Volume LXXXII, Part 2 **No. 157**

PROCEEDINGS
November 12, 13, 14, 15, 1995

BALANCING TRANSACTION COSTS AND RISK LOAD IN RISK SHARING ARRANGEMENTS

CLIVE L. KEATINGE

Abstract

In formulating efficient risk sharing arrangements, it is desirable to minimize both transaction costs and the risk load required by the participating insurers. A simple yet realistic model that explicitly incorporates both transaction costs and risk load is put forth in this paper. It is shown that, under very general conditions, the optimal risk sharing arrangement which results is constructed in layers. Remarkably simple expressions are given for the optimal boundaries between layers as well as each participating insurer's share of each layer. Several examples are included that illustrate the application of the model.

1. INTRODUCTION

This paper addresses the related subjects of optimal risk sharing and premium calculation. "Risk sharing" refers to an arrangement among various entities (in an insurance context, usually insureds, insurers, and reinsurers) to share in the payment of losses. "Premium calculation" refers to the process of figuring charges to add to expected losses to obtain premiums for a particular risk sharing arrangement. These charges take into account both the transaction costs (e.g., commission, brokerage, and overhead) and the risk load associated with a risk sharing arrangement. Optimal risk sharing and premium calculation have been discussed quite frequently in the actuarial literature. The primary feature of this paper that distinguishes it from most other treatments of these subjects is the explicit inclusion of transaction costs as an integral part of the model used to derive results.

The problem discussed in this paper is that of finding the risk sharing arrangement that minimizes the combined premium charged by all of the insurers sharing a particular risk.[1] In the model used to address this problem, we assume that each insurer charges a specified percentage of its own expected losses to account for transaction costs and a specified percentage of the variance of its own losses to account for risk load. These percentages may differ by insurer. In general, we expect insurers that tend to take on small amounts of expected losses for each risk (often reinsurers) to have transaction costs that are a larger percentage of their expected losses than insurers that tend to take on large amounts of expected losses for each risk. Likewise, we expect insurers that tend to take on a very large number of risks (often reinsurers) to have risk loads that are a smaller percentage of the variance of their losses than insurers that take on a small number of risks. More will be said about this later.

[1] For convenience, throughout this paper, the term "insurer" will be used to refer to any participant in a risk sharing arrangement. However, all the participants in a risk sharing arrangement need not be insurers. An insured may retain a portion of its own losses, a reinsurer may assume losses through a primary insurer, or the risk sharing could be in a noninsurance context. In the case of an insured retaining a portion of its own losses, although premium would not change hands, the insured would incur a cost in maintaining the additional capital and liquidity necessary to absorb the retained losses.

There are certainly other ways one could account for transaction costs in a risk sharing arrangement, and risk load has been a subject of ongoing debate for many years. The purpose of this paper is not to debate the merits of various methods of handling transaction costs and risk load. The model described in this paper is useful because it is simple enough to yield results that are mathematically tractable yet realistic enough to yield results that provide real insight.

2. THE PROBLEM

We begin with the usual formulation of the collective risk model. Let N denote the number of claims produced by a risk (or portfolio of risks) in a given time period. Let $X_1, X_2, X_3, \ldots$ denote the various claim sizes. We assume N, $X_1, X_2, X_3, \ldots$ to be mutually independent random variables and $X_1, X_2, X_3, \ldots$ to be identically distributed. If $S = X_1 + X_2 + \cdots + X_N$, then:

$$\begin{aligned}
\mathrm{E}[S] &= \mathrm{E}[N] \cdot \mathrm{E}[X], \qquad \text{and} \\
\mathrm{Var}[S] &= \mathrm{E}[N] \cdot \mathrm{Var}[X] + \mathrm{Var}[N] \cdot (\mathrm{E}[X])^2 \\
&= \mathrm{E}[N] \cdot \left\{ \mathrm{E}[X^2] - (\mathrm{E}[X])^2 + \frac{\mathrm{Var}[N]}{\mathrm{E}[N]} \cdot (\mathrm{E}[X])^2 \right\} \\
&= \mathrm{E}[N] \cdot \left\{ \mathrm{E}[X^2] + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1 \right) \cdot (\mathrm{E}[X])^2 \right\}.
\end{aligned}$$

Next, we assume that there are C insurers available to share in the payment of losses. Further, we assume that each insurer pays a predetermined percentage of each claim. These percentages may vary by claim size. Thus, each insurer has associated with it a payment function, $p(x)$, which can vary between 0 and 1, that indicates the percentage of each claim that the insurer will pay.[2] If S_i designates the total losses paid by the ith insurer and

[2]Payment may also be based on the sum of all claims arising from each occurrence. For convenience, the term "claim" will be used throughout this paper.

$p_i(x)$ is the payment function for the ith insurer, then:

$$\mathrm{E}[S_i] = \mathrm{E}[N] \cdot \mathrm{E}[p_i(X) \cdot X], \qquad \text{and}$$

$$\mathrm{Var}[S_i] = \mathrm{E}[N] \cdot \left\{ \mathrm{E}[(p_i(X) \cdot X)^2] + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1 \right) \cdot (\mathrm{E}[p_i(X) \cdot X])^2 \right\}.$$

Let ϕ_i be the percentage of its own expected losses charged by the ith insurer to account for transaction costs, and let ψ_i be the percentage of the variance of its own losses charged by the ith insurer to account for risk load.[3] Then the combined premium charged by all of the insurers sharing the risk is:

$$\begin{aligned} M &= \sum_{i=1}^{C} (\mathrm{E}[S_i] + \phi_i \cdot \mathrm{E}[S_i] + \psi_i \cdot \mathrm{Var}[S_i]) \\ &= \mathrm{E}[S] + \sum_{i=1}^{C} (\phi_i \cdot \mathrm{E}[S_i] + \psi_i \cdot \mathrm{Var}[S_i]). \end{aligned}$$

The problem is to find the payment functions for each of the C insurers that minimize M subject to the constraints that:

$$0 \le p_i(x) \le 1 \qquad \text{and} \qquad \sum_{i=1}^{C} p_i(x) = 1.$$

3. THE SOLUTION

The solution is given here without proof. The proof is provided in the Appendix. First, we assume that the C insurers have been arranged so that the following relation holds:

$$\phi_1 \le \phi_2 \le \cdots \le \phi_C.$$

The solution then involves the familiar concept of layering. The optimal risk sharing arrangement is organized into C layers.

[3] ϕ_i is dimensionless and ψ_i has dimension $\$^{-1}$ (if we are working with dollars). For convenience, $\$^{-1}$ will be omitted throughout this paper.

The first layer (from zero to the first layer boundary) is allocated entirely to the first insurer.

The second layer is allocated to the first two insurers in the following proportions:

$$\text{Insurer 1:} \quad \frac{1/\psi_1}{1/\psi_1 + 1/\psi_2}, \qquad \text{and}$$

$$\text{Insurer 2:} \quad \frac{1/\psi_2}{1/\psi_1 + 1/\psi_2}.$$

Thus, for a claim that penetrates the second (but not the third) layer, the first insurer pays the entire portion of the claim that falls below the first layer boundary and a fraction of the portion above it. The second insurer pays a fraction of the portion of the claim above the first layer boundary.

