A STUDY OF RISK ASSESSMENT

RICHARD G. WOLL

Much attention has recently been directed towards the subject of risk assessment in private passenger automobile insurance.

In 1975, SRI International, a research organization, was commissioned to do a major study of insurance classification, or risk assessment. They defined a measure of its efficiency and developed a procedure to utilize this measure for automobile accident frequencies based on the assumption that individual accident experience was Poisson distributed. Based on this analysis they concluded that current pricing and selection practices in automobile insurance did a poor job of creating homogeneous groups of risks [1].

Shortly after the release of the SRI Report, the Massachusetts State Rating Bureau (SRB) addressed the same subject and concluded that current automobile risk assessment practices were not only ineffective but that their use generated side effects that were detrimental to society. They recommended that traditional actuarial rates, based on expected costs, should be modified on the basis of subjective judgments about what was "fair" or what would contribute to the welfare of society [2].

Even more recently, changing social values and arguments like those above helped to create a situation where an NAIC task force condemned present automobile risk assessment practices and concluded that:

> ". . . sex and marital status are seriously lacking in justification and are subject to strong public opposition, and should therefore be prohibited as classification factors." [3]

The fact that such an essential aspect of insurance has come into question indicates a need for more knowledge and a better understanding of how we can measure class homogeneity. It is the contention of this paper that the SRI procedure is based on an oversimplified model of reality and will understate the effectiveness of any risk assessment system because it assumes that no random or stochastic elements affect an individual's exposure to loss. An alternative model of the loss generating process is suggested and a more general measure of class homogeneity is developed which makes use of individual risk experience and the findings of credibility theory.

RISK ASSESSMENT MODELS

The purpose of risk assessment is to partition a risk population into groups whose members have a similar expectation of loss. This requires the assumption that such groups exist and that it is possible to distinguish them.

The best indication that these assumptions are reasonable is the fact that persistent classification differentials do exist and form the basis for present risk assessment systems. This attests to the fact that insureds differ from each other in *consistent* and *predictable* ways.

In order to study risk assessment, it is necessary to focus on the loss generating process. In this paper, the analysis will concentrate on that process as it affects the frequency of automobile accidents.

INDIVIDUAL RISK MODEL

We begin by assuming that the probability of loss for an individual within any period of time is determined by the nature and quality of that individual's driving experience. We will call the expected number of accidents resulting from any set of circumstances, *exposure*, and will use ϕ_i to denote the exposure for an individual i associated with a particular set of circumstances. We consider ϕ_i to be a function of driving environment, amount of driving, and driver characteristics.

More formally, we designate the function:

$$\phi_i = W(E,A,C)$$

where:

(1) E = Driving environment
(2) A = Amount of driving
(3) C = Driver characteristics

Since the value of ϕ_i is determined by individual circumstances which, in turn, are affected by all the uncertainties of daily life, *we consider ϕ_i itself to be the result of a stochastic process which is independent with respect to time.*

We assume further that the actual number of accidents arising from a particular value of ϕ_i is determined by a Poisson process with a parameter equal to ϕ_i [4]. This means that the conditional distribution of claims for the ith individual is:

$$g(X_i \mid \phi_i) = \frac{(\phi_i)X_i}{X_i!} \quad e^{-\phi_i}$$

where X_i is a random variable denoting the actual number of accidents, given a particular value of ϕ_i, and $g(X_i \mid \phi_i)$ is the conditional probability density function (pdf) of X_i given ϕ_i. If we denote the distribution function of ϕ_i as $V(\phi_i)$, we can write the unconditional pdf of X_i as:

$$g(X_i) = \int_{\phi} \frac{(\phi_i)X_i}{X_i!} e^{-\phi_i} \, dV(\phi_i)$$

It can be seen, given these assumptions, that the actual distribution of accidents for the ith individual is compound-Poisson with moments which can be expressed in terms of the exposure function as follows [5]:

$$E(X_i) = E(\phi_i) \tag{1}$$

$$Var(X_i) = E(\phi_i) + Var(\phi_i) \tag{2}$$

Since the exposure function is independent in time, it is also possible to express the mean and variance of ϕ_i for different time intervals as follows:

$$E_t(\phi_i) = t \times E(\phi_i) \tag{3}$$

$$Var_t(\phi_i) = t \times Var(\phi_i) \tag{4}$$

where t represents the ratio of the time interval of interest to that used to define ϕ_i. The mean and variance of X_i for such a time interval are thus:

$$E_t(X_i) = t \times E(\phi_i) \tag{5}$$

$$Var_t(X_i) = t \times [E(\phi_i) + Var(\phi_i)] \tag{6}$$

GROUP RISK PROCESS

When we consider a group of individuals, we are interested in the unconditional distribution of X which can be thought of as the actual number of accidents happening to an individual selected at random from the group. This requires knowledge about the individual risk process, and the distribution of individual expected losses, the distribution of $E(X_i)$.

We begin by using the random variable M to denote the distribution of expected losses between individuals and define its distribution function as $U(m)$. The function $U(m)$ has been designated the "structure" function and can be thought of as a description of the structure of expected loss differences throughout the given population [6]. It should be evident that the value of M for a particular individual is equal to $E(\phi_i)$, the expectation of the individual's exposure function. That is:

$$M = E(\phi_i)$$

In this context the distribution of X_i, the accident frequency for the i th risk, is conditional on the value of M and we denote its pdf as follows:

$$h(X|m_i) = g(X_i)$$

The unconditional pdf of X is thus:

$$h(X) = \int_m h(X|m)\, dU(m)$$

The mean of this distribution is equal to $E(M)$ and the variance is equal to the variance of the expected accident frequencies, $Var(M)$, plus the expected value of the variance for each individual, $E(Var(X|m_i))$ [7]. Thus:

$$E(X) = E(M) \tag{7}$$

$$Var(X) = Var(M) + E(Var(X|m_i)) \tag{8}$$

$$= Var(M) + E[E(\phi_i) + Var(\phi_i)]$$

$$= Var(M) + E(M) + E[Var(\phi_i)] \tag{9}$$

Thus the unconditional variance of X is equal to the sum of the mean and variance of the structure function plus the average variance of the individual exposure functions.

We can observe the effect of time on the moments of the accident distribution by noting first that it acts as a scaling factor with respect to the moments of the expected loss distribution [8]. That is:

$$E_t(M) = t \times E(M) \tag{10}$$

$$Var_t(M) = t^2 \times Var(M) \tag{11}$$

When we consider the moments of X, the distribution of actual accident frequencies for different time intervals, we get the following:

$$E_t(X) = E_t(M) \tag{12}$$

$$Var_t(X) = Var_t(M) + E[Var_t(X|m_i)] \tag{13}$$

$$= t^2 \times Var(M) + t \times E[E(\phi_i) + Var(\phi_i)]$$

$$= t^2 \times Var(M) + t \times [E(M) + E(Var(\phi_i))] \tag{14}$$

Thus the variance of accident frequencies for a group of individuals is a quadratic function of time with respect to the structure function variance and a linear function of time with respect to the expected variance of individual accident frequencies!

EFFICIENCY STANDARDS

In 1960, R.A. Bailey introduced the idea of evaluating risk assessment systems by comparing the coefficient of variation for classification relativities to the coefficient of variation of the distribution of individual expected losses [9]. L.H. Roberts suggested, in turn, that a ratio of variances resulting in what he called a "coefficient of determination," rather than a ratio of coefficients of variation, might be preferable [10]. Both Bailey and Roberts were interested in what is now termed "class plan efficiency" from the viewpoint of competition.

Sixteen years later, SRI International suggested using the variance measure proposed by Roberts as a way of measuring what percentage of what is ultimately possible has been achieved [11]. It is a measure of how well the system does relative to the ideal situation where the value of M for each individual is known.

It is important to realize that risk assessment represents a partition of the structure function and that the variance of M can be separated into two components related to such a partition:

(a) *Between cell variance* $= BVAR_m$

(b) *Within cell variance* $= WVAR_m$

Thus:

$$Var(M) = BVAR_m + WVAR_m$$

In these terms, the SRI measure can be expressed as:

$$Efficiency = \frac{BVAR_m}{BVAR_m + WVAR_m} = \frac{1}{1 + \dfrac{WVAR_m}{BVAR_m}} \tag{15}$$

SRI International uses the variance produced by the risk assessment system partition to estimate $BVAR_m$. To estimate $Var(M)$, they assume that the distribution of claims for an individual risk, $g(X_i)$ is Poisson and that $U(m)$ is gamma distributed. This in turn leads to the conclusion that:

$$Var(M) = Var(X) - E(X)$$

Thus the SRI procedure consists of measuring classification variance and dividing it by the difference between the estimated mean and variance of the actual claims distribution [12].

Note that the terms $BVAR_m$ and $WVAR_m$ at this point refer to partitions within the structure function. The within class variance term, $WVAR_m$, refers to the average variance of M within the cells of the partition produced by the risk assessment system, while the between class variance term, $BVAR_m$, refers to the variance of expected loss frequencies between the cells. They refer to the variance of expected loss, not actual loss.

The SRI measure is not the only one which can be used for this purpose. Millicent Treloar, a statistical research analyst with the NAII, has noted:

"If efficiency were expressed as:

$$\frac{BVAR_m}{WVAR_m}$$

we would have a measure which increases as the spread of class relativities and class homogeneity increase. We would also have a quantity of known distribution (an F distribution) by which we could make inferences about the extent of spread of class relativities (and homogeneity). Further, this quantity is that which is employed in classic statistics applications to classification problems dating back to R. Fisher (1936).

"It is most desirable to utilize a measure of efficiency which has a known distribution when one desires to make statements of confidence about a particular value." [13]

MEASURING RISK ASSESSMENT EFFICIENCY-AN EXAMPLE

Before proceeding further with this exposition, a simple example may help to clarify what is meant by risk assessment efficiency. Suppose we have a risk population with the following structure function:

$$U(m) = \begin{cases} 10m & m = .01, .02, .03, \ldots, .10 \\ 0 & \text{otherwise} \end{cases}$$

$E(M)$ in this case is .055 and $Var(M)$ is .000825.

We can illustrate this structure as follows:

	Group									
	1	2	3	4	5	6	7	8	9	10
Accident Frequency:	.01	.02	.03	.04	.05	.06	.07	.08	.09	.10

Suppose that we decide to partition this population into two classes and that our first attempt to do so assigns groups 1, 3, 5, 7, and 9 to the first class and the remainder to the other class. The two classes would look as follows:

	Group									
Class 1	1		3		5		7		9	
Class 2		2		4		6		8		10

This is not a very impressive partition, and the statistics show it:

	Weight	Mean	Variance: Within Group	Variance: Between Group	Eff.
Class 1	.5	.050	.000800	XX	XX
Class 2	.5	.060	.000800	XX	XX
Total	1.0	.055	.000800	.000025	3%

We learn more about our population, and succeed in producing a better partition:

	Group									
Class 1	1	2			5		7	8		
Class 2			3	4		6			9	10

The statistics for this group verify the fact that it is better:

	Weight	Mean	Variance: Within Group	Variance: Between Group	Eff.
Class 1	.5	.046	.000744	XX	XX
Class 2	.5	.064	.000744	XX	XX
Total	1.0	.055	.000744	.000081	10%

Continuing our efforts, we come up with a further improvement:

	Group									
Class 1	1	2	3		5		7			
Class 2				4		6		8	9	10

The statistics on this partition are as follows:

	Weight	Mean	Variance: Within Group	Variance: Between Group	Eff.
Class 1	.5	.036	.000464	XX	XX
Class 2	.5	.074	.000464	XX	XX
Total	1.0	.055	.000464	.000361	44%

Finally, one more plan is produced which divides the risk population as follows:

	Group									
Class 1	1	2	3	4	5					
Class 2						6	7	8	9	10

The statistics for this two partition set are quite impressive:

	Weight	Mean	Variance: Within Group	Variance: Between Group	Eff.
Class 1	.5	.030	.000200	XX	XX
Class 2	.5	.080	.000200	XX	XX
Total	1.0	.055	.000200	.000625	76%

This set of partitions provides a qualitative idea of what risk assessment efficiency means. It shows that greater efficiency, given the same number of partitions, generally means a greater spread of expected class relativities. This can be seen if one observes the class relativities which result from the partitions just presented:

	Partition 1	2	3	4
Lower Class	.91	.84	.65	.55
Higher Class	1.09	1.16	1.35	1.45

USING RELATIVITIES

In many situations, the variance of expected loss relativities produced by a particular risk assessment partition is more convenient to calculate than the variance of the actual expected loss estimates themselves. It would be convenient to express *Var (M)* in terms of relativities as well, so that direct comparisons can be made. Expected loss relativities are calculated by dividing the value of the random variable M by $E(M)$ and will be denoted by the symbol R. We can express the variance of R in terms of the variance of M as follows:

$$Var(R) = Var\left(\frac{M}{E(M)}\right) = \frac{Var(M)}{(E(M))^2}$$

We can determine the efficiency of any risk assessment partition by calculating the variance of the cell relativities produced by that system and then dividing by $Var(R)$. This, in turn is the same as multiplying by the squared mean of M divided by the variance of M. This quantity will henceforth be designated by the symbol BK such that:

$$BK = \frac{1}{Var(R)} = \frac{(E(M))^2}{Var(M)} = \frac{t^2 \times (E(M))^2}{t^2 \times Var(M)} + \frac{(E_t(M))^2}{Var_t(M)} \tag{16}$$

Thus BK is independent of time. It can also be seen that:

$$Var(M) = \frac{(E(M))^2}{BK} \tag{17}$$

It should be noted that BK is the inverse of the normalized variance of the structure function. Since we define homogeneity as the degree of similarity in expected losses for the members of any group, BK is a direct measure of the homogeneity of such a group. A high value of BK indicates a homogeneous group while a low value of BK indicates a relatively heterogeneous group.

