An Alternative Approach to Credibility for Large Account and Excess of Loss Treaty Pricing

Uri Korn, FCAS, MAAA

Abstract: This paper illustrates a comprehensive approach to utilizing and credibility weighting all available information for large account and excess of loss treaty pricing. The typical approach to considering the loss experience above the basic limit is to analyze the burn costs in these excess layers directly (see Clark 2011, for example). Burn costs are extremely volatile in addition to being highly right skewed, which does not perform well with linear credibility methods, such as Buhlmann-Straub or similar methods (Venter 2003). Using burn costs also involves developing and making a selection for each excess layer, which can be cumbersome. Also the formulas for calculating all of the correlations needed for determining the credibilities are complicated.

An alternative approach is shown that uses all of the available data in a more robust and seamless manner. Credibility weighting of the account's experience with the exposure cost for the basic limit is performed using Buhlmann-Straub credibility. Modified formulae are shown that are more suitable for this scenario. For the excess layers, the excess losses themselves are utilized to modify the severity distribution that is used to calculate the increased limit factors. This is done via a simple Bayesian credibility technique that does not require any specialized software to run. Such an approach considers all available information in the same way as analyzing burn costs, but does not suffer from the same pitfalls. Certain modifications are also illustrated to produce a method that does not differentiate between basic limit and the excess losses. Lastly, it is shown how the method can be improved for higher layers by leveraging Extreme Value Theory.

Keywords. Buhlmann-Straub Credibility, Bayesian Credibility, Loss Rating, Exposure Rating, Burn Cost, Extreme Value Theory

1. INTRODUCTION

This paper illustrates a comprehensive approach to utilizing and credibility weighting all available information for large account and excess of loss treaty pricing. The typical approach to considering the loss experience above the basic limit is to analyze the burn costs in these excess layers directly (see Clark 2011, for example). Burn costs are extremely volatile in addition to being highly right skewed, which does not perform well with linear credibility methods, such as Buhlmann-Straub or similar methods (Venter 2003). Using burn costs also involves developing and making a selection for each excess layer, which can be cumbersome. Also, the formulas for calculating all of the correlations needed for determining the credibilities are complicated.

An alternative approach is shown that uses all of the available data in a more robust and seamless manner. Credibility weighting of the account's experience with the exposure cost[1] for the basic limit

[1] Throughout this paper, the following definitions will be used:
Exposure cost: Pricing of an account based off of the insured characteristics and size using predetermined rates
Experience cost: Pricing of an account based off of the insured's actual losses. An increased limits factor is then usually applied to this loss pick to make the estimate relevant for a higher limit or layer.

is performed using Buhlmann-Straub credibility. Modified formulae are shown that are more suitable for this scenario. For the excess layers, the excess losses themselves are utilized to modify the severity distribution that is used to calculate the increased limit factors. This is done via a simple Bayesian credibility technique that does not require any specialized software to run. Such an approach considers all available information in the same way as analyzing burn costs, but does not suffer from the same pitfalls. Certain modifications are also illustrated to produce a method that does not differentiate between basic limit and the excess losses. Lastly, it is shown how the method can be improved for higher layers by leveraging Extreme Value Theory.

1.1 Research Context

Clark (2011) as well as Marcus (2010) and many others develop an approach for credibility weighting all of the available account information up an excess tower. The information considered is in the form of the exposure cost for each layer, the capped loss cost estimate for the chosen basic limit, and the burn costs associated with all of the layers above the basic limit up to the policy layer. Formulae are shown for calculating all of the relevant variances and covariances between the different methods and between the various layers, which are needed for calculating all of the credibilities.

This paper takes a different approach and uses the excess losses to modify the severity distribution that is used to calculate the ILF; this is another way of utilizing all of the available account information. This technique does not require the development and selection of a burn cost estimate for every excess layer. It also does not suffer from the problem of applying linear credibility methods, such as Buhlmann-Straub or similar methods, to highly skewed values, which can result in large errors (Venter 2003). Excess burn costs are definitely highly skewed.

1.2 Objective

The goal of this paper is to show how all available information pertaining to an account in terms of the exposure cost estimate and the loss information can be incorporated to produce an optimal estimate of the prospective cost.

1.3 Outline

Section 2 provides a review of account rating and gives a quick overview of the current approaches. Section 3 discusses credibility weighting of the basic layer loss cost, and section 4 shows strategies for credibility weighting the excess losses with the portfolio severity distribution. The end of this section shows simulation results to illustrate the relative benefit that can be achieved from this alternative

---

Burn Cost: Pricing of an excess account based off of the insured's actual losses in a non-ground up layer.

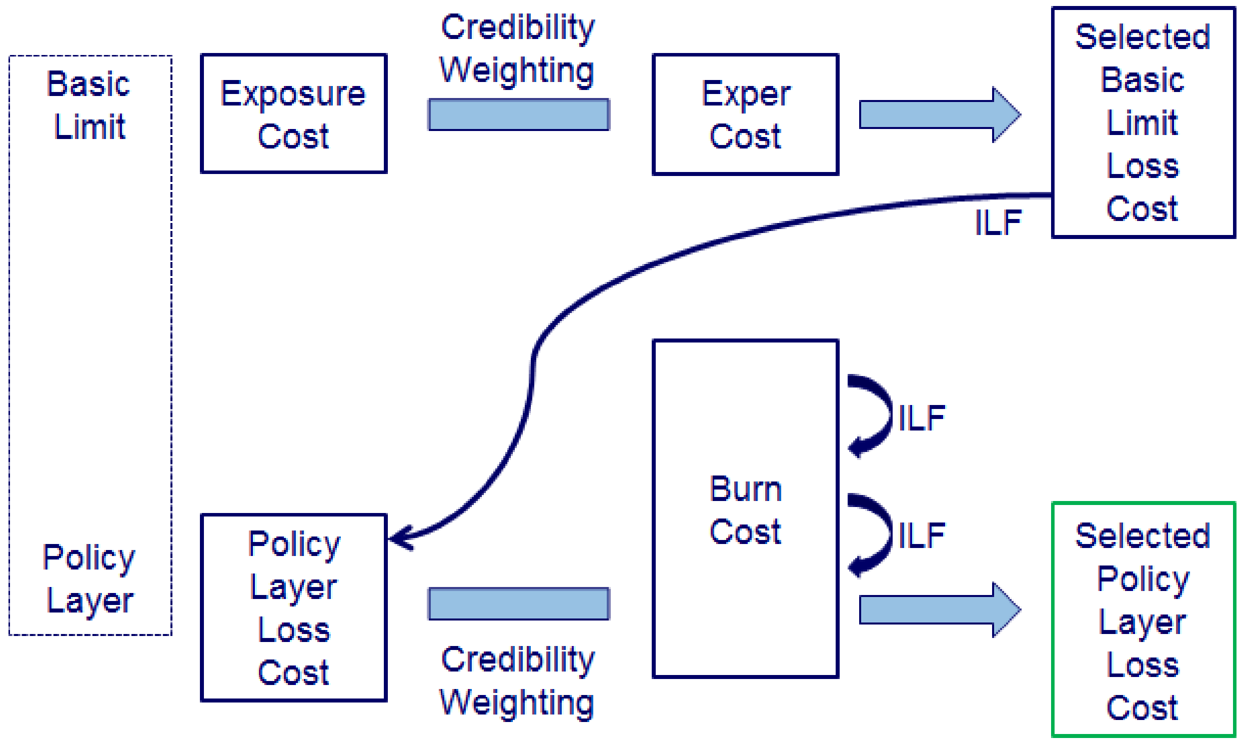
method, even with only a small number of claims. Section 5 shows how Extreme Value Theory can be leveraged to improve the method for high layers.

2. A BRIEF OVERVIEW OF ACCOUNT RATING AND THE CURRENT APPROACH

When an account is priced, certain characteristics about the account may be available, such as the industry or the state of operation. This information can be used to select the best exposure loss cost for the account, which is used as the a priori estimate for the account before considering the loss experience. The exposure loss cost can come from company data by analyzing the entire portfolio of accounts, from a large, external insurance services source, such as ISO or NCCI, from public rate filing information, from publicly available or purchased relevant data, or from judgment.

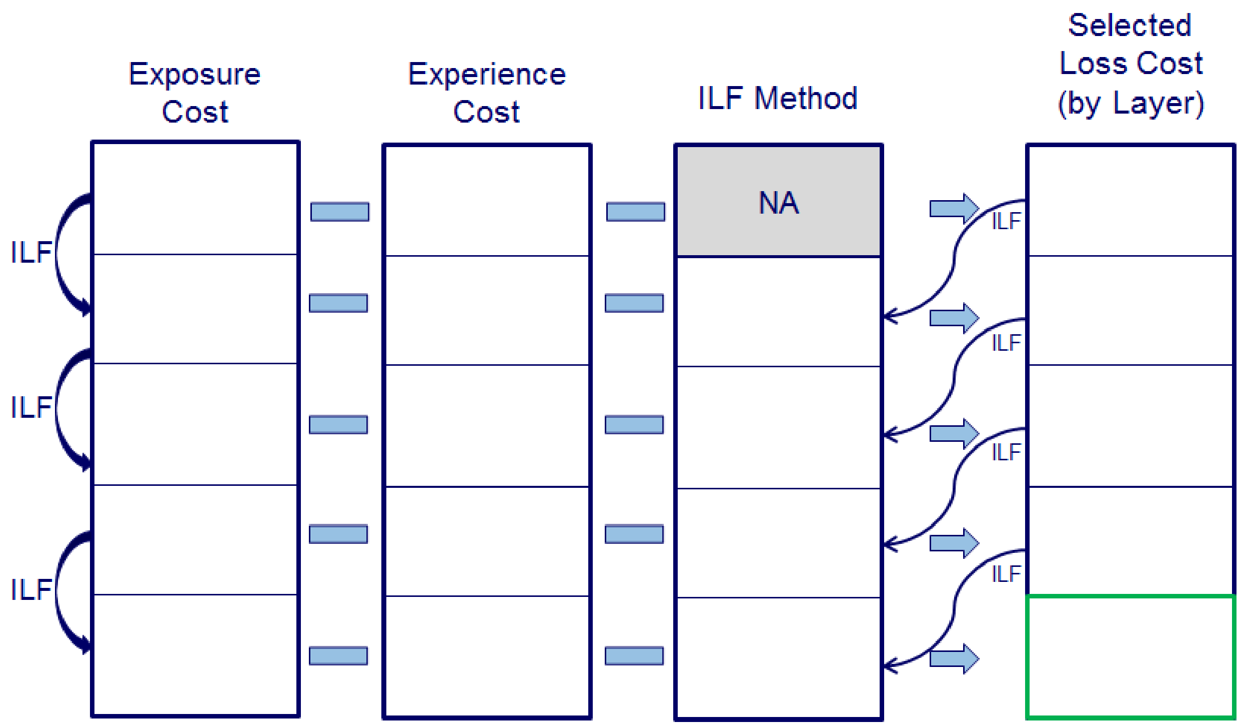
Very often, individual loss information is only available above a certain large loss threshold. Below this threshold, information is given in aggregate, which usually includes the sum of the total capped loss amount and the number of claims. More or less information may be available depending on the account. A basic limit is chosen, usually greater than the large loss threshold, as a relatively stable point in which to develop and analyze the account's losses. Once this is done, if the policy is excess or if the policy limit is greater than the basic limit, an ILF is applied to the basic limit losses to produce the loss estimate for the policy layer. It is also possible to look at the account's actual losses in the policy layer, or even below it but above the basic limit, which are known as the burn costs, as an another alternative estimate. The exposure cost is the most stable, but may be less relevant to a particular account. The loss experience is more relevant, but is usually more volatile, depending on the size of the account. The burn costs are the most relevant, but also the most volatile. Determining the amount of credibility to assign to each estimate can be difficult. Such an approach is illustrated in Figure 1 (where "Exper Cost" stands for the Experience Cost). The exact details pertaining to how the credibilities are calculated vary by practitioner.

