AMERICAN ACADEMY of ACTUARIES

Report on

Possible Proposed Test for Material Tail Risk

Presented by
the American Academy of Actuaries' Life Reserves Work Group Modeling Subgroup

To the National Association of Insurance Commissioners'
Life and Health Actuarial Task Force

September 2007

The American Academy of Actuaries is a national organization formed in 1965 to bring together, in a single entity, actuaries of all specializations within the United States. A major purpose of the Academy is to act as a public information organization for the profession. Academy committees, task forces and work groups regularly prepare testimony and provide information to Congress and senior federal policy-makers, comment on proposed federal and state regulations, and work closely with the National Association of Insurance Commissioners and state officials on issues related to insurance, pensions and other forms of risk financing. The Academy establishes qualification standards for the actuarial profession in the United States and supports two independent boards. The Actuarial Standards Board promulgates standards of practice for the profession, and the Actuarial Board for Counseling and Discipline helps to ensure high standards of professional conduct are met. The Academy also supports the Joint Committee for the Code of Professional Conduct, which develops standards of conduct for the U.S. actuarial profession.

Life Reserves Work Group

Dave Neve, F.S.A., M.A.A.A., Chair

Mary Bahna-Nolan, F.S.A., M.A.A.A.
Aryeh Bak, F.S.A., M.A.A.A.
Jeff Beckley, F.S.A., M.A.A.A.
Peter Boyko, F.S.A., M.A.A.A.
Arnold Dicke, F.S.A., M.A.A.A.
Armand DePalo, F.S.A., M.A.A.A.
Bob DiRico, F.S.A., M.A.A.A.
Todd Erkis, F.S.A., M.A.A.A.
Gary Falde, F.S.A., M.A.A.A.
Bruce Friedland, F.S.A., M.A.A.A.
Dieter Gaubatz, F.S.A., M.A.A.A.
Dale Hall, F.S.A., M.A.A.A.
Jane Hamrick, F.S.A., M.A.A.A.
Tom Kalmbach, F.S.A., M.A.A.A.
Barbara Lautzenheiser, F.S.A., M.A.A.A
Esther Milnes, F.S.A., M.A.A.A.
Greg Roemelt, F.S.A., M.A.A.A.
Alan Routhenstein, F.S.A, M.A.A.A.
Karen Rudolph, F.S.A., M.A.A.A.
Barry Shemin, F.S.A., M.A.A.A.
Steve Strommen, F.S.A., M.A.A.A.
Michael Villa, A.S.A., M.A.A.A.

Purpose of this document

This document has been prepared to facilitate discussion by members of LHATF regarding a possible test for material tail risk. The inclusion of such a test has been proposed as a standardized way to qualify for the stochastic testing exclusion. (This test might also be used to support the use of simplified approaches to determine the Reported Reserve.) This document presents one way to define such a test, along with modeling results of applying the test to four sample blocks of business. Input is sought from LHATF as to whether this definition of a test for material tail risk is felt to meet the regulatory need, and to help finalize a proposed approach that can be incorporated in the next draft of the Valuation Manual.

Background and purpose of the test

A fundamental element of the principles-based approach to reserves is to determine the extent to which actual financial results might vary from expectations, and to require a larger margin in reserves in cases where the financial results are more uncertain. Life insurance products with highly uncertain financial results are often considered to have "tail risk." The stochastic component of the principles-based approach is used to determine the extent to which a block of business has "tail risk," that is: that its financial results might vary depending on economic conditions.

However, the stochastic calculations can be voluminous and burdensome, and may not provide much value in cases where financial results are relatively insensitive to economic conditions. Because of this, a "stochastic testing exclusion" has been incorporated in the in the PBR proposal for life products. Business that qualifies for this exclusion could forgo the need to calculate the stochastic reserve and would hold just the deterministic reserve, or perhaps a reserve calculated using some simplified method.

To qualify a block of business for the stochastic testing exclusion, a company would need to demonstrate that the block does not have material "tail risk." The purpose of defining a test for material tail risk is to provide a standard approach to making that demonstration, without requiring full stochastic reserve calculations.

Description of a potential test

A test for material tail risk is designed to measure the variability of financial results by economic scenarios. However, to avoid the need for full stochastic testing, a small number of scenarios is used. The smaller set of scenarios are constructed in a manner to represent a wide range of economic conditions.

For each scenario, the assets required to fund the future benefits and expenses are determined using the GPVAD approach using anticipated experience assumptions. Once the asset requirements for each scenario are obtained, the test for material tail risk is a ratio that is calculated using the results. The ratio is:

$$\text{MTR test ratio} = \frac{\text{(Max. scenario asset requirement – Min. scenario asset requirement)}}{\text{(PV of benefits and expenses in “anticipated experience” scenario)}}$$

A high value for the ratio is an indication of tail risk. The NAIC may want to set some threshold value that distinguishes material tail risk from immaterial tail risk. Some modeling has been done (see later in this report) to help with setting that threshold.

It should be noted that the denominator of the ratio is not a reserve rather, it is a present value of benefits and expenses[1]. This denominator is used in place of a reserve because the use of a reserve in the denominator would always result in a high value for the ratio for products that have small reserves. And, in that case, a high value for the ratio would thus not necessarily be an indicator of material tail risk. Note, that for annual renewable business or business with increasing premiums, the reserve can be small but the present value of benefits can be high. To the extent such products don't exhibit what is typically considered tail risk, using a denominator based on benefits rather than reserves would properly classify them.

The potential test uses a set of 12 scenarios. Given the starting yield curve on the valuation date, the scenarios are created using the American Academy of Actuaries' Stochastic Scenario Generator and predefined sets of random numbers[2].

The rationale for this approach is twofold. First, the scenarios should be realistic, meaning that they could be produced by the generator. Second, it should be possible to measure in some way the likelihood of any scenario occurring.

One way to measure the likelihood of a scenario occurring is to measure the likelihood of its series of random shocks, that is, the random numbers used in the generator. Given any sequence of random numbers, their sum can be compared with a mean of zero and a standard error equal to the square root of the number of deviates in the sequence. With the mean and standard error, we can determine, in a crude way, where the sum of deviates in our sequence lies in the distribution of the sum of all such sequences.

For example, if we want a sequence that is always one standard error above average, we start with a value of 1.0 as the first deviate. The value of the nth deviate is the excess of the square root of n over the square root of n-1. So the second value is 1.414 – 1 = 0.414 and the third value is 1.732 - 1.414 = 0.318.

Generating interest rates

The American Academy of Actuaries' Interest Rate Generator uses three random numbers per period. These are:

1. A random shock to the 20-year Treasury rate

[1] One could consider this is a reserve plus the present value of premiums.
[2] Each random number is a sample from a normal distribution with mean zero and variance 1.

2. A random shock to the spread between 1-year and 20-year Treasury rates
3. A random shock to the volatility

In generating the scenarios for the test, zero shocks to volatility were used.

