

# Credibility and Other Modeling Considerations for Loss Development Factors

Uri Korn, FCAS, MAAA

---

## Abstract

In this paper, we discuss various credibility and modeling strategies for loss development factors. We present several improvements to the popular inverse power curve to help it better fit to the data. Using a basic approach to credibility weighting curves often produces results that do not lie in between the original curve and the overall average, as would be expected. We show a technique to deal with this issue. We also discuss how to model across continuous variables as well as show formulas for converting and modeling across different loss caps, retentions, and policy limits.

**Keywords.** Loss Development Factors, Credibility, Generalized Additive Models, Splines

---

## 1. INTRODUCTION

Unlike other businesses, due to the delay from the date an event occurs until it is reported and ultimately paid, insurance companies do not know whether their products are being adequately charged for or whether they were sold at a profit or at a loss. Actuaries rely on loss development patterns to help estimate all of this. A book of business is typically made up of insureds having different characteristics with policies written at various retentions and limits, all of which can be expected to have different development patterns. Not reflecting these differences properly can lead to inaccurate estimates of what is performing well and what is not, not to mention an inaccurate total result if there have been changes over time. But it is difficult to account for these differences since dividing up the data often results in portions that are too volatile to analyze on their own, especially under the typical approach which involves selecting a separate parameter for the loss development factor of each period.

This paper discusses some modeling strategies for loss development factors and for differentiating across segments while still leveraging the credibility across the divisions of the data. Specifically, this paper discusses the following:

1. An improved method of fitting curves to loss development factors
2. Credibility weighting techniques, both for individual LDFs and fitted curves. For fitted curves, ensuring that the credibility weighted results lie in between the original curve and the complement.
3. Methods to model across continuous variables
4. Formulas for converting and modeling across different loss caps, retentions, and policy

limits

All of the models in this paper can be implemented as full Bayesian models solved using Markov Chain Monte Carlo, or more simply in spreadsheets. This allows for easier adoption and also facilitates implementation in account rating engines where credibility weighting can be performed on an account's specific loss development pattern. Both versions will be discussed.

## **1.1 Outline**

To illustrate credibility weighting of loss development factors, three types of credibility models will be shown: the first involves credibility weighting each LDF individually, the second involves a parametric curve (the inverse power curve), and last can be thought of as a compromise between these two and involves using a Generalized Additive Model, which will be explained. For the latter two models, an approach is shown to help ensure that the resulting credibility weighted curves lie in between the original curve and the overall average, which is often violated using a more basic approach. We also discuss some improved strategies for fitting curves to loss development factors. After that, modeling across continuous variables as well as different retentions, policy limits, and loss caps is discussed.

## **1.2 Technical Background**

For the models that we will be discussing in this paper, we will be assuming that the variance of each LDF is inversely proportional to the volume of cumulative paid or reported losses in the previous age of the triangle. (So, for example, the variance of the first LDF is related to the losses from the first age.) This must be the case since the variance of the losses for a subsequent age for two equally sized accounts is equal to twice the variance of one of these accounts. To convert these losses into LDFs, they are divided by the sum of the losses in the previous age, and so the variance is divided by the square of this sum (since it is not a random variable). Thus, the denominator of the LDF variance for both equally sized accounts combined will be two squared greater than that of a single account. And so the variance of the combined LDF will be  $2 / 2^2$  or half of that of each of the accounts separately. Assuming any other relationship between variance and losses will not agree with this result and will lead to inconsistencies.

To calculate the variance of each LDF, the Buhlmann-Straub formula for the "within variance" can be used, using the previous cumulative losses as the weight (Dean 2005):

$$\hat{EPV} = \frac{\sum_{g=1}^G \sum_{n=1}^{N_g} W_{gn} (X_{gn} - \bar{X}_g)^2}{\sum_{g=1}^G (N_g - 1)} \quad (1.1)$$

Where  $G$  are the number of risks,  $N_g$  are the number of periods for group  $g$ ,  $W_{gn}$  is the weight for group  $g$  in period  $n$ ,  $X_{gn}$  is the value for group  $g$  in period  $n$ , and  $\bar{X}_g$  is the average value for group  $g$ .

Note how this formula multiplies by the weight, but does not divide by it. This is because this parameter is really more accurately described as a within variance factor, rather than the within variance for anything in particular. If we take a closer look at the Buhlmann-Straub credibility formula as well, we can see that this factor is divided by the total weight in order to calculate the actual within variance. The ratio of the between variance to the within variance determines the amount of credibility given to the risk.

$$Z = \frac{N}{N + W/A} = \frac{1}{1 + \frac{W/N}{A}} = \frac{1}{1 + V/A} \quad (1.2)$$

Where  $Z$  is the credibility given,  $N$  is the weight,  $A$  is the between variance,  $W$  is the between variance factor, and  $V$  is the actual between variance. So to calculate the variance of an individual LDF in the triangle, the within variance factor should be divided by the cumulative losses from the previous age. For the variance of the overall average LDFs for each age, the weight used should be the sum of the cumulative losses in the previous age that were used for calculating the LDF.

Once the within variance factors are calculated at every age, they should be smoothed by fitting a curve. A logarithmic curve on the logarithm of the age seems to provide a good fit. In practice, the first one or two ages may need to be removed, as well as the latter, extremely volatile tail portion to be able to fit more accurately.

We will be using the normal distribution to calculate the likelihood of each fitted loss development factor. Since we are allowing our variances to vary for each LDF, this approach is more similar to Kernel smoothing than to assuming that LDFs are actually normally distributed. Taking these two assumptions together of using normal distributions with variances inversely proportional to the losses produces the same result as taking a weighted average of the LDFs, as is commonly done, and so is consistent with traditional actuarial practice. This type of model works well in practice and is the easiest to implement and understand, although other assumptions are

possible as well.

Most of the credibility models presented in this paper use Bayesian credibility. These models can be implemented without the use of special Bayesian software. Since we are using the normal distribution for the LDFs and we will also be assuming that the prior distribution is normal (that is, the distribution of the hypothetical means for each group, in Bayesian terms), which is the common assumption, this is a conjugate prior and the resulting posterior distribution (the credibility weighted result) is normal as well. Using MLE returns the mode of a distribution, which will also be equal to the mean for the normal distribution, and so will return identical results to that produced using special Bayesian software. Further details will be discussed later.

## 2. Credibility Weighting Individual LDFs

The first LDF credibility model we will discuss involves credibility weighting the individual loss development factors. Only age-to-age factors should be used since age-to-ultimate factors have a high degree of dependency on one another.

