Simulation Models for Self-Insurance

by Trent R. Vaughn

Simulation Models for Self-Insurance

Trent R. Vaughn, FCAS

Abstract

Actuaries are increasingly utilizing simulation models in a variety of practical applications. Most of these applications have focused on the future operating results and financial condition of an insurance company. These simulation models are also applicable in the risk management field, focusing on the financial consequences of self-insurance. This paper discusses the special considerations and applications of applying simulations to self-insurance. The author also discusses the integration of these results with the long-term and short-term financial plan of the self-insured.

1. INTRODUCTION

Many large commercial insureds choose to finance pure risk exposures with a combination of a self-insured retention (SIR) and excess insurance protection. The insured's retention is usually stated as a given dollar amount per claim (or per occurrence). *Specific excess* insurance provides coverage against individual losses in excess of this retention, up to the insurer's policy limit.[1] *Aggregate excess* insurance provides coverage against aggregate retained losses exceeding a given amount.[2] This aggregate excess insurance only covers losses arising from the insured's SIR; individual large losses above the specific excess policy limit are not covered by the aggregate policy.[3]

The insured will consider many issues before deciding on the appropriate combination of SIR and excess insurance coverage. Most insureds will forecast the expected cost of a risk financing program over several years, incorporating this cost into the pro forma financial statements. However, the diligent financial manager is not interested only in the forecasted cost of the risk financing program, but also in the variability of this cost.

From this standpoint, one may perform a simulation of the costs of the program over several years. This simulation should incorporate the stochastic nature of all the relevant

---

[1]Donald S. Malecki et al., *Commercial Liability Risk Management and Insurance*, vol. II (Malvern, Pa: American Institute for Property and Liability Underwriters, 1986), 342.

[2]Ibid., 342-343.

[3]Peter W. Rice, *The Risk Funding and Self-Insurance Bulletins*, vol. I (Cincinnati: National Underwriter Co., 1992), Ec-6.

variables, including aggregate retained losses, insurance premiums, and administrative expenses. A realistic simulation model should also include advanced features, such as a stochastic model of inflation, business cycles, and dynamic control.

This paper considers the unique aspects of applying simulation models to self-insurance situations, as well as potential uses of these models. Section 2 outlines the simulation of aggregate retained losses. Section 3 discusses the simulation of administrative expenses and excess insurance premiums. Section 4 details applications of the model.

2. TECHNIQUES FOR SIMULATING AGGREGATE LOSSES

Simulating aggregate losses for self-insureds presents special difficulties. Aggregate retained losses usually represent the largest portion of each year's total risk financing cost. However, excess insurance premiums can be subject to profit share and experience rating arrangements. Thus, in addition to aggregate retained losses, the model must also simulate aggregate excess (insured) losses for each year.

Current Methods

The simplest and most-familiar method of simulating an aggregate loss distribution is straightforward simulation.[4] This method proceeds by first simulating the number of claims, then producing the dollar amount of each claim. The advantage of straightforward simulation is that the loss amount is available for each individual claim. Thus, when we are dealing with insurance contracts which attach on a claim-by-claim basis, this method can easily determine the gross and

[4]James N. Stanard, "A Simulation Test of Prediction Errors of Loss Reserve Estimation Techniques," *PCAS LXXII*, 1985, 124-148.

ceded (insured) aggregate loss amounts. However, Daykin, et al., point out that straightforward simulation is "only applicable in the (rare) cases where the number of claims is fairly small and the claim size d.f. is easy to handle."[5]

In contrast to straightforward simulation, Daykin, et al., offer a "short cut" method, utilizing the Wilson-Hilferty (WH) formula.[6] Unfortunately, the inaccuracy of the WH formula increases dramatically as the skewness of the aggregate loss distribution increases.[7] Most simulation models circumvent this problem by defining both the individual claim size and aggregate claim amount variables net of claim-by-claim insurance (or reinsurance) arrangements. This convention significantly reduces the skewness of the aggregate claim amount distribution, allowing the WH formula to provide an acceptable approximation. Yet, by defining the relevant distributions net of insurance, we lose valuable information regarding individual large losses.

