Understanding Split Credibility

Ira Robbin, PhD
Assistant Vice-President and Senior Pricing Actuary
Endurance US Insurance Operations

Abstract:
Under an experience-rating plan with a primary-excess split, losses are divided into per occurrence primary and excess components. Then an estimate of total loss is obtained by adding together credibility-weighted estimates of these components. But why should this splitting procedure lead to a better estimate? As proved in this paper, the general answer is that a split will improve on a no-split plan only if it produces a proportionately different allocation of process and parameter risk between the components. When that happens, credibility rises on the component that gets the relatively greater portion of parameter risk and falls on the other one. Unfortunately, a primary-excess split does not always lead to such a differential allocation. This is demonstrated in the paper using a Count-Severity model of loss based on conditionally Poisson claim counts, conditionally exponential severities, and Gamma priors. In some cases the split yields no improvement at all and in others it produces a primary-layer credibility that is smaller than the excess-layer credibility. Such an inversion of primary and excess credibilities can never happen under the Workers Compensation split rating plan that has been successfully used in the United States. The paper explains why these contrary results might occur by analyzing parameter variance relations between small and large claims that follow from the Count-Severity model. The disquieting conclusion is that split credibility will struggle to be effective if losses are modeled with a single claim count distribution and a single claim severity distribution that is subject to scale parameter uncertainty. The paper ends with brief discussion of areas for future research.
Keywords: Credibility, Experience Rating

1. INTRODUCTION

In an experience-rating plan with a primary-excess split, such as that promulgated by the National Council on Compensation Insurance (NCCI) [5] for rating Workers Compensation risks in the United States[1], losses are divided into per occurrence primary and excess components. An experience-adjusted estimate of total loss is then obtained by adding together credibility-weighted estimates of these components[2]. Since the splitting of losses and the introduction of separate credibility weights entails additional complication and expense, it is not unreasonable to ask what, if anything, is gained by use of a split plan versus a no-split plan.

To answer this question, we start by examining two widely accepted, but contradictory, intuitions. One is that a well-chosen primary-excess split could produce two components, each relatively less volatile and thus, presumably, more predictable than the total[3] . The contrary belief is that the excess layer loss is inherently more volatile than the total loss[4]. It turns out this is correct: excess loss has greater relative variability than primary loss. To be more precise, as we show in the Appendix, the excess loss has a process risk coefficient of variation (CV) at least as large as the corresponding CV for total loss. So there is no way a primary-excess split can produce two less volatile components.

Further, if we step back a bit to gain perspective, we see that the entire discussion of volatility is insufficient on its own to lead to any definitive conclusions about credibility. To put it succinctly, "low volatility" is not synonymous with "high credibility". Rather, credibility is conceptually the weight given to observed data, as opposed to the weight given to prior belief. It depends, not only on volatility (process risk), but also on the uncertainty in our initial belief (parameter risk). Therefore, when we try to assess credibility in a split plan, we need to examine not only the process variances, but also the parameter variances of each component. As well, when analyzing these components, it is insufficient to consider them in isolation. We are trying to estimate the total loss as the sum of estimates of these two components. To understand what causes inaccuracy in our final estimate, we cannot ignore their covariances[5].

Once process risk, parameter risk, and covariance are all included, we see there is no way to split losses so as to: i) produce two components that <u>each</u> have a higher credibility than the original total, and ii) produce credibility weighted estimates which when added together produce a more accurate estimate of the total. Intuitively, any such result can only occur if some of the process risk or process covariance disappears in the splitting process. Of course, no part of the total variance disappears: it is instead allocated to the components or to their covariances. Based on the realization that splitting leads to allocations of the process and parameter variances to the component variances and covariances, our understanding is that an effective split credibility plan necessarily entails a trade-off in which one component gets a higher credibility and the other a lower credibility than the total. The key to achieving an effective plan is to define components so as to separate a less predictable portion of losses from a more predictable one. The less predictable portion is then assigned a lower credibility, and the more predictable portion, a higher one. The split is effective when it allows us to raise credibility on a portion of the loss that is more predictable and to reduce credibility on a portion that is less so. To put it another way, a split will work if it helps us concentrate on the signal and ignore the noise.

To show that this intuition is in fact correct, we will start first by examining a no-split experience-rating plan. We will develop the important formula that the use of optimal credibility reduces the expected square error of the estimate of the mean (the parameter risk) by a ratio equal to that optimal credibility value. For example, if the optimal credibility is 40% and the original expected parameter variance is 100, the credibility-weighted estimate will have an expected square error of 60.

We will then turn to an arbitrary split plan, where the split is any way of dividing losses. We will derive optimal credibility formulas. Our formulas are equivalent to formulas previously presented by Mahler [3], though our notation is different. We will then study the reduction in mean square error when optimal credibility values are used. This will lead to a split credibility version of the error reduction formula. In the split model formula, we first allocate the original total parameter variance to the components. Each component gets its own variance plus the covariance. When optimal credibilities are used, the square error for each component is reduced by its credibility. For example, if the parameter variances are 60 and 20 respectively for two components that have a covariance of 10, the parameter variance for the total will be 100. The allocations of the total are therefore 70 (60+10) and 30 (20+10). If the optimal credibilities are 60% and 10%, then the mean square error is 70*(100%-60%) + 30*(100%-10%) = 28+27 = 55.

To study, in general, what might be gained by adopting a split plan, we will take the difference in the square errors of the optimal non-split and split plans. Based on the resulting formula, we will show split credibility is most effective is when the two components have relatively different proportions of process and parameter risk. The component with the relatively larger proportion of parameter risk than process risk ends up with a credibility larger than the credibility for the un-split losses.

Armed with this general understanding, we will explore primary-excess split plans. If losses are not split, the volatility arising from large claims will usually generate significant process risk. When a split is introduced, most of that process risk gets allocated to the excess layer. However, we can say nothing about whether the split is effective unless we also know how the parameter risk gets allocated. That in turn depends on the structure of the loss model and its priors. With an arbitrary loss model and arbitrary priors, there is no reason the split could not allocate a proportion of the parameter risk that is smaller to, equal to, or greater than the proportion of the process risk allocated to the excess layer. As a result, we arrive at the disappointing conclusion that introducing a split does not generally improve accuracy. Further, in general, there is nothing to prevent layer credibility inversions: the excess layer could have more credibility than the primary layer.

We will present a Claim Count-Claim Severity model as a specific example of a model in which such contrary results can occur. In the model, claim counts are assumed to be conditionally Poisson with a Gamma prior. The claim severities are conditionally exponential and also have a Gamma prior. This model of losses is not new or unusual; it is an application of the Collective Risk Model (CRM) presented by Heckman and Meyers [1]. Mahler [3] also derived credibility equations for it.

