NCCI's 2014 Excess Loss Factors

by Dan Corro and Yen-Chieh Tseng

Abstract

An excess loss factor is a measure of expected loss that is in excess of a given per-occurrence limit. The National Council on Compensation Insurance (NCCI) uses excess loss factors in its retrospective rating plan as well as in aggregate and class ratemaking.

NCCI computes annual updates of excess loss factors by state and hazard group for certain limits ranging from $10,000 to $10,000,000. These annual updates are filed with regulators in 37 NCCI states.

Periodically, NCCI reviews the methodology behind these annual updates. Such a review concluded in 2014 and made significant changes to the methodology used. This paper describes the new methodology and highlights some improvements over prior NCCI methodologies.

Key Words

Excess ratio, excess loss factor, dispersion, Extreme Value Theory, claim severity distribution, Bayesian statistical model

1. Introduction

NCCI uses the term *excess ratios* or *excess loss factors* (ELFs) to refer generally to ratios where the numerator is some measure of expected amount of losses in excess of a given limit, and the denominator is a corresponding measure related to expected unlimited losses.[1] Periodically, NCCI reviews and updates the ELF methodology. The 2014 update made many improvements over the 2004 update. These improvements are discussed in this paper.

Most private employers in the United States are required to provide workers compensation coverage to pay for lost wages and medical expenses arising from work injuries. *Hazard groups*, as described by Robertson (2009), are collections of workers compensation classifications that have relatively similar ELFs over a broad range of limits. NCCI produces and files ELFs by state and hazard group for various per-occurrence loss limitations; NCCI also publishes the data and calculations used to produce the ELFs.

1.1 Overview of ELF Framework

Gillam (1991) details NCCI's general framework for computing excess ratios by hazard group and for individual states. The 2014 update makes changes to each component in the framework while keeping the general concepts. This section gives an overview of how those components fit together in the 2014 update while highlighting changes. Figure 1 in Section 1.3 provides a schematic representation of the components and relationships that are described in Section 1.

For a given per-claim dollar limit, the per-claim excess ratio for each state-hazard-group combination is a weighted average of excess ratios over five different *claim groups*. The claim groups used in the 2014 update are groupings based on reported injury types and other claim characteristics. Although these claim groups play a similar role as the claim types in Gillam (1991), the injury types that form each grouping are different.

A severity distribution is fit for each claim group, and state claim group excess ratios are scaled to reflect the average severity of claims for the claim group in that state and hazard group. The weights used for averaging the excess ratios across claim groups are the share of total losses in the state and hazard group for each claim group.

The excess ratios on a per-claim basis for each state-hazard-group dollar limit are then converted to a per-occurrence basis and loaded with the appropriate expenses to calculate the desired type of ELF.

This article discusses the following:

1. Data used for the update

---

[1] Strictly speaking, the term *excess loss factors* (ELFs) refers to the ratio of expected amount of losses in excess of a given limit to standard premium in tables published by NCCI. There are similar ELF-like ratios where the numerator may include or exclude allocated loss adjustment expense (ALAE) or where the denominator may be either the loss cost or the full standard premium. The term excess ratio refers to a ratio where the numerator is the expected amount of losses in excess of a given limit and the denominator is the expected unlimited losses. Appendix A.1 provides some definitions stated in mathematical terms.

2. Development of individual claims, and groups of claims, to ultimate
3. Derivation of countrywide and state ELF curves by claim group
4. Modeling of claim counts and average severities by state, hazard group, and claim group
5. Derivation of state ELF curves including ALAE
6. Adjustment of per-claim ELFs to per-occurrence ELFs
7. Implementation of the new methodology

1.2 Highlights of Changes from Previous Updates

This section contains a brief summary highlighting changes from previous updates.

Gillam and Couret (1997) and Corro and Engl (2006) have established the necessity of modeling ultimate losses that reflect loss development on an individual claim basis. They refer to this individual claim development as *dispersion*. To reflect dispersion, Gillam and Couret used different closed-form gamma distributed divisors to model the development of open and closed claims separately; Corro and Engl replaced each open claim by 173 values reflecting developed loss amounts and adjusted for reopened claims. The 2014 update adds refinements to this idea, including reflecting differences in development by size of claim.

