WHY LARGER RISKS HAVE SMALLER INSURANCE CHARGES

IRA ROBBIN

Abstract

The insurance-charge function is defined as the excess ratio (the ratio of expected loss excess of an attachment point to the expected total loss) and is expressed as a function of the entry ratio (the ratio of the attachment to the total loss expectation). Actuaries use insurance-charge algorithms to price retrospective rating maximums and excess of aggregate coverages. Many of these algorithms are based on models that can be viewed as particular applications of the Collective Risk Model (CRM) developed by Heckman and Meyers [4]. If we examine the insurance-charge functions for risks of different sizes produced by these models, we will find invariably that the insurance charge for a large risk is less than or equal to the charge for a small risk at every entry ratio. The specific purpose of this paper is to prove that this must be so. In other words, we will show the assumptions of the CRM force charge functions to decline by size of risk. We will take a fairly general approach to the problem, develop some theory, and prove several results along the way that apply beyond the CRM.

We will first prove that the charge for a sum of two non-negative random variables is less than or equal to the weighted average of their charges. We will extend that result to show that under certain conditions, the charge for a sum of identically distributed, but not necessarily independent, samples declines with the sample size. The extension is not entirely straightforward, as the desired result cannot be directly derived using simple in-

duction or straightforward analysis of the coefficient of variation (CV).

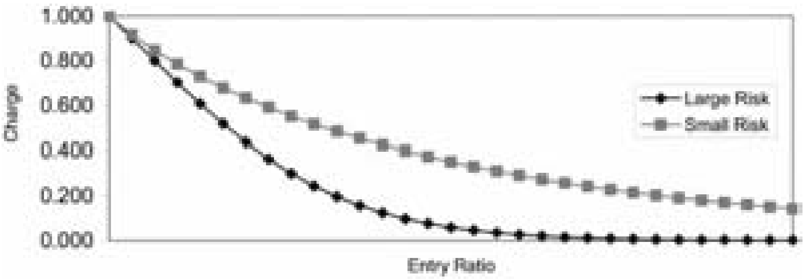
To explore size and charge in some generality, we will define the construct of a risk-size model. A risk-size model may initially be viewed as a collection of non-negative random variables whose sizes are defined by their expectation values. Given an appropriate measure on the risks of a particular size, we will be able to regard the cumulative distribution and the charge as well-defined functions of risk size. In a complete and continuous model, there are risks of every size and the cumulative distribution is a continuous function of risk size. We will first show that the charge declines with size if any risk can be decomposed into the independent sum of smaller risks in the model. Then we will employ the usual Bayesian construction to introduce parameter risk and extend the result to models that are not decomposable. This is an important extension, because actuaries have long known from study of Table M that a large risk is not the independent sum of smaller ones. In particular, our result implies that charges decrease with size in the standard contagion model of the Negative Binomial used in the CRM. Finally, we will introduce severity, prove our result assuming a fixed severity distribution, and then extend it to cover the type of parameter uncertainty in risk severity modeled in the CRM. Thus we will arrive at the conclusion that the assumptions of the CRM force charges to decline by size of risk.

ACKNOWLEDGEMENT

The author gratefully acknowledges the valuable contributions made by Michael Singer, Dave Clark, and Jim Adams. Dave was an excellent sounding board for ideas and his insightful responses greatly aided this work. Michael served admirably as proofreader and editor. His suggestions significantly improved the paper. Jim

FIGURE 1

ACTUARIAL INTUITION ABOUT CHARGES AND RISK SIZE



is a business analyst who merits honorable mention as a closet actuary whose examples, thought experiments, and practical observations helped motivate this work.

1. INTRODUCTION

Our broad objective is to study the dependence of insurance-charge functions on risk size. We would like to arrive at some sufficiently general conditions that will force charges to obey the actuarial intuition that larger risks ought to have smaller insurance charges. More precisely, we would like to show that the assumptions of the Collective Risk Model (CRM) [4] lead to decreasing charges by size of risk. In keeping with standard actuarial terminology, the insurance charge refers to the excess ratio, not the absolute dollar amount, and the charge is viewed as a function of the entry ratio. When we say the charge is smaller, we mean that the excess ratio is less than or equal to its initial value at every entry ratio (see Figure 1).

Before proving this holds under certain conditions, we should note that no one has published any article disputing it. Neither does the literature contain any example with actual data for which it fails to hold. In practice, it is implicitly assumed to be true or turns out to be true under the assumptions made for a particular model. Under the procedure promulgated by the National Council on Compensation Insurance (NCCI) [6], a column of insurance charges is selected for a given risk based on its expected losses. The columns of charges have been constructed to effectively guarantee that a large risk will always be assigned smaller insurance charge values than a small one [3].

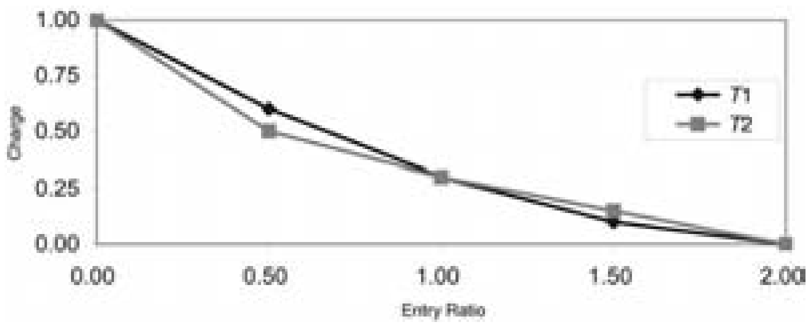
But why should this property hold? The basic intuition is that the excess ratio is related to the propensity of a distribution to take on relatively extreme values. When a large risk can be viewed as the independent sum of smaller risks, the Law of Large Numbers will apply and the likelihood of relatively extreme outcomes will decline.

Looking at the coefficient of variation (CV), the ratio of the standard deviation to the mean, supports these intuitions. When a large risk is the sum of independent, identically distributed smaller risks, the CV declines as risk size increases. Since the square of the CV is directly related to the integral of the insurance charge [8], these arguments suggest the insurance charge should also decline with risk size. However, this does not constitute a strict proof. The counterexample in Exhibit 1 shows that the CV does not uniquely determine the charge at every entry ratio. In that example, the risk with the smaller CV has a larger charge at some entry ratios (see Figure 2). In conclusion, because the arrow goes the wrong way, we cannot use a CV argument to arrive at a relatively trivial proof.

Instead we will use some numeric tricks and inequalities relating limited expected values to rigorously show that insurance charges do indeed decline as risk size increases when a large risk can be decomposed into a sum of independent smaller ones. This is a useful result, but alone it is insufficient for our larger purpose. Actuaries have long known that a large risk in practice has a distribution different than that resulting from the independent summation of smaller risks [5]. So we will go further and

FIGURE 2

RANDOM VARIABLE WITH SMALLER CV HAS LARGER CHARGE AT SOME ENTRY RATIOS



extend our result to models in which independent decomposability is only conditionally true. To do this, we will follow the usual Bayesian construction and view the true mean of a risk as a random variable having a prior distribution. A family of priors will then be used to define a family of unconditional distributions. This introduces parameter risk. Assuming an arbitrarily decomposable conditional model and priors having charges that decline with size, we will show the charges for the unconditional model decline with the unconditional risk size. However, the unconditional risk-size model will not be decomposable, and, due to the parameter uncertainty introduced via the priors, the CV for an unconditional risk will not tend towards zero as risk size becomes infinite.

We then turn to aggregate loss distributions that are generated by sampling claim counts and sampling claim severities in the manner described by the CRM. We will first show that an aggregate loss model inherits decomposability from its underlying claim count model, assuming severities are independently sampled from a fixed severity distribution. This leads to the conclusion that decomposable counts and independent fixed severities produce a model in which charges decrease by size of risk. We will then extend this result to cover the type of parameter risk in severity and frequency that is modeled in the CRM.

A reader versed in stochastic-process theory might observe the concept of decomposability is quite similar to the idea of "infinite divisibility" used in connection with Levy processes ([1], [7]). However, after some thought, we decided not to employ the terminology or results of stochastic theory. Though there is an analogy between increasing the risk size in a size-of-risk model and increasing the time in a stochastic process, we wish to maintain relevant distinctions between the two operations. In stochastic processes, the major concern is how a random variable changes over time [9] and the cumulative effect of possible jumps over a time interval. In short, it is the study of sample paths. For example, $N(t)$ might be the number of times a particular event has occurred as of time t and we might assume $N(t)$ is a right continuous function of t. The distribution of $N(t)$ would be a probabilistic summary of the number of events that have occurred as of time t, averaged over the space of sample paths (equipped with appropriate measure). In risk-size models, we are concerned with how risks of different size relate to one another; but there is no real analogue to the space of sample paths. This is not to say that many of the results to be presented here could not have been proven by applying stochastic-process theory after appropriately accounting for the distinction between time dependent paths and risk size. However, we will leave that work as a task for others who are more knowledgeable about stochastic-process theory. Also, we will not adopt the terminology of stochastic process theory since this might confuse the discussion of risk size. As well, in keeping with our actuarial focus, we will tend to make whatever reasonable assumptions we need, even though some of these could possibly be proved from previous assumptions or from more minimalist hypotheses. For instance, some of the assumptions that will be made about differentiability of our models with respect to risk size might be replaceable with more general and less restrictive statements or perhaps could be derived from the decomposability property. Since none of the assumptions put us outside the CRM, we leave the more abstract development along these lines as a topic of future research.

To maintain focus in the main exposition, many basic definitions and important foundational results have been relegated to the Appendices. The reader may be well advised to review these before proceeding much further.

In the end, most actuaries will find nothing surprising in what we will prove. But we will have rigorously established that actuarial intuition about insurance charges does indeed hold true for some fairly general classes of risk-size models, including the CRM.

2. THE INSURANCE CHARGE FOR A SUM OF RANDOM VARIABLES

We start by studying inequalities for insurance charges of sums. Our first result is that the insurance charge for the sum of two non-negative random variables is bounded by the weighted average of their insurance charges.

2.1. Charge for Sum Bounded by Weighted Average of Charges

Suppose T_1 and T_2 are non-negative random variables with means, μ_1 and μ_2, which are positive. Then, it follows that:

$$\varphi_{T_1+T_2}(r) \leq \frac{\mu_1}{\mu_1+\mu_2}\varphi_{T_1}(r) + \frac{\mu_2}{\mu_1+\mu_2}\varphi_{T_2}(r). \tag{2.1}$$

Proof Applying definition A.1 from the Appendix, we write

$$\begin{aligned}\varphi_{T_1+T_2}(r) &= \frac{\mathrm{E}[\max(0,(T_1+T_2)-r(\mu_1+\mu_2))]}{\mu_1+\mu_2}\\ &= \frac{\mathrm{E}[\max(0,(T_1-r\mu_1)+(T_2-r\mu_2))]}{\mu_1+\mu_2}.\end{aligned}$$

Next, we use the subadditivity property of the "max" operator to get $\max(0, A+B) \leq \max(0, A) + \max(0, B)$. We apply this to the previous equation and do some simple algebra to find:

$$\begin{aligned}\varphi_{T_1+T_2}(r) &\leq \frac{\mu_1}{\mu_1+\mu_2}\frac{\mathrm{E}[\max(0,(T_1 - r\mu_1)]}{\mu_1} \\ &\quad + \frac{\mu_2}{\mu_1+\mu_2}\frac{\mathrm{E}[\max(0,(T_2 - r\mu_2)]}{\mu_2} \\ &= \frac{\mu_1}{\mu_1+\mu_2}\varphi_{T_1}(r) + \frac{\mu_2}{\mu_1+\mu_2}\varphi_{T_2}(r).\end{aligned}$$

Note this result applies even if T_1 and T_2 are not independent, as it follows from subadditivity of the max operator and the linearity of the expectation operator. This result leads directly to the proof that the summation of two identically distributed risks leads to a smaller insurance charge.

2.2. Summation of Two Identically Distributed Variables Reduces the Charge

Suppose T_1 and T_2 are identically distributed and let T denote a random variable with their common distribution. Then

$$\varphi_{T_1+T_2}(r) \leq \varphi_T(r). \tag{2.2}$$

Proof From 2.1 it follows that

$$\varphi_{T_1+T_2}(r) \leq \tfrac{1}{2}\varphi_{T_1}(r) + \tfrac{1}{2}\varphi_{T_2}(r) = \varphi_T(r). \tag{2.3}$$

Note that if T_1 and T_2 were perfectly correlated, then A.6 would apply and summing would be equivalent to doubling and this would not change the charge. In Exhibit 2, we show a discrete example with two cases: one in which the two variables are independent and the other in which they are correlated. Not surprisingly, the sum of independent variables does have a lower charge than the charge for the sum when the variables are correlated. Yet, even when correlation exists, the charge for the sum is less than or equal to the charge for T.