Also, when generating scenarios for the test, upward shocks to the 20-year Treasury were associated with downward shocks to the spread, making the yield curve less steep (or potentially inverted).

Generating equity returns

The American Academy of Actuaries' Equity Generators (C3 phase 2) provides two random numbers per period. These are:

1. A random shock to make the return more or less than the mean
2. A random shock to the volatility

This potential test used zero shocks to volatility in defined scenarios. However, since the mean is a function of the volatility in the American Academy of Actuaries' model, we needed to select a mean and a volatility to use.

With that in mind, the random numbers that define the scenarios were set up as follows:

Scenario 1 – Pop up, high equity
Interest rate shocks that maintain the cumulative shock at the 90% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 2 – Pop up, low equity
Interest rate shocks that maintain the cumulative shock at the 90% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 3 – Pop down, high equity
Interest rate shocks that maintain the cumulative shock at the 10% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 4 – Pop down, low equity
Interest rate shocks that maintain the cumulative shock at the 10% level (1.282 standard errors).
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 5 – Up/down, high equity, Scenario 7 – Down/up, high equity
Interest rate shocks that, for each five-year period are consistently in the same direction. The cumulative shock for each 5-year period is at the 90% level.
Equity returns that maintain the cumulative equity return at the 90% level.

Scenario 6 – Up/down, low equity, Scenario 8 – Down/up, low equity
Interest rate shocks that, for each five-year period are consistently in the same direction. The cumulative shock for each 5-year period is at the 90% level.
Equity returns that maintain the cumulative equity return at the 10% level.

Scenario 9 – Anticipated experience

This scenario starts with today's yield curve and then slowly applies the mean reversion within the generator to eventually arrive at the generator's mean reversion point (i.e., 5.4% 20-yr. Treasuries). The transition from here to there is slow and smooth, based on exponential decay.
All shocks are zero.

Scenario 10 – Inverted yield curves

Zero shocks to long term rates and equities.
Shocks to the spread between short and long rates that are consistently in the same direction for each three-year period. The shocks for the first three-year period are in the direction of reducing the spread (usually causing an inverted yield curve). Shocks for each subsequent three year period alternate in direction.

Scenario 11 – Volatile equity returns

Zero shocks to interest rates
Shocks to equity returns that are consistently in the same direction for each two-year period, and then switch directions.

Scenario 12 – Deterministic scenario for valuation – premium pattern test

Uniform downward shocks each month for 20 years, sufficient to get down to the 80% point on the distribution of 20 year shocks. After 20 years, shocks are at a level that keeps the cumulative shock at the 80% level (or the 20% level, depending on how you look at it).

Scenario 12 is not directed at testing sensitivity to investment returns. Rather, it is directed at sensitivity to premium payment patterns. For flexible premium products, four patterns are tested (i.e., single premium, level premium, minimum premium, and no future premium) and the one creating the greatest asset requirement is included when measuring the range of results in the numerator of the test ratio.

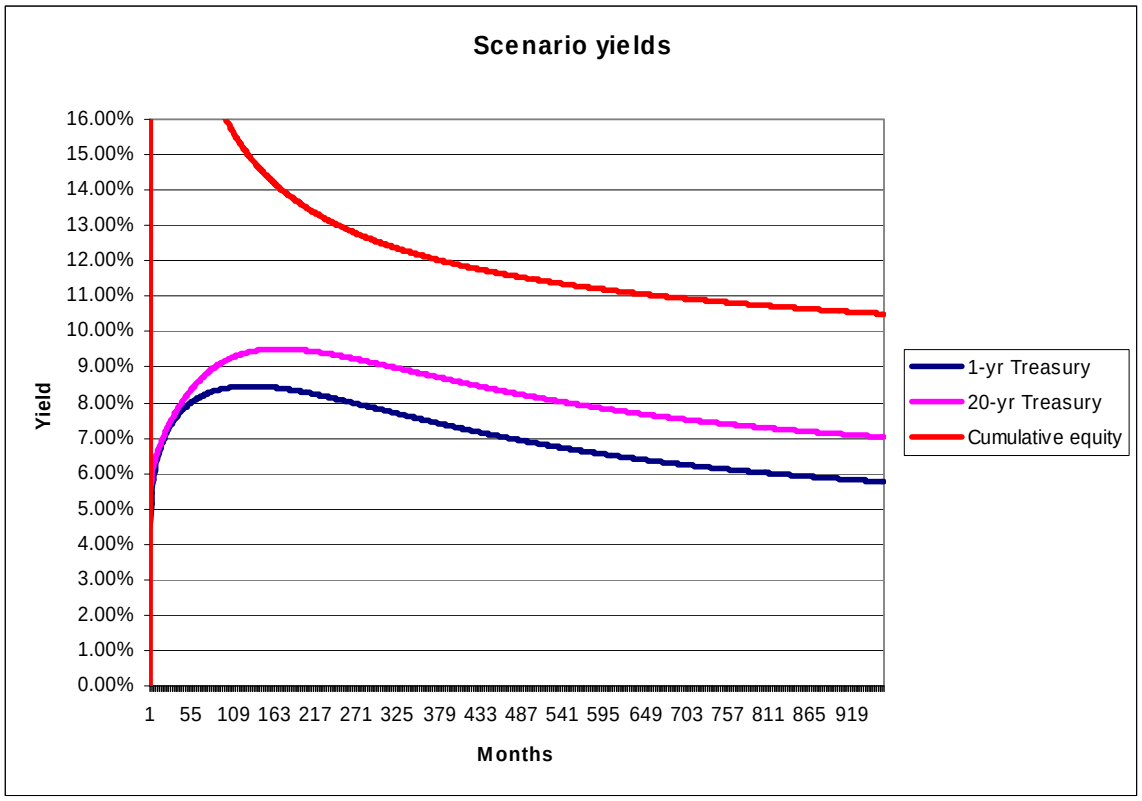
It needs to be pointed out that the deterministic scenario for valuation has not yet been defined by the LRWG, but this is one possible definition that is reasonably consistent with values that have been discussed.

Also, none of these stochastic generators have been finalized by the Economic Scenario Work Group. This work is based on a preliminary working draft.