To credibility weight individual LDFs, Buhlmann-Straub credibility can be used. The “between variance” should be calculated for every age. The formula for calculating the between variance is shown below (Dean 2005):

$$V\hat{HM} = \frac{\sum_{g=1}^G W_g (\bar{X}_g - \bar{X})^2 - (G-1)EPV}{W - \frac{\sum_{g=1}^G W_g^2}{W}} \tag{2.1}$$

Where all terms are the same as above,  $\bar{X}$  is the average value across all groups,  $W_g$  is the sum of the weights for group  $g$ , and  $W$  is the sum of the weights across all groups.

Either a curve should be fit to these points similar to the within variance factors, or they can be determined as a constant factor of the within variance factors.

The downside to this simple credibility approach is that each LDF is treated individually, and not as part of a curve. This discards a large amount of useful information about the relationship between the LDFs that can be used to improve the accuracy of our selections. As mentioned in England and Verrall 2002 (and others), selecting parameters for every single age is over-parameterized, meaning that more parameters are being chosen than necessary, which will increase

the prediction variance. Therefore, using a curve is highly recommended. A problem with using a curve along with credibility, however, is that normal credibility weighting techniques often result in curves that do not lie between the original curve and the overall mean, as we would hope. The next section discusses a technique to address this issue.

### **3. Applying Credibility to the Inverse Power Curve**

The inverse power curve is a well-known method used to help smooth LDFs (Sherman 1984). This curve can be fit by using the regression equation, as mentioned in the paper<sup>1</sup>:

$$\log(LDF - 1) = A + B \times \log(t) \tag{3.1}$$

Where  $A$  and  $B$  are the regression coefficients and  $t$  is the age. In our experience, it can be a useful tool to help smooth out some LDFs, especially in the latter portion of the curve, although it often has trouble fitting the entire curve. This depends on the type of business being modeled, however. Also, solving this regression equation using ordinary regression gives too much weight to the tail portion of the curve as mentioned in Lowe et al (1985). This is not an issue if using an extrapolation from earlier more stable points to predict later ages in the curve, as is commonly done, but can be an issue when attempting to fit to the entire curve. An improved way to fit this curve will be discussed.

Instead of using ordinary regression to solve for the parameters, maximum likelihood estimation (MLE) can be used instead. (This can be implemented using a weighted GLM as well.) Doing this makes the weights by age more appropriate and helps provide a better fit than simple regression. MLE can be performed either on each individual LDF or on the weighted averages by age; the results will be the same if the normal distribution is used. If using the weighted averages, the losses used for calculating the variances should be the sum of the cumulative losses in the previous period that were used in calculating the average LDF. The only real reason to use the individual LDFs is when constructing a full Bayesian model (which solves for the within and between variances as part of the model). The log-likelihoods of the fitted LDFs for each age (using equation 3.1) should be added together, and this sum should be maximized. In practice, it may help to exclude the LDFs that do not provide a good fit to the curve or that are too volatile, such as the first few LDFs or the latter portion of the curve that is very sparse and volatile.

---

<sup>1</sup> The paper actually uses the logarithm of the inverse of the age, but the regression equations are equivalent.

## *Credibility and Other Modeling Considerations For Loss Development Factors*

To perform credibility weighting among the segments, normally Bayes' formula would be used:  $f(\text{Parameters} \mid \text{Data}) = f(\text{Data} \mid \text{Parameters}) \times f(\text{Parameters})$ , or equivalently,  $\text{Posterior}(\text{Parameters}, \text{Data}) = \text{Likelihood}(\text{Data}, \text{Parameters}) \times \text{Prior}(\text{Parameters})$ , and the regression parameters would be used for calculating this prior likelihood component, which is the credibility component of the likelihood. However, doing so often results in poor behaving curves, as mentioned. Instead, we suggest reparameterizing the curve as will be explained.

The first step is to invert the regression equation (3.1) to solve for the LDFs. Doing so results in the following equations. Since there are two parameters, we need two LDFs at two separate ages to solve for them.

$$B = \frac{\log\left(\frac{LDF_1 - 1}{LDF_2 - 1}\right)}{\log\left(\frac{t_1}{t_2}\right)} \quad (3.2)$$

$$A = \log(LDF_1 - 1) - B \log(t_1) \quad (3.3)$$

Now, given any two LDFs of the fitted curve, we can solve for the original regression parameters. And given these, we can calculate all of the LDFs of the curve. Ignoring the middle step, we can construct the entire curve from these two LDFs. Since the entire curve can be defined by these two LDFs, we can consider these as the parameters of the curve. This being the case, we can alternatively calculate the prior likelihood, that is the credibility component of the likelihood, using these new parameters instead. When doing so, the between variance used should be consistent with our new LDF parameters and not the original regression parameters. This between variance can be estimated by calculating the between variance of the actual LDFs using equation 2.1 above<sup>2</sup>. The ages for these LDF parameters can be selected as being equally spaced along the ages used to perform the fit, but they can be tweaked if needed. Note that even though we are only performing the credibility weighting on two LDFs of the curve, we are still credibility weighting the entire curve, since changing these LDFs affects the entire curve as they are the new curve parameters.

The equations should be inverted when implementing a full Bayesian model. For an MLE model, it is not necessary to actual invert the equations (assuming we are not implementing a

---

2 Note that including the likelihood of the between variance of every LDFs, as was done in the previous section, is not correct here, since following Bayes' equation, only the parameters should be included in the prior likelihood; in the above section, each LDF was considered separately and so is considered a parameter. This is not the case here.

multidimensional model or modeling across continuous variables, which will be discussed later), since we can solve for any fitted LDF along the curve and use this LDF to calculate the prior likelihood component that we need, effectively “pretending” that we have inverted the equation. Doing this will yield the exact same results as if we had actually performed the inversion. For the MLE model, the complement of credibility for each parameter should be taken from the results from fitting a curve to all segments combined, and not the actual empirical LDFs. (For the Bayesian model, the complement is determined as part of the model.) To summarize, the log-likelihood for this model is:

$$\begin{aligned}
 & \sum_{d=\text{ Durations}} N(\text{ Fitted LDF}_d, \text{ Actual LDF}_d, \text{ Within Variance}_d) \\
 + & \sum_{c=2 \text{ Durations Selected For Credibility Weighting}} N(\text{ Fitted LDF}_c, \text{ Average LDF}_c, \text{ Between Variance}_c)
 \end{aligned} \tag{3.4}$$

Where  $N(A, B, C)$  is the logarithm of the probability density function (PDF) of a normal distribution at  $A$ , with mean and variance of  $B$  and  $C$ , respectively. The fitted LDFs are determined from the inverse power curve equation (3.1). The parameters of this equation are determined by a maximization routine that maximizes this likelihood. The within variances are calculated by dividing the within variance factor by the cumulative paid or reported losses, as mentioned. (As a practical matter if implementing with MLE, a minimum value, such as  $1 \times 10^{-20}$  should be set for the values of the normal PDFs, so that they are not too close and rounded to zero, which will cause errors with the logarithm function. Also, for the earlier ages, it sometimes improves the fit to use the within variance factor at a point a few ages later.)

