BY YVONNE C. CHUEH

Efficient Stochastic Modeling

Scenario Sampling Enhanced by Parametric Model Outcome Fitting

THIS ARTICLE EXAMINES SAMPLING EFFECTIVENESS by fitting the sample-run distribution to a parametric probability density and then compares the resulting density with the full-run distribution. Using the Accumulation Stochastic Economic Model (ASEM) as an example, it shows that the sampling may be enhanced through the parametric fitting.

ASEM is equipped with the flexibility to model simultaneously both the general account and separate accounts of deferred annuities business. The ASEM models the financial consequences of externally specified scenarios for a typical accumulation business. It captures the dynamic relationship among interest rate scenarios and other driving factors in accordance with the actual experience of the accumulation line of business. For example, disinvestment and disintermediation risks were considered by incorporating dynamic withdrawals, surrenders, fund transfers, and stochastic yields and returns. ASEM model outputs can be used to analyze ending surplus, economic value added, embedded value added (EVA), sensitivity, interaction relationships among drivers, and interest rate sampling methods.

Model Features

Some basic model features of ASEM are summarized as follows:

■ ASEM is a simplified asset-liability model with interactions between the asset model and the liability model.

■ ASEM develops annual pre-tax projections for both fixed and variable funds to the full 30-year period. Iterative formulas and dynamic relationships among drivers are applied to the beginning-of-year account values.

■ Any external scenario generator can easily be applied to ASEM. Current ASEM uses the stochastic variance interest rate model as a yield model. The bond and equity return model is a one-factor auto regression model that builds upon the yield model in accordance with historical data of correlation and volatility.

■ Three initial-value inputs into this model are initial fixed account value, initial variable account value, and initial annual interest rates.

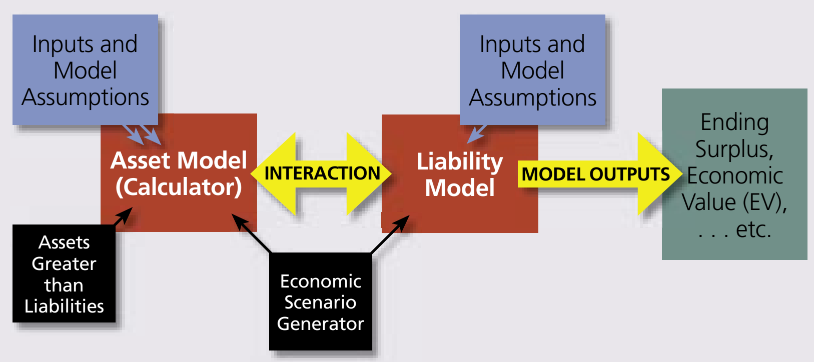
■ Key assumptions and information are applied to the model drivers, such as a °guaranteed floor rate for fixed account, investment spreads for fixed and variable accounts, surrender rates, surrender charges, fund transfers, deposits, withdrawals, portfolio rates, and new money rates.

■ Fixed-account earned rates for the projection period are developed using existing historical and projected cash flows, new money rates, old portfolio rates, and investment assumptions.

■ The model provides an intermediate summary of the projections, such as annual total margin, surplus, probability of ruin, ruin duration, economic value, and ending surplus.

■ A sensitivity test on each model driver can be obtained by fixing all driver assumptions except one or two. The report shows the impact of each model driver on model projections.

FIGURE 1. ASEM Model



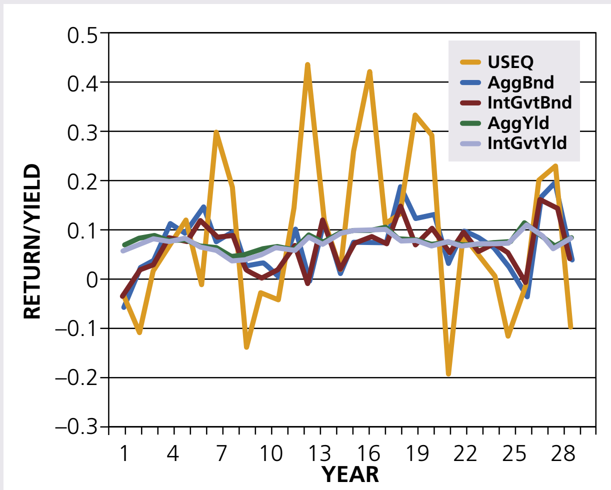
■ An investigation of interaction among two or three model drivers can be generated.