The third layer is allocated to the first three insurers in the following proportions:

$$\text{Insurer 1:} \quad \frac{1/\psi_1}{1/\psi_1 + 1/\psi_2 + 1/\psi_3},$$

$$\text{Insurer 2:} \quad \frac{1/\psi_2}{1/\psi_1 + 1/\psi_2 + 1/\psi_3}, \qquad \text{and}$$

$$\text{Insurer 3:} \quad \frac{1/\psi_3}{1/\psi_1 + 1/\psi_2 + 1/\psi_3}.$$

One insurer is then added in each successive layer until the top layer has all of the C insurers participating in the following proportions:

$$\text{Insurer } i: \quad \frac{1/\psi_i}{1/\psi_1 + 1/\psi_2 + \cdots + 1/\psi_C}.$$

Thus, for low layers, which contribute much more to expected losses than to variance, only the insurers with the smallest ϕ_is participate. For high layers, where variance is a much more important consideration than expected losses, many insurers participate in order to better reduce the variance. Within a particular layer, the insurers with the smallest ψ_is get the largest shares.[4]

We now address the issue of the location of the layer boundaries. Let the layer boundaries be $l_1 \leq l_2 \leq \cdots \leq l_{C-1}$. Then each l_j is given by the solution of the following equation:

$$l_j + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1\right) \cdot \mathrm{E}[X; l_j] - \sum_{i=1}^{j} \frac{\phi_{j+1} - \phi_i}{2 \cdot \psi_i} = 0$$

where $\mathrm{E}[X; l_j]$ is the expected value of X limited at l_j.[5] The relationship between adjacent l_js can be expressed as follows:

$$(l_j - l_{j-1}) + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1\right) \cdot (\mathrm{E}[X; l_j] - \mathrm{E}[X; l_{j-1}])$$

$$- \frac{\phi_{j+1} - \phi_j}{2} \cdot \sum_{i=1}^{j} \frac{1}{\psi_i} = 0.$$

The first thing to observe about these equations is that the l_js depend on the claim count distribution only through the ratio of the variance to the mean. If this ratio is 1, as it is with the Poisson distribution, the l_js are independent of the claim size distribution. If this ratio is less than 1, the more severe the claim size distribution, the higher the l_js will be. If this ratio is greater than 1, the more severe the claim size distribution, the lower the l_js will be.

The second of the above equations shows that if ϕ_j is associated with the insurer just added in a given layer and ϕ_{j+1} is

---

[4]This same type of layering arrangement was derived by Bühlmann and Jewell [1] in the context of a model based on exponential utility functions.

[5]This equation can be easily solved for l_j using Newton's method. Note that the derivative of the left side of the equation with respect to l_j is simply:

$$1 + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1\right) \cdot (1 - F(l_j))$$

where $F(x)$ is the cumulative distribution function of X.

associated with the insurer to be added in the layer above, then the greater the difference between them, the greater the width of the former layer will be. In other words, if the insurer to be added in the layer above charges a much greater percentage of its expected losses than the most expensive of the insurers participating on a given layer, a large increment will be required to reach a point where the reduction in variance provided by the addition of the next insurer is worthwhile. On the other hand, if $\phi_j = \phi_{j+1}$, then the width of the layer will be 0, and both insurers will be added at the same time. In the extreme case where all of the ϕ_is are equal to one another, all of the l_js will be 0, so there effectively will be only one layer, with all C of the insurers participating. This reflects the well-known result that if transaction costs do not depend on how a risk is shared, then a quota share arrangement is optimal.

Another noteworthy aspect of the above equations is that a given l_j is only affected by the ψ_is associated with insurers on layers below it. Thus, the ψ_is associated with insurers to be added in higher layers have no effect on the location of a particular l_j. It is also clear that smaller ψ_is will result in higher l_js. In other words, if insurers do not charge large percentages of the variances of their losses, variance reduction is less of a priority than it would otherwise be, and the points at which insurers are added can be higher.

The optimal risk sharing arrangement described above minimizes the combined premium charged by all of the insurers sharing a risk. In order to calculate each insurer's premium, expressions are needed for $\mathrm{E}[p_i(X) \cdot X]$ and $\mathrm{E}[(p_i(X) \cdot X)^2]$, the first and second moments of each insurer's own claim payment distribution. Since the optimal risk sharing arrangement is constructed in layers, the needed expressions are as follows:

$$\mathrm{E}[p_i(X) \cdot X] = \sum_{j=i}^{C} r_{ij} \int_{l_{j-1}}^{l_j} (1 - F(x))\,dx, \qquad \text{and}$$

$$\mathrm{E}[(p_i(X)\cdot X)^2] = \sum_{j=i}^{C}\int_{l_{j-1}}^{l_j}\left[\sum_{m=i}^{j-1} r_{im}\cdot(l_m - l_{m-1}) + r_{ij}\cdot(x - l_{j-1})\right]^2 f(x)dx,$$

where:

$l_0 = 0$ and $l_C = \infty$,

$f(x)$ = probability density function of X, and

r_{ij} = ith insurer's share of the jth layer as defined above.

If claims are censored by a policy limit, a term must be added to the expression for the second moment to take into account the spike of probability at the policy limit. However, the equation used to calculate the l_js is not affected by a policy limit. Insurers that participate only on layers that fall completely above a policy limit are effectively not needed in the optimal risk sharing arrangement.

4. EXAMPLES

The application of the results presented in the previous section will be illustrated with several examples. The claim size distributions used in the examples are Mixed Paretos.[6] Each distribution is the weighted average of two Paretos, one of which has a relatively thick tail and one of which has a relatively thin tail. The density and distribution functions of the Mixed Pareto are as follows:

$$f(x) = \frac{Q1\cdot B1^{Q1}}{(x+B1)^{Q1+1}}\cdot(1-P) + \frac{Q2\cdot B2^{Q2}}{(x+B2)^{Q2+1}}\cdot P, \qquad \text{and}$$

$$F(x) = 1 - \left(\frac{B1}{x+B1}\right)^{Q1}\cdot(1-P) - \left(\frac{B2}{x+B2}\right)^{Q2}\cdot P.$$

[6]Mixed Pareto distributions are used in the Insurance Services Office increased limits procedure.

TABLE 1

KEY STATISTICS FOR OPTIMAL RISK SHARING
TYPICAL GENERAL LIABILITY RISK

$B1$ = 25,000	Layer 1: 0 – 75,677				Policy Limit = 1,000,000	
$B2$ = 5,000	Layer 2: 75,677 – 255,814					
$Q1$ = 1.25	Layer 3: 255,814 – 562,307					
$Q2$ = 3.25	Layer 4: 562,307 – 1,000,000					
P = .80						
Var[N]/E[N] = 2	Insurer 1	Insurer 2	Insurer 3	Insurer 4	Total	
Layer 1 Share	100.0%				100.0%	
Layer 2 Share	45.5	54.5%			100.0	
Layer 3 Share	27.0	32.4	40.6%		100.0	
Layer 4 Share	17.5	21.1	26.3	35.1%	100.0	
						No Risk Sharing
Expected Loss	9,814	2,589	1,057	414	13,874	13,874
ϕ Charge	491	259	158	83	991	694
ψ Charge	315	107	38	8	468	1,401
Total Charge	806	366	196	91	1,459	2,095
Percentage	8.2%	14.1%	18.5%	22.0%	10.5%	15.1%

The charges in the examples are calculated assuming that E[N] = 1. Charges for other values of E[N] can be found simply by multiplying the charges shown by E[N].

