number of scenarios or use different time-steps than those from the prescribed generator – e.g., if the company has a hedge program that requires frequent rebalancing. Proprietary scenario generators can ensure that these specific requirements are more easily met and tailored to each company's risk-management needs.

The Working Group subsequently discussed a governance structure similar to Oliver Wyman's eventual recommendation that was exposed on December 1, 2017, which drew heavily from the scenario governance framework current implemented in VM-20. Specifically, the governance structure would:

i. Define the same prescribed scenario generator as that in VM-20;

ii. Require that all separate account funds be mapped to a linear combination of the funds produced by the prescribed scenario generator;

iii. Allow the use of proprietary scenario generators for funds that cannot be appropriately mapped to a simple linear combination of funds from the prescribed scenario generator, subject to disclosure requirements and the general principle of "no consistent outperformance without higher risk".

However, because of the rationale outlined above, participants attending the Working Group supported an additional provision whereby the "standardized" funds generated by the prescribed scenario generator may be replaced by comparable funds generated by a proprietary scenario generator as long as certain properties were satisfied. To this end, three types of properties were proposed as potential options:

i. **Calibration criteria**: similar to the governance of projected US diversified equity fund returns under the existing AG 43 framework, proprietary scenarios need to meet a set of calibration criteria around their tail characteristics – as well as correlations constraints among different asset classes;

ii. **Simple option valuation**: the CTE Amount for the payoff from a panel of standardized options – e.g., simple equity puts and calls – would be calculated under both proprietary scenarios and scenarios generated by the prescribed generator. The proprietary scenarios may be used if and only if the CTE Amounts from the two calculations are sufficiently close;

iii. **Full portfolio valuation**: the CTE Amount for the full portfolio would be calculated under both proprietary scenarios and scenarios generated by the prescribed generator. The proprietary scenarios may be used if and only if the CTE Amounts from the two calculations are sufficiently close.

In light of the substantial impact that economic scenarios have on the CTE Amount, Oliver Wyman ultimately decided to proceed with the third proposal – i.e., full portfolio valuation – in the final recommendations in order to provide regulators due robustness in governing a paramount input into the calculation.

8.2. Reinsurance Issues Working Group

As outlined in Section 2, the Reinsurance Issues Working Group aimed to ensure that the recommended framework revisions:

i. Do not inadvertently create unintended and undue consequences – whether financial or operational in nature – for reinsurers of such portfolios;

ii. Are sufficiently clear and well-specified such that reinsurers of such portfolios may unambiguously apply the intended methodologies therein to their assumed portfolios.

The Working Group was attended by both QIS II participants and select non-participating reinsurers who were invited to attend in light of having significant assumed VA exposure. Because non-participants were permitted to attend, Oliver Wyman only presented materials related to QIS II that were made available to the public prior to each Working Group meeting.

Discussions within the Working Group drove the development of two specific additional recommendations within the redlined AG 43 document:

i. Certain reinsurers do not have visibility into the direct writer's general account earned rate; for purposes of determining the fixed account crediting rate that should be used in the Standard Projection Method, Oliver Wyman added a prescribed assumption for such reinsurers in section A3.2)E);

ii. Reinsurers within the Working Group reflected that the minimum maintenance expense assumption prescribed for the Standard Scenario in Oliver Wyman's 2016 recommendations were substantially higher than actual expenses incurred for policies for which the company is not responsible for administration – e.g., liability assumed in rider-only reinsurance transactions. To address this issue, Oliver Wyman benchmarked the actual maintenance expenses incurred for such policies among the participants of the Working Group and included a separate minimum maintenance expense assumption for such policies in section A3.2)F)2) of the redlined AG 43 document.

9. VAIWG meetings

Throughout QIS II, Oliver Wyman held a series of trilateral meetings – typically on a weekly basis – between Oliver Wyman, regulators from the VAIWG, and the QIS II participants, with additional participation from the NAIC and representatives from the ACLI. Given the often-technical nature of the subject matter being studied in QIS II, these meetings aimed to provide the VAIWG additional details across the full range of topics within QIS II, beyond those shared in the public VAIWG meetings to facilitate maximum understanding for purposes of tweaking the proposed testing parameters. In addition, these trilateral meetings served as a forum for members of the VAIWG to provide Oliver Wyman guidance on the direction of the QIS, as well as preliminary thoughts on certain potential revisions to the framework.

The series of VAIWG meetings began on March 15, 2017 with a discussion of the scope of QIS II and the desired communication protocol throughout the process. Subsequently, Oliver Wyman covered seven topics with the VAIWG and QIS II participants across 21 trilateral meetings that spanned until the end of October 2017:

i. General framework properties

ii. Capital markets scenarios for the CTE calculation

iii. Economic scenario generator for the CTE calculation

iv. Reflection of revenue sharing in the CTE calculation

v. Governance of CDHS in the CTE calculation

vi. Standard Scenario purpose and construct

vii. Policyholder behavior

The remainder of this section summarizes the content discussed and the meetings held in each topic area. Note that the presentations held to discuss results from each Testing Cycle have been excluded from this section of the document, as they are discussed in greater detail in Sections 4, 5, and 6.