To generalize Equation (2.2), suppose we take samples $(T_1, T_2, \ldots)$ of a random variable, T. We define $S(1,2,\ldots,n) = T_1 + T_2 + \cdots + T_n$. We make the assumption that such sums are *sample selection independent* by which we mean that the distribution of $S(1,2,\ldots,n)$ is the same as the distribution of $S(i_1, i_2, \ldots, i_n)$ where $(i_1, i_2, \ldots, i_n)$ is any ordered n-tuple of distinct positive integers. Note this does not require the T_i to be independent of one another, but it does force the distribution of $T_1 + T_2$, for example, to be the same as $T_1 + T_3$, $T_2 + T_3$, $T_{21} + T_{225}$, or the sum of any pair of distinct variables in our original sample. This would imply that there is a common correlation between any two of our samples. Under the assumption of sample selection independence, we will show the insurance charge for S_n declines as n increases. While it might seem that there ought to be some simple induction proof based on Equation (2.1), the only quick extension is that the charge for S_{nm} is less than or equal to the charge for S_n. To arrive at the general proof, we will use a numeric grouping trick and properties of the "min" operator.

2.3. Insurance Charge for Sample Selection Independent Sums Declines with Sample Size

$$\varphi_{S_{n+1}}(r) \leq \varphi_{S_n}(r). \tag{2.4}$$

Proof For $k = 1, 2, \ldots, n+1$, define

$$S(\sim k/n+1) = T_1 + T_2 + \cdots + T_{k-1} + T_{k+1} + \cdots + T_{n+1}.$$

In other words, $S(\sim k/n+1)$ is the sum of the n out of the first $n+1$ trials obtained by excluding the kth trial. For example, $S(\sim 2/3) = T_1 + T_3$. With this notation, we may write the following formula

$$n \cdot S_{n+1} = \sum_{k=1}^{n+1} S(\sim k/n+1).$$

When $n = 2$, this formula says

$$\begin{aligned} 2 \cdot S_3 &= 2(T_1 + T_2 + T_3) \\ &= (T_2 + T_3) + (T_1 + T_3) + (T_1 + T_2) \\ &= S(\sim 1/3) + S(\sim 2/3) + S(\sim 3/3). \end{aligned}$$

The formula implies

$$\begin{aligned} n(n+1)\mathrm{E}\left[\frac{S_{n+1}}{n+1}; r\right] &= \mathrm{E}[n \cdot S_{n+1}; n(n+1)r] \\ &= \mathrm{E}\left[\sum_{k=1}^{n+1} S(\sim k/n+1); n(n+1)r\right]. \end{aligned}$$

Next, we apply the inequality $\min(A+B, C+D) \geq \min(A,B) + \min(C,D)$ repeatedly to get

$$\mathrm{E}\left[\sum_{k=1}^{n+1} S(\sim k/n+1); n(n+1)r\right] \geq \sum_{k=1}^{n+1} \mathrm{E}[S(\sim k/n+1); nr].$$

Since the T_i are identically distributed and sample selection independent, it follows that

$$\mathrm{E}[S(\sim k/n+1); nr] = \mathrm{E}[S_n; nr]$$

and thus that

$$\mathrm{E}\left[\sum_{k=1}^{n+1} S(\sim k/n+1); n(n+1)r\right] \geq (n+1)\mathrm{E}[S_n; nr].$$

Connecting inequalities and factoring out n from the right hand expectation, we obtain

$$n(n+1)\mathrm{E}\left[\frac{S_{n+1}}{n+1}; r\right] \geq (n+1)n\mathrm{E}\left[\frac{S_n}{n}; r\right].$$

Assuming without loss of generality for the purpose at hand that $\mathrm{E}[T] = 1$, this inequality implies $(1 - \varphi_{S_{n+1}}(r)) \geq (1 - \varphi_{S_n}(r))$. The result then follows.

While the proof is rather abstract and the algebra of our numeric trick can be confusing, it is easy to see how it all works in any simple example.

EXAMPLE 1:

$$\varphi_{S_3}(r) \le \varphi_{S_2}(r).$$

Proof Consider

$$\begin{aligned} 6\mathrm{E}\left[\frac{S_3}{3};r\right] &= \mathrm{E}[2\cdot S_3;6r] = \mathrm{E}[(T_1+T_2+T_3)+(T_1+T_2+T_3);6r] \\ &= \mathrm{E}[(T_1+T_2)+(T_1+T_3)+(T_2+T_3);6r] \\ &\ge \mathrm{E}[T_1+T_2;2r]+\mathrm{E}[T_1+T_3;2r]+\mathrm{E}[T_2+T_3;2r] \\ &= 3\mathrm{E}[S_2;2r]. \end{aligned}$$

Thus we have

$$\mathrm{E}\left[\frac{S_3}{3};r\right] \ge \mathrm{E}\left[\frac{S_2}{2};r\right]$$

and the desired conclusion follows.

It is important to understand that we have *not* proved that any way of adding risks together reduces the charge. For example, if we had a portfolio of independent risks with small charges, and then added another risk with a large charge function, the addition of that risk could well result in a new larger portfolio with a larger charge. However, that would violate our assumption that the risks were identically distributed. Also, if we had two identically distributed risks, initially independent, and then added a third risk, but while doing so combined their operations so that all the risks were now strongly correlated, the charge might well increase. This is not a counterexample to our result, because our construction does not allow one to change correlations in the middle of the example.

3. RISK-SIZE MODELS

We need to introduce some precision in our discussion to at least guarantee that there is a well-defined notion of the insurance charge for a particular risk size. To start, we initially ignore parameter risk so that we can unambiguously identify the size of a risk with its expectation value. We then define a risk-size model, **M**, as a collection of non-negative random variables each having a finite non-negative mean. We index a random variable within such a model by its mean. We then use the risks of a particular size in the model to define the cumulative distribution, limited expected value, and insurance charge at that size. We let $\mathbf{M}_\mu$ be the set of risks in **M** of size μ and we suppose there is a measure Σ_μ on $\mathbf{M}_\mu$. We then define the cumulative distribution as a function of risk size via: $F_{\mathbf{M}}(t \mid \mu) = \mathrm{E}[F_T(t) \mid T \in \mathbf{M}_\mu]$ where the expectation is taken with respect to Σ_μ. Similarly, we define limited expected values and insurance charges as functions of risk size. We say **M** is well-defined if the measures give rise to a well-defined cumulative distribution for every $\mathbf{M}_\mu$ that is non-empty. We say a well-defined model is complete if there is a cumulative distribution for the model at every size. Unless otherwise noted, we henceforth assume all models are well-defined and complete. We define **M** to be size continuous at $t > 0$ if $F_{\mathbf{M}}(t \mid \mu)$ is a continuous function of μ and nth order size differentiable at $t > 0$ if $F_{\mathbf{M}}(t \mid \mu)$ has a nth order partial derivative with respect to μ, for $\mu > 0$. Note that **M** could be size continuous and differentiable even if all the random variables in **M** are discrete.

In the simplest case, each $\mathbf{M}_\mu$ consists of a single random variable that we denote as T_μ, and the measure, Σ_μ, assigns a mass point of 100% to this random variable. We say this is a unique size model and we use $F_{T_\mu}(t)$, the cumulative distribution function for the unique risk of size, μ, to define $F_{\mathbf{M}}(t \mid \mu)$, the cumulative distribution function at t for the model at size μ. Similarly we use the limited expected value and charge function of T_μ. to define the limited expected value and charge for the model **M** at size μ. To simplify notation when working with a unique size model, we may sometimes write $F_{T_\mu}(t)$ in place of $F_{\mathbf{M}}(t \mid \mu)$.

Next we define the notions of closure under independent summation, and decomposability in a unique size model.

3.1. Definitions of Independence, Closure and Decomposability in a Unique Size Model

Given a unique size model, **M**, and assuming $\mu_1 > 0$, $\mu_2 > 0$, we say

> M is *closed under independent summation* if $T_{\mu_1} \in \mathbf{M}$ and $T_{\mu_2} \in \mathbf{M}$ implies their independent sum, $T_{\mu_1} + T_{\mu_2}$, is also in **M**. Note these could well be independent samples of the same random variable. (3.1a)

> M is *arbitrarily decomposable* if for any positive μ, μ_1, and μ_2 with $\mu = \mu_1 + \mu_2$, there exist $T_\mu \in \mathbf{M}$, $T_{\mu_1} \in \mathbf{M}$, and $T_{\mu_2} \in \mathbf{M}$ such that the independent sum, $T_{\mu_1} + T_{\mu_2}$ has the same distribution as T_μ. (3.1b)

Unless there is need for greater specificity, we will usually say "closed" instead of "closed under independent summation." In a closed complete model, we can add identically distributed random samples of any given risk in the model and still stay in the model.

Arbitrary decomposability is a strong condition. It says that any way of splitting the mean of a risk into a sum leads to a decomposition of that risk into the independent sum of smaller risks in the model. To simplify terminology when no confusion should ensue, we may hereafter refer to "arbitrarily decomposable" models as simply "decomposable." We will show that charges decrease with size in a decomposable unique model.

First, we observe:

3.2. Decomposability Equivalence to Closure in Unique Size Model

$$\mathbf{M} \text{ is decomposable} \Leftrightarrow \mathbf{M} \text{ is closed.} \tag{3.2}$$

Proof We prove one direction and leave the other as an exercise.

"$\Rightarrow$" Omitted.

"$\Leftarrow$" Since $\mathbf{M}$ is complete, there exist $T_\mu \in M$, $T_{\mu_1} \in M$, and $T_{\mu_2} \in M$. By assumption, $\mathbf{M}$ is closed under independent summation. So the independent sum, $T_{\mu_1} + T_{\mu_2}$, is in $\mathbf{M}$. Taking expectations one has $\mathrm{E}[T_{\mu_1} + T_{\mu_2}] = \mu_1 + \mu_2$. In a unique size model, we know $T_{\mu_1+\mu_2}$ has the unique distribution in $\mathbf{M}$ with $\mathrm{E}[T_{\mu_1+\mu_2}] = \mu_1 + \mu_2$.

If we assume size differentiability in a decomposable model, we can obtain some results constraining the behavior of the cumulative distribution and the limited expected value function when these are viewed as functions of risk size.

3.3. Inequalities for Risk Size Partials in Decomposable Models

If $\mathbf{M}$ is a continuously differentiable decomposable risk-size model, then

$$\frac{\partial F_{T_\mu}}{\partial \mu} \leq 0 \tag{3.3a}$$

$$1 \geq \frac{\partial \mathrm{E}[T_\mu;t]}{\partial \mu} \geq 0 \tag{3.3b}$$

$$\frac{\partial^2 \mathrm{E}[T_\mu;t]}{\partial \mu^2} \leq 0. \tag{3.3c}$$

Proof Applying the definition of decomposability, we derive

$$F_{T_\mu}(t) = \Pr(T_\mu \leq t) \geq \Pr(T_\mu + T_{\Delta\mu} \leq t) = F_{T_\mu + T_{\Delta\mu}}(t) = F_{T_{\mu+\Delta\mu}}(t).$$

F is thus a decreasing function of the risk size and Equation (3.3a) follows.

To prove the partial exceeds zero in Equation (3.3b), we use the decomposability property to derive

$$\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t] = \mathrm{E}[T_\mu + T_{\Delta\mu};t] - \mathrm{E}[T_\mu;t] \geq 0.$$

To prove the partial is less than unity, we similarly derive

$$\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t] = \mathrm{E}[T_\mu + T_{\Delta\mu};t] - \mathrm{E}[T_\mu;t]$$
$$\leq \mathrm{E}[T_\mu;t] + \mathrm{E}[T_{\Delta\mu};t] - \mathrm{E}[T_\mu;t] = \mathrm{E}[T_{\Delta\mu};t] \leq \Delta\mu.$$

It follows that

$$\frac{\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t]}{\Delta\mu} \leq 1$$

and this leads immediately to our result.

As for Equation (3.3c), we claim that with our continuous differentiability assumption, it suffices to show that $\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t]$ is a decreasing function of μ for any $\Delta\mu > 0$. This is sufficient because, if it is true, we can then use an argument based on the Mean Value Theorem to show that the first partial derivative with respect to risk size is decreasing. A decreasing first partial derivative forces the second partial to be less than or equal to zero.