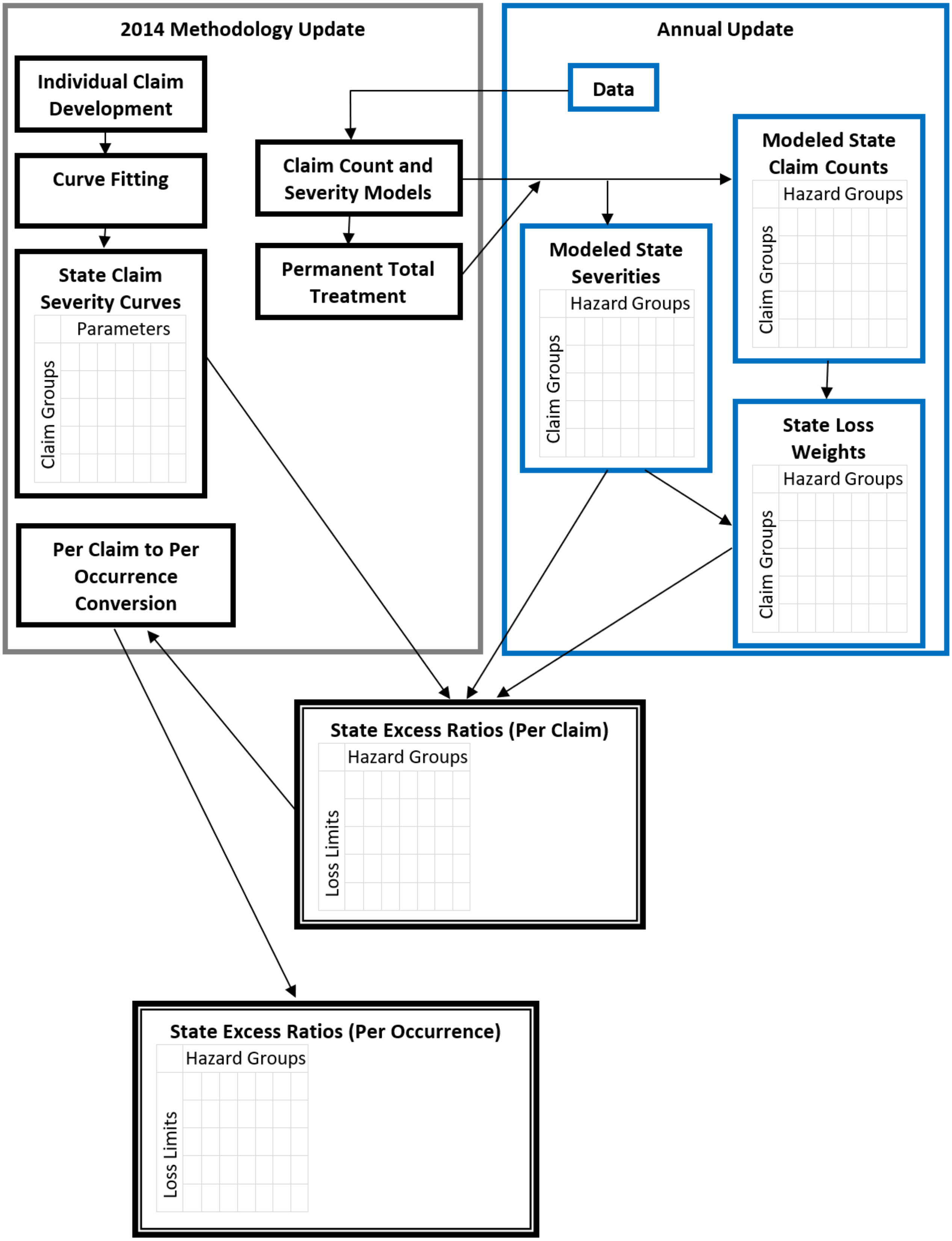
Another major theme is fitting distributions to each of the groupings of claims to the data. Gillam and Couret used, for each of their injury type groupings, maximum likelihood to fit the entire claim severity distribution. Corro and Engl used, for each injury type, the empirical distribution for the body of the distribution and fit a mixed exponential form to the tail, with parameters estimated by minimizing the sum of squared differences. The 2014 update, for each claim group, fits the body of the claim severity distribution to a mixed lognormal form. A Pareto form is fit to the tail, using Extreme Value Theory (EVT) as described by McNeil (1997).

Finally, in terms of reflecting state differences in claim severity distributions for each state and groupings of claims, Gillam and Couret used one countrywide claim severity distribution for all states for each of their injury type groupings. Corro and Engl created state-specific claim severity distributions for each injury type by credibility-weighting the claim severity distributions calculated using state-specific claims experience with the countrywide claim severity distributions. The 2014 update reflects state differences via an adjustment related to the coefficient of variation of the state's distribution for each claim group.

1.3 ELF Framework and the 2014 Update

Figure 1 shows the major components of the 2014 update that are described in Sections 1.1 and 1.2, as well as their relationships within the general framework of the annual update of ELF values.

Figure 1. Major components of the 2014 methodology and of the annual update of ELFs



2. Organization of Data

2.1 Overview

The 2014 update uses claim data on a case-incurred basis for 36 states reported under NCCI's ***Statistical Plan for Workers Compensation and Employers Liability Insurance*** ("unit data"). Compared to the 2004 update, the 2014 update takes advantage of:

- Additional data elements, namely the reported injured part of body, and open/closed claim status
- Claim data at maturities from 6th through 10th report; previously, only data through 5th report was available

These changes reduce the uncertainty in several loss development analyses, relative to previous updates.

