Insurance Modeling and Stochastic Cash Flow Scenario Testing:

Effective Sampling Algorithms to Reduce Number of Runs

By Yvonne C. Chueh

Abstract

With the proposed RBC requirements for variable annuities with guarantees (so called C3 Phase II) requiring stochastic testing, there should be a lot of interest in the approaches to reducing the number of runs. This article provides strategies to reduce number of runs in stochastic insurance modeling as well as cash flow testing when the interest scenarios are stochastic. Three interest rate sampling algorithms and a computer software program SALMS that performs sampling are introduced.

Introduction

One of the greatest challenges of stochastic insurance modeling for large insurance businesses is the run-time. Using a complete stochastic asset/liability model to analyze a large block of business is often too time consuming to be practical. Efficient stochastic modeling can be achieved by applying effective interest rate sampling algorithms that are presented in this article. The algorithms were tested on a simplified asset/liability model as well as a commercial asset/liability model using assets and liabilities of Aetna Insurance Company of America (AICA) in 1999. Another methodology using the New York 7 scenarios is proposed and could become an enhancement to the Model Regulation on cash flow testing, thus requiring all companies to do stochastic cash flow testing in a uniform, non-onerous manner. A beta version PC software (SALMS) designed to perform interest rate sampling algorithms in this article is available to modelers upon request.

First Algorithm

The algorithms are independent of interest rate generator so any multi-factor (or the less popular one-factor) generator can be considered. However, to keep the sampling algorithms simple yet effective, only short-term (one-year) interest rates are needed to perform sampling.

We first define a distance measure that is similar to the Euclidean distance and then demonstrate the steps to select representative (pivot) scenarios. A probability is then assigned to each representative scenario.

<u>Definition of Distance</u>

The distance between a given 30-year interest rate path and a pivot interest rate path is defined as

$$D_1 = \sqrt{\sum_{t=1}^{30} \left(i_t - i_t^P\right)^2 \cdot v^t}\ .$$

where i_t , t=1,2,….,30 is an interest rate path consisting of short-term (one-year) interest rates for 30 years .

i_t^P , t=1,2, … ,30 is a pivot (i.e. representative)interest rate path consisting of short-term (one-year)interest rates for 30 years.

v is a weight factor with a value between 0 and 1 that distinguishes the relative importance of interest rate of each projecting year. The author felt that the value is around the one-year discount rate and its value used in this article is 1/1.06.

To illustrate how to perform the sampling algorithm, we select 100[1] representative interest rate paths out of a large number, say N, of interest rate paths by taking the following steps:

Step1. Choose an arbitrary interest rate path out of the N simulated ones and call it Pivot #1.
Step2. Calculate the distances from Pivot#1 to the remaining N-1 interest rate paths.

Step3. Name the interest rate path with the largest distance to Pivot #1 as Pivot #2. Randomly decide among ties.

Step4. Calculate the distances of the N-2 non-pivot interest rate paths to Pivot #1 and Pivot #2.
Step5. Assign each of the N-2 interest rate paths to the closest of Pivot #1 or Pivot #2, thus forming 2 disjoint sets of interest rate paths. Flip a coin if the distances are equal. Each of the N-2 interest rate paths now has a unique distance to its pivot scenario.

Step6. Rank these N-2 distances in descending order. The interest rate path producing the top distance is called Pivot #3. (Break ties randomly)

Step7. Follow the above procedure to select the additional 96 pivot scenarios, Pivot #4, Pivot #5,…, Pivot #100.

Step8. If the number of interest rate paths associated to a pivot scenario is N_k, then assign a probability of N_k /1000 to this pivot scenario.

Second Algorithm

[1] In practice, running 100 stochastic scenarios is manageable and sometimes the maximum. It's not unusual to take more than one hour to project cash flows for one scenario using the most powerful computer available to the industry.

A distance measure between an interest rate path and a Pivot scenario is given as below. Note that this distance measure does not require a weighting factor. Instead, the interest rate for a given year is discounted by the interest rates in the earlier years.

$$D_2 = \sqrt{\sum_{t=1}^{30} \left(\prod_{k=1}^{t} \frac{1}{1+i_k} - \prod_{k=1}^{t} \frac{1}{1+i_k^P} \right)^2}$$

where i_t , t=1,2,….,30 is an interest rate path consisting of short-term (one-year) interest rates for 30 years .

i_t^P , t=1,2,….,30 is a pivot (i.e. representative)interest rate path consisting of short-term (one-year)interest rates for 30 years.

Repeat the sampling algorithm in the previous section to select a sample of representative interest rate paths. Assign a probability to each of the representative interest rate paths as before.

Third Algorithm

In this section, a much quicker but still effective way to sample interest scenarios is introduced by first giving a new distance definition and then taking a sample in a uniform way. This method assigns an equal probability to scenarios.