#### **4. Applying Credibility to an Additive Model**

The following model can be described as a combination of the inverse power curve and the Generalized Additive Model suggested by England and Verrall 2001 in which they use this to model the incremental paid or reported loss amounts. Here, however, we will be modeling on the actual LDFs instead, as we did for the inverse power curve, because it works better with credibility weighting. It also involves solving for fewer parameters so that it can be implemented from within spreadsheets.

Before we begin explaining how this model works, we will briefly explain splines and additive models. An ordinary regression model has a dependent variable that is a linear function of one or

more predictive variables and has the form:

$$Y_i = \sum^i B_i X_i$$

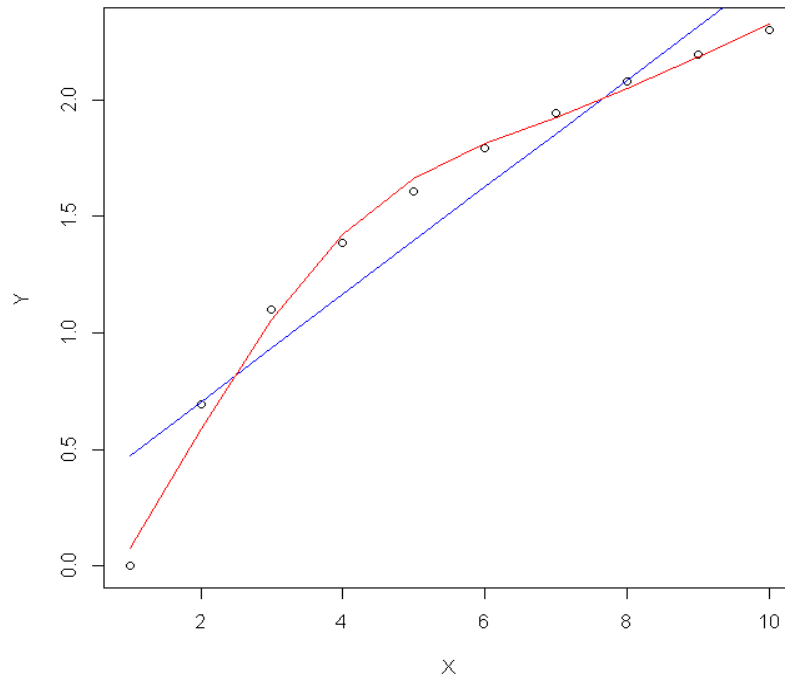
An additive model, instead of just linear functions, allows for any function, and has the form:

$$Y_i = \sum^i f(X_i)$$

Usually, some type of smoothing function is used that helps adapt the curve to the actual data, even if the relationship is not perfectly linear. Cubic splines are a very common choice since they do a good job of adapting the curve to the data and results in nice, smooth curves.

For example, if we were trying to fit a regression model to the data below and were not able to find a simple transformation of the independent variable that nicely fit the data, such as a logarithm, we might consider using an additive model. The results using a linear regression model (blue line) versus an additive model (red line) are shown below. Note how the additive model nicely adapts the shape of the curve to the data.





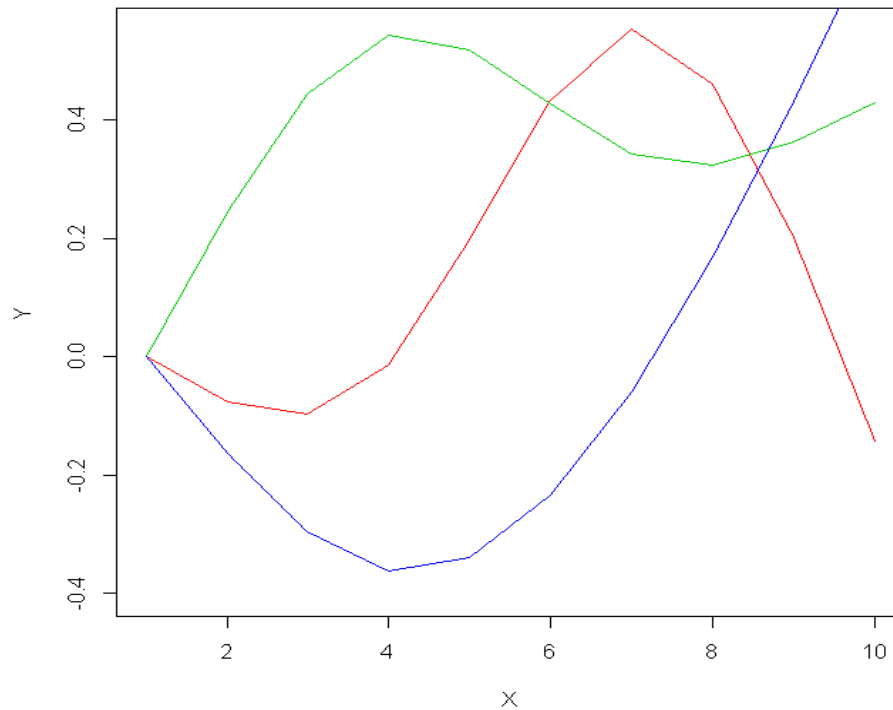
An additive model can also be implemented with an ordinary regression model using splines<sup>3</sup>. These are functions that generate multiple new numeric sequences based off of the original sequence that can be used for smoothing. These new variables can then be plugged into a standard regression model with the same result as an additive model. The example shown above used three natural cubic spline transformations off of the numeric sequence from one to ten and are graphed below. Each of these resulting curves represents one of the degrees of freedom of the spline. By multiplying each curve by a coefficient and adding the results together, smooth curves can be fit to data of multiple forms and shapes. The benefits of this approach is that splines work better for our credibility procedure and it also allows additive models to be implemented from within spreadsheets<sup>4</sup>. A full discussion of additive models is outside of the scope of this paper.

---

3 A full additive model may also help decide how many new variables should be generated (known as the degrees of freedom), etc. but this is not crucial.