Figure 1: Current Approach



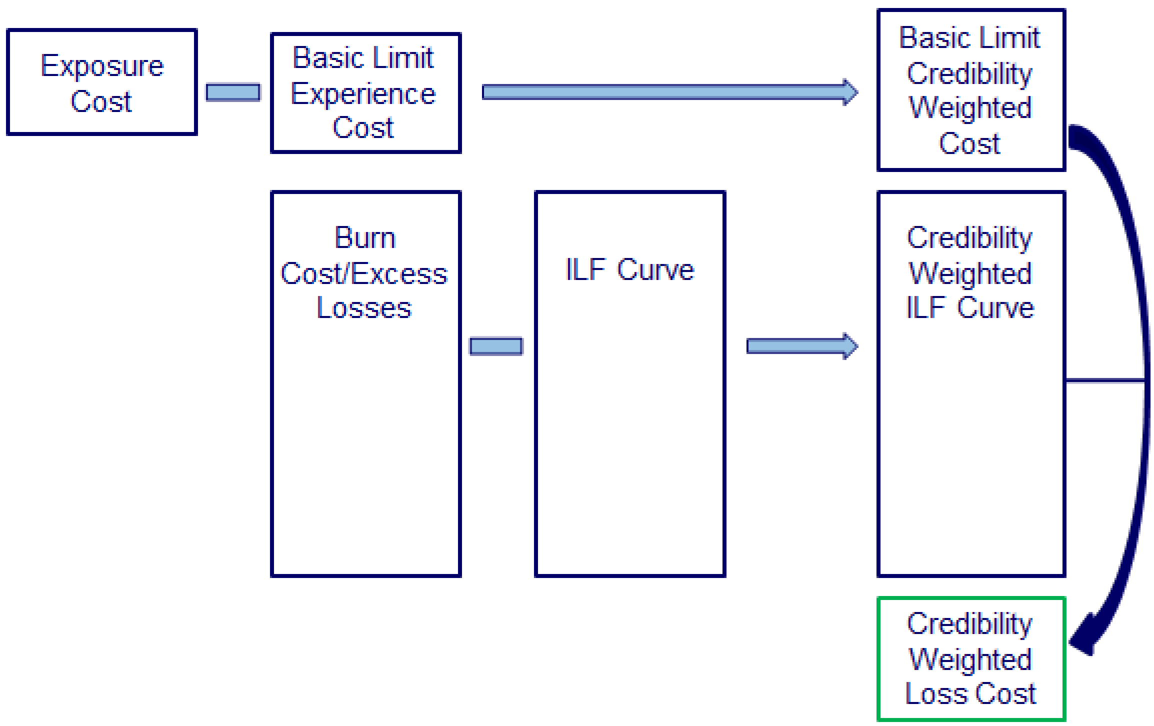
Clark (2011) developed a comprehensive approach to utilizing all of the data. For the basic limit, a selection is made based off of a credibility weighting between the exposure cost and the loss rating cost. For each excess layer, a credibility weighting is performed between the exposure cost multiplied by the appropriate ILF, the actual loss cost in the layer (i.e., the burn cost), and the previous layer's selection multiplied by the appropriate ILF. Formulas are shown for calculating all relevant variances and covariances, which are needed for estimating the optimal credibilities for each method in each layer. For further details on this method, refer to the paper. This approach is illustrated in Figure 2.

Figure 2: Clark's Method



The approach discussed in this paper is illustrated in Figure 3. It can be seen that all of the data that is used in Clark's approach is used here as well. Credibility weighting of the basic limit is expounded upon in the next section. Credibility weighting of the excess layers is discussed in section 4.

Figure 3: Proposed Method



3. CREDIBILITY WEIGHTING THE BASIC LAYER

Buhlmann-Straub credibility can be used to perform credibility weighting between the exposure cost and the account's actual losses in the basic layer. But account pricing is a bit different from the typical scenario of credibility weighting various segmentations in three ways:

1. Each item being credibility weighted has a different a priori loss cost (since the exposure costs can differ based on the class, etc.), that is, the complements are not the same. This also puts each account on a different scale. A difference of $1000 may be relatively large for one account, but not as large for another.

2. The expected variances differ between accounts since their losses may be capped at different amounts. The standard Buhlmann-Straub formulae assume that there is a fixed relationship between the variance and the exposures.

3. Additional information is available that can be used to improve the estimates in the form of exposure costs and ILF distributions, which can be used to calculate some of the expected values and variances.

Credibility can be performed either on the frequency and severity separately, or on the combined aggregate losses. The former will be discussed first. Accounting for trend and development is discussed at the end of the section. A related but off topic question of choosing the optimal capping point for the basic limit is discussed in Appendix D.

3.1 Separate Frequency and Severity

Splitting up frequency and severity often results in more robust estimates, although requires slightly more work than combining them. This section assumes that separate frequency and severity exposure estimates are available. If only a loss cost is available, the frequency can be calculated by dividing out the average capped severity using the appropriate ILF distribution (after removing legal expenses, if relevant). If various rating factors are applied to the initial loss cost estimate, for each one, it will need to be determined what percentage applies to the frequency versus the severity. Most factors are usually more related to frequency and so this can serve as the default assumption.

3.1.1 Frequency

It is usually assumed that the variance of the frequency is proportional to the mean (such as in Generalized Linear Models). Since the complements of credibility are different for each account, the expected variances are expected to differ as well, even for the same number of exposures. The Buhlmann-Straub within and between variance formulae assume a constant variance per number of exposures, but they can be modified to take this situation into account by dividing the variance component (that is, the square of the differences) by the frequency mean[2]. Doing this, the variances are calculated as a percentage of the expected frequency. The formulae are as follows:

$$\widehat{EPV} = \frac{\sum_{g=1}^{G} \sum_{n=1}^{N_g} e_{gn} \left(f_{gn} - \bar{f}_g \right)^2 / \bar{F}_g}{\sum_{g=1}^{G} \left(N_g - 1 \right)} \tag{3.1}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} e_g \left(\bar{f}_g - \bar{F}_g \right)^2 / \bar{F}_g - (G - 1)\,\widehat{EPV}}{e - \frac{\sum_{g=1}^{G} e_g^2}{e}} \tag{3.2}$$

[2] The proof of this formula is that it is similar to treating the number of exposures divided by the frequency as the weight in the formula, which is expected to be inversely proportional to the variance

Where *EPV* is the expected value of the process variance, or the "within variance", and *VHM* is the variance of the hypothetical means, or the "between variance". *G* is the number of groups, *N* is the number of periods, *e* is the number of exposures, f_{gn} is the frequency (per exposure) for group *g* and period *n*, $\bar{f}_g$ is the average frequency for group *g*, and $\bar{F}_g$ is the expected frequency for group *g* using the exposure costs. If the exposure frequency used comes from an external source, it can be seen that any overall error between it and the actual loss experience will increase the between variance and will thus raise the credibility given to the losses, which is reasonable. If this is not desired, the actual average frequency from the internal experience can be used instead in the formulae even if it is not used during the actual pricing. It can be seen that if the exposure frequency, $\bar{F}_g$, is the same for every account, these terms will cancel out in the resulting credibility calculations and the formulae will be identical to the original.