It is easy to calculate the efficiency of the different partitions shown in the example above when we know BK. Since we know that $E(M) = .055$ and $Var(M) = .000825$, we have:

$$BK = \frac{(.055)^2}{.000825} = 3.67$$

Since BK is 3.67, we can determine the efficiency of these partitions by calculating the variance of the class relativities that they produce and multiplying the result by

3.67. Shown below are the efficiency estimates for each of these partitions calculated in this manner.

	Partition			
	1	2	3	4
(1) Class Relativities				
Lower Class	.91	.84	.65	.55
Higher Class	1.09	1.16	1.35	1.45
(2) Variance of (1)	.0081	.0268	.1193	.2066
(3) Efficiency				
(2) x 3.67 x 100	3%	10%	44%	76%

The value of *BK* can also be calculated within each class wherein it measures the variance between individuals within that class. In this case, it is a direct measure of class homogeneity!

We can observe the improvements in class homogeneity in the example by calculating the average value of *BK* for each class within each partition. It should be noted that since the average value of *BK* is the inverse of the average of the normalized variance for each class, one first obtains the normalized variance for each class by taking the inverse of *BK*. These values are then averaged and the inverse of the result is then taken. This point becomes more intuitive if one notes that if any single class were perfectly homogeneous, the variance in expected losses for members of that class would be zero and *BK* would be infinite. Clearly, a direct average of *BK* itself could lead to absurd results.

Shown below are the average *BK* values for each partition; these values are calculated in the appropriate manner:

		Partition			
	Population	1	2	3	4
BK Value	3.67	3.69	3.75	4.52	7.89
Efficiency	0%	3%	10%	44%	76%

We can see that more efficient partitions produce more homogeneous class cells.

THE SRI MEASURE OF RISK ASSESSMENT EFFICIENCY

In the example given above, the structure function, *U(m)*, was known. In reality, *U(m)* cannot be observed and must be estimated. *Only the mean and variance of M, however, are necessary for measuring risk assessment efficiency and class homogeneity.*

It is possible to estimate the moments of X, the actual claims distribution, by observing actual data. These estimates are of use in estimating the moments of M. It was shown earlier in formula (12) that:

$$E_t(X) = E_t(M)$$

$$Var_t(X) = E_t(M) + Var_t(M) + t \times E[Var(\phi_i)]$$

Thus:

$$Var_t(M) = Var_t(X) - E_t(X) - t \times E[Var(\phi_i)] \tag{18}$$

Both the SRI Report and the Massachusetts Rating Bureau studies assumed that the distribution of X_i was Poisson and thus $Var(X_i) = E(X_i) = E(\phi_i)$. Since $Var(X_i)$ is also equal to $E(\phi_i) + Var(\phi_i)$, this necessarily implies that $Var(\phi_i)$ is equal to zero in all cases. *In other words, these models assume that there are no elements of chance affecting exposure to loss.* This assumption makes it possible to simplify the formula for $Var(M)$ given above, since $E[Var\ \phi_i)]$ is also equal to zero.

Thus:

$$Var_t(M) = Var_t(X) - E_t(X)$$

We can express this result in terms of BK:

$$\begin{aligned} Var_t(M) = \frac{(E_t(M))^2}{BK} &= Var_t(X) - E_t(X) \\ &= Var_t(X) - E_t(M) \end{aligned}$$

Thus:

$$Var_t(X) = E_t(M) + \frac{(E_t(M))^2}{BK}$$

If we look at the SRI terminology, we see that they write $Var_t(X)$ as follows:

$$Var_t(X) = mt + \frac{(mt)^2}{K}$$

where $mt = E_t(M)$. It can be seen that in this case, BK and K are equal [14].

K can be expressed in all cases (not just the Poisson case) as a function of "excess" variance, the difference between the mean and variance of X, as follows:

$$K = \frac{(E_t(X))^2}{Var_t(X) - E_t(X)} \tag{19}$$

since $E_t(M) = E_t(X)$.

Given these assumptions, *the SRI procedure for assessing the efficiency of a risk assessment system is to calculate the observed mean and variance of the distribution of actual losses for the risk population,* $Var_t(X)$, *and then to estimate the variance of expected losses, the variance of the structure function, by subtracting the population mean from the population variance:*

$$Var_t(M) = OV - OM$$

where OM and OV are the observed mean and variance of the actual claim distribution.

The SRI method thus "solves" the problem of measuring $Var_t(M)$ by assuming that $E[Var_t(X|m_i)] = E_t(M) = E_t(X)$ and thus that all of the "excess variance" is due to the variance of expected losses. That is:

$$\begin{aligned} Var_t(M) &= Var_t(X) - E[Var_t(X|m_i)] \\ &= Var_t(X) - E_t(X) \end{aligned}$$

If, in fact, $E[Var_t(X|m_i)]$ *is not equal to* $E_t(X)$ *then the SRI method will not work!*

MASSACHUSSETTS DEVELOPMENTS

In 1977, the State Rating Bureau disregarded the preliminary nature of the SRI conclusions and made them the basis for a severe indictment of current risk assessment practices. They declared that pricing groups of risks according to their expected loss costs was improper and should be prohibited [15]. In doing so they relied heavily on the SRI conclusions about the efficiency of the risk assessment process:

"current risk assessment schemes in automobile insurance resolve only a small fraction of the uncertainty about individual expected losses." [16]

They pointed out, further, that the SRI study claimed that the "fraction explained" was about 30% and observed that:

"with so much of the difference in expected loss among individuals unresolved, heterogeneous classes are unavoidable." [17]

Finally, they threw cold water on the idea that the way to improve the situation is to improve the class plan. They did this on both practical grounds and because they felt that using certain kinds of information in risk assessment might be socially undesirable [18].

A significant part of the SRB's effort in 1977 was a study of Massachusetts data in order to gain some idea of the expected loss variance in each rating class. Data pertaining to collision coverages was published in a paper on merit rating and is included in this paper as Exhibit I [19].

This data is more suitable than the data used by the SRI to examine the process of risk assessment in insurance because:

(1) It is insurance data.
(2) It represents a complete cross section of insurance business.
(3) It shows differences in homogeneity by class.

MERIT RATING

The intent of the exhibit published by the SRB was to show that class and territorial relativities in Massachussetts should be modified because of the impact of merit rating. In making this point, merit rating data had to be generated through a simulation process since no actual data about individual risk experience existed in Massachussetts at that time. I generated the same data through a computer simulation of the following formula:

$$P(X=x) = \int_m m\, g(x|m)\, dU(m)$$

where $U(m)$ is a gamma distribution function and $g(x|m_i)$ is the Poisson probability of having X claims given the parameter m_i. See the Appendix for a further description of the simulation process.

It is worth noting that the pdf:

$$P(M=m_i|x) = \frac{\int\limits_{m_i} P(X=x|m_i)\, dU(m_i)}{\int\limits_{m} P(X=x|m)\, dU(m)}$$

where x is a discrete number of claims, represents the accident likelihood distribution for risks which have had x claims and is also a gamma distribution. This fact has been pointed out by several commentators including the SRI [20].

It is particularly important to note that the ratio:

$$\frac{E(M|x)}{E(M|0)} = \frac{\int m\ P(M=m|x)\ d(m)}{\int m\ P(M=m|0)\ d(m)}$$

or the ratio of expected loss frequencies for risks who have had x losses and those who have had none, given the assumption that the structure function is gamma, is:

$$\alpha(x)/\alpha(0) = 1 + x/K \tag{20}$$

where $\alpha(x)$ is the expected mean for those risks with x claims. In particular, the ratio of expected means for risks with one claim during this interval of time, and those with none, is:

$$\alpha(1)/\alpha(0) = 1 + 1/K \tag{21}$$

which is dependent only upon the coefficient of variation of expected losses for the subpopulation under consideration.

Exhibit II shows a grid of data by class by merit rating category generated by the simulation process mentioned above. Part 3 of Exhibit II shows the ratio of expected means for risks with X claims divided by the expected mean for risks with 0 claims. The following table reproduces these ratios for risks with one claim:

Class	$\alpha(1)/\alpha(0)$
00	1.98
10	1.57
12	1.46
15	1.48
20	1.54
22	1.54
24	1.58
26	1.37
30	1.35
31	2.76
40	1.37
42	1.29
50	1.44

Using these ratios, it is easy to compute *BK* values for each class using the formula:

$$BK = \frac{\alpha(0)}{\alpha(1) - \alpha(0)} \tag{22}$$

The following table shows the *BK* values estimated in this manner compared to those underlying the simulation:

Class	*BK*	Estimate
00	1.03	1.03
10	1.75	1.75
12	2.15	2.16
15	2.06	2.06
20	1.96	1.91
22	1.95	1.90
24	1.77	1.75
26	2.72	2.72
30	2.83	2.83
31	0.58	0.58
40	2.76	2.74
42	3.51	3.51
50	2.28	2.27

These results show that it is possible to use accident history data rather than the SRI assumptions to estimate class homogeneity.

AN ALTERNATIVE RISK ASSESSMENT MODEL

The SRI method for estimating class plan efficiencies carries with it the implication that the purpose of risk assessment is to determine each risk's exact exposure to loss.

If one considers the nature of the events that determine exposure to loss, it seems more reasonable to assume that exposure is only determinable in a stochastic sense. It was stated earlier that exposure to automobile accidents is determined by the following elements:

(1) Driving environment
(2) Amount of driving
(3) Driver characteristics

Each of these elements is affected by the uncertainties of daily life and should be regarded as random in nature. There are differences in exposure expectations between risks—the success of the current risk assessment system is ample evidence of that—but it seems clear that $Var(\phi_i)$, the variance of the exposure function of the individual risk, is likely to be significantly greater than zero and thus the variance of the individual accident distribution, $Var(X_i|\phi_i)$, has to be greater than its mean. This follows from formula (2), the formula for the variance of the individual claims distribution given earlier:

$$VAR(X_i) = E(\phi_i) + Var(\phi_i)$$

In order to estimate the impact that exposure variance might have on the SRI method for estimating risk assessment efficiency, a comparative set of estimates will be calculated, assuming:

(1) $Var(\phi_i) = 0$
(2) $Var(\phi_i) = .0625 \times (E(\phi_i))^2$

For each case, we can calculate the moments of X given these assumptions about $Var(\phi_i)$ and the facts about the structure function used in the example given earlier in this paper. In the example, $E(M)$ was .055 and $Var(M)$ was .000825.

In the first case, using formula (8):

$$\begin{aligned} Var(X) &= Var\ (M) + E[Var(X|m_i)] \\ &= Var(M) + E(M) + E[Var(\phi_i)] \\ &= .000825 + .055 + 0 \\ &= .055825 \end{aligned}$$

Using the SRI method:

$$\begin{aligned} Var(M) &= Var(X) - E(X) \\ &= .055825 - .055 = .000825 \end{aligned}$$

and:

$$\begin{aligned} BK &= \frac{E(M)^2}{Var(M)} \\ &= .055^2 \div .000825 = 3.67 \end{aligned}$$

We can estimate the efficiency of the various partitions in the example, given the SRI assumptions, by multiplying the variance of the class relativities they produce by 3.67.