The significance of an interest rate path is defined as

$$S = \sqrt{\sum_{t=1}^{30} \left(\prod_{k=1}^{t} \frac{1}{1+i_k} \right)^2} .$$

where i_k , k = 1,2,….,30 is an interest rate path consisting of short-term (one-year) interest rates for 30 years .

Select 100 representative interest rate paths out of a large number, say N, of interest rate paths by taking the following steps:

Step1. Calculate the significance measure of all the N interest rate paths based on the above formula.

Step2. Sort the interest rate paths by significance measure in ascending order. Denote by $I_{[1]}, I_{[2]}, \ldots, I_{[N]}$.

Step3. Choose the 100 representative interest rate paths as follows: For k = 1, 2,…,100, the k-th interest rate path is $I_{[10k-5]}$.

The probabilities of the selected representative interest rate paths are set equal (1/100 in this case).

The sampling steps proposed above can be generalized to select n paths instead of just 100 paths.

These sampling methods (algorithms) discussed above have a broad use in any stochastic modeling that involves stochastic interest rates. Another use of these algorithms could be in stochastic models that have stochastic scenarios in the form of time series (e.g. exchange rate, inflation rate, equity return, . . . , etc.) Instead of running a large number of stochastic scenarios, actuaries can use the above sampling idea to reduce the number of runs while preserving the behavior of tail distributions.

Another Strategy: Using The New York 7 Scenarios

In contrast to the previously discussed sampling algorithms, this approach is directly linked to the long existing approach testing seven extreme interest scenarios.

The author believes that this method shows the greatest promise for a standardized stochastic cash flow testing procedure to be required for all companies.

The New York 7 Asset Approach is implemented as follows:

1. First, run a commercial asset modeling software system to project asset cash flows of a company's current assets for only the "required New York 7" scenarios.
2. Next, generate a universe of stochastic interest rate scenarios by an appropriate stochastic interest rate model and map each scenario to one of the New York 7 scenarios, using *the relative present value distance formula.*
3. Now for each stochastic interest rate scenario, use the initial asset cash flows determined by its associated New York 7 interest rate scenario. However, we continue to use the specific interest rate scenario for the reinvestment of liability cash flows, principal repayments, coupons, and any "borrowing". This would be done in the asset/liability model (TAS or PTS) where the run time problem is not an issue.

Since it is the asset projection process that takes most of the run time, the use of only 7 asset projection scenarios involves the same run time as is currently incurred in deterministic cash flow testing. This approach preserves the extreme asset projections since the New York 7 scenarios were originally designed to test extreme conditions. This approach is not only time efficient but capable of generating a distribution with tails comparable with a full stochastic run (as shown in Appendix: Figures 5 and 1 and Table 5). By assigning each scenario to one of the New York 7 we don't overemphasize the extreme interest rate scenarios since each of the New York 7 is automatically given a probability proportional to the number of stochastic scenarios assigned to it. With the approximated asset cash flows assigned to every stochastic scenario, each scenario can then properly interact with its associated liabilities and allow the asset/liability model to project financial outcomes in the future years.

In order to demonstrate the effectiveness of the *New York 7 asset approach*, we first perform a full stochastic run of Economic Value (EV) in our ASEM model by inputting the baseline assumptions to the ASEM model and then running an asset/liability projection based on 1500 stochastic interest rate scenarios (Appendix: Figure 1). Then we repeat the process, limiting asset cash flow projections to only the required New York 7 scenarios. As described above, we project liability cash flows still based on the full 1500 stochastic scenarios. It can be observed that the probability distribution shown in Figure 5 closely resembles Figure 1.

In order to take a closer look at the above two distributions, we compare the percentiles of the probability distributions of Economic Value.

Table A. Comparison of Percentiles of EV Distributions for Full Run and Efficient New York 7 Asset Approach

Percentile	**1st**	**5th**	**10th**	**15th**	**20th**	**25th**	**30th**	**50th**
1500/1500 (Full Run)	23442	24897	26013	26558	26886	27060	27182	27552
1500/NY7	23460	25171	26102	26679	26984	27130	27237	27567
% Difference	0.08%	1.10%	0.34%	0.46%	0.36%	0.26%	0.21%	0.05%

Table A. (Continued)

Percentile	**70th**	**75th**	**80th**	**85th**	**90th**	**95th**	**99th**
1500/1500 (Full Run)	27979	28256	28516	29088	29617	30552	32087
1500/NY7	27999	28258	28585	29071	29666	30558	32054
% Difference	0.07%	0.01%	0.24%	-0.06%	0.16%	0.02%	-0.10%

Table A shows that the percentage difference between any percentile in the table does not exceed 1.1%. This data strongly supports that the two distributions derived from two different stochastic runs, one full run and the other by using our efficient New York 7 asset modeling approach, have very similar patterns at the tails.

Table B. K-S Test Results, New York 7 Asset Approach

K-S Test	NY 7 Asset	D_n	y	Pr(Y>y)
Overall	N_1=1500, N_2=1500	0.03000	0.82158	0.50946
Two tails	N_1=409, N_2=412	0.01267	0.18147	1.00000

Table B compares the K-S test results for both overall and tail distributions. The results strongly support the claim that the sampled distribution using New York 7 Asset Approach agrees with that of the full run.