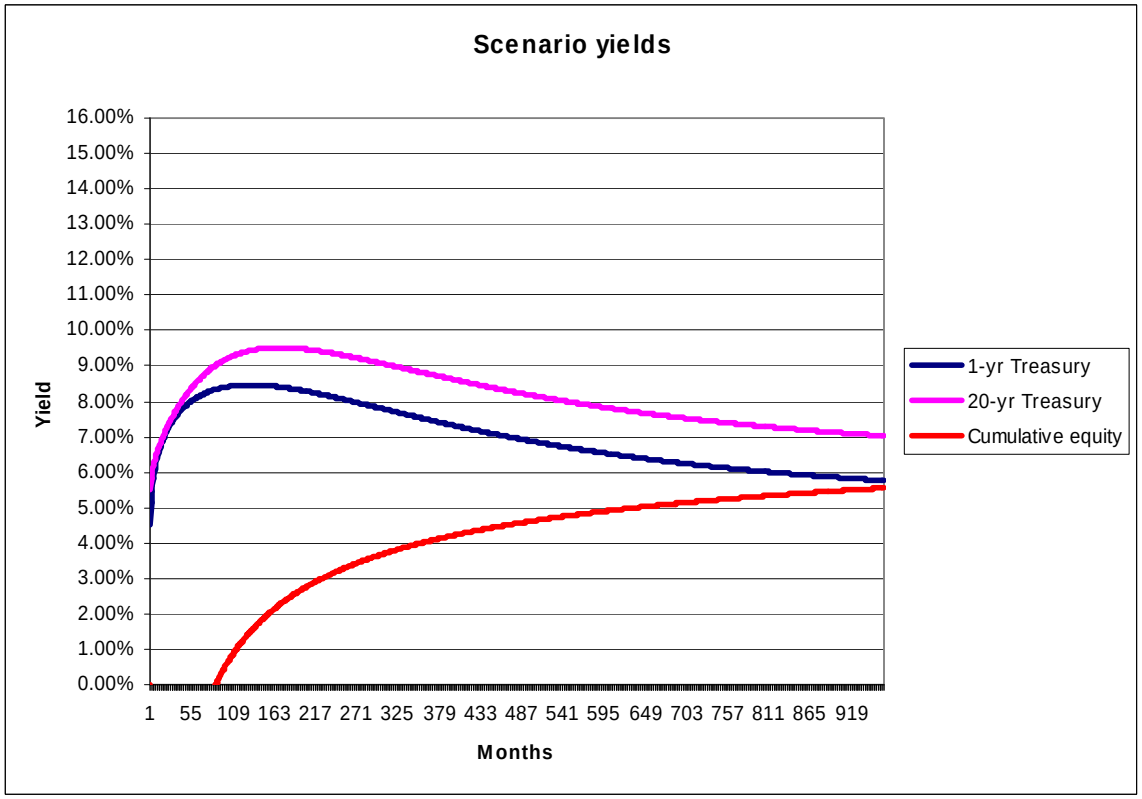
Finally, consideration was given by the Modeling Subgroup of the LRWG to determine the scenarios for the test by selecting outliers from a stochastically generated set of scenarios. That approach was not selected for two reasons. First, it was determined that a larger number of sample scenarios would be required. Second, the process for choosing the scenarios might need to be product-specific, thereby requiring multiple sets of scenarios and a very long technical process of picking them. The Modeling Subgroup of the LRWG is satisfied that the set of 12 scenarios defined above provides an appropriate test.

The graphs that follow show how the 12 defined scenarios could evolve given a starting yield curve that includes a 4.5% one-year rate and a 5.5% 20-year rate.