To gain insight into the interaction between split rating and the Gamma-Poisson, Gamma-Exponential model, we will analyze what the model implies about the parameter variance relations between the number of Small and Large claims. These groupings are defined by whether a claim is below or above the split point. We will find in most cases of interest that the model forces the contagion for the number of Large claims to be greater than the contagion for the number of Small ones. The model also leads to a significant correlation between the expected number of Small and Large claims. From this we can see that embedded in the structure of the model is a tendency to allocate relatively more parameter risk to the excess layer. This makes it harder to achieve the differential allocation of process and parameter risk by layer that is needed to make split credibility effective. We will generalize and assert that similar problems would likely be encountered whenever the loss model is based on a single claim count random variable and a single claim severity random variable, where the prior on severity reflects only uncertainty about the overall scale of that distribution. We will then speculate that a Mixed Claim Type model, one having different types of claims with different severities, might allow us to relax the tendency to allocate relatively large amounts of parameter risk to the excess layer. Finally we will suggest future work to fully implement the Small Claim-Large Claim approach, to develop a Mixed Claim Type model, and to analyze Workers Compensation losses.

2. NO-SPLIT CREDIBILITY

We start with a general no-split plan. Let A be the random variable representing actual historical loss. We suppose A is dependent on a possibly multi-dimensional parameter,θ, and define $\mu(\theta) = E[A(\theta)]$ and $\sigma^2(\theta) = Var(A(\theta))$. Let h be the prior distribution of θ and use h to define $E = E[\mu(\theta)]$, $\sigma^2=E[\sigma^2(\theta)]$ and $\tau^2=Var(\mu(\theta))$. Under this notation, σ^2 is a measure of the process risk and τ^2 is a measure of parameter risk. We also set $\lambda^2 = \sigma^2 + \tau^2$ so that λ^2 is the total variance of A.

In this construction each risk has a particular θ value that we have no way of knowing in advance. Our initial knowledge is only of the distribution of θ. Given an observation of A, we could use Bayes

Theorem to obtain the posterior distribution, h(θ|A). From this, we could in principle compute the conditional expected value, E[μ(θ)|A], and define the mod, M, via M = E[μ(θ)|A]/E, where E is the a priori mean. However, the conditional expected value may be difficult to compute. So, what actuaries have done as a practical matter is to find the best linear estimator. The formula is:

$$M = z\frac{A}{E} + (1-z)$$

There are several ways to define what is meant by "best", but, for our purposes, it suffices to use the most common definition and define an estimator as best if it produces the least mean square error. It is

$$z^* = \frac{\tau^2}{\tau^2 + \sigma^2} = \frac{\tau^2}{\lambda^2}$$

well known that the linear estimator with minimal mean square error is given by using credibility:

For comparison purposes, we also want to know the mean square error that is produced when this optimal z is used. We first derive a formula for the mean square error. To simplify the derivation, we will omit the "*" denoting that credibility is optimal. We write:

$$\begin{aligned}
\varepsilon^2 &= E[\left(zA + (1-z)E - \mu(\theta)\right)^2] \\
&= z^2 \cdot E[\left(A - \mu(\theta)\right)^2] + (1-z)^2 \cdot E[\left(E - \mu(\theta)\right)^2] \\
&= z^2\sigma^2 + (1-z)^2\tau^2
\end{aligned}$$

In obtaining this expression, various cross terms vanish under the assumption that the sampling deviation of actual results from the mean for a risk is independent of the deviation of the risk mean from the population mean. Specifically, we have assumed that:

$$E[(A-\mu(\theta))(E-\mu(\theta))]=0$$

When we plug in the optimal z, we find the minimum mean square error for the no-split linear estimator is given as:

$$\varepsilon_0^2(NS)=\left(\frac{\tau^2}{\tau^2+\sigma^2}\right)^2\cdot\sigma^2+\left(\frac{\sigma^2}{\tau^2+\sigma^2}\right)^2\cdot\tau^2=\left(\frac{\tau^2\sigma^2}{\left(\tau^2+\sigma^2\right)^2}\right)\cdot(\tau^2+\sigma^2)=\frac{\tau^2\sigma^2}{\left(\tau^2+\sigma^2\right)}$$

The "NS" label stands for "No-Split". Observe this minimal error can be written as:

$$\varepsilon_0^2(NS)=\frac{\tau^2\sigma^2}{\tau^2+\sigma^2}=\tau^2\left(1-\frac{\tau^2}{\lambda^2}\right)=\tau^2\left(1-z^*\right)$$

Here we have reintroduced the notation z* to indicate that the formula is only valid when the optimal credibility is used. Since the initial square error before any observations are made is τ^2, this equation says use of the optimal credibility value reduces the initial mean square error by a proportion that is equal to that optimal credibility value.

3. GENERAL SPLIT PLAN CREDIBILITIES

Assume A can be written as the sum of two loss random variables: $A= A_1 +A_2$. In this generality, the split is not necessarily between primary and excess losses: it could be any way of splitting losses. We suppose each A_i is dependent on a possibly multi-dimensional parameter, θ, and define $\mu_i(\theta) = E[A_i(\theta)]$ and $\sigma_i^2(\theta) = Var(A_i(\theta))$. Also let $C(\theta) = Cov(A_1(\theta),A_2(\theta))$. Assume h is the prior distribution of θ and use h to define $E_i =E[\mu_i(\theta)]$, $\sigma_i^2=E[\sigma_i^2(\theta)]$, $\rho= E[C(\theta)]$, $\tau_i^2=Var(\mu_i(\theta))$, $\lambda_i^2 =\sigma_i^2 + \tau_i^2$, and $\pi = Cov(\mu_1(\theta),\mu_2(\theta))$. Note that, in addition to the process and parameter risk terms for each loss component, we have also defined expected process covariance and parameter covariance terms. Set $\kappa = \rho +\pi$ so that κ is the total covariance. Define σ^2 as the total process variance, τ^2 as the total parameter variance, and λ^2 as

the total variance. We observe that $\sigma^2 = \sigma_1^2 + \sigma_2^2 + 2\rho$, $\tau^2 = \tau_1^2 + \tau_2^2 + 2\pi$, and $\lambda^2 = \lambda_1^2 + \lambda_2^2 + 2\kappa$. As before, we will find the least mean square error linear estimator. The Mod formula is:

$$M = \frac{z_1 A_1 + (1 - z_1)E_1 + z_2 A_2 + (1 - z_2)E_2}{E}$$

We now derive a formula for the mean square error:

$$\begin{aligned}
\varepsilon^2 &= E[\left(z_1 A_1 + (1 - z_1)E_1 - \mu_1(\theta) + z_2 A_2 + (1 - z_2)E_2 - \mu_2(\theta)\right)^2] \\
&= z_1^{\ 2} \cdot E[\left(A_1 - \mu_1(\theta)\right)^2] + (1 - z_1)^2 \cdot E[\left(E_1 - \mu_1(\theta)\right)^2] \\
&+ z_1^{\ 2} \cdot E[\left(A_2 - \mu_2(\theta)\right)^2] + (1 - z_2)^2 \cdot E[\left(E_2 - \mu_2(\theta)\right)^2] \\
&+ 2 z_1 z_2 E[C(\theta)] + 2(1 - z_1)(1 - z_2) Cov(\mu_1(\theta), \mu_2(\theta))
\end{aligned}$$

