# THE COMPETITIVE MARKET EQUILIBRIUM RISK LOAD FORMULA FOR INCREASED LIMITS RATEMAKING

GLENN G. MEYERS

*Abstract*

*Insurance Services Office, Inc. (ISO) has adopted a new risk load formula which is to become effective with 1991 advisory increased limits filings. This paper describes the underlying rationale of the new risk load formula. This formula differs from previous ISO formulas in that: (1) it is derived from competitive market assumptions; and (2) it recognizes the risks faced by the insurer in estimating the price of its product; i.e., parameter uncertainty. After the derivation of the formula, the paper will discuss considerations to be made by the insurer when using the formula. These considerations include excess-of-loss reinsurance.*

## 1. INTRODUCTION

It is a common observation that, as the policy limit increases, the premium for a casualty insurance policy rises faster than its expected cost. This observation fits well with the economic principles of supply and demand. Policies with higher limits are perceived as being more risky. Insurers are more reluctant to sell them and insureds are more anxious to buy them. In the language of increased limits ratemaking, the additional premium to cover this increased risk is called the risk load. A risk load which rises faster than the expected cost as the policy limit increases is necessary if higher policy limits are to be made available.

In the late 1970s, Insurance Services Office, Inc. (ISO) introduced increased limits factors which were calculated with an explicit formula for the risk load.[1] In the years that followed, the formula has been refined

[1] See *Report of the Increased Limits Subcommittee: A Review of Increased Limits Ratemaking*, Insurance Services Office, Inc., 1980. The work done by ISO was based on Miccolis [1].

or revised a number of times, often with considerable debate. A major part of the debate has centered around whether or not the risk load formula met the demands of a competitive marketplace.

Effective with 1991 advisory increased limits filings, the ISO risk load formula has undergone still another change. As is the case with all advisory filings, each insurer must make its own decision to accept or modify the contents.

This paper describes the underlying rationale of the new risk load formula. This formula differs from previous ISO formulae in that: (1) it is derived from economic assumptions about the competitive market; and (2) it recognizes the risks faced by the insurer in estimating the price of its product; i.e., parameter uncertainty.

Table 1 illustrates the basic steps involved in the calculation of increased limits factors (ILFs).

TABLE 1

| Policy Limit | Average Severity | ILF without Risk Load | Risk Load | ILF with Risk Load | Percent Risk Load |
|---|---|---|---|---|---|
| $ 25,000 | $ 8,202 | 1.00 | $ 281 | 1.00 | 3.42% |
| 50,000 | 10,660 | 1.30 | 393 | 1.30 | 3.69 |
| 100,000 | 13,124 | 1.60 | 542 | 1.61 | 4.13 |
| 250,000 | 16,255 | 1.98 | 844 | 2.02 | 5.19 |
| 300,000 | 16,854 | 2.05 | 929 | 2.10 | 5.51 |
| 400,000 | 17,780 | 2.17 | 1,087 | 2.22 | 6.11 |
| 500,000 | 18,484 | 2.25 | 1,235 | 2.32 | 6.68 |
| 750,000 | 19,726 | 2.40 | 1,580 | 2.51 | 8.01 |
| 1,000,000 | 20,579 | 2.51 | 1,903 | 2.65 | 9.25 |
| 2,000,000 | 22,543 | 2.75 | 3,094 | 3.02 | 13.72 |

The policy limit refers to the maximum indemnity amount that will be paid for a single accident (or occurrence, in ISO terminology). The average severity is the average occurrence severity when subject to the given policy limit. The increased limits factor without risk load is the average severity at the increased limit divided by the average severity at the basic limit (usually $25,000). The risk load will be calculated on a per occurrence basis. The increased limits factor with risk load is the sum of the average severity and the risk load at the increased limit divided by the corresponding sum at the basic limit. Usually, loss adjustment expenses are included in the increased limits factor calculation but they will be ignored in this paper since they are not at issue here.

This paper will proceed by first developing the underlying economic rationale for the risk load formula. Next comes the description of the insurance risk. The risk load formula will then be derived, followed by considerations to be made by insurers when using the formula. These considerations include excess-of-loss reinsurance.

## 2. THE INCOMPLETENESS OF UTILITY THEORY

The original ISO risk load formula was based on the variance of the insured's losses. One possible economic basis for this formula comes from utility theory.[2] There are (at least) two questions addressed by utility theory that are relevant to insurance markets. The first question is: How much is a person willing to pay for insurance covering an uncertain loss? Utility theory provides an answer to this question by calculating a price so that the utility of insuring is equal to the expected utility of not insuring.[3] This mathematical exercise is usually not relevant in practice since the competitive nature of the insurance market often makes insurance available for less than the insured is willing to pay.

The second question is: How much premium must an insurer receive in order to be persuaded to take on the uncertain liability of an insurance policy? Utility theory provides an answer to this question by calculating a price so that the expected utility of not insuring the additional risk is equal to the expected utility of insuring.[4] This can be less than the price actually charged. The premium charged will be set by competitive market

[2] This is described by Bowers, Gerber, Hickman, Jones and Nesbitt [2]. For an exponential utility function and normal loss distribution, the variance-based risk load can be derived from utility theory (page 11). Exercise 1.10a shows that the variance based risk load can be used as an asymptotic approximation for any loss distribution or utility function.

[3] Bowers, *et al.*, *op. cit.*, Equation 1.3.1.

[4] Bowers, *et al.*, *op. cit.*, Equation 1.3.5.

forces or government regulation. If the chargeable premium is less than the insurer's utility calculation indicates, the insurer will not sell the policy.

Thus it can be seen that utility theory provides an upper and lower bound for the price of an insurance policy based on the risk preferences of the insurer and the insured. The actual price of the insurance policy depends upon market conditions; i.e., the supply and demand for insurance. The new risk load formula improves on the old by taking insurance market conditions into account. However, it should be noted that the supply and demand for insurance is influenced collectively by the attitudes toward risk of the insurers and the insureds.

## 3. THE INSURANCE MARKET

Insurance is a precondition for a great deal of economic activity. Financing for home and automobile ownership is usually contingent on obtaining insurance. Commercial enterprises can be liable for sums that could cripple the business operation. For example, employers are financially responsible for injuries to employees on the work premises and, in most instances, are required to purchase workers compensation insurance. Because insurance is a practical necessity, the demand for insurance might be assumed to be relatively inelastic. However, there is anecdotal evidence of insureds reducing, or even dropping, their coverage during periods of rapid price increases.

Property-casualty insurance companies in the United States number well over 1,000. These companies range from small specialty companies to large multiline companies. Entry to the insurance market is generally easy, and no single company has a dominant share of the market.

Some limitation to the supply of insurance comes from state regulators. They are interested in the solvency of the insurance companies under their jurisdiction and thus require the insurance company to have funds (i.e., surplus) available to pay for any excess of claim payments over collected premium. Surplus requirements usually are a function of the annual premium of the insurance company, although a more refined view holds that the required surplus should be a function of the variability of the total loss payments. James Stone [3], and R. Beard, T. Pentikainen, and E. Pesonen [4] provide some discussion of this view. Recently, the National Association of Insurance Commissioners formed a working group to develop risk-based capital and surplus requirements for insurers [5]. The total surplus provided by investors to the insurance companies (and consequently the total supply of insurance) depends upon the relative profitability of insurance and other investments.

The market structure of insurance, ease of entry into the business, and dependence of supply upon profitability indicate that the supply of insurance should be very elastic. Evidence of this proposition abounds in the several jurisdictions where regulatory price restraints have led to shortages in the voluntary insurance market.

The options available to an insurer in this environment are limited. The insurer has limited control over the prices of its products, because they are determined either by the competition or by government price regulation. The insurer *can* establish goals on how much insurance to write (within limits prescribed by state regulators). A multiline insurer can establish goals on how much insurance to write in each of several lines of insurance.

We summarize the above discussion by making the following assumptions. Admittedly, these assumptions may be somewhat stronger than the above discussion justifies, but it is believed that they are reasonable in light of the goal of deriving a workable risk load formula.[5]

1. The insurance market is competitive and efficient. The risk load cannot be influenced by the actions of a single insurer; i.e., insurers are price-takers, not price-makers.
2. The demand for each line/limit combination is known and fixed. That is, in deciding how much insurance to purchase, people and firms do not consider the cost of insurance.
3. Each insurer can decide how much insurance to write in each line of business and policy limit.
4. Each insurer is an efficient manager of its insurance portfolio. For the purpose of this paper, this means that each insurer will write the line/limit combinations in such a way as to maximize

[5] This paper has not addressed a large segment of financial theory which has been applied to the pricing of insurance policies, and which may have some bearing on the validity of these assumptions. A discussion of these issues is beyond the scope of this paper, but some issues are addressed separately by the author [6].

its total risk load subject to a constraint on the variance of its insurance portfolio.

5. The result of all insurers competing for business, as described above, will be an equilibrium characterized by the supply of insurance equaling the demand for insurance for each line/limit combination.