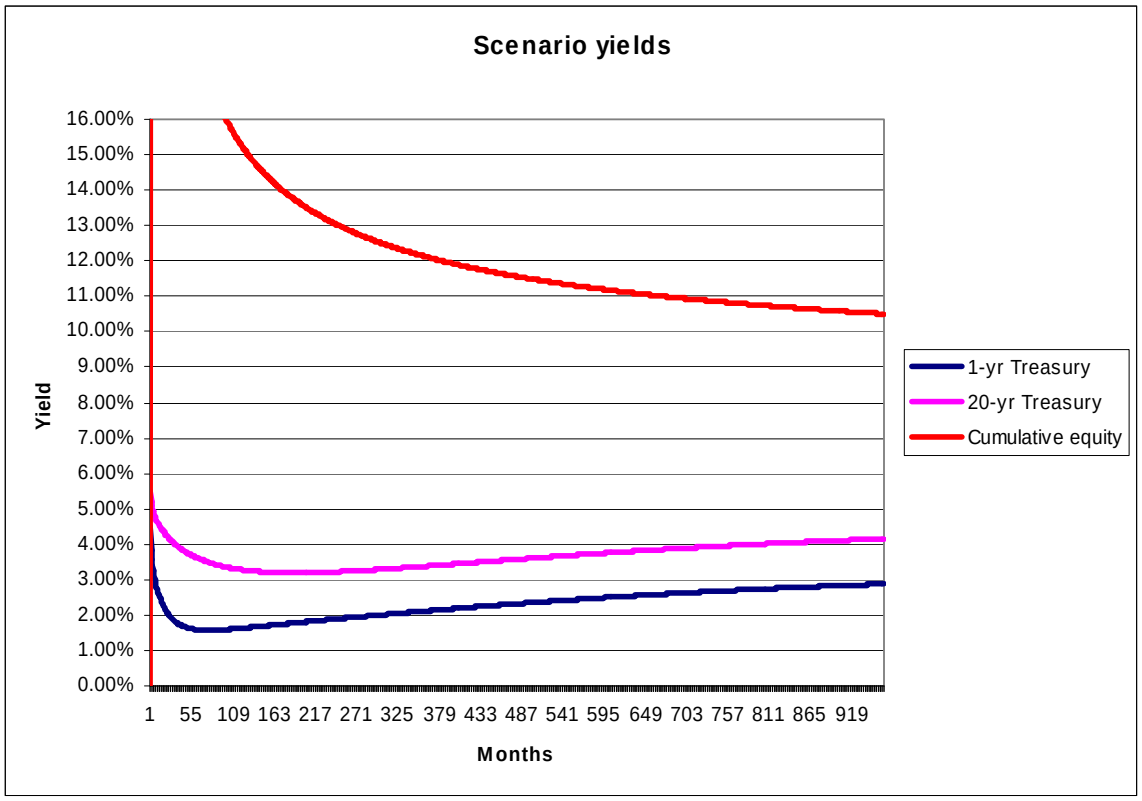
Scenario 1 – Popup, high equity



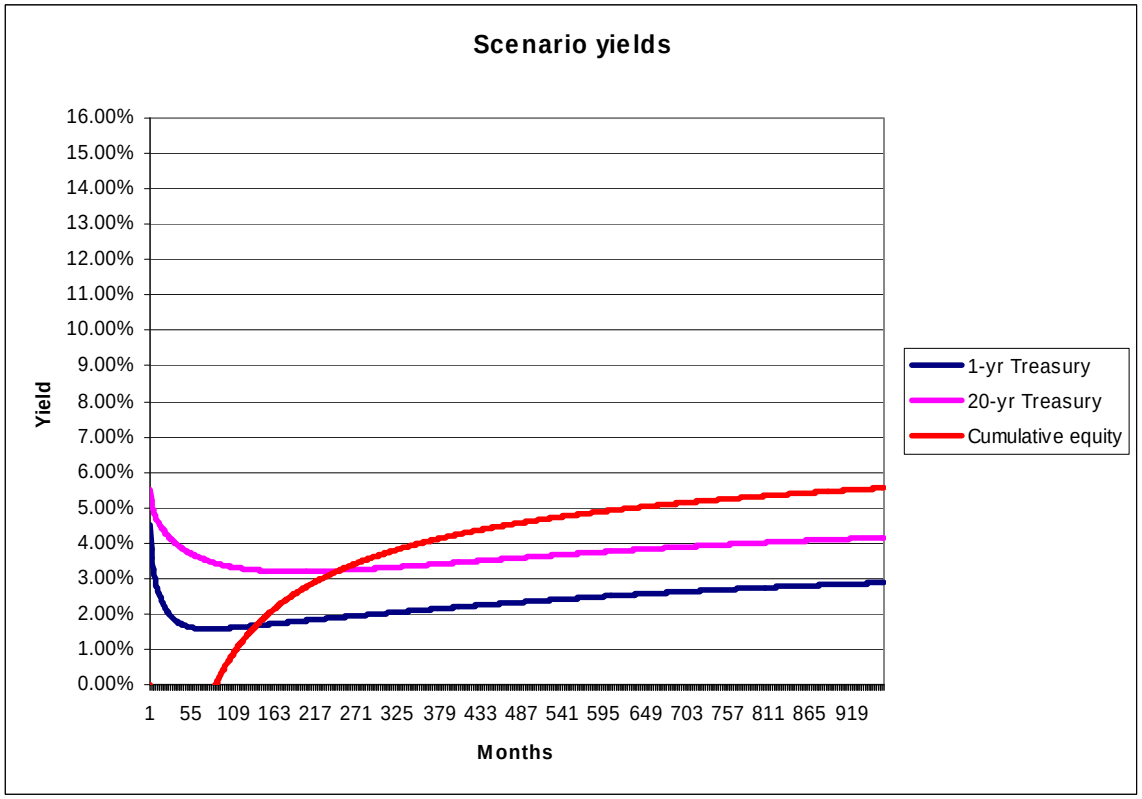
Scenario 2 – Popup, low equity



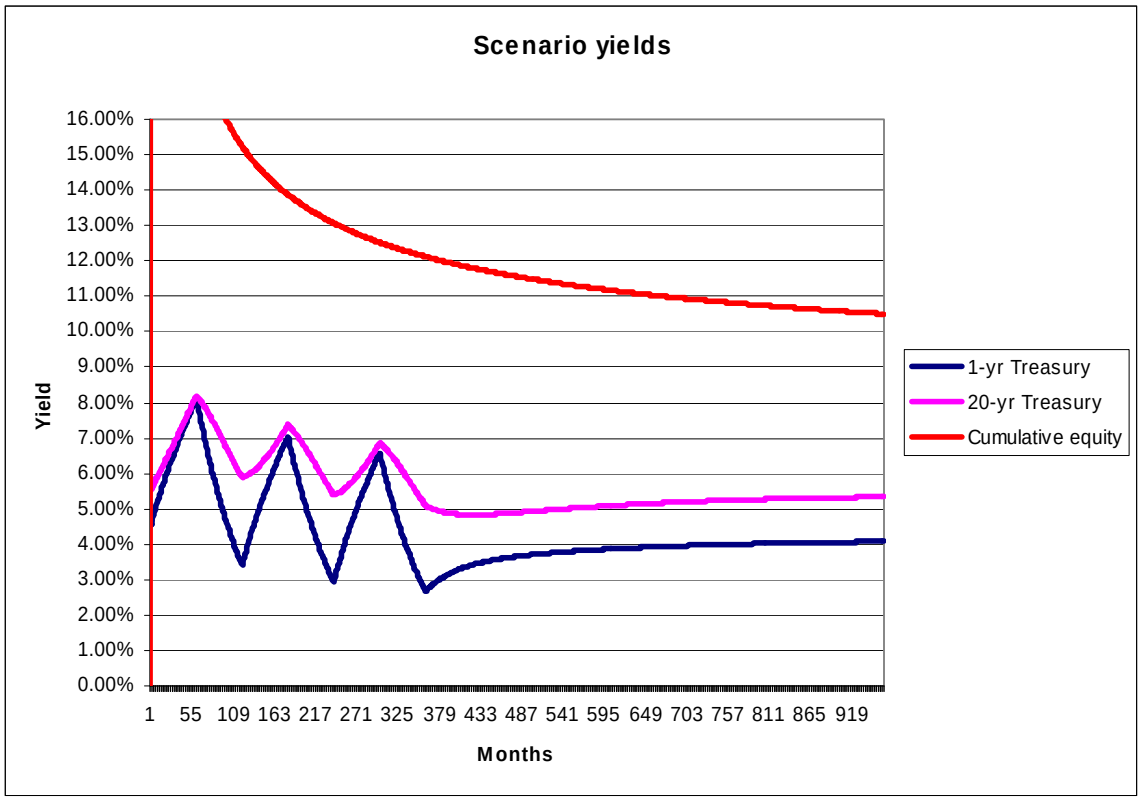
Scenario 3 – Popdown, high equity



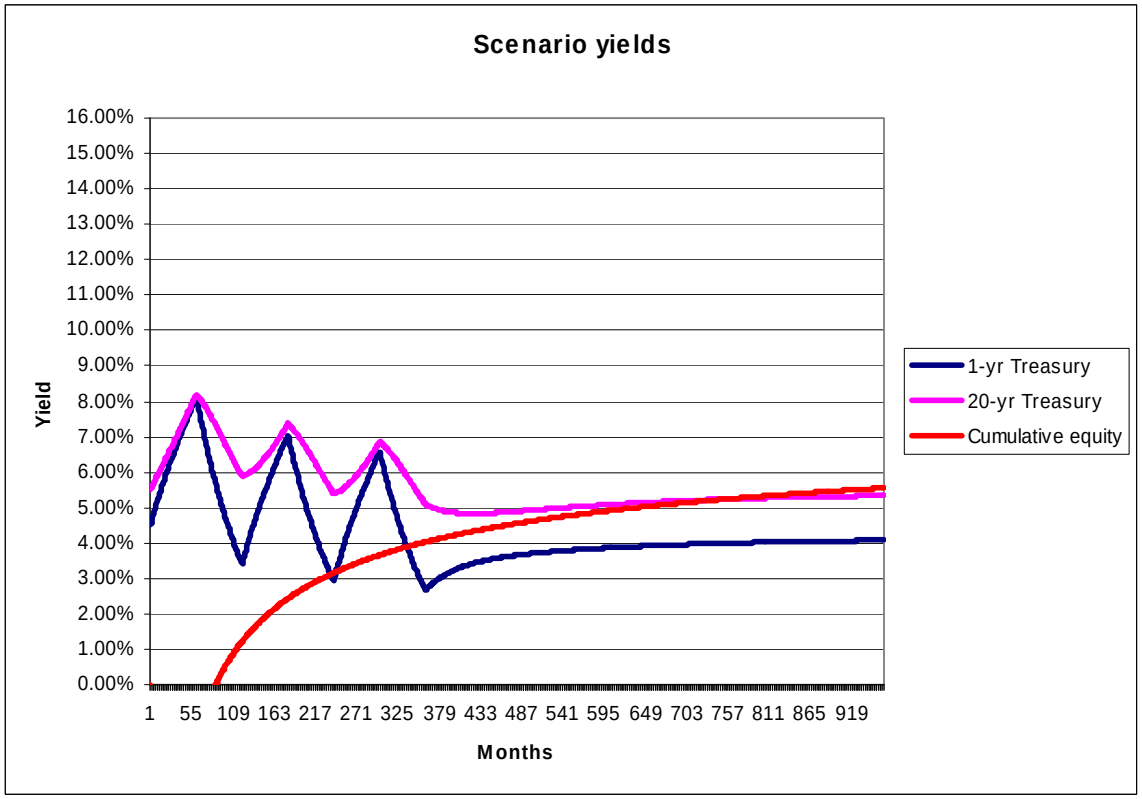
Scenario 4 – Popdown, low equity



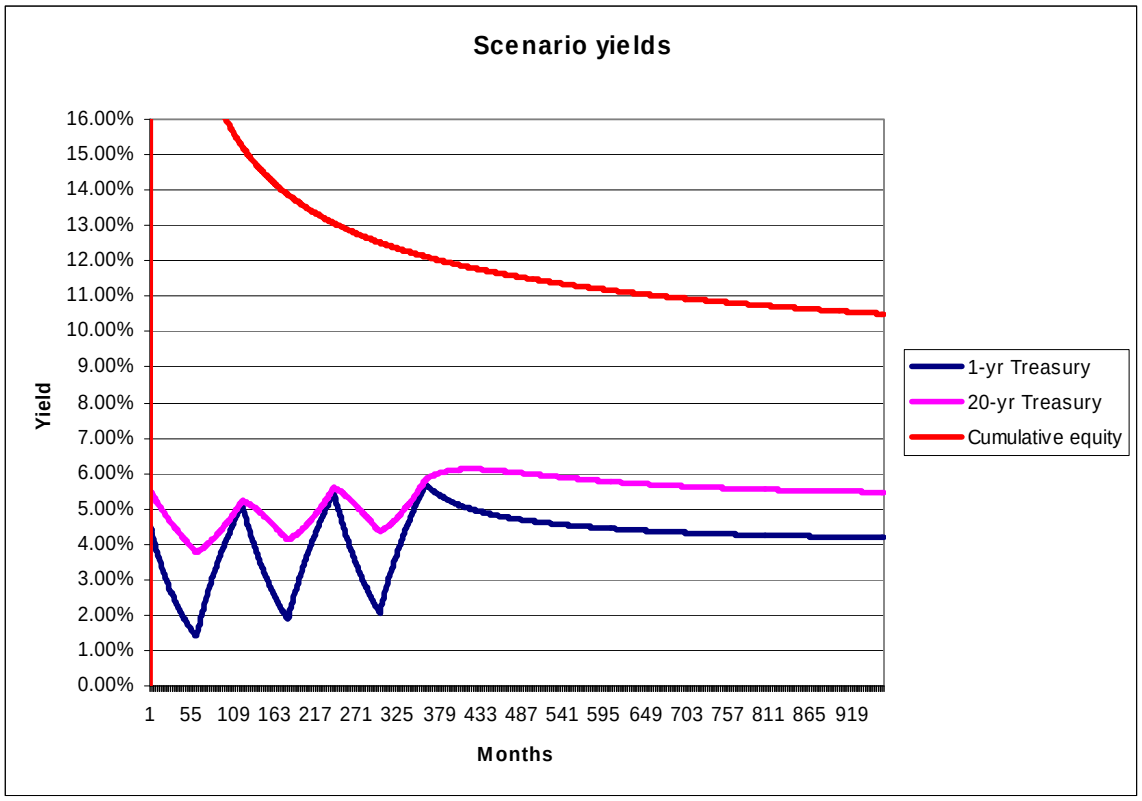
Scenario 5 – Up/down, high equity



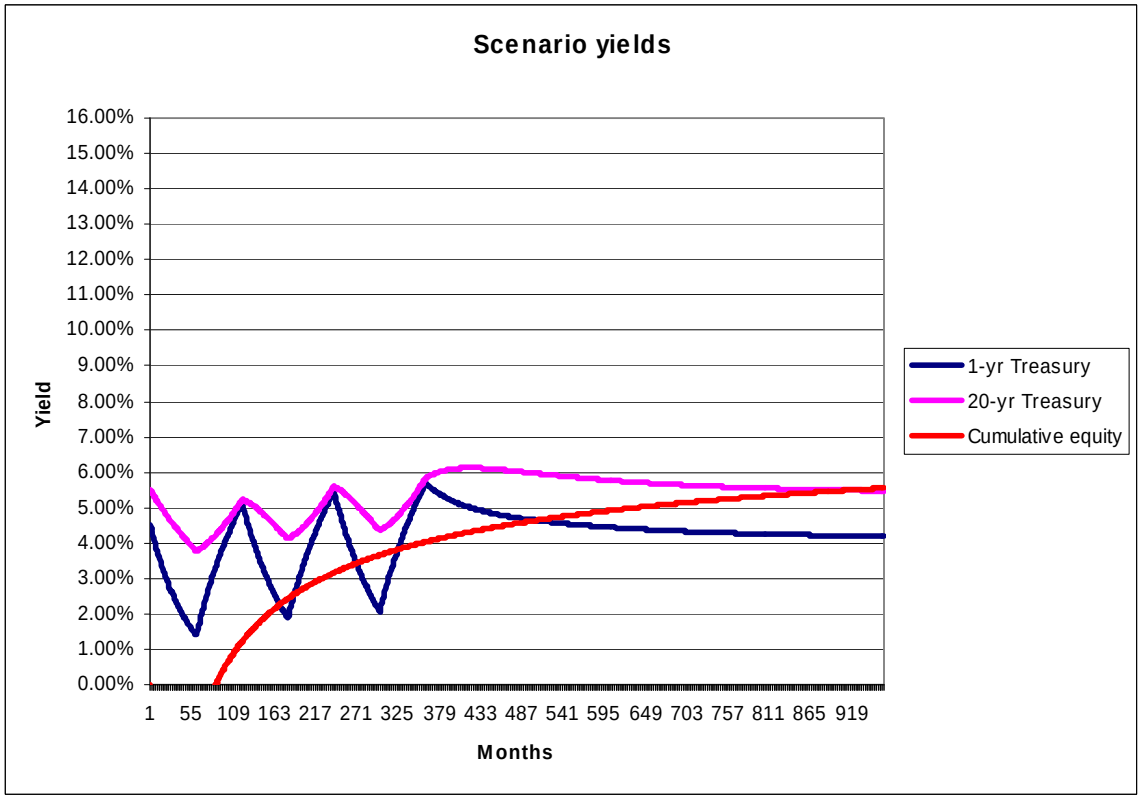
Scenario 6 – Up/down, low equity



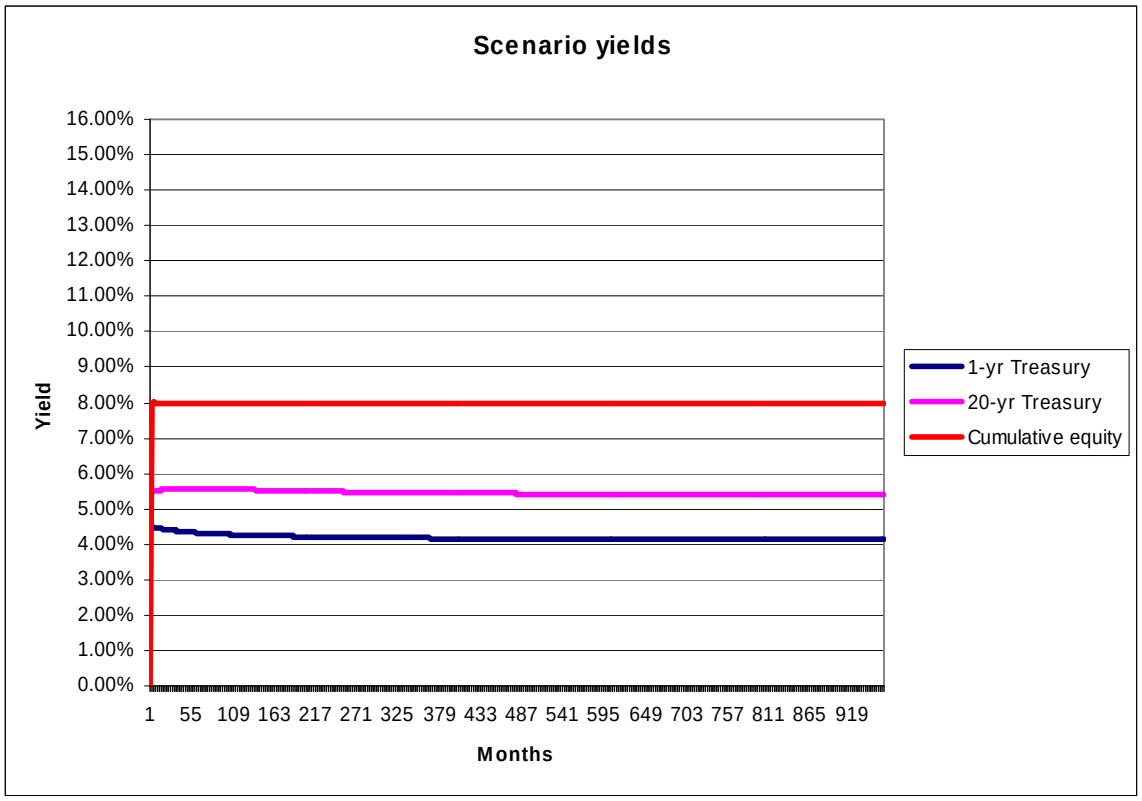
Scenario 7 – Down/up, high equity



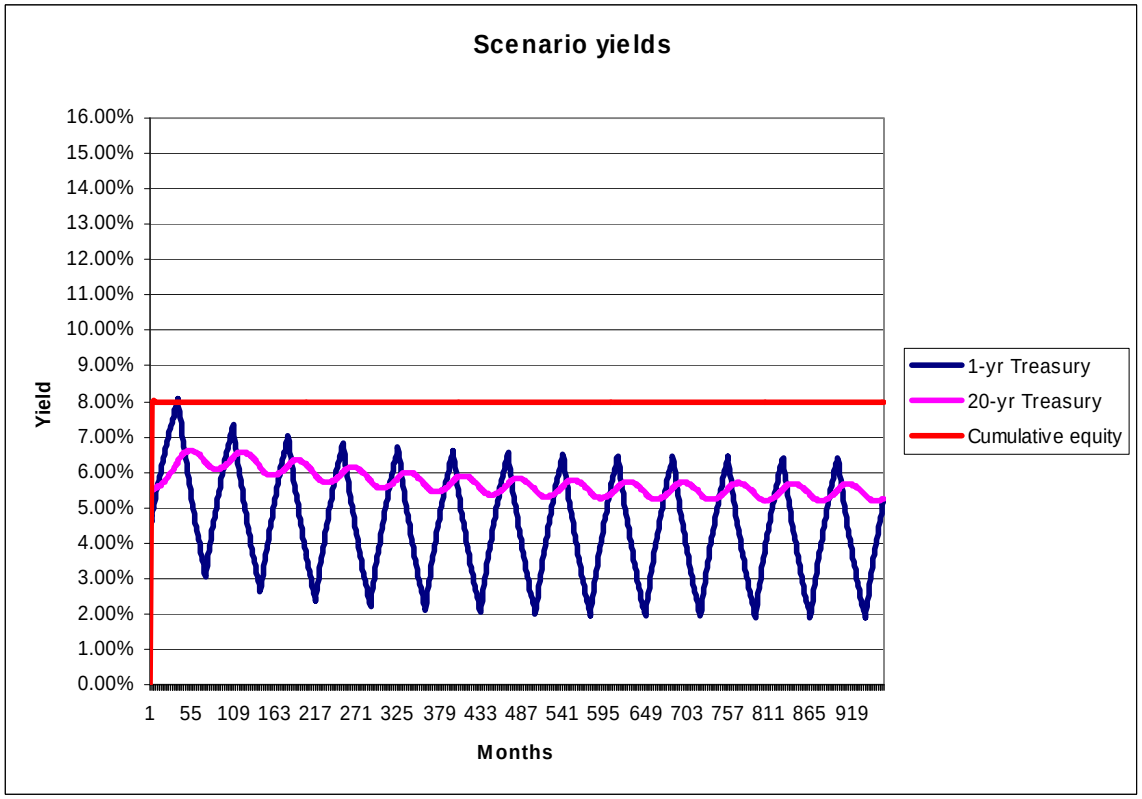
Scenario 8 – Down/up, low equity



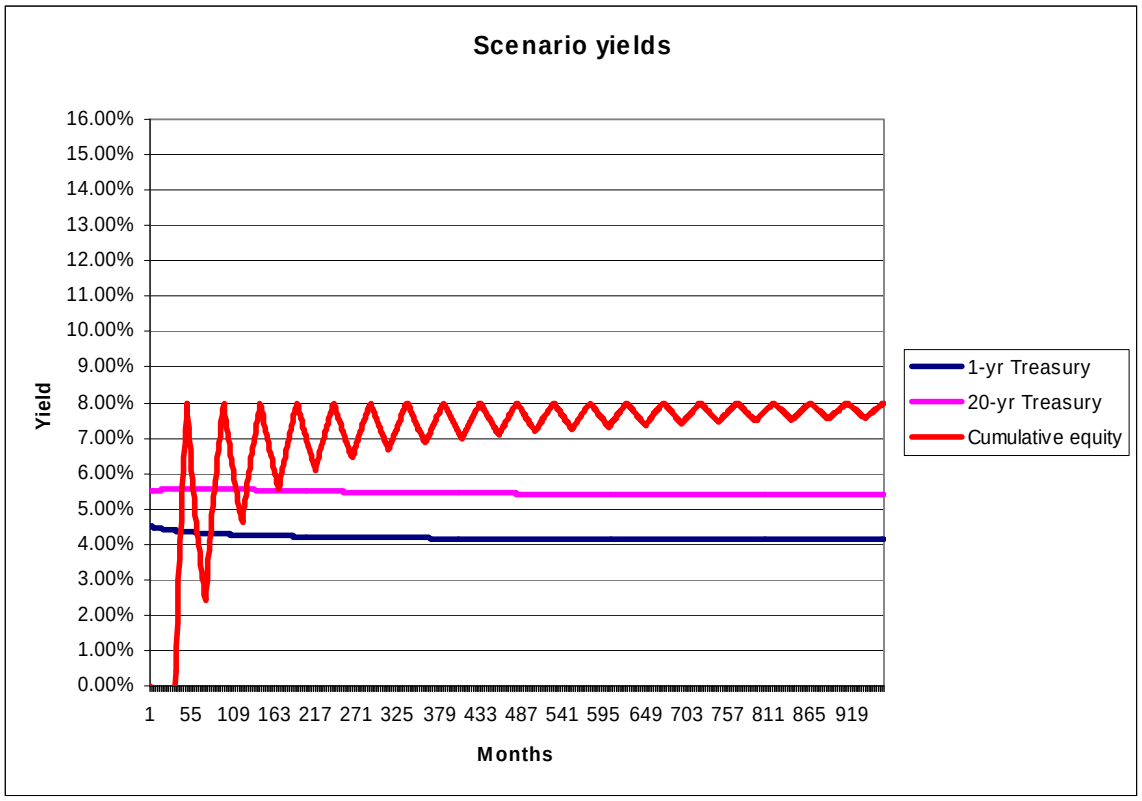
Scenario 9 – Anticipated experience