In each of the examples, we will assume that there are six insurers available to share the risk. The ϕ_is for the six insurers are .05, .10, .15, .20, .25, and .30, and the ψ_is for the six insurers are $.30 \cdot 10^{-6}$, $.25 \cdot 10^{-6}$, $.20 \cdot 10^{-6}$, $.15 \cdot 10^{-6}$, $.10 \cdot 10^{-6}$, and $.05 \cdot 10^{-6}$, respectively. More will be said later about how these values might be estimated.

Table 1 shows the key statistics for the optimal risk sharing arrangement for what might be considered a typical general liability risk with a $1,000,000 policy limit. Note that only four insurers are required in this case. For comparative purposes, charges are also shown for the case in which there is no risk sharing and In-

TABLE 2

KEY STATISTICS FOR OPTIMAL RISK SHARING LARGE POLICY

$B1$ = 25,000	Layer 1: 0 – 75,677	Policy Limit = 10,000,000
$B2$ = 5,000	Layer 2: 75,677 – 255,814	
$Q1$ = 1.25	Layer 3: 255,814 – 562,307	
$Q2$ = 3.25	Layer 4: 562,307 – 1,036,058	
P = .80	Layer 5: 1,036,058 – 1,760,102	
$\mathrm{Var}[N]/\mathrm{E}[N] = 2$	Layer 6: 1,760,102 – 10,000,000	

	Insurer 1	Insurer 2	Insurer 3	Insurer 4	Insurer 5	Insurer 6	Total	No Risk Sharing
Layer 1 Share	100.0%						100.0%	
Layer 2 Share	45.5	54.5%					100.0	
Layer 3 Share	27.0	32.4	40.6%				100.0	
Layer 4 Share	17.5	21.1	26.3	35.1%			100.0	
Layer 5 Share	11.5	13.8	17.2	23.0	34.5%		100.0	
Layer 6 Share	6.8	8.2	10.2	13.6	20.4	40.8%	100.0	
Expected Loss	10,099	2,932	1,485	986	822	984	17,308	17,308
ϕ Charge	505	293	223	197	206	295	1,719	866
ψ Charge	401	193	124	94	86	109	1,007	8,722
Total Charge	906	486	347	291	292	404	2,726	9,588
Percentage	9.0%	16.6%	23.4%	29.5%	35.5%	41.1%	15.7%	55.4%

surer 1 takes the entire risk. Risk sharing in this example results in a savings of 4.6% of expected losses.

Table 2 shows the statistics for the optimal risk sharing arrangement for a risk identical to that underlying Table 1 except with a policy limit of \$10,000,000. In this case, all six insurers are required, and risk sharing results in a savings of 39.7% of expected losses.[7]

[7]This is an illustration of how the absence of risk sharing in a model can result in very large risk loads at high policy limits. Robbin [2] has discussed the need to consider risk sharing when computing risk loads and has presented a simple model of risk sharing (allowing only quota share arrangements) that incorporates transaction costs (attributed to Klinker).

TABLE 3

Key Statistics for Optimal Risk Sharing Variance Equals Mean

$B1 = 25{,}000$	Layer 1: 0 – 83,333				Policy Limit = 1,000,000	
$B2 = 5{,}000$	Layer 2: 83,333 – 266,667					
$Q1 = 1.25$	Layer 3: 266,667 – 575,000					
$Q2 = 3.25$	Layer 4: 575,000 – 1,000,000					
$P = .80$						
$\text{Var}[N]/\text{E}[N] = 1$						
	Insurer 1	Insurer 2	Insurer 3	Insurer 4	Total	
Layer 1 Share	100.0%				100.0%	
Layer 2 Share	45.5	54.5%			100.0	
Layer 3 Share	27.0	32.4	40.6%		100.0	
Layer 4 Share	17.5	21.1	26.3	35.1%	100.0	
						No Risk Sharing
Expected Loss	9,978	2,477	1,022	397	13,874	13,874
ϕ Charge	499	248	153	80	980	694
ψ Charge	301	103	36	8	448	1,343
Total Charge	800	351	189	88	1,428	2,037
Percentage	8.0%	14.2%	18.5%	22.2%	10.3%	14.7%

Table 3 shows the statistics for the optimal risk sharing arrangement for a risk identical to that underlying Table 1 except with $\text{Var}[N]/\text{E}[N]$ equal to 1. As noted in the previous section, in this case, the layer boundaries are independent of the claim size distribution. The results are similar to those shown in Table 1.

Table 4 shows the statistics for the optimal risk sharing arrangement for a risk identical to that underlying Table 1 except with smaller $Q1$ and $Q2$ parameters. This adjustment thickens the tail of the claim size distribution, thus making risk sharing more important. The layer boundaries change very little, but risk sharing results in a savings of 9.7% of expected losses.

TABLE 4

KEY STATISTICS FOR OPTIMAL RISK SHARING SMALLER $Q1$ AND $Q2$ PARAMETERS

$B1 = 25{,}000$
$B2 = 5{,}000$
$Q1 = 0.75$
$Q2 = 2.75$
$P = .80$
$\mathrm{Var}[N]/\mathrm{E}[N] = 2$

Layer 1: 0 – 72,930
Layer 2: 72,930 – 248,026
Layer 3: 248,026 – 548,936
Layer 4: 548,936 – 1,000,000

Policy Limit = 1,000,000

	Insurer 1	Insurer 2	Insurer 3	Insurer 4	Total	No Risk Sharing
Layer 1 Share	100.0%				100.0%	
Layer 2 Share	45.5	54.5%			100.0	
Layer 3 Share	27.0	32.4	40.6%		100.0	
Layer 4 Share	17.5	21.1	26.3	35.1%	100.0	
Expected Loss	17,352	8,338	4,807	2,397	32,894	32,894
ϕ Charge	868	834	721	479	2,902	1,645
ψ Charge	910	450	199	54	1,613	6,039
Total Charge	1,778	1,284	920	533	4,515	7,684
Percentage	10.2%	15.4%	19.1%	22.2%	13.7%	23.4%

5. AGGREGATION

To this point, we have assumed that risk sharing is done on a claim by claim (or occurrence by occurrence) basis. Each insurer participating on a risk pays a predetermined percentage of each claim, with the percentage depending on the size of the claim. However, the model can also be applied to situations where a number of claims are aggregated together before being allocated to each insurer. If the claims are independent of one another, algorithms are available that may be used to calculate an aggregate distribution from the underlying claim count and claim size distributions, or a simulation technique may be used.

The only change to the model involves the equation for the layer boundaries. If claims are aggregated together over definite time periods, there will be only one "claim" per time period.