In the second case, we have:

$$\begin{aligned} Var(X) &= Var(M) + E[Var(X|m_i)] \\ &= Var(M) + E(M) + E[Var(\phi_i)] \end{aligned}$$

Since $Var(\phi_i) = .0625 \times (E(\phi_i))^2$ and $E(\phi_i) = M$ we have:

$$\begin{aligned} E[Var(\phi_i)] &= .0625 \times E(M^2) \\ &= .0625 \times [Var(M) + (E(M))^2] \\ &= .0625 \times (.000825 + .055^2) \end{aligned}$$

Thus:

$$\begin{aligned} Var(X) &= .000825 + .055 + .000241 \\ &= .056066 \end{aligned}$$

Again applying the SRI method, we calculate *Var (M)* and *BK*:

$$\begin{aligned} Var(M) &= .056066 - .055 = .001066 \\ BK &= .055^2 \div .001066 = 2.84 \end{aligned}$$

Since *Var (M)* is really .000825 and *BK* is really 3.67, it can be seen that the use of the SRI method does not provide an accurate picture of the effectiveness of risk assessment. If we were to use the *BK* estimate of 2.84 to evaluate the efficiency of the partitions used in the example, we would be 23% too low!

The fact is that the SRI method is not really an estimate of risk assessment efficiency at all. It is, in fact, an estimate of the lower bound of that efficiency. If $Var\ (\phi_i)$ for *any* risk is greater than zero, then the SRI estimate will be too low.

It is not the SRI measure that fails, but the assumption that it is possible to estimate the variance of expected losses, $Var_t\ (M)$, by subtracting the mean of the actual loss distribution, $E_t\ (X)$, from its variance, $Var_t\ (X)$. What is needed is some other method for estimating $Var_t\ (M)$.

Since the structure function itself cannot be directly observed, any inferences that can be made about its characteristics must come from observation of actual claims experience. We know that for any group:

$$Var_t\ (X) \ = \ Var_t\ (M) \ + \ E[Var_t\ (X|m_i)]$$

or, since we have a partition of the risk population achieved by the expected losses for each member:

$$TVAR_x \ = \ BVAR_x \ + \ WVAR_x$$

where $BVAR_x$ is the variance between risks and $WVAR_x$ is the expected value of the within risk variance.

It is particularly important to avoid confusing the concepts of between variance and within variance as used here with their use in the SRI efficiency measure. The total variance term used above refers to the variance of actual losses, $Var_t\ (X)$, while the total variance term used in the SRI measure refers to the variance of expected losses, $Var_t\ (M)$. The within variance term used above refers to the variance of individual losses while the within variance term used in the SRI measure is the variance in expected losses remaining within each partition created by a risk assessment system. It is interesting to note that $BVAR_x$ taken with respect to the distribution of actual losses is identical to $TVAR_m$ taken with respect to the distribution of expected losses. That is:

$$BVAR_x \ = \ TVAR_m \ = \ Var_t\ (M)$$

Since $BVAR_x \ = \ Var_t\ (M)$, it can also be expressed in terms of BK by using formula (17) as follows:

$$BVAR_x \ = \ \frac{(E_t\ (M))^2}{BK} \tag{23}$$

and can continue to express the variance of X, or $TVAR_x$, in terms of "excess variance," as follows:

$$TVAR_x = Var_t(X) = E_t(M) + \frac{(E_t(M))^2}{K} \tag{24}$$

See formula (19) for the definition of K as a function of "excess variance."

We know that $Var_t(X_i)$ for any individual is either equal to or greater than $E_t(X_i)$. Thus we also know that $WVAR_x$, or $E[Var_t(X_i)]$, is greater than or equal to $E_t(M)$ since:

$$E[Var_t(X_i)] \geqslant E[E_t(X_i)] = E_t(M)$$

We can, therefore, express $WVAR_x$ in terms of "excess variance" as well, using the quantity WK as the index of the degree to which $WVAR_x$ exceeds $E_t(M)$:

$$WK = \frac{(E_t(M))^2}{WVAR_x - E_t(M)}$$

and thus:

$$WVAR_x = E_t(M) + E[Var_i(\phi_i)] = E_t(M) + \frac{(E_t(M))^2}{WK} \tag{25}$$

We can now write:

$$TVAR_x = WVAR_x + BVAR_x$$

$$E_t(M) + \frac{(E_t(M))^2}{K} = \left\{ E_t(M) + \frac{(E_t(M))^2}{WK} \right\} + \left\{ \frac{(E_t(M))^2}{BK} \right\}$$

and thus:

$$\frac{1}{K} = \frac{1}{WK} + \frac{1}{BK}$$

$$K = \frac{WK \times BK}{WK + BK}$$

$$WK = \frac{BK \times K}{BK - K}$$

$$BK = \frac{WK \times K}{WK - K}$$

These formulae provide insight into the limits of both BK and WK. We know that K, WK and BK must all be positive (since we have concluded that $Var_t(X_i) \geqslant E_t(X_i)$ and $Var_t(X) \geqslant E_t(X)$) and we see that as either WK or BK approaches K, the other increases without bound. Thus we conclude that K is a lower bound for both variables, and there is no upper bound. It is interesting to note that when WK increases without bound, $WVAR_x$ becomes equal to $E_t(M)$, and $BVAR_x$ to:

$$\frac{(E_t(M))^2}{K}$$

When these conditions obtain, $g_t(X_i)$ becomes a Poisson distribution.

We see, therefore, that the Poisson case is a limiting case of the class of all compound-Poisson individual risk distributions.

Since $WVAR_x$ is at a minimum when the simple Poisson case obtains, $BVAR_x$ is at a maximum, BK is at a minimum, and estimates of risk assessment efficiency are minimized. *When WK possesses a finite value, estimates based on the simple Poisson assumption will invariably be understated.*

A GAMMA-NEGATIVE BINOMIAL SIMULATION

It was found in studying the Massachussetts data under the Poisson assumptions that claims history data gave a good estimate of BK, the index of population or subpopulation homogeneity. A simulation was run under the assumption that $Var(\phi_i)$ was not equal to zero in order to find out whether it was still possible to use the ratio method to get a good estimate of BK and thus of the variance of expected losses. In the simulation, $g_t(X_i)$ was assumed to be negative binomial with a variance equal to:

$$Var_t[X|E(\phi_i)] = \{t \times E(\phi_i)\} + \{\frac{(t \times E(\phi_i))^2}{10}\}$$

Thus:

$$Var_t(\phi_i) = \frac{(t \times E(\phi_i))^2}{10}$$

The results of this simulation are shown in Exhibit III. The value 10 was chosen for the denominator of the second term in the above equation because it seemed to provide results that were reasonably similar to those achieved in the Poisson simulation but which differed enough to provide a reasonable picture of how exposure variance might affect the observable characteristics of the risk population.

The results of this simulation compared with the gamma-Poisson case are as follows:

(1) *The number of risks with 0, 1, 2, . . . claims in a three year period is virtually the same in both instances!* Part 2 for both Exhibits II and III shows this distribution within each class for both cases. Shown below are the statewide claims distributions for each case along with negative binomial distributions possessing the same mean and variance.

	Compound Distributions			
	Gamma/Poisson		Gamma/Neg. Bin.	
Number of Claims	Actual	Neg. Binomial	Actual	Neg. Binomial
0	.652	.652	.653	.652
1	.247	.246	.249	.246
2	.074	.074	.073	.074
3	.020	.020	.020	.020
4	.005	.005	.005	.005
5	.001	.001	.001	.001

(2) *The ratio of expected losses for groups having x accidents in a three year period, $\alpha(x)$, to those having none, $\alpha(0)$, is substantially lower in the negative binomial case than it is in the Poisson case.* Part 1 of Exhibits II and III shows the values of $\alpha(x)$ within each class for the two simulations, while Part 3 shows their relativity to the 0 accident class. Part 4 shows the relativities to the class mean frequency. It is interesting to note that in both cases the frequency of claims in classes 20 and 22 is so high that even risks with one claim are better than the average for the class and should be charged a rate below the class average!

Shown below are the statewide expected frequencies and their relativities to the expected frequency of the group with zero accidents:

	Compound Distributions			
Number of	Gamma/Poisson		Gamma/Neg. Bin.	
Claims	Frequency	Relativity	Frequency	Relativity
X	$\alpha(X)$	$\alpha(X)/\alpha(0)$	$\alpha(X)$	$\alpha(X)/\alpha(0)$
0	.126	1.00	.135	1.00
1	.199	1.58	.191	1.42
2	.274	2.16	.247	1.83
3	.351	2.78	.304	2.26
4	.433	3.43	.363	2.69
5	.522	4.13	.425	3.15

These results can be explained by the fact that the expected loss distribution underlying the negative binomial case has less variance than that underlying the Poisson case. This is consistent with a model which assumes that more of the total population variance is explained by the variance of the individual risk processes, and less by the variance between risks. This is also evident in the K and BK values resulting from each case. Column 1 from Part 1 of Exhibits II and III shows the BK values underlying the class expected loss distributions in each case, while column 2 shows the K value underlying the actual distribution of claim frequencies. These two columns should be identical in Exhibit II, the Poisson case, but the limitations of the simulation process resulted in slight differences.

In Exhibit III, the value BK of the expected loss distribution over the entire state is 2.22 while the value K of the claim frequency distribution is 1.68. Thus it can be seen that if the negative binomial assumption used in the example is a better picture of reality than the Poisson assumption, a given class plan will actually be 32% more efficient than the SRI methodology would indicate. This difference in class plan efficiency estimates can be observed when we test the efficiency of rates based on the claim frequencies shown in the SRB exhibit (Exhibit I). Shown below are the class relativities and their variance:

Class	Relativity (R_i)	Distribution ($Prob\,(R_i)$)	Variance
00	1.067	3.5%	XX
10	0.938	58.0	XX
12	0.889	10.2	XX
15	0.726	9.3	XX
20	2.213	0.1	XX
22	2.192	0.2	XX
24	1.514	1.0	XX
26	1.313	7.4	XX
30	1.067	2.8	XX
31	0.807	1.3	XX
40	1.621	1.9	XX
42	1.800	2.9	XX
50	1.319	1.3	XX
Total	1.000	100.0%	.053

From the formula given earlier for estimating the efficiency of a risk assessment system:

$$Efficiency = BK \times \Sigma\ [(R_i - 1)^2 Prob\,(R_i)]$$

we see that this class plan would be 8.9% efficient if *BK* were equal to 1.68, (the Poisson case), and 11.8% efficient if *BK* were equal to 2.22 (the negative-binomial case).

(3) *The efficiency of a merit rating plan is reduced in the negative binomial case, compared with the Poisson case.* The total efficiency of rates based on the indicated frequencies shown in Exhibit III Part 1 is 26.4%, while it would be 28.9% if rates were based on Exhibit II Part 1, generated from the gamma-Poisson model. This is all the more surprising since the class plan by itself (without claims history) is more effective in the negative binomial case. This effect is due, of course, to the reduced variance underlying the accident likelihood distribution shown in Exhibit III. The efficiency contribution of the claims history portion of such a rating plan is 14.6% in the negative binomial case and 20.0% in the Poisson case!

(4) *The ratio of expected frequencies for risks with one claim in three years to those with none is still a good indicator of class homogeneity, but not quite as good as in the Poisson case.* Shown below are the ratios for each class, the indicators of class *BK* values based on them, and the actual *BK* values underlying the simulation:

Class	$\alpha(1)/\alpha(0)$	*BK* (Est)	*BK* (Actual)
00	1.770	1.30	1.25
10	1.410	2.44	2.33
12	1.317	3.15	3.02
15	1.338	2.96	2.85
20	1.378	2.65	2.48
22	1.380	2.63	2.47
24	1.412	2.43	2.30
26	1.230	4.35	4.09
30	1.219	4.57	4.34
31	2.462	0.68	0.66
40	1.227	4.40	4.12
42	1.158	6.34	5.88
50	1.294	3.41	3.22
Total	1.419	2.39	2.22

These ratios give a reasonably good estimate of the *BK* values underlying the accident likelihood distribution, *but are definitely biased.*

CREDIBILITY THEORY AND RISK ASSESSMENT

It seems evident that dividing risks into groups according to the number of claims they have experienced over a particular period of time and then observing the results over a subsequent period can provide insight into class homogeneity and the efficiency of risk assessment.

There is a need, however, for a better understanding of the way that individual experience and expected loss distributions relate to each other.

It has long been recognized that in many instances greater rate accuracy can be gained by utilizing both group information and individual risk experience. Credibility theory was developed, in part, as a tool for combining these two sources of information.

In *Mathematical Models in Risk Theory* H. Bühlmann discussed Bayesian methods for estimating the expected losses for an individual risk given its actual losses. He pointed out that most such methods require knowledge of the parametric distributions of the individual risk processes and of the structure function. Since such knowledge is lacking in most practical applications, Bühlmann suggested the use of formulae based on linear approximations of the theoretically correct quantities. In effect, he suggested that the theoretically correct quantities could be approximated by a straight line fitted to the regression of expected losses over actual losses, using the method of least squares.

We can represent such a line as follows:

$$E\,(M|x) = a + bx$$

where the linear expression on the right side of the equation represents the line of best fit of the regression of expected losses over actual losses [21].