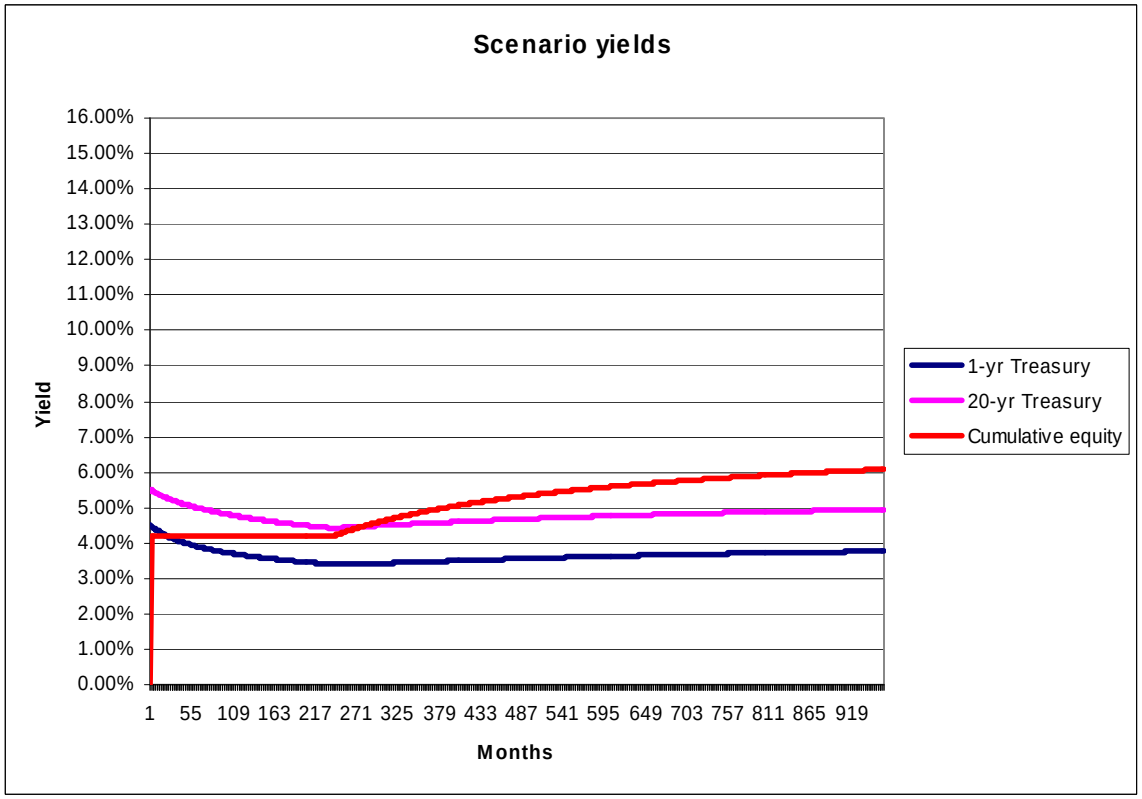
Scenario 10 – Volatile short rates, inverted yields



Scenario 11 – Volatile equity returns



Scenario 12 – Deterministic (test premium patterns)



Results of modeling the test on four products

We tested four products that were featured in product reports sent to LHATF.

- 20-year Level Term
- Secondary-guarantee Universal Life
- Accumulation Universal Life
- Participating Whole Life

Because the results of the test could depend on the level of starting interest rates, we performed the test using two sets of the selected scenarios, based on two different starting yield curves:

- Low: 1-yr Treasury is 3%, 20-yr Treasury is 4%
- High: 1-yr Treasury is 6%, 20-yr Treasury is 7%

These starting points were selected because they are about equidistant above and below the mean reversion point currently being discussed for the updated American Academy of Actuaries' Interest Rate Generator.

In calculating the scenario asset requirements, anticipated experience assumptions were used because the test is likely to be completed before margins on such assumptions are determined. This is because the result of the test may qualify the block for the use of simplified methods, under which either assumptions or margins are specified by regulation, and would not need to be set by the actuary. We wish to avoid having the actuary determine assumption margins that may never need to be used. The term "asset requirement" is used, instead of "reserve," because the assumption margins that would be included when calculating a reserve are not included here.

The results of the testing of these four products appear in the table below. Details of the scenario results and other background on the modeling procedure used are contained in the Appendix.

MTR Ratios by Product

	Starting Yield Curve	
Product	High	Low
Secondary Guar UL	20%	21%
20yr level term	7%	5%
Accumulation UL	3%	6%
Par Whole Life	1%	1%

Three observations stand out when viewing these results:

1. The test clearly flags the SGUL product tested as having the highest degree of tail risk in both interest-rate environments.