4 One way this can be done is to generate the spline numeric sequences outside of the spreadsheet and then paste them in. Natural cubic splines can be generated in R using the ns method of the splines package. For example, the following code can be used to generate a spline with three degrees of freedom (that is, equivalent of three variables) starting at the second age, ending at the 20<sup>th</sup>, but having a tail that goes out to the 40<sup>th</sup>, all on a log scale as we will mention a bit later:

```
library(splines)
ns(log(2:40), Boundary.knots=c(log(2),log(20)), df=3)
```



Having described the benefits of additive models and how they work, it becomes clear why we would want to use them in a loss development model; it is very difficult to find a nice parametric shape that fits nicely to the entire curve. Additive models solve this problem by adapting the curve to the data. A downside, however, is that they can sometimes over-fit. When fitting volatile data with splines, it is often necessary to remove the later, more volatile points from the fitting. (When this is done, the spline should only be constructed for the actual ages being fit, although the tail can be extended further. See the code in the footnote above for how to generate a spline like this in R.) These models may have trouble fitting to the first one or two LDFs as well, and these may need to be removed from the fitting and selected outside of this model.

This method works best when the spline is generated on the logarithm of age, and so is very similar to the inverse power curve, but with additional smoothing to help fit the data even better. We refer to this model as the smoothed inverse power curve. The data we examined worked best using a spline with three degrees of freedom and so we will assume this is the case below, but a different number can be used as well.

Compared to the inverse power curve, the spline/additive model usually provides a much better fit to the data. Also, having more parameters, the credibility weighting occurs at more points in the curve and is often better behaved. This makes it better able to handle situations when a curve

## *Credibility and Other Modeling Considerations For Loss Development Factors*

intersects with the credibility complement, for example; the resulting credibility weighted splines curve usually does a good job of staying in between the original curve and the complement. The points at which credibility is being done can also be tweaked, which can help sometimes if needed.

The regression equation used here is:

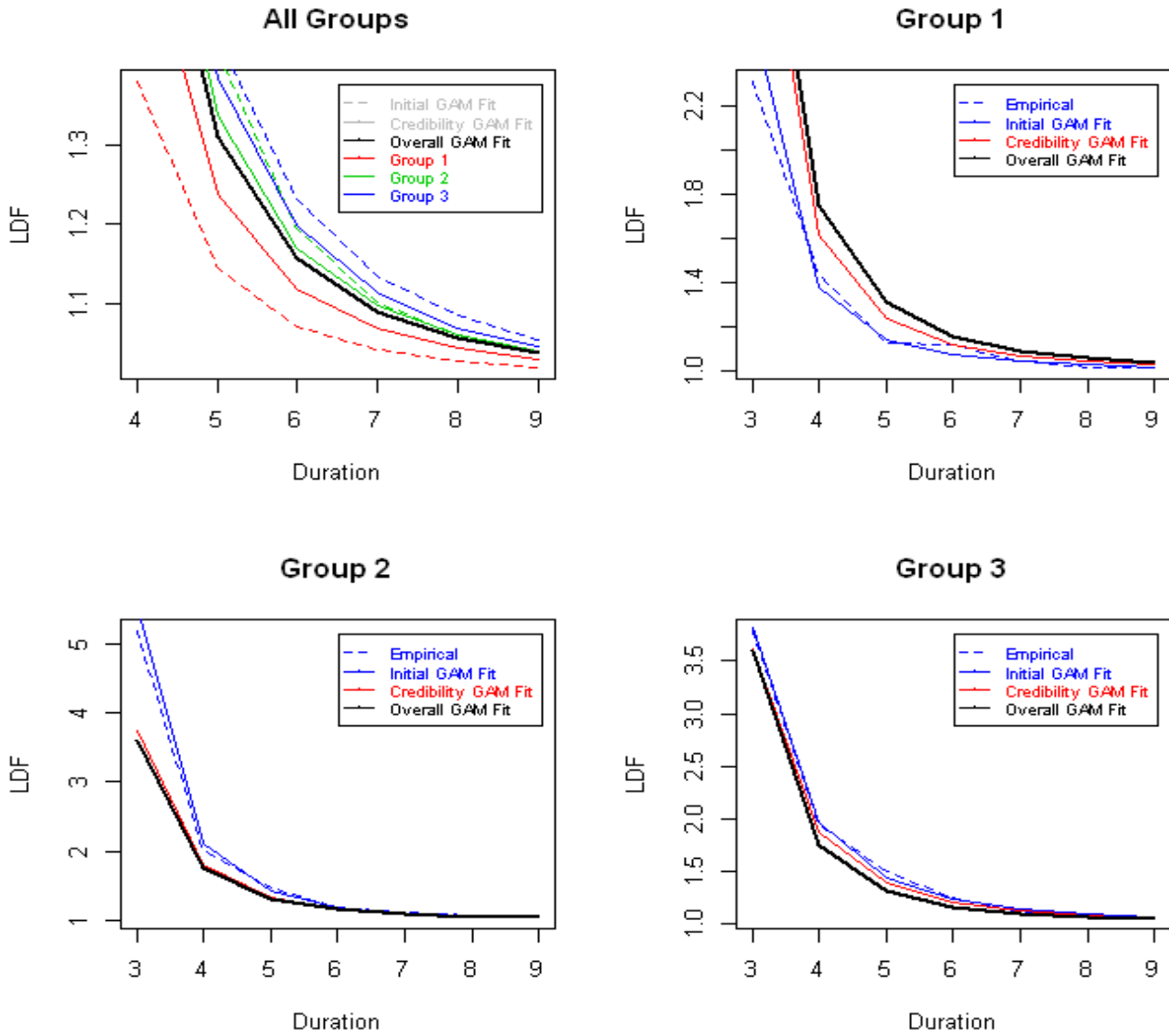
$$\log(LDF - 1) = A + B \times s(\log(t))$$

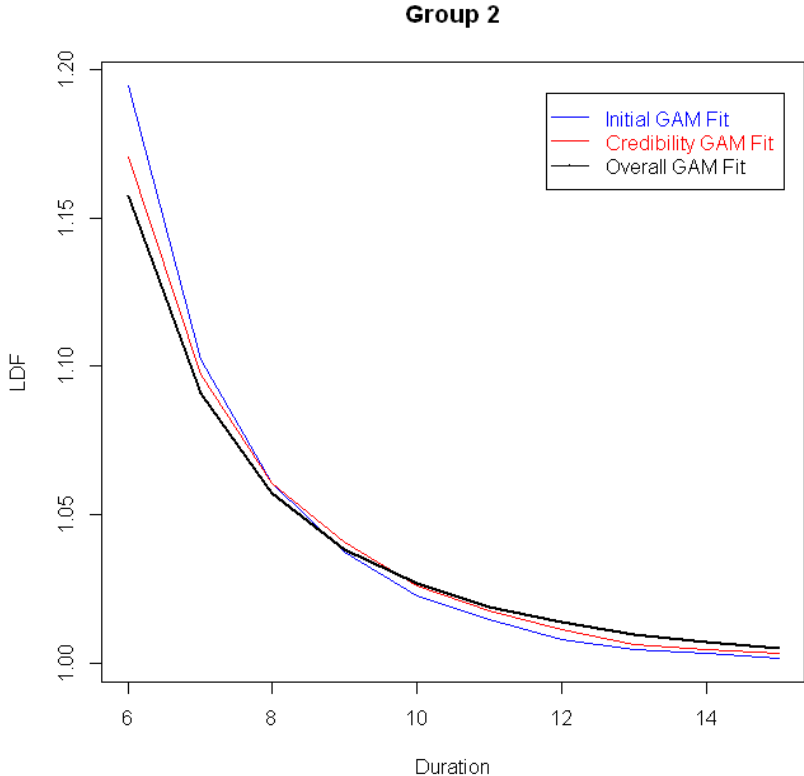
Where  $A$  and  $B$  are the regression coefficients,  $t$  is the age, and  $s$  is a smoothing cubic spline function. Using splines, this can be extended to:

$$\log(LDF - 1) = A + Bt^1 + Ct^2 + Dt^3 \tag{4.1}$$

Where  $A$  is the intercept of the curve and  $B$ ,  $C$ , and  $D$  are the slope parameters for each of the generated spline variables,  $t^1$ ,  $t^2$ , and  $t^3$  on the logarithm of the age. (We used superscripts to denote the different spline variables.) To fit such a model, similar to the inverse power curve, we use maximum likelihood to solve for the parameters. Implementing credibility is also very similar to what we did for the inverse power curve, except that here the prior likelihood should be calculated for four different LDFs instead of two since this regression equation has a total of four parameters including the intercept.

Below are some examples of implementing credibility weighting with this method on real data. (The resulting LDFs have been modified so as not to reveal any propriety information.) The spline curves provide a good fit to the data and the credibility weighted curves lie in between the original and the overall. The second graph zooms in on the second segment to show that the credibility curve does a good job of staying in between the original and the overall curve even when these two curves cross over each other. (All of these fits were produced in Excel using the Solver feature to maximize the log-likelihood.)





The equations for inverting these curves are shown in Appendix A for both three and four parameter curves.

### 5. Multidimensional Credibility Models

The models we have been discussing up to this point were one dimensional models as the credibility weighting was done across a single variable. A multidimensional model can also be constructed that considers the differences across more than a one variable. Assuming that our two variables are state and industry, a two dimensional model can be built by defining a relationship for the (inverted) LDF parameters, such as the following:

$$\log(LDF_{s,i}) = \text{Intercept} + \text{State Coefficient}_s + \text{Industry Coefficient}_i \tag{5.1}$$

We used a log-link here to make the relationship multiplicative, since this usually works best for multidimensional models. The total log-likelihood would be calculated by summing up the log-likelihood of each fitted LDF and the log-likelihood of the priors for each coefficient. The equation

would be as follows:

$$\begin{aligned}
 & \sum_{s=States} \sum_{i=Industries} N(\text{Fitted LDF}_{s,i}, \text{Actual LDF}_{s,i}, \text{Within Variance}_{s,i}) \\
 & \quad + \sum_{s=States} N(\text{State Coef}_s, 0, \text{State Between Variance}) \\
 & \quad + \sum_{i=Industries} N(\text{Industry Coef}_i, 0, \text{Industry Between Variance})
 \end{aligned} \tag{5.2}$$

Where, once again,  $N(A, B, C)$  is the logarithm of the probability density function evaluated at  $A$ , with a mean of  $B$  and variance of  $C$ . Each state and industry coefficient is credibility weighted back towards zero, which pushes each curve back towards the intercept, which will be the complement of credibility. It is also possible to add an interaction term for state and industry and have that credibility weighted back towards zero as well. This will give the model more flexibility to better reflect the differences of state-industry combinations that differ from the average.

This type of model can be solved without the use of special Bayesian software as well. Since MLE parameters are known to be approximately normal, and since the prior distribution is normal, the posterior (credibility weighted result) will be approximately normal as well. And, as we discussed, this type of model can be solved with MLE. However, unlike the one dimensional models discussed above, every segment needs to be maximized together. Because of this, the number of parameters may be too many to have accurately solved with a maximization routine, but this depends on the circumstance.

## 6. Modeling across Continuous Variables

For modeling LDFs across most continuous predictive variables in the data, such as account size or retention (which will also be discussed more thoroughly in the next section), we would usually not want to credibility weight these curves back towards the mean, since we usually expect there to be an order to these curves, and credibility weighting would just bump up the lower curves and bump down the higher ones. Instead, we can define a relationship between the different curves that depends on the continuous variable. To implement, the inverted, reparameterized version of the curves should be used. For each group, the LDF parameters should be set to a function of the continuous variable. For example, the following formula can be used to determine the LDF parameters for each retention group:

$$\log(\text{LDF Parameter}) = \text{Intercept} + \exp(\text{Coefficient}) \times \log(\text{Retention}) \quad (6.1)$$

We took the exponent of the coefficient to guarantee that the actual coefficients used are positive and so will result in curves that can only increase by retention, which is usually the expectation.

As an alternative, sometimes it works better to simply constrain the LDF parameters to the expected order. This can be implemented by setting each LDF parameter for each group to the LDF parameter of the group below it plus the exponent of another parameter. The exponent is used to ensure that the difference is positive. For both of the approaches mentioned, when working with volatile data, setting a minimum value for the slope or difference parameters to something small, such as 1% or lower, often produces results that fall more in line with expectations.

## 7. Loss Caps, Retentions, and Policy Limits

Besides for the strategy mentioned in the previous section, when modeling across different loss caps, retentions, and/or policy limits, we can leverage information from the severity distribution to help define the relationships between the groups. This method assumes that the loss severity distribution has already been estimated. It also requires claim count development factors. Our approach differs from that in Sahasrabuddhe 2010, which suggests using the severity distribution to modify the actual data of the triangle; here we convert the LDFs themselves. Note that this strategy uses the regular (non-inverted, that is) versions of the curves (unless credibility weighting is being done as well.)

We will start off with the following relationship mentioned in Siewert 1996 (although in a slightly different syntax). This formula simply states that loss development consists of the arrival of new claims as well as increased severity of both the existing and new claims.

$$LDF_t = CCDF_t \times SDF_t \quad (7.1)$$

Where CCDF is the claim count development factor and SDF is the severity development factor, which accounts for the increase in the average claim severity as a year matures. Flipping the equation around, this becomes:

$$SDF_t = \frac{LDF_t}{CCDF_t} \quad (7.2)$$

We will use these relationships to demonstrate modeling across different loss caps assuming that we are basing the LDFs of other, less stable caps on one particular, more stable cap. We will then generalize to include retentions and policy limits and also allow modeling of all groups simultaneously.