In obtaining this expression, we have assumed that the sampling deviation of actual results from the mean for each variable is independent of the deviation of the risk mean from the population mean for both variables. This implies that various cross terms vanish. Note that the square error formula has other terms do not vanish but which depend on the process and parameter covariance. These terms are present in Mahler's formula, though in different notation. Simplifying to express the terms as polynomials of the credibilities, we have:

$$\varepsilon^2 = \tau^2 + z_1^2 \lambda_1^2 - 2 z_1 (\tau_1^2 + \pi) + z_2^2 \lambda_2^2 - 2 z_2 (\tau_2^2 + \pi) + 2 z_1 z_2 \kappa$$

We take partials with respect to the credibility parameters:

$$\begin{aligned}
\frac{\partial \varepsilon^2}{\partial z_1} &= 2 z_1 \lambda_1^2 - 2(\tau_1^2 + \pi) + 2 z_2 \kappa \\
\frac{\partial \varepsilon^2}{\partial z_2} &= 2 z_2 \lambda_2^2 - 2(\tau_2^2 + \pi) + 2 z_1 \kappa
\end{aligned}$$

Setting the partials equal to zero, we obtain the system of equations:

$$z_1\lambda_1^2 + z_2\kappa = (\tau_1^2 + \pi)$$
$$z_2\lambda_2^2 + z_1\kappa = (\tau_2^2 + \pi)$$

Solving we find:

$$z_1 = \frac{\lambda_2^2(\tau_1^2 + \pi) - \kappa(\tau_2^2 + \pi)}{D}$$
$$z_2 = \frac{\lambda_1^2(\tau_2^2 + \pi) - \kappa(\tau_1^2 + \pi)}{D}$$
$$\text{where } D = \lambda_1^2\lambda_2^2 - \kappa^2$$

As Mahler [3] noted, these are not really credibilities in the traditional sense, since in this generality one of them could be negative or have a value above unity. It can be shown with some messy high-school level algebra that the minimal mean square error for the split plan is given as:

$$\varepsilon_0^2(SP) = \tau^2 - \frac{1}{D}\left(\lambda_2^2(\tau_1^2 + \pi) + \lambda_1^2(\tau_2^2 + \pi) - 2\kappa(\tau_1^2 + \pi)(\tau_2^2 + \pi)\right)$$

This can further be reduced to:

$$\varepsilon_0^2(SP) = (\tau_1^2 + \pi)(1 - z_1^*) + (\tau_2^2 + \pi)(1 - z_2^*)$$

Here we have reintroduced the "*" denoting optimal credibility to emphasize that the formula is only valid when optimal credibility values are used. The “SP” indicates the formula is for a split plan. This formula extends the error reduction formula from the no-split case. The initial mean square error is the parameter risk. It is split, with each component taking its own parameter variance and the parameter covariance. Since the total portion of the parameter variance is two times the parameter covariance, each component is allocated half of the parameter covariance contribution to the total parameter variance. These allocations are then reduced in proportion to the respective optimal credibility values.

4. WHEN DOES THE SPLIT REDUCE MEAN SQUARE ERROR?

To study whether a split plan reduces minimum mean square error, we first define the reduction in minimal mean square error as the difference between the mean square errors: $\Delta(\varepsilon_0^2) = \varepsilon_0^2(\text{NS}) - \varepsilon_0^2(\text{SP})$. We find:

$$\begin{aligned}\Delta\varepsilon_0^2 &= \tau^2(1-z^*) - (\tau_1^2+\pi)(1-z_1^*) - (\tau_2^2+\pi)(1-z_2^*)\\ &= (\tau_1^2+\pi)(z^*-z_1^*) - (\tau_2^2+\pi)(z^*-z_2^*)\end{aligned}$$

We immediately see that error improvement requires that one of the split plan credibility values to be larger than the credibility from the original no-split plan. In this generality, where the split is arbitrary and not necessarily between primary and excess losses, there is no reason why the split plan should reduce mean square error.

We will use this equation to derive an intuitively accessible formula that summarizes what is required for a split to reduce the minimal mean square error. But, before proving our main result, it is useful to consider a simple example to hone our intuition.

Example

Consider a split plan where the two components have the same process variance, the same parameter variance, and perfect parameter correlation. In other words, assume $\sigma_1^2 = \sigma_2^2$, $\tau_1^2 = \tau_2^2$ and $\pi = \tau_1\tau_2$. It follows that $\lambda_1^2 = \lambda_2^2$ and that $z_1^* = z_2^*$. We derive:

$$z_1^* = \frac{\lambda_2^2(\tau_1^2+\pi) - \kappa(\tau_2^2+\pi)}{\lambda_1^2\lambda_2^2 - \kappa^2} = \frac{\lambda_1^2(\tau_1^2+\pi) - \kappa(\tau_1^2+\pi)}{(\lambda_1^2-\kappa)(\lambda_1^2+\kappa)} = \frac{(\tau_1^2+\pi)}{(\lambda_1^2+\kappa)}$$

$$z^* = \frac{\tau^2}{\lambda^2} = \frac{\tau_1^2+\tau_2^2+2\pi}{\lambda_1^2+\lambda_2^2+2\kappa} = \frac{(2\tau_1^2+2\pi)}{(2\lambda_1^2+2\kappa)} = \frac{(\tau_1^2+\pi)}{(\lambda_1^2+\kappa)}$$

Since all the optimal credibilities are equal, it follows that the split plan did not reduce mean square error. Note this result holds no matter what the process covariance is between the components. So, for example a split where each component is equal to half the loss gains us nothing. Neither does a plan

where we toss a fair coin to decide if a claim belongs to one component or the other. Of course, there is no intuitive reason to expect either of these split plans could improve the accuracy of our credibility weighted estimate of the mean. More generally, our intuition is that a split cannot improve the accuracy of the final estimate if it does not meaningfully use additional information beyond that which was used for the no-split plan.