Validating of Sampling Algorithms Using Commercial Asset/Liability Models

Independent validation and testing of the proposed sampling algorithms have been conducted against the assets and liabilities of Aetna Insurance Company of America (AICA) using commercial asset/liability models, BondEdge and TAS. Three algorithms (1. Uniform method–the *third* sampling algorithm, 2. Pivot method–the *second* sampling algorithm, and 3. NY7– New York 7 asset approach) were tested against 500 interest rate scenarios on the AICA accumulation business. One hundred interest rate scenarios were sampled based on respective sampling algorithms. The test statistic measured is the present value of the ending surplus. The AICA portfolio contains a relatively diverse set of assets and hence is proper to be used in testing. Here is the portfolio composition:

- Treasuries: 2.4%
- Passthroughs: 5.6%
- CMOs: 7.6%
- ABS: 7.8%
- Industrials: 31.6%
- Electric & Gas: 3.2%
- Telephone: 2.6%
- Finance: 28.3%
- International: 4.4%
- Cash: 0.8%
- Other: 5.7%

A summary of asset data is listed as follows:

- Number of Assets: 102
- Average Quality: A1
- Market Weighted Average Coupon: 6.556%
- Market Value: $127,872,000
- Modified Duration: 5.79
- Effective Duration: 5.33
- Convexity: 0.19

Table C summarizes the test results. The results based on AICA assets and liabilities are very promising. The large P-values for all the three sampling algorithms support the goodness of the fit to the full stochastic distribution of the present value of the ending surplus.

Table C. Testing Results on the AICA Portfolio

Sampling Algorithm	**Uniform**	**Pivot**	**NY 7**

Results	**K-S Test: Total P-Value**	>98%	>85%	>75%
	K-S Test: Tails P-Value	>98%	>95%	>88%

K-S Test: Total = Kolmogorov-Smirnov test on the full sample.
K-S Test: Tails = K-S test on just the bottom 20% and top 20% sample values.

Conclusion

The sampling algorithms developed and analyzed in this article are applicable to all business lines of the insurance industry. This article provides a practical and statistically sound methodology for even small companies to perform stochastic analysis of their business. The NY 7 methodology in particular is something companies could easily adopt.

Software Program SALMS

The SALMS (Stochastic Asset Liability Model Sampling) software is based on the above algorithms to assist sampling stochastic interest rates. SALMS features three algorithms to sample with: Modified Euclidean Distance Method, Relative Present Value Distance Method and the Significance Method.

The SALMS has the flexibility that users may create a universe of interest rates using the SALMS software or load a pre-existing universe of short-term yearly rates. The stochastic interest rate model used in the SALMS is the Stochastic Variance Model developed by the SOA c-3 risk task force.

SALMS was developed at Central Washington University in Ellensburg, Washington. The system requirements are as follows:

1. CD-Rom drive
2. 256 MB RAM
3. At least 100 Megabytes (MB) of disk space. Universe files can reach up to 55 MB.
4. Windows XP, 2000, ME, 98, 98SE
5. Windows Scripting Host 5.6.

To create a new universe and sample from it go to the main menu and select Create New Universe or from the File menu select New Universe.

Figure 1 - SALMS main menu

Input the values for universe size, first year interest rate, 20-year interest rate, minimum and maximum interest rate.

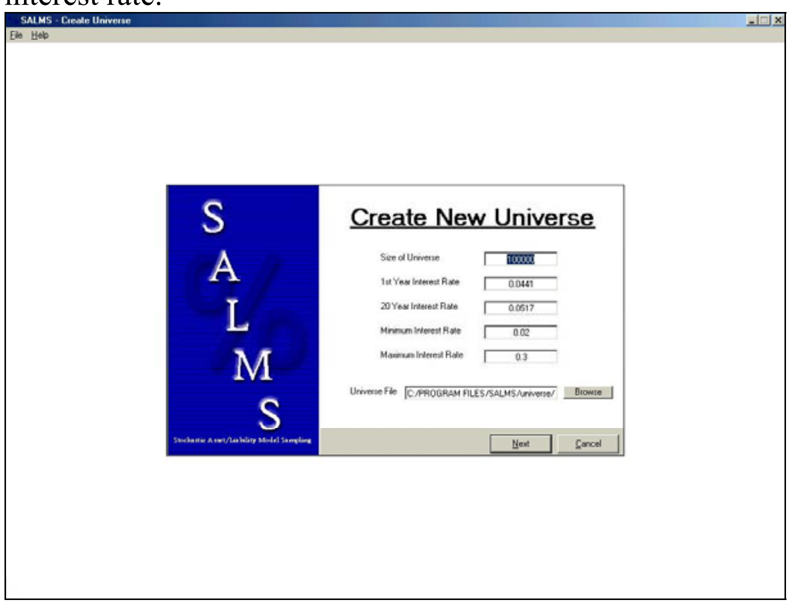

Figure 2 - SALMS create new universe page

On this same page, give the location and name of the file you want to store the universe to. This can be done by typing it into the field or click browse and entering a file name. Once these have been entered, click the next button.