For a particular cap,  $c1$ :

$$SDF_t(c1) = \frac{LEV_T(c1)}{LEV_t(c1)} \tag{7.3}$$

Where  $LEV_t(x)$  is the limited expected value at  $x$  at age  $t$  and  $T$  is infinity (although  $t+1$  can be used to convert age-to-age factors as well). If we have an assumption for how severity development affects claims, we can use this to derive the LDFs. For now, we will assume that all uncapped losses increase on average by the same multiplicative factor as a year matures, and we adjust the loss severity distribution by a scale adjustment. To explain, most distributions have a way of modifying one of the parameters which causes each claim to increase or decrease by the same multiplicative factor. For example, the  $\mu$  parameter of the lognormal distribution is a log-scale parameter and adding the natural logarithm of 1.1, for example, will increase each claim by 10%. For a mixed exponential distribution, each  $\theta$  parameter would be multiplied by 1.1. We rewrite equation 7.3 to show the scale parameters instead of the ages, where  $LEV(\theta; c1)$  is the LEV with a scale parameter of  $\theta$  at a cap of  $c1$ <sup>5</sup>:

$$SDF_t(c1) = \frac{LEV(\theta; c1)}{LEV(\theta/a_t; c1)} \tag{7.4}$$

Since we know the SDF and can calculate  $LEV(\theta; c1)$ , we can back into the factor,  $a$ , that satisfies this equation. If losses are uncapped, there are no policy limits, and  $c1$  is infinity, then the factor,  $a$ , would equal the SDF. Otherwise, it will be slightly higher. Once we have the loss severity distribution at time  $t$ , we can use this to derive the severity development factor at another loss cap,  $c2$ :

---

<sup>5</sup> As a side note, applying this SDF to the claims of each year can also be used as a strategy for developing the severity distribution to ultimate when fitting increased limit factors.



$$SDF_t(c2) = \frac{LEV(\theta; c2)}{LEV(\theta/a_t; c2)} \quad (7.5)$$

And we can then use this to calculate the loss development factor, at a loss cap of  $c2$ :

$$LDF_t(c2) = CCDF_t \times SDF_t(c2) \quad (7.6)$$

The above assumed that all claims were ground up. If this is not the case, and there is a retention (assuming that it is uniform across all policies, for now), the average severities can be calculated as:

$$\frac{LEV(AP + Cap) - LEV(AP)}{s(AP)} \quad (7.7)$$

Where  $AP$  is the retention. We divided by the survival function at the retention to produce the average severities conditional on having a claim above the retention, which is consistent with the claims we observe in the triangle.

For modeling across different retentions, the strategy changes slightly since the claim counts are not at the same level. We can control for this by making the average severities for a retention conditional of having a claim at another retention by dividing by the survival function at this retention. When converting the severity development factor to an LDF, the claim count development factor at this retention should be used. The formulas are as follows, where  $SDF(x, y, \text{Relative to } z)$  is the severity development factor at a retention of  $x$ , a cap of  $y$ , and expressed relative to the claim counts of retention  $z$ .

$$SDF_t(AP2, Cap, \text{Relative to } AP1) = \frac{(LEV_T(AP2 + Cap) - LEV_T(AP2)) / s_T(AP1)}{(LEV_t(AP2 + Cap) - LEV_t(AP2)) / s_t(AP1)} \quad (7.8)$$

$$LDF_t(AP2, Cap) = CCDF_t(AP1) \times SDF_t(AP2, Cap, \text{Relative to } AP1) \quad (7.9)$$

If just converting from LDFs of one retention to another once the  $a$  factors are already known, the

*Credibility and Other Modeling Considerations For Loss Development Factors*

SDFs can also be expressed relative to first dollar CCDFs (even if these are not available) to simplify the formula. The formula is:

$$LDF_i(AP2, Cap2) = LDF_i(AP1, Cap1) \times \frac{SDF(AP2, Cap2, \text{Relative to } 0)}{SDF(AP1, Cap1, \text{Relative to } 0)} \quad (7.10)$$

To leverage credibility in the claim count development factors as well, claim counts from one retention can be converted to another using this formula:

$$CCDF_i(AP2) = CCDF_i(AP1) \times \frac{s_T(AP2)/s_T(AP1)}{s_i(AP2)/s_i(AP1)} \quad (7.11)$$

To give an example, if the (age-to-ultimate) LDF for a particular age for a group of ground-up policies with limits of one million is 1.3 and the claim count development factor is 1.2, this would indicate that the severity development factor is  $1.3 / 1.2 = 1.083$ , using equation 7.2. Assuming the severity distribution (at ultimate) is lognormal with mu and sigma parameters of 10 and 2, respectively, we can back into the scale adjustment using equation 7.5. Using an  $a$  value of 1.115 produces the desired SDF of 1.083. Using this same equation and same value of  $a$ , the SDF for a group of policies with limits of five million is equal to 1.101, which implies an LDF of  $1.101 \times 1.2 = 1.321$ , slightly higher than the original LDF of 1.3, as expected. Similarly, using equations 7.8 and 7.9 for excess policies, the SDF for a group of policies with retentions and policy limits of one million, expressed relative to group up claim counts, is 1.144, which implies an LDF of 1.372.

In the above discussion, we assumed that every policy is written at the same retention or policy limit. For a more realistic scenario with different limits and retentions within each group, the average expected severity should be calculated across all policies. If we assume that the expected frequency of each policy is equal to the (on-level) premium divided by the expected average severity, the average severity is equal to the total premium divided by the total number of expected claims, or:

$$\frac{\sum_i \text{Premium}_i}{\sum_i \text{Premium}_i / \text{Expected Average Severity}_i} \quad (7.12)$$

If the retentions or policy limits within each group are not too far apart, it is possible that calculating the average severities using a premium weighted average limit and/or retention may not be too far off. This strategy can also be used to adjust LDFs if there are shifts in the average retentions and/or limits by year. We ignored the effects of trend in the above which can easily be added by applying a de-trending factor to the scale parameter of the loss severity distribution.

As we mentioned, the above discussion was geared towards converting LDFs from one retention/limit/cap to another, but it is also possible to model across all groups simultaneously using these relationships. To do so, instead of backing into the  $a$  factors, since we are using MLE, they can be included in the parameters being maximized. A curve can be fit to the age-to-age  $a$  factors, and then they can be multiplied together to produce the age-to-ultimate factors, which are needed to calculate the SDFs. If simultaneously fitting the CCDFs, parameters will be needed for these as well (at one particular retention). The log-likelihoods can then be calculated and added up across all LDFs and CCDFs, and this sum can be maximized. Alternatively, the claim counts can be ignored and the parameters being maximized can include the  $a$  factors and the parameters for the LDFs of one of the groups. Once we have all of these, the SDFs can be calculated and equation 7.10 can be used to convert these LDFs to different retentions, caps, and/or limits. This approach, however, does not utilize the data in the claim count develop factors and so is not as strong.