9.1. General framework properties

Two VAIWG meetings were held to discuss the regulators' desired properties of the overall statutory reserve and capital framework for VAs. The specifics of these meetings are summarized below in **Table 26**.

Table 26: VAIWG meetings on general framework properties

Date	Agenda topics	Outcome
May 10	• Review and agree on considerations that should be taken into account for evaluating testing results • Review key focal areas in Testing Cycle 2	Oliver Wyman received select feedback on the considerations proposed from select regulators, which were subsequently reflected in the criteria used to evaluate QIS results and discussed throughout this document
May 31	• Outline nature of balance sheet pro-cyclicality under the current VA statutory framework • Discuss the risk factors that should or should not influence VA funding requirements in a pro-cyclical manner	Key takeaways were as follows: i. Agreement that equity and interest rates, but not volatility or credit spreads, are appropriate risk factors that should influence VA funding requirements in a pro-cyclical manner ii. Support for regularly-scheduled review of

Date	Agenda topics	Outcome
		prevailing framework parameters, but not making regular, discretionary adjustments to the framework

9.2. Capital markets scenarios

Six VAIWG meetings were held to discuss the capital markets scenarios used to conduct the CTE calculation that should be tested as part of the QIS. The specifics of these meetings are summarized below in **Table 27**.

Table 27: VAIWG meetings on capital markets scenarios in the CTE calculation

Date	Agenda topics	Outcome
March 31	• Discuss the equity calibration criteria that should be tested in QIS II – specifically, whether the equity calibration criteria should be linked to prevailing interest rates	The VAIWG requested that an additional meeting be held to continue the discussion before making a decision
April 12	• Continue discussions on the equity calibration criteria that should be tested in QIS II • Potential development of a hybrid governance model for Standard Scenario behavior assumptions	The VAIWG made the following decisions: i. A direct relationship between equities and interest rates should not be tested in QIS II ii. Alternative equity calibration criteria with higher volatility and/or lower mean returns should be tested in QIS II iii. Regulators are open to assessing a potential "hybrid governance model"
April 19	• Review specifics of alternative equity calibration criteria to be tested in QIS II • Discuss the process for testing the alternative equity calibration criteria in QIS II – specifically, in Testing Cycle 2	Oliver Wyman provided illustrations for the VAIWG on potential alternative equity calibration criteria – including (i) higher dispersion across scenarios and (ii) lower cumulative returns The VAIWG decided that Oliver Wyman should test in Testing Cycle 2, via an internal model, sensitivities of the CTE calculation under the revised framework to alternative sets of equity calibration criteria – including those with (i) higher dispersion across scenarios and (ii) lower cumulative returns in the left tail
June 14	• Provide an overview of characteristics of economic scenarios generated by the Academy scenario generator used in VM-20 and their ability to capture historical market dynamics • Outline the historical economic data used in calibrating the current Academy generator in VM-20	No decisions were reached, as the meeting was primarily educational in nature
July 19	• Discuss alternative equity calibration criteria to be tested in Testing Cycle 3	The VAIWG decided that the following criteria should be tested: i. The current equity calibration criteria ii. Equity calibration criteria calibrated based on historical US equity returns, but with an

Date	Agenda topics	Outcome
		extended timeframe that includes data from the Great Depression
August 16	• Outline characteristics of the alternative equity calibration criteria to be tested in Testing Cycle 3 • Discuss details of the Testing Cycle 3 specifications	The VAIWG approved the proposed Testing Cycle 3 specifications and alternative equity calibration criteria to be tested

In 2016, Oliver Wyman recommended, via Proposals 5A and 5B, that alternative calibration criteria or scenarios for equities and interest rates be tested. Specifically, Oliver Wyman proposed that:

i. Interest rate scenarios should be generated using a modified version of the VM-20 interest rate generator, modified to allow for negative interest rates and greater volatility in low interest rate environments; and

ii. Mean equity returns across all stochastic scenarios should be linked to the interest rate mean reversion parameter in the VM-20 interest rate generator as of the valuation date.

Prior to the start of the QIS II, during the public VAIWG meeting held on November 29, 2016, the VAIWG decided that while additional testing should be conducted on the use of the VM-20 interest rate generator in a subsequent QIS, the modifications to the VM-20 interest rate generator suggested by Oliver Wyman should not be tested. Correspondingly, QIS II testing was conducted only with an unmodified version of the VM-20 interest rate generator, and no further discussions around interest rate scenarios were held.

Instead, discussions on the topic of capital markets scenarios in QIS II focused on the equity scenarios that should be tested, as the VAIWG did not make a decision prior to the QIS on whether the linkage of the mean equity return to interest rates should be tested.

In 2016, using the exhibits outlined in **Figure 36** and **Figure 37** of this document, Oliver Wyman noted that:

i. There is substantial “scenario generation risk” embedded in a reserve and capital framework that relies on what are commonly referred to as “real-world” economic scenarios, as such scenarios fundamentally reflect a subjective view of the distribution of potential future market outcomes;

ii. Even if such scenarios are calibrated against historical data, there is material parameter uncertainty in the calibration, as changes to the calibration window used or inclusion of additional data – e.g., from other geographies – can often cause large changes in the parameters;

iii. The risk is exacerbated if the assumptions and parameters underlying such scenarios are not reviewed – and potentially recalibrated – on a frequent basis against emerging data and evolving market conditions;

iv. The automatic recalibration of the mean reversion parameter within the VM-20 interest rate generator represents an example where such period recalibration is used to mitigate the scenario generation risk;

v. However, no similar mechanism exists for the equity calibration criteria within the current AG 43, which has – with the low interest rate conditions in recent years – effectively reduced the regulatory capital requirement for equity risk taken via variable annuity products over the years.