Therefore, the variance-to-mean ratio of the claim count distribution must be set at zero, and the equation reduces to:

$$l_j - \mathrm{E}[X;l_j] - \sum_{i=1}^{j} \frac{\phi_{j+1} - \phi_i}{2 \cdot \psi_i} = 0.$$

An advantage of aggregating independent claims together before allocating them to insurers is that claims considered as a group are more predictable than claims considered individually. As a result, more of the expected losses can remain in the lower layers with insurers with lower ϕ_is, thus resulting in a lower combined premium for each risk. The larger the number of claims aggregated, the greater the effect will be.

A lower bound for the combined premium may be easily computed. First, note that the combined charge for transaction costs cannot be lower than the total expected losses multiplied by ϕ_1, which we have assumed to be the smallest of the ϕ_is. Second, as alluded to earlier, if transaction costs are disregarded, a quota share arrangement is optimal, with each of the insurers being allocated relative shares inversely proportional to their ψ_is. Therefore, a lower bound for the combined premium may be obtained by assuming that all of the expected losses are allocated to the lowest layer and all of the variance is allocated to the highest layer. This lower bound is thus:

$$\mathrm{E}[S] + \phi_1 \cdot \mathrm{E}[S] + \sum_{i=1}^{C} \psi_i \cdot \left(\frac{1/\psi_i}{\sum_{k=1}^{C} 1/\psi_k} \right)^2 \cdot \mathrm{Var}[S]$$

$$= \mathrm{E}[S] + \phi_1 \cdot \mathrm{E}[S] + \frac{1}{\sum_{i=1}^{C} 1/\psi_i} \cdot \mathrm{Var}[S].$$

Table 5 shows how this lower bound compares to results without risk sharing and with risk sharing on a claim by claim basis for the risks underlying Tables 1 and 2.

TABLE 5

TOTAL CHARGE AS A PERCENTAGE OF EXPECTED LOSSES

Policy Limit	No Risk Sharing	Claim By Claim	Aggregate Lower Bound
1,000,000	15.1%	10.5%	5.7%
10,000,000	55.4	15.7	8.4

Given that it is possible to lower the combined premium by aggregating claims together before allocating them to each insurer, one might conclude this should always be done. However, this may not always be the best approach. For example, if claims are aggregated together, an insurer participating on a high layer can be affected by a large number of small claims in addition to one large claim, which may not be desirable. In some cases, the overhead associated with aggregate coverage may result in larger transaction costs. The model cannot account for all the practical realities that must be considered. Also, to reduce the combined premium by a significant amount, it may be necessary to aggregate together a very large number of claims.

Finally, it should be noted that many risk sharing contracts exist that aggregate together only the portion of claims in specified layers. For example, in many retrospective rating contracts, losses below a given retention are aggregated together before determining coverage, while the insurer pays the portion of any claim that falls above the retention. As another example, a reinsurer may provide coverage only if the sum of all losses that fall in a given layer exceeds a given aggregate retention, while the ceding insurer retains all losses in this layer below the aggregate retention as well as all claims that fall completely below the layer. The hybrid nature of these contracts makes them difficult to analyze. However, the expression giving a lower bound for the combined premium remains valid.

6. PARAMETER ESTIMATION

This section addresses several points that must be considered if we wish to estimate actual values for the ϕ_is and ψ_is. As shown previously, each l_j is given by the solution of the following equation:

$$l_j + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1\right) \cdot \mathrm{E}[X; l_j] - \sum_{i=1}^{j} \frac{\phi_{j+1} - \phi_i}{2 \cdot \psi_i} = 0.$$

Note that only differences of ϕ_is appear in this equation. If all the ϕ_is were increased by the same amount, the solution to the equation would not change. This reflects the fact that what matters are differences in transaction costs among insurers. If there are some costs (e.g., agents' commissions) that are incurred regardless of how a risk is ultimately shared among insurers, then these costs have no effect on the optimal risk sharing arrangement.

If risk sharing is accomplished through reinsurance, the difference between ϕ_i for a primary insurer and ϕ_i for a reinsurer should reflect the additional transaction costs (e.g., brokerage and overhead) that are incurred as a result of the reinsurance contract. Reinsurers that take on small amounts of expected losses for each risk, such as those that tend to take on high layers, can be expected to have larger ϕ_is than reinsurers that take on large amounts of expected losses for each risk, such as those that tend to take on low layers.

The estimation of ψ_i for an insurer should generally be somehow based on the variability of the insurer's overall results. A simple estimation method is illustrated here. Suppose an insurer estimates that its aggregate loss distribution for the next year (for losses retained) has a mean of \$50,000,000 and a standard deviation of \$5,000,000 (and thus a variance of 25,000,000,000,000 $\$^2$). Suppose further that the insurer decides that it needs half of the standard deviation, or \$2,500,000, as risk load. Then, in order to generate the required amount of risk load, its ψ_i should be calculated as follows:

$$\psi_i = \frac{2{,}500{,}000}{25{,}000{,}000{,}000{,}000} = .10 \cdot 10^{-6}.$$

Since variance is additive for independent risks (or independent blocks of risks), if the insurer uses this ψ_i when calculating the risk load for each of its independent risks (or independent blocks of risks), the required amount of risk load will be generated.

It may be difficult to obtain estimates of ϕ_i and ψ_i for each insurer participating in a risk sharing arrangement. However, if a primary insurer is simply interested in finding the retention below which it should retain 100% of every risk, and if the insurer is willing to assume that $\mathrm{Var}[N]/\mathrm{E}[N]$ is 1, then the equation at the beginning of this section simplifies to:

$$l_1 = \frac{\phi_2 - \phi_1}{2 \cdot \psi_1}.$$

Thus, the insurer needs only an estimate of the additional transaction costs associated with the most inexpensive acceptable reinsurance available and an estimate of its own ψ_i. For example, if $\phi_2 - \phi_1$ is estimated to be .05 and ψ_1 is estimated to be $.10 \cdot 10^{-6}$, then:

$$l_1 = \frac{.05}{2 \cdot .10 \cdot 10^{-6}} = \$250{,}000.$$

The final topic of this section is the effect of trend in claim sizes. When a trend factor T is applied to a claim size distribution, we expect the optimal layer boundaries to be multiplied by T. If we examine the equation at the beginning of this section, we see that this will occur if all of the ψ_is are divided by T. Since the variance of each insurer's losses is multiplied by T^2 when T is applied to a claim size distribution, each insurer's risk load would be multiplied by $(1/T) \cdot T^2 = T$. Since each insurer's expected losses would also be multiplied by T, each insurer's risk load as a percentage of expected losses would remain constant. If nothing else changes, this is the desired result. Thus, whenever a claim size distribution is trended, the ψ_is must be "detrended." The ϕ_is are not affected.