To show $\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t]$ is decreasing, we first use the additivity and independence assumptions to write the convolution formula:

$$F_{T_{\mu+\Delta\mu}}(t) = F_{T_\mu + T_{\Delta\mu}}(t) = \int_0^t dF_{T_\mu}(s) \cdot F_{T_{\Delta\mu}}(t-s).$$

This implies

$$G_{T_{\mu+\Delta\mu}}(t) - G_{T_\mu}(t) = \int_0^t dF_{T_\mu}(s) \cdot G_{T_{\Delta\mu}}(t-s).$$

Using

$$\mathrm{E}[T;t] = \int_0^t ds\, G_T(s)$$

we derive

$$\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t] = \int_0^t dx \int_0^x dF_{T_\mu}(y) \cdot G_{T_{\Delta\mu}}(x-y).$$

Switching orders of integration, we have

$$\begin{aligned}\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t] &= \int_0^t dF_{T_\mu}(y) \int_y^t dx\, G_{T_{\Delta\mu}}(x-y) \\ &= \int_0^t dF_{T_\mu}(y) \mathrm{E}[T_{\Delta\mu};t-y].\end{aligned}$$

Next, we integrate by parts and evaluate terms to obtain

$$\begin{aligned}\mathrm{E}[T_{\mu+\Delta\mu};t] - \mathrm{E}[T_\mu;t] &= F_{T_\mu}(0) \cdot \mathrm{E}[T_{\Delta\mu};t] \\ &\quad + \int_0^t dy\, F_{T_\mu}(y) \cdot G_{T_{\Delta\mu}}(t-y).\end{aligned}$$

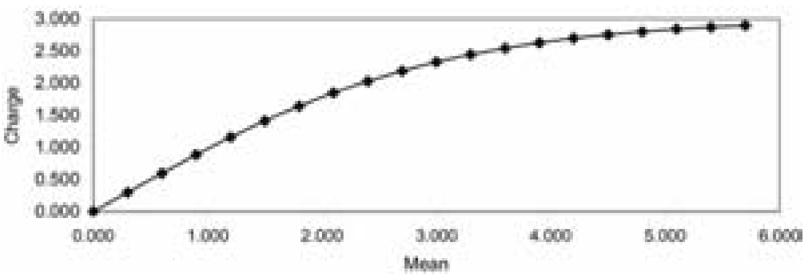
Applying Equation (3.3a), we can conclude that this decreases with μ, thus proving our result.

Note that these results apply to discrete as well as continuous distributions and that the proof does not require $F_{T_\mu}(0)$ to equal zero. Exhibit 3 shows how Poisson, Negative Binomial, and Gamma limited expected values vary as the mean changes. See Figure 3 for a graph of Poisson limited expected values at the limit 3.0 that vary as a function of risk size.

Now inequality 2.4 will be used to show that decomposable models must have charges that decrease with risk size.

FIGURE 3

POISSON LIMITED EXPECTED VALUES $E[T_\mu;3]$



3.4. *Charges Decrease with Risk Size in Decomposable Models*

Suppose **M** is a decomposable model. Then:

$$\mu_1 < \mu_2 \Rightarrow \varphi_{T_{\mu_1}} \geq \varphi_{T_{\mu_2}}. \tag{3.4}$$

Proof Since the charge function is a continuous function of risk size, it suffices to prove the result when μ_1 and μ_2 are rational. Assuming rationality, there exists θ such that $\mu_1 = m_1\theta$ and $\mu_2 = m_2\theta$ where m_1 and m_2 are integers with $m_1 < m_2$. Using completeness and decomposability, it follows that there exists T_θ in **M** and that

$$T_{\mu_1} = (T_\theta)_1 + (T_\theta)_2 + \cdots + (T_\theta)_{m_1}$$

$$T_{\mu_2} = (T_\theta)_1 + (T_\theta)_2 + \cdots + (T_\theta)_{m_1} + \cdots + (T_\theta)_{m_2}$$

where the sums are of independent samples of T_θ. Then, via inequality 2.4, the charge for T_{μ_1} exceeds the charge for T_{μ_2}.

We now apply Equation (3.4) to prove that insurance charges decrease with risk size for several families of distributions commonly used in insurance models.

3.5. Charges Decrease with Size in Poisson, Negative Binomial, and Gamma Models

The insurance charge decreases by size of risk in each of the following models:

Poisson:

$$\mathbf{M} = \{N \sim \text{Poisson}(\mu) \mid \mu > 0\}. \tag{3.5a}$$

Negative Binomial with common q:

$$\mathbf{M} = \{N \sim \text{Negative Binomial}(\alpha, q) \mid q \text{ is fixed and } \alpha > 0\}. \tag{3.5b}$$

Gamma with common scale parameter λ:

$$\mathbf{M} = \{T \sim \text{Gamma}(\alpha, \lambda) \mid \lambda \text{ is fixed and } \alpha > 0\}. \tag{3.5c}$$

Proof With the given restrictions, it can be easily shown that each of the families is a unique size model that has well-defined charges. It is also readily seen that each is decomposable. The results then follow from Equation (3.4).

Exhibit 4 shows columns of charges for risks of different sizes for Poisson random variables, Negative Binomials with common failure rate parameter, and Gammas with common scale parameter.

In a decomposable model, the charge decreases to the smallest possible charge as risk size goes to infinity.

3.6. Charge for an Infinitely Large Risk Equals Smallest Possible Charge in Decomposable Model

Suppose $\mathbf{M}$ is a differentiable decomposable model. Then

$$\varphi_{T_\mu} \to \varphi_0 \quad \text{as} \quad \mu \to \infty \quad \text{where} \quad \varphi_0(r) = \max(0, 1-r). \tag{3.6}$$

Proof It suffices to show $\varphi_{T_{n\mu}} \to \varphi_0$ as n approaches ∞ for arbitrary fixed μ. We use a CV argument. Consider $\mathrm{Var}(T_{n\mu}) = n\mathrm{Var}(T_\mu)$ for a decomposable model. Thus,

$$\mathrm{CV}^2(T_{n\mu}) = \mathrm{Var}(R_{n\mu}) = \frac{\mathrm{Var}(T_\mu)}{n\mu^2} \to 0 \qquad \text{as} \quad n \to \infty.$$

By Equation (A.11), this implies

$$\int_0^\infty dr\, \varphi_{R_{n\mu}}(r) \to \frac{1}{2}.$$

Using $\int_0^\infty dr\, \varphi_0(r) = \frac{1}{2}$, the result follows.

It is not difficult to construct a risk-size model in which charges decrease with size, even though it is not decomposable.

EXAMPLE 2: Charges Decrease by Size in a Non-Decomposable Model
Define T_μ to be the distribution having probability mass $p = \mu/(\mu+1)$, at $t = \mu + 1$ and mass $1 - p$, at $t = 0$. It follows that $\mathrm{E}[T_\mu] = \mu$ and

$$\varphi_{T_\mu}(r) = 1 - \frac{r}{r_\mu} \qquad \text{for} \quad 0 \le r < r_\mu \quad \text{where} \quad r_\mu = \frac{\mu+1}{\mu}.$$

Since r_μ declines as μ increases, the charge function declines with risk size. Yet, the independent sum of two members of this family always yields a distribution with three mass points and thus a sum that is not even in the family.

4. PARAMETER UNCERTAINTY IN RISK-SIZE MODELS

To introduce parameter uncertainty, we now suppose a model in which there may be many random variables sharing a common *expected* mean, but whose *actual* true means are uncertain. Because we do not know in advance the true mean of a risk in such a model, we use the a priori expected mean to define risk size. If we let θ represent the true mean of a risk, and μ the a priori mean, then M_μ consists of all the risks with prior mean equal to μ. While M_μ can therefore contain many risks each having a true mean, θ, which is not equal to μ, we do insist that Σ_μ, the measure, is defined so that μ is the average mean over all risks in M_μ. Following the usual Bayesian construction and the CRM, we will restrict our attention to models in which we may represent Σ_μ using a prior distribution, $H(\theta \mid \mu)$.

Before going further, it is instructive to see how such a construction can be used to model the combined effects of population parameter uncertainty and population heterogeneity.

EXAMPLE 3: Population Uncertainty and Heterogeneity
Consider a model in which the group of risks of size 100 actually consisted of risks whose true means were 90, 100, and 110. Assume we have no way of determining the true mean of any risk in advance of an experiment. Suppose there are two possible states of the world, "L" and "H," each of which has an equal random chance of occurring. If "L" applies, then we will sample from a subgroup of low risks, half of which have a true mean of 90 and half which have a true mean of 100. If "H" applies, the sampling subgroup will consist of an even split of high risks with true means of 100 and 110. We then take independent samples with replacement. In this analogy, the two possible states of the world correspond to population parameter uncertainty and the mix of risks in each state corresponds to heterogeneity of the population. To carry the analogy further, suppose the expected losses excess of 120 are 4, 10, and 18 for risks whose true means are 90, 100, and 110 respectively. If we are in state "L," our sampling will produce an average excess loss of 7, while the average excess loss will be 14 if we are in state "H." The average over all replications of this sampling process over all states will be 10.5. With a prior distribution of 25%, 50%, and 25% for risks with true means of 90, 100, and 110 respectively, we will duplicate this result. Note the correct average charge for the population

exceeds that charge for a risk whose mean equals the average mean of the population.

We will now study a specific class of risk-size models with parameter uncertainty that are constructed by applying priors to the risk sizes in a conditionally decomposable model. Under the usual Bayesian construction, the (prior probability) weighted average of the conditional distributions generates the unconditional distribution. If the family of priors itself forms a well-defined risk-size model, the family of resulting unconditional distributions will also be a well-defined risk-size model. If the risk model of priors is sufficiently well-behaved, we will be able to derive conclusions about the insurance charges of the unconditional distributions. Suppose the priors have charges that decrease, not necessarily strictly, with the unconditional risk size. We will then show the resulting unconditional distributions must also have charges that decrease with unconditional risk size.

To begin the mathematical development of this construction, let $T(\theta)$ be a non-negative random variable parametrized by θ such that $\mathrm{E}[T(\theta)] = \theta$. Suppose the family $\{T(\theta) \mid \theta > 0\}$ is differentiable with respect to θ and that it has insurance charges that decrease with risk size. Now we view the parameter θ as a random variable Θ and let $H(\theta) = H(\theta \mid \mu)$ denote its cumulative distribution. Assume $H(0) = 0$ and that Θ has density $h(\theta) = H'(\theta)$, which is continuously differentiable. Let $\mathrm{E}[\Theta]$ be finite. Since $\mathrm{E}[T(\theta)] = \theta$, it follows that the unconditional risk size, $\mathrm{E}[T(\Theta)]$, is equal to the mean, $\mathrm{E}[\Theta]$, of the parameter distribution. We often will use $\mu = \mathrm{E}[\Theta]$ to simplify notation. Finally, we let $\varphi_{T(\Theta)}$ denote the unconditional insurance charge. Given these definitions, the usual Bayesian construction leads to:

4.1. Unconditional Insurance Charge Formula

$$\varphi_{T(\Theta)}(r) = \frac{1}{\mu} \int d\theta\, h(\theta) \theta \varphi_{T(\theta)}\left(\frac{r\mu}{\theta}\right). \tag{4.1}$$

Proof Omitted.

The Bayesian construction implies there is some parameter uncertainty about the true mean of any risk. Under one interpretation, the final charge value is the (prior) probability-weighted average of the dollar charge over all values of the true mean, divided by the expected value of the true mean. Under another interpretation, we are dealing with a population of risks whose true overall mean we know, even though there is some parameter uncertainty regarding the mean of any particular risk in the population. The prior then represents the spread in the population and the formula arrives at the correct average charge for the population. It is also important to note that, as in our example, the charge for an average risk is not the same as (and is usually lower than) the weighted average charge for the population of risks. These two interpretations correspond to two types of parameter risk. The first expresses our uncertainty about the overall mean of a population, while the second expresses our uncertainty about the parameter dispersion or heterogeneity of a population. While a hierarchical, "double-integral" model could be used to separately delineate their effects, Example 3 shows that with respect to insurance charges, Equation (4.1) can be used to model both types of parameter risk together. For other applications such as in credibility theory, it may be important to maintain a distinction between these sources of parameter risk.

Now let $\mathbf{Q}$ be a family of Θ random variables and $\mathbf{M}(\mathbf{Q})$ the associated set of unconditional random variables. One quick, but important, result can be obtained assuming all the priors are scaled versions of a single distribution. Thus all the priors have the same insurance-charge function. Assume we have a conditional model in which charges decline with size. Then we show that applying the scaled priors to this model generates an unconditional model in which the insurance charge declines as a function of the unconditional risk size.