Specifically, Table 1 gives a summary of the data used for the major analyses in the 2014 update:

Table 1. Data used in analysis

Analysis	Policy Years Spanning	Valuations (reports)
Curve Fitting	2000 to 2005[2]	6th through 10th
Claim Group Loss Development Factors[3]	2000 to 2009	1st through 10th
Development by Size of Claim[4]	2000 to 2005	4th through 10th
Per-Claim to Per-Occurrence Model	2000 to 2009	1st through 10th
Calculation of Severities and Loss Weights for Annual Update	Five policy periods underlying approved rate or loss cost filing	1st through 5th
Initial Values for Permanent Total Claims[5]	2000 to 2005[6]	6th through 10th

2.2 Claim Groups

In the 2014 update, NCCI categorizes each claim into one of the following five claim groups:

- Fatal
- Permanent Total
- Likely to Develop—Permanent Partial and Temporary Total ("Likely to Develop")

[2] Except for Florida, where only policy periods 2004 and 2005 were used, due to a major legislative reform in 2003.
[3] These factors were selected by NCCI ratemaking staff during their annual selection of loss development factors by state and injury type and were not modified for this analysis.
[4] For any claim open at any report between 4th and 10th report, all available link ratios (where the claim was open at the starting report) between 4th and 10th report were used. Claims that changed claim groups between reports were excluded from the regression analysis.
[5] This analysis also used exposure information from unit data corresponding to the claim data.
[6] Except for Florida, where only policy periods 2004 and 2005 were used, due to a major legislative reform in 2003.

- Not Likely to Develop—Permanent Partial and Temporary Total ("Not Likely to Develop")
- Medical Only

This assignment is based on the injury type[7] and, for the Likely to Develop and Not Likely to Develop groups, certain other claim characteristics. These characteristics are based on NCCI class ratemaking injury type groupings as described by Daley (2012). They consist of the injured part of body and open/closed claim status reported in unit data at the claim's 1st report and latest report.

The Likely to Develop claim group consists of the Permanent Partial and Temporary Total claims for which the injured part of body, together with the open/closed status, is one that indicates a greater likelihood of upward development in the claim value over time. The Not Likely to Develop claim group consists of the remaining Permanent Partial and Temporary Total claims.

The use of the Likely to Develop and Not Likely to Develop groupings reduces the effect of claims moving between the Permanent Partial and Temporary Total injury types. This helps improve the accuracy of loss development estimates. Claims in the Likely to Develop claim group tend to be more severe than those in the Not Likely to Develop claim group.

In contrast, the grouping of claims used by NCCI in the 2004 update were the injury types themselves:

- Fatal
- Permanent Total
- Permanent Partial
- Temporary Total
- Medical Only

The use of claim groups allows for changes in the distribution between them to be reflected automatically in the annual updates of ELFs.

3. Individual Claim Development

3.1 Overview: A Two-Step Approach

Where actuarial central estimates are of primary concern, it may suffice to determine the average development for a group of similar claims. However, the impact of development on individual claims is not uniform. Some open claims will have ultimate values much higher than originally estimated from

[7] The following is a brief description of the defining characteristic of each injury type:

- Fatal—death to an insured worker
- Permanent Total—not expected to ever be able to work again
- Permanent Partial—able to work after a recovery period, but with a permanent injury causing a reduction in function
- Temporary Total—able to work after a recovery period with no reduction in function
- Medical Only—able to work after receiving medical treatment and no wage replacement benefits

applying a uniform development factor, and some lower. Gillam and Couret (1997) incorporated this phenomenon, which they referred to as dispersion. Mahler (1998) shows how the work on dispersion by Gillam and Couret fit into a more general mathematical framework.

Dispersion increases the variance of the severity distribution and produces higher ELFs than would otherwise result from applying a uniform development factor to all individual claims.

Additionally, past NCCI studies, such as Evans (2011), showed that individual claim development varies by size of claim, specifically that smaller claims tend to have proportionally greater upward development than large claims. The 2014 ELF update is the first to reflect this finding.

In the 2014 update, NCCI uses a two-step approach to reflect individual claim development. The first step reflects development and dispersion through 10^{th} report. In this step, the average development applied to claims varies by size of claim, where the mean of the distribution of the development factor associated with a smaller open claim exceeds that of a larger open claim. Additional development and dispersion past 10^{th} report is applied in a second step that does not vary with the size of claim.

The 2014 use of claim information at 6^{th} through 10^{th} reports, not yet available during the 2004 update, incorporates data that goes further out to ultimate, thereby reducing the amount of extrapolation going into the final ELF. This improves the estimate of dispersion and, in turn, the projection of individual claim development to ultimate.

3.2 Step 1: Size of Claim and Loss Development Through 10^{th} Report

The reported incurred value of each claim in a state, claim group, and report combination is divided by the average claim size for the state, claim group, and report combination. The result is called an *entry ratio* because the mean across all claims in the cohort after this adjustment is 1. This normalization to entry ratios allows comparing claims across states and reports at a common basis as well as allows the pooling of claims from different states and reports to create a countrywide pool of claims.

For each state-claim group-report combination, NCCI constructs linear models that relate open claim development factors (LDFs) with claim size. These regression models have the logarithm of the LDF as the dependent variable and the log size of claim as the lone explanatory variable, as follows:

$$\ln(LDF) = Intercept + Coefficient \cdot \ln(Size\ of\ Claim\ in\ Entry\ Ratio) + \epsilon$$

A negative coefficient relates an increase in the size of claim with a decrease in the mean of the log LDF. As such, it relates changes in the size of claim to changes in the μ parameter of the lognormal LDF distribution for the claim. Similarly, the standard error of the regression estimates the standard deviation of the error and is used to estimate the σ parameter. We apply this regression by size of claim through 10^{th} report; beyond 10^{th} report, we assume no relationship between size of claim and individual claim development.