7. THE REINSURANCE MARKET

This section is a brief discussion of a few issues that relate to how the model and its results fit into the actual workings of the reinsurance market, within which most risk sharing among insurers takes place. First, to this point, no mention has been made of allocated loss adjustment expenses (ALAE). If ALAE is included with losses before being allocated to layer, ALAE may be incorporated into the model by using a claim size distribution that is based on the sum of losses and ALAE. If ALAE is allocated to layer in the same proportions as the losses, ALAE is not easily incorporated into the model. However, setting aside any practical considerations, this treatment of ALAE is less efficient from a risk sharing perspective than including ALAE with losses. A clear illustration of this occurs when a ceding insurer incurs a large amount of ALAE in defending a claim on which ultimately no payment is made. In this case, risk sharing does not occur; the ceding insurer pays the entire ALAE amount.

In the examples presented earlier, the insurers with the larger ϕ_is, presumably reinsurers, were also assumed to have the smaller ψ_is. An examination of the model shows that this relationship does not necessarily have to hold. Although large reinsurers may indeed have small ψ_is, there is also room for reinsurers with large ψ_is. They would simply receive smaller shares of the layers on which they participate.

One apparent drawback of the model is that, in order to apply it, we must assume that a set number of insurers are available to participate in a risk sharing arrangement. In reality, numerous insurers and reinsurers may be competing to participate on a particular risk. In the examples presented earlier, we assumed that there was only one insurer with a ϕ_i of .05 available to participate in the risk sharing arrangement. In reality, there may be numerous insurers with ϕ_is of .05 available to participate in the risk sharing arrangement. If the model were strictly applied, all of the insurers would participate, each receiving a relatively small share of the expected losses. However, if this were to occur, it is doubtful that the ϕ_is of these insurers would remain at .05. It is likely that their transaction costs as a percentage of their expected losses would increase.

This illustrates an implicit assumption underlying the model, namely that each insurer's ϕ_i is reasonable given the amount of expected losses taken on by each insurer for a particular risk. Too many insurers participating in a risk sharing arrangement simply drives up the ϕ_is for all of them.[8] At some point, this offsets the reduction in variance achieved by incorporating extra insurers on a risk. If a number of insurers with ϕ_is of .05 were to compete for a particular risk, in reality only one of them would end up participating on the risk. The higher layers would be left to the reinsurers, with larger ϕ_is, that specialize in taking on small amounts of expected losses for each risk.

Thus, for purposes of finding the optimal risk sharing arrangement and its associated premiums, we can assume that a limited number of insurers are available to participate. It would certainly be possible to construct the ϕ_i for each insurer as a function of the expected losses it takes on, instead of as a fixed value. However, the danger in doing this is that the mathematical complications introduced may obscure any additional insight that might be achieved. The allure of the model as it stands is that it captures the essential features of the problem being addressed, yet is still simple enough to yield a tractable solution.

8. CONCLUSION

In formulating risk sharing arrangements, if transaction costs are minimized without accounting for risk load, then the conclusion is that risk sharing should not take place. If risk load is minimized without accounting for transaction costs, then the conclusion is that every risk should be shared pro rata among as many insurers as possible. Clearly, neither conclusion is correct. The model described in this paper provides a workable way to find the risk sharing arrangement that strikes the best balance between the two competing goals of the minimization of transaction costs and the minimization of risk load.

---

[8]An alternate point of view is that additional insurers bring with them additional fixed costs instead of larger ϕ_is. Either way, the effect is the same.

REFERENCES

[1] Bühlmann, H. and W. S. Jewell, "Optimal Risk Exchanges," *ASTIN Bulletin* 10, 1979, pp. 243–262.

[2] Robbin, I., Discussion of Meyers: "The Competitive Market Equilibrium Risk Load Formula for Increased Limits Ratemaking," *PCAS* LXXIX, 1992, pp. 367–384.

APPENDIX

1. PRELIMINARIES

The problem addressed in this appendix is that of finding the set of payment functions $\{p_i(x)\}$ for the C insurers that minimizes:

$$
\begin{aligned}
M &= \mathrm{E}[S] + \sum_{i=1}^{C} (\phi_i \cdot \mathrm{E}[S_i] + \psi_i \cdot \mathrm{Var}[S_i]) \\
&= \mathrm{E}[N] \cdot \left[\mathrm{E}[X] + \sum_{i=1}^{C} \left(\phi_i \cdot \mathrm{E}[p_i(X) \cdot X] \right.\right. \\
&\quad \left.\left. + \psi_i \cdot \left\{ \mathrm{E}[(p_i(X) \cdot X)^2] + \left(\frac{\mathrm{Var}[N]}{\mathrm{E}[N]} - 1 \right) \cdot (\mathrm{E}[p_i(X) \cdot X])^2 \right\} \right) \right],
\end{aligned}
$$

subject to the constraints that:

$$
0 \le p_i(x) \le 1 \qquad \text{and} \qquad \sum_{i=1}^{C} p_i(x) = 1.
$$

There are three basic steps to the proof of the solution, corresponding to the remaining three sections of this appendix. In the first step we show that any set of payment functions minimizing M must satisfy the condition that an insurer which pays a given amount on a claim of a given size pays at least as much on claims of all larger sizes. This implies that the number of insurers participating in the payment of a claim may not decrease (and may very well increase) as the size of the claim increases. In the second step, we use the method of Lagrange multipliers to find a condition that must be satisfied by any set of payment functions minimizing M given that the expected losses allocated to each insurer are fixed at certain amounts. It can then be deduced that the only risk sharing arrangement satisfying both these conditions is a layering arrangement with one insurer added at each successively higher layer. Finally, in the third step we find the layering arrangement minimizing M without restricting the amount of expected losses allocated to each insurer.

Similar reasoning applies regardless of whether the claim size distribution is discrete, continuous, or mixed. However, to make the proof easy to follow, we use a discrete formulation in the first two steps and a continuous formulation in the third step. In the second step, we assume that claims may take on integral values from 1 to ∞ and that each possible value has positive probability. In the third step, we assume that the claim size distribution has a probability density function that is positive everywhere. The assumption of positivity does not restrict the generality of the solution, because any probability or probability density function that vanishes in some places can be approximated by a function that is positive everywhere, yet where the contribution to M from points or intervals that actually have zero probability is arbitrarily small. Thus, with the proviso that the payment functions may take on arbitrary values where the probability or probability density function of the claim size distribution is zero, the solution holds for any claim size distribution with finite mean and variance (which is necessary for the problem to make sense).

2. A FIRST NECESSARY CONDITION

We will now show that if M is at a minimum and x_L and x_R are any two possible claim sizes such that $x_L < x_R$, then $p_i(x_L)x_L \leq p_i(x_R)x_R$ for each of the C insurers. In other words, any set of payment functions minimizing M must satisfy the condition that an insurer which pays a given amount on a claim of a given size pays at least as much on claims of all larger sizes.