4.2. Unconditional Charge Declines with Risk Size in Scaled Priors Model

If

$$\frac{\partial \varphi_{T(\theta)}}{\partial \theta} \leq 0 \qquad \text{and} \qquad \Theta_2 = (1+c)\Theta_1 \quad \text{for} \quad c > 0,$$

then

$$\varphi_{T(\Theta_2)}(r) \leq \varphi_{T(\Theta_1)}(r). \tag{4.2}$$

Proof We start by using Equation (4.1) to write

$$\varphi_{T(\Theta_2)}(r) = \frac{1}{\mu_2}\int_0^\infty d\theta\, h_2(\theta)\cdot\theta\cdot\varphi_{T(\theta)}\left(\frac{r\mu_2}{\theta}\right)$$

where

$$\mu_2 = \mathrm{E}[\Theta_2] = \int_0^\infty d\theta\, h_2(\theta)\cdot\theta.$$

Then consider

$$\mu_2 = (1+c)\mu_1$$

$$H_2(\theta) = \Pr(\Theta_2 \leq \theta) = \Pr((1+c)\Theta_1 \leq \theta)$$

$$= \Pr\left(\Theta_1 \leq \frac{\theta}{1+c}\right) = H_1\left(\frac{\theta}{1+c}\right)$$

$$h_2(\theta) = h_1\left(\frac{\theta}{1+c}\right)\frac{1}{1+c}.$$

Substituting, rewrite the integral as

$$\varphi_{T(\Theta_2)}(r) = \frac{1}{\mu_1(1+c)}\int_0^\infty d\theta\, h_1\left(\frac{\theta}{1+c}\right)\frac{1}{1+c}$$
$$\cdot\,\theta\cdot\varphi_{T(\theta)}\left(\frac{r\mu_1(1+c)}{\theta}\right).$$

Then change variables using $\eta = (\theta/1 + c)$ to get

$$\varphi_{T(\Theta_2)}(r) = \frac{1}{\mu_1}\int_0^\infty d\eta\, h_1(\eta)\cdot \eta \cdot \varphi_{(T(\eta(1+c))}\left(\frac{r\mu_1}{\eta}\right).$$

Because we assumed the conditional insurance charges were decreasing with risk size, it follows that

$$\varphi_{T(\eta(1+c))} \le \varphi_{T(\eta)}.$$

This leads to

$$\varphi_{T(\Theta_2)}(r) \le \frac{1}{\mu_1}\int_0^\infty d\eta\, h_1(\eta)\cdot \eta \cdot \varphi_{T(\eta)}\left(\frac{r\mu_1}{\eta}\right) = \varphi_{T(\Theta_1)}(r).$$

Why does this result make intuitive sense, despite the fact that scaling up the prior doesn't change its insurance charge? The answer is that scaling up does raise the mean of the prior so that larger conditional risks have more weight in the weighted average signified by the integral. Since the large conditional risks have smaller charges, the net effect of scaling up the prior is to reduce the charge of the unconditional distribution. Note that this result did not depend on decomposability of the conditional model, merely that the conditional risk-size model had charges that decrease with size.

For an example, consider the following:

EXAMPLE 4: Gamma Contagion on Conditional Poissons
Let $T(\theta)$ be conditionally Poisson with parameter, θ. Suppose $\theta = \mu\cdot\nu$ where ν is Gamma distributed with shape parameter α, and scale parameter $\lambda = \alpha$, so that $E[\nu] = 1$. The variable, ν, introduces parameter uncertainty and $\mathrm{Var}(\nu) = (1/\alpha) = c$ is called the contagion [4] parameter for claim counts. It follows that Θ_μ, the random variable for θ, is Gamma distributed with shape parameter α, and scale parameter $\lambda = \alpha/\mu$, so that $E[\Theta_\mu] = \mu$. The unconditional distribution $T(\Theta_\mu)$, is Negative Binomial with failure rate probability parameter $q = (1+\lambda)^{-1} = \mu/(\mu+\alpha)$. If $\mathbf{Q} = \{\Theta_\mu \mid \mu > 0\}$ and $\mathbf{M}(\mathbf{Q})$ is the resulting set of Negative Bino-

mials. It follows as a consequence of Equation (4.2) that charges decrease with risk size in **M**(**Q**).

Note **M**(**Q**) consists of Negative Binomials with a common shape but different failure rate parameters. This is different from our previous decomposable Negative Binomial risk-size model, from Equation (3.5b), in which all the variables had a common failure rate parameter. Our result is that charges decline with size in the standard Gamma-Poisson claim-count contagion model. Note this model is not closed under independent summation. Further, observe that the square of the coefficient of variation, $\mathrm{CV}^2 = \mathrm{Var}(\Theta_\mu)/\mu^2 = 1/(\alpha q) = c + (1/\mu)$, decreases toward the contagion, and not zero, as the risk size grows infinite. See Exhibit 5 for tables of charges for Negative Binomials as defined in Example 3.

Next we will consider two priors with the same mean. Assume these priors are acting on a continuously differentiable decomposable conditional risk-size model. We will show that, under certain conditions, if one prior has a smaller insurance charge, then its resulting unconditional random variable also has a smaller insurance charge. In order to prove this, we will first use integration by parts to express the unconditional insurance charge in terms of an integral of a product of the risk partials of the limited expected values of the conditional model and the prior.

4.3. Unconditional Charge Formula

$$\varphi_T(\Theta)(r) = 1 - \frac{1}{\mu}\int_0^\infty d\theta \frac{\partial \mathrm{E}[\Theta;\theta]}{\partial\theta} \cdot \frac{\partial \mathrm{E}[T(\theta);r\mu]}{\partial\theta}. \tag{4.3}$$

Proof We write

$$\varphi_{T(\Theta)}(r) = 1 - \frac{1}{\mu}\int_0^\infty d\theta\, h(\theta)\cdot \mathrm{E}[T(\theta);r\mu].$$

Then we perform integration by parts as follows to derive

$$\int_0^\infty d\theta\, h(\theta)\mathrm{E}[T(\theta);r\mu] = -(1-H(\theta))\mathrm{E}[T(\theta);r\mu]\Big|_{\theta=0}^{\theta=\infty}$$

$$+\int_0^\infty d\theta(1-H(\theta))\frac{\partial \mathrm{E}[T(\theta);r\mu]}{\partial\theta}$$

$$=\int_0^\infty d\theta(1-H(\theta))\frac{\partial \mathrm{E}[(T(\theta);r\mu]}{\partial\theta}.$$

The result follows since

$$1-H(\theta)=\frac{\partial \mathrm{E}[\Theta;\theta]}{\partial\theta}.$$

With this we can now show a smaller charge for the prior leads to a smaller charge for the unconditional distribution at a common risk size.

4.4. Smaller Charges for the Prior Lead to Smaller Unconditional Charges—Size Fixed

If

$$\mu=\mathrm{E}[\Theta_1]=\mathrm{E}[\Theta_2] \qquad \text{and} \qquad \frac{\partial^2\mathrm{E}[T(\theta);r\mu]}{\partial\theta^2}\le 0$$

then

$$\varphi_{\Theta_2}\le\varphi_{\Theta_1} \quad \text{implies} \quad \varphi_{T(\Theta_2)}\le\varphi_{T(\Theta_1)}. \tag{4.4}$$

Proof We use Equation (4.5) to obtain

$$\varphi_{T(\Theta_1)}(r)-\varphi_{T(\Theta_2)}(r)$$

$$=-\frac{1}{\mu}\int_0^\infty d\theta\left(\frac{\partial\mathrm{E}[\Theta_1;\theta]}{\partial\theta}-\frac{\partial\mathrm{E}[\Theta^2;\theta]}{\partial\theta}\right)\frac{\partial\mathrm{E}[T(\theta);r\mu]}{\partial\theta}.$$

We integrate by parts to obtain

$$\mu(\varphi_{T(\Theta_1)}(r) - \varphi_{T(\Theta_2)}(r))$$
$$= -(\mathrm{E}[\Theta_1;\theta] - \mathrm{E}[\Theta_2;\theta])\,\frac{\partial \mathrm{E}[T(\theta); r\mu]}{\partial\theta}\bigg|_{\theta=0}^{\theta=\infty}$$
$$+ \int_0^\infty d\theta(\mathrm{E}[\Theta_1;\theta] - \mathrm{E}[\Theta_2;\theta])\frac{\partial^2 \mathrm{E}[T(\theta); r\mu]}{\partial\theta^2}.$$

Since $\mathrm{E}[\Theta_1] = \mathrm{E}[\Theta_2] = \mu$ and the first partial of the limited expected value is bounded by unity as per Equation (3.3b), it follows that the first term vanishes and we have

$$\mu\cdot(\varphi_{T(\Theta_1)}(r) - \varphi_{T(\Theta_1)}(r))$$
$$= \int_0^\infty d\theta(\mathrm{E}[\Theta_1;\theta] - \mathrm{E}[\Theta_2;\theta])\frac{\partial^2 \mathrm{E}[T(\theta); r\mu]}{\partial\theta^2}.$$

We use the expectation formula for the insurance charge, A1, to arrive at the formula

$$\mathrm{E}[\Theta;\theta] = \mu\left(1 - \varphi_{T(\Theta)}\left(\frac{\theta}{\mu}\right)\right).$$

We then use this to substitute into the previous integral to yield

$$\varphi_{T(\Theta_1)}(r) - \varphi_{T(\Theta_2)}(r)$$
$$= \int_0^\infty d\theta\left(\varphi_{\Theta_2}\left(\frac{\theta}{\mu}\right) - \varphi_{\Theta_1}\left(\frac{\theta}{\mu}\right)\right)\frac{\partial^2 \mathrm{E}[T(\theta); r\mu]}{\partial\theta^2}.$$

The result then follows immediately from the assumptions of the proposition.

To gain a better understanding of the formulas, consider the following example.

EXAMPLE 5: Poisson Conditionals and Exponential Priors
Let $T(\theta)$ be Poisson. We leave it as an exercise for the reader to

show

$$\frac{\partial F(n \mid \theta)}{\partial \theta} = -f(n \mid \theta)$$

$$\frac{\partial \mathrm{E}[T(\theta);n]}{\partial \theta} = F(n-1 \mid \theta)$$

$$\frac{\partial^2 \mathrm{E}[T(\theta);n]}{\partial \theta^2} = -f(n-1 \mid \theta).$$

Suppose the prior on θ is an exponential with mean μ, so that $1 - H(\theta) = \exp(-\theta/\mu)$. Applying Equation (4.3), we derive

$$\varphi_{T(\Theta)}\left(\frac{n}{\mu}\right) = \int_0^\infty d\theta\, e^{-\theta/\mu} \cdot e^{-\theta} \frac{\theta^{n-1}}{(n-1)!} = \left(\frac{\mu}{\mu+1}\right)^n.$$

To see this is correct, we apply the prior to the conditional density and integrate to obtain the unconditional density

$$f_{T(\Theta)}(n) = \frac{1}{\mu+1}\left(\frac{\mu}{\mu+1}\right)^n.$$

We recognize this as a Geometric density. It is an exercise in summation formulas to then verify the insurance charge associated with this density is in fact the same as the one just derived using Equation (4.3).

We may now put the results from Equations (4.2) and (4.4) together to show that decreasing charges by risk size for the priors acting on a differentiable decomposable conditional family lead to unconditional charges that decrease with risk size.

4.5. Charges Decrease by Size for Model Based on Decomposable Conditionals with Priors that Decrease by Size

Suppose $\mathbf{M} = \{T(\theta) \mid \theta > 0\}$ is a differentiable decomposable risk-size model and let $\mathbf{Q}$ be a risk-size model with unique random variables, $\{\Theta_\mu\}$ such that $\mathrm{E}[\Theta_\mu] = \mu$.

If $\varphi_{\Theta_2} \leq \varphi_{\Theta_1}$ when $\mu_1 < \mu_2$, then

$$\varphi_{T(\Theta_2)} \leq \varphi_{T(\Theta_1)}. \tag{4.6}$$

Proof Let $\mu_2 = (1+c)\mu_1$ where $c > 0$ since $\mu_2 > \mu_1$. Via Equation (4.2) we have

$$\varphi_{T((1+c)\Theta_1)} \leq \varphi_{T(\Theta_1)}.$$

Since $\varphi_{\Theta_2} \leq \varphi_{\Theta_1} = \varphi_{(1+c)\Theta_1}$, we can use Equation (4.4) to show

$$\varphi_{T(\Theta_2)} \leq \varphi_{T((1+c)\Theta_1)}.$$

Note it is valid to apply Equation (4.4) since the second partial is negative for differentiable decomposable models via Equation (3.3c). Connecting the two inequalities leads to the desired result.

Next we extend these results to aggregate loss distributions.

5. INSURANCE CHARGES FOR LOSSES

We define an aggregate loss random variable as the compound distribution generated by selecting a claim count from a claim count distribution and then summing up that number of severities, where each claim severity is drawn from a severity distribution. We associate a risk with a particular count distribution and a particular severity distribution. When we talk about the aggregate losses for a risk, we mean the aggregate losses generated by samples appropriately drawn from its count and severity distributions according to our protocols. Suppose we have a collection of risks whose claim count distributions form a risk-size model. If we make appropriate assumptions about the severities of our risks, the aggregate-loss random variables for these risks will also constitute a risk-size model. We will show under certain conditions that, if the claim counts have charge functions that decrease by size of risk, then so do the charge functions for the aggregate losses.

Beginning the mathematical exposition, let N be the random variable representing the number of claims for a particular risk and let $p_N(n) = \Pr(N = n)$. Use a non-negative random variable X, with finite mean μ_X, to represent claim severity. Assume X

has finite variance and write $\tau^2 = \mathrm{Var}(X)$. Let X_i be the ith in a sequence of trials of X. Define the aggregate loss random variable for the risk via $T(N,X) = X_1 + X_2 + \cdots + X_N$. Note in this process of generating results, we are generating losses for a particular risk. When no confusion should result, we will write T instead of $T(N,X)$. Using Equation (A.6), we know the random variable, (T/μ_X), has the same insurance charge as T. This means we can assume $\mu_X = 1$ for the purpose at hand without loss of generality. When there are exactly n claims, define $T(n,X) = X_1 + X_2 + \cdots + X_n$. Assuming $\mu_X = 1$, it follows that $\mathrm{E}[T(n,X)] = n$. To simplify notation, we may write $\varphi_{n/X}$ or even φ_n in place of $\varphi_{T(n,X)}$.