In the above regression model for the left-side dependent variable $\ln(LDF)$, we replaced the $\ln(Size\ of\ Claim\ in\ Entry\ Ratio)$ term on the right-hand side with the following transformation:

$$f(x) = \begin{cases} x - 1, & x < 1 \\ \ln x, & x \geq 1 \end{cases}$$

This transformation was found to produce a reasonable value for the claim size entry ratio explanatory variable. A straight log transformation worked well for entry ratios above unity, but it produced aberrant behavior for small entry ratios near zero, where the log transform approaches negative infinity. The use of the linear transform for entry ratios below unity was adopted to address this. More details on the parameters of the application of development by size of claim can be found in Appendices B.1–B.3.

3.3 Step 2: Loss Development and Dispersion Beyond 10^{th} Report

NCCI reflects dispersion beyond 10th report by treating each open claim at ultimate as a lognormal distribution. In contrast, each closed claim is treated as a point mass. Because claim information is not available past a 10th report in unit data, we determined claim closure rates past 10th report by selecting maximum additional durations of claims beyond 10th report by claim group. We also reviewed claims data from NCCI Financial Data Call 31 (Large Loss and Catastrophe) for information past a 10th report such as observed dispersion and claim closure.

More details on the parameters of the lognormal dispersion models can be found in Appendices B.1, B.2, and B.4.

3.4 Final Adjustments to the Combined Development

To account for claims reopening, the closed claim share of total losses is adjusted downward, and the open claim share of total losses is adjusted correspondingly upward, with the adjustments varying by claim group.[8]

As a final step, the total developed expected loss for open claims is balanced by state, report, and claim groups to an open-only LDF. This is calculated from the state's LDFs for both open and closed claims underlying NCCI rate and loss cost filings. These LDFs are adjusted to an open-only basis, using the empirical percentage of losses for claims that are open by state, report, and claim group.

4. Curve Fitting

4.1 Overview

For each claim group, NCCI pools the developed and dispersed claims data for all 36 available states to determine a countrywide claim severity distribution. This claim severity distribution has a mixture of two

[8] The adjustment for Fatal and Permanent Total claim groups is 0.5%; for Likely to Develop and Not Likely to Develop, 1.0%; and for Medical Only, 2.5%.

lognormal distributions for the body and a generalized Pareto distribution for the tail. Parameters for the mixed lognormal distribution are determined by best fit to the data for regions where there was enough data to be credible. Parameters for the generalized Pareto distribution and the splice points are selected according to Extreme Value Theory, specifically by choosing Hill estimator parameters using peak-over-threshold charts for the right-hand tail region where data was sparse.

To produce the state-specific distributions for each claim group, NCCI adjusts the lognormal parameters of the countrywide distribution using a statistic for approximating the coefficient of variation (CV) for each state relative to countrywide. We refer to this statistic as the *R-value* and use this R-value to adjust the parameters of the mixed lognormal curves; we do not adjust the tail of the Pareto distribution by state.

In contrast, the 2004 update, represented the body of the distribution using empirical excess ratio tables with the tail represented by a mixed exponential distribution, each determined by state and injury type. The advantages of the 2014 changes include the following:

- The countrywide distribution is much more resistant to outliers
- The body of the distribution has a more compact representation via a closed functional form
- There is a simple adjustment of the countrywide distribution to a state distribution to reflect changes in the shape of the state distributions
- The mixed lognormal distributions fit the body closely to the expected excess ratios resulting from developed and dispersed claim data
- The generalized Pareto distribution fits the tail more closely

4.2 Form of Body of Claim Severity Distribution

Excess ratios behave well with mixtures of distributions, in the sense that the excess ratio function for a mixture can be expressed as a weighted average of the excess ratio functions of the component distributions. Additionally, lognormal distributions are generally a reasonable choice to represent claim severity distributions and have closed-form expressions for excess ratios and related values that are reasonably easy to work with.

Since our dispersion method treats open claims at ultimate as a lognormal distribution and closed claims as point masses, the results of the dispersion calculation can be regarded as a mixture of the resulting lognormals and point masses. This naturally leads to representing the body of the excess ratio curve by using a mixture of lognormal distributions. Analysis of goodness-of-fit and related metrics showed that a mixture of two lognormal distributions provided sufficient accuracy to represent the body of the curve.

A nonlinear model routine is used to fit the excess ratio of a mixture of lognormal distributions to 5,000 excess ratio values determined from the development and dispersion model. The routine first uses maximum likelihood to fit a single lognormal and then uses those parameters to select starting values for fitting a mixture of two lognormal distributions through an iterative process designed to minimize the sum of squared differences.

4.3 Form of Tail of Claim Severity Distribution

The tail of the curve is represented as a generalized Pareto distribution, based on the peak over threshold (POT) method from Extreme Value Theory (EVT) as described by McNeil (1997). Pareto-tailed distributions are easy to work with when calculating the excess ratio. The point of transition between the main body and the Pareto tail is called the "splice point."

More details on the tail selection can be found in Appendices A.1 and A.2.