Equivalent formulae can also be used that utilize the actual claims counts instead of the frequency. These can be obtained by multiplying the numerator and denominator by the exposures.

$$\widehat{EPV} = \frac{\sum_{g=1}^{G} \sum_{n=1}^{N_g} (c_{gn} - \bar{c}_g)^2 / \bar{C}_g}{\sum_{g=1}^{G} (N_g - 1)} \tag{3.3}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} (\bar{c}_g - \bar{C}_g)^2 / \bar{C}_g - (G - 1)\,\widehat{EPV}}{e - \frac{\sum_{g=1}^{G} e_g^2}{e}} \tag{3.4}$$

Where *c* are the actual claim counts for the account and *C* are the claim counts using the exposure information for the account. This between variance can be calculated using a sample of actual accounts. The formula assumes that the between variance of accounts is proportional to the expected mean as well, which is a reasonable assumption. If the exposure costs used in the between variance formula (that is $\bar{F}$ or $\bar{C}$) do not utilize the same data being used to calculate the between variance (that is *f* or *c*), the bias correction component in the denominator can be removed and the formula becomes[3]:

[3] For a proof of this, refer to the appendix in Dean 2005

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} \; e_g \, (\bar{f}_g \; - \; \bar{F}_g \,)^2 \, / \, \bar{F}_g \; - \; (G \; - \; 1) \, \widehat{EPV}}{e} \tag{3.5}$$

Once the within and between variances are calculated, the credibility assigned to an account can be calculated as normal:

$$k \; = \; \frac{\widehat{EPV}}{\widehat{VHM}} \tag{3.6}$$

$$Z \; = \; \frac{e}{e \; + \; k} \tag{3.7}$$

If only claims above a certain threshold are being considered, and this threshold can differ by account, then different formulae are needed and are shown below. These formulae work by calculating the within and between variances on the excess frequencies, but then converting them to ground up variances before combining them so that all variances are at the same level. A full explanation is shown in Appendix C. For these formulae to work, the frequencies should be expressed relative to one unit of exposure. So if, for example, the exposure unit is $100,000 of revenue, an account/year with $500,000 of revenue should be counted as 5 exposures. The survival probability at the threshold is shown as *p*.

$$\widehat{EPV} \; = \frac{\sum_{g=1}^{G} \; \sum_{n=1}^{N_g} \; \{ \, [\, e_{gn} \, (f_{gn} \; - \; \bar{f}_g \; \times \; p_{gn})^2 \, / \, (\, \bar{F}_g \times \, p_{gn} \,) \; - \; 1] \, / \, p_{gn} \; + \; 1 \, \}}{\sum_{g=1}^{G} \; (\, N_g \; - \; 1 \,)} \tag{3.8}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} \; e_g \, \{ \, (\, \bar{f}_g \; - \; \bar{F}_g \; \times \; p_g \,)^2 \, / \, (\, \bar{F}_g \; \times \; p_g \,) \; - \; [\, (\, \frac{G \; - \; 1}{G} EPV \; - \; 1 \,) \, \widehat{\times \; p_g} \; + \; 1 \,] \, / \, e_g \, \} \, / \, p_g}{e \; - \; \frac{\sum_{g=1}^{G} \; e_g^2}{e}} \tag{3.9}$$

To calculate the credibility for an account, *k* can be calculated as below. Higher retentions (which have lower *p* values) will result in higher values of *k*, and thus lower credibility values.

$$\mathrm{k} = \frac{(\,EPV - 1\,) \times p + 1}{VHM \times p} \tag{3.10}$$

3.1.2 Severity

Buhlmann-Straub credibility can be used to perform credibility weighting on the account's basic limit average severity as well. The common assumption for severity is that the standard deviation is proportional to the mean, or equivalently, that the variance is proportional to the mean squared (such as in Generalized Linear Models, for example). The formulae can be modified to calculate the variances as a percentage of the square of the expected capped severity. The Buhlmann-Straub formulae that account for this are below:

$$\widehat{EPV} = \frac{\sum_{g=1}^{G} \sum_{n=1}^{N_g} c_{gn}\,(\,s_{gn} - \bar{s}_g\,)^2 \,/\, \bar{S}_g^2}{\sum_{g=1}^{G}\,(\,N_g - 1\,)} \tag{3.11}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} c_g\,(\,\bar{s}_g - \bar{S}_g\,)^2 \,/\, \bar{S}_g - (G - 1)\,\widehat{EPV}}{c - \frac{\sum_{g=1}^{G} c_g^2}{c}} \tag{3.12}$$

Where *c* is the claim count, s_{gn} is the average severity for group *g* and period *n*, $\bar{s}_g$ is the average severity for group *g* across all years, and $\bar{S}_g$ is the expected average severity for group *g* using the ILF distribution. These formulae use the actual (i.e., undeveloped) claim count as the weights, which is appropriate as the variance of the average severity equals the variance of the severity divided by the claim count.

For these formulae to work, however, the capping point would need to be the same for every account. They also do not take advantage of all available information, as the ILF distribution can be used to estimate the expected volatility. Modified formulae are shown below. The derivation of these formulae is shown in Appendix A.

$$\widehat{EPV}_{g,cap} = \frac{LEV2(cap) - LEV(cap)^2}{\bar{S}^2} \tag{3.13}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} c_g \left[(\bar{s}_g - \bar{S}_g)^2 / \bar{S}_g^{\,2} - \frac{(G-1)\,\widehat{EPV}_{g,cap}}{G\, c_g} \right]}{c - \frac{\sum_{g=1}^{G} c_g^2}{c}} \tag{3.14}$$

Where *LEV(x)* is the limited expected value at *x*. A separate EPV is calculated for each account, while a common VHM is used across all accounts regardless of the expected severity or capping point. The credibility for each account can then be calculated as normal using the account's calculated EPV and the portfolio calculated VHM.

$$k = \frac{\widehat{EPV}_{g,cap}}{\widehat{VHM}} \tag{3.15}$$

$$Z = \frac{c}{c + k} \tag{3.16}$$

Note that the final credibility depends on the number of reported claims and the EPV, which depends on the capping point. Higher capping points will produce higher EPV values and thus will be assigned lower credibility and vice versa.

3.2 A Combined Loss Cost Approach

Sometimes, it may be more desirable to develop and perform credibility on the aggregate losses with the frequency and severity combined. The common assumption for aggregate losses is that the variance is proportional to the average taken to some power between one and two (as in the Tweedie distribution), although these equations are far less sensitive to the power used than in GLM modeling. A common assumption is to set this power to 1.67 (Klinker 2011). Alternatively, it may be an acceptable simplification to assume that the standard deviation is roughly proportional to the mean and to use a power of two. The formulae for aggregate losses are as follows:

$$\widehat{EPV} = \frac{\sum_{g=1}^{G} \sum_{n=1}^{N_g} e_{gn} \left(l_{gn} - \bar{l}_g \right)^2 / \bar{L}_g^p}{\sum_{g=1}^{G} \left(N_g - 1 \right)} \tag{3.17}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} e_g \left(\bar{l}_g - \bar{L}_g \right)^2 / \bar{L}_g^p - (G - 1)\,\widehat{EPV}}{e - \frac{\sum_{g=1}^{G} e_g^2}{e}} \tag{3.18}$$

Where l_{gn} is the loss cost (per exposure) for group g and period n, $\bar{l}_g$ is the average loss cost for group g, $\bar{L}_g$ is the expected exposure loss cost for group g, and p is the Tweedie power used.

Similar to severity, this within variance formula does not handle different capping points and also does not leverage information from the severity distribution. Modified formulae are shown below. The derivation is shown in Appendix B.

$$\widehat{EPV}_{g,cap} = \frac{\left[LEV2(cap) - LEV(cap)^2 \right] \times \bar{F} + \widehat{EPV_f} \times \bar{F} \times \bar{S}^2}{\bar{L}^P} \tag{3.19}$$

$$\widehat{VHM} = \frac{\sum_{g=1}^{G} e \left[\left(\bar{l}_g - \bar{L}_g \right)^2 / \bar{L}_g^{\ P} - \frac{(G - 1)\,\widehat{EPV}_{g,cap}}{G\, e_g} \right]}{e - \frac{\sum_{g=1}^{G} e_g^2}{e}} \tag{3.20}$$

Similar to the above, if the loss costs come from an external source, the bias correction in the denominator can be removed. The credibility can now be calculated as normal. Similar to severity, higher loss caps will result in less credibility being assigned and vice versa.

3.3 Accounting for Trend and Development

Accounting for trend is relatively straightforward. All losses should be trended to the prospective year before all of the calculations mentioned above. The basic limit as well as the large loss threshold are trended as well, with no changes to procedure due to credibility weighting.

To account for development, a Bornhuetter-Ferguson method should not be used since it pushes each year towards the mean and thus artificially lowers the volatility inherent in the experience. Instead, a Cape Cod-like approach can be used, which allows for a more direct analysis of the experience itself. This method compares the reported losses against the "used" exposures, which results in the chain ladder estimates for each year, but the final result is weighted by the used exposures, which accounts for the fact that more volatility is expected in the greener years (Korn 2015a).

For frequency, the development factor to apply to the claim counts and the exposures is the claim count development factor. For severity, the actual claim count should be used since these are the exposures for the current estimate of the average severity. The actual average severity still needs to be developed though, since it has a tendency to increase with age. Severity development factors can be calculated by dividing the loss development factors by the claim count development factors (Siewert 1996), or the severity development can be analyzed directly to produce factors. The total exposures for each group should be the sum of the used exposures across all years.

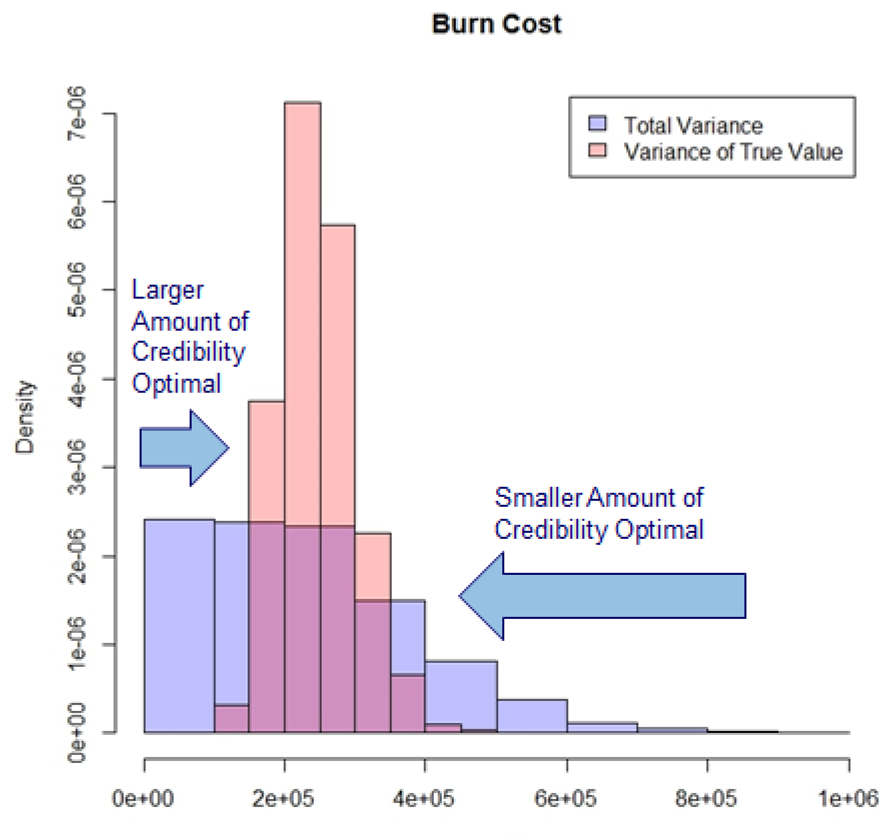
4. CONSIDERING THE EXCESS LOSSES

4.1 Introduction

Another source of information not considered in the basic layer losses are the excess losses, that is, the losses greater than the basic limit. These losses can be used to calculate the burn cost in each excess layer above the basic limit. After applying the appropriate ILF to each, if relevant, these values can serve as alternative loss cost estimates as well. In this type of approach, each of these excess layers needs to be developed separately, and credibility needs to be determined for each, which can be cumbersome. Calculating the credibility of each method in each layer requires the calculation of each variance as well as all of the correlations between them.