The insurance charge for $T(N,X)$ can be decomposed as a weighted sum of the insurance charges for $T(n,X)$, each evaluated at an appropriately scaled entry ratio.

5.1. Count Decomposition of the Insurance Charge for Aggregate Loss

$$\varphi_{T(N,X)}(r) = \frac{1}{\mu_N} \sum_{n=1} p_N(n) \cdot n \cdot \varphi_{T(n,X)}\left(\frac{r\mu_N}{n}\right). \qquad (5.1)$$

Proof Left as an exercise for the reader.

While one could get some general results by working with the claim count decomposition and using discrete distribution analogues of integration by parts, the proofs are a bit messy. Instead, we will employ the simpler strategy of deriving properties of compound distributions from their claim count models. Then we can use the results of Chapters 3 and 4 to arrive at relatively painless conclusions about charges for the aggregate loss models. We start by proving that if the claim count model is decomposable, then so is the resulting compound distribution model. To do this, we will first make the following severity assumptions:

5.2. Fixed Independent Severity

A compound risk-size model has independent fixed severity if:

i) all risks share a common severity distribution, X.

ii) $\{X_1, X_2, \ldots, X_N\}$ is an independent set.

iii) X_i is independent of N. (5.2)

iv) X_i is independent of θ, where θ is the true mean of N for a risk.

Given these severity assumptions and a decomposable claim count model, we can show the aggregate loss model is also decomposable.

5.3. Aggregate Loss Model Inherits Decomposability from Claim Count Model Assuming Fixed Independent Severity

If $\mathbf{M}_N$ is a decomposable claim count model and the compound model, $\mathbf{M}_{T(N,X)} = \{T(N,X) \mid N \in \mathbf{M}_N\}$ has fixed independent severity, then $\mathbf{M}_{T(N,X)} = \{T(N,X) \mid N \in \mathbf{M}_N\}$ is also decomposable. (5.3)

Proof Recall we have assumed without loss of generality that $\mathrm{E}[X] = 1$. Thus $\mathrm{E}[T(N,X)] = \mathrm{E}[N]\mathrm{E}[X] = \mathrm{E}[N]$. Given $\theta > 0$, completeness of $\mathbf{M}_N$ implies there exists a unique $N(\theta) \in \mathbf{M}_N$ such that $\mathrm{E}[N(\theta)] = \theta$. It follows that $\mathrm{E}[T(N(\theta),X)] = \theta$. Thus, $\mathbf{M}_{T(N,X)}$ is complete. Now let $T(N(\theta_1),X)$ and $T(N(\theta_2),X)$ be in $\mathbf{M}_{T(N,X)}$. Then using our severity assumptions we can show $T(N(\theta_1),X) + T(N(\theta_2),X) = T(N(\theta_1) + N(\theta_2),X)$. Since $\mathbf{M}_N$ is closed under independent summation, it follows that $N(\theta_1) + N(\theta_2) = N(\theta_1 + \theta_2)$ and $N(\theta_1 + \theta_2) \in \mathbf{M}_N$. Therefore, $T(N(\theta_1) + N(\theta_2),X) \in \mathbf{M}_{T(N,X)}$, proving that $\mathbf{M}_{T(N,X)}$. is closed under independent summation. Now we apply Equation (3.2) to finish the proof.

Using this, it follows as a direct application of Equation (3.4) that the aggregate loss model has charges that decrease with risk size.

5.4. Decomposable Claim Counts Imply Aggregate Loss Model Has Charges that Decrease with Risk Size Assuming Fixed Independent Severity

If $\mathbf{M}_N$ is a differentiable decomposable risk-size model for claim counts and the compound model, $\mathbf{M}_T(\mathbf{N},X) = \{T(N,X) \mid N \in \mathbf{M}_N\}$ has fixed independent severity, then $\mathbf{M}_T(\mathbf{N},X)$ has charges that decrease with risk size. (5.4)

Proof Via Equation (5.2), $\mathbf{M}_{T(N,X)}$ is decomposable and the introduction of fixed independent severity does not affect differentiability with respect to risk size. The result then follows from Equation (3.4).

Note in this result that all risks share a common severity distribution and there is no parameter uncertainty regarding this severity distribution. While adding severity to the model does lead to larger charges for all risks, the result says that under these assumptions charges for aggregate loss still decline by size of risk.

We now apply Statement 5.4 to prove that insurance charges decrease with risk size for several classes of distributions commonly used in insurance models.

5.5. Charges for Aggregate Loss Decrease with Risk Size When Counts Are Poisson or Negative Binomial (Fixed Failure Rate)

Assuming claim sizes are independently and identically distributed and independent of the claim count, the insurance charge decreases as the size of a risk is increased in each of the following models, $\mathbf{M}_{T(N,X)}$, where $\mathbf{M}_N$ is:

Poisson:

$$\mathbf{M} = \{N \in \text{Poisson}(\mu) \mid \mu \geq 0\} \tag{5.5a}$$

Negative Binomial with common q:

$$\mathbf{M} = \{N \in \text{Negative Binomial}(\alpha, q) \mid q \text{ is fixed and } \alpha \geq 0\}. \tag{5.5b}$$

Proof Apply Statement 5.4.

We now introduce parameter uncertainty regarding the mean severity for a risk. We follow the CRM [4] as shown in Appendix B and assume severity may vary from risk to risk only due to a scale factor. Let β be a positive continuous random variable with density, $w(\beta)$, such that $\mathrm{E}[1/\beta] = 1$ and $\mathrm{Var}(1/\beta) = b$. The parameter b is called the mixing parameter.

We let X be a fixed severity distribution that does not change by risk. The severity distribution for a particular risk Y is obtained by first randomly selecting a β and then using the formula $Y = X/\beta$. Under this construction, each risk has a particular β that does not change from claim to claim. We will prove that when the selection of β is independent of θ, it follows that the compound model on decomposable counts has charges that decrease with size. To ensure clarity, we first define the notion of independent severity with scale parameter uncertainty by risk. This provides with us a terminology for describing the severity model just presented.

5.6. Independent Severity with Scale Parameter Uncertainty

A compound risk-size model has independent severity with scale parameter uncertainty if

i) each particular risk has a particular β and associated severity distribution, $Y = X/\beta$, where X is fixed for all risks and b is a positive continuous random variable with $\mathrm{E}[1/\beta] = 1$ and $\mathrm{Var}(1/\beta) = b$.

ii) $\{X_1, X_2, \ldots, X_N\}$ is an independent set.

iii) X_i is independent of N.

iv) The selection of β for a risk is independent of θ, where θ is the true mean of N for a risk.

v) The selection of β for a risk is independent of μ, where μ is the a priori mean of N for a risk. (5.6)

We now show

5.7. Decomposable Claim Counts Imply Aggregate Loss Model Has Charges that Decrease with Risk Size Assuming Independent Severity with Scale Parameter Uncertainty

If $\mathbf{M}_N$ is a differentiable decomposable risk-size model for claim counts and the compound model, $\mathbf{M}_{T(N,Y)} = \{T(N,Y) \mid N \in \mathbf{M}_N, Y = X/\beta\}$, has independent severity with scale parameter uncertainty, then $\mathbf{M}_{T(N,Y)}$ has charges that decrease with risk size. (5.7)

Proof By the usual Bayesian conditioning and using the independent severity assumptions, we can show the charge for the model at risk size θ is given via:

$$\varphi_{T(N(\theta),Y)}(r) = \int_0^{\infty} d\beta\, w(\beta)(1/\beta)\varphi_{T(N(\theta),X)}(r\beta).$$

Using Statement 5.5, we know the integrands decrease by size of risk, and the result follows using the independence of β and θ.

Even though the aggregate model in Statement 5.7 is based on decomposable counts, it will not be decomposable. Indeed, the aggregate model is not even a unique model as the introduction of the scale parameter leads to an infinite number of risks with the same a priori expected aggregate loss and thus the same size. We also need to be careful in interpreting the order in which the scaling parameter is averaged over the population. Suppose each risk in a decomposable count model has an exponential severity distribution and the prior on severity is a Gamma. It follows that the unconditional severity over all risks is Pareto distributed. If we then construct a model where each risk has this Pareto as its severity, we will have a decomposable model that is different and has different charges than the one in Statement 5.7.

Next we reprise the work done in Chapter 4 and extend our result to aggregate loss models in which the counts are subject to parameter uncertainty. We start with decomposable counts and then introduce a family of prior distributions on the mean claim counts, such that the priors constitute a risk-size model. If the priors have charges that decrease, not necessarily strictly, with risk size and if the compound model has independent severity with scaling parameter uncertainty, then the aggregate model has charges that decline with risk size. This is the key result of the paper.

5.8. Unconditional Aggregate Loss Model Charges Decrease with Risk Size Assuming Counts Based on Decomposable Conditionals with Priors that Decrease by Size and Independent Severity with Scaling Parameter Uncertainty

> Assume $\mathbf{M}_N$ is a differentiable decomposable claim-count model parameterized by θ such that $\mathrm{E}[N(\theta)] = \theta$. Let $\mathbf{Q} = \{\Theta_\mu\}$ be a complete set of priors on θ having charges that decrease with risk size. Let Y denote risk severity and suppose the aggregate loss model, $\mathbf{M}_T$, has independent severity with scale parameter uncertainty. Then $\mathbf{M}_T$ has charges that decrease with risk size. (5.8)

Proof The charge in the model for risks of size μ is given by

$$\varphi_{M_T(\mu)}(r) = \int_0^\infty d\beta\, w(\beta)(1/\beta)\frac{1}{\mu}\int_0^\infty d\theta\, h(\theta \mid \mu) \cdot \theta \cdot \varphi_{T(N(\theta),X)}(r\beta\mu/\theta).$$

Using the same integration by parts argument made in proving Statement 4.5, we can show the integral

$$\frac{1}{\mu}\int_0^\infty d\theta\, h(\theta \mid \mu)\cdot\theta\cdot\varphi_{T(N(\theta),X)}(r\beta\mu/\theta)$$

declines as a function of μ for any fixed β. The result then follows directly using the independence of β with μ.

Note the assumption of independence between β and θ allows us to integrate over the priors for severity and the priors for claim counts in any order. Thus we need our assumption that β is independent of θ as well as μ.

6. CONCLUSION

We started by proving some general inequalities for the insurance charge of a sum of identically distributed random variables. We used some numeric grouping techniques to show the key basic result that the charge for such a sum declines with the sample size. We then introduced the construct of a risk-size model. We showed that charges decline with size in a decomposable model. We introduced parameter risk with a family of Bayesian priors. We demonstrated that if the priors had decreasing charges by size and they acted on a decomposable conditional model, the resulting unconditional model has charges that decline by size of risk. We showed this to be true, even though the resulting unconditional models were not decomposable. Then we extended our results to aggregate loss models by adding severity and making some independence assumptions. We finally extended our result to aggregate loss models in which severity is subject to scale parameter risk. Though our final model is based on conditionally decomposable claim counts, the parameter risk on both counts and severity produce a model that is not decomposable.

The CRM is based on conditional Poisson counts and has severities that satisfy our independence assumptions. The parameter scale uncertainty in the CRM is the same as in our model. Thus the CRM satisfies the assumptions of our key result in Statement 5.8 and therefore it will generate charges that decline by size of risk. This is what we set out to prove.

The latest NCCI Table M was produced with the Gamma-Poisson claim-count model, where, to fit the data, the contagion declines with risk size [3]. This does not imply that the latest Table M is based on a decomposable model, but rather that the straight CRM model with constant contagion by size may lead to an overstatement of the charges for large risks.

In our unconditional model, we found charges were not forced to asymptotically approach the lowest possible charge function, $\varphi_0(r) = \max(0, 1 - r)$, as risk size tends to infinity. While we have shown $\varphi_0(r)$ is indeed the charge function for "an infinitely large risk" in a decomposable model, in our unconditional count model the charge for a very large risk approaches the charge for the prior of that risk.

Though severity increases insurance charges, the introduction of severity did not cause our size versus charge relation to fail. Intuitively this is because risk size is driven by the expected claim count. In short, severity does increase the insurance charge, but it does not change the relation between charge and size in the models we have developed here. While in the actual derivation of the latest Table M, severity did vary a bit by size in order to reconcile against the expected losses and fitted frequencies [3], the variation was not sufficient to cause inversions of the declining charge by size of risk relation. It is a topic of future research to understand how far severity assumptions may be relaxed before such inversions would occur.