Suppose that for some x_L and x_R, $x_L < x_R$ and $p_i(x_L)x_L > p_i(x_R)x_R$ for at least one insurer. Let one of these insurers have index 1 and let $D_1 = p_1(x_L)x_L - p_1(x_R)x_R$. D_is associated with the other insurers may then be selected such that the following conditions are satisfied:

$$D_i = 0, \qquad \text{if} \quad p_i(x_L)x_L \geq p_i(x_R)x_R,$$

$$p_i(x_L)x_L - p_i(x_R)x_R \leq D_i \leq 0, \qquad \text{if} \quad p_i(x_L)x_L < p_i(x_R)x_R,$$

$$\text{and} \qquad D_1 + \sum_{i=2}^{C} D_i = 0.$$

Now let an alternate set of payment functions $\{p_i^*(x)\}$ be defined as follows:

$$p_i^*(x) = p_i(x) - \frac{f(x_R)}{f(x_L) + f(x_R)} \cdot \frac{D_i}{x_L}, \qquad \text{if} \quad x = x_L,$$

$$p_i^*(x) = p_i(x) + \frac{f(x_L)}{f(x_L) + f(x_R)} \cdot \frac{D_i}{x_R}, \qquad \text{if} \quad x = x_R, \qquad \text{and}$$

$$p_i^*(x) = p_i(x), \qquad \text{otherwise,}$$

where $f(x)$ is the probability function of the claim size distribution. Then:

$$\begin{aligned} &p_i^*(x_L)x_L f(x_L) + p_i^*(x_R)x_R f(x_R) \\ &= \left[p_i(x_L) - \frac{f(x_R)}{f(x_L) + f(x_R)} \cdot \frac{D_i}{x_L}\right] x_L f(x_L) \\ &+ \left[p_i(x_R) + \frac{f(x_L)}{f(x_L) + f(x_R)} \cdot \frac{D_i}{x_R}\right] x_R f(x_R) \\ &= p_i(x_L)x_L f(x_L) + p_i(x_R)x_R f(x_R). \end{aligned}$$

Thus, $\mathrm{E}[p_i^*(X)\cdot X] = \mathrm{E}[p_i(X)\cdot X]$. Also:

$$
\begin{aligned}
&(p_i^*(x_L)x_L)^2 f(x_L) + (p_i^*(x_R)x_R)^2 f(x_R)\\
&\quad = \left(\left[p_i(x_L) - \frac{f(x_R)}{f(x_L)+f(x_R)}\cdot\frac{D_i}{x_L}\right]x_L\right)^2 f(x_L)\\
&\quad\quad + \left(\left[p_i(x_R) + \frac{f(x_L)}{f(x_L)+f(x_R)}\cdot\frac{D_i}{x_R}\right]x_R\right)^2 f(x_R)\\
&\quad = (p_i(x_L)x_L)^2 f(x_L) + (p_i(x_R)x_R)^2 f(x_R)\\
&\quad\quad - 2D_i[p_i(x_L)x_L - p_i(x_R)x_R]\cdot\frac{f(x_L)\cdot f(x_R)}{f(x_L)+f(x_R)}\\
&\quad\quad + D_i^2\cdot\frac{f(x_L)\cdot f(x_R)}{f(x_L)+f(x_R)}.
\end{aligned}
$$

Since $D_1 = p_1(x_L)x_L - p_1(x_R)x_R$:

$$
\begin{aligned}
&(p_1^*(x_L)x_L)^2 f(x_L) + (p_1^*(x_R)x_R)^2 f(x_R)\\
&\quad = (p_1(x_L)x_L)^2 f(x_L) + (p_1(x_R)x_R)^2 f(x_R)\\
&\quad\quad - D_1^2\cdot\frac{f(x_L)\cdot f(x_R)}{f(x_L)+f(x_R)}\\
&\quad < (p_1(x_L)x_L)^2 f(x_L) + (p_1(x_R)x_R)^2 f(x_R).
\end{aligned}
$$

For $i \neq 1$, since $p_i(x_L)x_L - p_i(x_R)x_R \leq D_i \leq 0$:

$$
\begin{aligned}
&(p_i^*(x_L)x_L)^2 f(x_L) + (p_i^*(x_R)x_R)^2 f(x_R)\\
&\quad \leq (p_i(x_L)x_L)^2 f(x_L) + (p_i(x_R)x_R)^2 f(x_R)\\
&\quad\quad - D_i^2\cdot\frac{f(x_L)\cdot f(x_R)}{f(x_L)+f(x_R)}\\
&\quad \leq (p_i(x_L)x_L)^2 f(x_L) + (p_i(x_R)x_R)^2 f(x_R).
\end{aligned}
$$

Thus, $\mathrm{E}[(p_1^*(X)\cdot X)^2] < \mathrm{E}[(p_1(X)\cdot X)^2]$ and for $i \neq 1$, $\mathrm{E}[(p_i^*(X)\cdot X)^2] \leq \mathrm{E}[(p_i(X)\cdot X)^2]$.

Therefore, the alternate set of payment functions $\{p_i^*(x)\}$ produces a smaller value of M than that produced by the original set of payment functions. Hence, if M is at a minimum, $p_i(x_L)x_L$ may not be greater than $p_i(x_R)x_R$ for any insurer.

3. A SECOND NECESSARY CONDITION

We will now show that the optimal risk sharing arrangement must be constructed in layers, with one insurer added at each successively higher layer.

To ensure that $0 \le p_i(x) \le 1$, let $p_i(x) = z_i^2(x)$. We will then optimize each $z_i(x)$, which for notational convenience will be written as simply z_i. Also for notational convenience, let $v = \mathrm{Var}[N]/\mathrm{E}[N]$. In the long expression for M, we will drop the leading factor $\mathrm{E}[N]$ and the leading term $\mathrm{E}[X]$ in the brackets since neither one will have an effect on the solution. Thus, we are left with the problem of minimizing:

$$M_1 = \sum_{i=1}^{C}(\phi_i \cdot \mathrm{E}[z_i^2 \cdot X] + \psi_i \cdot \{\mathrm{E}[(z_i^2 \cdot X)^2] + (v-1)\cdot(\mathrm{E}[z_i^2 \cdot X])^2\})$$

subject to the constraint that:

$$\sum_{i=1}^{C} z_i^2 = 1.$$

Now let $Z_i = \mathrm{E}[z_i^2 \cdot X]$. For now, we will assume that the Z_is are fixed. Later, we will find optimal values for the Z_is. Thus, we want to minimize:

$$M_1 = \sum_{i=1}^{C}\left(\phi_i \cdot Z_i + \psi_i \cdot \left\{\sum_{x=1}^{\infty}(z_i^4 x^2 f(x)) + (v-1)\cdot Z_i^2\right\}\right)$$

$$= \sum_{i=1}^{C}(\phi_i \cdot Z_i + \psi_i \cdot (v-1)\cdot Z_i^2) + \sum_{i=1}^{C}\sum_{x=1}^{\infty}\psi_i z_i^4 x^2 f(x)$$

subject to the constraints that:

$$\sum_{i=1}^{C} z_i^2 = 1 \qquad \text{and for each } i, \qquad \sum_{x=1}^{\infty} z_i^2 x f(x) = Z_i.$$