Proposed Hybrid Method

This section presents a hybrid simulation method, which combines the advantages of both current methods, while avoiding the inherent drawbacks. The model is briefly described, with mathematical details contained in Appendix A.

First, simulate the claim number variable, **k**=k, which is often assumed to have a mixed Poisson distribution. Next, utilizing the (gross) claim size distribution, determine the probability, p_e, that an individual claim will exceed the insured's retention. Then, generate the number of

[5]C.D. Daykin, T. Pentikainen, and M. Pesonen, *Practical Risk Theory for Actuaries* (London: Chapman & Hall, 1994), 143.

[6]Ibid., 144-145.

[7]Ibid., 132.

excess claims (k_e=k_c) as a random draw from a binomial distribution with parameters n=k and p=p_e. The number of claims that are strictly less than the retention is then given by $k_p = k - k_e$.

The aggregate losses pertaining to the k_p primary claims can then be simulated by means of the conditional WH generator, utilizing the censored (at the insured's retention) claim size distribution. The conditional WH generator slightly modifies the WH procedure, providing the aggregate loss amount when the number of claims is known (k_p in this case).[8] Lastly, the dollar amount of each of the k_e excess claims can then be simulated utilizing the conditional (given a claim amount in excess of the retention) claim size distribution and straightforward simulation.

The portion of each excess claim which exceeds the policy limit of the specific excess insurance protection is considered part of the retained (by the insured) loss. Thus, aggregate retained losses are equal to the sum of the following: (1) aggregate losses pertaining to the k_p primary claims, (2) the number of excess claims (k_e) multiplied by the insured's retention, and (3) the portion of any excess claim which exceeds the policy limit of the specific excess. Aggregate losses from sources (1) and (2) may also be subject to an aggregate excess policy.

The portion of each excess claim within the policy limit of the insurance is attributed to the loss experience of the specific excess contract. Hence, the aggregate losses attributed to the specific excess policy are easily determined by summing the amount of each excess claim between the insured's retention and the policy limit.

Provided that the insured's SIR is reasonably high, the number of excess claims which must be individually simulated (k_e) is sufficiently small. Moreover, the conditional claim size distribution (given that the claim is larger than the SIR) can be easily determined for claim size d.f.'s that are expressed in either analytic or tabular form. In fact, even in cases where the tabular

[8]Ibid., 146.

method is used for the claim size d.f., the tail of the distribution is still often expressed in analytic form.[9]

Finally, note that the aggregate losses for the k_p primary claims are simulated by means of the conditional WH-generator. The skewness of this aggregate loss distribution is usually low enough for the WH formula to provide an acceptable level of accuracy, since the tail of the claim size distribution has been removed by specific excess insurance.

Treatment of Allocated Loss Adjustment Expenses

Specific excess insurance contracts are usually designed to handle allocated loss adjustment expenses (ALAE) in one of two ways. First, the ALAE amount can be added to the indemnity loss prior to the application of the insured's SIR and policy limit. If this treatment of ALAE is employed, the individual size of loss distribution utilized in the simulation should be based on the total amount of indemnity loss and ALAE per claim.

Alternatively, the ALAE for each claim is prorated according to each party's share of the indemnity loss. In this case, the claim size distribution pertains only to indemnity loss; ALAE must be simulated separately--for both the k_p primary claims and the k_e excess claims.

For the k_p primary claims, the entire ALAE amount is attributed to the insured. This ALAE amount is dependent on both the number and the aggregate dollar amount of primary claims. Additionally, ALAE is incurred on claims which close with no payment; this expense will exhibit much year-to-year variation.