The fifth assumption requires additional discussion. This assumption should be viewed as an operational one. It was made to provide a useful tool to insurers. One can seriously question if insurance prices have ever been in equilibrium in recent history. If they do reach equilibrium, it is at best short-lived. The underwriting cycle is often presented as evidence of instability in insurance prices.

## 4. THE VARIABILITY OF INSURER LOSSES

We shall use the collective risk model with parameter uncertainty to describe the variability of insurer losses for a given line and policy limit. This model is described by the following algorithm.

1. Select $\chi$ at random from a distribution with mean 1 and variance $c$.
2. Select the occurrence count, $K$, at random from a distribution with mean $\chi \cdot n$ and variance $\chi \cdot n \cdot (1 + d)$.
3. Select $\alpha$ at random from a distribution with mean 1 and variance $a$.
4. Select occurrences, $Z_1, Z_2, \ldots, Z_K$, at random from a distribution with mean $\alpha \cdot \mu$ and variance $\alpha^2 \cdot \sigma^2$.
5. The total loss is given by: $X = \sum_{j=1}^{K} Z_j$

Actuaries have long recognized that a major part of the risk to insurers is that of estimating the cost of the insurance product. The technical term for this estimate of risk is parameter uncertainty. The random variables $\chi$ and $\alpha$ are introduced to model parameter uncertainty for the occurrence count and the occurrence severity distribution, respectively.

The expected occurrence count, $n$, will be used to quantify exposure. It will be very important to specify how the variance of the insurer-loss depends on exposure. Consider the case of a single unit of exposure. If there is no parameter uncertainty, we set

$$\mathrm{Var}[K] = 1 + d. \tag{4.1}$$

If we move to $n$ independent units of exposure, we have

$$\mathrm{Var}[K] = n \cdot (1 + d), \tag{4.2}$$

and[6]

$$\mathrm{Var}[X] = n \cdot (\sigma^2 + \mu^2 \cdot (1 + d)). \tag{4.3}$$

It is important to note that the variance is a linear function of exposure when there is no parameter uncertainty.

When parameter uncertainty is introduced, the variance of the total loss is given by[7]

$$n \cdot u + n^2 \cdot v, \tag{4.4}$$

where

$$u = (\mu^2 \cdot (1 + d) + \sigma^2) \cdot (1 + a), \tag{4.5}$$

and

$$v = \mu^2 \cdot (a + c + a \cdot c). \tag{4.6}$$

When there is parameter uncertainty, the variance is a quadratic function of exposure. In practice, the values of $a$ and $c$ are relatively small and thus $u$ is noticeably larger than $v$. For a small exposure; i.e., small $n$, parameter uncertainty is barely noticeable. However, as the exposure increases, parameter uncertainty becomes increasingly important.

---

[6] A special case of Equation 4.4 when $a = c = 0$.

[7] Demonstrated in Appendix A.

In order to incorporate Assumption 4 of Section 3, one must calculate the variance of the entire insurance portfolio. Use subscripts ranging from 1 to $m$ to identify the parameters (e.g., $n_i$, $c_i$) of the various line/limit combinations. Different values of the subscript may denote completely different lines of insurance, such as commercial auto or products liability, or different policy limits within the same line. The parameters associated with the occurrence count distribution or the parameter uncertainty will be the same for each policy limit within a line of insurance. The occurrence severity distribution will be adjusted for each policy limit.

The variance for the entire portfolio of insurance is given by

$$\mathrm{Var}\left[\sum_{i=1}^{m} X_i\right] = \sum_{i=1}^{m} \sum_{j=1}^{m} \mathrm{Cov}[X_i, X_j]. \tag{4.7}$$

There are three cases to consider in the evaluation of $\mathrm{Cov}[X_i, X_j]$.

Case 1. $i = j$

In this case, $\mathrm{Cov}[X_i, X_j] = \mathrm{Var}[X_i]$.

Case 2. $i \neq j$, but the increased limits table of $i$ is the same as that of $j$.

In this case, $X_i$ and $X_j$ will have the same underlying occurrence severity distribution, and the uncertainty random variables $\chi$ and $\alpha$ will be the same for $X_i$ and $X_j$. However, $X_i$ and $X_j$ will be conditionally independent given $\chi$ and $\alpha$.

Case 3. $i \neq j$ and the increased limits table for $i$ is different from that of $j$.

In this case, assume that $X_i$ and $X_j$ are completely independent. Thus, $\mathrm{Cov}[X_i, X_j] = 0$.

The expressions for the covariance become:

$$\mathrm{Cov}[X_i,X_j] = n_i \cdot u_i + n_i^2 \cdot v_{ii}, \tag{4.8}$$
for Case 1; and

$$\mathrm{Cov}[X_i,X_j] = n_i \cdot n_j \cdot v_{ij}, \tag{4.9}$$
for Cases 2 and 3 ($v_{ij} = 0$ for Case 3).

The exact expressions for $\{u_i\}$ and $\{v_{ij}\}$ are given in Appendix A. Suffice it to say in the main text that they are similar to $u$ and $v$ in Equations 4.5 and 4.6.

At this point, it becomes more efficient to express our results in matrix notation. Set the column vector $\mathbf{U} = \{u_i\}$ and the matrix $\mathbf{V} = \{v_{ij}\}$. Also set the column vector $\mathbf{n} = \{n_i\}$. We then have:

$$\mathrm{Var}\left[\sum_{i=1}^{m}\sum_{j=1}^{m} X_{ij}\right] = \sum_{i=1}^{m}\sum_{j=1}^{m} \mathrm{Cov}[X_i,X_j] = \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}$$

Note that if there is no parameter uncertainty (i.e., $a = c = 0$), then $\mathbf{V} = \mathbf{0}$ but $\mathbf{U} \neq \mathbf{0}$. For this reason we say that $\mathbf{U}$ quantifies the process risk and $\mathbf{V}$ quantifies the parameter risk.

## 5. THE COMPETITIVE MARKET EQUILIBRIUM RISK LOAD FORMULA

Let the column vector $\mathbf{R} = \{r_i\}$ be the risk load per expected occurrence. As stated in the assumptions, the insurer attempts to maximize its total risk load, denoted by $\mathbf{n}^{\mathrm{T}} \cdot \mathbf{R}$ in matrix notation, subject to the constraint that the variance of its total insurance portfolio cannot exceed a preset amount, $A^2$. The variance constraint is a function of the size (or surplus) of the insurer and of various other risks (such as investment risk) faced by the insurer. Since the market is competitive, the insurer cannot control $\mathbf{R}$, but it can control $\mathbf{n}$, the amount it insures in each line/limit combination. Mathematically, the problem the insurer faces can be expressed as follows.[8]

[8] The problem posed here is similar to that posed by R. E. Brubaker [7].

Choose **n** to maximize

$$\mathbf{n}^{\mathrm{T}} \cdot \mathbf{R}$$

subject to the constraint that[9]

$$\mathbf{n}^{\mathrm{T}} \cdot \mathbf{U} + \mathbf{n}^{\mathrm{T}} \cdot \mathbf{V} \cdot \mathbf{n} = A^2.$$

It is shown in Appendix B that **n** satisfies the equation[10]

$$\mathbf{n} = \frac{1}{2} \cdot \mathbf{V}^{-1}\left(\frac{\mathbf{R}}{\lambda} - \mathbf{U}\right) \tag{5.1}$$

where $$\lambda = \sqrt{\frac{\mathbf{R}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{R}}{4 \cdot A^2 + \mathbf{U}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{U}}}\,. \tag{5.2}$$

It would be useful to consider some simple examples at this point. Let's consider an insurer who writes four independent lines of insurance with parameters $d$ and $a$ set equal to zero. The remaining parameters are given in the first three columns of Table 2 below. The vector **u** is calculated using Equation 4.5. The matrix **V** is a diagonal matrix with the diagonal elements calculated by Equation 4.6. The variance constraint, $A^2$, was set equal to $10^{14}$ indicating that the insurer has sufficient surplus to cover a loss portfolio with a standard deviation of \$10,000,000. Using Equation 5.2[11] we obtain $\lambda = 1.952 \times 10^{-8}$.

Using the given risk loads for each line, in the column headed by $r$, one can then use Equation 5.1 to calculate the exposure, **n**, for each line to maximize the total risk load obtained by the insurer.

---

[9] Philip E. Heckman in his paper "Some Unifying Remarks on Risk Load" (submitted for publication) has derived an alternative formulation which produces the same result. His formulation has the insurer minimizing the variance, $\mathbf{n}^{\mathrm{T}} \cdot \mathbf{U} + \mathbf{n}^{\mathrm{T}} \cdot \mathbf{V} \cdot \mathbf{n}$, is subject to the constraint that $\mathbf{n}^{\mathrm{T}} \cdot \mathbf{R} = \mathbf{P}$.

[10] Note that the matrix **V** may not have an inverse. If this is the case, interpret $\mathbf{x} = \mathbf{V}^{-1} \cdot \mathbf{y}$ as one of the many solutions of $\mathbf{y} = \mathbf{V} \cdot \mathbf{x}$. This case is treated rigorously in Appendix B.