As such, Oliver Wyman suggested investigating a similar automatic recalibration mechanism as that used in the VM-20 interest rate generator, but applied to the equity calibration criteria. Specifically, Oliver Wyman proposed that the equity calibration criteria at each point be linked to changes in the mean reversion parameter from the VM-20 interest rate generator on an annual basis, as illustrated in **Figure 38.**

Figure 36: exhibit used in the recommendations presentation on August 23, 2016 providing rationale for Proposal 5B

Appendix G. Proposal 5B | Evaluate alternative calibration criteria for equities and other market risk factors

As previously discussed, we believe that funding requirements should be calculated based on a market-informed "real-world" tail measure

Market-consistent ⟶ "Real-world"

	1 Market-consistent	2 Market-informed "real-world" tail	3 Fixed "real-world" tail measure
Description	• Fair value of all contract cash flows • Theoretical price to hedge all market risk with no hedge ineffectiveness	• Assets required to defease liabilities in tail scenarios • Scenario generator follows historical data, but updated regularly based on prevailing conditions	• Assets required to defease liabilities in tail scenarios • Scenario generator follows historical data and does not change across reporting periods
Advantages	• Ensures regulators can elect at any time to fully de-risk the portfolio	• Follows current statutory construct • Reduces risk of being "incorrect" on scenario generation	• Aligns with current statutory construct • Low impact on current funding and reduces liability pro-cyclicality
Limitations	• Raises funding requirements • Introduces pro-cyclicality in liabilities – though can be offset by hedging • Departs substantially from current statutory framework	• Reliant on "correctness" of scenario generator – though risk is mitigated through regular updates • Misaligned with conditions in which hedge assets are priced, but again mitigated through regular updates	• Reliant on "correctness" of scenario generator and risks under-funding if scenario generator is "incorrect" • Misaligned with prevailing conditions in which hedge assets are priced
		Proposed revised framework	Most similar to current framework

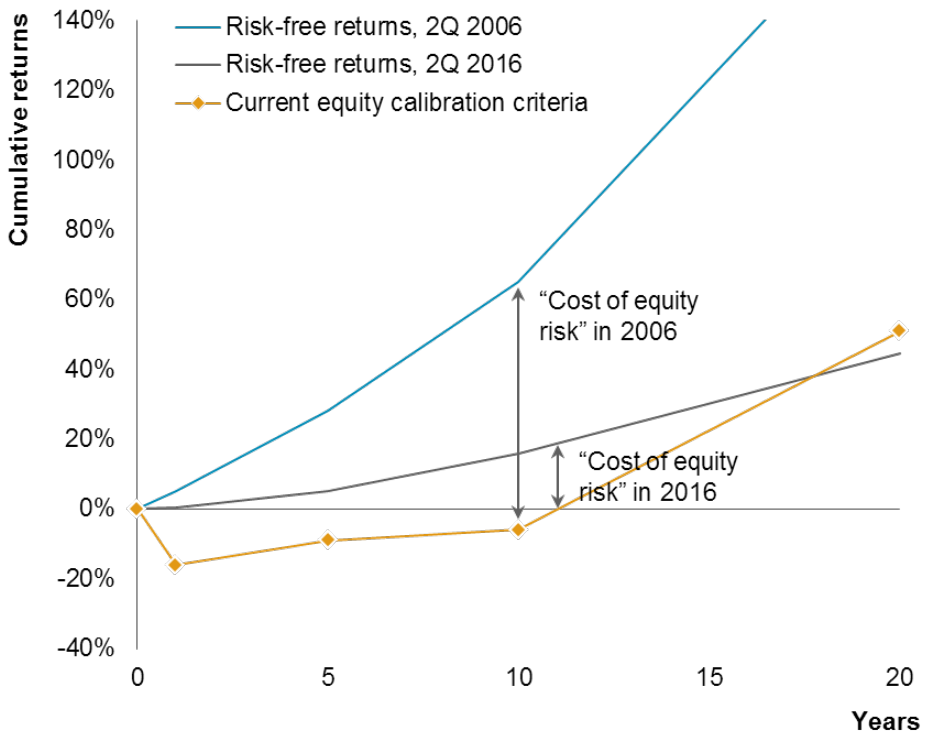
However, "real-world" scenarios reflect a subjective view of potential market outcomes – with normative parameters influencing the likelihood of adverse outcomes

Figure 37: exhibit used in the recommendations presentation on August 23, 2016 providing rationale for Proposal 5B

Appendix G. Proposal 5B | Evaluate alternative calibration criteria for equities and other market risk factors

Persistently-low interest rates in recent years have created a lower "cost" to back guarantees with equities under current equity calibration criteria

Equity calibration criteria at the 5th percentile vs. risk-free returns
Cumulative returns