For the k_e excess claims, the ALAE amount must be simulated for each individual claim, then prorated according to the corresponding indemnity amount. This requires a bivariate model,

[9]Ibid., 75.

where the probability distribution of ALAE is dependent on the size of the indemnity loss. Details are provided in Hogg and Klugman.[10]

Determining the Payment Pattern

The simulation model discussed above is easily modified to produce aggregate retained losses by accident year and payment year. The expected claim count for each accident year is first allocated by payment year based on the estimated probability of a claim settling in the nth payment year. Aggregate losses for each combination of accident year/payment year are then simulated separately. This procedure is discussed in detail in Daykin, et al.[11]

A necessary refinement to the Daykin procedure is to allow the individual claim amount distribution to vary with the payment year. For instance, the mean and variance of the claim severity distribution will usually increase as the payment year (relative to the accident year) increases.[12]

The ALAE amount for each combination of accident year/payment year is simulated as discussed above.

---

[10]Robert V. Hogg and Stuart A. Klugman, *Loss Distributions* (New York: John Wiley & Sons, 1984), 167-177.

[11]Daykin, Pentikainen, and Pesonen, 296-298.

[12]Emanuel Pinto and Daniel F. Gogol, "An Analysis of Excess Loss Development," *PCAS LXXIV*, 1987, 245-247.

ɿulation of Losses for Multiple Accident Years

For most applications, the simulation will extend over a period of several accident years. ɿ analyzing multiple accident years, we must consider both changes in the claim severity distribution and the claim number distribution.

Changes in the claim severity distribution are elicited by claims inflation. In order to create a realistic model of uncertainty, we need to introduce a stochastic model of claims inflation. The loss severity distribution for each accident year/payment year is adjusted by the simulated claims inflation rate, as discussed in Daykin, et al.[13]

This stochastic model of inflation is a crucial submodel in the self-insurance context; great care should be exercised in accurately modeling inflation's future variability. Many insureds will find that exposures which are subject to extremely volatile inflation are not well suited for self-insurance.[14]

Also, in extending the model for several years, the number of claims will be affected by the growth in the self-insured portfolio. In this context, the number of claims by accident year is best modeled as the product of the claim frequency and the number of exposure units. The actual claim frequency (claims per exposure unit) will depend on trends, cycles and short-term (stochastic) fluctuation.[15] The number of exposure units is largely determined by the business plan of the insured, and may be treated as a deterministic variable. With this treatment, we are modeling the variation in risk financing costs for a given business plan.

---

[13]Daykin, Pentikainen, and Pesonen, 282-283.

[14]Rice, Ar-10.

[15]Daykin, Pentikainen, and Pesonen, 40-41.

3. SIMULATION OF ADMINISTRATIVE EXPENSES AND EXCESS INSURANCE PREMIUMS

Administrative Expenses

For self-insureds, administrative expenses can be a substantial portion of overall risk financing cost. These administrative expenses include the cost of services such as MIS, loss control, actuarial, and claims handling expenses that are not allocable to a specific claim. Usually, allocated loss adjustment expenses are included in the aggregate loss distribution, as the amount of these expenses is closely related to the dollar amount of indemnity loss (Section 2). A realistic long-term simulation should consider the stochastic behavior of these administrative expenses.

The model should separate in-house expenses from the cost of services purchased externally. In-house expenses, which largely consist of salaries, fringe benefits, and overhead items, may increase through time broadly in line with general earnings inflation and the growth of the self-insured portfolio. Other in-house expenses, such as the cost of MIS equipment, are most accurately modeled as step-increase expenses,[16] with the number of incurred claims determining the position of the steps. For example, a certain number of claims may exceed the capacity of the current management information system, prompting additional investment.

External costs may exhibit much more fluctuation, as they are affected by the specific forces of supply and demand for these types of services. These costs should also increase broadly in line with earnings inflation and growth of the portfolio, but with larger year-to-year variations.

Additionally, the model should include the cost of assessments, fees and taxes which are often levied on self-insurers. For example, firms which self-insure workers compensation

---

[16]Ibid., 323.

.posures are often required to contribute to second-injury funds.[17] A detailed discussion of the ature and extent of these assessments is beyond the scope of this paper.

Lastly, one can also monitor varying levels of loss control efforts. For example, the model may specify that if aggregate retained losses (insured losses may also be included) exceed a certain threshold, this will trigger a strategic impact variable,[18] representing the cost of improved loss control efforts. Presumably, this loss control effort would reduce the future frequency and/or severity of retained losses. Such a feature requires careful consideration, as not all exposures will respond to loss control efforts.