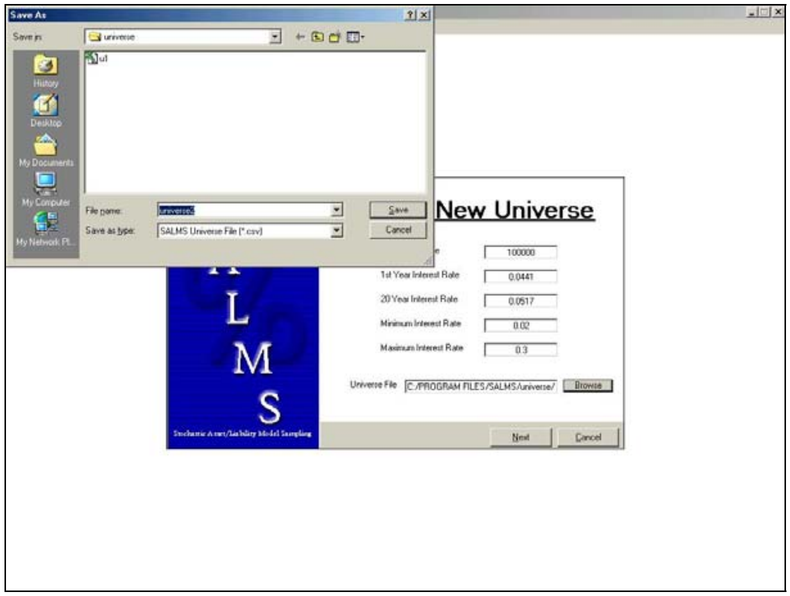

Figure 3 - SALMS create new universe page with save dialog box

On the Sample Size page, enter in the location and file name for the results. You can do this by typing it in or clicking browse. Then enter the sample size. Once these have been entered, click the next button.

Figure 4 - SALMS sample size page

This page allows you to choose the algorithm you would like to sample with.

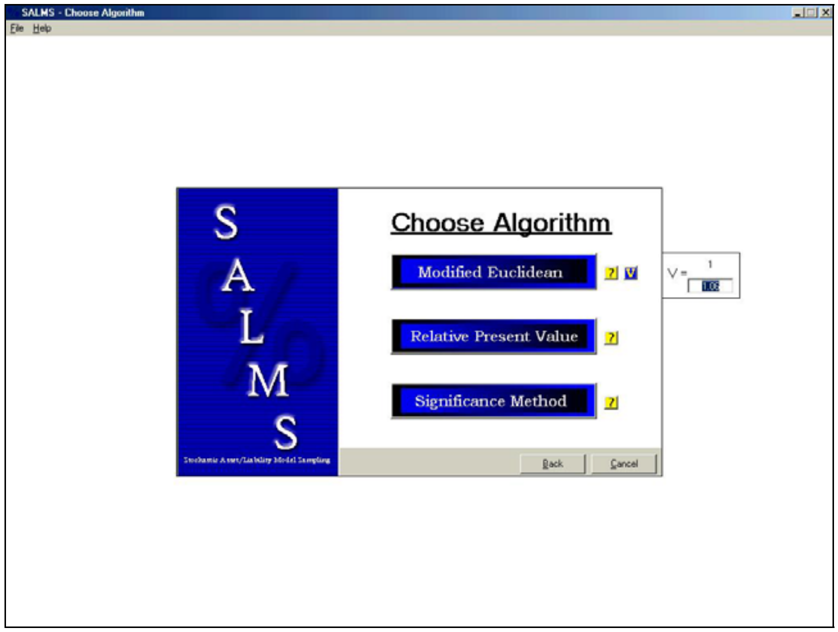

Figure 5 - SALMS choose algorithm screen

SALMS will notify you when it is creating a new universe and sampling from it. The sampling is done when the results screen appears.