We assumed here that every claim increases by the same amount using a scale factor adjustment, but since we are backing into (or maximizing) the adjusted value of  $LEV_i(x)$  using the SDF, this procedure allows for any sort of parameter transformations. For example, for excess losses modeled with a one- or two-parameter Pareto, allowing the alpha parameter to vary instead of the Beta parameter, which is a scale parameter, has the effect of increasing or decreasing the tail of the distribution.

## **8. Individual Account Credibility**

Most of the models discussed above can be implemented relatively simply without use of any specialized software. This allows the use of these credibility models in account rating engines, often implemented in spreadsheets. The complement of credibility for each account should be the selected LDFs for the portfolio with the between variances and the within variance factors calculated at the portfolio level. The between variance should represent the variance of the differences across accounts and can be calculated by looking at a sampling of accounts. The within variance can be calculated by dividing the within variance factor by the account's losses and credibility weighted LDFs can be produced.

## **9. Conclusion**

In this paper we discussed several loss development models that perform very well in practice and that are relatively simple to implement. Using these models will allow for more accurate differentiation between risks that properly reflects the differences in the patterns in which losses arrive.

**Appendix A**

The following are the equations for inverting a splines regression equation with four parameters in total, an intercept and a spline with three degrees of freedom. The equation for this curve is as follows:

$$\log(LDF - 1) = A + Bt^1 + Ct^2 + Dt^3$$

Where, once again, superscripts here denote each resulting spline variable. The following substitution variables will be used:

$$X = \log\left(\frac{LDF_1 - 1}{LDF_2 - 1}\right) - \frac{t_1^1 - t_2^1}{t_3^1 - t_4^1} \times \log\left(\frac{LDF_3 - 1}{LDF_4 - 1}\right)$$

$$I = (t_1^2 - t_2^2) - \frac{(t_3^2 - t_4^2) \times (t_1^1 - t_2^1)}{t_3^1 - t_4^1}$$

$$J = (t_1^3 - t_2^3) - \frac{(t_3^3 - t_4^3) \times (t_1^1 - t_2^1)}{t_3^1 - t_4^1}$$

$$Y = \log\left(\frac{LDF_1 - 1}{LDF_3 - 1}\right) - \frac{t_1^1 - t_3^1}{t_2^1 - t_4^1} \times \log\left(\frac{LDF_2 - 1}{LDF_4 - 1}\right)$$

$$K = (t_1^2 - t_3^2) - \frac{(t_2^2 - t_4^2) \times (t_1^1 - t_3^1)}{t_2^1 - t_4^1}$$

$$L = (t_1^3 - t_3^3) - \frac{(t_2^3 - t_4^3) \times (t_1^1 - t_3^1)}{t_2^1 - t_4^1}$$

The equations for each of the four parameters are:

$$D = \frac{X - Y \times I / K}{J - L \times I / K}$$

$$C = \frac{(t_1^1 - t_2^1) \left[ \log\left(\frac{ldf_3 - 1}{ldf_4 - 1}\right) - D(t_3^3 - t_4^3) \right] - (t_3^1 - t_4^1) \left[ \log\left(\frac{ldf_1 - 1}{ldf_2 - 1}\right) - D(t_1^3 - t_2^3) \right]}{(t_3^2 - t_4^2)(t_1^1 - t_2^1) - (t_1^2 - t_2^2)(t_3^1 - t_4^1)}$$

$$B = \frac{\log\left(\frac{LDF_1 - 1}{LDF_2 - 1}\right) - C(t_1^2 - t_2^2) - D(t_1^3 - t_2^3)}{t_1^1 - t_2^1}$$

$$A = \log(LDF_1 - 1) - B t_1^1 - C t_1^2 - D t_1^3$$

To help facilitate implementation, the R code for performing this inversion is as follows, where *xx* represents *x* above, etc., *spline1*, *spline2*, and *spline3* are vectors which contain the three spline transformations, *age* is a vector which contains the four ages being used for the parameters, *b3* represents *D*, *b2* represents *C*, etc., and *g* represents the group index of each LDF curve:

```
xx[g] <- log( ( ldf.param1[g] - 1 ) / ( ldf.param2[g] - 1 ) ) - ( ( spline1[age[1]] - spline1[age[2]] ) / ( spline1[age[3]] - spline1[age[4]] ) ) * log( ( ldf.param3[g] - 1 ) / ( ldf.param4[g] - 1 ) )
ii[g] <- ( spline2[age[1]] - spline2[age[2]] ) - ( ( spline2[age[3]] - spline2[age[4]] ) * ( spline1[age[1]] - spline1[age[2]] ) ) / ( spline1[age[3]] - spline1[age[4]] )
jj[g] <- ( spline3[age[1]] - spline3[age[2]] ) - ( ( spline3[age[3]] - spline3[age[4]] ) * ( spline1[age[1]] - spline1[age[2]] ) ) / ( spline1[age[3]] - spline1[age[4]] )
yy[g] <- log( ( ldf.param1[g] - 1 ) / ( ldf.param3[g] - 1 ) ) - ( ( spline1[age[1]] - spline1[age[3]] ) / ( spline1[age[2]] - spline1[age[4]] ) ) * log( ( ldf.param2[g] - 1 ) / ( ldf.param4[g] - 1 ) )
kk[g] <- ( spline2[age[1]] - spline2[age[3]] ) - ( ( spline2[age[2]] - spline2[age[4]] ) * ( spline1[age[1]] - spline1[age[3]] ) ) / ( spline1[age[2]] - spline1[age[4]] )
ll[g] <- ( spline3[age[1]] - spline3[age[3]] ) - ( ( spline3[age[2]] - spline3[age[4]] ) * ( spline1[age[1]] - spline1[age[3]] ) ) / ( spline1[age[2]] - spline1[age[4]] )
b3[g] <- ( xx[g] - yy[g] * ii[g] / kk[g] ) / ( jj[g] - ll[g] * ii[g] / kk[g] )
b2[g] <- ( ( spline1[age[1]] - spline1[age[2]] ) * ( log( ( ldf.param3[g] - 1 ) / ( ldf.param4[g] - 1 ) ) ) - b3[g] * ( spline3[age[3]] - spline3[age[4]] ) ) - ( spline1[age[3]] - spline1[age[4]] ) * ( log( ( ldf.param1[g] - 1 ) / ( ldf.param2[g] - 1 ) ) ) - b3[g] * ( spline3[age[1]] - spline3[age[2]] ) ) ) / ( ( spline2[age[3]] - spline2[age[4]] ) * ( spline1[age[1]] - spline1[age[2]] ) - ( spline2[age[1]] - spline2[age[2]] ) * ( spline1[age[3]] - spline1[age[4]] ) )
b1[g] <- ( log( ( ldf.param1[g] - 1 ) / ( ldf.param2[g] - 1 ) ) ) - b2[g] * ( spline2[age[1]] - spline2[age[2]] ) - b3[g] *
```