Excess Premiums and Business Cycles

Many articles and publications have dealt extensively with the issue of the insurance underwriting cycle. Empirical evidence suggests that the underwriting results of the insurance industry, measured by the combined ratio, fluctuate in an irregular cyclical fashion.[19] While self-insurance offers some protection against the underwriting cycle, self-insureds and captives are still exposed to the risk of a hard market in the area of excess insurance or reinsurance.

In the short-term, the commercial premium for excess coverage is easily determined. For a guaranteed cost policy, the amount and timing of the payments are fixed in advance, subject only to possible audit adjustments. However, over the long-term, the price and availability of excess coverage are subject to the vagaries of the insurance underwriting cycle. In fact, evidence

[17]Rice, vol. II, Da-1 through Dwy-1.

[18]Daykin, Pentikainen, and Pesonen, 323-324.

[19]Steven P. D'Arcy, "Investment Issues in Property-Liability Insurance," *Foundations of Casualty Actuarial Science* (New York: Casualty Actuarial Society, 1990), 510-514.

exists, in the context of reinsurance, that the underwriting cycle is more severe for excess business.[20] Moreover, this excess insurance cycle will not necessary coincide with the underwriting cycle in the primary market.

A realistic long-term model should incorporate the cyclical aspects of excess coverage premiums. Methods for accomplishing this are discussed in Daykin, et al.[21] In particular, the model should consider the potentially harsh consequences of a hard market occurring early in the life of a self-insurance program, possibly leaving the insured without excess protection at reasonable prices and terms.

Excess Premiums and Experience Rating

Many insureds purchase numerous layers of specific excess coverage, each from a different insurer. In this manner, the insured is not financially reliant entirely on one carrier. However, the disadvantage of such an approach is that conflicts often arise between the various insurers.[22] Moreover, the total insurance coverage may be more costly, as the insurers are duplicating certain underwriting functions.

A reliable long-term model should recognize that future premiums for lower level coverage will be affected by the loss experience incurred under these contracts, in addition to the premium effects caused by the market cycle. Possible mechanisms for accomplishing this are

---

[20]R.C. Reinarz et al., *Reinsurance Practices*, vol. II (Malvern, Pa.: Insurance Institute of America, 1990), 90-91.

[21]Daykin, Pentikainen, and Pesonen, 327-356.

[22]Rice, vol. I, Ec-4.

xperience rating and exponential smoothing,[23] which can be incorporated into the model. The higher level excess contracts may be regarded as more of a pure risk situation; insurers will not usually require "payback" of a large loss, nor should the insured expect a refund from favorable experience.[24]

Since the simulation model outlined in Appendix A produces individual information on each excess claim, determining the experience of each insured layer is possible. However, in long-term analysis, one should consider the effect of delayed payment of claims ("run-off error"). Generally, the experience under excess contracts takes much more time to develop and assess profitability. In fact, the insurer may not acquire reliable knowledge regarding the true profitability of the contract until many years after it has expired. Thus, the time lag between the loss experience and its reflection in premium rates is usually longer for excess coverage.

The model builder may also incorporate the risk that very adverse loss experience will leave the insured without coverage at any price, forcing complete self-insurance.

Dynamic Control - Modifying the Level of Excess Insurance Protection over Time

Generally, the insured will gradually increase the retention and policy limit of the specific excess roughly in lockstep with inflation. However, the amount of excess protection desired and available is also dependent on market conditions and the insured's loss experience.

In practice, the management of self-insured firms is constantly fine tuning the risk financing strategy. When the market is soft, many insureds purchase additional insurance

[23]Daykin, Pentikainen, and Pesonen, 180-183.

[24]Richard F. Gilmore, "Planning and Managing a Reinsurance Program," *Reinsurance*, ed. R.W. Strain (New York: College of Insurance, 1980), 384-385.

coverage (through higher limits and/or reduced deductibles). Likewise, insureds often rely more heavily on self-insurance when the market hardens, out of desire or necessity.