[11] For a diagonal matrix, **V**,

$$\mathbf{x}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{x} = \sum_{i=1}^{m} x_i^2/v_{ii}.$$

TABLE 2

| $\mu$ | $\sigma$ | $c$ | $u$ | diag $\mathbf{V}$ | $r$ | $n$ |
|---|---|---|---|---|---|---|
| 10,000 | 30,000 | 0.010 | $1.000 \times 10^9$ | $1.000 \times 10^6$ | 250.00 | 5,904 |
| 20,000 | 100,000 | 0.010 | $1.040 \times 10^{10}$ | $4.000 \times 10^6$ | 500.00 | 1,902 |
| 10,000 | 30,000 | 0.030 | $1.000 \times 10^9$ | $3.000 \times 10^6$ | 250.00 | 1,968 |
| 20,000 | 100,000 | 0.030 | $1.040 \times 10^{10}$ | $1.200 \times 10^7$ | 500.00 | 634 |

Suppose, for the sake of discussion, that all insurers are identical to the one described by this example. If the total exposure demanded by all insureds was proportional to the exposure provided by the insurer described by Table 2, the market would be in equilibrium. However, if the total exposure demanded by all insureds was the same for each line of insurance, the market would not be in equilibrium. There would be a surplus of the first line, and a shortage of the last line.

Consider, instead, the case where our insurer is given the risk loads described by Table 3. All other conditions described above are the same. Using Equation 5.2, we obtain $\lambda = 2.017 \times 10^{-8}$. Equation 5.1 then gives the exposures needed to maximize the total risk load. If all insurers are identical and the total exposure demanded by all insureds was equal for each line of insurance, the market would be in equilibrium.

TABLE 3

| $\mu$ | $\sigma$ | $c$ | $u$ | diag $\mathbf{V}$ | $r$ | $n$ |
|---|---|---|---|---|---|---|
| 10,000 | 30,000 | 0.010 | $1.000 \times 10^9$ | $1.000 \times 10^6$ | 90.28 | 1,738 |
| 20,000 | 100,000 | 0.010 | $1.040 \times 10^{10}$ | $4.000 \times 10^6$ | 490.25 | 1,738 |
| 10,000 | 30,000 | 0.030 | $1.000 \times 10^9$ | $3.000 \times 10^6$ | 230.50 | 1,738 |
| 20,000 | 100,000 | 0.030 | $1.040 \times 10^{10}$ | $1.200 \times 10^7$ | 1051.13 | 1,738 |

The above examples illustrate the question to be addressed: *What risk load will result in market equilibrium?*

Assume that insurers 1, 2, . . . , $g$ are seeking to maximize their total risk load by employing the strategy indicated by Equations 5.1 and 5.2. Assume further that **U** and **V** are the same for all insurers, but the $j^{\text{th}}$ insurer has its own vector $\mathbf{n}(j)$ and its own $\lambda_j$. Also make the normative assumptions that: (1) all insurers are participating in all lines; and (2) the risk loads are the same for all insurers. (We will relax these two assumptions later.) We want to find the vector **R** that exists when the market is in equilibrium. Under equilibrium, the total insurance demanded must equal total insurance supplied which is given by:

$$\sum_{j=1}^{g} \mathbf{n}(j) = \frac{1}{2} \cdot \sum_{j=1}^{g} \mathbf{V}^{-1}\left(\frac{\mathbf{R}}{\lambda_j} - \mathbf{U}\right)$$

$$= \frac{1}{2} \cdot \mathbf{V}^{-1} \cdot \mathbf{R} \sum_{j=1}^{g} \frac{1}{\lambda_j} - \frac{g}{2} \cdot \mathbf{V}^{-1} \cdot \mathbf{U}.$$

Define

$$\bar{\lambda} \equiv \frac{g}{\sum_{j=1}^{g} \frac{1}{\lambda_j}} \tag{5.3}$$

and

$$\bar{\mathbf{n}} \equiv \frac{\sum_{j=1}^{g} \mathbf{n}(j)}{g} . \tag{5.4}$$

We then have that

$$\bar{\mathbf{n}} = \frac{1}{2} \cdot \mathbf{V}^{-1}\left(\frac{\mathbf{R}}{\bar{\lambda}} - \mathbf{U}\right) . \tag{5.5}$$

Solving for **R** yields

$$\mathbf{R} = \bar{\lambda} \cdot (\mathbf{U} + 2 \cdot \mathbf{V} \cdot \bar{\mathbf{n}}). \tag{5.6}$$

Some discussion about $\bar{\lambda}$ is in order. $\bar{\lambda}$ is the result, through Equations 5.2 and 5.3, of the variance constraints of all the insurance companies. While the variance constraints may provide a general description of how insurance companies operate, they are not sufficiently explicit to use Equations 5.2 and 5.3. However, it is possible to express $\bar{\lambda}$ in more concrete terms. Multiplying both sides of Equation 5.6 by $\bar{\mathbf{n}}^T$ yields:

$$\bar{\lambda} = \frac{\bar{\mathbf{n}}^T \cdot \mathbf{R}}{\bar{\mathbf{n}}^T \cdot (\mathbf{U} + 2 \cdot \mathbf{V} \cdot \bar{\mathbf{n}})} \equiv \frac{\text{Average Total Risk Load}}{\bar{\mathbf{n}}^T \cdot (\mathbf{U} + 2 \cdot \mathbf{V} \cdot \bar{\mathbf{n}})} \cdot \tag{5.7}$$

The average total risk load can be derived from external considerations such as the overall profitability of the insurance industry.

## 6. INDIVIDUAL INSURER PRICING DECISIONS

Equation 5.6 was derived as a description of insurance market pricing. This section discusses its applicability as a tool for insurers to determine the price at which they will offer insurance.

Recall that Equation 5.6 was derived by making certain normative economic assumptions, namely that: (1) all insurers are participating in all lines/limits; and (2) the risk loads are the same for all insurers. One can argue that these assumptions are appropriate in the long run when the less efficient companies have been weeded out. Large multiline insurers are generally regarded as more efficient users of capital. Also, it is the total price of the product that is subject to competitive pressures. The marketplace might allow an insurer with an expense advantage to charge a greater risk load. But, in the long run, the insurers with an expense advantage should dominate the market.

As sensible as these normative assumptions may seem, they do not describe today's insurance market. Small specialty insurers are common and often successful. Direct writers consistently have held an expense advantage over the agency companies. While direct writers are growing, agency companies are concentrating on niches where they can provide superior service. To use this risk load formula as a pricing tool in today's marketplace, one should investigate what happens when the normative assumptions are relaxed.

First, relax the assumption that all insurers are participating in all lines/limits. Let $\mathbf{n}(j)$ be the exposure vector for the $j^{\text{th}}$ company and let $\mathbf{I}_j$ be a diagonal matrix with the $i^{\text{th}}$ diagonal element equal to 1 or 0 depending on whether or not the $j^{\text{th}}$ company writes insurance in the $i^{\text{th}}$ line/limit. As in the derivation of Equation 5.6, the total insurance demanded must equal the total insurance supplied which is given by

$$\sum_{j=1}^{g} \mathbf{n}(j) \cdot \mathbf{I}_j = \frac{1}{2} \cdot \sum_{j=1}^{g} \mathbf{V}^{-1} \left(\frac{\mathbf{R}}{\lambda_j} - \mathbf{U}\right) \cdot \mathbf{I}_j. \tag{6.1}$$

The effect of the $\mathbf{I}_j$ is to eliminate the $j^{\text{th}}$ company's contribution to the line/limits it does not insure. Multiplying both sides of this equation by $\mathbf{V}$ and reordering some terms yields

$$\mathbf{R} \cdot \left(\sum_{j=1}^{g} \frac{\mathbf{I}_j}{\lambda_j}\right) = \mathbf{U} \cdot \left(\sum_{j=1}^{g} \mathbf{I}_j\right) + 2 \cdot \mathbf{V} \cdot \left(\sum_{j=1}^{g} \mathbf{n}(j) \cdot \mathbf{I}_j\right). \tag{6.2}$$

Let:

$\ell(i)$ be the set of insurers who write line/limit $i$,

$g_i$ = number of insurers who write line/limit $i$,

$$\overline{\lambda}_i = \frac{g_i}{\sum_{j \in \ell(i)} \frac{1}{\lambda_j}},$$

$$\overline{\mathbf{n}}_i = \frac{\sum_{j=1}^{g} \mathbf{n}(j) \cdot \mathbf{I}_j}{g_i}.$$

Then the $i^{\text{th}}$ component of Equation 6.2 can be written in the form

$$\mathbf{R}_i = \overline{\lambda}_i \cdot (\mathbf{U}_i + 2 \cdot (\mathbf{V} \cdot \overline{\mathbf{n}}_i)_i) \tag{6.3}$$

which resembles Equation 5.6, except that $\overline{\lambda}_i$ and $\overline{\mathbf{n}}$ can be different for each line $i$. $\overline{\mathbf{n}}_i$ can be interpreted as the average exposure vector over all companies that write line/limit $i$. The risk load multiplier, $\overline{\lambda}_i$, can be interpreted as the average $\lambda_j$ over all companies who write line/limit $i$. In effect, this means that the risk load multiplier is strongly influenced by competitors.

