PROFIT/CONTINGENCY LOADINGS AND SURPLUS:

RUIN AND RETURN IMPLICATIONS

By Gary G. Venter

As its name suggests, the profit/contingency loading in the insurance rate serves two purposes: (1) to aid solvency by absorbing some degree of fluctuation in loss experience and (2) to provide a suitable average return to the underwriter.

If solvency were considered alone, the loading amount could be reduced as long as surplus were increased commensurately. However, the desire to receive an adequate return on surplus will tend to keep the ratio of loading to surplus from becoming too low. It seems therefore that there should exist an equilibrium point where (a) surplus and the loading together give sufficient protection against insolvency, (b) the loading is high enough to yield the desired average return rate and (c) for any lower loading amount any surplus selected will violate either (a) or (b). The most competitive premium consistent with the ruin and return goals will result from this minimum loading.[1]

By formulating ruin and return objectives it will be found that such a point does exist and that it can be expressed in terms of statistics of the portfolio loss distribution. The solvency objective will be specified by giving a maximal acceptable probability of ruin for a set time period, while profitability goals will be assumed to be in terms of the desired rate of return on the surplus invested.

To highlight the interplay of surplus and the profit/contingency loading, these considerations will exclude investment income and expenses. That is, the return rate will be premiums minus losses over surplus, or

$$R = \frac{W - L}{S} \qquad (1)$$

where W is premiums, L is losses, R is the rate of return and S is surplus.

The total amount of loading dollars and the necessary surplus will first be determined for a company's entire portfolio of insured risks. A method of allocating the loading amount to contract will then be discussed.

TOTAL PORTFOLIO

Taking the expected values of both sides of (1) yields:

$$E(R) = \frac{W - E(L)}{S} \tag{2}$$

or

$$W = E(L) + S \cdot E(R) \tag{3}$$

This expresses the premium as the expected losses plus the profit/contingency load of $S \cdot E(R)$.

Accepting a probability of ruin of ε means the insurer wants $P\{W + S \leq L\} \leq \varepsilon$, where P is the probability operator.

By (3) this becomes $P\{E(L) + S \cdot E(R) + S \leq L\} \leq \varepsilon$ or:

$$P\left\{\frac{S + S\,E(R)}{\sigma_L} \leq \frac{L - E(L)}{\sigma_L}\right\} \leq \varepsilon \tag{4}$$

where σ_L represents the standard deviation of L. Let T_ε be defined as the smallest x such that $P\{x \leq \frac{L - E(L)}{\sigma_L}\} \leq \varepsilon$, i.e., T_ε is the number of standard deviations above expected losses one must go to have a probability of only ε of experiencing higher actual losses. Since T_ε is the smallest such x, the probability of ruin criterion (4) becomes:

$$\frac{S + S \cdot E(R)}{\sigma_L} \geq T_\varepsilon \tag{5}$$

If we were to graph the relationship between surplus (x - axis) and the load SE(R) (y - axis), any combination of surplus and loading that is equivalent to a point on or above the line $SE(R) = T_\epsilon \sigma_L - S$ would therefore meet the ruin criterion.

This acceptable region is further restricted by allowing only those combinations of surplus and loading which meet management's return goals.

For instance, consider the case where the insurer has a fixed return percentage aim E(R) = a. Then (5) will be satisfied as long as $S \geq T_\epsilon \sigma_L/(1 + a)$, and the loading aS will be minimized at the minimum value of S in this region, namely at $S = T_\epsilon \sigma_L/(1 + a)$. The minimum value taken is then $SE(R) = aS = aT_\epsilon \sigma_L/(1 + a)$.

Instead of requiring a fixed rate of return, a more general approach would be for management to seek to have the expected return rate increase or decrease with the uncertainty of return that is inherent in its current mix of business. For example, it may attempt to maintain the relationship:

$$E(R) \geq a + b\,\sigma_R \qquad (6)$$

where σ_R is the standard deviation of the rate of return, and a and b are fixed constants. This can be interpreted as requiring the expected return to be at least as great as some risk free rate a plus b times the standard deviation of the return rate. As (1) implies that $\sigma_R = \sigma_L/S$, $a + b\,\sigma_R$ will increase as S decreases. Thus, (6) will allow the underwriter a choice of return rates depending in part on the surplus invested. Substituting σ_L/S for σ_R in (6) gives $SE(R) \geq aS + b\,\sigma_L$ which constrains the risk load to be equal to or greater than the linear function of surplus $aS + b\,\sigma_L$. This relationship is graphed along with the ruin requirement (5) on Exhibit 1. The points in the cross-hatched area above both lines represent surplus/loading combinations which satisfy both the ruin and the return requirements. As is apparent from the graph, the minimum acceptable risk load SE(R), which generates the minimum acceptable premium E(L) + SE(R), occurs where the two lines intersect. This point can be determined by simultaneously solving the two equations:

$$SE(R) = T_\varepsilon \sigma_L - S \qquad \text{and}$$

$$SE(R) = a S + b \sigma_L$$

for S and E(R). Doing this yields $S = \frac{T_\varepsilon - b}{a + 1} \sigma_L$ and $SE_R = \frac{a T_\varepsilon + b}{a + 1} \sigma_L$, which expresses surplus and the loading in terms of the constants a and b and the statistics T_ε and σ_L of the portfolio aggregate loss distributions.

Formulating the return goal as a linear function of the standard deviation of the return rate is consistent with several investment studies as further referenced in Appendix 1. If some other profitability goal is followed it will still be possible in most cases to find a minimum loading point in the region of loading/surplus pairs satisfying the ruin and return requirements. It should be noted, however, that if the return objective allows for the return amount to decrease as surplus increases in some interval then the minimum point may occur strictly above the ruin boundary $SE(R) = T_\varepsilon \sigma_L - S$. As an example of this, if $T_\varepsilon = 3.1$ and a return criterion $E(R) \geq .04 + .36 \sigma_R^2$ is followed, the return boundary will be given by $SE(R) = .04 S + .36 \sigma_L^2/S$. This boundary takes its minimum of $.24 \sigma_L$ at at $S = 3 \sigma_L$. Since this point also satisfies the ruin requirement $SE(R) \geq T_\varepsilon \sigma_L - S$ it is the minimum loading possible under both objectives.

Exhibit 2 graphs this situation.

ALLOCATION TO CONTRACTS

Each insurance contract can be treated as if it were an insurer's entire portfolio in itself, in order to determine the surplus and profit/contingency loading required to support the contract on a stand alone basis. The sum of the stand alone loadings and surpluses over all the insurer's contracts will usually exceed the load and surplus needed for the portfolio in its entirity, because of the benefit of pooling. Thus, the purchasers of insurance will in general be charged less than their stand alone loading requirement. The stand alone load is what it would cost to transfer the risk if pooling were not possible. Thus, a reasonable principle for the allocation of the portfolio loading to contract would be to give each risk a uniform percentage reduction from his stand alone loading amount. This credits each insured proportionally for the salutary effect of pooling.

Non-pooled risk transfers are not common commercial transactions when pooling is available, so the price assigned to such contracts should be based on sound theory. The profit criterion $E(R) \geq a + b\,\sigma_R$ has a fair amount of support and thus seems appropriate for this application. As mentioned above, the stand alone loading under this criterion is $\frac{a\,T_\varepsilon + b}{a + 1}\sigma_L$, where now T_ε and σ_L refer to the individual contract's loss distribution. The actual loading for each risk will be proportional to this amount, with the constant of proportionality determined as the ratio of the loading required for the entire portfolio to the sum of the stand alone loadings for the contracts in the portfolio. For most insurers a model of the distribution of risks in the portfolio will be needed to estimate this constant. An example for a simple portfolio is given in Appendix 2.

Each contract's loading can also be expressed as being proportional to $(1 + \frac{a}{b} T_\varepsilon)\sigma_L$ (by dividing the stand alone loading by $\frac{b}{a+1}$). Numerically this would be $(1 + .15\ T_\varepsilon)\ \sigma_L$ using the estimates a = .047 and b = .316 mentioned in Appendix 1. It is important to realize that this is not a standard deviation loading, because for a given ε, T_ε will usually depend on the higher moments of the loss distribution, and thus may vary significantly among contracts.

It should be noted as well that the profit/contingency loading developed by this approach is not additive. That is, the sum of the loads for two separate contracts will not in general equal the load for a single contract that encompasses both risks. Additivity is sometimes considered advantageous for a loading system on the grounds of ease of calculation, but it is not clear that it is better in principle. I would like to argue that a loading system should not be additive in the sense above, essentially because combining contracts is itself a form of pooling that reduces risk. Thus, a large insured with many homogeneous exposures should have a smaller proportional load than an insured with a single similar exposure unit.

In favor of additivity, it could be argued that a collection of smaller risks equal in size to a single large risk should have the same total loading, because the contribution to the portfolio loss distribution would be the same. This argument does not consider the demand side of the pricing mechanism, however. Even if the supplier of insurance views writing a large contract and several small accounts as equivalent propositions, the small insureds would probably be willing to pay more proportionally above their expected losses to transfer their risk. This is because fluctuations in losses would be proportionately greater for the smaller insured, and because the assets available to absorb these fluctuations would probably be smaller.