Burn costs are also right skewed, which do not perform well with linear credibility methods, as mentioned. To get a sense of why this is so, consider Figure 4, which shows the distribution of the burn cost in a higher layer produced via simulation. The majority of the time, the burn cost is only slightly lower than the true value (the left side of the figure). A smaller portion of the time, such as when there has been a large loss, the burn cost is much greater than the true value (the right side of the figure). For cases where the burn cost is lower than the true value and not that far off, a larger amount of credibility can be assigned to the estimate on average than when it is greater that the true value and is very far off. That is why linear credibility methods that assign a single weight to an estimate do not work well in this case.

Figure 4: Example of a Burn Cost Distribution



As an alternative, instead of examining the burn costs directly, the excess losses can be leveraged to modify the severity distribution that is used to calculate the increased limit factor. Such an approach considers all available information just as the direct burn cost approach does. It is also more robust as mentioned.

This remainder of this section discusses an implementation of this method and addresses various potential hurdles.

4.2 Method of Fitting

The first question to consider is what is the best fitting method when only a small number of claims, often only in summarized form, are available To answer this question a simulation was performed with only 25 claims and a large loss threshold of 200 thousand. See the following footnote for more details on the simulation[4]. For the maximum likelihood method, the full formula from section 4.8 was used but without the credibility component, which is discussed later. The bias and root mean square error (RMSE) was calculated by comparing the fitted limited expected values against the actual. The results are shown below in Table 1.

Table 1: Performance of Different Fitting Techniques

Method	Bias	RMSE (Thousands)
MLE	4.7%	194
CSP Error Squared	16.5%	239
CSP Error Percent Squared	13.5%	243
CSP Binomial	8.9%	209
LEV Error Percent Squared	55.2%	282
Counts Chi-Square	41.4%	256

CSP stands for conditional survival probability. The methods that utilized this sought to minimize either the squared error or the square of the percentage error, or performed MLE on these probabilities using a binomial distribution. Another method sought to minimize the squared percentage errors of the fitted and actual LEVs. The final method shown looked at the number of excess claims in each layer and sought to minimize the chi-squared statistic. It can be seen that the maximum likelihood has both the lowest bias and the lowest root mean square error (RMSE). (Using credibility will reduce the bias further to small amounts, as is shown with the simulated results in

[4] A lognormal was simulated with mean mu and sigma parameters of 11 and 2.5, respectively. The standard deviation of the parameters was 10% of the mean values. The policy attachment point and limit was both 10 million.

section 4.10.) It is also the most theoretically sound and the best for incorporating credibility, as is explained in the following section. For all of these reasons, maximum likelihood is used as the fitting method for the remainder of this paper.

Before deriving the likelihood formula for aggregate losses, first note that instead of applying an ILF to the basic limit losses, it is also possible to simply multiply an account's estimated ultimate claim count by the limited average severity calculated from the same severity distribution. The advantage of using an ILF is that it gives credibility to the basic limit losses, as shown below, where N is the estimated claim count for the account and $LEV(x)$ is the limited expected value calculated at x:

$$\begin{aligned} &Capped\ Losses\ \times ILF(Policy\ Layer) \\ &= N\ \times LEV_{Account}\ (Loss\ Cap)\ \times \frac{LEV_{Portfolio}(Policy\ Layer)}{LEV_{Portfolio}(Loss\ Cap)} \\ &= N\ \times LEV_{Portfolio}\ (Policy\ Layer)\ \times \frac{LEV_{Account}\ (Loss\ Cap)}{LEV_{Portfolio}\ (Loss\ Cap)} \end{aligned} \tag{4.1}$$

So applying an ILF is the same as multiplying an account's claim count by the portfolio estimated limited expected value at the policy layer, multiplied by an experience factor of the account's actual capped losses divided by the expected capped loss. This last component gives (full) credibility to the account's experience.

If individual claim data is only available above a certain threshold, which is often the case, there are three pieces of information relevant to an account's severity: the sum of the capped losses, the number of losses below the large loss threshold, and the number and amounts of the losses above the threshold. If the ILF method is used, the first component is already accounted for, and so only the two latter items should be considered[5]. (Including the first component when using an ILF actually produces slightly worse estimates, as is shown in the simulation results in section 4.10.) The claims below the threshold are left censored (as opposed to left truncated or right censored, which actuaries are more used to), since we are aware of the presence of each claim but do not know its exact value, similar to the effect of a policy limit. Maximum likelihood estimation can handle left censoring similar to how it handles right censoring. For right censored data, the logarithm of the survival function at

[5] Note that even though there may be some slight correlation between the sum of the capped losses and the number of claims that do not exceed the cap, as mention by Clark (2011), these are still different pieces of information and need to be accounted for separately.

the censoring point is added to log-likelihood. Similarly, for a left censored point, the logarithm of the cumulative distribution function at the large loss threshold is added to the log-likelihood. This should be done for every claim below the large loss threshold and so the logarithm of the CDF at the threshold should be multiplied by the number of claims below the threshold. Expressed algebraically, the formula for the log-likelihood is:

$$\sum_{x=Claims>LLT} PDF(x) + n \times CDF(LLT) \tag{4.2}$$

Where *LLT* is the large loss threshold, *PDF* is the logarithm of the probability density function, *CDF* is the logarithm of the cumulative density function, and n is the number of claims below the large loss threshold. The number of claims used in this calculation should be on a loss-only basis and claims with only legal payments should be excluded from the claim counts, unless legal payments are included in the limit and are accounted for in the ILF distribution. If this claim count cannot be obtained directly, factors to estimate the loss-only claim count will need to be derived for each duration.

As an example, assume that an account had a total of ten claims, three of which were above the large loss threshold of $100 thousand with the following values: $200 thousand, $500 thousand, and one million. The log-likelihood using a lognormal distribution for mu and sigma parameters of 10 and 2 would equal the following:

log(lognormal-pdf(200000, 10, 2)) + log(lognormal-pdf(500000, 10, 2))

+ log(lognormal-pdf(1000000, 10, 2)) + 7 x log(lognormal-cdf(100000, 10, 2))

Where *lognormal-pdf(a, b, c)* is the lognormal probability density function at *a* with mu and sigma parameters of *b* and *c,* respectively, and *lognormal-cdf(a, b, c)* is the lognormal cumulative density function at *a* with mu and sigma parameters of *b* and *c,* respectively.

4.3 Method of Credibility Weighting

Bayesian credibility will be used to incorporate credibility into the severity fit. This method performs credibility on each of the distribution parameters simultaneously while fitting the distribution and so is optimal to another approach that may attempt to credibility weight already fitted parameters. This method can be implemented without the use of specialized software. The distribution of maximum likelihood parameters is assumed to be approximately normally distributed. A normally distributed prior distribution will be used (which is the complement of credibility, in Bayesian terms), which is the common assumption. This is a conjugate prior and the resulting posterior distribution (the credibility weighted result, in Bayesian terms) is normally distributed as well. Maximum likelihood estimation (MLE) returns the mode of the distribution, which will also return the mean in the case, since the mode equals the mean for a normal distribution. So, this simple Bayesian credibility model can be solved using just MLE (Korn 2015b). It can also be confirmed that the resulting parameter values are almost identical whether MLE or specialized software is used.

To recap, the formula for Bayesian credibility is f(Posterior) ~ f(Likelihood) x f(Prior), or f(Parameters | Data) ~ f(Data | Parameters) x f(Parameters). When using regular MLE, only the first component, the likelihood, is used. Bayesian credibility adds the second component, the prior distribution of the parameters, which is what performs the credibility weighting. The prior used for each parameter will be a normal distribution with a mean of the portfolio parameter. The equivalent of the within variances needed for the credibility calculation to take place are implied automatically based on the shape of the likelihood function and do not need to be calculated, but the between variances do, which is discussed in section 4.4. This prior log-likelihood should be added to the regular log-likelihood. The final log-likelihood formula for a two parameter distribution that incorporates credibility is as follows:

$$\sum_{x=Claims > LLT} PDF(x, p1, p2) + n \times CDF(LLT, p1, p2) + Norm(p1, Porfolio\ p1, Between\ Var1) + Norm(p2, Porfolio\ p2, Between\ Var2) \tag{4.3}$$

Where *PDF(x, p1, p2)* is the logarithm of the probability density function evaluated at *x* and with parameters, *p1* and *p2*; *CDF(x, p1, p2)* is the logarithm of the cumulative density function evaluated at *x* and with parameters, *p1* and *p2*; and *Norm(x, p, v)* is the logarithm of the normal probability distribution function evaluated at *x*, with a mean of *p*, and a variance of *v*. *Portfolio p1* and *Portfolio p2* are the portfolio parameters for the distribution and *Between Var 1* and *Between Var 2* are the between variances for each of the portfolio parameters.

Using the same example from the previous section and assuming that the mu and sigma parameters for the portfolio are 11 and 3 with between variances of 1 and 0.5, respectively, the log-likelihood at mu and sigma parameters of 10 and 2 would be equal to the following:

log(lognormal-pdf(200000, 10, 2)) + log(lognormal-pdf(500000, 10, 2))

+ log(lognormal-pdf(1000000, 10, 2)) + 7 x log(lognormal-cdf(100000, 10, 2))

+ log(normal-pdf(10, 11, 1)) + log(normal-pdf(2, 3, 0.5))

Where *normal-pdf(a, b, c)* is the probability density function of a normal distribution at *a* with a mean of *b* and variance of *c*.