■ The probability distributions of EV and EVA can be obtained under stochastic model driver scenarios.

■ This model provides a support tool to study interest rate sampling methods with the advantage of great flexibility of making frequent changes on actuarial and economic assumptions. The sampling techniques have been proved to be effective in reducing run time for stochastic models while capturing the tail behavior.

YVONNE C. CHUEH is an assistant professor of actuarial science and statistics at Central Washington University in Ellensburg, Wash. She can be contacted at chueh@cwu.edu.

FIGURE 2. Economic Scenario



Efficient Scenario Sampling

After Longley-Cook's "Probabilities of 'Required 7' Scenarios," I proposed three efficient sampling methods to reduce the number of runs for stochastic modeling of assets and liabilities. Longley-Cook extended these sampling algorithms from interest rate paths to equity returns. His paper applied one of my sampling methods, the Significance Method, to reduce the number of runs and also to verify the efficiency of the sampling approach. To sample the tail distribution adequately, it's possible to modify the distance formula to suit different asset-liability model risks or business features. However, the modified distance formula has to satisfy the basic mathematical requirement for distance definition, such as triangle inequality, and be tested against a comparable business model.

To verify the effectiveness of the sampling, I used a different

FIGURE 3A. Full Run

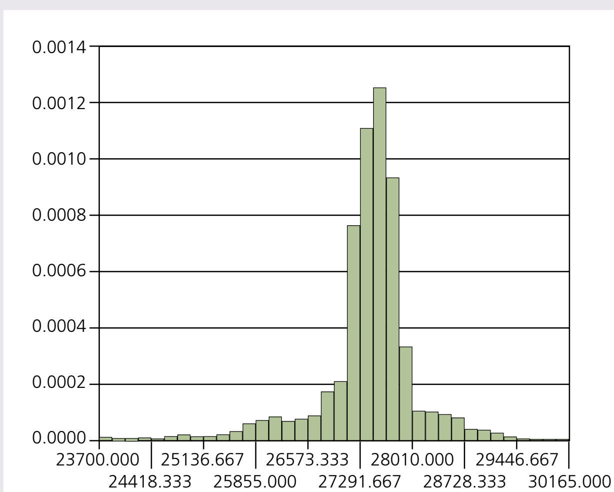


FIGURE 3B. Sample Run

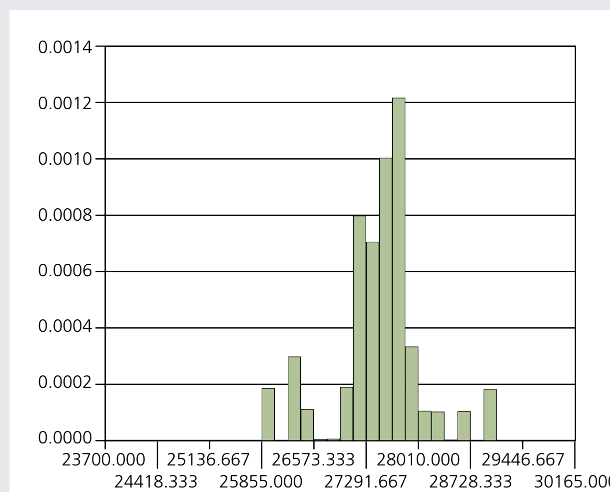


FIGURE 3C. Parametric Fitting

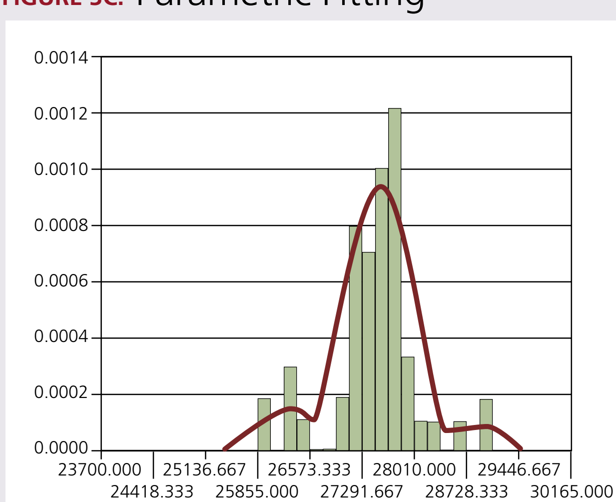


FIGURE 3D. Parametric Model

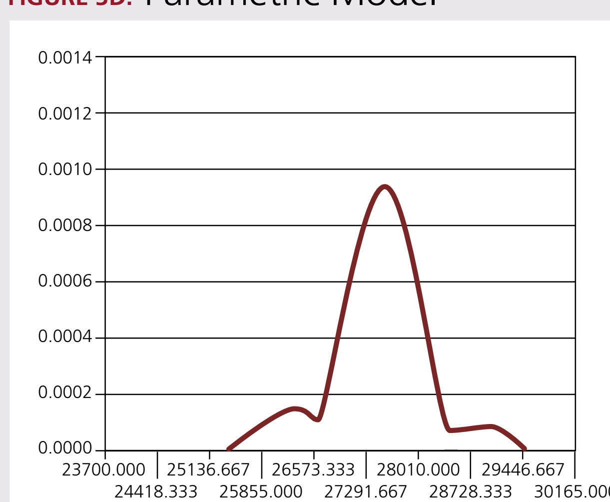


TABLE 1. Scenario Statistics