So we conclude, having proved there is a fairly large class of risk-size models in which charges decline with size of risk. This class includes widely used actuarial models such as the CRM and Table M. Also, we have developed some theoretical constructs that should also provide a solid foundation for future research.

REFERENCES

[1] Bertoin, Jean, *Lévy Processes*, Cambridge, U.K.: Cambridge University Press, 1996, pp. 1–8, 187–214.

[2] Feller, William, *An Introduction to Probability Theory and Its Applications Volume I*, Third Edition, New York, NY: John Wiley and Sons, 1968, Chapter XVII, pp. 444–482.

[3] Gillam, W. R., "The 1999 Table of Insurance Charges," *PCAS* LXXXVII, 2000, pp. 188–218.

[4] Heckman, Phillip E. and Meyers, Glenn G., "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX, 1983, pp. 27–31.

[5] Hewitt, Jr., Charles C., "Loss Ratio Distribution—A Model," *PCAS* LIV, 1967, pp. 70–88.

[6] National Council on Compensation Insurance, *Retrospective Rating Plan Manual for Workers Compensation and Employers Liability Insurance*, 2001.

[7] Pittman, James W., "Levy Process and Infinitely Divisibility Law," Lecture Notes, Lecture 26, Spring 2003.

[8] Robbin, Ira, "Overlap Revisited: The Insurance Charge Reflecting Loss Limitation Procedure," *Pricing*, Casualty Actuarial Society Discussion Paper Program, 1990, Vol. II, pp. 809–850.

[9] Ross, Sheldon M., *Stochastic Processes*, Second Edition, New York, NY: John Wiley and Sons, 1996, pp. 41–162, 356–403.

APPENDIX A

BASIC INSURANCE CHARGE THEORY

Let T be a non-negative random variable having finite positive mean, μ, cumulative distribution function F, and tail probability function $G = 1 - F$. Define the normalized random variable R associated with T via $R = T/\mu$.

Perhaps the most compact mathematical definitions of the charge and saving can be given by taking the expected values of "min" and "max" operators.

A.1. Charge and Saving Functions Defined using Min and Max Expectations

Charge:

$$\varphi(r) = \frac{\mathrm{E}[\max(0, T - r\mu)]}{\mu} = \frac{\mathrm{E}[T - \min(T, r\mu)]}{\mu} = 1 - \frac{\mathrm{E}[T; r\mu]}{\mu} \tag{A.1a}$$

Saving:

$$\psi(r) = \frac{\mathrm{E}[\max(0, r\mu - T)]}{\mu} = r - \frac{\mathrm{E}[T; r\mu]}{\mu}. \tag{A.1b}$$

The definitions can be simplified even further by using the normalized random variable.

A.2. Charge and Saving Definitions using the Normalized Random Variable

Charge:

$$\varphi(r) = \mathrm{E}[\max(0, R - r)] = \mathrm{E}[R - \min(R, r)] = 1 - \mathrm{E}[R; r] \tag{A.2a}$$

Saving:

$$\psi(r) = \mathrm{E}[\max(0, r - R)] = r - \mathrm{E}[R; r]. \tag{A.2b}$$

The charge and saving can also be expressed in terms of integrals.

A.3. Insurance Charge and Saving Functions Defined using Integrals

$$\begin{aligned}\varphi(r) &= \frac{1}{\mu}\int_{r\mu}^{\infty} dF_T(t)(t-r\mu) = \int_{r}^{\infty} dF_R(s)(s-r)\\ &= \frac{1}{\mu}\int_{r\mu}^{\infty} dt\, G_T(t) = \int_{r}^{\infty} ds\, G_R(s) \qquad \text{(A.3a)}\\ \psi(r) &= \frac{1}{\mu}\int_{0}^{r\mu} dF_T(t)(r\mu-t) = \int_{0}^{r} dF_R(s)(r-s)\\ &= \frac{1}{\mu}\int_{0}^{r\mu} dt\, F_T(t) = \int_{0}^{r} ds\, F_R(s). \qquad \text{(A.3b)}\end{aligned}$$

When the random variable is discrete, these are viewed as Reimann integrals and are interpreted as sums.

Many basic properties can be proved directly from the definitions using simple properties of integrals, minimum operators, and expectations.

A.4. Insurance Charge: Basic Properties

φ is a continuous function of r. (A.4a)

φ is a decreasing function of r, which is strictly decreasing when $\varphi(r) > 0$. (A.4b)

$\varphi(0) = 1$ and, as $r \to \infty$, $\varphi(r) \to 0$. (A.4c)

$\varphi_0(r) \le \varphi(r) \le 1$, where $\varphi_0(r) = \max(0, 1-r)$. (A.4d)

With the definitions in A.2, one can show the charge and saving are related by a simple formula.

A.5. Relation Between Charge and Saving

$$\varphi(r) - \psi(r) = 1 - r. \tag{A.5}$$

Proof Use Equations (A.2a) and (A.2b) to write $\varphi(r) - \psi(r) = 1 - \mathrm{E}[R;r] - (r - \mathrm{E}[R;r]) = 1 - r$.

It is straightforward to see that multiplication of the underlying random variable by a scalar does not change its insurance charge.

A.6. Scaling a Random Variable Does Not Change Its Charge

For $c > 0$,

$$\varphi_T(r) = \varphi_{cT}(r). \tag{A.6}$$

Proof Left as an exercise for the reader.

The charge can never decline too rapidly between any two points.

A.7. Insurance Charge Slope Between Two Points Greater Than -1

If $s > r$ then

$$\frac{\varphi(s) - \varphi(r)}{s - r} \geq -1. \tag{A.7}$$

Proof Using Equation (A.4), write

$$\begin{aligned}\varphi(s) - \varphi(r) &= -\int_r^s du\, G_R(u) = \int_r^s du(-1 + F_R(u)) \\ &= -(s - r) + \int_r^s du\, F_R(u).\end{aligned}$$

Therefore,

$$\frac{\varphi(s) - \varphi(r)}{s - r} \geq -1 + \frac{1}{s - r}\int_r^s du\, F_R(u) \geq -1.$$

Further, the insurance charge must always be concave up. This means the insurance charge curve never goes above a straight line drawn between any two points on the curve.

A.8. Insurance Charge is Concave Up

$$\varphi(wr + (1-w)s) \leq w\varphi(r) + (1-w)\varphi(s) \qquad \text{for} \quad 0 \leq w \leq 1. \tag{A.8}$$

Proof Using the general property of the "min" operator

$$\min(A + B, C + D) \geq \min(A, C) + \min(B, D)$$

we derive

$$\begin{aligned}\min(R, wr + (1-w)s) \geq{}& \min(wR, wr)\\ &+ \min((1-w)R, (1-w)s).\end{aligned}$$

Factoring out w and $(1-w)$ respectively, yields

$$\min(R, wr + (1-w)s) \geq w \cdot \min(R, r) + (1-w)\min(R, s).$$

Using Equation (A.2) repeatedly, we find

$$\begin{aligned}\varphi(wr + (1-w)s) &= 1 - \mathrm{E}[R; wr + (1-w)s]\\ &\leq w + (1-w) - w\mathrm{E}[R, r] - (1-w)\mathrm{E}[R; s]\\ &= w(1 - \mathrm{E}[R; r]) + (1-w)(1 - \mathrm{E}[R; s])\\ &= w\varphi(r) + (1-w)\varphi(s).\end{aligned}$$

Though the charge and saving functions are continuous, they need not always be differentiable. However, when they are, the following formulas hold.

A.9. Insurance Charge and Saving First Derivatives

If φ or ψ is known to be differentiable at r, then

$$\frac{d\varphi}{dr}(r) = -G_R(r) \qquad \text{and} \qquad \frac{d\left(\varphi\left(\frac{t}{\mu}\right)\right)}{dt} = -\frac{G_T(t)}{\mu} \tag{A.9a}$$

$$\frac{d\psi}{dr}(r) = F_R(r) \qquad \text{and} \qquad \frac{d\left(\psi\left(\frac{t}{\mu}\right)\right)}{dt} = \frac{F_T(t)}{\mu}. \tag{A.9b}$$

Proof Apply Equation (A.4) and the Fundamental Theorem of Calculus.

If R has a density function at a point r, one can take second derivatives.

A.10. Insurance Charge and Saving Second Derivatives

If R has a well-defined density f_R at a point r, then

$$\frac{d^2\varphi}{dr^2}(r) = f_R(r) \tag{A.10a}$$

$$\frac{d^2\psi}{dr^2}(r) = f_R(r). \tag{A.10b}$$

Proof Take derivatives of the first derivatives shown in Equation (A.9a) and (A.9b).

The variance can be expressed in terms of an integral of the insurance charge.

A.11. Variance Formula using the Integral of the Insurance Charge

If $r^2 G_R(r) \to 0$ as $r \to \infty$, then

$$\mathrm{Var}(T) = \mu^2 \mathrm{Var}(R) = \mu^2 \left(2 \int_0^\infty dr\, \varphi(r) - 1 \right). \qquad \text{(A.11)}$$

Proof It suffices to prove the result when T is a continuous random variable. Integrate by parts twice and use the assumptions to derive

$$\int_0^\infty dr\, \varphi(r) = -rG_R(r)\Big|_{r=0}^{r=\infty} + \int_0^\infty dr\, rG_R(r)$$

$$= 0 + \frac{r^2}{2} G_R(r)\Big|_{r=0}^{r=\infty} + \int_0^\infty dr \frac{r^2}{2} f_R(r) = \frac{1}{2}\mathrm{E}[R^2].$$

Therefore, one has $2\int_0^\infty dr\, \varphi(r) = \mathrm{E}[R^2]$.

Then, using the definition of the variance along with the fact that $\mathrm{E}[R] = 1$, one can write

$$\mathrm{Var}(R) = \mathrm{E}[R^2] - (\mathrm{E}[R])^2 = 2 \int_0^\infty dr\, \varphi(r) - 1.$$

Note the coefficient of variation, CV, is given as the square root of $\mathrm{Var}(R)$.

The result, Equation (A.11), is intuitive since to have a large insurance charge a random variable must take on extreme values with some significant probability. This means it has a relatively large CV. The converse is not true. If $\mathrm{CV}(R_1) \geq \mathrm{CV}(R_2)$, then $\varphi_1(r)$ must exceed $\varphi_2(r)$ on average, but not necessarily at every entry ratio. Exhibit 1 shows a discrete counterexample in which one random variable has a larger charge at some entry ratios, even though it has a smaller CV than another random variable.

APPENDIX B

COLLECTIVE RISK MODEL SUMMARY

The quick summary uses notation that is equivalent to, but not always identical with, the notation used by Heckman and Meyers [4].

We start with the claim count model and define the number of claims as a counting random variable, N. Let θ be the conditional expected number of claims so that $\mathrm{E}[N|_\theta] = \theta$. We also write $N(\theta)$ to denote the conditional claim count distribution.

Let μ be the unconditional mean claim count. To introduce parameter uncertainty, we let ν be a non-negative random variable with $\mathrm{E}[\nu] = 1$ and $\mathrm{Var}(\nu) = c$. The parameter c is called the contagion. To model unconditional claim counts, we first select a value of ν at random and then randomly select a claim count N from the distribution $N(\theta)$ where $\theta = \nu\mu$. Heckman and Meyers assume N is conditionally Poisson, so that it follows that

B.1. Unconditional Claim Count Mean and Variance

$$\mathrm{E}[N] = \mathrm{E}[\mathrm{E}[N(\theta) \mid \theta = \nu\mu]] = \mathrm{E}[\nu\mu] = \mu\mathrm{E}[\nu] = \mu. \quad \text{(B.1a)}$$

$$\begin{aligned}\mathrm{Var}(N) &= \mathrm{E}[\mathrm{Var}(N(\theta) \mid \theta = \nu\mu)] + \mathrm{Var}(\mathrm{E}[N(\theta) \mid \theta = \nu\mu]) \\ &= \mathrm{E}[\mu\nu] + \mathrm{Var}(\mu\nu) = \mu\mathrm{E}[\nu] + \mu^2\mathrm{Var}(\nu) = \mu + \mu^2 c.\end{aligned} \quad \text{(B.1b)}$$

If ν is Gamma distributed, the unconditional claim count distribution is Negative Binomial.

We now add severity to the model. We let $X(\lambda)$ be the conditional claim severity random variable defined so that $\mathrm{E}[X(\lambda)] = \lambda$. Heckman and Meyers model severity parameter uncertainty by assuming the shape of the severity distribution is known but there is uncertainty about its scale. Let β be a positive random variable such that $\mathrm{E}[1/\beta] = 1$ and $\mathrm{Var}(1/\beta) = b$. We call b the mixing parameter. Suppose γ is the unconditional expected severity and let $\mathrm{Var}(X(\gamma)) = \tau^2$ To model severity, we first take a sample from β and then sample the scaled severity distribution $Y = X(\gamma)/\beta$.

To generate aggregate losses T, we first independently sample the number of claims N, and the scaling parameter, β. Then we independently draw N samples from the severity random variable $Y = X(\gamma)/\beta$. The aggregate loss T is the sum of these N severity samples.