4.4 Calculating the Between Variance of the Parameters

Calculation of the variances used for the prior distributions can be difficult. The Buhlmann-Straub formulae do not work well with interrelated values such as distribution parameters. MLE cannot be used either as the distributions of the between variances are usually not symmetric and so the mode that MLE returns is usually incorrect and is very often at zero. A Bayesian model can be built, but this requires a specialized expertise that not everyone has and will not be discussed here. Another technique is to use a method similar to ridge regression which estimates the between variances using cross validation.

This method is relatively straightforward to explain and is quite powerful as well[6]. Possible candidate values for the between variance parameters are tested and are used to fit the severity distribution for each risk on a fraction of the data, and then the remainder of the data is used to evaluate the resulting fitted distributions. The between variance combination with the highest out-of-sample total likelihood is chosen. The calculation of the likelihood on the test data should not include the prior/credibility component. The fitting and testing for each set of parameters should be run multiple times until stability is reached, which can be verified by graphing the results. The same training and testing samples should be used for each set of parameters as this greatly adds to the stability of this approach. Simulation tests using this method (with two thirds of the data used to fit and the remaining one third to test) on a variety of different distributions are able to reproduce the actual between variances on average, which shows that the method is working as expected. Repeated n-fold cross validation can be used as well, but will not be discussed here.

4.5 Distributions with More Than Two Parameters

If the portfolio distribution has more than two (or perhaps three) parameters, it may be difficult to apply Bayesian credibility in this fashion. The method can still be performed as long as two

[6] One advantage of this approach over using a Bayesian model is that this method works well even with only two or three groups, whereas a Bayesian model tends to overestimate the prior variances in these cases. Though not relevant to this topic, as many accounts should be available to calculate the between variances, this is still a very useful method in general for building portfolio ILF distributions.

"adjustment parameters" can be added that adjust the original parameters of the severity distribution. For a mixed distribution, such as a mixed exponential or a mixed lognormal, one approach is to have the first adjustment parameter apply a scale adjustment, that is, to modify all claims by the same factor. The second adjustment parameter can be used to shift the weights forwards and backwards, which will affect the tail of the distribution if the individual distributions are arranged in order of their scale parameter. To explain the scale adjustment, most distributions have what is known as a scale parameter which can be used to adjust all claims by the same factor. For the exponential distribution, the theta parameter is a scale parameter, and so multiplying this parameter by 1.1, for example, will increase all claim values by 10%. For the lognormal distribution, the mu parameter is a log-scale parameter, and so to increase all claims by 10%, for example, the logarithm of 1.1 can be added to this parameter. For a mixed distribution, the scale parameter of each of the individual distributions should be adjusted.

One way to implement this is as follows, using the mixed exponential distribution as the example:

$$\theta_i' = \theta_i \times exp(Adj1) \tag{4.4}$$

$$R_i = W_i \times exp(i \times Adj2) \tag{4.5}$$

$$W_i' = R_i \,/ \sum R \tag{4.6}$$

Where *Adj1* and *Adj2* are the two adjustment parameters, *i* represents each individual distribution within the mixed exponential ordered by the theta parameters, *R* is a temporary variable, and *W* are the weights for the mixed distribution. Adjustment parameters of zero will cause no change, positive adjustment parameters will increase the severity, and negative adjustment parameters will decrease the severity.

4.6 Separate Primary and Excess Distributions

Sometimes a separate severity distribution is used for the lower and upper layers and they are then joined together in some fashion to calculate all relevant values. One way to join the distributions is to use the survival function of the upper distribution to calculate all values conditional on the switching point (that is, the point at which the first distribution ends and the second one begins), and then use the survival function of the lower distribution to convert the value to be unconditional again from ground up. The formulae for the survival function and for the LEV for values in the upper layer, assuming a switching point of p are as follows:

$$S(x) = S_U(x) \,/\, S_U(p) \times S_L(p) \tag{4.6}$$

$$LEV(x) = [LEV_U(x) - LEV_U(p)] \,/\, S_U(p) \times S_L(p) + LEV_L(p) \tag{4.7}$$

Where U indicates using the upper layer severity distribution and L indicates using the lower layer severity distribution. More than two distributions can be joined together in the same fashion as well.

Using this approach, both the lower and upper layer severity distributions can be adjusted if there is enough credible experience in each of the layers to make the task worthwhile. When adjusting the lower distribution, values should be capped at the switching point (and the survival function of the switching point should be used in the likelihood formula for claims greater than this point). When adjusting the upper distribution, only claim values above the switching point can be used and so the data should be considered to be left truncated at this point. Even if no or few claims pierce this point, modifying the lower layer severity distribution still affects the calculated ILF and LEV values in the upper layer since the upper layer lies on top of the bottom one using this or a similar approach.

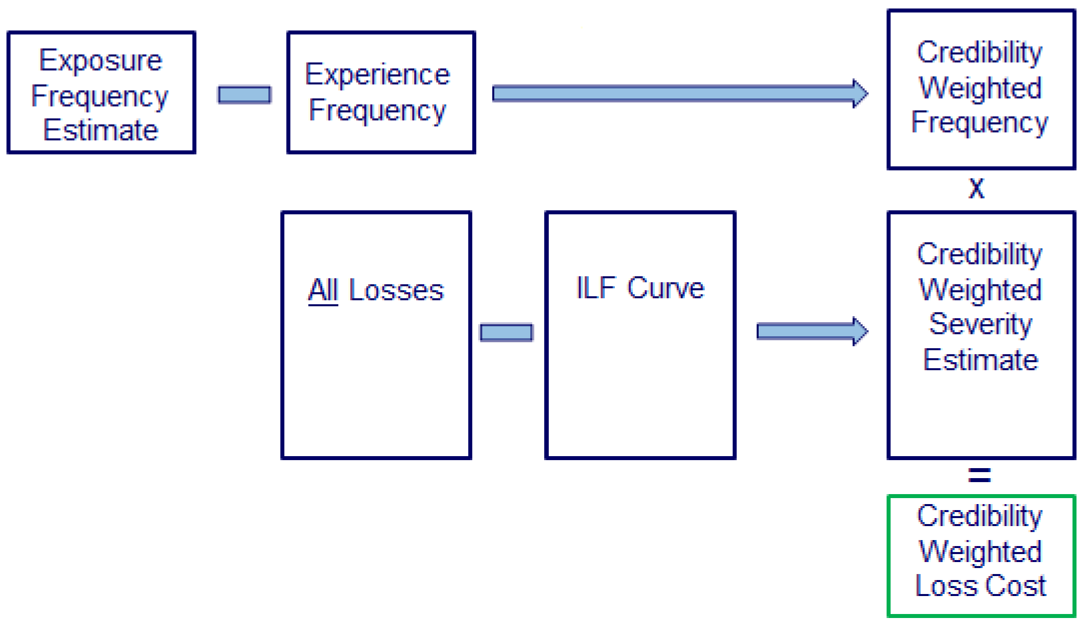
4.7 An Alternative when Maximum Likelihood Cannot be Used

Depending on the environment a pricing system is implemented in, an optimization routine required to determine the maximum likelihood may be difficult to find. An alternative is to calculate the log-likelihood for all possible parameter values around the expected using some relatively small increment. The exponent of these log-likelihoods minus the maximum log-likelihood can then be calculated to produce likelihoods that do not round to zero. These can then be used as the relative weights, and a weighted average of the parameter values can be calculated. The result should be very close to the MLE estimate.

4.8 An Alternative that Involves Combining All Severity Information Together

Using the approach mentioned thus far, the basic limit average severity is credibility weighted using the Buhlmann-Straub method, either by itself or included with the frequency in the aggregate losses, and the excess losses are credibility weighted using the Bayesian method mentioned. It is possible to simplify this procedure and incorporate both the basic limit severity as well as the excess severity in the same step. This can be accomplished by including the average capped severity in the likelihood formula used to modify the severity distribution. Then, instead of applying an ILF, the limited average severity in the policy layer can calculated from this credibility weighted severity distribution. Multiplying this by a (credibility weighted) frequency estimate produces the final result. (If only an exposure loss cost is available, this cost can be divided by the expected average severity in the layer to produce a frequency estimate.) This approach is illustrated below in Figure 5.

Figure 5: Proposed Approach Without a Basic Limit



Utilizing central limit theorem, it can be assumed that the average capped severity is approximately normally distributed. (Performing simulations with a small number of claims and a Box-Cox test justifies this assumption as well.) For very small number of claims, it is possible to use a Gamma distribution instead, although in simulation tests, using a Gamma does not seem to provide any benefit. The expected mean and variance of this normal or Gamma distribution can be calculated with the MLE parameters using the limited first and second moment functions of the appropriate distribution. The variance should be divided by the actual claim count to produce the variance of the average severity. For a normal distribution, these parameters can be plugged in directly; for a Gamma distribution, they can be used to solve for the two parameters of this distribution. The likelihood formula for this approach is as follows, including the credibility component:

$$\sum_{x=Claims > LLT} PDF(x, p1, p2) + n \times CDF(LLT, p1, p2) + Norm(Average\ Capped\ Severity, \mu, \sigma^2) + Norm(p1, Portfolio\ p1, Between\ Var1) + Norm(p2, Portfolio\ p2, Between\ Var2) \tag{4.8}$$

Where μ and σ^2 are calculated as:

$$\mu = LEV(Basic\ Limit, p1, p2)$$

$$\sigma^2 = [\ LEV2(Basic\ Limit, p1, p2) - LEV(Basic\ Limit, p1, p2)^2\ ] \,/\, m$$

Average Capped Severity is the average severity at the basic limit calculated from the account's losses, *n* is the number of claims below the large loss threshold, *m* is the total number of claims, and LEV2 is the second moment of the limited expected value. As above, *PDF*, *CDF*, and *Norm* are the <u>logarithms</u> of the probability distribution function, cumulative distribution function, and the normal probability density function respectively.