Commentary

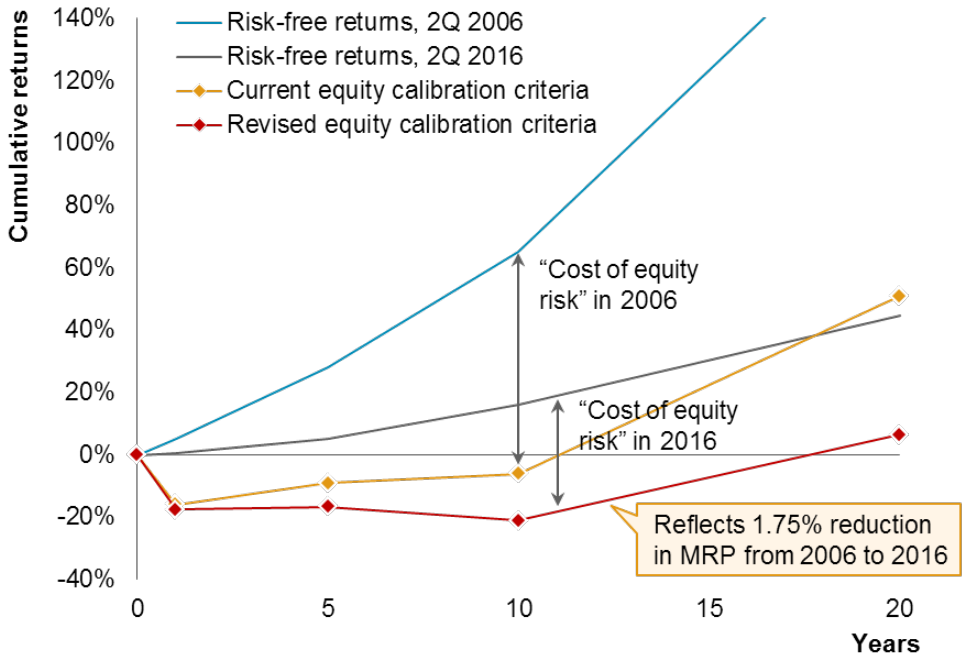
- In the chart on the left, we define "cost of equity risk" as the difference in cumulative returns between:
  - The 5th percentile equity calibration criteria under current AG43 framework
  - A risk-free asset
- Because the persistently-low interest rate environment, the "cost of equity risk" 10 years from Valuation Date has declined considerably since 2006
  - **With 2006 interest rates**: ~40%
  - **With 2016 interest rates**: ~20%
- As a result, the funding requirement implications of holding equity vs. risk-free assets in the underlying funds have become more similar in recent years

Figure 38: exhibit used in the recommendations presentation on August 23, 2016 providing rationale for Proposal 5B

Appendix G. Proposal 5B | Evaluate alternative calibration criteria for equities and other market risk factors

We support investigating updating the equity calibration criteria annually based on the change in long-term interest rate, per the NAIC's MRP formula

Equity calibration criteria at the 5th percentile vs. risk-free returns
Cumulative returns



Commentary

- Under the proposed approach, the equity calibration criteria **R** at each point would be updated via the following formula:

$$\mathbf{R}' = \frac{\mathbf{R}}{\exp(-\Delta\text{MRP} \times \text{T})}$$

  Where T is the number of years
- The proposed approach maintains the "cost of equity risk" over time – and is self-adjusting as economic conditions change
- Short-term interest rate movements would have limited impact on equity calibration criteria, as MRP only moves meaningfully for sustained interest rate stress
- Hence, this mechanism would be relatively slow-moving and primarily intended to capture potential "regime shifts"
- A critical decision will be the "anchor MRP" used to link to current calibration criteria

However, we note that this is a fundamental framework change with potentially large implications; hence, additional study and impact assessments are needed before this proposal is formally introduced

QIS II delved further into this topic through a series of trilateral VAIWG meetings, during which:

i. Oliver Wyman outlined additional rationale – including original research on historical US equity return data and their relationship with prevailing interest rates – supporting the 2016 recommendation in a document delivered to the VAIWG and QIS participants on March 17;

ii. The ACLI compiled a list of counterarguments against Oliver Wyman's 2016 recommendation in a document delivered to the VAIWG and Oliver Wyman on March 29;

iii. Oliver Wyman drafted a brief response to the ACLI document on March 30 to support the VAIWG discussion on March 31.

Oliver Wyman's primary arguments in supporting the 2016 recommendation were as follows:

i. **Support from academic theory**: many academic models and frameworks link excess return – i.e., return in excess of the risk-free rate – with non-diversifiable risk, with the implication that changes in risk-free interest rates should result in equal changes in the expected nominal returns of a risky asset if the risk premium and correlation structure are not interest rate-sensitive;

ii. **Support from empirical data**: historical US equity data underlying the calibration of the current VM-20 economic scenario generator – and therefore the equity calibration criteria – show meaningful correlation with long interest rates. For instance, both the 10-year US Treasury rate and the 20-year US Treasury rate exhibit a correlation in excess of 75% with the subsequent 15-year S&P 500 total return from 1970 onward,[3] indicating that lower long interest rates at a given point in time – e.g., the valuation date – are associated with lower subsequent long-term equity returns after that point in time;