To find the z_is minimizing M_1 for any given values of the Z_is, it is sufficient to minimize:

$$M_2 = \sum_{i=1}^{C} \sum_{x=1}^{\infty} \psi_i z_i^4 x^2 f(x)$$

subject to the above constraints. If any of the Z_is are zero, the corresponding z_is must be identically zero. These Z_is are disregarded in what follows. Because of the constraints on the z_is, we must introduce Lagrange multipliers and consider:

$$M_3 = \sum_{i=1}^{C} \sum_{x=1}^{\infty} \psi_i z_i^4 x^2 f(x) + \sum_{x=1}^{\infty} \left(\lambda(x) \sum_{i=1}^{C} z_i^2 \right) + \sum_{i=1}^{C} \left(\mu_i \sum_{x=1}^{\infty} z_i^2 x f(x) \right).$$

A necessary condition for M_2 to be at a constrained minimum is that there exist a function $\lambda(x)$ (i.e., a separate multiplier for each possible claim size) and C constants μ_i such that, for each z_i and each possible claim size:

$$\frac{\partial M_3}{\partial z_i} = 4\psi_i z_i^3 x^2 f(x) + 2\lambda(x) z_i + 2\mu_i z_i x f(x)$$

$$= 2z_i[2\psi_i z_i^2 x^2 f(x) + \lambda(x) + \mu_i x f(x)] = 0.$$

If $z_i \neq 0$, then:

$$\frac{\lambda(x)}{\psi_i} = -2z_i^2 x^2 f(x) - \frac{\mu_i}{\psi_i} x f(x).$$

For a given claim size, we may sum over the insurers for which $z_i \neq 0$ to get:

$$\lambda(x) \sum_{\substack{k=1 \\ z_k \neq 0}}^{C} \frac{1}{\psi_k} = -2x^2 f(x) - \sum_{\substack{k=1 \\ z_k \neq 0}}^{C} \frac{\mu_k}{\psi_k} x f(x).$$

Solving for $\lambda(x)$ and substituting back yields:

$$\frac{\partial M_3}{\partial z_i} = 2 z_i x f(x) \left[2\psi_i z_i^2 x + \frac{-2x - \sum\limits_{k=1, z_k \neq 0}^{C} \mu_k / \psi_k}{\sum\limits_{k=1, z_k \neq 0}^{C} 1/\psi_k} + \mu_i \right] = 0.$$

If $z_i \neq 0$, then:

$$z_i^2 x = \left[\frac{x + \sum\limits_{k=1, z_k \neq 0}^{C} \mu_k / 2\psi_k}{\psi_i \sum\limits_{k=1, z_k \neq 0}^{C} 1/\psi_k} - \frac{\mu_i}{2\psi_i} \right].$$

From the first necessary condition, we know that if M_2 is at a constrained minimum, an insurer that pays a given amount on a claim of a given size must pay at least as much on claims of all larger sizes. Thus, the number of insurers participating in the payment of a claim may not decrease (and may very well increase) as the size of the claim increases. The difference function of $z_i^2 x$ with respect to x within a range of claim sizes with the same participating insurers is:

$$z_i^2 (x+1) - z_i^2 x = \frac{1/\psi_i}{\sum\limits_{k=1, z_k \neq 0}^{C} 1/\psi_k}.$$

Since this expression is not dependent on x, we may conclude that the optimal risk sharing arrangement must be constructed in layers, with one insurer added at each successively higher layer. Each participating insurer's share of a particular layer is then given by:

$$\frac{1/\psi_i}{\sum\limits_{k=1, z_k \neq 0}^{C} 1/\psi_k}.$$

Recall that we have been assuming that $Z_i = \mathrm{E}[z_i^2 \cdot X]$ is fixed for each insurer. Given the Z_is, we now have enough information to determine in which order the insurers are added and the boundaries between the layers without finding explicit values for the μ_is. The highest layer has all of the insurers participating with shares that have been determined above. The highest layer boundary, l_{C-1}, is determined by moving it down from ∞ until the allocation of expected losses for one insurer, given by its Z_i, has been satisfied. That insurer is then dropped from further participation and the next layer boundary down, l_{C-2}, is determined by moving it down from l_{C-1} until the allocation of expected losses for another insurer has been satisfied. This procedure is continued until all the layer boundaries have been determined.

The risk sharing arrangement described above minimizes M given that the expected losses allocated to each insurer are fixed at certain amounts. It remains to find the risk sharing arrangement that minimizes M without any restrictions on the amount of expected losses allocated to each insurer. To do this we must find the optimal set of Z_is. Each possible set of Z_is is associated with a set of layer boundaries, and vice versa. It is more convenient to focus on finding the optimal set of layer boundaries. The optimal set of Z_is will then directly follow.

4. OPTIMAL LAYER BOUNDARIES

Since we know that the optimal risk sharing arrangement is constructed in layers, we may write M_1 as follows:

$$
\begin{aligned}
M_1 = \sum_{i=1}^{C} \Bigg\{ & \phi_i \cdot \sum_{j=i}^{C} \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} \int_{l_{j-1}}^{l_j} G(x)\,dx \\
& + \psi_i \cdot (\nu - 1) \cdot \left(\sum_{j=i}^{C} \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} \int_{l_{j-1}}^{l_j} G(x)\,dx \right)^2 \\
& + \psi_i \cdot \sum_{j=i}^{C} \int_{l_{j-1}}^{l_j} \left(\sum_{m=i}^{j-1} \frac{1/\psi_i}{\sum_{k=1}^{m} 1/\psi_k} (l_m - l_{m-1}) \right. \\
& \left. + \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} (x - l_{j-1}) \right)^2 f(x)\,dx \Bigg\},
\end{aligned}
$$

where $l_0 = 0$, $l_C = \infty$, and $G(x) = 1 - F(x)$ where $F(x)$ is the cumulative distribution function of X. At this point, we do not know in which order the insurers should be added in successively higher layers. The above expression, with insurers indexed according to the order in which they are added, could apply to any ordering of insurers. Differentiating with respect to a particular

l_j yields:

$$
\begin{aligned}
\frac{\partial M_1}{\partial l_j} = & \sum_{i=1}^{j} \phi_i \cdot \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} \cdot G(l_j) - \sum_{i=1}^{j+1} \phi_i \cdot \frac{1/\psi_i}{\sum_{k=1}^{j+1} 1/\psi_k} \cdot G(l_j) \\
& + \sum_{i=1}^{j} 2 \cdot \psi_i \cdot (\nu - 1) \cdot \left(\sum_{m=i}^{C} \frac{1/\psi_i}{\sum_{k=1}^{m} 1/\psi_k} \int_{l_{m-1}}^{l_m} G(x)\,dx \right) \\
& \cdot \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} \cdot G(l_j) \\
& - \sum_{i=1}^{j+1} 2 \cdot \psi_i \cdot (\nu - 1) \cdot \left(\sum_{m=i}^{C} \frac{1/\psi_i}{\sum_{k=1}^{m} 1/\psi_k} \int_{l_{m-1}}^{l_m} G(x)\,dx \right) \\
& \cdot \frac{1/\psi_i}{\sum_{k=1}^{j+1} 1/\psi_k} \cdot G(l_j) \\
& + \sum_{i=1}^{j} \left(2 \cdot \psi_i \cdot \frac{1/\psi_i}{\sum_{k=1}^{j} 1/\psi_k} \cdot G(l_j) \cdot \sum_{m=i}^{j} \frac{1/\psi_i}{\sum_{k=1}^{m} 1/\psi_k} \cdot (l_m - l_{m-1}) \right) \\
& - \sum_{i=1}^{j+1} \left(2 \cdot \psi_i \cdot \frac{1/\psi_i}{\sum_{k=1}^{j+1} 1/\psi_k} \cdot G(l_j) \cdot \sum_{m=i}^{j} \frac{1/\psi_i}{\sum_{k=1}^{m} 1/\psi_k} \cdot (l_m - l_{m-1}) \right)
\end{aligned}
$$