Using the example from sections 4.2 and 4.3 assuming that the average capped severity at $100 thousand is $70 thousand, the log-likelihood at mu and sigma parameters of 10 and 2 would be calculated as follows:

μ = lognormal-lev(100000, 10, 2)

σ^2 = [lognormal-lev2(100000, 10, 2) – lognormal-lev(100000, 10, 2)^2] / 10

log-likelihood = log(lognormal-pdf(200000, 10, 2)) + log(lognormal-pdf(500000, 10, 2))

+ log(lognormal-pdf(1000000, 10, 2)) + 7 x log(lognormal-cdf(100000, 10, 2))

+ log(normal-pdf(70000, μ, σ^2)

+ log(normal-pdf(10, 11, 1)) + log(normal-pdf(2, 3, 0.5))

Where *lognormal-lev(a, b, c)* is the lognormal limited expected value and *lognormal-lev2(a, b, c)* is the second moment of the lognormal limited expected value at *a* with mu and sigma parameters of *b* and *c*, respectively.

4.9 Accounting for Trend and Development

Both the losses and the large loss threshold should be trended to the prospective year before performing any of the above calculations. Using the likelihood formulae above (4.3 and 4.8), it is possible to account for different years of data with different large loss thresholds by including the parts from different years separately. Or alternatively, all years can be grouped together and the highest large loss threshold can be used.

There is a tendency for the claim severity of each year to increase with time since the more severe claims often take longer to settle. The claims data needs to be adjusted to reflect this. A simple approach is to apply the same amount of adjustment that was used to adjust the portfolio data to produce the final ILF distribution, whichever methods were used. With this approach, the complement of credibility used for each account should be the severity distribution before adjustment, and then the same parameter adjustments that were used at the portfolio level can be applied to these fitted parameters.

Another simple method is to assume that severity development affects all layers by the same factor. (This is the implicit assumption if loss development factors and burn costs are used.) The severity development factor for each year can be calculated by dividing the (uncapped) LDF by the claim count development factor, or it can be calculated directly from severity triangles. Each claim above the large loss threshold as well as the threshold itself should then be multiplied by the appropriate factor per year before performing any of the credibility calculations mentioned. Many more methods are possible as well that will not be discussed here.

4.10 Simulation

A simulation was conducted to help demonstrate the benefit this method can provide even with only a small number of claims. The results are shown in Tables 2-6. Results of using aggregate claim data with the likelihood formulae discussed in this paper as well as using the individual claim data were both calculated. Both of the aggregate likelihood methods, with and without the basic limits portion, were used. The errors were calculated on the total estimated losses for the policy layer. Tables 2-4 show the results of using a lognormal severity distribution: the first shows a lower excess layer, the second shows a higher excess layer, and the last shows a primary layer with a self insured retention. Table 5 shows the results for a mixed exponential and Table 6 shows the results for a mixed lognormal. (Simulations were also conducted with Gamma and Pareto distributions as well with similar results, but are not shown here for the sake of brevity.) Refer to the following footnote for more details on how the simulation was conducted[7]. All simulations used only 25 ground up claims.

[7] For the lognormal distribution, mean mu and sigma parameters of 11 and 2.5 were used, respectively. The standard deviation as well as the prior standard deviation assumed was 10% of the mean parameter values. The large loss threshold was 200 thousand, which translated to an average of 8.1 claims above the threshold. For the mixed exponential, the

Table 2: Lognormal Distribution with Attachment Point and Limit of $2 Million

Method	Bias: LEV Method	RMSE: LEV Method (Millions)	RMSE Relative to Portfolio ILF Method	Bias: ILF Method	RMSE: ILF Method (Millions)	RMSE Relative to Portfolio ILF Method
Portfolio	0.0%	3.05	+48.2%	0.0%	2.06	0.0%
Account Only (Full Credibility)	-0.5%	1.89	-8.4%	-0.5%	1.91	-7.2%
Credibility - Individual Claims	1.3%	1.41	-31.6%	3.1%	1.48	-28.3%
Credibility - Aggregate, Including Capped Sum	-0.3%	1.43	-30.5%	2.8%	1.49	-27.7%
Credibility - Aggregate, NOT Including Capped Sum	2.1%	1.46	-29.2%	3.6%	1.47	-28.7%

following mean mu values were used: 2863.5, 22215.7, 89355.0, 266664.3, 1108333.2, 3731510.8, 9309907.8, 20249975.1, 51141863.9, 230000000.0 and the following weights were used: 0.378297, 0.327698, 0.19941, 0.080178, 0.012106, 0.001764, 0.000362, 0.000125, 0.000048, 0.000012. The large loss threshold was 30 thousand which translated to an average of 8 claims above the threshold. The standard deviation of the adjustment parameters was 1 and 0.5. For the mixed lognormal, the mu parameters were 8 and 12, the sigma parameters were 2.5 and 2.7, and the weights were 75% and 25%. The large loss threshold was 25 thousand, which translated to an average of 8.7 claims above the threshold. The standard deviation of the adjustment parameters was 1 and 0.5. Simulating with certain mean parameter values and standard deviations would result in an average policy layer LEV that differed from the LEV calculated from the mean parameters, and so using these mean parameter values as the complement of credibility would cause a bias. Using prior values that result from fitting all of the data together would also not be exact as the data from multiple draws of a certain distribution with different parameter values would not necessarily be a perfect fit to that distribution, and a bias would show up as well. Instead, the prior parameters used for credibility weighting were adjusted together so that the result from using the average LEV would be unbiased. This is only an issue for simulation and would not be an issue in practice.

Table 3: Lognormal Distribution With Attachment Point and Limit of $10 Million

Method	Bias: LEV Method	RMSE: LEV Method (Millions)	RMSE Relative to Portfolio ILF Method	Bias: ILF Method	RMSE: ILF Method (Millions)	RMSE Relative to Portfolio ILF Method
Portfolio	0.0%	5.73	+28.1%	0.0%	4.47	0.0%
Account Only (Full Credibility)	5.2%	4.70	+5.1%	5.2%	4.73	+5.8%
Credibility - Individual Claims	3.4%	3.07	-31.4%	5.8%	3.20	-28.5%
Credibility - Aggregate, Including Capped Sum	2.7%	3.09	-30.9%	6.3%	3.24	-27.6%
Credibility - Aggregate, NOT Including Capped Sum	4.9%	3.10	-30.8%	7.2%	3.19	-28.6%

Table 4: Lognormal Distribution With a Limit of $2 Million and an SIR of $50 Thousand

Method	Bias: LEV Method	RMSE: LEV Method (Millions)	RMSE Relative to Portfolio ILF Method	Bias: ILF Method	RMSE: ILF Method (Millions)	RMSE Relative to Portfolio ILF Method
Portfolio	0.0%	5.16	85.0%	0.0%	2.79	0.0%
Account Only (Full Credibility)	-0.9%	2.63	-5.7%	-0.9%	2.69	-3.4%
Credibility - Individual Claims	0.2%	2.17	-22.2%	1.2%	2.33	-16.6%
Credibility - Aggregate, Including Capped Sum	-1.7%	2.27	-18.7%	0.7%	2.35	-15.7%
Credibility - Aggregate, NOT Including Capped Sum	0.7%	2.36	-15.4%	1.3%	2.31	-17%

Table 5: Mixed Exponential Distribution With a Limit of $2 Million and an SIR of $100 Thousand

Method	Bias: LEV Method	RMSE: LEV Method (Millions)	RMSE Relative to Portfolio ILF Method	Bias: ILF Method	RMSE: ILF Method (Millions)	RMSE Relative to Portfolio ILF Method
Portfolio	0.0%	1,007	+28.0%	-0.5%	787	0.0%
Account Only (Full Credibility)	8.6%	817	+3.8%	10.2%	848	+7.8%
Credibility - Individual Claims	-0.3%	561	-28.6%	0.5%	584	-25.8%
Credibility - Aggregate, Including Capped Sum	-4.3%	582	-26.1%	-2.2%	599	-23.9%
Credibility - Aggregate, NOT Including Capped Sum	-1.6%	590	-24.9%	-1.1%	588	-25.3%

Table 6: Mixed Lognormal Distribution With an Attachment Point and Limit of $10 Million

Method	Bias: LEV Method	RMSE: LEV Method (Millions)	RMSE Relative to Portfolio ILF Method	Bias: ILF Method	RMSE: ILF Method (Millions)	RMSE Relative to Portfolio ILF Method
Portfolio	0.0%	3.53	+31.8%	-0.8%	2.68	0.0%
Account Only (Full Credibility)	-18.3%	2.73	+1.9%	-20.1%	2.70	+0.6%
Credibility - Individual Claims	3.7%	1.98	-26.2%	5.5%	2.09	-22.1%
Credibility - Aggregate, Including Capped Sum	1.5%	2.06	-23.0%	4.6%	2.17	-19.1%
Credibility - Aggregate, NOT Including Capped Sum	3.8%	2.08	-22.3%	5.2%	2.11	-21.2%

As the results show, this method is able to provide substantial benefit over the basic approach of applying an ILF to the capped loss estimate, even with only a small number of claims. The biases are very low as well. For the lognormal distributions, the sigma parameter was multiplied by $n / (n - 1)$, where n is the claim count, which is a well-known adjustment for reducing the MLE bias of the normal and lognormal distributions[8]. (The biases are slightly larger for the higher excess accounts, but this is within an acceptable range for these layers, given the high estimation volatility.) To further reduce the bias, it is possible to conduct a simulation to estimate the approximate bias factor and then divide out the bias factor from each account's loss cost estimate, although this should not be necessary most of the time.

As expected, the LEV method (with the aggregate losses) is able to perform better than the ILF method (without the aggregate losses), since it also takes into account the credibility of the basic limit losses. Also, the LEV method performs best when taking into account the basic limit losses since more information is being included. The ILF method performs better when this is not included, since this information is already captured from applying an ILF, and including it in the likelihood double counts this piece of information. Although, the difference is not drastic.

5. USING EXTREME VALUE THEORY

A method of estimating the excess severity potential was illustrated in the previous section whereby an account's losses are credibility weighted against the portfolio loss distribution. Normally, extrapolating a severity distribution past the range of available data is not recommended (unless the distribution is a distribution that allows extrapolating such as the commonly used single parameter Pareto). But in this case, it is justified even if the account's losses are well below the policy layer since the portfolio's loss data is being considered as well, which hopefully contains some losses near the layer being estimated. However, using an account's losses to predict the expected severity of a higher excess layer would not be recommended if full credibility was being assigned to the account's losses.

An alternative which may yield more accurate results is to work with a Generalized Pareto Distribution (GPD), which is the statistically recommended method of predicting severity potential in excess of the available data. Based on the Peak Over Threshold method of Extreme Value Theory, excess severity potential can be estimated by fitting a GPD to the loss data above a certain threshold.