4.4 Splicing of Body and Tail of Countrywide Claim Severity Distribution

When combined with the splice point and shape parameter for the Pareto tail, the weights and parameters of the lognormal mixture provide a complete specification of the excess ratio curve as well as the corresponding claim severity distribution. This representation of the countrywide excess ratio curve for a claim group requires the following eight values:

- Two μ parameters, one for each of the two lognormal distributions
- Two σ parameters, one for each of the two lognormal distributions
- One weight parameter for the mixture of the lognormal distributions
- One splice point parameter
- Two parameters for the generalized Pareto distribution

Formulas for the excess ratio for the lognormal and Pareto distributions provide closed-form expressions that are well-behaved and easy to work with for determining excess ratios and related values.

More details on the countrywide curves can be found in Appendix C.1.

4.5 Relationship Between Countrywide and State Distributions

For a distribution of a given form, the CV is among the most useful statistics for determining the shape of the distribution. The CV is especially useful when working with excess ratios, because there is a closed form expression relating the area under the excess ratio curve with the CV of the claim severity distribution. This relationship motivates using the CV to adjust the countrywide curves for each claim group to a state level. We considered three possibilities related to the CV: the ordinary sample CV of untransformed losses, the sample CV of the logarithm of the losses, and the standard deviation of log losses. We chose to use the standard deviation of log losses because it produces the least bias and is most resistant to outliers. Additionally, our analysis suggested that it is appropriate to assign a credibility, based on the claim count volume, to the standard deviation of the log losses.

Note that the σ parameter of a lognormal claim severity distribution (the standard deviation of log losses) is related to the CV of the severity distribution:

$$CV^2 + 1 = e^{\sigma^2}$$

Each state-claim group combination is assigned a credibility based on claim count volume. The complement of credibility is the standard deviation of log losses for that claim group countrywide. The resulting credibility-adjusted state relativity factor is called the R-value. We apply this R-value to the parameters of each of the countrywide mixture of lognormal distributions to determine state-specific excess ratio curves by state and claim group.

More details on adjusting the countrywide curves to a state level can be found in Appendix C.2.

5. Estimation of State Severities and Loss Weights Using Bayesian Statistical Models

5.1 Overview

Once we have excess ratio curves by state and claim group that are normalized to a mean of 1,[9] we then require two more sets of values to calculate the excess ratio at a given loss dollar limit in each state and hazard group. The first is the average cost per claim (called *severity*) for each claim group, and the second is the percentage of total loss dollars in each claim group (called *loss weights*).

The severity for each claim group is used to convert the loss limit from a dollar basis to an entry ratio basis. The excess ratio curve for the corresponding claim group is used to find the excess ratio for the claim group at the loss limit. These excess ratios by claim group are weighted together using state loss weights by claim group to obtain the desired by-state excess ratio for the given loss limit.

This is essentially the same general procedure as described in Gillam (1991).

During the annual review, NCCI calculates updated severities and loss weights by state, hazard group, and claim group. These empirical values are based on five policy periods of unit data underlying the most recently approved rate or loss cost filings. In the 2014 update, NCCI updated the methodology to calculate these severities and loss weights. The enhanced methodology uses multilevel/hierarchical statistical models to increase year-to-year stability while maintaining responsiveness to state and hazard group differences.

We use one Bayesian hierarchical model to estimate claim counts and another to estimate severities. We combine the results of the two models to produce loss weights.

We make an additional stabilizing adjustment due to the large year-to-year fluctuations inherent in the emergence of Permanent Total (PT) claims. Initial severities and claim counts for PTs are estimated via the same Bayesian hierarchical models as for the other claim groups. However, in annual updates of ELFs, the PT severity is trended forward, and the share of PT claim counts relative to all lost-time claim

---

[9] As explained in Section 3.2.

counts is kept constant. In this way, movement in the frequency of the PT claim group is stabilized, changing in proportion to the state's lost-time claim frequency.

In contrast, in the 2004 update, the severities and loss weights were calculated based on developed, trended, and on-leveled data (the same data that are inputs to the Bayesian models). Because this allowed for significant year-to-year fluctuations in the ELFs, the severities and loss weights were then reviewed during the annual review for reasonableness and year-to-year fluctuations. Ad hoc adjustments were made as appropriate, usually to the Permanent Total severities and loss weights. The resulting ELFs calculated in one year were then averaged with the result of the ELFs calculated in previous years as an additional stabilizing adjustment.

Both the 2004 and 2014 updates use five policy periods of data to calculate severities and loss weights. However, due to the use of Bayesian models to estimate claim counts and severities, as well as the separate treatment of Permanent Totals, none of those ad hoc adjustments or weighted averages are still needed. The 2004 update stabilizing adjustments are not made to ELFs produced under the 2014 update.