We are now ready to state our key result. This is a formula for the difference in minimal mean square error.

$$\Delta(\varepsilon_0^2) = \frac{1}{D\lambda^2}\left((\tau_1^2+\pi)(\sigma_2^2+\rho)-(\sigma_1^2+\rho)(\tau_2^2+\pi)\right)^2$$

This proof of this formula involves some laborious, but elementary, algebra. It leads immediately to a general intuitive understanding of split credibility. The essential idea is that for the split to be effective it must lead to a proportionately different allocation of the total process and total parameter variances. The most effective split possible would be to put all the process risk in one component, all the parameter risk in another, and do this in such a way that there was no process or parameter covariance between them. If that could be done, one component would have credibility of 100% and the other a credibility of 0%. Further, the mean square error of the resulting split credibility estimate would be zero. Realistically, it is impossible to do this, but it provides a strong intuition about what makes a split plan effective.

5. NO-SPLIT CREDIBILITY FOR LOSSES

Now we will derive credibility formulas for losses that arise from a model in which claim counts are generated by a single random variable and each claim severity is an independent sample from a single severity distribution. In this Count-Severity model, parameter risk arises from uncertainty about the claim count and the claim severity expectations.

Let N be the number of claims and write X(i) for the loss from the ith claim. Assume the X(i) are conditionally independent and identically distributed. Further suppose each X(i) is independent of the claim count. Assume each X(i) is an independent random sample of the severity random variable, X Define the actual loss, A, via: A= X(l)+ X(2)+ ...+ X(N). Now suppose N is parametrically dependent on a parameter, θ_N, and that X is parametrically dependent on a parameter, θ_X. Assume θ_N and θ_X have prior

distributions that are independent. We will abuse notation and usually drop the subscripts, N and X on θ. Define $\mu_N(\theta) = E[N|\theta]$, $\mu_X(\theta) = E[X|\theta]$, $\sigma_N^2(\theta) = Var(N|\theta)$, and $\sigma_X^2(\theta) = Var(X|\theta)$. Then take expectations and variances with respect to the priors to define $\mu_N = E[\mu_N(\theta)]$, $\mu_X = E[\mu_X(\theta)]$, $\sigma_N^2 = E[\sigma_N^2(\theta)]$, $\sigma_X^2 = E[\sigma_X^2(\theta)]$, $\tau_N^2 = Var(\mu_N(\theta))$, and $\tau_X^2 = Var(\mu_X(\theta))$.

We will now derive the process and parameter variance of loss using terms based on the claim count and claim severity. The conditional mean and variance are given by:

$$\mu_A(\theta) = \mu_N(\theta) \cdot \mu_X(\theta)$$
$$\sigma^2{}_A(\theta) = \mu_N(\theta) \cdot \sigma^2{}_X(\theta) + \sigma^2{}_N(\theta) \cdot (\mu_X(\theta))^2$$

Taking expectations with respect to the priors, we find the process and parameter variances:

$$\sigma^2{}_A = \mu_N \cdot \sigma^2{}_X + \sigma^2{}_N \cdot (\tau^2{}_X + \mu_X{}^2)$$
$$\tau^2{}_A = \tau^2{}_N \cdot \tau^2{}_X + \tau^2{}_N \cdot \mu_X{}^2 + \mu_N{}^2 \cdot \tau^2{}_X$$

Note the expected process variance contains a term that includes the severity parameter variance. Plugging these into the basic no-split credibility formula, we find the optimal credibility is given as:

$$z^* = \frac{\tau^2}{\tau^2 + \sigma^2} = \frac{\tau^2{}_N \cdot \tau^2{}_X + \tau^2{}_N \cdot \mu_X{}^2 + \mu_N{}^2 \cdot \tau^2{}_X}{\tau^2{}_N \cdot \tau^2{}_X + \tau^2{}_N \cdot \mu_X{}^2 + \mu_N{}^2 \cdot \tau^2{}_X + \mu_N \cdot \sigma^2{}_X + \sigma^2{}_N \cdot (\tau^2{}_X + \mu_X{}^2)}$$

If we assume N is conditionally Poisson so that $\sigma^2{}_N = \mu_N$, the process variance is:

$$\sigma^2{}_A = \mu_N \cdot (\sigma^2{}_X + \tau^2{}_X + \mu_X{}^2)$$

The resulting formula for optimal credibility is then given as:

$$z^* = \frac{\tau^2}{\tau^2 + \sigma^2} = \frac{\tau^2{}_N \cdot \tau^2{}_X + \tau^2{}_N \cdot \mu_X{}^2 + \mu_N{}^2 \cdot \tau^2{}_X}{\tau^2{}_N \cdot \tau^2{}_X + \tau^2{}_N \cdot \mu_X{}^2 + \mu_N{}^2 \cdot \tau^2{}_X + \mu_N \cdot (\sigma^2{}_X + \tau^2{}_X + \mu_X{}^2)}$$

Now N be Poisson with parameter $n\chi$, where $E[\chi] = 1$ and $Var(\chi) = c$. Under these assumptions, we have $\mu_N = n$, $\sigma^2{}_N = n$, and $\tau^2{}_N = cn^2$. Let X be conditionally exponential with mean $s\beta$ where $E[\beta] = 1$ and $Var(\beta) = b$. This implies $\mu_X = s$, $\sigma^2{}_X = E[s^2 \beta^2] = s^2 (1+b)$, and $\tau^2{}_X = Var(s\,\beta) = s^2 b$. It follows that $\mu_A = n \bullet s$ and:

$$\sigma^2{}_A = E[n\chi \cdot E[X^2 \mid \beta]] = ns^2 \cdot E[2\beta^2] = 2ns^2 \cdot (1+b)$$

$$\tau^2{}_A = Var(n\chi s\beta) = n^2 s^2 \cdot ((1+c)(1+b) - 1)$$

Thus the optimal credibility is given as:

$$z^* = \frac{n^2 s^2 \cdot ((1+c)(1+b) - 1)}{n^2 s^2 \cdot ((1+c)(1+b) - 1) + 2ns^2 (1+b)}$$
$$= \frac{n^2 \cdot ((1+c)(1+b) - 1)}{n^2 \cdot ((1+c)(1+b) - 1) + 2n(1+b)}$$

For a specific numerical example, suppose n =40, b = .01, and c=.04. Then the process variance is $2 \cdot 40 \cdot 1.01 = 80.8$ and the parameter variance is $1600 \cdot (1.0504-1) = 80.6$. Thus we find the credibility is roughly 50%.