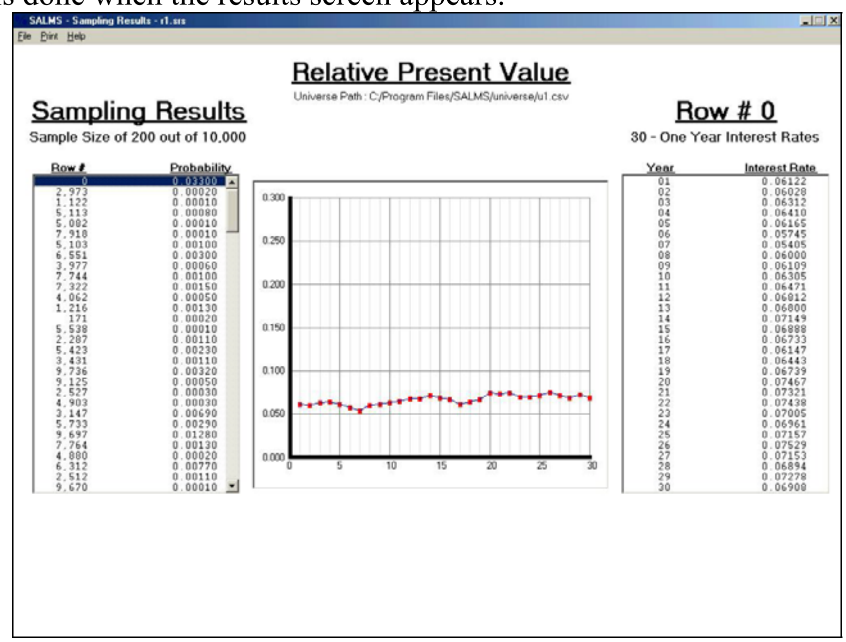

Figure 6 - SALMS results page

Appendix:

Validation and Testing for the Sampling Algorithms

This section uses the ASEM model (a simplified asset/liability model) as the basis to examine and validate our sampling methods. The interest rate generator used is called the Stochastic Variance Model. The results provide strong evidence to support the effectiveness and efficiency of the sampling algorithms.

The following graphs (Figures 1, 2, 3, and 4) show empirical probability distributions derived from four stochastic runs. One is a 1500–scenario full run and the other three are 50–scenario sample runs in the ASEM model calculating Economic Value (EV) using baseline assumptions. Figure 1 illustrates the EV distribution based on 1500 stochastic scenarios. Figures 2 to 4 illustrate the distribution of Economic Values (EV) based on the three sampling algorithms. Table 1 compares percentiles of the EV distributions, under the *first*, *second*, and *third* sampling algorithms respectively, against the full scenario run. Table 2 shows the Kolmogorov-Smirnov (K-S) Test results.

Figure 2 clearly shows that the sample points determined by the 50 selected interest rate scenarios using the *first sampling algorithm* cluster in the center of the EV distribution and is a weaker fit at the tails. The main reason for this is in the arbitrariness of the weighting factor V that directly affects the scenarios being selected. This factor V is related to the future lifetime of the contracts being modeled as well as the duration of the underlying assets and liabilities. It requires some experiments to determine the value for v. In testing the algorithms, we use $v = 1/1.06$.

The second distance formula ("*Relative Present Value Distance Formula*") avoids this subjective selection of a discount factor and demonstrates a significant improvement on the sample EV distribution as reflected in Figure 3

Figure 3 shows a more accurate EV distribution. The sample points are evenly spread through the interval (22,000, 32,000) and the distribution makes a smooth curve that is commonly seen on cumulative distribution functions. When compared with Figure 1, the graph based on 1500 runs, we observe an excellent fit on the overall distribution and tails.

Since the *Relative Present Value Distance Formula* makes such a large improvement, we wanted to see if the distance formula and the sampling procedures could be simplified while maintaining the improved accuracy. This motivated the *third sampling algorithm*: *Significance Method*. Figure 4 shows the cumulative distribution function of economic value (EV) based on the 50 selected representative scenarios using the *third sampling algorithm*.

We can observe that the 50 sample points from the *third sampling algorithm* form a smooth, solid distribution curve. However, some extreme sample points are missing so that the observed values of EV fall in a shorter interval (23,000, 33,000). This is anticipated since we selected 50 representatives by dividing the 1000 scenarios into 50 groups and then taking the mid-point of each group as a representative scenario. When the sampled percentiles at the left tail are less than those for the full run, over sampling at the left tail occurs. However, when the sampled percentiles at the left tail are greater than those for the full run, under sampling at the left tail occurs. Similarly, when the sampled percentiles at the right tail are greater than those for the full run, over sampling at the right tail occurs. However, when the sampled percentiles at the right tail are less than those for the full run, under sampling at the right tail occurs. When comparing the above three figures, 2 to 4, with figure 1, it is important to know whether or not the three sampling algorithms over sample the extreme scenarios. Perhaps the *third sampling algorithm* under-samples the extreme scenarios and thus misrepresents the tail-end distribution.

To answer the question of "over sampling or under sampling?", we turn to quantitative evidence: percentage difference of percentiles of two economic value distributions, one based on representative scenarios, the other based on a 1500 scenario full stochastic run. The numerical results based on each sampling algorithm are listed in Tables 1. The first two columns after the table heading list the percentiles of EV from two distributions: one from 1500 stochastic runs and the other from 50 representative runs using the *first* sampling algorithm.