It is well known that the slope of such an equation is equal to the covariance of the dependent and independent variables divided by the variance of the independent variable [22].

Thus:

$$b = \frac{Cov\,(M,X)}{Var\,(X)}$$

In turn:

$$\begin{aligned} Cov\,(M,X) &= E\,(M,X) - E\,(M)\,E\,(X) \\ &= E\,(M,X) - (E(M))^2 \end{aligned} \tag{26}$$

since $E(X) = E\,(M)$. (See formula (7).) Furthermore:

$$\begin{aligned} E\,(M,X) &= \int_0^\infty \sum_{j=0}^\infty M\,X_j\,P(M,X_j)\,dM \\ &= \int_0^\infty M \sum_{j=0}^\infty X_j\,P(M)\,P(X_j|M)\,dM \\ &= \int_0^\infty M\,P(M)\,E(X|M)\,dM \\ &= \int_0^\infty M^2\,P(M)\,dM \\ &= E(M^2) \end{aligned} \tag{27}$$

since $E(X|m) = m$. (Note that $m = E(\phi_i)$, and see formula (1).) Thus:

$$Cov\ (M, X) = E(M^2) - (E(M))^2 = Var\ (M) \tag{28}$$

and the slope of the credibility equation is:

$$b = \frac{Var\ (M)}{Var\ (X)} \tag{29}$$

Since the constant in a least squares regression line is equal to the mean of the dependent variable minus the slope of the line times the mean of the independent variable, we can express the constant in this case as (see note [22]):

$$\begin{aligned} a &= E(X) - b \times E(X) \\ &= E(X) - E(X) \times \frac{Var\ (M)}{Var\ (X)} \\ &= E(X) \times \{ 1 - \frac{Var\ (M)}{Var\ (X)} \} \end{aligned} \tag{30}$$

Thus the linear Bayesian formula for estimating expected losses for an individual risk, EL, given its actual experience, X, is the familiar credibility equation:

$$EL = E(X) \times (1 - Z) + X \times Z \tag{31}$$

where X is the observed experience for the risk and:

$$Z = \frac{Var\ (M)}{Var\ (X)} \tag{32}$$

If we interchange the order of integration and summation in formula (27), we can express $E\ (M, X)$ as follows:

$$\begin{aligned} E\ (M, X) &= \sum_{j=0}^{\infty} \int_0^{\infty} X_j\ M\ P(X_j, M)\ dM \\ &= \sum_{j=0}^{\infty} X_j\ P(X_j) \int_0^{\infty} M\ P(M|X_j)\ dM \\ &= \sum_{j=0}^{\infty} X_j\ P(X_j)\ E(M|X_j) \end{aligned} \tag{33}$$

The importance of this last expression lies in the fact that all of its components can be estimated from observable data. The quantity $P(X_j)$ can be estimated from the number of risks having X_j losses during any given observation period, while $E(M|X_j)$ can be estimated by observing those risks with X_j losses during a subsequent observation period. It should also be noted that since $Cov(M, X) = Var(M)$:

$$\begin{aligned} Var(M) &= E(M,X) - E(M)E(X) \\ &= \{\sum_{j=0}^{\infty} X_j P(X_j) E(M|X_j)\} - (E(X))^2 \end{aligned} \tag{34}$$

Thus Var (M) can be estimated by making two observations of a risk population, estimating $P(X_j)$ for all j from the first observation, $E(M|X_j)$ from the second, summing over all j, and then subtracting the square of the population mean. It should be noted that since the second observation is being used to estimate conditions prevailing during the first period, adjustments should be made to reflect any changes in conditions between the first and second periods, such as differences in the underlying population mean [23].

At this point we will define new terms which are helpful in estimating $Var(M)$ using the covariance method:

$$\alpha(X_j) = E(M|X_j)$$

and:

$$r(\alpha(X)_j)) = P(X_j)$$

We further define the term t as the adjustment factor reflecting those differences between the observation periods which affect the group as a whole.

Using these identities, we can estimate $Var(M)$ as follows:

$$\begin{aligned} Var(M) &= E(M,X) - E(M)E(X) \\ &= \{\sum_{j=0}^{\infty} X_j r(\alpha(X_j)) (a(X_m) \div t)\} - (E(X))^2 \end{aligned} \tag{35}$$

where $\alpha(X_j)$ is calculated from a subsequent observation period and is adjusted to conditions prevailing during the first. Note that:

$$E(X) = \sum_{j=0}^{\infty} X_j r(\alpha(X_j))$$

$$Var(X) = \{\sum_{j=0}^{\infty} X_j^2 r(\alpha(X_j))\} - (E(X))^2$$

C. Hewitt provided a useful example of a loss generating process and its relationship to Bayesian credibility theory which will be used to illustrate the relationships just discussed [24]. Mr. Hewitt's example used a die and spinner to create a population with four loss processes, all equally represented in the population. The following matrix shows the joint probability of each process and its outcome [25]:

	Outcome		
State	0	2	14
A_1B_1	.83333	.13889	.02778
A_1B_2	.83333	.08333	.08333
A_2B_1	.50000	.41667	.08333
A_2B_2	.50000	.25000	.25000

Suppose that we have been able to observe this population for three repetitions of this process (three "years") and wish to estimate the variance of expected losses by comparing the last repetition to the first two. We obtain the following matrix of joint probabilities:

Joint Probabilities of Loss—$(P(X_1X_2))$

3rd Year	2 Year Losses (X_1)					
Losses (X_2)	0	2	4	14	16	28
0	.35185	.16049	.03498	.08025	.02881	.01029
2	.08025	.06996	.02281	.02881	.01560	.00480
14	.04012	.02881	.00780	.02058	.00960	.00420
$r(\alpha(X_i))$	.47222	.25926	.06559	.12963	.05401	.01929
$\alpha(X_i)$	1.5294	2.0953	2.3608	2.6667	3.0667	3.5467

where the subscripts refer to observations made from the first and second periods respectively.

We can use this information to compute the mean, variance, and covariance of these outcomes and can estimate *Var (M)* by recognizing that the mean for the group during the second observation period will be only half of that during the first, since only one period of time is utilized for the second observation while two are utilized for the first. Thus t, in this case, will be .50 and using formula (35) we have:

1. $E(X)$	4.000
2. $Var(X)$	40.444
3. t	.500
4. $E(M,X)$	22.222
5. $Var(M)$	6.222
6. Z	.154
7. BK	2.571

We can compare the estimates of $E(M|X_j)$ generated by formula (31), the credibility equation,

$$E(M|X_j) = \{4.000 \times (1 - .154)\} + X_j \times .154$$

to the results obtained using Bayes theorem (since we have the necessary information in this simulation).

	$E(M\|X_j)$		
X_j	Bayesian	Credibility	Difference
0	3.0588	3.3846	0.3258
2	4.1906	3.6922	−0.4984
4	4.7216	4.0000	−0.7216
14	5.3334	5.5384	0.2050
16	6.1334	5.8460	−0.2874
28	7.0934	7.6922	0.5988
Total	4.0000	4.0000	0.0000

Hewitt made several observations about the nature of the credibility estimate compared to the true, or Bayesian, estimate, which are particulary cogent at this point. He observed that:

1. Credibility does not (necessarily) produce the optimum estimate while the Bayesian estimate is optimum.
2. Credibility does produce the "least-squares" fit to the optimum (Bayesian) estimates for all possible outcomes weighted by the respective probabilities of those outcomes.
3. Both estimates—credibility and Bayesian—are "in-balance" for all possible outcomes [26].

It can be seen from this example how these points apply. The credibility estimates are quite biased in most cases and thus are not optimum. They are in balance, however, since the expectation of the credibility estimates is equal to $E(X)$!

It is clear, therefore, that Var (M) can be estimated using observable data as long as at least two observations of the risk population can be made. These estimates are unbiased and do not require any assumptions about the nature of the loss processes for individual risks, or about the distribution of expected losses!

RATIO ESTIMATES

It was pointed out earlier in the paper that a reasonably accurate estimate of BK, and thus $Var\ (M)$, was obtained by the simple ratio of merit rating frequencies for risks with one accident to that for risks who were claim free. That is (see formula (22)):

$$BK = \frac{\alpha(0)}{\alpha(1) - \alpha(0)}$$

The credibility estimate for $\alpha(n)$ is:

$$\alpha(n) = \{ E(X) \times (1-Z) \} + \{ Z \times n \}$$

Thus:

$$\frac{\alpha(1)}{\alpha(0)} = \frac{\{ E(X) \times (1-Z) \} + \{ Z \}}{E(X) \times (1-Z)} = 1 + \{ \frac{Z}{1-Z} \times \frac{1}{E(X)} \}$$

and:

$$\frac{\alpha(1) - \alpha(0)}{\alpha(0)} = \frac{Var\ (M)}{E(X) \times \{E(X) + E[Var\,(\phi_i)]\}}$$

$$= \frac{Var\ (M)}{(E(X))^2 + E(X) \times E[Var\,(\phi_i)]}$$

$$= \frac{E(X)}{BK \times \{ E(X) + E[Var\,(\phi_i)] \}}$$

Thus:

$$\frac{\alpha(1)}{\alpha(0)} = 1 + \{ \frac{1}{BK} \times \frac{E(X)}{E(X) + E[Var\,(\phi_i)]} \}$$

Since the Poisson assumption is only valid when $E[Var(\phi_i)]$ is equal to zero, this simplifies to:

$$\frac{\alpha(1)}{\alpha(0)} = 1 + \frac{1}{BK}$$

It can be seen that this ratio test does produce unbiased estimates of BK when the Poisson assumptions hold. It is interesting to note that the ratio test is exact in the gamma-Poisson case, since it was shown earlier (see formula (20)) that the ratio of $\alpha(1)$ to $\alpha(0)$ was equal to one plus the inverse of BK (since in the Poisson case, K is equal to BK).

If the individual risk process is not Poisson, the ratio test will be biased by the amount:

$$\frac{E(X) + E[Var(\phi_i)]}{E(X)}$$

This explains why the results were biased when this test was applied to the negative binomial simulation where $E[Var(\phi_i)]$ was greater than zero.

CLAIM FREE DISCOUNT

If the regression of expected losses over actual losses is reasonably linear, which it usually is when only accident frequencies are involved, there is another convenient way to estimate $Var(M)$ using merit rating data.

To begin with, we note that:

$$\alpha(0) = E(X) \times (1 - Z)$$

Therefore:

$$E(X) - \alpha(0) = \frac{E(X) \times Var(M)}{Var(X)}$$

and:

$$Var(M) = Var(X) \times \left\{ 1 - \frac{\alpha(0)}{E(X)} \right\}$$

The quantity $\alpha(0) \div E(X)$ represents the ratio of expected losses for risks with claim free experience to the ratio of losses for all risks and thus the quantity in braces represents the claim free discount. We can see, therefore, that the variance of expected losses can be estimated by multiplying the variance of actual losses, $Var(X)$, by the claim free discount!

NORTH CAROLINA EXPERIENCE

The methods outlined above can be applied to actual data by assuming that the observed frequency of the events, X_1 and X_2 are unbiased estimators of the true joint probabilities of these events. The following table shows the experience of North Carolina drivers over a four year period, split between the first three years and the fourth year [27].

Second Period (X_2)	Number of Losses First Period (X_1)							
	0	1	2	3	4	5	6	7
0	2002577	295414	45203	7666	1441	300	82	25
1	104048	26776	6255	1577	375	83	20	4
2	5931	2362	811	247	80	30	13	7
3	438	231	102	34	11	10	0	3
4	30	16	12	2	3	1	0	1
5	5	9	3	2	0	0	0	0
$r(\alpha(X_j))$	.8445	.1298	.0209	.0038	.0008	.0002	.0000	.0000
$\alpha(X_j)$	.0555	.0994	.1574	.2300	.3037	.4175	.4000	.7750

We note the following facts:

1.	First period mean		.1874
2.	Second period mean		.0643
3.	$Var(X)$		.2316
4.	t	{Quotient of means for two periods}	.3432
5.	$E(M,X)$		.0688
6.	$Var(M)$		.0337
7.	Claim Free Discount	{1.0 − (.0555 ÷ (2))}	.1369
8.	$Var(M)$ from Claim Free Discount	{(7) × (3)}	.0317
9.	Z	{(6) ÷ (3)}	.1455
10.	BK	{from covariance formula}	1.0421
11.	K		.8656

We see therefore, that we have been able to estimate the homogeneity of the North Carolina driving population without having to make any estimates about a gamma-Poisson process. We note further that there is a significant difference between the two estimates of BK (since K is the SRI estimate of BK) and thus there is a clear indication that the SRI method does not accurately measure the homogeneity of the North Carolina population!