6. SPLIT CREDIBILITY FOR COUNT-SEVERITY MODEL LOSSES

Given a per occurrence split point, k, and an occurrence of size X, we define $X_p = \min(X,k)$ as the primary severity and $X_e = X - \min(X,k)$ as the excess severity. Observe under this definition, that X_e will have a mass point at zero equal to the probability that X is less than or equal to the split point. In other words, X_e here is not the conditional excess severity. We have adopted this approach so that the primary, excess, and total losses all have the same claim count distribution. This simplifies some derivations. We now define the actual primary loss, $A_p = X_p(1) + X_p(2) + ,.. + X_p(N)$ and the actual excess loss, $A_e = X_e(1) + X_e(2) + ... + X_e(N)$. The primary and excess process and parameter variances are given as:

$$\sigma^2{}_{A_p} = \mu_N \cdot \sigma^2{}_{X_p} + \sigma^2{}_N \cdot (\tau^2{}_{X_p} + \mu_{X_p}{}^2)$$

$$\sigma^2{}_{A_e} = \mu_N \cdot \sigma^2{}_{X_e} + \sigma^2{}_N \cdot (\tau^2{}_{X_e} + \mu_{X_e}{}^2)$$

$$\tau^2{}_{A_p} = \tau^2{}_N \cdot \tau^2{}_{X_p} + \tau^2{}_N \cdot \mu_{X_p}{}^2 + \mu_N{}^2 \cdot \tau^2{}_{X_p}$$

$$\tau^2{}_{A_e} = \tau^2{}_N \cdot \tau^2{}_{X_e} + \tau^2{}_N \cdot \mu_{X_e}{}^2 + \mu_N{}^2 \cdot \tau^2{}_{X_e}$$

We can derive the following formulas for the covariances:

$$\rho = E[Cov(A_p, A_e)] = (\sigma^2{}_N - \mu_N) \cdot \mu_{X_p} \cdot \mu_{X_e} + k \cdot \mu_{X_e} \cdot \mu_N$$

Here $\pi_X = Cov(E[X_p|\theta], E[X_e|\theta])$ denotes the parameter covariance of the primary and excess severities.

Assuming claim counts are conditionally Poisson, the process variance terms simplify to:

$$\sigma^2{}_{A_p} = \mu_N \cdot (\sigma^2{}_{X_p} + \tau^2{}_{X_p} + \mu_{X_p}{}^2)$$

$$\sigma^2{}_{A_e} = \mu_N \cdot (\sigma^2{}_{X_e} + \tau^2{}_{X_e} + \mu_{X_e}{}^2)$$

$$\rho = k \cdot \mu_N \cdot \mu_{X_e}$$

We will now apply our priors to evaluate the required terms. We already have the formulas for the variances of the claims counts. Using our assumption that severity is conditionally exponential, we may write the formulas for the conditional means of the primary and excess severities.

$$\mu_{X_p}(\theta) = s\beta\left(1-\exp(-k/(s\beta)\right)$$
$$\mu_{X_e}(\theta) = s\beta\exp(-k/(s\beta)$$

We can also derive the formulas for the conditional severity process variances:

$$\sigma^2_{X_p}(\theta) = \int_0^k dx\ x^2\cdot(s\beta)^{-1}\cdot\exp(-x/(s\beta)) + k^2\exp(-k/(s\beta)) - s^2\beta^2\left(1-\exp(-k/(s\beta)\right)^2$$
$$= s^2\beta^2 - 2s\beta\cdot\exp(-k/(s\beta)) - 2s^2\beta^2\cdot\exp(-2k/(s\beta))$$

$$\sigma^2_{X_e}(\theta) = \int_k^\infty dx\ (x-k)^2\cdot(s\beta)^{-1}\cdot\exp(-x/(s\beta)) - s^2\beta^2\cdot\exp(-2k/(s\beta))$$
$$= 2s^2\beta^2\cdot\exp(-k/(s\beta)) - s^2\beta^2\cdot\exp(-2k/(s\beta))$$

Now assume that β is such that γ=1/β is Gamma distributed. Let γ have shape parameter α and scale parameter λ such that E[γ]=$\alpha\cdot\lambda^{-1}$ and Var(γ) = $\alpha\cdot\lambda^{-2}$. It follows β has density:

$$h(\beta) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\beta^{-(\alpha+1)}\cdot\exp(-\lambda/\beta)$$

With this density, we can derive the unconditional severities and the process and parameter variances and covariances. To ensure the derivations are clearly understood, we will show the first one in some detail.

$$\mu_{X_p}(\theta) = E[\mu_{X_p}(\theta)] = E[(s\beta)\cdot(1-\exp(-k/(s\beta))]$$
$$= s\int_0^\infty d\beta\ \beta\cdot\frac{\lambda^\alpha}{\Gamma(\alpha)}\beta^{-(\alpha+1)}\exp(-\lambda/\beta) - s\int_0^\infty d\beta\ \beta\cdot\frac{\lambda^\alpha}{\Gamma(\alpha)}\beta^{-(\alpha+1)}\exp(-(\lambda+k)/(s\beta))$$
$$= s\frac{\lambda}{\alpha-1}\left(1-\left(\frac{\lambda}{(\lambda+k/s)}\right)^{\alpha-1}\right)$$

Recall we have also assumed that E[β]=1 and that Var(β)=b. It follows that λ=α-1 and α=2+1/b. Using this we have:

$$\mu_{X_p} = s\left(1-\left(\frac{\lambda}{(\lambda+k/s)}\right)^{\lambda}\right) = s\left(1-\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)}\right)$$

Similarly we can show:

$$\mu_{X_e} = s\left(\frac{\lambda}{(\lambda+k/s)}\right)^{\lambda} = s\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)}$$

Using similar methods we can derive the following formulas for the process and parameter variances:

$$\sigma^2_{X_p} = s^2\cdot(1+b) - 2sk\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)} - s^2\cdot(1+b)\left(1+\frac{2kb}{s(b+1)}\right)^{-(1/b)}$$

$$\sigma^2_{X_e} = 2s^2\cdot(1+b)\left(1+\frac{kb}{s(b+1)}\right)^{-(1/b)} - s^2\cdot(1+b)\left(1+\frac{2kb}{s(b+1)}\right)^{-(1/b)}$$

$$\tau^2_{X_p} = s^2\cdot(1+b)\left(1-2\left(1+\frac{kb}{s(b+1)}\right)^{-1/b} + \left(1+\frac{2kb}{s(b+1)}\right)^{-1/b}\right) - s^2\left(1-\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)}\right)^2$$

$$\tau^2_{X_e} = s^2\cdot(1+b)\left(1+\frac{2kb}{s(b+1)}\right)^{-1/b} - s^2\left(1+\frac{kb}{s(b+1)}\right)^{-2(1+1/b)}$$