iii. **Support from the existing AG 43 guidance**: section A5.1) of AG 43 stipulates that "as a general rule, funds with higher expected returns should have higher expected volatilities". The guidance and practice note subsequently outline three potential methods for ensuring that this property is satisfied – generating fund returns using a Capital Asset Pricing Model ("CAPM") approach, holding various funds to a consistent Sharpe ratio (referred to as the Market Price of Risk in AG 43), or generating correlated fund return simultaneously – all of which include the risk-free rate of return in their calculation. In the absence of any linkage between interest rate conditions on the valuation date and the equity scenarios, section A5.1) of AG 43 would be violated in high interest rate conditions as fixed income funds would outperform equity funds consistently with lower volatility in excess returns.

The ACLI document delivered on March 29 subsequently contended that:

i. **Academic theory is mixed**: no single theory explains the relationship between interest rates and equity returns, and there are several counteracting theories at play. The equity-interest rate linkage proposed by Oliver Wyman's recommendation 2016 would imply a single, prevailing theory that prevailing interest rates and subsequent equity returns must be highly correlated;

ii. **Empirical evidence is mixed when viewed over a longer period of time**: correlations between prevailing interest rates and subsequent equity returns vary over time, and data from 1871 to 1970 shows an inverse correlation between long interest rates and subsequent equity returns;

iii. **No relationship exists between interest rates and tail scenarios**: historical data demonstrates no relationship between interest rates and the likelihood of tail outcomes; in addition, tail scenarios generated by the current VM-20 economic scenario generator already includes many severe events and a range of correlations between interest rates and equity returns;

[3] At the time of the research, Oliver Wyman did not have access to S&P 500 total return data prior to 1970. However, re-conducting the analysis with S&P 500 price return since 1957 showed correlations in excess of 80% with both the 10-year and 20-year US Treasury rates at the beginning of the 15-year S&P return window.

iv. **Existing AG 43 guidance does not support recommendation**: the existing guidance within AG 43 should be interpreted as addressing the relationship between expected absolute returns and volatility, not expected excess returns and volatility. Additionally, AG 43 allows multiple approaches to be used in order to model fund returns;

v. **Recommendation would create a non-hedgeable risk**: linkage of equity scenarios to the interest rate mean reversion parameter would create an artificial and non-economic risk that is not hedgeable.

As indicated in **Table 27**, the VAIWG on April 12 ultimately decided not to test equity scenarios that are linked to the prevailing interest rate mean reversion parameter, having judged that the arguments put forth by Oliver Wyman as insufficient evidence to demonstrate a direct historical relationship between equity returns and interest rates.

However, at the same time, the VAIWG requested that equity scenarios with lower mean returns, higher volatility, or both lower mean returns and higher volatility should be tested in QIS II, thereby leading to the series of VAIWG meetings after April 12. To avoid arbitrariness in defining the alternative equity scenarios, Oliver Wyman proposed to generate the alternative equity scenarios by changing the scope of historical data used to calibrate the existing VM-20 economic scenario generator, while keeping all other aspects of the generator the same. Specifically:

i. The existing VM-20 economic scenario generator was calibrated using historical US equity return data from 1955 to 2003 – a similar timeframe as that used in Oliver Wyman's defense of the proposal to link equity returns with prevailing interest rate conditions;

ii. As historical data from 1871 to 1955 had been given strong consideration by the VAIWG in the discussion of linking equity returns with prevailing interest rate conditions, Oliver Wyman interpreted this to mean that the calibration of alternative equity scenarios should also contemplate data from this time period;

iii. The calibration of the existing VM-20 economic scenario generator relied on daily time series of US equity total return data to calibrate the stochastic volatility process. Accordingly, Oliver Wyman sought to obtain the longest-possible daily US equity total return time series between 1871 and 2016 for recalibrating the VM-20 generator and producing the alternative equity scenarios.

Ultimately, Oliver Wyman was successful only in obtaining a daily time series of S&P 500 Total Return data dating back to 12/30/1927, with daily data prior to that date unavailable. As such, Oliver Wyman used this historical time series to generate the alternative equity scenarios that were tested in Projection Set 24.

Additional details of the recalibration process are provided in the description of Projection Set 24, outlined in Section 6.1.1 of this document.

9.3. Economic scenario generator

Two VAIWG meetings were held to discuss economic scenario generators used in industry and potential regulatory governance mechanisms for proprietary economic scenario generators. The specifics of these meetings are summarized below in **Table 28**.

Table 28: VAIWG meetings on economic scenario generators

Date	Agenda topics	Outcome
May 17	• Provide regulators an overview of the range of proprietary economic scenario generators in use today • Collect initial views from regulators on governance of proprietary generators	The VAIWG requested additional educational materials on proprietary economic scenario generators – in particular, rationale for deviating from the VM-20 approach of prescribing the Academy generator for major fund types Regulators requested that the Economic

Date	Agenda topics	Outcome
		Scenario Generator Working Group provide an articulation of the rationale and a proposal for scenario generation governance
June 23	• Provide regulators additional education on proprietary economic scenario generators, as requested by the VAIWG in the meeting held on May 17	No decisions were reached, as the meeting was primarily educational in nature

9.4. Reflection of revenue sharing

The level of non-guaranteed revenue sharing recognized by insurers is governed by conflicting standards within the current AG 43 and C3 Phase II frameworks. AG 43 requires a reduction in the recognition of revenue sharing of up to 50% of the insurer's best-estimate, whereas while C3 Phase II requires that companies account for uncertainties in revenue sharing income, it does not specify explicit numerical limits. However, Oliver Wyman's revised C3 charge calculation, which uses a single distribution of scenario deficiencies to calculate both CTE High and statutory reserve, requires the selection of a single revenue sharing recognition approach.