Table 1
Comparison of Percentiles of Two EV Distributions:
1500 Runs versus 50 Representative Runs

EV Percentile	**1500 (Full Run)**	**50 Rep. First Algorithm**	**50 Rep. Second Algorithm**	**50 Rep. Third Algorithm**	**% Difference First Algorithm**	**% Difference Second Algorithm**	**% Difference Third Algorithm**
1st	23442	19404	22647	23861	-17%	-3.39%	1.79%
5th	24897	19404	24663	25720	-22%	-0.94%	3.31%
10th	26013	23587	25743	25889	-25%	-1.04%	-0.48%
15th	26558	26705	26503	26725	0.55%	-0.21%	0.63%
20th	26886	27183	26567	27026	1.10%	-1.19%	0.52%
25th	27060	27214	26706	27140	0.57%	-1.31%	0.29%
30th	27182	27244	27243	27280	0.23%	0.22%	0.36%
50th	27552	27365	27452	27612	-0.68%	-0.36%	0.22%
70th	27979	28065	27969	28169	0.31%	-0.04%	0.68%
75th	28256	28280	28150	28425	0.08%	-0.37%	0.60%
80th	28516	28321	28813	28522	-0.68%	1.04%	0.02%
85th	29088	29147	29165	29061	0.20%	0.26%	-0.09%
90th	29617	29739	29644	29545	0.41%	0.09%	-0.24%
95th	30552	30548	30495	30507	-0.01%	-0.18%	-0.15%
99th	32087	32241	32136	31817	0.48%	0.15%	-0.84%
Average of Absolute Value					4.62%	0.72%	0.68%

Table 1 and Figure 2 indicate that the low percentiles (1^{st}, 5^{th}, and 10^{th} percentiles) based on the *first* sampling algorithm are significantly over sampled. The high percentiles, the 90^{th} and 99^{th} percentile, is also slightly over sampled. As a result, the *first* sampling overstates the low percentiles by 17% to 25% and overstates the high percentile (the 99^{th} percentile only) by 0.48%.

Table 1 indicates that the *second* sampling method slightly over-samples both tails of the distribution. The 1st percentile is overstated by 3.39% of "true" value that is assumed to be the 1st percentile based on the 1500 stochastic runs. Considering the low percentage of all the differences (no more than 3.39%) and the small sample size selected (50 out of 1500) the overall distribution is well sampled by the *second sampling algorithm*.

Table 1 indicates slight under-sampling at both tails for the *third* sampling algorithm. The largest percentage difference occurs at the 5th percentile, at 3.31%. The percentage differences in general are lower and comparable with the *second sampling algorithm.* Thus the *third* sampling algorithm also well samples the distribution.

For each sampling algorithm, both the resulting overall EV distribution and the two-tail EV distribution were examined by the modified Kolmogorov-Smirnov (K-S) test. The modified K-S test involved comparing the cumulative probabilities of EV results from the sampled N_1 (50 and 100, respectively) scenarios with those from the universe of N_2 (here 1500) scenarios. When comparing two empirical distributions with a K-S test, $\sqrt{N}D_N$ is replaced with $\sqrt{\frac{N_1 N_2}{N_1 + N_2}} D_{N_1,N_2}$. One can use the following limiting probability to show the probability of matching the two empirical distributions when N_1 and N_2 >50:

$$\lim_{N\to\infty} \Pr(\sqrt{N}D_N > y) = 2\sum_{j=1}^{\infty}(-1)^{j+1}\exp(-2j^2y^2)$$

For the overall distribution, the K-S test indicated the probability that the overall distribution from the full run equals the distribution from the *first* sampling algorithm was low. The probability was only 0.02314 for the 50-scenario sample run and 0.09265 for the 100-scenario sample run. The *second* and *third* sampling algorithms had higher probability 0.96581 of matching.

For the distribution at two tails (i.e. test of fit for the bottom 20% and top 20% of the distribution), the probability of matching for the *first* sampling algorithm was increased dramatically to 0.99977. Thus we found no significant evidence against the hypothesis of matching, although our result was suspect since we had only N_1=37. All the three algorithms displayed a very high probability of matching the two tails of the distribution, although the results were suspect due to small N_1. It is interesting to see that although the *third* sampling algorithm is a lot simpler and quicker than the first two algorithms, the results for both overall and two-tail distributions seem promising.

Table 2 K-S Test Results, Three Sampling Algorithms

K-S Test	First Sampling Algorithm	D_n	y	Pr(Y>y)
Overall	N_1=50, N_2=1500	0.21467	1.49324	0.02314
Overall	N_1=100, N_2=1500	0.12800	1.23935	0.09265
Two tails	N_1=37, N_2=412	0.05933	0.34532	0.99977

Table 2 (Continued)

K-S Test	Second Sampling Algorithm	D_n	y	Pr(Y>y)
Overall	N_1=50, N_2=1500	0.11000	0.76517	0.60170

Overall	N_1=100, N_2=1500	0.08400	0.81333	0.52262
Two tails	N_1=22, N_2=412	0.04333	0.19803	1.00000

Table 2 (Continued)

K-S Test	Third Sampling Algorithm	D_n	y	Pr(Y>y)
Overall	N_1=50, N_2=1500	0.05800	0.40345	0.99683
Overall	N_1=100, N_2=1500	0.05133	0.49703	0.96581
Two tails	N_1=14, N_2=412	0.02933	0.10765	1.00000

Two- tail refers to the bottom 20% and top 20% of the distribution.