We now relax the assumption that the risk load for each insurer will be the same for each line/limit. Let $\mathbf{R}(j)$ be the risk load vector for the $j^{\text{th}}$ company. Setting the total insurance demanded equal to the total insurance supplied yields

$$\sum_{j=1}^{g} \mathbf{n}(j) \cdot \mathbf{I}_j = \frac{1}{2} \cdot \sum_{j=1}^{g} \mathbf{V}^{-1} \left( \frac{\mathbf{R}(j)}{\lambda_j} - \mathbf{U} \right) \cdot \mathbf{I}_j. \tag{6.4}$$

Multiplying both sides of this equation by $\mathbf{V}$ and reordering some terms yields

$$\sum_{j=1}^{g} \frac{\mathbf{R}(j) \cdot \mathbf{I}_j}{\lambda_j} = \mathbf{U} \cdot \left( \sum_{j=1}^{g} \mathbf{I}_j \right) + 2 \cdot \mathbf{V} \cdot \left( \sum_{j=1}^{g} \mathbf{n}(j) \cdot \mathbf{I}_j \right). \tag{6.5}$$

Since one cannot move the risk load vector outside the summation sign, a risk load equation with the form of Equation 5.6 or 6.3 is not possible. It is possible, however, for the risk load equation to be applicable to a segment of the line/limit's business. Consider, for example, the case when direct writers have an expense advantage and can command a higher profit. They will write as much insurance as is appropriate (perhaps governed by their variance constraints and Equations 5.1 and 5.2). Those insureds that remain will purchase their policies from agency companies. In effect, the line of insurance is segmented into two separate markets. One segment is serviced by the direct writers and the other segment is serviced by the agency companies. There may, or may not, be a qualitative difference between the two segments.

To summarize, the normative risk load formula given by Equation 5.6 may not be appropriate in all cases because of line specialization and/or segmentation. However, using Equation 5.6 with a risk load multiplier, $\overline{\lambda}$, that can vary by line of insurance may provide a usable risk load formula. The choice of $\bar{\lambda}$ will be influenced by competitive considerations. We will refer to Equation 5.6 as the Competitive Market Equilibrium (CME) risk load.

To date (mid-1991), ISO has filed the CME risk load for Commercial Auto, Premises/Operations General Liability, Products/Completed Operations, and Medical Malpractice. The same risk load multiplier is used for Commercial Auto Liability, Premises/Operations Liability, and Products/Completed Operations Liability. A different risk load multiplier is used for Medical Malpractice. The rationale for this is that, largely, the same companies compete for business in the first three lines but a different set of companies compete for business in the last line. It is likely that many insurers will be selecting their own risk load multipliers for each line of insurance.

## 7. AN ILLUSTRATIVE EXAMPLE

The risk loads in Table 1 were calculated by the CME formula. This section describes the calculations. Additional mathematical details are given in Appendix C. Since this paper was written to illustrate the concepts in the simplest way possible, the example shown below will not be identical to what ISO actually does in its advisory filings, but instead it will be a simpler analog.[12]

ISO publishes 19 separate increased limits tables for its standard commercial liability lines: three for Premises/Operations; three for Products/Completed Operations; and 13 for Commercial Auto. If ISO were to publish 10 increased limits factors for each table, there would be 190 separate line/limit combinations. At first glance it would appear that one has to work with a 190 × 190 matrix, **V**. But, as shown below, that is unnecessary.

[12] There are two simplifications. The first is that this example uses a two-parameter Pareto rather than a five-parameter truncated Pareto. The second is that this example uses a simpler block structure in the matrix, **V**, than is used in ISO filings. The more complicated block structure is necessary because ISO estimates the occurrence severity distribution with countrywide data grouped by increased limits table within line, but does its basic limits ratemaking on statewide (countrywide for Products/Completed Operations) data grouped by line.

If the increased limits ratemaking is done independently by table, $v_{ij} = 0$ when $i$ and $j$ represent different tables. If the subscripts for each table are entered consecutively, the matrix **V** has a block diagonal structure. This block diagonal structure of **V** makes possible a useful simplification. This is best illustrated by way of example. Suppose there are two lines of insurance, each with two policy limits. Equation 5.6 would give:

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{bmatrix} = \bar{\lambda} \cdot \left( \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + 2 \cdot \begin{bmatrix} v_{11} & v_{12} & 0 & 0 \\ v_{21} & v_{22} & 0 & 0 \\ 0 & 0 & v_{33} & v_{34} \\ 0 & 0 & v_{43} & v_{44} \end{bmatrix} \cdot \begin{bmatrix} \bar{n}_1 \\ \bar{n}_2 \\ \bar{n}_3 \\ \bar{n}_4 \end{bmatrix} \right) .$$

This equation produces the same results as:

$$\begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = \bar{\lambda} \cdot \left( \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + 2 \cdot \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \cdot \begin{bmatrix} \bar{n}_1 \\ \bar{n}_2 \end{bmatrix} \right) ,$$

and

$$\begin{bmatrix} r_3 \\ r_4 \end{bmatrix} = \bar{\lambda} \cdot \left( \begin{bmatrix} u_3 \\ u_4 \end{bmatrix} + 2 \cdot \begin{bmatrix} v_{33} & v_{34} \\ v_{43} & v_{44} \end{bmatrix} \cdot \begin{bmatrix} \bar{n}_3 \\ \bar{n}_4 \end{bmatrix} \right) .$$

This example demonstrates how, once $\bar{\lambda}$ is determined, the risk load equation can be applied to a single line of insurance without a detailed consideration of the other lines. The example given below illustrates how the formula works for a single line of insurance, but should be viewed in the above multiline context.

To construct an increased limits table with risk loads, one needs the following information:

1. The occurrence severity distribution, with uncertainty parameter $a$. (In our example, we use a Pareto distribution with cumulative distribution function:

$$S(z) = 1 - \left(\frac{b}{z + b}\right)^q,$$

with $b = 5{,}000$ and $q = 1.1$. For the uncertainty parameter we use $a = .001$.)

2. The parameters $d$ and $c$ of the occurrence count distribution. (Recall that $c$ is used to quantify parameter uncertainty in the count distribution. As illustrative values, we use $d = 0$ and $c = .02$.)

3. The exposure vector, $\bar{\mathbf{n}}$. (In practice, this can be estimated by first dividing the expected number of annual occurrences for a line by the number of insurers writing this line and using an all-industry policy limits distribution to distribute the expected claim count count to policy limit. The $\bar{\mathbf{n}}$ used is in Table 4 below.)

4. The risk load multiplier, $\bar{\lambda}$. (In this example, we used $\bar{\lambda} = 2 \times 10^{-7}$. In practice, this will be selected by individual insurers.)

The occurrence severity and count distributions are used to assemble the vector $\mathbf{U}$ and the matrix $\mathbf{V}$. The details of the calculations are provided in Appendix C. The risk load is then calculated using Equation 5.6 with the process risk vector defined as $\bar{\lambda} \cdot \mathbf{U}$, and the parameter risk vector defined as $\bar{\lambda} \cdot 2 \cdot \mathbf{V} \cdot \bar{\mathbf{n}}$. The results are in Table 4.

TABLE 4

| Policy Limit | Average Severity | ILF without Risk Load | Process Risk | Parameter Risk | ILF with Risk Load | Percent Risk Load | $\bar{n}$ |
|---|---|---|---|---|---|---|---|
| $ 25,000 | $ 8,202 | 1.00 | $ 28 | $253 | 1.00 | 3.42% | 2 |
| 50,000 | 10,660 | 1.30 | 64 | 330 | 1.30 | 3.69 | 2 |
| 100,000 | 13,124 | 1.60 | 135 | 407 | 1.61 | 4.13 | 10 |
| 250,000 | 16,255 | 1.98 | 339 | 505 | 2.02 | 5.19 | 2 |
| 300,000 | 16,854 | 2.05 | 404 | 524 | 2.10 | 5.51 | 24 |
| 400,000 | 17,780 | 2.17 | 533 | 553 | 2.22 | 6.11 | 2 |
| 500,000 | 18,484 | 2.25 | 659 | 575 | 2.32 | 6.68 | 70 |
| 750,000 | 19,726 | 2.40 | 965 | 615 | 2.51 | 8.01 | 8 |
| 1,000,000 | 20,579 | 2.51 | 1,262 | 641 | 2.65 | 9.25 | 70 |
| 2,000,000 | 22,543 | 2.75 | 2,391 | 703 | 3.02 | 13.72 | 10 |

## 8. THE RISK LOAD FOR EXCESS-OF-LOSS REINSURANCE

The conventional method of calculating increased limits factors for excess-of-loss reinsurance has been to subtract the ground-up increased limits factor for the retention point from the increased limits factor for the policy limit. For example, this method of calculating the increased limits factor for the layer between $500,000 and $1,000,000, using Table 4, yields the following:

TABLE 5

LAYERED INCREASED LIMIT FACTOR CALCULATION BY SUBTRACTION METHOD

| Policy Limit | Average Severity | ILF without Risk Load | Process Risk | Parameter Risk | ILF with Risk Load | Percent Risk Load |
|---|---|---|---|---|---|---|
| $ 25,000 | $ 8,202 | 1.00 | $ 28 | $253 | 1.00 | 3.42% |
| 500,000 | 18,484 | 2.25 | 659 | 575 | 2.32 | 6.68 |
| 1,000,000 | 20,579 | 2.51 | 1,262 | 641 | 2.65 | 9.25 |
| Layer $500,000 to $1,000,000 | $ 2,096 | 0.26 | $ 603 | $ 66 | 0.33 | 31.92% |

This method of calculating increased limits factors has the property that the price of a policy where the loss is shared between primary insurer and excess-of-loss reinsurer is the same as the price of a policy where the entire loss is retained by the primary insurer. From an economic point of view, it seems unlikely that the insurance market would supply both these options at the same price. There are two countervailing influences on the price which must be balanced. The first is the additional expense involved in reinsurance, and the second is the sharing of risk. Excess-of-loss reinsurance contracts are common because there is a sizable market segment for which the economic value of risk sharing is greater than the additional expense of reinsurance. The subtraction method of calculating increased limits factors for excess layers does not change layer prices to reflect the economic value of risk sharing when risks are so shared.[13]

The CME risk load applies for excess layers as well as for ground-up coverages. The formula presented in Appendix C has the lower and upper limits of the layer as input. Table 6 gives the result for the layer from $500,000 to $1,000,000.

TABLE 6

LAYERED INCREASED LIMIT FACTOR CALCULATION USING CME RISK LOAD

| Policy Limit | Average Severity | ILF without Risk Load | Process Risk | Parameter Risk | ILF with Risk Load | Percent Risk Load | $\bar{n}'$ | $\bar{n}$ |
|---|---|---|---|---|---|---|---|---|
| $ 25,000 | $ 8,202 | 1.00 | $ 28 | $253 | 1.00 | 3.42% | 2 | 2 |
| 500,000 | 18,484 | 2.25 | 659 | 575 | 2.32 | 6.68 | 70 | 90 |
| 1,000,000 | 20,579 | 2.51 | 1,262 | 641 | 2.65 | 9.25 | 70 | 50 |
| Layer $500,000 to $1,000,000 | $ 2,096 | 0.26 | $ 183 | $ 66 | 0.28 | 11.90% | 0 | 20 |

[13] The subtraction method is usually subject to judgmental revision. The author has found that most knowledgeable reinsurance actuaries will use the subtraction method on increased limits factors without the risk load (which is appropriate), and judgmentally add in their own risk load.

There are two observations that should be made about the CME risk load formula and layering. First, the total process risk load is reduced by layering, but the total parameter risk load remains the same. This is proved by Appendix D.[14] This reduction in the process risk load provides a quantification of the economic value of risk sharing. In the example above, the total process risk load is reduced from \$1,262 to \$842 (= 659 + 183). The final increased limits factor depends upon the total charge for reinsurance. Table 7 shows the increased limits factors after reinsurance for a variety of reinsurance expense charges for our example. If the reinsurance expense charge is less than \$420 per expected occurrence (our unit of exposure), the increased limits factor with reinsurance is less than the increased limits factor without reinsurance, and thus it is more economical to reinsure.

TABLE 7

INCREASED LIMITS FACTORS WITH EXCESS REINSURANCE
PRIMARY LIMIT—\$500,000

| Policy Limit | Average Severity | Process Risk | Parameter Risk | Reinsurance Expense Charge | ILF with Reinsurance |
|---|---|---|---|---|---|
| \$ 25,000 | \$ 8,202 | \$ 28 | \$253 | \$ 0 | 1.00 |
| 1,000,000 | 20,579 | 842 | 641 | 0 | 2.60 |
| 1,000,000 | 20,579 | 842 | 641 | 140 | 2.62 |
| 1,000,000 | 20,579 | 842 | 641 | 280 | 2.63 |
| 1,000,000 | 20,579 | 842 | 641 | 420 | 2.65 |
| 1,000,000 | 20,579 | 842 | 641 | 560 | 2.67 |

This example makes the very important point that the actuary should be aware of his company's reinsurance strategy when setting prices for increased limits.

The second observation has to do with the estimation of $\bar{\mathbf{n}}$. At first glance, it would seem necessary that the distribution of policy limits takes into account all excess-of-loss reinsurance arrangements. For ex-

[14] The result for process risk was originally demonstrated by Miccolis [1].

ample, in Table 6, the 70 units of exposure with a $1,000,000 policy limit could really consist of 50 units with no reinsurance, and 20 units with a primary insurer retention of $500,000 and an excess reinsurance policy covering the layer from $500,000 to $1,000,000. It is demonstrated in Appendix D that the CME risk load will be the same if we: (1) ignore excess reinsurance of the primary insurer; and (2) assume there is no reinsurance exposure in the excess limits. This is illustrated in the final two columns of Table 6.

There is one additional point to be discussed about layering: consistency. Consistency refers to the property that the price of a layer of constant width should not increase, as the initial attachment point increases. For example, the losses in the $250,000 excess of $750,000 layer will be no higher than the losses in the $250,000 excess of $500,000 layer. The consistency property states that the premium for the first layer should be no higher than the premium for the second layer. Since a loss in a higher layer is always less than or equal to a loss in a lower layer of equal width, it has been felt that increased limits factors should be consistent.

"Consistency tests" have historically been applied to increased limits factors using the subtraction method of calculating increased limits factors for layers. The justification for this practice only addresses losses. When consistency tests using the subtraction method have been applied to increased limits factors with risk loads, the consistency test would occasionally fail, and judgmental modifications to the increased limits factors were made.[15]

It is shown in Appendix E that the CME risk load will always produce consistent increased limits premiums.

[15] A discussion of the use of consistency tests is given by Rosenberg [8].

## REFERENCES

[1] Robert S. Miccolis, "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, p. 27.

[2] Newton L. Bowers, Hans U. Gerber, James C. Hickman, Donald A. Jones, and Cecil J. Nesbitt, *Actuarial Mathematics,* Chapter 1.

[3] James M. Stone, "A Theory of Capacity and the Insurance of Catastrophe Risks," (Parts I and II), *Journal of Risk and Insurance,* June, September, 1973.

[4] R. E. Beard, T. Pentikainen, and E. Pesonen, *Risk Theory,* Chapman and Hall, 1984, Chapter 7.

[5] Lisa S. Howard, "NAIC Considers Risk-Based Capital Requirements," *The National Underwriter,* September 17, 1990, p. 3.

[6] Glenn G. Meyers, Discussion of "Risk Loads for Insurers," *(PCAS* LXXVII, 1990) Unpublished Discussion.

[7] R. E. Brubaker, "A Constrained Profit Maximization Model for a Multiline Property/Liability Company," *Total Return Due A Property/Casualty Insurance Company,* CAS Discussion Paper Program, 1979, p. 28.

[8] Sheldon Rosenberg, Discussion of Miccolis: "On the Theory of Increased Limits and Excess-of-Loss Pricing," *PCAS* LXIV, 1977, p. 60.

[9] Anthony A. Ralston, *A First Course in Numerical Analysis,* McGraw-Hill Inc., 1965, p. 97.

## APPENDIX A

## DERIVATION OF VARIANCE FORMULAS

1. Unconditional variance of the occurrence count, $K$:

$$\begin{aligned}\mathrm{Var}[K] &= \mathrm{E}_{\chi}[\mathrm{Var}[K|\chi]] + \mathrm{Var}_{\chi}[\mathrm{E}[K|\chi]] \\ &= \mathrm{E}_{\chi}[\chi \cdot n \cdot (1 + d)] + \mathrm{Var}_{\chi}[\chi \cdot n] \\ &= n \cdot (1 + d) + n^2 \cdot c \qquad \text{(A.1)}\end{aligned}$$

2. Unconditional variance of the total loss, $X$, with parameter uncertainty for the occurrence count but without parameter uncertainty for severity:

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}_{K}[\mathrm{Var}[X|K]] + \mathrm{Var}_{K}[\mathrm{E}[X|K]] \\ &= \mathrm{E}_{K}[K \cdot \sigma^2] + \mathrm{Var}_{K}[K \cdot \mu] \\ &= n \cdot \sigma^2 + n \cdot \mu^2 \cdot (1 + d) + n^2 \cdot \mu^2 \cdot c \\ &\quad \text{(from Eq. A.1)} \qquad \text{(A.2)}\end{aligned}$$

3. Unconditional Variance for the total loss, $X$, with parameter uncertainty:

$$\begin{aligned}\mathrm{Var}[X] &= \mathrm{E}_{\alpha}[\mathrm{Var}[X|\alpha]] + \mathrm{Var}_{\alpha}[\mathrm{E}[X|\alpha]] \\ &= \mathrm{E}_{\alpha}[\alpha^2 \cdot (n \cdot \sigma^2 + n \cdot \mu^2 \cdot (1 + d) + n^2 \cdot \mu^2 \cdot c)] \\ &\quad + \mathrm{Var}_{\alpha}[n \cdot \mu \cdot \alpha] \quad \text{(from Eq. A.2)} \\ &= n \cdot (\mu^2 \cdot (1 + d) + \sigma^2) \cdot (1 + a) + \\ &\quad n^2 \cdot \mu^2 \cdot (a + c + a \cdot c) \\ &= u \cdot n + v \cdot n^2 \qquad \text{(A.3)}\end{aligned}$$

For the remainder of this appendix, replace Step 4 of the description of the collective risk model, in Section 4, with the following statement:

4′. Multiply the scale parameter of the occurrence severity distribution by $\alpha$ and select $Z_1, Z_2, \ldots, Z_K$ at random from the distribution.