As discussed above, the premium for lower level excess contracts is also influenced by the developing loss experience. Unfavorable results will prompt the excess carrier to increase the required premium, possibly with a time lag. If this premium becomes too high, the insured may choose to increase the attachment point of the coverage.

The final model may allow the attachment point (and possibly policy limit) of the excess coverage to respond to market conditions and loss experience. However, it is important to recognize the limitations on this flexibility. A significant self-insurance program, once in place, is not easy to abandon; staff and resources have been dedicated and may not be easily transferable to other areas.

4. APPLICATIONS

Presentation of Results

For each year of the planning horizon, the simulation model will produce a probability distribution for the total risk financing cost, on both a cash basis and an accounting basis. The primary distinction between the cash model and the accounting model is in the treatment of retained losses. In the cash model, the risk financing cost for each year includes only the paid retained losses; in the accounting model, the risk financing cost includes the total retained losses from claims occurring any time during the year, regardless of when the payment is made ("incurred losses"). Daykin, et al., discuss the relationship between the two methods of defining losses, including the treatment of run-off error and discounting in the accounting model.[25]

This distinction between paid and incurred losses is crucial, as most self-insureds are only permitted to deduct paid losses from each year's taxable income. When an insured initially switches from fully insured status to a self-insurance program, federal income taxes can be expected to increase, at least in the first few years of the program. This is especially true for self-insurance of long-tailed exposures, where the payment date may extend several years past the occurrence date.

The cash model is most useful for the applications discussed below. The accounting model is not discussed in detail, but it can be useful for demonstrating the potential impact of a risk financing program on future financial accounting statements.

Comparison of Alternative Programs

The simulation model is well suited for comparing the relative cost savings and cost stability of alternative risk financing solutions, over a period of several years. A fully insured strategy is a good "base case" scenario from which to compare alternatives.

A common criticism of simulation routines is that the results for alternative scenarios may be dramatically influenced by sampling error.[26] Daykin, et al., respond that "the impact of the distortion resulting from sampling errors can be greatly reduced by using the same sequence

---

[25]Daykin, Pentikainen, and Pesonen, 9-10.

[26]Sholom Feldblum, "Forecasting the Future: Stochastic Simulation and Scenario Testing," *Incorporating Risk Factors in Dynamic Financial Analysis* (Landover, Md: Casualty Actuarial Society, 1995), 156.

of random numbers for each of the concurrent simulations (i.e. always starting from the same specified seed of the primary random number generator)."[27]

Applications to Long-Term Financial Planning

Most business firms rely on a combination of debt and equity for long-term financing needs. The trade-off theory of capital structure proposes that business firms determine the optimal debt-equity ratio as a trade-off between interest tax shields and the costs of financial distress.[28] The risk financing decision interacts with this trade-off theory by affecting the taxes and business risk of the firm.

For instance, most financial economists agree that there is a moderate tax advantage to corporate borrowing for firms which are reasonably confident that they will be in a taxpaying position.[29] The cash basis simulation model displays the probability distribution of tax deductible risk financing costs for each year in the planning horizon. This distribution can be used to calculate the probability of these costs exceeding a given threshold at which the firm would not be able to fully utilize tax deductible interest payments during any given year in the future. If this probability is significant, the tax advantage of the debt is decreased.

Secondly, for a given level of debt, the probability of financial distress increases as the business risk of the firm increases. The cash basis simulation model can be used to estimate the probability of risk financing costs exceeding a given threshold, causing the firm to default (or

---

[27]Daykin, Pentikainen, and Pesonen, 154.

[28]Richard A. Brealey and Stewart C. Myers, *Principles of Corporate Finance* (New York: McGraw-Hill Book Co., 1991), 434.

[29]Ibid., 447.

have difficulty meeting) its principal and interest obligations during any year in the future. As this probability of financial distress increases, the expected cost of financial distress also increases.