The following table shows merit rating relativities from actual experience, the credibility indicated relativities, and the relativities indicated by the Poisson model (using the four year K value of .8656 [28]):

		Merit Rating Relativities		
X_j	$r(\alpha(X_j))$	Actual	Credibility	Poisson
0	.845	.864	.855	.822
1	.130	1.546	1.630	1.772
2	.021	2.448	2.406	2.722
3	.004	3.576	3.182	3.672
4	.001	4.722	3.958	4.622
5	.000	6.492	4.733	5.571
6	.000	6.220	5.509	6.521
Total	1.000	1.000	1.000	1.000

It can be seen that a merit rating procedure based on the Poisson assumptions would undercharge the 84.5% of the population who were claim free and would substantially overcharge the 15.1% with one or two claims.

The actual data also shows a noticeable departure from linearity for those groups with three or more claims, which suggests that the gamma distribution may not be an appropriate description of the structure of the distribution of expected losses for the North Carolina driving population!

MEASURING HETEROGENEITY

Various ways of estimating the variance of expected losses within an insurance population or subpopulation have been explored in this paper. In all cases, attention has been focused on estimating the variance of the structure function, *Var (M)*, since it is the measure of how much heterogeneity there actually is in the population. If one can measure *Var (M)* for any given group, one has a direct measure of the homogeneity of that group.

The first measure explored was that used by the SRI study which consisted of estimating *Var (M)* by substracting the mean of the actual loss experience, *E(X)*, from its variance. This measure can be thought of as the "excess variance" method. It has been shown that the use of this method requires the assumption that there are no random or stochastic elements affecting exposure to loss, ϕ_i. *If, in fact, this assumption is invalid then any conclusions about the effectiveness of current risk assessment practices based on this measure are not appropriate.*

The second measure consisted of estimating BK (and thus $Var\ (M)$ indirectly) by calculating the ratio of merit rating experience for risks with one accident and zero accidents respectively. This method was termed the "ratio method" and proved reasonably effective even when the Poisson assumption was not made. It was shown, however, that it would be biased by the ratio of the average within risk variance to the population mean:

$$\frac{E(X)\ +\ E[Var\,(\phi_i)]}{E(X)} = \frac{E[E(\phi_i)]\ +\ E[Var\,(\phi_i)]}{E[E(\phi_i)]}$$

A third method for estimating $Var\ (M)$ is to multiply the indicated claim free discount by the variance of the claims experience. That is:

$$Var\ (M) = CFD \times Var(X)$$

where CFD is the claim free discount. This measure was shown to be independent of the Poisson assumption, but it is dependent on the linearity of the regression of expected loss over actual loss. It gives reasonable results if the departure from linearity is not too great but can give poor results as in the Hewitt example where the difference between the actual and linear estimate of the claim free discount is approximately 10%. This situation is likely to exist in most pure premium applications.

The fourth method uses the relationship

$$E(M,\ X) = \sum_{j=0}^{\infty} X\ P(X_j)\ E(M|X_j)$$

and the fact that the expression on the right can be estimated from observable data taken over to successive periods of time to estimate $Var\ (M)$.

This measure is unbiased, is not affected by the linearity of the regression of expected losses over actual losses, and requires no assumptions about the distribution of losses for individual risks or the distribution of expected loss between members of the risk population. It is, however, subject to sampling variance and the possibility that the characteristics of groups selected on the basis of their loss experience may change with respect to the rest of the population from one period to the next. This would occur, for instance, if individual risk experience were not independent over time.

This method also provides a measure of heterogeneity of the distribution of expected losses over the entire period observed. If each observation period is two years, then the measure estimates $Var\ (M)$ where $M = E\ (\phi_i)$ is the expected loss for the ith risk over the entire four year period.

CONCLUSIONS

This paper has explored questions of risk assessment efficiency and class homogeneity. It has been shown that:

1. The SRI efficiency measure itself is an intuitively reasonable way to gain an overall idea of the effect of risk assessment on the variance within and between classes.

2. Since it is impossible to observe the structure function directly, it is necessary to make inferences about its nature using data which can be observed.

3. SRI International and the SRB "solve" the problem of estimating *Var (M)* by making use of the following relationship:

   $$Var_t\,(M) \;=\; Var_t\,(X) \;-\; E_t\,(X) \;-\; t \;\times\; E[Var\,(\phi_i)]$$

   (see formula (18)). They assume that there are no random, or stochastic, elements affecting exposure to loss and thus conclude that $E[Var\,(\phi_i)]$ is equal to zero. This conclusion makes it possible to use the "excess variance" method of determining *Var (M)* which consists of subtracting the observed mean of the actual loss experience from the variance. That is:

   $$Var\,(M) \;=\; OV \;-\; OM$$

   If there are, in fact, random elements associated with exposure, estimates using the "excess variance" method will be biased and misleading.

4. It is possible to estimate *Var (M)* without making arbitrary assumptions about the variance of exposure, ϕ_i, or the nature of the loss process and the shape of the structure function by observing actual experience over more than one period of time and utilizing the fact that:

   $$Var\,(M) \;=\; \{\sum_{j=0}^{\infty} X_j\, r(\alpha(X))\, (\alpha(X_j) \div t)\,\} \;-\; (E(X))^2$$

   where t represents the ratio of the average loss frequency for the first observation to that of the second observation, $r(\alpha(X_j))$ represents the probability that a risk will have X_j losses, and $\alpha(X_j)$ represents the expected losses of that group as estimated from a second observation period.

The purpose of risk assessment is to create homogeneous groups of insureds. The covariance method provides a readily available tool to measure group homogeneity directly, as long as credible subgroups are the object of measurement, and thus provides a way of measuring and monitoring the effectiveness of risk assessment. It provides an objective methodology for defining partitions of the insurance population and also builds in a mechanism for responding to changes in circumstances which might indicate a need for a different type of partitioning system.

The consequences of a lack of class homogeneity were pointed out by the Massachusetts State Rating Bureau:

> "If . . . classes are homogeneous, then each such class average is indeed typical of the expected loss associated with all policies in that class.
>
> "But when classes are heterogeneous, the mean expected loss for each class—however accurately it is estimated—is not at all typical of what each policy is expected to cost."[29]

In the future, actuaries will no longer be allowed to focus their attention exclusively on mean class rates without explicit concern about the types of classes that they are defining and working with. There is valid public concern about the possibility that "good" risks may be paying more than they should for their insurance, while "bad" risks are paying less. There is no evidence whatsoever that this is taking place, but our past inability to demonstrate that our classes are relatively homogeneous has troubled many reasonable people. Actuaries can hope to provide this reassurance only by developing objective measures and standards for class homogeneity. The methods and analyses presented in this paper should provide the basis for such objective measures and standards, and it is up to practicing actuaries to determine how they may be developed and applied.

APPENDIX
SIMULATING PRIVATE PASSENGER AUTOMOBILE EXPERIENCE RATING FREQUENCIES

The purpose of these simulations was to produce annual expected loss frequencies for groups of risks partitioned on the basis of the number of losses experienced during the prior three years. This can also be thought of as a way of generating the actual loss frequency expected during a fourth year, in which case it represents a simulation of the results of two observations of the population of interest.

The simulation procedure consists of the following steps:

1. Create a discretized structure function for the group or subgroup being analyzed.
   In this paper, gamma distributions were generated on the computer for each class shown on Exhibit I. The means of these gamma distributions were set equal to the means of the various classes. In the Poisson simulation, the variance of the gamma distributions was set equal to:

   $$\frac{m_i^2}{K}$$

   where the subscript refers to the class. In the negative-binomial simulation, the variance of X for each class was set equal to the variance of X in the Poisson simulation, so the variance of the gamma structure function was set equal to the following:

   $$\frac{m_i^2}{K} \times \frac{(10 - K)}{11}$$

   This adjustment reflects the fact that for each class:

   $$Var(\phi_i) = \frac{(t \times E(\phi_i))^2}{10}$$

   as shown on page 103 of the text.
   The result of this procedure for each simulation was a 62 by 13 matrix. The rows represent a partition of the domain of the structure function and the columns represent the 13 classes. Exhibit IV shows selected values from these matrices for each simulation. It can be seen, for example, that

that members of Class 10 had a 4.7% probability of having expected losses, $E(\phi_i)$, between .05 and .06. Note that in terms of the distributions discussed in the paper, each column of the matrix represents the discrete density function of the structure function, $U(m_i)$.

2. Calculate for each discrete value of m_i, the Poisson and negative binomial conditional probabilities of X losses given $t \times m_i$. This results in a 62 by 6 matrix where the rows represent the discrete values of m_i, as before, and the columns represent the values of X, $X = 0, 1, \ldots, 5$. The values in this matrix represent the probability of X accidents in three years, given an annual frequency rate, $E(\phi_i)$, equal to m_i. These values are shown in Exhibit V. If one refers to Exhibit V, Part 2, it can be seen that the negative binomial probability of being claim free for three years, given an annual expected frequency of .055, is 84.9%, while the Poisson probability for the same event shown on Part 1 of Exhibit V is 84.8%.
3. Calculate the matrix of $r(\alpha(X))$ values as follows (using matrix notation):

$$r(\alpha(X)) = U' \times H$$

where H is the matrix of conditional probabilities of X accidents given m_i and $h(X| m_i)$, and U' is the transpose of the structure function matrix, U.
4. Calculate the matrix of $\alpha(X)$ values by first defining the matrix W as being the product of the ith row of U and the scalar $m_i \div r(\alpha(X_i))$ and then taking the matrix product:

$$\alpha(X) = W' \times H$$

Clearly $\alpha(X)$ represents the following expectation:

$$\alpha(X) = E(m_i|X)$$

where X represents the accident experience during the prior three years.

Part 1 of Exhibits II and III shows the matrix of $\alpha(X)$ values while Part 2 shows the matrix of values of $r(\alpha(X))$. Shown below for illustrative purposes is data which can be used to generate values of $\alpha(X)$ and $r(\alpha(X))$ for the structure function provided in the partitioning example used in the paper. It will be recalled that the only possible values of m_i were .01, .02, .03, . . . ,.10 and that the third partition was 44% efficient. The probability of any particular value of m_i within each partition was .2. Thus we have the matrix U reflecting the structure function within each class and within the overall population, as follows:

	Structure Function: *(U)*		
	Class		Population
m_i	1	2	
.01	.2	.0	.1
.02	.2	.0	.1
.03	.2	.0	.1
.04	.0	.2	.1
.05	.2	.0	.1
.06	.0	.2	.1
.07	.2	.0	.1
.08	.0	.2	.1
.09	.0	.2	.1
.10	.0	.2	.1
Total	1.0	1.0	1.0

Assuming that $g(x_i)$ is negative binomial with an exposure variance equal to:

$$.0625 \times (E(\phi_i))^2$$

as discussed earlier (see page 99), the conditional probability matrix, *H*, assuming an initial three year observation period, is as follows:

	Probability of *X* claims given m_i: *(H)* Number of Claims *(X)*					
m_i	0	1	2	3	4	5
.01	.97045	.02910	.00045	.00000	.00000	.00000
.02	.94180	.05644	.00173	.00004	.00000	.00000
.03	.91401	.08211	.00376	.00012	.00000	.00000
.04	.88705	.10618	.00649	.00027	.00001	.00000
.05	.86091	.12873	.00983	.00051	.00002	.00000
.06	.83555	.14984	.01371	.00085	.00004	.00000
.07	.81096	.16956	.01810	.00131	.00007	.00000
.08	.78710	.18796	.02291	.00190	.00012	.00001
.09	.76396	.20511	.02811	.00262	.00019	.00001
.10	.74151	.22107	.03364	.00348	.00028	.00002

With this information, we begin by calculating the matrix of values of $\alpha(X)$:

	$\alpha(X)$					
Class	0	1	2	3	4	5
1	.03463	.04762	.05641	.06144	.06433	.06610
2	.07260	.07911	.08401	.08745	.08985	.09156
Total	.05254	.06813	.07727	.08281	.08647	.08905

Next we calculate the matrix of values of $r(\alpha(X))$:

	$r(\alpha(X))$					
Class	0	1	2	3	4	5
1	.89963	.09319	.00677	.00040	.00002	.00000
2	.80303	.17403	.02097	.00183	.00013	.00001
Total	.85133	.13361	.01387	.00111	.00007	.00000

From these two matrices it is now possible to make the following series of calculations:

	Class		
	1	2	Total
1. $E(X_1)$	.10800	.22200	.16500
2. $E(X_2)$	.03600	.07400	.05500
3. t	.33333	.33333	.33333
4. $E(M,X)$	.00528	.01782	.01155
5. $Cov(M,X)$	.00139	.00139	.00247
6. $Var(M)$	.00046	.00046	.00082
7. $Var_t(X)$	.11250	.22728	.17314
8. $E[Var(\phi_i)]$	.00011	.00037	.00024

The value of $E[Var(\phi_i)]$ is calculated using the following formula:

$$E[Var(\phi_i)] = \{ Var_t(X) - [E_t(X) + Var_t(M)] \} / t$$

In the above table, $E(X_1)$ is equal to $E_t(X)$, and the three year variance of the structure function, $Var_t(M)$, is nine times the one year variance (see pages 85—88). The details on how the values of $\alpha(X)$ and $r(\alpha(X))$ are put to use can be found on pages 107—116.