Formulas for the unconditional mean and variance of the aggregate loss are derived as follows

B.2. Unconditional Aggregate Loss Mean and Variance

$$\mathrm{E}[T] = \mathrm{E}[N]\mathrm{E}[Y] = \mathrm{E}[\nu\mu]\mathrm{E}[X(\gamma)/\beta] = \mu\mathrm{E}[\nu]\gamma\mathrm{E}[1/\beta] = \mu\gamma. \tag{B.2a}$$

$$\begin{aligned}
\mathrm{Var}(T) &= \mathrm{E}[\mathrm{Var}(T|_{\nu},\beta)] + \mathrm{Var}(\mathrm{E}[T \mid \nu,\beta]) \\
&= \mathrm{E}[\mathrm{E}[N \mid \mu\nu]\mathrm{Var}(Y \mid \beta) + \mathrm{Var}(N \mid \mu\nu)\mathrm{E}[Y \mid \beta]^2] \\
&\quad + \mathrm{Var}(\mu\nu\gamma/\beta) \\
&= \mu\mathrm{E}[\nu]\tau^2\mathrm{E}[(1/\beta)^2] + \mu\mathrm{E}[\nu]\gamma^2\mathrm{E}[(1/\beta)^2] \\
&\quad + (\mu\gamma)^2\mathrm{Var}(\nu/\beta) \\
&= \mu\tau^2(1+b) + \mu\gamma^2(1+b) + (\mu\gamma)^2\mathrm{Var}(\nu/\beta) \\
&= \mu(\tau^2+\gamma^2)(1+b) + \mu^2\gamma^2(b+c+bc).
\end{aligned} \tag{B.2b}$$

Heckman and Meyers assume that β has a Gamma distribution.

EXHIBIT 1

Discrete Example
Smaller CV Does Not Imply Smaller Charge

Statistics	$T1$	$T2$
Mean	4.00	4.00
Variance	8.00	7.20
Relative Variance = Var(R)	0.50	0.45
CV	0.71	0.67

Random Variable $T1$

Index i	Point t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	20.0%	0.00	0.00	20.0%	0.0%	80.0%	100.0%
2	2.00	20.0%	4.00	0.50	40.0%	10.0%	60.0%	60.0%
3	4.00	20.0%	16.00	1.00	60.0%	30.0%	40.0%	30.0%
4	6.00	20.0%	36.00	1.50	80.0%	60.0%	20.0%	10.0%
5	8.00	20.0%	64.00	2.00	100.0%	100.0%	0.0%	0.0%
Mean	4.00		24.00	1.00				

Random Variable $T2$								
Index i	Point t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	0.0%	0.00	0.00	0.0%	0.0%	100.0%	100.0%
2	2.00	60.0%	4.00	0.50	60.0%	0.0%	40.0%	50.0%
3	4.00	10.0%	16.00	1.00	70.0%	30.0%	30.0%	30.0%
4	6.00	0.0%	36.00	1.50	70.0%	65.0%	30.0%	15.0%
5	8.00	30.0%	64.00	2.00	100.0%	100.0%	0.0%	0.0%
Mean	4.00		23.20	1.00				

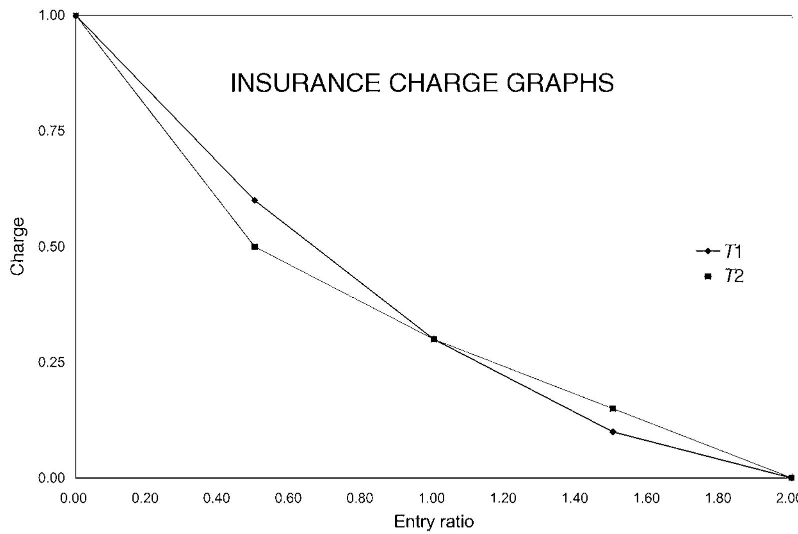


EXHIBIT 2

SHEET 1
CHARGE FOR INDEPENDENT SUM: DISCRETE EXAMPLE

Statistics	$T1$	$T2$	$T1+T2$
Mean	10.00	10.00	20.00
Variance	63.20	63.20	126.40
Relative Variance = Var(R)	0.63	0.63	0.32
CV	0.79	0.79	0.56

	$T1,T2$
Covariance	0.00
Correlation	0.00

Random Variable $T1$

Index i	$T1$ Value t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	20.0%	0.00	0.00	20.0%	0.0%	80.0%	100.0%
2	8.00	50.0%	64.00	0.80	70.0%	16.0%	30.0%	36.0%
3	16.00	20.0%	256.00	1.60	90.0%	72.0%	10.0%	12.0%
4	24.00	5.0%	576.00	2.40	95.0%	144.0%	5.0%	4.0%
5	32.00	5.0%	1,024.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	10.00	100.0%	163.20	1.00				

Random Variable $T2$

Index i	$T2$ Value t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	20.0%	0.00	0.00	20.0%	0.0%	80.0%	100.0%
2	8.00	50.0%	64.00	0.80	70.0%	16.0%	30.0%	36.0%
3	16.00	20.0%	256.00	1.60	90.0%	72.0%	10.0%	12.0%
4	24.00	5.0%	576.00	2.40	95.0%	144.0%	5.0%	4.0%
5	32.00	5.0%	1,024.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	10.00	100.0%	163.20	1.00				

Joint Density of $T1$ and $T2$

$T1$ Value	$T2$ Value 0.00	8.00	16.00	24.00	32.00	$T1$ Marginal
0.00	4.0%	10.0%	4.0%	1.0%	1.0%	20.0%
8.00	10.0%	25.0%	10.0%	2.5%	2.5%	50.0%
16.00	4.0%	10.0%	4.0%	1.0%	1.0%	20.0%
24.00	1.0%	2.5%	1.0%	0.3%	0.3%	5.0%
32.00	1.0%	2.5%	1.0%	0.3%	0.3%	5.0%
$T2$ Marginal	20.0%	50.0%	20.0%	5.0%	5.0%	100.0%

Random Variable $T1 + T2$

Index i	$T1 + T2$ t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	4.0%	0.00	0.00	4.0%	0.0%	96.0%	100.0%
2	8.00	20.0%	64.00	0.40	24.0%	1.6%	76.0%	61.6%
3	16.00	33.0%	256.00	0.80	57.0%	11.2%	43.0%	31.2%
4	24.00	22.0%	576.00	1.20	79.0%	34.0%	21.0%	14.0%
5	32.00	11.0%	1,024.00	1.60	90.0%	65.6%	10.0%	5.6%
6	40.00	7.0%	1,600.00	2.00	97.0%	101.6%	3.0%	1.6%
7	48.00	2.3%	2,304.00	2.40	99.3%	140.4%	0.8%	0.4%
8	56.00	0.5%	3,136.00	2.80	99.8%	180.1%	0.3%	0.1%
9	64.00	0.3%	4,096.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	20.00	100.0%	526.40	1.00				

EXHIBIT 2

SHEET 2
CHARGE FOR CORRELATED SUM: DISCRETE EXAMPLE

Statistics	$T1$	$T2$	$T1+T2$
Mean	10.00	10.00	20.00
Variance	63.20	63.20	201.60
Relative Variance = Var(R)	0.63	0.63	0.50
CV	0.79	0.79	0.71

	$T1,T2$
Covariance	37.60
Correlation	0.59

Random Variable $T1$

Index i	$T1$ Value t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	20.0%	0.00	0.00	20.0%	0.0%	80.0%	100.0%
2	8.00	50.0%	64.00	0.80	70.0%	16.0%	30.0%	36.0%
3	16.00	20.0%	256.00	1.60	90.0%	72.0%	10.0%	12.0%
4	24.00	5.0%	576.00	2.40	95.0%	144.0%	5.0%	4.0%
5	32.00	5.0%	1,024.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	10.00	100.0%	163.20	1.00				

Random Variable $T2$

Index i	$T2$ Value t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	20.0%	0.00	0.00	20.0%	0.0%	80.0%	100.0%
2	8.00	50.0%	64.00	0.80	70.0%	16.0%	30.0%	36.0%
3	16.00	20.0%	256.00	1.60	90.0%	72.0%	10.0%	12.0%
4	24.00	5.0%	576.00	2.40	95.0%	144.0%	5.0%	4.0%
5	32.00	5.0%	1,024.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	10.00	100.0%	163.20	1.00				

Joint Density of $T1$ and $T2$

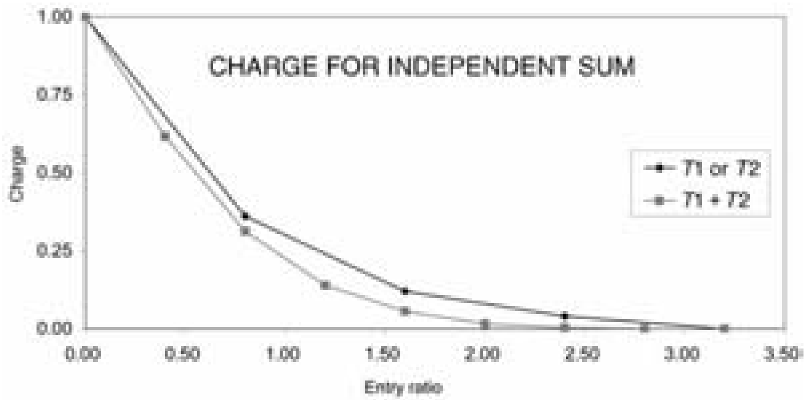
$T1$ Value	$T2$ Value 0.00	8.00	16.00	24.00	32.00	$T1$ Marginal
0.00	10.0%	7.0%	3.0%	0.0%	0.0%	20.0%
8.00	5.0%	37.0%	8.0%	0.0%	0.0%	50.0%
16.00	5.0%	5.0%	7.0%	2.0%	1.0%	20.0%
24.00	0.0%	0.0%	1.0%	2.0%	2.0%	5.0%
32.00	0.0%	1.0%	1.0%	1.0%	2.0%	5.0%
$T2$ Marginal	20.0%	50.0%	20.0%	5.0%	5.0%	100.0%

Random Variable $T1 + T2$

Index i	$T1+T2$ t_i	Density $f(t_i)$	Square t_i^2	Ratio r_i	CDF $F(r_i)$	Savings $\psi(r_i)$	Tail $G(r_i)$	Charge $\phi(r_i)$
1	0.00	10.0%	0.00	0.00	10.0%	0.0%	90.0%	100.0%
2	8.00	12.0%	64.00	0.40	22.0%	4.0%	78.0%	64.0%
3	16.00	45.0%	256.00	0.80	67.0%	12.8%	33.0%	32.8%
4	24.00	13.0%	576.00	1.20	80.0%	39.6%	20.0%	19.6%
5	32.00	7.0%	1,024.00	1.60	87.0%	71.6%	13.0%	11.6%
6	40.00	4.0%	1,600.00	2.00	91.0%	106.4%	9.0%	6.4%
7	48.00	4.0%	2,304.00	2.40	95.0%	142.8%	5.0%	2.8%
8	56.00	3.0%	3,136.00	2.80	98.0%	180.8%	2.0%	0.8%
9	64.00	2.0%	4,096.00	3.20	100.0%	220.0%	0.0%	0.0%
Mean	20.00	100.0%	601.60	1.00				