Three VAIWG meetings were held to discuss the methodologies for reflecting non-guaranteed revenue sharing within the CTE calculation that should be tested as part of the QIS. The specifics of these meetings are summarized below in **Table 29**. To inform the discussion, during QIS II the ACLI prepared an experience study to capture the cumulative experience regarding the risk to revenue sharing gained since the formulation of AG 43 and C3 Phase II.

Table 29: VAIWG meetings on reflection of revenue sharing

Date	Agenda topics	Outcome
May 24	• Provide regulators a review of industry revenue sharing experience, collected via a survey conducted by the ACLI • Discuss reflection of non-guaranteed revenue sharing in CTE calculation in light of industry experience	The VAIWG requested additional information from the industry on the survey to clarify specific points and provide greater segmentation The industry also requested that the VAIWG list and prioritize their concerns about the security of non-guaranteed revenue sharing
July 12	• Address items related to the revenue sharing survey shared in the meeting on May 24, as requested by the VAIWG • Collect additional regulator views on the alternative revenue sharing reflection methodologies that should be tested	The VAIWG decided that the following methods should be tested: i. The current AG 43 prescription for reflecting revenue sharing ii. The methodology in Projection Set 18, as outlined in Section 5.1.5, or a variant thereof
August 30	• Address follow-up items requested by regulators from prior meetings related to the revenue sharing survey • Discuss whether the Diversification Benefit Adjustment for the Standard Scenario is needed	Regulator feedback on both topics was mixed On the first topic some regulators maintained that the current revenue sharing method in AG 43 should be retained; on the second topic, one regulator supported removing all diversification benefit – similar to the current AG 43 Standard Scenario – while most did not provide views

Overall, Oliver Wyman believed that the ACLI study bolstered support for smaller reductions to companies' best-estimates of non-guaranteed revenue sharing. However, select regulators expressed concerns about the validity

of the ACLI study, in particular noting the absence of certainty that the study contained data only for non-guaranteed revenue sharing streams. In addition, select regulators expressed concerns regarding a potential risk of affiliated funds which, in the view of Oliver Wyman, justified a differentiation in the treatment between affiliated and non-affiliated funds. Specifically, the regulator concern arose in circumstances where the fund manager and insurance operating entity share a common management team or beneficial owner. In a situation where the fund revenue would be more valuable to the beneficial owner retained in the asset manager than ceded to the insurer, the beneficial owner may take efforts to retain the fund revenue within the asset manager.

Oliver Wyman incorporated the guidance received from these VAIWG meetings into the two alternative revenue sharing reflection methodologies tested in Testing Cycle 2, though we did not have time to test a methodology that differentiated between affiliated and non-affiliated funds. Nevertheless, we believe that the recommendation ultimately presented by Oliver Wyman on December 1, 2017 strikes a good balance between reflecting the experience data collected via the industry survey for non-affiliated funds and addressing the regulator concern regarding affiliated funds. Specifically, Oliver Wyman recommended applying the less stringent C3 Phase II guidelines on non-guaranteed revenue sharing to non-affiliated funds, but maintain the more stringent AG 43 restrictions for affiliated funds in light of the risk described and to incentivize affiliated fund managers to provide guarantees of revenue sharing to safeguard against this risk.

9.5. Governance of CDHS

Three VAIWG meetings were held to discuss potential regulatory governance structures for reflection of companies' CDHS within the CTE calculation. The specifics of these meetings are summarized below in **Table 30**.

Table 30: VAIWG meetings on governance of CDHS

Date	Agenda topics	Outcome
September 13	• Summarize mechanisms for governing CDHS reflection in the CTE calculation in current framework • Outline key challenges – and sources of model risk – in CDHS modeling	No decisions were reached, as the meeting was primarily educational in nature
October 4	• Discuss proposed mechanisms for governing CDHS reflection in the CTE calculation	The VAIWG supported including additional disclosures to govern CDHS and the general principles for projecting implied volatility Select regulators supported disallowing reserves to be reduced by "volatility gains" – i.e., difference between projected realized and implied volatility, though most regulators did not provide views on this proposal
October 25	• Continued discussion on the same topic as that covered on October 4	

Based on feedback from the VAIWG, Oliver Wyman upheld most of the initial recommendations around CDHS governance that were first presented to the VAIWG. Note that in response to the support received from select regulators that reserves should not be reduced by assumed "volatility gains", Oliver Wyman recommended that the Total Asset Requirement – instead of reserves – should not be reduced. We developed this recommendation in recognition that both the reserve and the C3 charge would undergo the CTE calculation in the recommended framework; accordingly, the CDHS governance structure should encompass both quantities in the form of a limitation on the resultant sum – i.e., the Total Asset Requirement.