Another test was conducted for the 100 scenarios chosen from the randomly generated 10,000 interest rate scenarios using the simple random sampling as well as the three sampling algorithms proposed in this article. We compared the resulting EV sampling distributions against that of the 10,000-scenario full run. The descriptive statistics and modified K-S test results are shown in Table 3.

Table 3 Descriptive Statistics of EV and K-S Test Results:
10,000-Scenario Full Run vs 100-Scenario Sample Run

	EV Statistics					**EV Percentage Difference**			
	10,000 Full Run	**Simple Random**	**First Algor.**	**Second Algor.**	**Third Algor.**	**Simple Random**	**First Algor.**	**Second Algor.**	**Third Algor.**
Min.	19691	21296	19691	19691	22228	8.15%	0.00%	0.00%	12.88%
1st Percentile	22624	22437	21807	22024	22697	-0.83%	-3.61%	-2.65%	0.32%
2.5th Percentile	23484	23206	22368	23612	23461	-1.19%	-4.75%	0.55%	-0.10%
5th Percentile	24207	24372	23950	24261	23716	0.68%	-1.06%	0.22%	-2.03%
10th Percentile	25081	25206	24463	24944	24133	0.50%	-2.46%	-0.55%	-3.78%
Q1	26485	26619	26392	26502	26069	0.50%	-0.35%	0.06%	-1.57%
Q2	27259	27188	27149	27164	27201	-0.26%	-0.40%	-0.35%	-0.21%
Q3	27626	27488	27624	27556	27579	-0.50%	-0.01%	-0.25%	-0.17%
90th Percentile	28258	27735	28421	28718	27906	-1.85%	0.58%	1.63%	-1.25%
95th Percentile	29186	28162	29225	28970	28982	-3.51%	0.13%	-0.74%	-0.70%
97.5th Percentile	29961	29383	30337	29567	29121	-1.93%	1.25%	-1.32%	-2.80%
99th Percentile	30827	30140	30861	30653	30061	-2.23%	0.11%	-0.56%	-2.49%
Max.	35214	30898	35214	35214	31049	-12.3%	0.00%	0.00%	-11.8%
Average of Absolute	26916	26472	26731	26837	26477	2.54%	1.11%	0.66%	2.94%

Value									

K-S Test (Overall)				
	Simple Random	**First Algorithm**	**Second Algorithm**	**Third Algorithm**
(N_1, N_2)	(100, 10000)	(100, 10000)	(100, 10000)	(100, 10000)
Probability F(100)=F(10000)	0.08327	0.53118	0.11552	0.5201
K-S Test (Two-Tail)				
	Simple Random	**First Algorithm**	**Second Algorithm**	**Third Algorithm**
N_1, N_2	(32,4001)	(71, 4001)	(66, 4001)	(36,4001)
Probability F(100)=F(10000)	0.94787	0.99857	0.99442	0.97044

Two- tail refers to the bottom 20% and top 20% of the distribution.

Based on the results in Table 3, it appears that the simple random sampling also gives good results (comparable with *third* algorithm) in terms of the sampled percentiles not too far way from those of the full run. The *first* and *second* sampling algorithms have superior results in percentiles. This does not surprise us since the algorithms attempt to select the low and high percentiles. We then expect that when the model outputs form a heavy-tail distribution, the *first* and *second* algorithms will do a good job. As the *first* and *second* sampling algorithms allow us to select the most extreme scenarios, they are likely to select the maximum and minimum economic values equal to those of the full run. In this test, they have. We can see the percentage differences (error) in the maximum and minimum economic values between the sample run and the full run were zero. All the three sampling algorithms passed the modified K-S test at 10% significance level. The s*econd* sampling algorithm had the lowest average (absolute) difference (error) in percentiles. The *first* and *third* sampling algorithms, however, had the higher probability of matching the full run based on the K-S test. Overall, the *second* algorithm performs best in the tails based on the test.

For the *first* and *second* sampling algorithms, the resulting representative (pivot) paths depend on the paths initially selected. In fact, the first pivot path determines the subsequent pivot selections. To assess the sensitivity of the sampling effectiveness to the initial selection of path, five samples using different initial (first) pivot scenario were examined under *first* and *second* sampling algorithm (*third* algorithm does not use pivot), respectively. Table 4 shows the percentage difference in EV statistics between full run and sample run under the *first* and *second* sampling algorithm respectively. Table 5 shows the modified K-S test results for each initial pivot scenario under the *first* and *second* sampling algorithm respectively.