---

[8] This adjustment cancels out most of the negative parameter bias. In this case, not applying this adjustment would have probably resulted in a lower overall bias. The positive part of the bias comes from the transformation involved in the LEV or ILF calculation, since even if the parameter mean value estimates are unbiased, applying a function to these estimates can create some bias, as Jensen's inequality states. As long as the parameter errors are not too great, the bias will remain small. The credibility weighting being performed reduces the parameter errors, and as a result, the bias as well.

(See McNeil 1997 for application to estimating loss severity.)

According to the theory, a GPD will better fit the data that is further into the tail, and so a higher threshold may provide a better fit. But there is a tradeoff since selecting a higher threshold will make less data available to analyze, which will increase the prediction variance. Looking at graphs of fitted versus empirical data is the typical way to analyze this trade off and to select a threshold. Although other methods are available. (See Scarrott & MacDonald 2012 for an overview.)

Returning to account rating, discarding losses below a certain threshold has an intuitive appeal as well. It is often debated how relevant smaller losses are to the severity potential of high excess layers. Using this method can help provide a statistical framework for evaluating the most relevant data to use for predicting expected excess loss potential. As to whether this method can be used in practice, testing a bunch of individual accounts in different commercial lines of business, the GPD seemed to provide a good fit to the accounts' losses above a certain threshold, even where a GPD may not be the ideal loss distribution for the portfolio.

To fit the GPD, the likelihood formula shown above should not be used, as the likelihood is simply the probability density function. Setting the threshold parameter to the appropriate value already takes the left truncation of the data into account, that is, that no losses below the threshold are being included. The fitted distribution will be conditional on having a claim of at least the threshold. Multiplying the calculated severity at the policy layer obtained from the fitted GPD by the expected excess claim count at the threshold will yield the loss cost. Formulas 3.8-3.10 shown above for excess frequency can be used to credibility weight the excess claim count estimate as well.

This method can be implemented using any type of distribution (or distributions, as explained in section 4.6) for the portfolio. Recall that Bayes' formula is being used for credibility: f(Parameters | Data) = f(Data | Parameters) x f(Parameters). Credibility is performed by calculating the prior likelihood on the parameters. It is also possible to reparameterize the distribution and use other new parameters instead. In this case, the logarithm of the instantaneous hazard function (that is, $f(x) / s(x)$) will be used for the new parameters, the same number as the number of parameters in the distribution. These are used since they are approximately normally distributed, work well, and are also not dependent on the selected threshold since they are conditional values. Using these values, it is possible to solve for the original distribution parameters since there are the same number of unknowns as equations. Then the original distribution parameters can be used to calculate any value from the distribution, such as PDFs and CDFs. Since this is the case, that any distribution value can be calculated from these new parameters, they can be thought of as the new parameters of the distribution, and the prior likelihood can be calculated on these new parameters instead. As a trick, instead of actually solving for the original parameters, we can effectively "pretend" that they were solved for by still using the original parameters as the input to the maximization routine but just calculating the prior likelihood on the new parameters, that is the instantaneous hazard values, since the results will be exactly the same. In practice, we suggest using the difference in the hazard functions for each addition parameter since it makes the parameters less correlated and seems to work better in simulation tests. In summary, the likelihood equation is as follows, assuming a two parameter distribution:

$$
\begin{gathered}
p1 = log(\,f(t_1)\,/\,s(t_1)\,) \\
p2 = log(\,f(t_1)\,/\,s(t_1) - f(t_2)\,/\,s(t_2)\,) \\
loglik = \sum_i GPD(\,x_i, \alpha, \beta, threshold\,) + N(\,p1, h\,\alpha v1\,) + N(\,p2, h\,\alpha v2\,)
\end{gathered}
\tag{5.1}
$$

Where *GPD* is the <u>logarithm</u> of the PDF of the GPD distribution, *N* is the <u>logarithm</u> of the normal PDF, t_1 and t_2 are the two points chosen to calculate the instantaneous hazards, x are the claim values, α and β are the fitted GPD parameters, *threshold* is the selected threshold, *h1* and *h2* are the credibility complements for the logarithm of the hazard functions from the portfolio, and *v1* and *v2* are the between variances for the complements.

The complement of each of these new parameters can be calculated from the portfolio distribution, even if it is not a GPD, and the between variances can be solved for in the same manner as described above in section 4.4. (This method is also helpful in allowing the threshold of the GPD to vary.)

Simulations were conducted using this method simulating from a lognormal distribution[9]. The results are shown below in Table 7.

Table 7: GPD performance for various layers

Layer (Limit xs Retention, In Millions)	Bias	Improvement in RMSE (From Using Portfolio Severity Estimate)
10 xs 10	+3.5%	34.9%
25 xs 25	+3.6%	31.3%
50 xs 50	+6.0%	28.3%

[9] Claims were simulated from a lognormal distribution with parameters 11 and 2.5, respectively. 1000 iterations were performed. The between standard deviation of the lognormal parameters was 1 and 0.5, respectively. This was used to calculate the between standard deviations of the transformed parameters, that is, the hazard functions. 50 claims were simulated, a threshold of 250 thousand was used, and there was an average of 14.5 claims above the threshold.

It is difficult to say how this method will perform relative to the more basic method explained above. This is an area for further research.

6. CONCLUSION

This paper discussed an alternative technique to considering all relevant information to produce the most accurate estimate of an account's loss expectation. In doing so, an actuary pricing an account or treaty can be confident that the most relevant and stable estimates are being used.

Appendix A

To derive the Buhlmann-Straub formulae for average capped severity, we first note that to calculate the variance of the average severity using the within variance formula shown above, the EPV needs to be divided by the claim count and then multiplied by the severity squared. (This can be seen from the fact that formula 3,11 multiplies by the claim counts as weights, but does not divide by them afterward.)

$$Var(Average\ Severity) = \frac{EPV \times \bar{S}^2}{c}$$

The formula for the variance of the average severity using the severity distribution is:

$$Var(Average\ Severity) = \frac{LEV2(cap) - LEV(cap)^2}{c}$$

Where LEV(x) is the limited expected value capped at x and LEV2 is the second moment of the limited expected value. Setting these two equations equal to each other and solving for the EPV produces the following:

$$EPV_{g,cap} = \frac{LEV2(cap) - LEV(cap)^2}{\bar{S}^2}$$

This EPV can be used in the VHM (between variance) formula instead of using the traditional Buhlmann-Straub EPV formula.

Appendix B

To derive the formula for the EPV for aggregate losses, we first need to derive the formula for the variance of the loss cost estimate. To calculate the variance of the loss cost, we first separate the frequency and severity components. The expected variance of the frequency is:

$$Var(f) = \frac{EPV_f}{e} \times \bar{F}$$

Where EPV_f is the within variance parameter for the frequency. We needed to divide by the exposures since, as explained by severity, the EPV formula multiplies by the exposures as weights but does not divide by them afterwards. We then multiplied by the expected frequency since the within variance formula we used was calculated as a percentage of the expected frequency. To get the expected variance of c, the claim count, this quantity needs to be multiplied by the square of exposures (since the claim count equals the frequency multiplied by the exposures):

$$Var(c) = \frac{EPV_f}{e} \times \bar{F} \times e^2 = EPV_f \times \bar{C}$$

Where $\bar{C}$ is the claim count expected using the exposure frequency.

The variance of the severity equals:

$$Var(s) = LEV2(cap) - LEV(cap)^2$$

Using the formula for aggregate variance, the variance of the aggregate losses can be calculated as:

$$Var(a) = [LEV2(cap) - LEV(cap)^2] \times \bar{C} + EPV_f \times \bar{C} \times \bar{S}^2$$

Where *a* are the aggregate losses. The variance of the loss cost per unit of exposure equals this divided by the exposures squared, which produces the following:

$$Var(l) = \frac{[LEV2(cap) - LEV(cap)^2] \times \bar{F} + EPV_f \times \bar{F} \times \bar{S}^2}{e}$$

This is the variance of the loss cost estimate. This will be used to back into the EPV parameter needed in the between variance formula. The variance of the loss cost (divided by the exposures) can also be written as the following, using the EPV parameter:

$$Var(L) = \frac{EPV_l}{e} \times L^p$$

Where EPV_t is the EPV of the loss cost. Setting these two equations equal to each other produces the following for the EPV:

$$EPV_{cap} = \frac{[LEV2(cap) - LEV(cap)^2] \times \bar{F} + EPV_f \times \bar{F} \times \bar{S}^2}{L^p}$$

Appendix C

To be able to convert the process variance from ground up to excess and vice versa, we first need to derive the formula for the excess variance-to-mean ratio relative to the ground up variance-to-mean ratio. To do this, the formula for aggregate variance can be used, where G is the aggregate cost, N is the claim count, and X is the severity: $V[G] = V[N]\ E[X]^2 + V[X]\ E[N]$.

In this case where the aggregate is the excess claim count, the severity is one if the claim pierces the threshold and zero otherwise. This severity follows a Bernoulli distribution which has a variance equal to p x (1 - p), where p is the probability of exceeding the threshold, which is the survival probability. Performing some algebra and then dividing by the excess frequency (E[N] x p) in order to derive the excess variance-to-mean ratio, produces the following[10]:

$$E[X] = p$$

$$V[X] = p\ (1 - p)$$

$$E[G] = Np$$

$$V[G] = V[N]\ p^2 + p\ (1 - p)\ N$$

$$XS\ VTM = V[G]\ /\ E[G] = V[N]\ p\ /\ N + 1 - p = GU\ VTM \times p + 1 - p$$

$$XS\ VTM = (GU\ VTM - 1) \times p + 1$$

Where *XS VTM* is the excess variance-to-mean ratio and *GU VTM* is the ground up variance-to-mean ratio. This is the formula that will be used for converting the process variance from ground up to excess.

To convert the between variance, the variance should be multiplied by the square of p, following the formula for the variance of the product of a constant (the expected value of p) and a random variable. To convert a between variance-to-mean ratio (as is used in the formulae of this paper), the numerator is multiplied by the square of p and the denominator is multiplied by p, and so the ratio needs to be multiplied by p.