5.2 Claim Count Model

The Bayesian hierarchical model to estimate claim counts specifies the following main effects and cross terms for claim frequencies for each state-policy period-claim group-hazard group combination:

- Claim group differences
- State differences
- Policy period differences
- Hazard group differences within each claim group
- Interactions between state and claim group differences
- Interactions between state and hazard group differences

The model performs the following major steps to get the expected claim count for each state-policy period-claim group-hazard group combination:

1. For each state and hazard group, start with the known exposure (payroll) by policy period
2. Adjusts the exposure to a common time period by multiplying by the corresponding policy period factor
3. Calculates an expected claim count for the given state-claim group-hazard group combination by multiplying together the main effects and cross terms

Details about the model, including model specification and additional structure for the parameters, can be found in Appendix D.1. Appendix Exhibit 1 provides an illustration for a sample state.

5.3 Severity Model

The Bayesian hierarchical model to estimate severity is similar to the claim count model; it specifies the following main effects and the cross term for claim severities for each state-claim group-hazard group combination:

- Base severity for claim group
- State differences
- Hazard group differences within each claim group
- Interactions between state and claim group differences

The model performs the following major steps to get the expected severity for each state-claim group-hazard group combination:

1. For each claim group, calculates a base severity over all states and hazard groups
2. Calculates the expected severity for the given state-claim group-hazard group combination by multiplying together the main effects and the cross term

Details about the model, including model specification and additional structure for the parameters, can be found in Appendix D.2. Appendix Exhibit 2 provides an illustration for a sample state.

5.4 Treatment of Permanent Total (PT) Claims

PT claims contribute to a significant portion of the ELFs, particularly at higher loss limits. PT claims also account for a comparatively small claim volume by state and hazard group and show high variability in loss amounts. This can combine to produce large year-to-year fluctuations in PT severities and loss weights. This, in turn, can result in large year-to-year fluctuations in the ELFs. In the past, NCCI averaged each year's indicated ELFs with the ELFs of prior years, as well as examined the severities and loss weights for all states, hazard groups, and injury types. Judgment was required where fluctuations in severities and loss weights led to large fluctuations in ELFs.

In the 2014 update, we increase the stability of severities and loss weights via the new Bayesian models. We calculate the claim counts and severities for PT claims separately from the other claim groups. For PTs, instead of the years of data used in the severity and claim count models for non-PT claim groups, we use the five years of data used to fit the excess ratio curves.

We generate the PT severities and claim counts based on two initial fixed values for each state and hazard group:

- An initial PT severity
- The initial ratio of PT claim counts to non-PT lost-time claim counts

To obtain the PT expected severity, we apply a two-stage trend to the initial PT severity.

In the 2014 update, initially the first stage used annual trend factors of 5.0% for indemnity and 6.7% for medical. These trend factors were the average annual changes from Accident Years 2002 to 2008 from NCCI's 2012 Countrywide Frequency and Severity Analysis. The period 2002 to 2008 was selected to avoid different frequency trends that occurred for different claim sizes before 2002 and the effects of the Great Recession on severities after 2008.

However, the observed indemnity and medical trends in subsequent years have decreased since the original choice of trend factors. For the annual review filed in 2016, NCCI has started to use a different first-stage trend. The first-stage trend is now a blend of six years. The original 5.0% indemnity and 6.7% medical trend are blended with newly selected trends of 2.0% indemnity and 3.0% medical. The newly selected trends will receive an additional year's worth of weight in each subsequent annual review.

The second stage uses state-specific trend factors for indemnity and medical separately, from the most recent state loss cost or rate filing.

To update the PT expected claim count each year, we multiply the sum of the claim counts for the Fatal, Likely, and Not-Likely claim groups by the initial ratio of PT claim counts to non-PT lost-time claim counts. This assumes that the ratio of PT claims to non-PT lost-time claims stays constant over time.

To obtain the PT loss weight, we combine the PT severities and claim counts.

Appendix Exhibits 3 and 4 in Appendix E provide illustrations for a sample state.

6. Treatment of Losses Including ALAE

6.1 Overview

In the 2004 update, the excess ratio curves were used for loss excluding ALAE and loss including ALAE were the same. To reflect the inclusion of ALAE, An ALAE factor was applied to only the most severe injury types (Fatal, Permanent Total, and Permanent Partial) when calculating severities on a basis including ALAE.

This changed with the 2014 update. NCCI reviewed paid ALAE in unit data and found that ALAE relative to loss is a smaller proportion of dollars for larger claims in the more "severe" claim groups. We reflect this by generating separate excess ratio curves on a loss including ALAE basis. We also apply ALAE factors that vary by claim groups when calculating severities that include ALAE. This generates a separate set of excess ratio curves shaped differently than the loss-only curves for the same claim group.

6.2 Excess Ratio Curves Including ALAE

In the 2014 update, NCCI generates countrywide excess ratio curves on a loss including ALAE basis. This is done by first adding an ALAE amount to each claim. For closed claims, we use the reported paid ALAE. For open claims, we analyzed development patterns in the ratio of paid ALAE to paid loss by claim group and size of claim. Based on that analysis, we adjust the paid ALAE to reflect differences by size of claim and claim group. For each open claim, we adjust the paid ALAE by multiplying it by a development factor that differs by claim group and claim size ranges. As a final step, we balance the total ALAE dollars to a target ALAE percentage by state and period.