Finally we turn to the parameter covariance of the severity. We can derive:

$$\pi_X = Cov\left(\beta s(1-\exp(-k/(s\beta)), \beta s \exp(-k/(s\beta))\right)$$
$$= s^2(1+b)\left(\left(1+\frac{kb}{s(b+1)}\right)^{-1/b} - \left(1+\frac{2kb}{s(b+1)}\right)^{-1/b}\right)$$
$$-s^2\left(1-\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)}\right)\left(1+\frac{kb}{s(b+1)}\right)^{-(1+1/b)}$$

We now have enough to compute the split credibilities. As shown in Exhibit 1, if b=.01 and c=.04, the primary credibility is roughly 80% and the excess credibility is approximately 20%. This compares with a credibility of 50% for total loss. We can see that this split has improved the accuracy of our overall estimate and not coincidentally led to a differential allocation of process and parameter risk. The excess loss has received a relatively larger share of the process risk and the primary loss was allocated a relatively larger share of the parameter risk. This is borne out when we evaluate the difference in the square errors between the split and no-split plans. As shown in Exhibit 1, this difference is 267. To gauge whether this improvement is meaningful, we examine the resulting parameter CV defined as the ratio of the resulting parameter standard deviation over the mean. We find the parameter CV drops from 16.0% for the un-split plan to 15.5% for the split plan. The percentage reduction is a quite modest 3.4%. Further, when we switch the values for b and c so that b=.04 and c=.01, we see in Exhibit 2 that the excess credibility increases to 70% and the primary credibility declines to around 15%. This inversion of primary and excess credibilities can never happen in the NCCI split rating plan[6]. In Exhibit 3, we present a grid of parameter CV percentage reduction values depending on b and c. We see zeros along the diagonal where b and c are equal. In those cases, the split does not improve accuracy.

Why does this troubling result occur? Is there something about the Count-Severity model in general or about the particular conditional distributions we have used that undercuts the potential effectiveness of the split plan and causes inversions of the primary and excess credibilities? One major clue in finding answers to these questions is to observe that the excess credibility is larger than the primary credibility when the relative parameter uncertainty about severity is larger than the parameter uncertainty about claim counts. In these scenarios, much of the overall parameter risk is due to uncertainty about the severity. The effect of a split is to deprive the primary layer of much of the information about severity and it is allocated relatively less parameter risk than the excess layer. As well, a split generally allocates relatively less process risk to the primary layer. By this logic, it becomes clear that there must be a point

at which these factors are in balance such that both layers get the same proportional allocation of process and parameter risk. In that case, the split does not improve the accuracy of the estimate. That explains the zeros we see in Exhibit 3. That the balance occurs where b and c are equal is related to our use of a split point that is equal to the mean severity. Nonetheless, the general reasoning holds: under this model there will be a point at which the relative allocations of process and parameter risk are in balance and the split plan will become ineffective. Further, when the severity parameter risk grows much larger than the parameter risk for claim counts, the relative allocation of parameter risk to the excess layer will be larger than its allocation of process risk. For those scenarios, the excess layer credibility will be larger than the primary layer credibility. From an a priori perspective, there is nothing anomalous or bizarre about these scenarios. Why can't we have, say, a 5% uncertainty about our mean claim counts and 10% uncertainty about our mean severity?

We can gain further insight by translating our original model which has a single distribution of claims counts and a single distribution of claim severities into a model that has separate but possibly correlated distributions for defined sets of Small and Large claims. Here a Small claim is one that falls below the split point and a Large one as a claim that falls above it. Parameter uncertainty in the original model gets translated to parameter uncertainty about the counts and severities of the Small and Large claims. It also determines the associated parameter covariances. Finally, this is used to compute the process variances, parameter variances, and covariances of primary and excess losses. The analysis of the excess layer is fairly easy as it is based solely on the Large claims.

For our current purposes, we will not carry through the entire Small Claim-Large Claim analysis, but merely derive results for the expected numbers of Small and Large claims. Using subscripts S and L for the Small and Large claims, we have:

$$\mu_{N_L} = E[n\chi \cdot \exp(-k/(\beta s))] = n \cdot \left(1 + \frac{kb}{s(b+1)}\right)^{-(2+1/b)}$$

$$\mu_{N_S} = E[n\chi \cdot (1 - \exp(-k/(\beta s))] = n \cdot \left(1 - \left(1 + \frac{kb}{s(b+1)}\right)^{-(2+1/b)}\right)$$

We can also derive the count parameter variances and covariance:

$$\tau_{N_S}^2 = Var\left(n\chi(1-\exp(-k/(s\beta)))\right)$$
$$= n^2 \cdot (1+c)\left(1-2\left(1+\frac{kb}{s(b+1)}\right)^{-(2+1/b)} + \left(1+\frac{2kb}{s(b+1)}\right)^{-(2+1/b)}\right) - \mu_{N_S}^2$$
$$\tau_{N_L}^2 = Var\left(n\chi \cdot \exp(-k/(s\beta))\right)$$
$$= n^2 \cdot (1+c)\left(1+\frac{2kb}{s(b+1)}\right)^{-(2+1/b)} - \mu_{N_L}^2$$

We define π_N as the covariance of the expected numbers of claims and derive:

$$\pi_N = Cov\left(\mu_{N_S}(\theta), \mu_{N_L}(\theta)\right)$$
$$= E\left[n^2\chi^2 \cdot \exp(-k/(s\beta)) \cdot \left(1-\exp(-k/(s\beta))\right)\right] - \mu_{N_S} \cdot \mu_{N_L}$$
$$= n^2(1+c)\left(\left(1+\frac{kb}{s(b+1)}\right)^{-(2+1/b)} - \left(1+\frac{2kb}{s(b+1)}\right)^{-(2+1/b)}\right) - \mu_{N_S} \cdot \mu_{N_L}$$

With these formulas, we can compute the claim count parameter variances and parameter covariance for the Small and Large claims associated with our Gamma-Poisson, Gamma-Exponential model of counts and severities. Exhibit 4 shows the Small and Large claim parameters that correspond to the Count-Severity model parameters shown in Exhibit 1. Observe the contagion for the Large claims is bigger than the contagion for the Small claims. This happened because uncertainty in the expected number of Small and Large claims flows from uncertainty about the total claim count and uncertainty about the severity distribution. Given that our split point is big enough to generate on average more Small claims than Large ones, the Large claims end up with relatively more parameter risk and thus a bigger contagion. Also observe there was a significant claim count parameter correlation. This happens because a variation in the mean claim count from the Count-Severity Model raises the mean claim count for both Low Severity and High Severity claims. Both these effects boost the allocation of parameter risk to the excess layer, while the splitting allocates a disproportionate amount of the process risk to it. This makes it difficult to achieve a differential allocation of parameter and process risk. Though this is but one example,

it is not too much of a stretch to generalize beyond it to any Count-Severity model in which the prior on severity reflects overall scaling uncertainty. The structure of such a model discourages differential risk allocation and therefore makes it difficult for split credibility to be effective.