*Credibility and Other Modeling Considerations For Loss Development Factors*

$$\frac{(\text{spline3}[\text{age}[1]] - \text{spline3}[\text{age}[2]])}{(\text{spline1}[\text{age}[1]] - \text{spline1}[\text{age}[2]])}$$

$$b0[g] <- \log(\text{ldf.param1}[g] - 1) - (b1[g] * \text{spline1}[\text{age}[1]]) - (b2[g] * \text{spline2}[\text{age}[1]]) - (b3[g] * \text{spline3}[\text{age}[1]])$$

For Excel, if the three spline transformations are on the top 3 rows of the spreadsheet for the four ages being used as parameters going across and starting at cell A1, and the four LDF parameters to be inverted are in cells A5 to D5, then the formulas for each of the intermediate parameters are as follows:

$$X = \text{LN}((A5-1)/(B5-1)) - (A1-B1)/(C1-D1) * \text{LN}((C5-1)/(D5-1))$$

$$I = (A2-B2) - ((C2-D2)*(A1-B1))/(C1-D1)$$

$$J = (A3-B3) - ((C3-D3)*(A1-B1))/(C1-D1)$$

$$Y = \text{LN}((A5-1)/(C5-1)) - (A1-C1)/(B1-D1) * \text{LN}((B5-1)/(D5-1))$$

$$K = (A2-C2) - ((B2-D2)*(A1-C1))/(B1-D1)$$

$$L = (A3-C3) - ((B3-D3)*(A1-C1))/(B1-D1)$$

If these formulas are placed on row 6 going across starting with column A, then the formulas for each of the curve parameters are as follows, assuming that these are placed in row 7 going across and starting from the column A:

$$D = \text{LN}(A5-1) - B7 * A1 - C7 * A2 - D7 * A3$$

$$C = (\text{LN}((A5-1)/(B5-1)) - C7 * (A2-B2) - D7 * (A3-B3)) / (A1-B1)$$

$$B = ((A1-B1) * (\text{LN}((C5-1)/(D5-1)) - D7 * (C3-D3)) - (C1-D1) * (\text{LN}((A5-1)/(B5-1)) - D7 * (A3-B3))) / ((C2-D2) * (A1-B1) - (A2-B2) * (C1-D1))$$

$$A = (A6 - D6 * B6 / E6) / (C6 - F6 * B6 / E6)$$

The equations for inverting a curve with only three parameters, an intercept and a spline with two degrees of freedom are shown below. The regression equation is as follows:

$$\log(LDF - 1) = A + Bt^1 + Ct^2$$

## *Credibility and Other Modeling Considerations For Loss Development Factors*

The following substitution variables are used:

$$X = t_1^1 - t_2^1$$

$$Y = t_1^2 - t_2^2$$

$$W = t_2^1 - t_3^1$$

$$Z = t_2^2 - t_3^2$$

The equations for inverting each of the variables are:

$$C = \frac{\frac{X}{W} \log\left(\frac{LDF_2}{LDF_3}\right) - \log\left(\frac{LDF_1}{LDF_2}\right)}{\frac{XZ}{W} - Y}$$

$$B = \frac{\frac{Y}{Z} \log\left(\frac{LDF_2}{LDF_3}\right) - \log\left(\frac{LDF_1}{LDF_2}\right)}{\frac{YW}{Z} - X}$$

$$A = \log(LDF_1) - B t_1^1 - C t_1^2$$

Formulas for R and Excel are not shown for this version since the equations are much less complicated than the four parameter version.



## 8. References

- [1] Dean, C., "Topics in Credibility Theory," Education and Examination Committee of the Society of Actuaries, 2005, <https://www.soa.org/files/pdf/c-24-05.pdf>
- [2] England P., and Verrall, R., "A Flexible Framework for Stochastic Claims Reserving," Proceedings of the Casualty Actuarial Society, 2001: LXXXVIII, pp. 1-38. <http://www.casact.org/pubs/proceed/proceed01/01001.pdf>
- [3] England, P. D. and Verrall, R. J. 2002. Stochastic Claims Reserving in General Insurance. British Actuarial Journal 8:443-544. [http://www.cassknowledge.com/sites/default/files/article-attachments/371~~richardverrall\\_-\\_stochastic\\_claims\\_reserving.pdf](http://www.cassknowledge.com/sites/default/files/article-attachments/371~~richardverrall_-_stochastic_claims_reserving.pdf)
- [4] Lowe, S. P., and Mohrman, D. F., "Extrapolating, Smoothing and Interpolating Development Factors [Discussion]," Proceedings of the Casualty Actuarial Society Casualty Actuarial Society, 1985: LXXII, pp. 182-189. <http://www.casact.org/pubs/proceed/proceed85/85182.pdf>
- [5] Sahasrabuddhe, R., "Claims Development by Layer: The Relationship between Claims Development Patterns, Trend and Claim Size Models," CAS E-Forum, Fall 2010, pp. 457-480 <https://www.casact.org/pubs/forum/10fforum/Sahasrabuddhe.pdf>
- [6] Sherman, R. E., "Extrapolating, Smoothing and Interpolating Development Factors," Proceedings of the Casualty Actuarial Society Casualty Actuarial Society, 1984: LXXI, pp. 122-155. <https://www.beanactuary.com/pubs/proceed/proceed84/84122.pdf>
- [7] Siewert, J. J., "A Model for Reserving Workers Compensation High Deductibles," CAS Forum, Summer 1996, pp. 217-244. <https://www.casact.org/pubs/forum/96sforum/96sf217.pdf>

## Biography of the Author

**Uri Korn** is an AVP & Actuary at Axis Insurance serving as the Research and Development support for all commercial lines of insurance. Prior to that, he was a Supervising Actuary at AIG in the Casualty pricing department. His work and research experience includes practical applications of credibility, trend estimation, increased limit factors, non-aggregated loss development methods, and Bayesian models. Uri Korn is a Fellow of the Casualty Actuarial Society and a member of the American Academy of Actuaries.