Supply and demand will thus interact to yield a higher profit/contingency loading percentage for the smaller insured in particular and more generally for the proportionately more hazardous contracts. The above would suggest that this is not a caprice of the marketplace but rather reflects the higher costs of capital and the proportionately greater surplus funds needed to absorb the proportionally greater fluctuations in actual experience.

APPENDIX I

The notion that the expected rate of return should increase linearly with the standard deviation as one moves to riskier investment portfolios has some support in the theory of capital markets. Sharpe in [6] demonstrates theoretically that under certain general assumptions on investor behavior, stock prices will adjust themselves to maintain a relationship of the form $E(R) = a + b\,\sigma_R$ for all efficient investment porfolios. An efficient portfolio is one which cannot be improved upon in one of the values of risk or return without deteriorating in the other.

In [7], Sharpe tested this theory using 10 years of return results for 34 mutual funds. While a degree of scattering about the linear relationship $E(R) = a + b\,\sigma_R$ was observed, the linearity was generally confirmed. Expanding the formulation to $E(R) = a + b\,\sigma_R + c\,\sigma_R^2$ produced only a slight improvement in the correlation coefficient (from .836 to .852).

Cooper in [2], Chapter 4, summarizes Sharpe's work and several related studies, and develops estimates for the coefficients a and b by investigating the risk/return relationships for 25 mutual funds during the fifteen year period 1957 to 1971. His result is $E(R) = .047 + .316\,\sigma_R$.

APPENDIX 2

EXAMPLE OF LOADING CALCULATION

An auto BI liability portfolio is assumed to consist of 100 single unit exposures plus a fleet of 100 vehicles. The expected frequency of each unit is 2% with a variance of 2% and the expected severity is $5000 with a standard deviation of $10,000. Then the expected losses are $100 with a variance of 2,500,000 square dollars by the formula: variance(loss) = variance(frequency) $(\text{expected severity})^2$ + (expected frequency) variance(severity) See [3] page 14. To estimate T_ε the gamma approximation to the loss distribution will be used as recommended by Seal [5]. The gamma parameters λ and r are determined by $\lambda = E(L)/\sigma_L^2 = \frac{1}{25000}$ and $r = \lambda\, E(L) = \frac{1}{250}$. Sums of independent gamma distributions are themselves gamma with the r_i parameters being additive (see [4] page 70). Thus, for the fleet $r = \frac{100}{250} = .4$ and for the entire portfolio $r = \frac{200}{250} = .8$. A constant λ is maintained throughout as required by $r = \lambda\, E(L)$. T_ε is determined by $\frac{\lambda^r}{\Gamma(r)} \int_0^{T_\varepsilon \sigma_L} y^{r-1} e^{-\lambda y}\, dy = 1 - \varepsilon$
The left hand side can be estimated by taking enough terms of the sum;

$$\frac{e^{-x} x^{r-1}}{\Gamma(r)} \sum_{i=0}^{\infty} \prod_{K=0}^{i} \frac{x}{r+K}$$

(see [1] page 262, formula 6.5.29), where $x = \lambda\, T_\varepsilon\, \sigma_L$. For $\varepsilon = .01$ and $\infty = 100$, x was determined iteratively, yielding the table below:

	One Exposure	Fleet	Portfolio
r	.004	.4	.8
$x = \lambda T_\varepsilon \sigma$	.04786	3.000	4.130
$\lambda\sigma = \sqrt{r}$	.06324	.6324	.8944
T_ε	.7568	4.744	4.617
σ	1581	15,810	23,360
$T_\varepsilon \sigma$	1197	75,000	103,200

Using a = .06 and b = .4 yields stand alone loading dollars of $664.30, $10,210, and $14,280 for the single risk, fleet and portfolio respectively. The total of the stand alone loading amounts is $74,640 so a factor of 14280 ÷ 74640 = .1913 can be applied to each case yielding $127 for each risk and $1953 for the fleet. The fleet then is charged 15 times as much loading as each individual risk. This compares to 10 times as much for a standard deviation loading and 100 times as much for a variance load.

The absolute magnitudes of the charges are probably unmarketable, which would indicate that a portfolio of 200 such risks would be too small to meet these particular ruin/return goals.

NOTES

1. The author is indebted to several conversations with Charles Hachemeister for this formulation of the problem and for a number of specific suggestions in developing the resulting study.

REFERENCES

[1] Abramowitz, M. and Stegun, I., editors, Handbook of mathematical functions, U. S. Department of Commerce (1964).