There are other loss models that might be more hospitable to the application of split credibility. One candidate is a Mixed Claim Type model in which there are different types of claims that have different severities. Assuming an undue amount of parameter covariance does not exist between the different claim types, parameter risk would not get preferentially allocated to the Large claims. Then it might be easier to achieve the desired differential allocation of process and parameter risk.

7. CONCLUSION

To summarize, we have shown that analysis of split experience rating requires analysis of the allocation of the process and parameter risk to the primary and excess layers. It is insufficient to focus on volatility alone. We have derived formulas for the mean square errors in the un-split and split plans. Then we found the mean square error of the estimate would be reduced if the split produced a differential allocation of process and parameter risk.

We have shown that a primary-excess split does not always improve accuracy, nor does it always produce a primary layer credibility that is bigger than the excess layer credibility. Using a partial implementation of the Small Claim-Large Claim approach, we saw that these disquieting results were not an artifact of our distributions or parameter choices but were inherent in the nature of our Count-Severity model. They arise from the process of modeling losses with a single claim count distribution in tandem with a single severity distribution that is subject to scale parameter uncertainty.

One task for the future is to fully implement the Small Claim-Large Claim approach. In principle, we could it to analyze the potential effectiveness of a split plan on any model of per risk losses. An additional topic for future research would be to develop and test the validity of a Mixed Claim Type model for per risk Workers Compensation losses. The existing classification of losses by Injury Type seems the most promising basis for such a model. Assuming such a model could be developed and the loss data was available to determine its parameters, an analysis could be done that might provide a stronger conceptual foundation for a plan that has been shown to work in practice.

Understanding Split Credibility

While we have had to grapple with tedious algebra and forbidding equations, and though we have encountered some disquieting results, we nonetheless believe we have achieved our main goal: we now have a better understanding of split credibility.

APPENDIX

We will show the process Coefficient of Variation (CV) of excess loss is at least as large as the process CV for total loss. Let N denote the claim count and let X denote the claim severity. Let A stand for total loss so that A = X(1) + X(2) + ...+X(N), where the X(i) are independent trials of X. Assuming N is Poisson, the mean and variance of A are given as:

$$\mu_A = \mu_N \cdot \mu_X$$

$$\sigma^2{}_A = \mu_N \cdot E[X^2]$$

Using CV_A to denote the process Coefficient of Variation of A, it follows that:

$$CV_A{}^2 = \frac{E[X^2]}{\mu_N \cdot \mu_X{}^2}$$

Given an attachment, k, define the excess severity, X_e , via X_e = X-min(X,k), and excess loss, A_e = $X_e(1) + X_e(2) + \ldots + Xe(N)$. The square of the process CV of excess loss is:

$$CV_{A_e}{}^2 = \frac{E[X_e{}^2]}{\mu_N \cdot \mu_{X_e}{}^2}$$

We will next take the derivative of the square of the process CV with respect to the attachment. To do this, we need formulas for the derivatives of the square of the expected excess severity and the expected square of excess severity:

$$\frac{\partial (E[X_e])^2}{\partial k} = 2\mu_{X_e} \frac{\partial}{\partial k} \int_k^\infty (x-k) dF_X(x) = 2\mu_{X_e} \cdot G_X(k)$$

$$\frac{\partial(E[X_e{}^2])}{\partial k} = \frac{\partial}{\partial k}\int_k^\infty (x-k)^2 dF_X(x) = -2\mu_{X_e}$$

The derivative of the square of the CV of excess loss is therefore:

$$\frac{\partial(CV_{A_e})^2}{\partial k} = \frac{-2(\mu_{X_e})^3 + 2E[X_e{}^2]\mu_{X_e}G_X(k)}{\mu_N(\mu_{X_e})^4}$$

This derivative will be non-negative if:

$$(\mu_{X_e})^2 \le E[X_e{}^2]\cdot G_X(k)$$

Now consider the extreme case in which the severity distribution above the attachment, k, consists of a single mass point of probability $G_X(k)$ at a point, y +k . It follows in that case that:

$$\mu_{X_e} = y\cdot G_X(k)$$

$$E[X_e{}^2] = y^2\cdot G_X(k)$$

So in the extreme case we have:

$$(\mu_{X_e})^2 = E[X_e{}^2]\cdot G_X(k)$$

This implies the derivative of the CV of excess loss is zero for the extreme case in which the tail consists of a point mass. In any other case, the inequality will hold strictly. We can thus conclude the derivative of the square of the process CV of excess loss is non-negative, where the derivative is taken with respect to the attachment point for excess loss. Since this is true for any attachment point, it follows that the process CV at any attachment is at least as large as the process CV of total loss. This is because the CV of total loss corresponds to an attachment of zero.

REFERENCES

[1] Phillip E. Heckman and Glenn G. Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions", PCAS LXX, 1983, pp 22-61.

[2] William R. Gillam, "Parameterizing The Workers Compensation Experience Rating Plan", PCAS LXXVI, 1989, pp. 21-56.

[3] Howard C. Mahler, Discussion of "An Analysis of Experience Rating", PCAS XXIV, 1987, pp 119-189.

[4] Glenn G. Meyers, "An Analysis of Experience Rating", PCAS LXXII, 1985, pp. 278-316.

[5] National Council on Compensation Insurance, *Experience Rating Plan Manual for Workers Compensation and Employers Liability Insurance*, March 2002.

[6] M.T.S. Teng, "Pricing Workers Compensation Large Deductible and Excess Insurance, CAS Forum, Winter 1994, pp. 413-437.

[7] Gary G. Venter, "Experience Rating - Equity and Predictive Accuracy", NCCI Digest, Vol. II Issue 1, April 1987, pp. 1-9.

ENDNOTES

[1] The NCCI Experience and Schedule Rating Plan [4] is a split plan.

[2] The actual plan uses a formula that contains ballast and weight values. Venter [7] shows this is equivalent to adding together separate credibility-weighted estimates of the primary and excess losses.

[3] Venter [7] wrote that "., both the primary and excess losses are less heavy-tailed than total losses: this seems obvious for primary losses. For excess losses, by eliminating the smaller portion, enough losses are eliminated to bring up the average value and to reduce the probability of a loss being a large multiple of the average. This makes the excess losses less heavy-tailed and thus more predictable than total losses."