Then, for each claim group, we developed, dispersed, and fit to curves the individual expected excess loss and ALAE claim amounts. This was done by the same procedure as described previously for losses not including ALAE. The same generalized Pareto tail distributions by claim group are spliced to the tail, as was done for losses not including ALAE. State curves are generated by adjusting these countrywide curves including ALAE using the R-value calculated from losses including ALAE, as was done for losses not including ALAE.

Consistent with the determination of excess ratios for losses without ALAE, the final state curves are generated by weighting those state curves with excess ratio curves of losses not including ALAE; the weights are related to how the state ALAE factor (updated during each state's annual review) compares to the overall countrywide ALAE factor of 1.127 used in the original fitting of the countrywide excess ratio curves on a loss-including-ALAE basis.

For a given dollar loss limit, the excess ratio for losses including ALAE can be greater than or less than the excess ratio for losses not including ALAE. For a given dollar limit, even when the excess ratio for losses including ALAE is less than the excess ratio for losses not including ALAE, the excess dollar amount of losses including ALAE is at least as great as the excess dollar amount of losses not including ALAE. Including ALAE cannot reduce the dollars in excess of any limit.

A further refinement is done to excess ratios including ALAE for consistency with those not including ALAE. Upper and lower bounds are applied to excess ratios including ALAE; these bounds are related to the excess ratios not including ALAE. The upper bound represents the case where the additional ALAE has full contribution to the excess on an including ALAE basis (i.e., an additional dollar of ALAE contributes to the excess including ALAE by one dollar); the lower bound represents the case where additional ALAE has no contribution to the excess on an including-ALAE basis (i.e., an additional dollar of ALAE does not contribute to the excess including ALAE).

The parameters for the excess ratio curves on a loss-including-ALAE basis are fixed (as are curves for loss excluding ALAE). However, these adjustments for the state ALAE factor, which may change during the annual review, as well as the upper and lower bounds, contribute to annual changes in ELFs for losses including ALAE.

More details on the specific excess ratio formulas for losses including ALAE can be found in Appendix F.

6.3 Severities Including ALAE

In the 2014 update, NCCI estimated countrywide ALAE percentages by claim group. These percentages are converted to relativities to a total countrywide ALAE percentage. Then, in the annual review, the percentages are applied to an overall state ALAE percentage. This yields state ALAE percentages by claim group. The state ALAE percentages by claim group are then applied to pure loss severities to obtain severities including ALAE. Table 2 shows the countrywide ALAE percentages by claim group. Appendix Exhibit 5 in Appendix G provides an illustration for a sample state.

Table 2. Countrywide ALAE percentages by claim group

Claim Group	ALAE Percentage (%)
Fatal	5.90
Permanent Total	7.82
Likely to Develop—Permanent Partial and Temporary Total	11.88
Not Likely to Develop—Permanent Partial and Temporary Total	11.32
Medical Only	13.20
Total	10.67

This represents another refinement to the methodology. In the 2004 update, an ALAE factor was calculated for the Fatal, Permanent Total, and Permanent Partial injury types so that the ALAE dollars assigned to those injury types balanced to the state total ALAE dollars. The Temporary Total and Medical Only injury types were not allocated any ALAE dollars. Note the use of injury types as opposed to claim groups when adjusting severities to include ALAE. The 2004 algorithm resulted in a higher ALAE percentage for more serious injury types, especially for fatal and permanent total, and this result was not supported by subsequent empirical investigation.

7. Per-Claim to Per-Occurrence Adjustment

7.1 Overview

Occurrences are claims from the same policy that arise from a single event. In the context of ELFs in general, dollar limits can apply either on a per-occurrence basis or on a per-claim basis. The countrywide excess ratio curve is based on a model of per-claim excess ratios at ultimate, while filed state ELF values are based on per-occurrence excess ratios at ultimate. Therefore, we need some way to convert excess ratios from a per-claim basis to a per-occurrence basis.

For converting per-claim excess ratios to per-occurrence excess ratios, NCCI constructed a table for use in all states that relates per-claim and per-occurrence excess ratios. We compared claim characteristics between claims that were reported as part of a multiclaim occurrence to those that were not. We also estimated the probability of a claim belonging to a multiclaim occurrence. That estimate was based on a comparison of the likelihood of injuries on the same policy having the same date of injury to what was observed in the actual data.

In contrast, in the 2004 update, a collective risk model approach was used, which produced much smaller differences between per-occurrence and per-claim excess ratios.

7.2 Claims from Multiclaim Occurrences vs. Single-Claim Occurrences

In the 2014 update, we considered using a collective risk model to aggregate individual claims into occurrences, as was done in the 2004 update. We rejected that approach because we observed positive correlation in claim size between claims within an occurrence (correlation coefficient of about 0.25), which violates the independence assumptions of the collective risk model.

Instead, we categorized occurrences as *singletons* and *multiples*, depending upon whether more than one claim arose from the occurrence. The catastrophe code reported in unit data identifies whether an individual claim on a given policy belongs to a multiple claim occurrence. For our purposes, the quality of catastrophe code data is suitable for identifying a subset of multiples but not the entire subset. We assumed that this subset of multiples identified from reported data is representative of all true multiples, and we looked at claim characteristics of this subset to draw conclusions about all multiples.