[2] Cooper, R., Investment return and property-liability insurance ratemaking, Irwin, Homewood Illinois (1974).

[3] De Vylder, F., Introduction aux theories actuarielles de credibilite, University Catholique de Louvain (1975), English translation by C. Hachemeister awaiting publication.

[4] Seal, H., "Approximations to risk theory's F (x, t) by means of the gamma distribution", Astin Bulletin 9, 213-218 (1977).

[5] Hastings, N. and Peacock, J., Statistical distributions, Butterworths, London (1975).

[6] Sharpe, W., "Capital asset prices: a theory of market equilibrium under conditions of risk", Journal of Finance 19, 425-442 (Sept. 1964).

[7] Sharpe, W., "Risk aversion in the stock market: some empirical evidence", Journal of Finance 20, 416-422 (Sept. 1965).

EXHIBIT 1

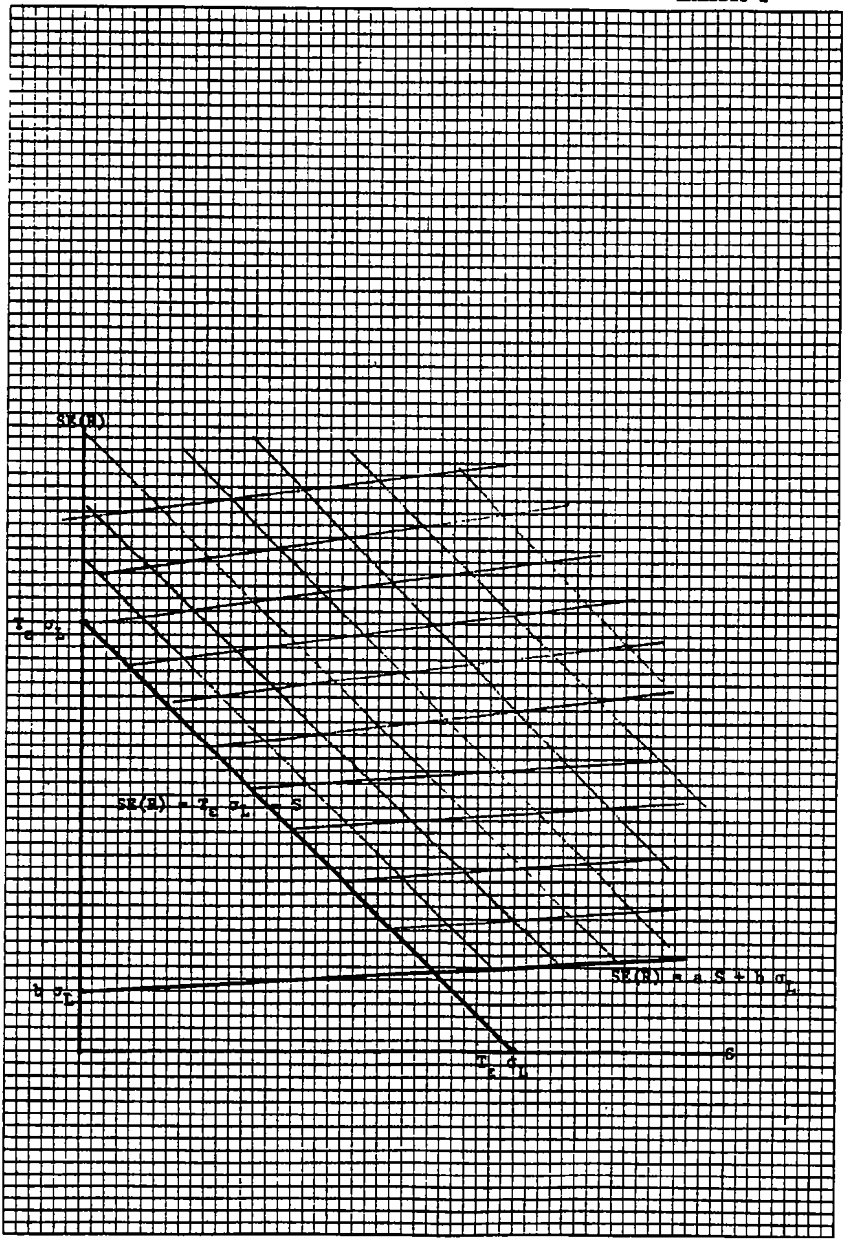
SE(H)
$T_e \sigma_L$
$SE(H) = T_e \sigma_L - S$
$b \sigma_L$
$SE(H) = a S + b \sigma_L$
$T_e \sigma_L$
S

EXHIBIT 2