These results can now be plugged into the Buhlmann-Straub formulae, which produces the results shown in the paper. The *EPV* formula works by calculating the excess variance-to-mean ratios, and then converting each to a ground up variance-to-mean ratio so that they are compatible before

[10] Thanks to Aaron Curry for helping me derive this formula

summing them together. The resulting *EPV* is then a ground up value. For the *VHM* formula, first the squared differences of the excess frequency are calculated. The *EPV*, which is the process variance, is converted to an excess EPV and then subtracted out from the total variance, which leaves over the squared differences related to the (excess) between variance. The results are then converted to ground up values by dividing by the survival probabilities so that they are compatible and are then combined as normal as in the original formula. The k credibility ratio can then be calculated for an excess threshold by converting the *EPV* and *VHM* values from ground up to excess and then using the original formula on these values[11].

[11] A simulation was also performed to double check the formulae and it was confirmed that the *EPV* and *VHM* values calculated were equivalent on average whether calculated from the ground up frequencies using the original formulae or from the excess frequencies using the revised formulae. It was also confirmed that the credibilities being calculated for the excess frequencies were optimal, and that either increasing or decreasing the credibility resulted in larger average errors from the known true values.

Appendix D

A related issue to loss rating credibility is selecting the most optimal basic limit to develop and multiply by an ILF. Choosing a lower basic limit helps the capped loss cost estimate be more stable, but involves the application of a more leveraged increased limit factor and also uses less of an account's individual experience (unless credibility weighting is being performed). The variance formulae and ideas discussed in this paper can be used to calculate the expected volatility of each capping point, and the capping point with the lowest volatility can then be chosen as the optimal.

The following formula can be used to calculate the variance of the loss cost estimate for the basic layer. (This formulae was discussed in Appendix B.)

$$Var(a) = [LEV2(cap) - LEV(cap)^2] \times \bar{C} + EPV_f \times \bar{C} \times \bar{S}^2$$

Note that the actual variance of an account's experience is not used, as this would be subject to a large amount of error due to volatility. Applying credibility as discussed will decrease this variance. Using the formula for a normal conjugate prior as an approximation, the inverse of the final variance equals the sum of the inverse of the within variance and the inverse of the between variance. (Bolstad 2007)

The variance of the ILF should include both the parameter estimation error as well as the variance of the differences between accounts to reflect the fact that a portfolio ILF may be less relevant for a specific account. The parameter uncertainty of the estimated parameters can be calculated by taking the matrix inverse of the Fisher information matrix, which contains the second derivatives of the likelihood. Most statistical packages have methods to calculate this automatically as well. These can be calculated numerically as well, as follows: the derivative of the likelihood can be calculated by taking the difference of the likelihood at a parameter value slightly above the MLE estimate and slightly below, and then dividing this by the difference of these two points chosen. Similarly, the second derivatives can be obtained by taking the derivative of these values.

Estimating the variance between accounts for the distribution parameters was discussed in section 4.4. The parameter variances for the within and between variances can be summed to derive the total parameter variance if credibility is not being performed. Otherwise, the total parameter variance can be calculated directly using the Fisher Information matrix resulting from the fitting the parameters using Bayesian credibility. Or it can be estimated by using the same variance formula for a conjugate normal that was mentioned. If the source of the ILFs are from ISO or a similar source, the parameter error will be difficult to calculate, but can be assumed to be small relative to the between companies parameter variance, which will need to be judgmentally selected.

Once the parameter variance has been obtained, it can be used to calculate the variance of the ILF using simulation. (The delta method can be used as well, but will not be discussed here.)

Once we have the variance of the capped loss pick and the variance of the ILF, the variance of the loss cost estimate for the account layer can be calculated by using the variance of a product formula: $V[A \times B] = V[A]\,V[B] + V[A]\,E[B]^2 + V[B]\,E[A]^2$. Using all of this, the total variance of each potential loss pick can be calculated at each capping level, and the capping level with the lowest variance can be selected. This variance depends on the number of ground up claims and the retention and limit of the policy being priced (for each ILF table, that is). A table of the optimal capping levels can be built by these dimensions and then the appropriate capping level can be looked up for each account, or alternatively, this value can also be computed in real time when pricing an account.

To help illustrate this method, a simulation was performed with varying amounts of ground up claims on a fictional account with both a retention and limit of one million (without considering credibility). Figure 6 shows the estimated variance of the final estimated loss cost (divided by the average variance, so that all of the variances can appear on the same graph) at various capping levels for different number of ground up claims. Note how the curves decrease initially and then start to increase. At lower capping levels, the total variance is higher due using less information about the account's actual losses as well as the increased uncertainty in the ILF. At higher capping levels, the variance starts increasing again due to more volatility in the capped loss pick. The point in between with the lowest variance is the optimal capping point. Note how the variance changes at a slower rate and is more stable with more ground up claims. Figure 7 summarizes the results and shows the optimal capping points for each amount of ground up claims. As expected, a higher capping level should be chosen for larger accounts with more ground up claims.

Figure 6: Relative variance by capping level

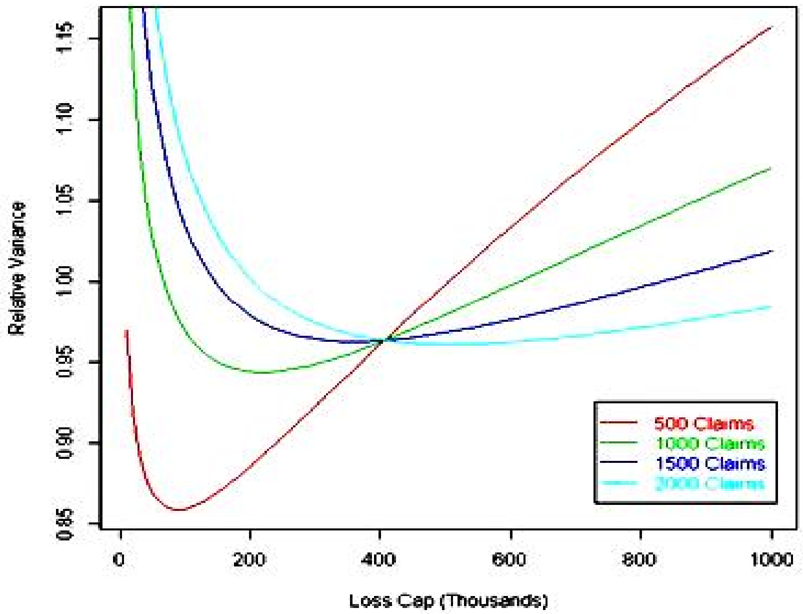
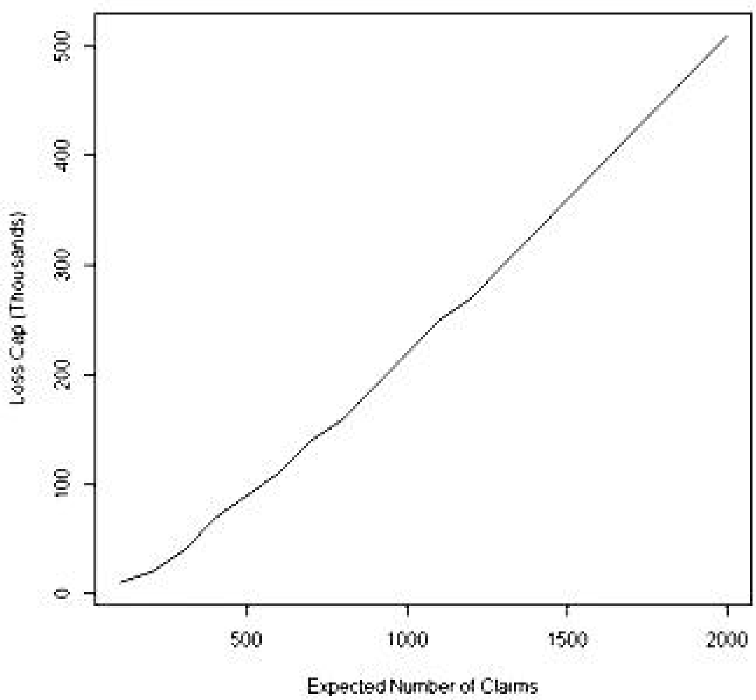


Figure 7: Optimal capping point



7. REFERENCES

[1] Bolstad, W., "Introduction to Bayesian Statistics (Second Edition)," 2007, p.106-108, 207-209

[2] Clark, D. R., "Credibility for a Tower of Excess Layers," Variance Casualty Actuarial Society, 2011, Volume 05, Issue 01, p.32-44. http://www.variancejournal.org/issues/05-01/32.pdf

[3] 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

[4] Klinker, Fred., "GLM Invariants." Casualty Actuarial Society E-Forum, Summer 2011. 2011. https://www.casact.org/pubs/forum/11wforumpt2/Klinker.pdf

[5] Korn, U. "Credibility for Pricing Loss Ratios and Loss Costs." CAS E-Forum, Fall 2015. http://www.casact.org/pubs/forum/15fforum/Korn.pdf

[6] Korn, U. "Credibility and Other Modeling Considerations for Loss Development Factors." CAS E-Forum, Summer 2015. http://www.casact.org/pubs/forum/15sumforum/Korn.pdf

[7] Marcus, L. F., "Credibility for Experience Rating, A Minimum Variance Approach," CAS E-Forum, Summer 2010. https://www.casact.org/pubs/forum/10suforum/Marcus.pdf

[8] McNeil, Alexander J. "Estimating the tails of loss severity distributions using extreme value theory." ASTIN bulletin 27.01 (1997): 117-137. http://dx.doi.org/10.2143/AST.27.1.563210

[9] Scarrott, Carl, and Anna MacDonald. "A review of extreme value threshold estimation and uncertainty quantification." REVSTAT–Statistical Journal10.1 (2012): 33-60. https://www.ine.pt/revstat/pdf/rs120102.pdf

[10] 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

[11] Venter, G., "Credibility Theory for Dummies," Casualty Actuarial Society Forum, Winter 2003, pp. 621–627. https://www.casact.org/pubs/forum/03wforum/03wf621.pdf

Biography of the Author

Uri Korn is the Industry Analytics Leader for AIG, Client Risk Solutions. Prior to that, he was an AVP & Actuary at Axis Insurance serving as the Research and Development support for all commercial lines of insurance. 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.