This technical modification is necessary to remove the effect of parameter uncertainty on the policy limit. Otherwise it is equivalent to the original Step 4.

4. $\mathrm{Cov}[X_i,X_j]$ with parameter uncertainty:

$$\mathrm{Cov}[X_i,X_j] = \mathrm{E}_{\alpha,\chi}[\mathrm{Cov}[X_i,X_j|\alpha,\chi]] + \mathrm{Cov}_{\alpha,\chi}[\mathrm{E}[X_i|\alpha,\chi],\ \mathrm{E}[X_j|\alpha,\chi]] \tag{A.4}$$

We now evaluate the first term of Equation A.4.

For Cases 2 and 3 ($i \neq j$), $\mathrm{Cov}[X_i,X_j|\alpha,\chi] = 0$, and so $\mathrm{E}_{\alpha,\chi}[\mathrm{Cov}[X_i,X_j|\alpha,\chi]] = 0$.

For Case 1 ($i = j$):

$$\begin{aligned}
\mathrm{E}_{\alpha,\chi}[\mathrm{Cov}[X_i,X_j|\alpha,\chi]] &= \mathrm{E}_{\alpha,\chi}[\mathrm{Var}[X_i|\alpha,\chi]] \\
&= \mathrm{E}_{\alpha,\chi}[\mathrm{E}_K[\mathrm{Var}[X_i|K,\alpha,\chi] + \mathrm{Var}_K[\mathrm{E}[X_i|K,\alpha,\chi]] \\
&= \mathrm{E}_{\alpha,\chi}[\mathrm{E}_K[K \cdot \mathrm{Var}[Z_i|\alpha,\chi]] + \mathrm{E}[Z_i|\alpha]^2 \cdot \mathrm{Var}_K[K|\chi]] \\
&= \mathrm{E}_{\alpha,\chi}[\chi \cdot n_i \cdot \mathrm{Var}[Z_i|\alpha] + \mathrm{E}[Z_i|\alpha]^2 \cdot \chi \cdot n_i \cdot (1 + d)] \\
&= n_i \cdot (\mathrm{E}_\alpha[\mathrm{Var}[Z_i|\alpha]] + \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha]^2] \cdot (1 + d)) \\
&= n_i \cdot (\mathrm{E}_\alpha[\mathrm{E}[Z_i^2|\alpha]] + \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha]^2] \cdot d) && \text{(A.5)} \\
&= n_i \cdot u_i && \text{(A.6)}
\end{aligned}$$

We now evaluate the second term of Equation A.4.

$$\begin{aligned}
&\mathrm{Cov}_{\alpha,\chi}[\mathrm{E}[X_i|\alpha,\chi],\mathrm{E}[X_j|\alpha,\chi]] \\
&= \mathrm{E}_{\alpha,\chi}[\mathrm{E}[X_i|\alpha,\chi] \cdot \mathrm{E}[X_j|\alpha,\chi] - \\
&\quad \mathrm{E}_{\alpha,\chi}[\mathrm{E}[X_i|\alpha,\chi]] \cdot \mathrm{E}_{\alpha,\chi}[\mathrm{E}[X_j|\alpha,\chi]] \\
&= \mathrm{E}_{\alpha,\chi}[\chi \cdot n_i \cdot \mathrm{E}[Z_i|\alpha] \cdot \chi \cdot n_j \cdot \mathrm{E}[Z_j|\alpha]] - \\
&\quad \mathrm{E}_{\alpha,\chi}[\chi \cdot n_i \cdot \mathrm{E}[Z_i|\alpha]] \cdot \mathrm{E}_{\alpha,\chi}[\chi \cdot n_j \cdot \mathrm{E}[Z_j|\alpha]] \\
&= n_i \cdot n_j \cdot ((1 + c) \cdot \\
&\quad \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha] \cdot \mathrm{E}[Z_j|\alpha]] - \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha]] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]]) && \text{(A.7)} \\
&\equiv n_i \cdot n_j \cdot v_{ij} && \text{(A.8)}
\end{aligned}$$

This derivation applies for Cases 1 and 2 (i.e., the increased limits table for $i$ and $j$ is the same). For Case 3, $v_{ij} = 0$.

Combining Equations A.6 and A.8:

$\mathrm{Cov}[X_i,X_j] = n_i \cdot u_i + n_i^2 \cdot v_{ii}$ for Case 1; and

$\mathrm{Cov}[X_i,X_j] = n_i \cdot n_j \cdot v_{ij}$ for Cases 2 and 3 ($v_{ij} = 0$ for Case 3).

where $u_i$ and $v_{ij}$ are given in Equations A.5 and A.7.

## APPENDIX B

## DERIVATION OF THE RISK LOAD FORMULA

Our problem is to choose $\mathbf{n}$ which maximizes

$$\mathbf{n}^{\mathrm{T}} \cdot \mathbf{R}$$

subject to the constraint that

$$A^2 = \mathbf{n}^{\mathrm{T}} \cdot \mathbf{U} + \mathbf{n}^{\mathrm{T}} \cdot \mathbf{V} \cdot \mathbf{n}.$$

This can also be expressed as maximizing

$$\sum_{i=1}^{m} n_i \cdot r_i$$

subject to the constraint that

$$A^2 = \sum_{i=1}^{m} n_i \cdot u_i + \sum_{i=1}^{m} \sum_{j=1}^{m} n_i \cdot n_j \cdot v_{ij} .$$

To solve this, use the method of Lagrange multipliers. Set

$$L = \sum_{i=1}^{m} n_i \cdot r_i + \lambda \cdot \left(A^2 - \sum_{i=1}^{m} n_i \cdot u_i - \sum_{i=1}^{m} \sum_{j=1}^{m} n_i \cdot n_j \cdot v_{ij}\right).$$

By setting $\partial L/\partial \lambda = 0$, we see that

$$A^2 = \mathbf{n}^{\mathrm{T}} \cdot \mathbf{U} + \mathbf{n}^{\mathrm{T}} \cdot \mathbf{V} \cdot \mathbf{n}. \tag{B.1}$$

By setting $\partial L/\partial n_i = 0$ for each $i$, we see that the solution vector $\mathbf{n} = \{n_i\}$ satisfies the equations

$$r_i = \lambda \cdot \left(u_i + 2 \cdot \sum_{j=1}^{m} n_j \cdot v_{ij}\right) \quad \text{for each } i.$$

Expressing this in matrix notation we have that $\mathbf{n}$ is a solution to the equation

$$\mathbf{R} = \lambda \cdot (\mathbf{U} + 2 \cdot \mathbf{V} \cdot \mathbf{n}).$$

At this stage, the derivation will be easier to follow if one assumes that $\mathbf{V}$ is nonsingular. At the end of this appendix, it will be indicated how the equations must be adjusted for the case when $\mathbf{V}$ is singular.

Solving the above equation for $\mathbf{n}$ yields

$$\mathbf{n} = \frac{1}{2} \cdot \mathbf{V}^{-1} \left( \frac{\mathbf{R}}{\lambda} - \mathbf{U} \right). \tag{B.2}$$

Substituting the expression for $\mathbf{n}$ in Equation B.2 into Equation B.1 and solving for $\lambda$ yields, after some algebra,

$$\lambda = \sqrt{\frac{\mathbf{R}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{R}}{4 \cdot A^2 + \mathbf{U}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{U}}} \tag{B.3}$$

$\mathbf{V}$ can be singular. Consider, for example, if the line and the limit for $X_i$ are the same as the line and limit for $X_j$, then $n_i$ and $n_j$ could be any two numbers with the same sum. If $\mathbf{V}$ is singular, Equations B.2 and B.3 must be interpreted and derived differently. We now indicate how to do this.