2. The results do not differ greatly between the interest-rate environments. For the Accumulation UL product tested, the higher ratio in the low interest-rate environment arises due to the minimum interest crediting guarantee. For this test, it was assumed that the company did not increase COI charges to offset losses due to the interest crediting guarantee (although we are not opining on whether this is an appropriate assumption).

3. The results for the Par Whole Life product tested did not differ between the interest-rate environments. This is because we assumed that mortality gains offset interest losses through the reduction of the dividend (although we are not opining on whether this is an appropriate assumption).

Appendix 1 – Notes on modeling

The four products that were modeled were the same as those on which our earlier reports focused. Detailed product descriptions will not be repeated here because they are available in those earlier reports.

A sample block of business was developed for purposes of testing each product. Each sample block included business issued over many years. The testing was performed on these sample blocks of business, not on individual policies.

A very simple starting investment portfolio and investment strategy was used. The starting portfolio consisted of a ladder of bonds maturing over the first 10 projected years. All bonds in the starting portfolio yielded 6% annually, net of investment expenses and defaults. The size of the starting portfolio was close to the reserve.

Free cash flow was modeled as being invested in 10-year bonds at a net spread (after expense and defaults) of 0.70% over 10-year Treasuries. When cash was not available, borrowing was simulated at an interest rate of 0.80% over the 90-day Treasury.

The effect of alternate investment strategies or hedging strategies was not studied. In particular, we did not model any equity investments, nor any products with benefits linked to equity returns.

Non-guaranteed elements were simulated by setting interest crediting rates equal to the portfolio yield, less the expected pricing spread, subject to contractual minimums. For par whole life, dividends were not allowed to go below zero, but negative contributions from interest were allowed to be offset by positive contributions from mortality.

Scenario reserves were calculated using the GPVAD methodology with a zero working reserve and anticipated experience assumptions.

Two different starting yield curves were modeled on the valuation date, in order to determine whether the result of the test was sensitive to the level of the starting yield curve. The two starting yield curves were as follows:

- Low: 1-yr Treasury is 3%, 20-yr Treasury is 4%
- High: 1-yr Treasury is 6%, 20-yr Treasury is 7%

The results of modeling each scenario for each product using each starting yield curve are shown in Exhibit 1.

Exhibit 1

Detailed results of scenario testing

		Secondary Guar UL		20-year level term	
		Scenario Asset Requirement		Scenario Asset Requirement	
Int Rates	Equity Returns	Set 1 (low)	Set 2 (high)	Set 1 (low)	Set 2 (high)
High (1 yr pop up)	High	217,960,533	173,071,140	34,655,446	29,346,538
High (1 yr pop up)	Low	217,960,533	173,071,140	34,655,446	29,346,538
Low (1 yr pop down)	High	376,807,021	300,221,251	40,559,206	36,420,032
Low (1 yr pop down)	Low	376,807,021	300,221,251	40,559,206	36,420,032
Up/down cycle	High	252,556,853	192,569,731	35,990,976	30,767,557
Up/down cycle	Low	252,556,853	192,569,731	35,990,976	30,767,557
Down/up cycle	High	316,211,769	246,843,710	39,524,564	35,141,302
Down/up cycle	Low	316,211,769	246,843,710	39,524,564	35,141,302
Best estimate	Best estimate	284,974,723	216,123,030	37,932,252	33,091,644
Inversion	Best estimate	264,224,594	202,351,802	36,872,464	31,883,525
Best estimate	Volatile	284,974,723	216,123,030	37,932,252	33,091,644
Deterministic	Deterministic	309,680,047	289,006,545	38,507,348	33,812,012
	Range:	158,846,488	127,150,111	5,903,759	7,073,494
	Denominator:	753,395,836	642,316,212	118,924,527	107,793,711
	Test Ratio:	21%	20%	5%	7%

		Accumulation UL		Participating whole life	
		Scenario Asset Requirement		Scenario Asset Requirement	
Int Rates	Equity Returns	Set 1 (low)	Set 2 (high)	Set 1 (low)	Set 2 (high)
High (1 yr pop up)	High	17,082,177	16,245,628	25,503,234	25,552,683
High (1 yr pop up)	Low	17,082,177	16,245,628	25,503,234	25,552,683
Low (1 yr pop down)	High	19,097,988	16,894,787	25,629,922	25,420,528
Low (1 yr pop down)	Low	19,097,988	16,894,787	25,629,922	25,420,528
Up/down cycle	High	17,265,768	16,917,003	25,532,334	25,656,157
Up/down cycle	Low	17,265,768	16,917,003	25,532,334	25,656,157
Down/up cycle	High	17,007,067	17,115,160	25,483,378	25,469,057
Down/up cycle	Low	17,007,067	17,115,160	25,483,378	25,469,057
Best estimate	Best estimate	17,003,645	16,838,394	25,454,357	25,552,691
Inversion	Best estimate	17,257,450	16,962,428	25,471,335	25,589,825
Best estimate	Volatile	17,003,645	16,838,394	25,454,357	25,552,691
Deterministic	Deterministic	16,874,421	17,179,288	25,434,764	25,532,514
	Range:	2,223,567	933,660	195,158	235,629
	Denominator:	37,106,808	34,453,171	60,443,926	52,656,527
	Test Ratio:	6%	3%	0.3%	0.4%

Appendix 2 – Issues still under consideration

When this report was initially discussed by the Life Reserve Work Group, several issues were raised that merit further consideration by the LHATF. The list below summarizes these issues.

1. The test, as described, includes both upside and downside risk. Perhaps only downside risk should be measured. The range of results in the numerator could be limited to those scenarios whose asset requirement is greater than or equal to that for the anticipated experience scenario.

2. Should the premium pattern test in scenario 12 be included? Perhaps it should be excluded and the focus should only be placed on the variability resulting from the economic scenario.

3. To what grouping should the test be applied? If applied at the block of business level, several plans of insurance might be included in the block of business, and some of those plans might, if measured separately, exhibit material tail risk.

4. What value of the test ratio indicates material tail risk? Should there be a single value or should it vary somehow by policy duration or by the level of the starting yield curve? One idea here is that the range of scenario asset requirements might be smaller than the margin that would be included in the reported reserve. Note that the scenario asset requirements used in the test are smaller than scenario reserves because no assumption margins are used in their calculation. If the dollar effect of those assumption margins could be estimated, one could compare it with the range of scenario asset requirements to determine whether reserve margins would cover the range. The result of that comparison might play a role in the test.

5. How often should the test be carried out to continue to qualify for the stochastic testing exclusion? Annually, as with cash flow testing, or less often?

6. Is the test robust if the starting yield curve is much higher or lower than was tested for this report?

7. Should the test be performed using anticipated experience assumptions, or should prudent estimate assumptions be required?