NOTES AND REFERENCES

[1] SRI International, *The Role of Risk Classification in Property and Casualty Insurance: A Study of the Risk Assessment Process*. (1976). (Note that there were three volumes issued as a part of this project: an Executive Summary, the Final Report, and a Supplement to the Final Report.)

The most important discussion of the SRI Report findings on class homogeneity and risk assessment efficiency is found on pp. 81–82. of the Final Report: "within each group there remains a wide range of accident likelihoods. The risk assessment process is still imprecise for individual insureds . . . "

[2] The attitude of the Massachusetts Division of Insurance, of which the State Rating Bureau was a part, was expressed most strongly in "Insurance Rates and Social Policy." This paper was presented at the 1977 hearings on 1978 Massachusetts Auto Rates conducted by Commissioner Stone.

The SRB recommendations about departing from actuarial rates are found in another paper presented at the 1977 hearings: "Identifying Equitable Insurance Premiums for Risk Classes: An Alternative to the Classical Approach" by Dr. J. Ferreira, Jr. found as Chapter IV of *Automobile Insurance Risk Classification: Equity and Accuracy* issued in 1978 by the Massachusetts Divison of Insurance.

[3] "Report of the Rates and Rating Procedures Task Force of the Automobile Insurance (D3) Subcommittee," November 1978, p. 6.

[4] A brief description of the assumptions underlying the use of the Poisson process is provided on pp. 175–176 and 212–213 of the Supplement to the SRI Report. Another discussion of the importance of the compound-Poisson process in risk theory is given in *Mathematical Methods in Risk Theory* by H. Bühlmann (Springer-Verlag, 1970). Dr. Bühlmann discusses what he calls "infinitely divisible" probability distributions and makes the statement that *for distributions defined on the non-negative integers, every infinitely divisible characteristic function is compound-Poisson!* (See pp. 69–73)

Intuitive support for the Poisson assumption can be derived by consideration of the fact that the limit of a binomial process taken over shorter and shorter time intervals is Poisson.

[5] The moments of the compound-Poisson process can be derived from the fact that the unconditional expectation of a random variable can be expressed in terms of conditional expectations. If we let μ_i represent the conditional mean of the ith state, we can express the unconditional variance of such a random variable, X, as:

$$\begin{aligned} Var\,(X) &= E\,\{\, E(X^2|\mu_i)\,\} \;-\; \{\, E[E\,(X|\mu_i)]\,\}^2 \\ &= E\,\{\, Var\,(X|\mu_i) \;+\; [\, E\,(X|\mu_i)]^2\,\} \;-\; \{\, E[E\,(X|\mu_i)]\,\}^2 \\ &= E\,[Var\,(X|\mu_i)] \;+\; E(\mu_i)^2 \;-\; (E(\mu_i))^2 \\ &= E\,[Var\,(X|\mu_i)] \;+\; Var\,(\mu_i) \end{aligned}$$

since $E(X|\mu_i) = \mu_i$. In the compound-Poisson distribution, the mean and variance of the conditional Poisson process are both equal to the parameter of the process, ϕ_i, and thus:

$$Var\,(X) = E(\phi_i) + Var\,(\phi_i)$$

[6] Bühlmann, p. 65.

[7] See note [5]. In this case the conditional mean is m_i.

[8] If m_i represents the average losses for the ith risk during a single period of time, then $t \times m_i$ will represent the average losses for t units of time. If $Var\,(M)$ represents the variance of the structure function for a single period of time, then $Var\,(t \times M)$ will represent that variance for t units time. Thus:

$$Var_t\,(M) = t^2 \times Var\,(M)$$

[9] R.A. Bailey, "Any Room Left for Skimming the Cream?" *PCAS* XLVII (1960), p. 30.

[10] See discussion by L.H. Roberts of Bailey, op. cit. p. 213.

[11] SRI Final Report pp. 46–55, and Supplement pp. 200–203.

[12] The SRI procedure is discussed further on pages 93—95 of this paper.

[13] Private letter from M. Treloar to R.G. Woll, June 1978.

[14] The mean and variance of the accident distribution are defined in the SRI Supplement, p. 177.

[15] Ferreira, p. 110. Dr. Ferreira states: ". . . it is recommended that the factors other than class means be considered in setting 1978 auto insurance rates in Massachusetts. *It is further recommended that consideration be given to the homogeneity of such classes* and that either the method incorporated in this paper or an approach incorporating its basic principles be used *in place of the traditional actuarial method for determining class and territorial differentials*." (Emphasis added)

[16] Ibid. p. 85.

[17] Ibid.

[18] Ibid. p. 86.

[19] J. Ferreira, Jr. "Merit Rating and Automobile Insurance," *Automobile Insurance Classification: Equity and Accuracy*, Chapter III, p. 69.

[20] SRI Supplement, pp. 205–206.

[21] Bühlmann, pp. 100–103.

[22] For example, see Hoel, Port, and Stone, *Introduction to Statistical Theory*, (Houghton Mifflin, 1971), p. 115. They show that if we write the regression equation as:

$$Y = a + b(X - \bar{X})$$

then:

$$a = \bar{Y}$$

and:

$$b = \rho \frac{s_y}{s_x}$$

where s_Y and s_X are the standard deviations of Y and X respectively, and ρ is the correlation coefficient of Y and X. Thus:

$$b = \frac{Cov\,(Y, X)}{\sqrt{Var\,(Y)\,Var\,(X)}} \times \frac{\sqrt{Var\,(Y)}}{\sqrt{Var\,(X)}}$$

$$= \frac{Cov(Y,X)}{Var\,(X)}$$

Note that the constant term becomes:

$$a = \bar{Y} - \frac{Cov\,(Y, X)}{Var\,(X)} \times \bar{X}$$

In other words, the constant term is equal to the mean of the dependent variable minus the mean of the independent variable times the slope of the regression line.

[23] I am indebted to Dr. D. Rosenfield of Arthur D. Little (ADL) who helped me realize how the covariance of M and X could be utilized to estimate $Var\ (M)$.

[24] C. C. Hewitt, Jr., "Credibility for Severity," *PCAS* LVII (1968), pp. 148–171.

[25] Ibid. p. 150.

[26] Ibid. p. 152. The probabilities shown in the table are taken from the description of the die and spinner probabilities assuming independence.

[27] R. Stewart and R.J. Campbell, "The Statistical Association between Past and Future Accidents and Violations," (1970). This study is very useful for analyzing the concepts discussed in this paper, since it contains the data used in this paper along with other combinations of observation periods and driver groups. It is not insurance data, and it is hard to know how indicative results based on such data might be of actual insurance results.

[28] This value of K was calculated by estimating the values of $r(\alpha(X_i))$ from four year North Carolina data. First, values of $E(X)$ and $Var\ (X)$ were calculated:

$$E(X) = \sum_{j=0}^{\infty} X_j\, r\,(\alpha(X_j))$$

$$Var\ (X) = \{\sum_{j=0}^{\infty} X_j^2\, r(\alpha(X_j))\} - (E(X))^2$$

and then K was set equal to the following:

$$\begin{aligned} K &= \frac{(E(X))^2}{Var\ (X) - E(X)} \\ &= \frac{.2517^2}{.3249 - .2517} \\ &= .8656 \end{aligned}$$

[29] J. Ferreira, Jr. "Identifying Equitable Insurance Premiums for Risk Classes: An Alternative to the Classical Approach," p. 82.

EXHIBIT I

EFFECT OF MERIT RATING ON CLASS RELATIVITIES*

Driver Class	1975 Exposure %	Observed 1975 Average Claim Frequency (× 100)	k** Value	Observed 1975 Driver-Class Relativity (before)	Predicted Driver-Class Relativity (after 3 years of Merit Rating)	Percent Change (6) − (5) × 100
(1)	(2)	(3)	(4)	(5)	(6)	(7)
00	3.55	17.31	1.025	1.067	1.016	− 4.8%
10	58.11	15.21	1.752	.938	.952	+ 1.5%
12	10.27	14.42	2.160	.889	.924	+ 3.9%
15	9.34	11.78	2.066	.726	.780	+ 7.4%
20	.07	35.90	1.902	2.213	1.770	−20.0%
22	.24	35.56	1.894	2.192	1.759	−19.8%
24	1.00	24.55	1.747	1.514	1.362	−10.0%
26	7.37	21.30	2.720	1.313	1.272	− 3.1%
30	2.77	17.31	2.842	1.067	1.085	+ 1.7%
31	1.28	13.09	0.570	.807	.745	− 7.7%
40	1.92	26.30	2.739	1.621	1.480	− 8.3%
42	2.94	29.19	3.493	1.800	1.650	− 8.3%
50	1.33	21.40	2.272	1.319	1.262	− 4.3%
State -wide Average	100.0	16.22	1.969***	1.00	1.00	0.0%

*Reproduced by permission of Massachusetts State Rating Bureau.

**The value of k is an estimate of class homogeneity. (The square root of the reciprocal of k is the coefficient of variation of the claim frequency distribution underlying the class.)

***The actual statewide value of k is 1.685 (See Exhibit II, Part 1).

EXHIBIT II

MASSACHUSETTS: POISSON SIMULATION

PART 1

Expected Claim Frequencies: $\alpha(X)$

				$\alpha(X)$					
Class	*BK*	*K*	*E(X)*	0	1	2	3	4	5
00	1.022	1.025	.173	.115	.227	.340	.451	.563	.675
10	1.753	1.752	.152	.121	.190	.258	.327	.396	.464
12	2.162	2.160	.144	.120	.176	.231	.287	.343	.398
15	2.068	2.066	.118	.101	.149	.198	.247	.296	.345
20	1.872	1.904	.359	.227	.349	.471	.592	.714	.837
22	1.865	1.895	.356	.225	.347	.468	.589	.711	.833
24	1.739	1.747	.245	.172	.272	.371	.469	.568	.667
26	2.722	2.720	.213	.173	.236	.299	.363	.426	.488
30	2.845	2.842	.173	.146	.198	.249	.301	.352	.404
31	0.568	0.570	.131	.077	.214	.350	.486	.622	.758
40	2.734	2.739	.263	.204	.279	.354	.428	.502	.576
42	3.488	3.493	.292	.233	.300	.367	.434	.500	.567
50	2.271	2.272	.214	.167	.240	.314	.387	.460	.533
Total	1.684	1.685	.162	.126	.199	.274	.351	.433	.522

PART 2

Distribution within Class and Merit Rating Category: $r(\alpha(X))$

	$r(\alpha(X))$						
Class	0	1	2	3	4	5	Weight
00	.657	.226	.077	.026	.009	.003	.035
10	.667	.241	.069	.018	.004	.001	.580
12	.674	.243	.064	.015	.003	.001	.102
15	.722	.218	.049	.010	.002	.000	.093
20	.428	.292	.153	.072	.032	.014	.001
22	.431	.291	.152	.071	.031	.013	.002
24	.542	.280	.114	.042	.015	.005	.010
26	.563	.292	.103	.031	.008	.002	.074
30	.621	.273	.081	.028	.005	.001	.028
31	.742	.172	.055	.019	.007	.003	.013
40	.500	.306	.128	.045	.015	.004	.019
42	.458	.320	.144	.053	.017	.005	.029
50	.568	.284	.103	.032	.009	.003	.013
Total	.652	.247	.074	.020	.005	.001	1.000

EXHIBIT II

MASSACHUSETTS: POISSON SIMULATION

PART 3

Merit Rating Relativities to Claim Free Rate: $\alpha(X) \div \alpha(0)$

	$\alpha(X) \div \alpha(0)$					
Class	0	1	2	3	4	5
00	1.000	1.982	2.963	3.936	4.908	5.887
10	1.000	1.571	2.141	2.711	3.279	3.844
12	1.000	1.462	1.925	2.388	2.850	3.309
15	1.000	1.483	1.966	2.451	2.937	3.425
20	1.000	1.540	2.076	2.610	3.147	3.691
22	1.000	1.542	2.080	2.616	3.156	3.701
24	1.000	1.579	2.154	2.726	3.297	3.873
26	1.000	1.368	1.735	2.102	2.467	2.831
30	1.000	1.351	1.702	2.054	2.406	2.756
31	1.000	2.758	4.526	6.278	8.028	9.793
40	1.000	1.368	1.734	2.098	2.460	2.823
42	1.000	1.288	1.575	1.859	2.143	2.430
50	1.000	1.441	1.882	2.320	2.756	3.193
Total	1.000	1.576	2.164	2.777	3.425	4.125