As of Dec. 31, 2002; Current Int. Gvt. Yield is 2.33%; Current Aggregate Yield is 4.06%

1-YEAR			CORRELATION				
	Mean	**Stdev**	**USEQ**	**AggBnd**	**IntGvtBnd**	**AggYld**	**IntGvtYld**
USEQ	2.5%	16.9%	1.00				
AggBnd	0.2%	4.4%	0.18	1.00			
IntGvtBnd	–0.4%	2.9%	0.06	0.95	1.00		
AggYld	5.2%	1.1%	–0.11	–0.96	–0.96	1.00	
IntGvtYld	3.8%	1.4%	–0.04	–0.92	–0.96	0.98	1.00
5-YEAR			**CORRELATION**				
	Mean	**Stdev**	**USEQ**	**AggBnd**	**IntGvtBnd**	**AggYld**	**IntGvtYld**
USEQ	6.0%	17.2%	1.00				
AggBnd	3.9%	2.6%	0.19	1.00			
IntGvtBnd	3.0%	2.2%	0.13	0.87	1.00		
AggYld	6.5%	1.4%	–0.03	–0.66	–0.39	1.00	
IntGvtYld	5.7%	1.8%	0.04	–0.57	–0.33	0.97	1.00
10-YEAR			**CORRELATION**				
	Mean	**Stdev**	**USEQ**	**AggBnd**	**IntGvtBnd**	**AggYld**	**IntGvtYld**
USEQ	7.5%	17.8%	1.00				
AggBnd	5.3%	2.7%	0.24	1.00			
IntGvtBnd	4.5%	2.9%	0.21	0.90	1.00		
AggYld	6.5%	1.4%	–0.04	–0.43	–0.17	1.00	
IntGvtYld	5.9%	1.8%	0.03	–0.34	–0.11	0.97	1.00
20-YEAR			**CORRELATION**				
	Mean	**Stdev**	**USEQ**	**AggBnd**	**IntGvtBnd**	**AggYld**	**IntGvtYld**
USEQ	8.4%	18.9%	1.00				
AggBnd	6.0%	2.9%	0.22	1.00			
IntGvtBnd	5.3%	3.4%	0.23	0.93	1.00		
AggYld	6.5%	1.4%	0.00	–0.28	–0.09	1.00	
IntGvtYld	6.0%	1.8%	0.04	–0.22	–0.05	0.97	1.00
30-YEAR			**CORRELATION**				
	Mean	**Stdev**	**USEQ**	**AggBnd**	**IntGvtBnd**	**AggYld**	**IntGvtYld**
USEQ	8.5%	17.6%	1.00				
AggBnd	6.3%	2.9%	0.24	1.00			
IntGvtBnd	5.6%	3.6%	0.24	0.94	1.00		
AggYld	6.5%	1.4%	–0.01	–0.26	–0.11	1.00	
IntGvtYld	6.0%	1.8%	0.03	–0.21	–0.08	0.97	1.00

external scenario generator to model the interest rate and equity return risks. Parametric probability models were sought that best fit the empirical distribution of the sample run. The resulting parametric probability curve was compared with the empirical distribution of the stochastic full run.