The distribution of risk financing costs during each year of the planning horizon will also influence the timing of debt and equity issues. For example, if an insured switches from fully insured status to self-insurance, next year's expected cash outflow may be dramatically reduced. Depending on the variability of these risk financing costs, the insured may want to invest part of the savings in an increased level of net working capital. However, the balance can be invested in operations or utilized to reduce long-term debt or equity.

Applications to Short-Term Financial Planning

The cash basis model can also be utilized to demonstrate the impact of the risk financing program on short-term cash planning. Although most simulation models are designed to display results at annual intervals, a simulation can be created to demonstrate the variability of risk financing costs during the next 12 months. This model can be extremely useful to insureds in determining net working capital requirements. For example, the model may reveal the significant probability of a required cash outflow which would exceed the insured's cash balance, line of credit, and other sources of short-term borrowing (such as "stretching payables").[30] This may prompt the insured that an increase in net working capital is required.

This discussion exposes an often-overlooked cost of self-insurance. Self-insurance increases the short-term variability in cash outflow, often increasing net working capital requirements.

---

[30]Ibid., 736.

Relationship to Insured's Business Plan

The discussion above analyzes the variability in risk financing costs, relative to the insured's projected cash flows net of risk financing costs. However, the insured's cash flow from operations (net of risk financing costs) is dependent on the risk financing costs. For example, if risk financing costs are exaggerated by an unusually high rate of claims inflation, this high inflation will also affect other areas of the insured's operation (for example, revenues and cost of goods sold).

If the insured utilizes a Monte Carlo simulation model for other aspects of the business, this model can be integrated with the risk financing model. The overall model should recognize that claims inflation reflects general price inflation as well as social inflation.

If the insured does not utilize a Monte Carlo simulation, the results of the risk financing model are still valuable in demonstrating the variability involved in alternative risk financing solutions.

5. CONCLUSION

Actuaries are increasingly utilizing simulation routines for a variety of practical applications; most of these applications focus on the future operating results and financial performance of insurance companies. Simulation models are also applicable in the risk management field, focusing on the financial consequences of self-insurance of pure risk exposures.

The application of simulation models to self-insurance presents unique challenges. Excess insurance will protect the insured from a catastrophic frequency or severity of losses, but premiums for this excess coverage are subject to outside influences. Also, many insureds are capable of mitigating their loss exposures through loss control programs. A realistic model should reflect management's ability to respond to changing conditions in loss experience and in the insurance marketplace.

These simulation models will assist the risk manager in evaluating the cost and cash flow variability inherent in alternative risk financing strategies. These results can be incorporated into the overall financial plan of the insured. In this manner, simulation models will serve as a useful planning tool for risk managers, and as an important marketing tool for insurers.

Appendix A - Hybrid Simulation Model for Aggregate Losses

Step 1: Simulate random value of **q**=q, the mixing variable.

Step 2: Simulate random value of **k**=k, the number of claims from a mixed Poisson distribution with mean nq.

Step 3: Determine the probability, p_e, that a given claim will exceed the insured's retention, M. This probability is given by 1 - S(M), where S(x) is the (gross) claim size distribution.

Step 4: Simulate the number of excess claims, $\mathbf{k_e}=k_e$ as a random draw from a Binomial distribution with parameters n = k and p = p_e.

Step 5: For each of the k_e claims, simulate the amount of the claim utilizing the conditional claim size d.f., given that a claim exceeds M. This conditional d.f. is given by:

[S(x) - S(M)] divided by [1 - S(M)] for x greater than or equal to M

Step 6: Simulate the aggregate amount of loss with respect to claims that are strictly less than the retention. The conditional claim size distribution, given that a claim is less than the retention, is given by:

S(x) divided by S(M) for x less than or equal to M

Next, calculate the moments of this conditional distribution. These moments, as well as the number of claims less than M, k_p = k - k_e, are inputs in the conditional WH generator. This conditional WH generator will produce a simulated aggregate amount of losses for claims which are strictly less than M.

Step 7: The aggregate amount of losses under M is given by the aggregate losses from step 6 plus k_e x M.

Step 8: The aggregate amount of losses in excess of M is given by the sum of the k_e loss amounts minus k_e x M.