9.6. Standard Scenario purpose and construct

Three VAIWG meetings were held to discuss the purpose and construct of the Standard Scenario. The specifics of these meetings are summarized below in **Table 31**.

Table 31: VAIWG meetings on Standard Scenario purpose and construct

Date	Agenda topics	Outcome
April 26	• Discuss the regulatory purpose of the Standard Scenario • Discuss the conservatism margin that should be embedded within Standard Scenario behavior assumptions	The VAIWG made the following decisions: i. The Standard Scenario should govern actuarial assumptions, modeling choices, and other components where the CTE calculation incorporates material company judgment ii. Behavioral assumptions prescribed for the Standard Scenario should be calibrated to a "Prudent Estimate" level iii. Where there is no industry data, prescribed assumptions do not need to use the worst-case scenario and can leverage data from analogous behaviors where available In addition, the VAIWG elaborated that the prescribed behavioral assumptions should be set in a way such that they are fit for the purpose of "catching outliers"
May 3	• Discuss types of standardized market paths that should be tested for the Standard Scenario in Testing Cycle 2	The VAIWG agreed that four types of market paths should be tested: i. "Stress and recovery" without CDHS, where market paths are not linked to the forward curve and an initial shock is applied ii. "Stress and recovery" with CDHS, where market paths are not linked to the forward curve and an initial shock is applied iii. "Stressed intrinsic value", where market paths are linked to the forward curve on the valuation date and an initial shock is applied iv. "Intrinsic value", where market paths are linked to the forward curve on the valuation date, but no initial shock is applied
August 30	• Address follow-up items requested by regulators from prior meetings related to the revenue sharing survey • Discuss whether the Diversification Benefit Adjustment for the Standard Scenario is needed	Regulator feedback on both topics was mixed On the first topic some regulators maintained that the current revenue sharing method in AG 43 should be retained; on the second topic, one regulator supported removing all diversification benefit – similar to the current AG 43 Standard Scenario – while most did not provide views

Decisions reached during these VAIWG meetings – particularly around the purpose of the Standard Scenario – led Oliver Wyman to develop the recommendation of replacing the current Standard Scenario construct to the Standard Projection Method outlined in the redlined AG 43 document.

Based on the observations of Testing Cycle 2 results, as outlined in Section 5 of this document, Oliver Wyman judged that the current Standard Scenario construct was ill-equipped to meet its stated objectives. Subsequent results from Testing Cycle 3, as outlined in Section 6 of this document, galvanized our view and supported the recommendation of the Standard Projection Method as a superior means of achieving the stated objectives of the Standard Scenario.

9.7. Policyholder behavior

Two VAIWG meetings were held to discuss policyholder behavior experience and assumptions within the Standard Scenario calculation. The specifics of these meetings are summarized below in **Table 32**.

Table 32: VAIWG meetings on policyholder behavior

Date	Agenda topics	Outcome
June 7	• Provide regulators an overview of policyholder behavior dynamics • Outline key industry observations on policyholder behavior trends • Discuss valuation impact of different assumptions	No decisions were reached, as the meeting was primarily educational in nature
August 3	• Discuss results of the QIS II Experience Study, as outlined in Section 6.2.3.3 • Outline proposed Testing Cycle 3 specifications • Discuss potential alternative constructs for the Standard Scenario, including a construct with a market path that has a company-specific initial stress that is calibrated to the company's CTE results	Key takeaways were as follows: i. No opposition from regulators on the proposed Testing Cycle 3 specifications ii. No oppositions from regulators on the proposed Standard Scenario behavioral assumptions, with one piece of feedback on the prescribed lapse rates for moderately in-the-money policies; the feedback was subsequently incorporated by Oliver Wyman in developing the recommendations iii. No opposition from the regulators against testing the Standard Scenario construct with a company-specific initial stress

In addition, on August 3, 2017, Oliver Wyman presented the full set of finalized QIS II Experience Study results, as well as the revised recommendations for the prescribed behavioral assumptions for the Standard Projection Method, to the VAIWG and QIS II participants in an in-person meeting in Philadelphia, PA. Specifics of the QIS II Experience Study and results thereof are detailed further in Section 6.2.3.3 of this document.

10. Limitations of QIS II

QIS II was structured to evaluate the financial impact of a large number of recommended framework revisions across multiple market environments and ensure the tractability of conducting a standardized set of computations by companies with diverse capabilities. As such, while QIS II allowed Oliver Wyman to converge on a set of recommendations with completed parameterizations, it necessarily had several limitations:

i. The QIS II participants covered only a portion of the total VA industry;

ii. To meet the QIS II timeline, some participants needed to make modeling simplifications or use valuation systems that were different from those used for actual statutory reporting, while other participants elected to opt out of specific tests altogether;

iii. To reduce computational burden, only a small number of market conditions were ultimately tested – i.e., 2016YE market conditions, two parallel yield curve shocks, and one joint equity-interest rate shock;

iv. The portfolios tested were only those of QIS II participants as of 2016YE; potential future portfolio compositions – e.g., if each participant's in-force portfolio were rolled forward in time – were not tested as a result of limitations in participants' computational capabilities;

v. While the immediate statutory balance sheet impact of Oliver Wyman's recommendations was assessed, downstream impact from the recommendations – including potential changes to risk management and pricing strategies – were not tested given recommendations were still in development.