Table 4 Sensitivity Test of Initial Selection of Pivot Scenario – EV Statistics

First Algorithm					
Percentage Difference Between Full Run and Sample Run					
Initial Pivot Selection	**1**	**2**	**3**	**4**	**5**

EV	**Min**	0.00%	0.00%	0.00%	0.00%	0.00%
	1st Percentile	-12.52%	-12.72%	-12.72%	-11.62%	-11.62%
	2.5th Percentile	-13.49%	-11.17%	-8.21%	-13.62%	-12.48%
	5th Percentile	-9.68%	-8.14%	-8.08%	-10.52%	-7.48%
	10th Percentile	-8.09%	-6.80%	-6.95%	-6.77%	-6.40%
	Q1	-1.97%	-2.16%	-4.30%	-2.72%	-2.79%
	Q2	1.08%	0.65%	0.94%	0.69%	0.56%
	Q3	8.36%	6.85%	6.47%	6.86%	6.53%
	90th Percentile	12.76%	13.11%	12.44%	12.76%	12.76%
	95th Percentile	12.40%	12.40%	12.40%	12.40%	12.40%
	97.5th Percentile	11.02%	11.02%	11.02%	11.02%	11.02%
	99th Percentile	9.63%	9.63%	9.63%	9.63%	9.63%
	Max	0.00%	0.00%	0.00%	0.00%	0.00%
Average of Absolute Value		7.77%	7.28%	7.17%	7.59%	7.21%
Second Algorithm						
EV Percentage Difference Between Full Run and Sample Run						
Initial Pivot Selection		**1**	**2**	**3**	**4**	**5**
EV	**Min**	0.00%	0.00%	0.00%	0.00%	0.00%
	1st Percentile	-10.33%	-10.59%	-5.94%	-10.59%	-10.12%
	2.5th Percentile	-10.93%	-8.77%	-6.95%	-10.23%	-7.25%
	5th Percentile	-10.54%	-7.07%	-6.79%	-7.99%	-6.78%
	10th Percentile	-6.83%	-5.73%	-6.77%	-6.79%	-6.79%
	Q1	-3.83%	-2.83%	-4.00%	-3.47%	-3.47%
	Q2	-0.11%	0.06%	-0.13%	-0.51%	-0.51%
	Q3	4.32%	3.51%	3.42%	3.07%	3.09%
	90th Percentile	8.23%	8.27%	8.09%	6.27%	7.64%
	95th Percentile	8.04%	8.87%	8.87%	7.76%	6.56%
	97.5th Percentile	7.69%	8.03%	8.03%	7.26%	6.89%
	99th Percentile	7.13%	7.13%	7.13%	7.13%	7.13%
	Max	0.00%	0.00%	0.00%	0.00%	0.00%
Average of Absolute Value		6.00%	5.45%	5.09%	5.47%	5.10%

Table 5 Sensitivity Test of Initial Selection of Pivot Scenario – Modified K-S Test

K-S Test (Overall) N_1**=100,** N_2**=10000**			
Probability F (100) =F (10000)			
		First Algorithm	**Second Algorithm**
Initial Pivot Selection	**1**	0.19673	0.00773
	2	0.48714	0.21052
	3	0.09429	0.10245
	4	0.15875	0.14111
	5	0.29880	0.05803
Average		0.24714	0.10397

Standard Deviation		0.15329	0.07763
K-S Test (Two-Tail) N_2=4001			
Probability F (100) =F (10000)			
		First Algorithm (N_1)	**Second Algorithm (N_1)**
Initial Pivot Selection	**1**	0.99798 (69)	1.00000 (68)
	2	0.99241 (67)	1.00000 (68)
	3	0.75105 (65)	0.99588 (64)
	4	0.98194 (68)	0.99999 (62)
	5	0.94460 (69)	1.00000 (66)
Average		0.93360	0.99917
Standard Deviation		0.10415	0.00184

Two- tail refers to the bottom 20% and top 20% of the distribution.

EV percentiles under both sampling algorithms do not appear very sensitive to the initial pivot according to the results in Table 4. We observe that the EV percentiles sampled under both sampling algorithms are quite stable regardless of the initial pivot. *Second* algorithm consistently shows a better fit to the full run than *first* algorithm in terms of percentiles. Table 5 indicates that although the overall fit of EV distributions under both sampling algorithms seem sensitive to the initial pivot (*First* algorithm is more sensitive than *second* algorithm), the fit of two-tail EV distributions are not sensitive to the initial pivot. For both overall and two-tail EV distributions, *second* algorithm appears to be a more stable and effective sampling approach than *first*.

Figure 1
Cumulative Distribution Function for Economic Value:
1500 Scenarios
Figure 2
Cumulative Distribution Function for Economic Value:
50 Representative Scenarios, First Sampling Algorithm

Figure 3
Cumulative Distribution Function for Economic Value:
50 Representative Scenarios, Second Sampling Algorithm
Figure 4
Cumulative Distribution Function for Economic Value:
50 Representative Scenarios, Third Sampling Algorithm

Figure 5
Cumulative Distribution Function for Economic Value: New York 7 Approach