PART 4

Merit Rating Relativities to Class Mean: $\alpha(X) \div E(X)$

	$\alpha(X) \div E(X)$					
Class	0	1	2	3	4	5
00	0.663	1.313	1.963	2.608	3.252	3.901
10	0.793	1.246	1.699	2.151	2.601	3.050
12	0.833	1.219	1.604	1.990	2.375	2.757
15	0.854	1.267	1.679	2.093	2.508	2.925
20	0.634	0.976	1.315	1.654	1.995	2.339
22	0.635	0.979	1.321	1.661	2.004	2.350
24	0.702	1.108	1.512	1.913	2.314	2.719
26	0.810	1.108	1.405	1.702	1.997	2.293
30	0.846	1.143	1.440	1.737	2.035	2.331
31	0.591	1.631	2.676	3.712	4.747	5.791
40	0.776	1.061	1.345	1.627	1.908	2.189
42	0.799	1.029	1.258	1.486	1.713	1.941
50	0.779	1.123	1.467	1.809	2.148	2.489
Total	0.779	1.228	1.686	2.164	2.669	3.214

EXHIBIT III

MASSACHUSETTS: NEGATIVE BINOMIAL SIMULATION

PART 1

Expected Claim Frequencies: $\alpha(X)$

				$\alpha(X)$					
Class	*BK*	*K*	*E(X)*	0	1	2	3	4	5
00	1.245	1.017	.173	.124	.220	.312	.400	.485	.568
10	2.331	1.748	.152	.128	.181	.232	.283	.332	.380
12	3.019	2.153	.144	.127	.167	.207	.245	.284	.321
15	2.853	2.059	.118	.105	.141	.176	.210	.244	.277
20	2.473	1.836	.359	.257	.354	.447	.536	.622	.707
22	2.462	1.830	.356	.255	.351	.444	.533	.619	.704
24	2.299	1.729	.245	.189	.267	.342	.415	.485	.553
26	4.094	2.712	.213	.186	.229	.271	.312	.353	.393
30	4.344	2.831	.173	.156	.190	.223	.256	.289	.322
31	0.658	0.565	.131	.084	.207	.325	.437	.545	.650
40	4.125	2.727	.263	.224	.274	.324	.372	.420	.465
42	5.888	3.487	.292	.257	.298	.337	.376	.415	.452
50	3.221	2.265	.214	.180	.233	.285	.336	.385	.433
Total	2.219	1.679	.162	.135	.191	.247	.304	.363	.425

PART 2

Distribution within Class and Merit Rating Category: $r(\alpha(X))$

	$r(\alpha(X))$						
Class	0	1	2	3	4	5	Weight
00	.656	.229	.076	.025	.009	.003	.035
10	.666	.243	.068	.017	.004	.001	.580
12	.674	.244	.063	.015	.003	.001	.102
15	.721	.219	.048	.010	.002	.000	.093
20	.426	.297	.153	.070	.031	.031	.001
22	.429	.296	.152	.070	.030	.013	.002
24	.540	.283	.113	.041	.015	.005	.010
26	.562	.294	.102	.030	.008	.002	.074
30	.620	.274	.080	.020	.005	.001	.028
31	.740	.175	.055	.019	.007	.003	.013
40	.498	.309	.127	.045	.014	.004	.019
42	.456	.323	.144	.052	.017	.005	.029
50	.567	.287	.102	.032	.009	.003	.013
Total	.651	.249	.073	.020	.005	.001	1.000

EXHIBIT III

MASSACHUSETTS: NEGATIVE BINOMIAL SIMULATION

PART 3

Merit Rating Relativities to Claim Free Rate: $\alpha(X) \div \alpha(0)$

	$\alpha(X) \div \alpha(0)$					
Class	0	1	2	3	4	5
00	1.000	1.770	2.510	3.221	3.906	4.569
10	1.000	1.410	1.811	2.203	2.586	2.959
12	1.000	1.317	1.628	1.934	2.235	2.531
15	1.000	1.338	1.669	1.996	2.317	2.635
20	1.000	1.378	1.740	2.087	2.425	2.756
22	1.000	1.380	1.743	2.092	2.432	2.764
24	1.000	1.412	1.809	2.193	2.563	2.923
26	1.000	1.230	1.456	1.678	1.896	2.112
30	1.000	1.219	1.434	1.646	1.858	2.070
31	1.000	2.462	3.868	5.209	6.497	7.744
40	1.000	1.227	1.449	1.666	1.876	2.081
42	1.000	1.158	1.313	1.464	1.613	1.757
50	1.000	1.294	1.580	1.860	2.133	2.398
Total	1.000	1.419	1.834	2.256	2.691	3.150

PART 4

Merit Rating Relativities to Class Mean: $\alpha(X) \div E(X)$

	$\alpha(X) \div E(X)$					
Class	0	1	2	3	4	5
00	0.718	1.272	1.803	2.314	2.806	3.283
10	0.843	1.189	1.527	1.858	2.181	2.495
12	0.880	1.159	1.433	1.702	1.967	2.227
15	0.894	1.196	1.492	1.784	2.071	2.355
20	0.718	0.989	1.249	1.499	1.741	1.979
22	0.719	0.992	1.253	1.504	1.748	1.987
24	0.771	1.088	1.395	1.690	1.976	2.253
26	0.873	1.073	1.270	1.464	1.655	1.843
30	0.898	1.095	1.288	1.479	1.669	1.860
31	0.640	1.574	2.474	3.332	4.156	4.954
40	0.850	1.043	1.231	1.415	1.594	1.768
42	0.880	1.018	1.155	1.288	1.419	1.546
50	0.843	1.090	1.332	1.567	1.797	2.021
Total	0.831	1.179	1.523	1.873	2.235	2.616

EXHIBIT IV
PART 1A

MASSACHUSETTS: POISSON SIMULATION

Gamma Structure Function Probabilities: $U(m_i)$

	Class					
m_i	00	10	12	15	20	22
.00005	.00052	.00000	.00000	.00000	.00000	.00000
.00055	.00483	.00024	.00005	.00011	.00003	.00004
.00550	.04873	.01276	.00629	.01121	.00228	.00237
.01500	.05215	.02783	.01937	.03109	.00573	.00591
.02500	.04960	.03663	.03031	.04525	.00842	.00864
.03500	.04702	.04213	.03866	.05454	.01064	.01089
.04500	.04452	.04542	.04463	.05995	.01250	.01278
.05500	.04212	.04711	.04856	.06239	.01408	.01436
.06500	.03983	.04763	.05080	.06262	.01541	.01570
.07500	.03764	.04729	.05167	.06125	.01652	.01681
.08500	.03557	.04632	.05146	.05876	.01745	.01773
.09500	.03360	.04489	.05043	.05554	.01821	.01849
.10500	.03174	.04314	.04877	.05187	.01883	.01910
.11500	.02997	.04118	.04667	.04796	.01932	.01957
.12500	.02830	.03908	.04427	.04399	.01969	.01993
.13500	.02672	.03690	.04168	.04007	.01997	.02019
.14500	.02523	.03471	.03898	.03629	.02015	.02036
.15500	.02381	.03252	.03626	.03269	.02025	.02045
.16500	.02248	.03038	.03356	.02932	.02028	.02046
.17500	.02122	.02830	.03093	.02619	.02025	.02042
.18500	.02003	.02630	.02840	.02332	.02016	.02031
.19500	.01890	.02438	.02599	.02069	.02002	.02016
.20500	.01784	.02256	.02371	.01831	.01984	.01997
.21500	.01683	.02084	.02157	.01616	.01962	.01973
.22500	.01588	.01922	.01957	.01423	.01937	.01947
.23500	.01499	.01769	.01772	.01251	.01909	.01917
.24500	.01414	.01627	.01600	.01097	.01878	.01886
.25500	.01335	.01494	.01443	.00960	.01846	.01852
.26500	.01259	.01371	.01299	.00839	.01811	.01816
.27500	.01188	.01256	.01167	.00732	.01775	.01779
.28500	.01121	.01150	.01047	.00638	.01738	.01741
.29500	.01058	.01051	.00938	.00555	.01700	.01701
.30500	.00998	.00961	.00839	.00483	.01661	.01662
.31500	.00942	.00877	.00750	.00419	.01621	.01621
.32500	.00888	.00800	.00669	.00363	.01581	.01580
.33500	.00838	.00729	.00596	.00315	.01541	.01539
.34500	.00791	.00665	.00531	.00273	.01500	.01498
.35500	.00746	.00605	.00472	.00236	.01460	.01457
.36500	.00704	.00551	.00420	.00204	.01420	.01416
.74000	.00754	.00117	.00032	.00005	.03628	.03561
.84000	.00420	.00041	.00008	.00002	.02413	.02359
.94000	.00234	.00014	.00003	.00000	.01586	.01545

EXHIBIT IV
PART 1B

MASSACHUSETTS: POISSON SIMULATION

Gamma Structure Function Probabilities: $U(m_i)$

	Class						
m_i	24	26	30	31	40	42	50
.00005	.00000	.00000	.00000	.01429	.00000	.00000	.00000
.00055	.00012	.00000	.00000	.03789	.00000	.00000	.00001
.00550	.00605	.00078	.00103	.13601	.00045	.00005	.00219
.01500	.01339	.00393	.00558	.08564	.00231	.00045	.00765
.02500	.01817	.00827	.01204	.06508	.00497	.00138	.01315
.03500	.02166	.01297	.01902	.05369	.00798	.00279	.01812
.04500	.02426	.01760	.02568	.04604	.01108	.00460	.02243
.05500	.02617	.02188	.03159	.04038	.01411	.00670	.02603
.06500	.02756	.02569	.03652	.03595	.01696	.00899	.02894
.07500	.02851	.02893	.04039	.03234	.01957	.01137	.03122
.08500	.02911	.03160	.04320	.02933	.02189	.01376	.03292
.09500	.02942	.03368	.04503	.02676	.02391	.01609	.03410
.10500	.02949	.03523	.04597	.02453	.02561	.01831	.03482
.11500	.02937	.03627	.04615	.02258	.02701	.02036	.03516
.12500	.02909	.03685	.04569	.02086	.02812	.02222	.03515
.13500	.02868	.03703	.04469	.01932	.02895	.02387	.03486
.14500	.02816	.03686	.04327	.01793	.02952	.02529	.03433
.15500	.02755	.03639	.04152	.01669	.02986	.02649	.03360
.16500	.02687	.03567	.03954	.01555	.02998	.02745	.03272
.17500	.02614	.03474	.03740	.01452	.02992	.02819	.03171
.18500	.02537	.03364	.03516	.01357	.02969	.02872	.03061
.19500	.02457	.03242	.03287	.01271	.02931	.02905	.02943
.20500	.02375	.03110	.03058	.01191	.02881	.02919	.02820
.21500	.02292	.02971	.02833	.01117	.02819	.02916	.02695
.22500	.02208	.02827	.02614	.01049	.02749	.02897	.02568
.23500	.02124	.02682	.02403	.00986	.02672	.02864	.02440
.24500	.02041	.02535	.02201	.00927	.02588	.02819	.02314
.25500	.01958	.02390	.02010	.00873	.02500	.02763	.02190
.26500	.01877	.02248	.01831	.00822	.02409	.02698	.02068
.27500	.01797	.02108	.01663	.00775	.02315	.02625	.01950
.28500	.01719	.01973	.01507	.00731	.02220	.02546	.01835
.29500	.01643	.01843	.01362	.00690	.02124	.02462	.01724
.30500	.01569	.01717	.01229	.00651	.02029	.02374	.01618
.31500	.01497	.01598	.01106	.00615	.01934	.02282	.01516
.32500	.01428	.01484	.00994	.00581	.01840	.02189	.01418
.33500	.01360	.01376	.00892	.00549	.01748	.02095	.01326
.34500	.01295	.01273	.00799	.00519	.01658	.02000	.01238
.35500	.01233	.01177	.00714	.00491	.01570	.01906	.01154
.36500	.01172	.01087	.00638	.00465	.01485	.01812	.01075
.74000	.01337	.00285	.00046	.00655	.00979	.01131	.00467
.84000	.00725	.00099	.00010	.00403	.00432	.00470	.00190
.94000	.00389	.00033	.00005	.00250	.00186	.00189	.00076