EXHIBIT 2

SHEET 3



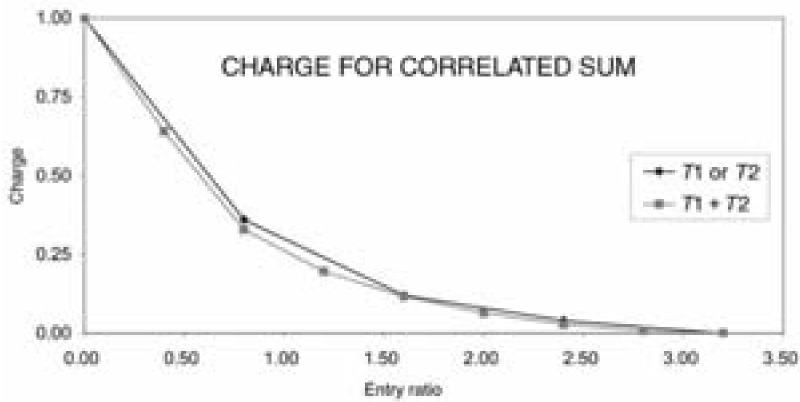


EXHIBIT 3

SHEET 1
POISSON LIMITED EXPECTED VALUES AND PARTIALS

	Limit = 3.000			
Mean	LEV	Numerical 1st Partial of LEV	Numerical 2nd Partial of LEV	Theoretical 2nd Partial of LEV
0.300	0.300	0.988		
0.600	0.596	0.959	–0.099	–0.099
0.900	0.884	0.910	–0.163	–0.165
1.200	1.157	0.845	–0.215	–0.217
1.500	1.410	0.770	–0.250	–0.251
1.800	1.641	0.690	–0.267	–0.268
2.100	1.848	0.609	–0.269	–0.270
2.400	2.031	0.531	–0.261	–0.261
2.700	2.191	0.458	–0.245	–0.245
3.000	2.328	0.391	–0.224	–0.224
3.300	2.445	0.330	–0.201	–0.201
3.600	2.544	0.277	–0.177	–0.177
3.900	2.627	0.231	–0.154	–0.154
4.200	2.697	0.191	–0.132	–0.132
4.500	2.754	0.158	–0.113	–0.112
4.800	2.801	0.129	–0.095	–0.095
5.100	2.840	0.105	–0.079	–0.079
5.400	2.872	0.085	–0.066	–0.066
5.700	2.897	0.069	–0.055	–0.054
6.000	2.918			

EXHIBIT 3

Sheet 2
Negative Binomial Limited Expected Values and Partials

Fixed $q = 0.750$

	Limit = 3.000		
Mean	LEV	Numerical 1st Partial of LEV	Numerical 2nd Partial of LEV
0.300	0.231	0.724	
0.600	0.448	0.679	−0.150
0.900	0.651	0.635	−0.146
1.200	0.842	0.592	−0.142
1.500	1.020	0.551	−0.136
1.800	1.185	0.512	−0.131
2.100	1.339	0.475	−0.125
2.400	1.481	0.439	−0.118
2.700	1.613	0.406	−0.112
3.000	1.734	0.374	−0.105
3.300	1.847	0.344	−0.099
3.600	1.950	0.317	−0.093
3.900	2.045	0.291	−0.087
4.200	2.132	0.266	−0.081
4.500	2.212	0.244	−0.075
4.800	2.285	0.223	−0.070
5.100	2.352	0.204	−0.064
5.400	2.413	0.186	−0.060
5.700	2.469	0.169	−0.055
6.000	2.520		

EXHIBIT 3

Sheet 3
Gamma Limited Expected Values and Partials

Scale = 1.000

	Limit = 3.000		
Mean	LEV	Numerical 1st Partial of LEV	Numerical 2nd Partial of LEV
0.300	0.294	0.981	
0.600	0.582	0.959	−0.074
0.900	0.860	0.927	−0.107
1.200	1.126	0.885	−0.141
1.500	1.376	0.833	−0.173
1.800	1.607	0.773	−0.200
2.100	1.819	0.707	−0.221
2.400	2.010	0.637	−0.234
2.700	2.180	0.565	−0.240
3.000	2.328	0.493	−0.237
3.300	2.455	0.425	−0.229
3.600	2.564	0.361	−0.214
3.900	2.654	0.302	−0.196
4.200	2.729	0.249	−0.176
4.500	2.790	0.203	−0.154
4.800	2.839	0.163	−0.133
5.100	2.877	0.129	−0.112
5.400	2.908	0.101	−0.093
5.700	2.931	0.079	−0.076
6.000	2.949	0.060	

EXHIBIT 4

Sheet 1
Poisson Insurance Charges

Entry Ratio	Mean					
	0.500	1.000	1.500	2.000	2.500	3.000
0.000	1.000	1.000	1.000	1.000	1.000	1.000
0.100	0.961	0.937	0.922	0.914	0.908	0.905
0.200	0.921	0.874	0.845	0.827	0.816	0.810
0.300	0.882	0.810	0.767	0.741	0.725	0.715
0.400	0.843	0.747	0.689	0.654	0.633	0.630
0.500	0.803	0.684	0.612	0.568	0.562	0.550
0.600	0.764	0.621	0.534	0.508	0.490	0.470
0.700	0.725	0.558	0.467	0.449	0.419	0.397
0.800	0.685	0.494	0.423	0.389	0.348	0.339
0.900	0.646	0.431	0.379	0.330	0.302	0.282
1.000	0.607	0.368	0.335	0.271	0.257	0.224
1.100	0.567	0.341	0.290	0.238	0.211	0.189
1.200	0.528	0.315	0.246	0.206	0.165	0.153
1.300	0.488	0.289	0.202	0.174	0.141	0.118
1.400	0.449	0.262	0.175	0.141	0.117	0.094
1.500	0.410	0.236	0.155	0.109	0.093	0.076
1.600	0.370	0.209	0.136	0.095	0.068	0.057
1.700	0.331	0.183	0.117	0.080	0.057	0.042
1.800	0.292	0.156	0.098	0.066	0.047	0.034
1.900	0.252	0.130	0.079	0.052	0.036	0.025
2.000	0.213	0.104	0.060	0.038	0.025	0.017
2.100	0.204	0.096	0.053	0.032	0.021	0.014
2.200	0.195	0.088	0.047	0.027	0.016	0.010
2.300	0.186	0.080	0.040	0.022	0.012	0.007
2.400	0.177	0.072	0.034	0.017	0.008	0.005
2.500	0.168	0.063	0.027	0.011	0.007	0.004
2.600	0.159	0.055	0.020	0.010	0.005	0.003
2.700	0.150	0.047	0.015	0.008	0.004	0.002
2.800	0.141	0.039	0.014	0.006	0.002	0.001
2.900	0.132	0.031	0.012	0.005	0.002	0.001
3.000	0.123	0.023	0.010	0.003	0.001	0.000

EXHIBIT 4

Sheet 2
Negative Binomial Insurance Charges

Fixed $q = 0.750$

Entry Ratio	Mean					
	0.500	1.000	1.500	2.000	2.500	3.000
0.000	1.000	1.000	1.000	1.000	1.000	1.000
0.100	0.979	0.963	0.950	0.940	0.931	0.925
0.200	0.959	0.926	0.900	0.879	0.863	0.850
0.300	0.938	0.889	0.850	0.819	0.794	0.775
0.400	0.917	0.852	0.800	0.759	0.726	0.713
0.500	0.897	0.815	0.750	0.698	0.677	0.656
0.600	0.876	0.778	0.700	0.658	0.628	0.600
0.700	0.856	0.741	0.656	0.617	0.580	0.548
0.800	0.835	0.704	0.625	0.577	0.531	0.506
0.900	0.814	0.667	0.594	0.537	0.495	0.464
1.000	0.794	0.630	0.562	0.496	0.460	0.422
1.100	0.773	0.609	0.531	0.468	0.425	0.390
1.200	0.752	0.587	0.500	0.440	0.390	0.359
1.300	0.732	0.566	0.469	0.412	0.364	0.327
1.400	0.711	0.545	0.445	0.384	0.338	0.301
1.500	0.691	0.524	0.424	0.356	0.313	0.277
1.600	0.670	0.502	0.403	0.336	0.287	0.253
1.700	0.649	0.481	0.382	0.316	0.268	0.231
1.800	0.629	0.460	0.362	0.296	0.249	0.214
1.900	0.608	0.439	0.341	0.276	0.230	0.196
2.000	0.587	0.417	0.320	0.257	0.212	0.178
2.100	0.577	0.404	0.306	0.243	0.198	0.165
2.200	0.566	0.391	0.292	0.228	0.184	0.151
2.300	0.555	0.377	0.278	0.214	0.170	0.138
2.400	0.545	0.364	0.264	0.200	0.156	0.127
2.500	0.534	0.351	0.250	0.186	0.146	0.117
2.600	0.523	0.337	0.236	0.176	0.136	0.107
2.700	0.512	0.324	0.223	0.166	0.126	0.098
2.800	0.502	0.310	0.213	0.156	0.116	0.090
2.900	0.491	0.297	0.203	0.146	0.108	0.083
3.000	0.480	0.284	0.194	0.135	0.101	0.075

EXHIBIT 4

Sheet 3
Gamma Insurance Charges

Scale = 1.000

Entry Ratio	Mean					
	0.500	1.000	1.500	2.000	2.500	3.000
0.000	1.000	1.000	1.000	1.000	1.000	1.000
0.100	0.917	0.905	0.902	0.901	0.900	0.900
0.200	0.847	0.819	0.809	0.804	0.802	0.801
0.300	0.785	0.741	0.723	0.713	0.708	0.705
0.400	0.729	0.670	0.644	0.629	0.620	0.614
0.500	0.679	0.607	0.572	0.552	0.539	0.530
0.600	0.633	0.549	0.507	0.482	0.465	0.453
0.700	0.591	0.497	0.449	0.419	0.399	0.384
0.800	0.553	0.449	0.397	0.363	0.340	0.323
0.900	0.517	0.407	0.350	0.314	0.289	0.270
1.000	0.484	0.368	0.308	0.271	0.244	0.224
1.100	0.453	0.333	0.271	0.233	0.205	0.185
1.200	0.425	0.301	0.239	0.200	0.172	0.152
1.300	0.399	0.273	0.210	0.171	0.144	0.124
1.400	0.374	0.247	0.184	0.146	0.120	0.101
1.500	0.351	0.223	0.161	0.124	0.100	0.082
1.600	0.330	0.202	0.142	0.106	0.083	0.066
1.700	0.310	0.183	0.124	0.090	0.068	0.053
1.800	0.291	0.165	0.108	0.077	0.056	0.043
1.900	0.274	0.150	0.095	0.065	0.046	0.034
2.000	0.258	0.135	0.083	0.055	0.038	0.027
2.100	0.243	0.122	0.073	0.046	0.031	0.022
2.200	0.228	0.111	0.063	0.039	0.026	0.017
2.300	0.215	0.100	0.055	0.033	0.021	0.014
2.400	0.202	0.091	0.048	0.028	0.017	0.011
2.500	0.191	0.082	0.042	0.024	0.014	0.009
2.600	0.180	0.074	0.037	0.020	0.011	0.007
2.700	0.169	0.067	0.032	0.017	0.009	0.005
2.800	0.160	0.061	0.028	0.014	0.007	0.004
2.900	0.150	0.055	0.024	0.012	0.006	0.003
3.000	0.142	0.050	0.021	0.010	0.005	0.003

EXHIBIT 5

CHARGES FOR GAMMA-POISSON CONTAGION MODEL

$q = \mu/(\mu + \alpha)$, $\alpha = 2.00$, Contagion = .50

Entry Ratio	Mean					
	0.500	1.000	1.500	2.000	2.500	3.000
0.000	1.000	1.000	1.000	1.000	1.000	1.000
0.100	0.964	0.944	0.933	0.925	0.920	0.916
0.200	0.928	0.889	0.865	0.850	0.840	0.832
0.300	0.892	0.833	0.798	0.775	0.759	0.748
0.400	0.856	0.778	0.731	0.700	0.679	0.677
0.500	0.820	0.722	0.663	0.625	0.621	0.612
0.600	0.784	0.667	0.596	0.575	0.562	0.547
0.700	0.748	0.611	0.538	0.525	0.504	0.488
0.800	0.712	0.556	0.499	0.475	0.446	0.441
0.900	0.676	0.500	0.459	0.425	0.406	0.393
1.000	0.640	0.444	0.420	0.375	0.366	0.346
1.100	0.604	0.419	0.380	0.344	0.326	0.312
1.200	0.568	0.393	0.341	0.313	0.286	0.278
1.300	0.532	0.367	0.302	0.281	0.259	0.245
1.400	0.496	0.341	0.274	0.250	0.233	0.218
1.500	0.460	0.315	0.253	0.219	0.206	0.194
1.600	0.424	0.289	0.232	0.200	0.180	0.171
1.700	0.388	0.263	0.210	0.181	0.163	0.150
1.800	0.352	0.237	0.189	0.163	0.146	0.134
1.900	0.316	0.211	0.168	0.144	0.129	0.119
2.000	0.280	0.185	0.146	0.125	0.112	0.103
2.100	0.270	0.174	0.135	0.114	0.101	0.092
2.200	0.259	0.163	0.124	0.103	0.090	0.081
2.300	0.249	0.152	0.113	0.092	0.079	0.071
2.400	0.238	0.141	0.102	0.081	0.069	0.062
2.500	0.228	0.130	0.091	0.070	0.062	0.055
2.600	0.218	0.119	0.080	0.064	0.055	0.048
2.700	0.207	0.107	0.070	0.058	0.048	0.042
2.800	0.197	0.096	0.065	0.052	0.042	0.037
2.900	0.186	0.085	0.059	0.045	0.038	0.033
3.000	0.176	0.074	0.054	0.039	0.033	0.028