Stochastic Model Outputs

The major purpose of building stochastic models is to analyze the financial impacts of future uncertainties on business, especially future economic scenarios. To this end, actuaries are interested in the probability distribution of the model outputs that often include the key financial measures of business, such as ending surplus (ES) and economic or embedded value (EV).

In finding the best parametric probability models to fit the empirical data (empirical distributions of EV and ES), I used maximum likelihood estimate (MLE) technique within Microsoft Excel spreadsheets. I also used a computer software package, *Mathematica*, to calculate the parametric moments of the probability distributions and *S-Plus* (statistical software) to test the goodness of fit. Conditional tail expectations (CTE) were suggested by Hardy as a better tool to examine the tail behavior of a skewed probability density curve. *Mathematica* calculated the theoretical CTE using the parameter values obtained through Microsoft Excel's optimization tool, *Solver.*

Test the Sampling Algorithm

In testing scenario sampling methods, we chose the method I call "significance method" not only because it's the easiest to apply among the three but also because it results in equal probability for each selected scenario. The other two sampling methods can be tested following the similar procedure presented in this article.

To test the sampling method under the current low interest rate and negative equity return environment, we use a scenario generator that simulates stochastic Treasury and bond yields and correlated equity returns. The generator parameters are based on the five major asset sectors from 1973 to 2002. The mean, standard deviation, and correlations observed in the scenarios are summarized in Table 1.

TABLE 2. Parameter Values of MLE Probability Models Derived from EV Sample-Run Distribution

	WEIGHT	μ	σ	θ	α	γ	τ
LogNormal	0.6500	10.2317	0.0243				
LogNormal	0.3500	10.2010	0.0220				
Transformed Beta				27999.9908	17.6050	18.4131	12.1078
Inverse Transformed Beta				30000.0136	5.6620		18.0616
Inverse Gaussian	0.9430	27489.1261		61340922.1369			
Inverse Gaussian	0.0570	26303.0693		61340921.7238			
Transformed Gamma				480.5062	127.0642		1.1983

TABLE 3. Goodness of Fit (Moments)

MLE Probability Models vs. EV Sample-Run Distribution

MLE Probability Distribution Model	Moment	Parametric (Goodness of Fit)	Empirical	% Diff in Fit
Transformed Beta (excellent fit) Likelihood 2.1814E-169	1st	27422.69999	27421.50631	0.0044%
	2nd	752324665.1	752268428.2	0.0075%
	3rd	2.06483E+13	2.06463E+13	0.0097%
	4th	5.66956E+17	5.66891E+17	0.0113%
	neg 1st	3.64817E-05	3.64839E-05	-0.0059%
	neg 2nd	1.33148E-09	1.33167E-09	-0.0138%
	neg 3rd	4.86162E-14	4.8628E-14	-0.0241%
	neg 4th	1.77588E-18	1.77654E-18	-0.0372%
Mixed Inverse Gaussian (excellent fit) Likelihood 2.15375E-169	1st	27421.49604	27421.50631	0.0000%
	2nd	752274687.3	752268428.2	0.0008%
	3rd	2.06531E+13	2.06463E+13	0.0330%
	4th	5.67264E+17	5.66891E+17	0.0657%
	neg 1st	3.64878E-05	3.64839E-05	0.0109%
	neg 2nd	1.3321E-09	1.33167E-09	0.0328%
	neg 3rd	4.86598E-14	4.8628E-14	0.0655%
	neg 4th	1.77848E-18	1.77654E-18	0.1092%
Inverse Transformed Gamma (good fit) Likelihood 8.4492E-171	1st	27399.78561	27421.50631	-0.0792%
	2nd	751197632.6	752268428.2	-0.1423%
	3rd	2.06074E+13	2.06463E+13	-0.1884%
	4th	5.65665E+17	5.66891E+17	-0.2163%
	neg 1st	3.65182E-05	3.64839E-05	0.0943%
	neg 2nd	1.33436E-09	1.33167E-09	0.2026%
	neg 3rd	4.87855E-14	4.8628E-14	0.3239%
	neg 4th	1.78466E-18	1.77654E-18	0.4571%
Mixed Log Normal (good fit) Likelihood 2.1419E-151	1st	27482.33426	27421.50631	0.2218%
	2nd	755860038.9	752268428.2	0.4774%
	3rd	2.08048E+13	2.06463E+13	0.7675%
	4th	5.73086E+17	5.66891E+17	1.0927%
	neg 1st	3.6415E-05	3.64839E-05	-0.1887%
	neg 2nd	1.32707E-09	1.33167E-09	-0.3450%
	neg 3rd	4.83997E-14	4.8628E-14	-0.4694%
	neg 4th	1.76654E-18	1.77654E-18	-0.5627%
Transformed Gamma (not good at left tail) Likelihood 6.0866E-187	1st	27373.71274	27421.50631	-0.1743%
	2nd	753427622.9	752268428.2	0.1541%
	3rd	2.08501E+13	2.06463E+13	0.9869%
	4th	5.8012E+17	5.66891E+17	2.3336%
	neg 1st	3.6733E-05	3.64839E-05	0.6828%
	neg 2nd	1.35681E-09	1.33167E-09	1.8878%
	neg 3rd	5.03966E-14	4.8628E-14	3.6371%
	neg 4th	1.88245E-18	1.77654E-18	5.9614%