$$= 2 \cdot G(l_j) \cdot \left[\frac{\sum_{i=1}^{j} (\phi_i / 2\psi_i) + (\nu - 1) \cdot \int_0^{l_j} G(x) dx + l_j}{\sum_{i=1}^{j} 1/\psi_i} \right.$$

$$\left. - \frac{\sum_{i=1}^{j+1} (\phi_i / 2\psi_i) + (\nu - 1) \cdot \int_0^{l_j} G(x) dx + l_j}{\sum_{i=1}^{j+1} 1/\psi_i} \right]$$

$$= 2 \cdot G(l_j) \cdot \frac{1/\psi_{j+1}}{\left(\sum_{i=1}^{j} 1/\psi_i\right)\left(\sum_{i=1}^{j+1} 1/\psi_i\right)}$$

$$\cdot \left[\sum_{i=1}^{j} \frac{\phi_i}{2\psi_i} - \phi_{j+1} \cdot \sum_{i=1}^{j} \frac{1}{2\psi_i} + (\nu - 1) \cdot \int_0^{l_j} G(x) dx + l_j \right].$$

For M_1 to be meaningful, we must have $0 \leq l_1 \leq l_2 \leq \cdots \leq l_{C-1} \leq \infty$. This will be referred to as the admissible region. The first thing to note is that if any of the l_js are near infinity, M_1 will not be at a minimum, since its derivative with respect to this l_j would be positive, thus indicating that M_1 is increasing as l_j approaches infinity.

For M_1 to be at a minimum, the derivative of M_1 with respect to each l_j in the interior of the admissible region must be zero. We will now determine what conditions must be satisfied by any l_js at the boundary of the admissible region when M_1 is at a minimum. An l_j is at the boundary of the admissible region if it is coincident with another l_j or with zero.

First take the case where two or more l_js are coincident with one another at a nonzero point. Let the point of coincidence be called l_s (s will be the index of the first of the l_js which coincide at l_s). Suppose that n of the l_js coincide at l_s. Then, at a minimum, the derivative of M_1 with respect to l_s must be zero. If this were not so, all n of the l_js that coincide at l_s could be either increased or decreased slightly to yield a smaller value of M_1. The derivative of M_1 with respect to l_s is simply the sum of the derivatives with respect to the n l_js that coincide at l_s:

$$\frac{\partial M_1}{\partial l_s} = 2 \cdot G(l_s) \cdot \left[\frac{\sum_{i=1}^{s} (\phi_i / 2\psi_i) + (v-1) \cdot \int_0^{l_s} G(x)dx + l_s}{\sum_{i=1}^{s} 1/\psi_i} - \frac{\sum_{i=1}^{s+n} (\phi_i / 2\psi_i) + (v-1) \cdot \int_0^{l_s} G(x)dx + l_s}{\sum_{i=1}^{s+n} 1/\psi_i} \right]$$

$$= 2 \cdot G(l_s) \cdot \frac{\sum_{i=s+1}^{s+n} 1/\psi_i}{\left(\sum_{i=1}^{s} 1/\psi_i\right)\left(\sum_{i=1}^{s+n} 1/\psi_i\right)} \cdot \left[\sum_{i=1}^{s} \frac{\phi_i}{2\psi_i} - \frac{\sum_{i=s+1}^{s+n} \phi_i/\psi_i}{\sum_{i=s+1}^{s+n} 1/\psi_i} \sum_{i=1}^{s} \frac{1}{2\psi_i} + (v-1) \cdot \int_0^{l_s} G(x)dx + l_s \right].$$

Note that the factor in brackets above is identical to the corresponding factor in the expression for the derivative of M_1 with respect to a single l_j except that ϕ_{j+1} is replaced by a weighted average of ϕ_{j+1}s. If the n ϕ_{j+1}s corresponding to the n l_js are not all equal to one another, then at least one of these ϕ_{j+1}s must be smaller than the weighted average. If these insurers are reordered so that an insurer with a ϕ_{j+1} smaller than the weighted average is placed first, then if the derivative with respect to l_s is zero (which implies that the factor in brackets must be zero), the derivative of M_1 with respect to the corresponding l_j will be greater than zero. This implies that if this first l_j of those coincident at l_s is moved down slightly, a smaller value of M_1 will result. Therefore, we conclude that, if M_1 is at a minimum, two or more l_js may not be coincident with one another at a nonzero point unless their corresponding ϕ_{j+1}s are all equal to one another.

We now move to the case where the first n of the l_js coincide at zero. If $\phi_1, \ldots, \phi_{n+1}$ are not all equal to one another, then at least one of these ϕ_{j+1}s must be greater than or equal to all the others, and strictly greater than at least one of the others. If these insurers are reordered so that this insurer is placed last, then the derivative with respect to the corresponding l_j will be less than zero. This implies that if this last l_j of those coincident at zero is moved up slightly, a smaller value of M_1 will result. Therefore, we conclude that if M_1 is at a minimum, the first n l_js may not coincide at zero unless $\phi_1 = \phi_2 = \cdots = \phi_{n+1}$.

The above arguments imply that, if M_1 is at a minimum, the derivative of M_1 with respect to each l_j must be zero. Therefore, for each of the l_js, the following equation must be satisfied:

$$\sum_{i=1}^{j} \frac{\phi_i}{2\psi_i} - \phi_{j+1} \cdot \sum_{i=1}^{j} \frac{1}{2\psi_i} + (v-1) \cdot \int_0^{l_j} G(x)\,dx + l_j$$

$$= l_j + (v-1) \cdot \int_0^{l_j} G(x)\,dx - \sum_{i=1}^{j} \frac{\phi_{j+1} - \phi_i}{2 \cdot \psi_i} = 0.$$

The relationship between adjacent l_js may be expressed as follows:

$$(l_j - l_{j-1}) + (v-1) \cdot \int_{l_{j-1}}^{l_j} G(x)dx - \frac{\phi_{j+1} - \phi_j}{2} \cdot \sum_{i=1}^{j} \frac{1}{\psi_i} = 0.$$

It is clear from this equation that $l_j - l_{j-1}$ will be positive if and only if ϕ_{j+1} is greater than ϕ_j, and that l_j will be equal to l_{j-1} if and only if ϕ_{j+1} is equal to ϕ_j. Thus, to ensure a solution to these equations in the admissible region, the insurers must be added in an order such that their ϕ_js are nondecreasing. Furthermore, since the order in which insurers with identical ϕ_js are added does not affect the solution, there is only one solution, which we conclude must yield the point at which M_1, and hence M, assumes its minimum value.