Throughout QIS II, we have used data supplied by participating companies in completing our analyses. While Oliver Wyman has considered such data for overall reasonableness and conducted a series of data verification exercises throughout the QIS as discussed in this document, we did not review each company's data in detail and cannot guarantee the accuracy of each participant's submissions.

In addition, given time constraints, a number of recommendations were not tested fully in QIS II, as outlined below in **Table 33**. However, we note that these recommendations were mostly those that do not affect a company's statutory balance sheet directly, such as disclosure requirements.

For simplicity, we have elected to enumerate these recommendations in a manner consistent with their presentation during the VAIWG call on December 1, 2017.

Table 33: Oliver Wyman's recommendations that were not tested fully in QIS II

Recommendation		Description
3	Allow companies to use proprietary scenario generators if – and only if – they do not reduce TAR	Testing of the revised framework was conducted only with a set of prescribed scenarios generated using the VM-20 economic scenario generator; companies were permitted to use proprietary scenarios only for those funds that they could not map to the fund types from the VM-20 scenario generator
4	Introduce principles to govern implied volatility scenario generation, with a "safe harbor" approach provided	Implied volatility governance was not tested, and companies were permitted to use their existing methods for projecting implied volatility scenarios However, Oliver Wyman believes that this would have only affected a small number of participants that have both (i) a CDHS and (ii) option strategies in their CDHS modeling
7	Follow VM-20 guidance on general account asset projections, with additional constraint on borrowing cost	VM-20 guidance on general account asset projections was fully tested in QIS II. However, the additional constraint on borrowing cost – i.e., that the borrowing cost may not exceed the general account reinvestment rate – was not tested
10	Differentiate treatment of non-	Different reflections of non-guaranteed revenue sharing income

Recommendation		Description
	guaranteed revenue sharing income by affiliated funds vs. non-affiliated funds	were tested – including the methodology that Oliver Wyman ultimately recommended be applied to non-affiliated funds. However, the differentiation in methodology between affiliated and non-affiliated funds was not tested
18	Calculate C3 as the difference between total statutory reserve and CTE 95 on same distribution	The full C3 charge calculation was tested in QIS II except for a small part – the application of a cap, equal to the amount of non-admitted DTAs attributable to the VA portfolio, to the tax-effected difference between the statutory and tax reserves. Instead, QIS II testing assumed that the cap was infinite
19	Permit smoothing to be conducted on the C3 charge, but not on TAR	The smoothing mechanism was not tested, as though different market conditions were tested, QIS II evaluated only portfolios compositions as of 2016YE and did not include roll-forward portfolios due to constraints on participants' computational capabilities
20+	All disclosure requirements	Though Oliver Wyman illustrated the disclosure requirements that were recommended, the production of these disclosures was not tested given time constraints and given that they do not directly affect participants' actual financial statements
25+	Asset admissibility limits	These recommendations were not tested in light of their intuitive impacts on the statutory financials – i.e., the non-admitted portion of these assets that become newly-admitted would flow directly into the company's surplus
27	Endorse hedge accounting for interest rate derivatives that are part of VA hedge programs	This recommendation was not tested for several reasons: i. It is an elective standard and not required; ii. The hedge accounting proposal is under the purview of the Statutory Accounting Practices Working Group ("SAPWG") and not the VAIWG; iii. Evaluation of the proposal would require evaluating roll-forward balance sheets through time, which was not conducted due to constraints on participants' computational capabilities
28	Allocate aggregate reserve to seriatim level based on PV of Accumulated Product Cash Flows	This recommendation was not tested given that it does not affect participants' actual statutory financial statements – only the assignment of statutory reserves to each contract

Qualifications, assumptions and limiting conditions

This report is for the exclusive use of the Oliver Wyman client named herein. This report is not intended for general circulation or publication, nor is it to be reproduced, quoted or distributed for any purpose without the prior written permission of Oliver Wyman. There are no third party beneficiaries with respect to this report, and Oliver Wyman does not accept any liability to any third party.

Information furnished by others, upon which all or portions of this report are based, is believed to be reliable but has not been independently verified, unless otherwise expressly indicated. Public information and industry and statistical data are from sources we deem to be reliable; however, we make no representation as to the accuracy or completeness of such information. The findings contained in this report may contain predictions based on current data and historical trends. Any such predictions are subject to inherent risks and uncertainties. Oliver Wyman accepts no responsibility for actual results or future events.

The opinions expressed in this report are valid only for the purpose stated herein and as of the date of this report. No obligation is assumed to revise this report to reflect changes, events or conditions, which occur subsequent to the date hereof.

All decisions in connection with the implementation or use of advice or recommendations contained in this report are the sole responsibility of the client. This report does not represent investment advice nor does it provide an opinion regarding the fairness of any transaction to any and all parties.

Oliver Wyman
1166 Avenue of the Americas, 29th Floor
New York, NY 10036
NAIC.QIS@oliverwyman.com