Figure 2 shows one of the scenarios generated. Sampling was applied in 2,000 scenarios using one-year yields in distance formula. The resulting empirical distributions of EV for both the full run (2,000 scenarios) and the sample run (50 selected scenarios) were compared. Figures 3(a) to 3(d) illustrate the idea of efficient modeling by scenario sampling and parametric probability fitting. Because of the run-time issue, Figure 3(a) is not available in reality. Figure 3(b) obtained from the sample scenario run is used to estimate the full-run distribution shown in Figure 3(c).

The two-sample Kolmogorov-Smirnov test was performed to test the null hypothesis that the CDF of EV in the full run equals the CDF of EV in the sample run. The ks = 0.1105, p-value = 0.5777. We failed to reject the null hypothesis at a high confidence level (99 percent).

Applying the Sampling Method

In applying the scenario sampling technique to the stochastic modeling of assets and liabilities, it's desirable to produce a parametric probability model of key financial or performance measure (e.g., EV, ES) derived from the sample run.

Past research by Robbins, Cox, and Phillips has emphasized the utility of parametric models in risk analysis and executive decision-making. The techniques of parametric model fit were also extensively discussed and demonstrated by Klugman and Panjer. Conditional tail expectation (CTE) has more advantages than other measures in evaluating the tail distribution

$$CTE = E[X|0<X<u] = \frac{\int_0^v xf(x)dx}{\int_0^v xf(x)dx}$$

where f(x) is the probability density function and u is a percentile.

In the initial attempt of model fit, normal, lognormal, weibull, and exponential probability distributions were applied to fit the data using both maximum likelihood and least-squares techniques. The fits were poor, with all the p-values equal to zero. These models should be rejected at any significance level.

To apply the sampling method to stochastic models and obtain a parametric EV probability model for a more practical risk analysis, we first fitted the empirical EV distribution of the sample run to a parametrical probability model using a maximum likelihood estimate (MLE) technique. Then we calculated the conditional tail expectations using model parameters.

The resulting model parameters were also used to predict the left tail of the full run to see how well the sampling method works. These parametric CTEs derived from the 50-sample run were compared with the empirical CTEs of the EV distribution of the 2,000 random economic scenarios.

An exhaustive list of mixed, transformed, and inverse parametric probability models were used to fit the sample data. The resulting parametric models, with their parameter values, are in Table 2. The top five models that fit the data well are listed in Table 3.

We used the *Solver* in Microsoft Excel 2000 to solve for the parameter values. The resulting parameter values are in Table 2. We used the software *Mathematica* to calculate the first four moments, both positive and negative, in order to evaluate how well the MLE models fit the empirical moments (data). The results are excellent. (See Table 3).

The top five models that gave the best CTE fit (to predict the full run) are listed in Table 4. It makes sense to evaluate the low percentiles alone because not only was the empirical EV distribution left skewed but it's the left tail the manager is concerned with.

Note that the EV distribution depends on the assets and liabilities, investment policies, and economic environment modeled by the scenario generators. As a result, a different stochastic asset-liability model may have a different form of parametric EV probability distribution.

Conclusion

Both the parametric model fit and sampling method ("Significance Method") give excellent results on the left tail of EV distribution. It's reasonable to assume that the right tail will fit well, too, since both positive and negative moments are excellent fits. The results are promising for modelers who are interested in parametric model outputs instead of just traditional empirical ones. As found in this study, the sampling effectiveness in the tail distributions has been further enhanced by fitting a parametric probability model.