EXHIBIT IV
PART 2A

MASSACHUSETTS: NEGATIVE BINOMIAL SIMULATION

Gamma Structure Function Probabilities: $U(m_i)$

	Class					
m_i	00	10	12	15	20	22
.00005	.00011	.00000	.00000	.00000	.00000	.00000
.00055	.00182	.00002	.00000	.00000	.00000	.00000
.00550	.03052	.00404	.00122	.00289	.00042	.00044
.01500	.04128	.01436	.00730	.01472	.00177	.00184
.02500	.04356	.02435	.01642	.02958	.00345	.00358
.03500	.04400	.03275	.02625	.04334	.00525	.00542
.04500	.04352	.03931	.03539	.05426	.00706	.00727
.05500	.04252	.04409	.04309	.06184	.00882	.00907
.06500	.04120	.04728	.04902	.06621	.01050	.01078
.07500	.03970	.04910	.05312	.06780	.01208	.01237
.08500	.03808	.04979	.05551	.06715	.01353	.01383
.09500	.03641	.04955	.05639	.06480	.01485	.01516
.10500	.03472	.04858	.05601	.06125	.01604	.01636
.11500	.03303	.04705	.05461	.05692	.01710	.01741
.12500	.03137	.04511	.05243	.05215	.01803	.01833
.13500	.02975	.04289	.04968	.04721	.01883	.01913
.14500	.02817	.04047	.04656	.04231	.01951	.01979
.15500	.02664	.03795	.04321	.03757	.02007	.02034
.16500	.02518	.03538	.03977	.03311	.02052	.02079
.17500	.02377	.03283	.03633	.02898	.02088	.02112
.18500	.02242	.03033	.03296	.02522	.02114	.02136
.19500	.02114	.02791	.02973	.02182	.02131	.02152
.20500	.01991	.02559	.02668	.01879	.02140	.02159
.21500	.01875	.02339	.02382	.01610	.02141	.02159
.22500	.01765	.02132	.02118	.01375	.02136	.02152
.23500	.01660	.01938	.01875	.01169	.02125	.02139
.24500	.01561	.01757	.01654	.00991	.02108	.02120
.25500	.01467	.01590	.01454	.00838	.02087	.02097
.26500	.01378	.01435	.01275	.00706	.02061	.02069
.27500	.01294	.01294	.01114	.00593	.02031	.02038
.28500	.01215	.01164	.00971	.00498	.01997	.02003
.29500	.01140	.01045	.00844	.00416	.01961	.01965
.30500	.01070	.00937	.00732	.00347	.01922	.01925
.31500	.01004	.00839	.00634	.00289	.01881	.01882
.32500	.00941	.00751	.00547	.00241	.01838	.01838
.33500	.00883	.00670	.00472	.00200	.01794	.01793
.34500	.00827	.00598	.00406	.00165	.01748	.01746
.35500	.00776	.00533	.00349	.00137	.01702	.01698
.36500	.00727	.00474	.00299	.00113	.01655	.01650
.74000	.00558	.00036	.00007	.00000	.03403	.03323
.84000	.00281	.00009	.00000	.00000	.02060	.02002
.94000	.00141	.00004	.00000	.00000	.01222	.01181

EXHIBIT IV
PART 2B

MASSACHUSETTS: NEGATIVE BINOMIAL SIMULATION

Gamma Structure Function Probabilities: $U(m_i)$

	Class						
m_i	24	26	30	31	40	42	50
.00005	.00000	.00000	.00000	.00773	.00000	.00000	.00000
.00055	.00001	.00000	.00000	.02699	.00000	.00000	.00000
.00550	.00154	.00003	.00005	.11876	.00001	.00000	.00025
.01500	.00554	.00048	.00075	.08325	.00021	.00001	.00182
.02500	.00974	.00182	.00302	.06593	.00083	.00005	.00477
.03500	.01369	.00422	.00716	.05572	.00199	.00019	.00862
.04500	.01724	.00756	.01287	.04855	.00372	.00051	.01295
.05500	.02035	.01162	.01960	.04307	.00593	.00111	.01740
.06500	.02300	.01610	.02670	.03866	.00854	.00205	.02170
.07500	.02520	.02071	.03356	.03500	.01141	.00337	.02567
.08500	.02698	.02521	.03974	.03188	.01442	.00507	.02917
.09500	.02837	.02938	.04490	.02918	.01744	.00714	.03214
.10500	.02941	.03308	.04886	.02682	.02039	.00952	.03455
.11500	.03012	.03620	.05158	.02472	.02316	.01215	.03639
.12500	.03056	.03870	.05307	.02285	.02569	.01493	.03768
.13500	.03074	.04055	.05344	.02117	.02794	.01779	.03847
.14500	.03071	.04176	.05283	.01965	.02986	.02063	.03880
.15500	.03049	.04238	.05139	.01826	.03144	.02337	.03871
.16500	.03010	.04245	.04929	.01700	.03268	.02594	.03827
.17500	.02958	.04203	.04670	.01585	.03358	.02828	.03752
.18500	.02895	.04120	.04376	.01479	.03416	.03035	.03653
.19500	.02823	.04002	.04060	.01382	.03442	.03209	.03532
.20500	.02743	.03856	.03734	.01292	.03441	.03351	.03396
.21500	.02657	.03687	.03407	.01209	.03414	.03458	.03248
.22500	.02566	.03502	.03085	.01133	.03364	.03531	.03091
.23500	.02472	.03306	.02776	.01061	.03295	.03570	.02929
.24500	.02377	.03103	.02482	.00995	.03208	.03578	.02764
.25500	.02279	.02898	.02207	.00934	.03108	.03557	.02599
.26500	.02182	.02693	.01952	.00877	.02996	.03509	.02435
.27500	.02085	.02492	.01719	.00824	.02876	.03438	.02275
.28500	.01989	.02296	.01506	.00774	.02749	.03347	.02118
.29500	.01894	.02108	.01315	.00728	.02618	.03238	.01967
.30500	.01801	.01928	.01143	.00684	.02484	.03116	.01823
.31500	.01710	.01757	.00990	.00644	.02348	.02982	.01684
.32500	.01622	.01597	.00855	.00606	.02214	.02840	.01553
.33500	.01537	.01447	.00736	.00571	.02080	.02692	.01429
.34500	.01454	.01307	.00631	.00537	.01950	.02540	.01312
.35500	.01374	.01178	.00540	.00506	.01822	.02388	.01203
.36500	.01297	.01059	.00461	.00477	.01699	.02235	.01101
.74000	.00925	.00063	.00018	.00553	.00406	.00342	.00175
.84000	.00428	.00003	.00000	.00322	.00125	.00078	.00051
.94000	.00195	.00016	.00000	.00188	.00036	.00032	.00013

EXHIBIT V
PART 1

MASSACHUSETTS: POISSON SIMULATION

Probability of X Claims Given m_i: (H)

	Number of Claims: (X)					
m_i	0	1	2	3	4	5
.00005	.99985	.00015	.00000	.00000	.00000	.00000
.00055	.99835	.00165	.00000	.00000	.00000	.00000
.00550	.98364	.01623	.00013	.00000	.00000	.00000
.01500	.95600	.04302	.00097	.00001	.00000	.00000
.02500	.92774	.06958	.00261	.00007	.00000	.00000
.03500	.90032	.09453	.00496	.00017	.00000	.00000
.04500	.87372	.11795	.00796	.00036	.00001	.00000
.05500	.84789	.13990	.01154	.00063	.00003	.00000
.06500	.82283	.16045	.01564	.00102	.00005	.00000
.07500	.79852	.17967	.02021	.00152	.00009	.00000
.08500	.77492	.19760	.02519	.00214	.00014	.00001
.09500	.75201	.21432	.03054	.00290	.00021	.00001
.10500	.72979	.22988	.03621	.00380	.00030	.00002
.11500	.70822	.24434	.04215	.00485	.00042	.00003
.12500	.68729	.25773	.04833	.00604	.00057	.00004
.13500	.66698	.27013	.05470	.00738	.00075	.00006
.14500	.64726	.28156	.06124	.00888	.00097	.00008
.15500	.62814	.29208	.06791	.01053	.00122	.00011
.16500	.60957	.30174	.07468	.01232	.00152	.00015
.17500	.59156	.31057	.08152	.01427	.00187	.00020
.18500	.57407	.31861	.08841	.01636	.00227	.00025
.19500	.55711	.32591	.09533	.01859	.00272	.00032
.20500	.54064	.33249	.10224	.02096	.00322	.00040
.21500	.52466	.33841	.10914	.02346	.00378	.00049
.22500	.50916	.34368	.11599	.02610	.00440	.00059
.23500	.49411	.34835	.12279	.02886	.00509	.00072
.24500	.47951	.35244	.12952	.03173	.00583	.00086
.25500	.46533	.35598	.13616	.03472	.00664	.00102
.26500	.45158	.35901	.14271	.03782	.00752	.00120
.27500	.43823	.36154	.14914	.04101	.00846	.00140
.28500	.42528	.36362	.15545	.04430	.00947	.00162
.29500	.41271	.36525	.16162	.04768	.01055	.00187
.30500	.40052	.36647	.16766	.05114	.01170	.00214
.31500	.38868	.36730	.17355	.05467	.01292	.00244
.32500	.37719	.36776	.17928	.05827	.01420	.00277
.33500	.36604	.36787	.18486	.06193	.01556	.00313
.34500	.35523	.36766	.19026	.06564	.01698	.00352
.35500	.34473	.36714	.19550	.06940	.01848	.00394
.36500	.33454	.36632	.20056	.07320	.02004	.00439
.74000	.10861	.24111	.26763	.19805	.10992	.04880
.84000	.08046	.20276	.25548	.21460	.13520	.06814
.94000	.05961	.16809	.23701	.22278	.15706	.08858

EXHIBIT V
PART 2

MASSACHUSETTS: NEGATIVE BINOMIAL SIMULATION

Probability of X Claims Given m_i: (H)

	Number of Claims: (X)					
m_i	0	1	2	3	4	5
.00005	.99985	.00015	.00000	.00000	.00000	.00000
.00055	.99835	.00165	.00000	.00000	.00000	.00000
.00550	.98365	.01620	.00015	.00000	.00000	.00000
.01500	.95609	.04283	.00106	.00002	.00000	.00000
.02500	.92800	.06908	.00283	.00008	.00000	.00000
.03500	.90082	.09360	.00535	.00022	.00001	.00000
.04500	.87451	.11649	.00853	.00045	.00002	.00000
.05500	.84904	.13782	.01230	.00080	.00004	.00000
.06500	.82438	.15768	.01659	.00127	.00008	.00000
.07500	.80051	.17615	.02132	.00188	.00013	.00001
.08500	.77740	.19331	.02644	.00263	.00021	.00001
.09500	.75502	.20922	.03189	.00353	.00032	.00002
.10500	.73334	.22395	.03761	.00459	.00046	.00004
.11500	.71235	.23757	.04357	.00581	.00063	.00006
.12500	.69202	.25013	.04972	.00719	.00084	.00009
.13500	.67232	.26169	.05602	.00872	.00110	.00012
.14500	.65324	.27232	.06244	.01041	.00141	.00016
.15500	.63476	.28205	.06893	.01225	.00177	.00022
.16500	.61684	.29094	.07547	.01424	.00218	.00029
.17500	.59949	.29903	.08204	.01637	.00265	.00037
.18500	.58266	.30637	.08860	.01864	.00318	.00047
.19500	.56636	.31301	.09514	.02103	.00378	.00058
.20500	.55055	.31897	.10164	.02356	.00444	.00072
.21500	.53523	.32431	.10808	.02619	.00516	.00088
.22500	.52038	.32905	.11443	.02894	.00595	.00105
.23500	.50598	.33322	.12070	.03180	.00681	.00125
.24500	.49202	.33687	.12686	.03474	.00773	.00148
.25500	.47848	.34002	.13290	.03778	.00872	.00174
.26500	.46534	.34270	.13881	.04089	.00979	.00202
.27500	.45261	.34494	.14459	.04408	.01092	.00233
.28500	.44025	.34677	.15022	.04733	.01212	.00267
.29500	.42827	.34820	.15571	.05064	.01338	.00305
.30500	.41664	.34927	.16103	.05400	.01471	.00345
.31500	.40536	.34999	.16620	.05740	.01611	.00389
.32500	.39442	.35039	.17121	.06084	.01757	.00437
.33500	.38380	.35049	.17604	.06431	.01909	.00488
.34500	.37349	.35030	.18071	.06780	.02067	.00543
.35500	.36348	.34985	.18520	.07130	.02230	.00601
.36500	.35377	.34915	.18952	.07482	.02400	.00663
.74000	.13468	.24466	.24446	.17765	.10489	.05335
.84000	.10567	.21269	.23546	.18957	.12401	.06989
.94000	.08339	.18343	.22192	.19527	.13960	.08598