[4] For example, Teng [6] argues that Workers Compensation Large Dollar Deductible and Excess programs are riskier than Full Coverage programs due to the greater variability of excess losses.

[5] Calculations have been done without including covariance. In particular, when Gillam [2] went to compute parameters for the NCCI Workers Compensation split experience-rating plan, he chose to omit the covariance terms from Mahler's equations. His decision was based on a simplifying assumption that was "defensible more on the basis of it usefulness than its veracity". Given the practical focus of Gillam's work, this may have been a reasonable choice, but that should not be read as a theoretical justification for ignoring covariance.

[6] Using the formula from Venter [7], $z_e = w \bullet z_p$ where w is the weighting value, we see that the design of the plan forces the excess credibility to always be less than the primary credibility.

DISCLAIMERS

This views expressed are solely those of the author and are not presented as the opinions of his current or prior employers. The author denies responsibility for any damages incurred as a result of reliance on the material presented in this paper.

Exhibit 1

Example of Split Plan vs No-Split Plan
Severity Mixing Parameter Smaller than Claim Count Contagion

Inputs		
Variable	**Notation**	**Value**
Mean Claim Count	n	40.000
Mean Severity	s	10.000
Severity Mxing Parameter	b	0.010
Claim Count Contagion	c	0.040
Split Point	k	10.000
Split Point to Mean Severity Rat	k/s	1.000

Results				
	Total	**Primary**	**Excess**	**Split Plan**
Claim Counts				
Mean	40.000	40.000	40.000	
Process Variance	40.000	40.000	40.000	
Parameter Variance	64.000	64.000	64.000	
Parameter CV	0.200	0.200	0.200	
Severity				
Mean	10.000	6.303	3.697	
Process Variance	101.000	12.847	61.203	
Parameter Variance	1.000	0.069	0.547	
Parameter CV	0.100	0.042	0.200	
Process Covariance				
Parameter Covariance				0.192
Loss				
Mean	400.000	252.123	147.877	
Process Variance	8,080	2,106	3,017	
Parameter Variance	8,064	2,657	1,785	
Parameter CV	0.224	0.204	0.286	
Process Covariance				1,479
Parameter Covariance				1,811
Total Covariance				3,290
Total Variance	16,144	4,763	4,802	
Credibility				
Numerator	8,064	9,623,602	2,430,852	
Denominator	16,144	12,048,713	12,048,713	
Optimal z	50.0%	79.9%	20.2%	
Error				
Initial MSE	8,064			8,064
Initial Parameter CV	0.224			0.224
MSE with Optimal z	4,036			3,770
Final Parameter CV	0.159			0.154
CV Improvement %				
From Initial to No-split	29.3%			
From No-split to Split				3.35%

Exhibit 2

Example of Split Plan vs No-Split Plan
Severity Mixing Parameter Larger than Claim Count Contagion

Inputs		
Variable	**Notation**	**Value**
Mean Claim Count	n	40.000
Mean Severity	s	10.000
Severity Mxing Parameter	b	0.040
Claim Count Contagion	c	0.010
Split Point	k	10.000
Split Point to Mean Severity Rat	k/s	1.000

Results				
	Total	**Primary**	**Excess**	**Split Plan**
Claim Counts				
Mean	40.000	40.000	40.000	
Process Variance	40.000	40.000	40.000	
Parameter Variance	16.000	16.000	16.000	
Parameter CV	0.100	0.100	0.100	
Severity				
Mean	10.000	6.252	3.748	
Process Variance	104.000	12.723	64.657	
Parameter Variance	4.000	0.260	2.258	
Parameter CV	0.200	0.082	0.401	
Process Covariance				
Parameter Covariance				0.741
Loss				
Mean	400.000	250.063	149.937	
Process Variance	8,320	2,083	3,239	
Parameter Variance	8,064	1,046	3,874	
Parameter CV	0.224	0.129	0.415	
Process Covariance				1,499
Parameter Covariance				1,572
Total Covariance				3,072
Total Variance	16,384	3,129	7,112	
Credibility				
Numerator	8,064	1,894,863	8,995,885	
Denominator	16,384	12,817,500	12,817,500	
Optimal z	49.2%	14.8%	70.2%	
Error				
Initial MSE	8,064			8,064
Initial Parameter CV	0.224			0.224
MSE with Optimal z	4,095			3,855
Final Parameter CV	0.160			0.155
CV Improvement %				
From Initial to No-split	28.7%			
From No-split to Split				2.98%

Exhibit 3

Comaprison of No-Split vs Split Plans For Collective Risk Model

Percentage Improvement in Parameter CV							
c	**b**						
	0.0025	0.0100	0.0225	0.0400	0.0625	0.0900	0.1225
0.0025	0.0%	1.7%	4.5%	6.9%	8.4%	9.2%	9.3%
0.0100	1.7%	0.0%	1.1%	3.0%	4.6%	5.7%	6.2%
0.0225	4.8%	1.2%	0.0%	0.6%	1.7%	2.7%	3.4%
0.0400	7.6%	3.4%	0.7%	0.0%	0.3%	0.9%	1.4%
0.0625	9.9%	5.4%	2.1%	0.5%	0.0%	0.1%	0.4%
0.0900	11.5%	7.1%	3.4%	1.3%	0.3%	0.0%	0.0%
0.1225	12.7%	8.4%	4.6%	2.1%	0.8%	0.2%	0.0%
0.1600	13.6%	9.3%	5.4%	2.8%	1.3%	0.5%	0.2%
0.2500	14.8%	10.5%	6.6%	3.9%	2.2%	1.2%	0.6%

Exhibit 4

Conversion of Collective Risk Count-Severity Loss Model To Claim Count Parameters of Mixed Claim Type Model

Count-Severity Model Parameters		
Variable	**Notation**	**Value**
Mean Claim Count	n	40.000
Mean Severity	s	10.000
Severity Mixing Parameter	b	0.010
Claim Count Contagion	c	0.040
Split Point	k	10.000
Split Point to Mean Severity Ratio	k/s	1.000

Mixed Claim Type Model		
Variable	**Notation**	**Value**
Small Claims		
Mean Claim Count	n_S	25.357
Parameter Variance		27.917
Claim Count Contagion	c_S	0.043
Large Claims		
Mean Claim Count	n_L	14.643
Parameter Variance		10.773
Claim Count Contagion	c_L	0.050
Parameter Dependence		
Claim Count Parameter Covariance	π_N	12.655
Claim Count Parameter Correlation		0.730