First consider the case where the equation $\mathbf{V} \cdot \mathbf{r} = \mathbf{R}$ has infinitely many solutions. Let $\mathbf{r}$ be any one of the solutions. Let $\mathbf{K}$ be a matrix whose columns span the linear space of vectors, $\mathbf{x}$, such that $\mathbf{V} \cdot \mathbf{x} = 0$. Then every solution, $\mathbf{y}$, of the equation $\mathbf{V} \cdot \mathbf{y} = \mathbf{R}$ can be written in the form

$$\mathbf{y} = \mathbf{K} \cdot \mathbf{s} + \mathbf{r},$$

where $\mathbf{s}$ is a column vector with dimension equal to the number of rows of $\mathbf{K}$. Similarly, every solution, $\mathbf{z}$, of the equation $\mathbf{V} \cdot \mathbf{z} = \mathbf{U}$ can be written in the form

$$\mathbf{z} = \mathbf{K} \cdot \mathbf{t} + \mathbf{u}.$$

Let the vectors $\mathbf{r}$, $\mathbf{u}$, $\mathbf{s}$, and $\mathbf{t}$ be given. Define

$$\mathbf{V}^{-1} \cdot \mathbf{R} \equiv \mathbf{K} \cdot \mathbf{s} + \mathbf{r},$$

and

$$\mathbf{V}^{-1} \cdot \mathbf{U} \equiv \mathbf{K} \cdot \mathbf{t} + \mathbf{u}.$$

Using the alternative definitions and carefully working through the steps in deriving Equations B.2 and B.3 for the nonsingular case will yield the same identical equations.

Note that the $\mathbf{n}$ in Equation B.2 will depend on the choices of the vectors $\mathbf{r}$, $\mathbf{u}$, $\mathbf{s}$, and $\mathbf{t}$. However, the Lagrange multiplier, $\lambda$, will be the same in all cases since, for all vectors $\mathbf{s}$:

$$\begin{aligned}\mathbf{R}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{R} &= \mathbf{R}^{\mathrm{T}} \cdot (\mathbf{K} \cdot \mathbf{s} + \mathbf{r}) \\ &= (\mathbf{V} \cdot \mathbf{r})^{\mathrm{T}} \cdot (\mathbf{K} \cdot \mathbf{s}) + \mathbf{R}^{\mathrm{T}} \cdot \mathbf{r} \\ &= \mathbf{r}^{\mathrm{T}} \cdot (\mathbf{V} \cdot \mathbf{K}) \cdot \mathbf{s} + \mathbf{R}^{\mathrm{T}} \cdot \mathbf{r} \\ &= \mathbf{R}^{\mathrm{T}} \cdot \mathbf{r}.\end{aligned}$$

Thus, $\mathbf{R}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{R}$ is independent of the particular solution, $\mathbf{y}$, of $\mathbf{V} \cdot \mathbf{y} = \mathbf{R}$. A similar statement can be made about $\mathbf{U}^{\mathrm{T}} \cdot \mathbf{V}^{-1} \cdot \mathbf{U}$. The uniqueness of $\lambda$ follows from Equation B.3.

If $\mathbf{V}$ is singular, it is possible for there to be no solution to the equation $\mathbf{V} \cdot \mathbf{y} = \mathbf{R}$; i.e., the system of equations is inconsistent. Consider, for example, the case where the line and limit for $X_i$ are the same as the line and limit for $X_j$, but $r_i \neq r_j$. In this case, it is clear what to do. If $r_i > r_j$, set $n_j = 0$, since one gets more premium in line $i$ than in line $j$, with the same amount of risk. In general, it will be possible to eliminate various line/limits without reducing $\mathbf{n}^{\mathrm{T}} \cdot \mathbf{R}$ and obtain a consistent set of equations.

Eliminating line/limits can also be appropriate even when $\mathbf{V}$ is nonsingular. It is possible for a (generally small) company to solve Equation B.2 and have negative exposures indicated for certain line/limits. Since an insurer sells insurance rather than buys insurance, this cannot happen. The solution is to eliminate line/limits when negative exposures are indicated.

An actual procedure for eliminating line/limits will not be specified here. However, it is clear that optimal solutions satisfying $n_i \geq 0$ for all $i$, and $A^2 = \mathbf{n}^T \cdot \mathbf{U} + \mathbf{n}^T \cdot \mathbf{V} \cdot \mathbf{n}$ will always exist (a continuous function will always have a maximum on a closed set). The method of Lagrange multipliers determines the optimal solution on the subset of line/limits, $i$, for which $n_i > 0$.

## APPENDIX C

## FORMULAE UNDERLYING THE ILLUSTRATIVE EXAMPLE

Let $f(\alpha)$ be the probability density function for $\alpha$. From Equations A.6 and A.8, it follows that:

$$u_i = \mathrm{E}_\alpha[\mathrm{E}[Z_i^2|\alpha]] + d \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha]^2]$$

$$= \int_0^\infty \mathrm{E}[Z_i^2|\alpha]f(\alpha)\mathrm{d}\alpha + d \cdot \int_0^\infty [\mathrm{E}[Z_i|\alpha]^2]f(\alpha)\mathrm{d}\alpha;$$

$$v_{ij} = (1 + c) \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha] \cdot \mathrm{E}[Z_j|\alpha]] - \mathrm{E}_\alpha[\mathrm{E}[Z_i|\alpha]] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]]$$

$$= (1 + c) \cdot \int_0^\infty \mathrm{E}[Z_i|\alpha] \cdot \mathrm{E}[Z_j|\alpha]f(\alpha)\mathrm{d}\alpha -$$

$$\int_0^\infty \mathrm{E}[Z_i|\alpha]f(\alpha)\mathrm{d}\alpha \cdot \int_0^\infty \mathrm{E}[Z_j|\alpha]f(\alpha)\mathrm{d}\alpha.$$

Let:

$$f(\alpha) = \frac{1}{\sqrt{2\pi a}} \cdot e^{-\frac{(\alpha - 1)^2}{2a}}.$$

The Hermite-Gauss three-point quadrature formula gives the approximations:[16]

$$u_i \approx \frac{1}{6} \cdot \mathrm{E}[Z_i^2|\alpha_1] + \frac{2}{3} \cdot \mathrm{E}[Z_i^2|\alpha_2] + \frac{1}{6} \cdot \mathrm{E}[Z_i^2|\alpha_3]$$

$$+ d \cdot \left(\frac{1}{6} \cdot \mathrm{E}[Z_i|\alpha_1]^2 + \frac{2}{3} \cdot \mathrm{E}[Z_i|\alpha_2]^2 + \frac{1}{6} \cdot \mathrm{E}[Z_i|\alpha_3]^2\right).$$

[16] The standard change of variables was used. See Ralston [9].

$$v_{ij} \approx (1 + c) \cdot \left(\frac{1}{6} \cdot \mathrm{E}[Z_i|\alpha_1] \cdot \mathrm{E}[Z_j|\alpha_1] + \frac{2}{3} \cdot \mathrm{E}[Z_i|\alpha_2] \cdot \mathrm{E}[Z_j|\alpha_2] + \frac{1}{6} \cdot \mathrm{E}[Z_i|\alpha_3] \cdot \mathrm{E}[Z_j|\alpha_3]\right)$$

$$-\left(\frac{1}{6} \cdot \mathrm{E}[Z_i|\alpha_1] + \frac{2}{3} \cdot \mathrm{E}[Z_i|\alpha_2] + \frac{1}{6} \cdot \mathrm{E}[Z_j|\alpha_3]\right)$$

$$\cdot \left(\frac{1}{6} \cdot \mathrm{E}[Z_j|\alpha_1] + \frac{2}{3} \cdot \mathrm{E}[Z_j|\alpha_2] + \frac{1}{6} \cdot \mathrm{E}[Z_j|\alpha_3]\right),$$

where:

$$\alpha_1 = 1 - 1.224745 \cdot \sqrt{2a}; \quad \alpha_2 = 1; \quad \alpha_3 = 1 + 1.224745 \cdot \sqrt{2a}.$$

The occurrence severity distribution used in this paper is the Pareto distribution with c.d.f.:

$$S(z|\alpha) = 1 - \left(\frac{\alpha \cdot b}{z + \alpha \cdot b}\right)^q$$

Let $LL_i$ and $UL_i$ be the respective lower and upper policy limits corresponding to $i$. Then:[17]

$$\mathrm{E}[Z_i|\alpha] = \int_{LL_i}^{UL_i} (z - LL_i) \cdot \mathrm{d}S(z|\alpha) + (UL_i - LL_i) \cdot (1 - S(UL_i|\alpha))$$

$$= \int_{LL_i}^{UL_i} (1 - S(z|\alpha))\mathrm{d}\alpha$$

$$= \frac{(\alpha b)^q}{q - 1}\left(\frac{1}{(LL_i + \alpha b)^{q-1}} - \frac{1}{(UL_i + \alpha b)^{q-1}}\right), \quad q \neq 1,$$

and

[17] The proofs of lemmas E.1 and E.2 in Appendix E may provide some help in evaluating these integrals.

$$E[Z_i^2|\alpha] = \int_{LL_i}^{UL_i} (z - LL_i)^2 \cdot dS(z|\alpha) + (UL_i - LL_i)^2 \cdot (1 - S(UL_i|\alpha))$$

$$= \frac{2(\alpha b)^q}{q-1}\left[\frac{1}{q-2}\left(\frac{1}{(LL_i + \alpha b)^{q-2}} - \frac{1}{(UL_i + \alpha b)^{q-2}}\right) - \frac{UL_i - LL_i}{(UL_i + \alpha b)^{q-1}}\right], q \neq 1,2.$$