This article summarizes a stochastic model and verifies that efficient modeling can be achieved by using a good sampling method. The robustness of other sampling methods can be tested following the similar procedures of parametric model fitting and testing using widely available computer technology and computation software. For interested modelers, a beta version of Visual Basic application program AMOOF that fits a variety of mixed distributions can be downloaded at: http://www.cwu.edu/~chueh/AMOOF.html.

An innovative and challenging application of sampling methods to efficient stochastic modeling has been described in a recent paper by Alastair Longley-Cook in the 2003 Stochastic Modeling Symposium. His creative application and impressive success are worthy of attention of stochastic modelers. ●

TABLE 4. Goodness of Fit (CTE)

Derived Parametric Probability Models vs. Full-Run Distribution

Parametric Model 50 Sample Run	Sample Run Percentiles	Parametric CTE E[EV| EV<kth sample percentile] 50 Sample Run	Empirical CTE = [Average EV | EV<kth sample percentile] 2000 Full Run	% diff
Mixed Lognormal	1%	25746.46	25447.82	1.16%
	2%	25747.36	25447.82	1.16%
	3%	25770.50	25479.61	1.13%
	4%	25803.60	25509.79	1.14%
	5%	25952.30	25673.56	1.07%
Transformed Beta	1%	25758.50	25447.82	1.21%
	2%	25759.50	25447.82	1.21%
	3%	25783.40	25479.61	1.18%
	4%	25817.60	25509.79	1.19%
	5%	25971.90	25673.56	1.15%
Inverse Transformed Gamma	1%	25816.10	25447.82	1.43%
	2%	25817.00	25447.82	1.43%
	3%	25839.70	25479.61	1.39%
	4%	25872.20	25509.79	1.40%
	5%	26017.80	25673.56	1.32%
Mixed Inverse Gaussian	1%	25774.66	25447.82	1.27%
	2%	25775.55	25447.82	1.27%
	3%	25798.90	25479.61	1.24%
	4%	25832.22	25509.79	1.25%
	5%	25982.48	25673.56	1.19%
Transformed Gamma	1%	24812.60	25447.82	-2.56%
	2%	24813.30	25447.82	-2.56%
	3%	24832.50	25479.61	-2.61%
	4%	24859.90	25509.79	-2.61%
	5%	24983.40	25673.56	-2.76%

TABLE 5. Goodness of Fit (CTE)

Sample-Run vs. Full-Run Distribution

CTE	Sample Run	Full Run	% diff
1%	25949.75	25447.82	1.97%
2%	25949.75	25447.82	1.97%
3%	25949.75	25479.61	1.85%
4%	25975.15	25509.79	1.82%
5%	25975.15	25673.56	1.17%

References

[1] Chueh, Yvonne C., "Efficient Stochastic Modeling: From Scenario Sampling to Parametric Model Fitting Utilizing ASEM as an Example," Symposium of Stochastic Modeling, 1-40, by Canadian Institute of Actuaries, Actuarial Foundation, and Society of Actuaries 2003.

[2] Chueh, Yvonne, "Efficient Stochastic Modeling for Large and Consolidated Insurance Business: Interest Rate Sampling Algorithms," *North American Actuarial Journal*, 2002.

[3] Hardy, Mary R., "A regime-Switching Model of Long-Term Stock Returns," *North American Actuarial Journal*, 2001.

[4] Johnson, Richard A., *Probability and Statistics for Engineers* 6th ed.; Prentice-Hall 2000.

[5] Klugman, Stuart A., Panjer, Harry H., Willmot, Gordon E., *Loss Models*, Wiley 1998.

[6] Longley-Cook, Alastair G., "Probabilities of 'Required 7' Scenarios (and a Few More)," *The Financial Reporter* (July 1997)

[7] Longley-Cook, Alastair, "Efficient Stochastic Modeling Utilizing Representative Scenarios: Application to Equity Risks," working paper, 2003.

[8] Robbins, Edward L., Cox, Samuel H., Philips, Richard D., "Application of Risk Theory to Interpretation of Stochastic Cash-Flow-Testing Results," *North American Actuarial Journal*, 1997.

Ms. Chueh's complete paper is available at www.contingencies.com.