# APPENDIX D

## DEMONSTRATION OF RISK REDUCTION BY LAYERING

From Equation 5.6:

$$r_i = \bar{\lambda} \cdot (u_i + 2 \cdot (\mathbf{V} \cdot \bar{\mathbf{n}})_i).$$

Without loss of generality, it can be assumed that $Z_1$, $Z_2$, and $Z_3$ represent the occurrence severities in the layers from $L$ to $H$, $L$ to $M$, and $M$ to $H$, respectively. To demonstrate risk reduction by layering, it must be shown that $r_1 > r_2 + r_3$. This will be done by showing that $u_1 > u_2 + u_3$ and that $(\mathbf{V} \cdot \bar{\mathbf{n}})_1 = (\mathbf{V} \cdot \bar{\mathbf{n}})_2 + (\mathbf{V} \cdot \bar{\mathbf{n}})_3$.

$$\begin{aligned}
u_1 &= \mathrm{E}_\alpha[\mathrm{E}[Z_1^2|\alpha] + \mathrm{E}[Z_1|\alpha]^2 \cdot d] \\
&= \mathrm{E}_\alpha[\mathrm{E}[(Z_2 + Z_3)^2|\alpha] + \mathrm{E}[Z_2 + Z_3|\alpha]^2 \cdot d] \\
&= \mathrm{E}_\alpha\Big(\mathrm{E}[Z_1^2|\alpha] + 2 \cdot \mathrm{E}[Z_2 \cdot Z_3|\alpha] + \mathrm{E}[Z_3^2|\alpha] \\
&\qquad + (\mathrm{E}[Z_2|\alpha]^2 + 2 \cdot \mathrm{E}[Z_2|\alpha] \cdot \mathrm{E}[Z_3|\alpha] + \mathrm{E}[Z_3|\alpha]^2) \cdot d\Big) \\
&> \mathrm{E}_\alpha\Big(\mathrm{E}[Z_2^2|\alpha] + \mathrm{E}[Z_3^2|\alpha] + (\mathrm{E}[Z_2|\alpha]^2 + \mathrm{E}[Z_3|\alpha]^2) \cdot d\Big) \\
&= u_2 + u_3.
\end{aligned} \tag{D.1}$$

$$\begin{aligned}
v_{1j} &= (1 + c) \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_1|\alpha] \cdot \mathrm{E}[Z_j|\alpha]] \\
&\qquad - \mathrm{E}_\alpha[\mathrm{E}[Z_1|\alpha]] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]] \\
&= (1 + c) \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_2 + Z_3|\alpha] \cdot \mathrm{E}[Z_j|\alpha]] \\
&\qquad - \mathrm{E}_\alpha[\mathrm{E}[Z_2 + Z_3|\alpha]] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]] \\
&= v_{2j} + v_{3j}.
\end{aligned} \tag{D.2}$$

It then follows that:

$$(\mathbf{V} \cdot \bar{\mathbf{n}})_1 = (\mathbf{V} \cdot \bar{\mathbf{n}})_2 + (\mathbf{V} \cdot \bar{\mathbf{n}})_3. \tag{D.3}$$

Equation D.1 shows that the total process risk is reduced by layering, while Equation D.3 shows that the total parameter risk remains constant with layering.

Equation D.2 makes possible a simplification in the tabulation of $\bar{\mathbf{n}}$. Let $p$ represent the average exposure for those of whom the upper layer is covered by one company, and the lower layer is covered by another company. Define $\bar{\mathbf{n}}'$ so that $\bar{\mathbf{n}}'_1 = \bar{\mathbf{n}}_1 + p$, $\bar{\mathbf{n}}'_2 = \bar{\mathbf{n}}_2 - p$ and $\bar{\mathbf{n}}'_3 = \bar{\mathbf{n}}_3 - p$. Since $v_{ij} = v_{ji}$, it follows from Equation D.2 that $\mathbf{V} \cdot \bar{\mathbf{n}} = \mathbf{V} \cdot \bar{\mathbf{n}}'$. In effect, this means that one can ignore the effect of excess reinsurance when estimating $\bar{\mathbf{n}}$, since the risk load will be the same as it would be if excess reinsurance were taken into account.

# APPENDIX E

## DEMONSTRATION OF CONSISTENCY

Let $Z$ be a random variable with cumulative distribution function, $S(z)$. Let the layer moment functions be given by:

$$M_1(a,h) = \int_a^{a+h} (z - a) \cdot \mathrm{d}S(z) + h \cdot (1 - S(a + h));$$

$$M_2(a,h) = \int_a^{a+h} (z - a)^2 \cdot \mathrm{d}S(z) + h^2 \cdot (1 - S(a + h)).$$

Lemma E.1. $M_1(a,h)$ is a decreasing function of $a$.

Integration by parts yields:

$$M_1(a,h) = -(z - a) \cdot (1 - S(z))\big|_a^{a+h} + \int_a^{a+h} (1 - S(z)) \cdot \mathrm{d}z + h \cdot (1 - S(a + h))$$

$$= -h \cdot (1 - S(a + h)) + \int_a^{a+h} (1 - S(z)) \cdot \mathrm{d}z + h \cdot (1 - S(a + h))$$

$$= \int_a^{a+h} (1 - S(z)) \cdot \mathrm{d}z.$$

$$\frac{\mathrm{d}M_1(a,h)}{\mathrm{d}a} = (1 - S(a + h)) - (1 - S(a))$$

$$= S(a) - S(a + h) < 0.$$

Thus, $M_1(a,h)$ is a decreasing function of $a$.

Lemma E.2. $M_2(a,h)$ is a decreasing function of $a$.

Integration by parts yields:

$$M_2(a,h) = -(z-a)^2 \cdot (1 - S(z))\big|_a^{a+h} + \int_a^{a+h} (1 - S(z)) \cdot \mathrm{d}z + h^2 \cdot (1 - S(a+h))$$

$$= -h^2 \cdot (1 - S(a+h)) + 2 \cdot \int_a^{a+h} (z-a) \cdot (1 - S(z)) \cdot \mathrm{d}z + h^2 \cdot (1 - S(a+h))$$

$$= 2 \cdot \int_a^{a+h} (z-a) \cdot (1 - S(z)) \cdot \mathrm{d}z$$

$$= 2 \cdot \int_a^{a+h} z \cdot (1 - S(z)) \cdot \mathrm{d}z - 2 \cdot a \cdot M_1(a,h)$$

$$\frac{\mathrm{d}M_2(a,h)}{\mathrm{d}a} = 2 \cdot (a+h) \cdot (1 - S(a+h)) - 2 \cdot a \cdot (1 - S(a)) - 2 \cdot a \cdot \frac{\mathrm{d}M_1(a,h)}{\mathrm{d}a} - 2 \cdot M_1(a,h)$$

$$= 2 \cdot h \cdot (1 - S(a+h)) - 2 \cdot \int_a^{a+h} (1 - S(z)) \cdot \mathrm{d}z$$

$$= 2 \cdot \int_a^{a+h} (S(z) - S(a+h)) \cdot \mathrm{d}z$$

$< 0$ since $S$ is an increasing function.

Thus, $M_2(a,h)$ is a decreasing function of $a$.

We now turn to establishing the consistency of: (1) the expected loss; (2) the process risk; and (3) the parameter risk. Without loss of generality, one can assume that $Z_1$ is the occurrence severity for the layer from $a_1$ to $a_1 + h$ and $Z_2$ is the occurrence severity from the layer from $a_2$ to $a_2 + h$ with $a_1 < a_2$.

1. The consistency of the expected loss:

$$\mathrm{E}[Z_1] = \mathrm{E}_\alpha[M_1(a_1,h|\alpha)] > \mathrm{E}_\alpha[M_1(a_2,h|\alpha)] = \mathrm{E}[Z_2].$$

2. The consistency of process risk:

$$\begin{aligned} u_1 &= \mathrm{E}_\alpha[M_2(a_1,h|\alpha) + d \cdot M_1(a_1,h|\alpha)^2] \\ &> \mathrm{E}_\alpha[M_2(a_2,h|\alpha) + d \cdot M_1(a_2,h|\alpha)^2] \\ &= u_2 \end{aligned}$$

3. The consistency of parameter risk:

$$\begin{aligned} v_{1j} &= (1 + c) \cdot \mathrm{E}_\alpha[M_1(a_1,h|\alpha) \cdot \mathrm{E}[Z_j|\alpha]] \\ &\quad - \mathrm{E}_\alpha[M_1(a_1,h|\alpha)] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]] \\ &> (1 + c) \cdot \mathrm{E}_\alpha[M_1(a_2,h|\alpha) \cdot \mathrm{E}[Z_j|\alpha]] \\ &\quad - \mathrm{E}_\alpha[M_1(a_2,h|\alpha)] \cdot \mathrm{E}_\alpha[\mathrm{E}[Z_j|\alpha]] \\ &= v_{2j}. \end{aligned}$$

It then follows that:

$$(\mathbf{V} \cdot \overline{\mathbf{n}})_1 > (\mathbf{V} \cdot \overline{\mathbf{n}})_2.$$