Such conclusions include the proportion of occurrences with exactly two claims, three claims, etc., and differences in claim characteristics between singletons and multiples. For example, multiples have a higher mean severity, have a higher proportion of fatalities, and are more likely to be caused by an auto accident. Additionally, we found claim size of claims within a multiclaim occurrence to have a positive correlation of 0.25.

More details on the claim characteristics for multiples can be found in Appendix H.

7.3 Probability of a Claim Belonging to a Multiclaim Occurrence

Because the subset of identified multiples did not include all multiples, we used an indirect simulation approach to measure the proportion of all claims that belonged to a multiple occurrence. We looked at claims data for the decade 2000–2009 but excluded claims whose injuries were reported to occur on a Monday or on the weekend. This exclusion avoids effects from possible clustering of claims reporting on Mondays for claims occurring on the weekend. For each claim, we calculated the number of days from policy effective date to the date of injury, excluding weekends and Mondays, and called this the *time index*. For any pair of distinct claims, we considered two events:

A. They are from the same policy
B. They have the same time index

If there were no multiple occurrences, then these events should be independent of one another. The data revealed a positive correlation. We used a simulation routine that grouped claims on the same policy into occurrences. The simulation was run, varying the assumed probability that a claim is the first claim within an occurrence. As that probability increased, so did the correlation between events A and B. The runs for which the correlation was closest to the observed correlation indicated that about 2% of claims belonged in a multiple occurrence and that the average number of claims in a multiple occurrence is 2.71.

We used this to express the countrywide excess ratio curve on a per-occurrence basis in terms of two excess ratio curves, both on a per-claim basis—one curve that reflects the frequency and severity by claim group for claims from singleton occurrences and another curve that reflects the characteristics of claims from multiple occurrences.

More details on the comparison of singleton and multiple occurrence claims and the excess ratio formulas can be found in Appendix H.

8. Implementation

The updated ELF values[10] and the underlying methodology were filed in NCCI Item Filing R-1408 on June 16, 2014, and have been approved since December 3, 2014,[11] for effective dates in 2014–2015.

ELF values are also used in NCCI ratemaking; the results of the updated values and methodology were incorporated into the loss cost and rate filings with effective dates in 2015–2016.

Additionally, as the 2014 ELF update provides methodology on determining the claim severity distribution, results from the update are being used in the update to the NCCI table of insurance charges ("Table M") currently in progress.

8.1 Impact

The impacts from implementing the methodology in the 2014 update vary by state, occurrence limit, and hazard group. Because the state excess ratio curves in the 2014 update are obtained by adjusting the countrywide excess ratio curves, there are some patterns that generally hold across states. NCCI compared the ELFs calculated in the previous methodology with those calculated with the methodology in the 2014 update and found the following:

- At loss limits below $3 million, most excess ratios from the 2014 update are higher than in the previous update
- At loss limits above $3 million, the excess ratios from the 2014 update are increasingly lower than in the previous update

9. Conclusion

The 2014 update makes significant theoretical and practical improvements in the methodology while keeping much of the existing framework and fundamental ideas. The improvements result in a more accurate treatment of losses including ALAE, tails of the claim severity distributions that are more accurate with additional theoretical grounding, and increased year-to-year stability in the ELFs.

---

[10] Specifically, the Excess Loss Pure Premium Factors (ELPPFs) and Excess Loss and Allocated Expense Pure Premium Factors (ELAEPPFs).

[11] Item R-1408 was filed and approved in all NCCI states except Arizona, Florida, Idaho, Iowa, and Virginia, where they were filed and approved in each state's loss cost/rate filing.

10. Acknowledgments

Many staff at NCCI contributed to the 2014 ELF update, including John Robertson, Jon Evans, Chris Laws, and Casey Tozzi. We also thank the NCCI Individual Risk Rating Working Group for their discussion and input.

Appendix

Appendix A.1—Mathematics of Excess Ratios

Given a claim severity distribution with the CDF $F(x)$, we define the following standard terms:

- Mean $\equiv \mu_F = \int_0^\infty xf(x)dx$
- Variance $\equiv \sigma_F^2 = \int_0^\infty (x-\mu_F)^2 f(x)dx$
- Standard Deviation $\equiv \sigma_F = \sqrt{\sigma_F^2}$
- Survival Function $\equiv S_F(x) = 1 - F(x) = \int_x^\infty f(y)dy$
- Excess Ratio Function $\equiv R_F(x) = \frac{\int_x^\infty (y-x)f(y)dy}{\int_0^\infty yf(y)dy} = \frac{\int_x^\infty S_F(y)dy}{\mu_F}$
- Mean Residual Lifetime $\equiv MRL_F(x) = \frac{\int_x^\infty (y-x)f(y)dy}{\int_x^\infty f(y)dy} = \frac{\mu_F R_F(x)}{S_F(x)}$.

When $\mu_F = 1$, we have the equation:

$$\int_0^\infty R_F(x)dx = \frac{(1+{CV_F}^2)}{2} \quad \text{where} \quad CV_F = \text{Coefficient of Variation} = \frac{\sigma_F}{\mu_F}.$$

We use Extreme Value Theory (EVT) to select a generalized Pareto distribution tail (select the splice point a and parameters m and b—see the following section). This was done for each claim group via the following steps:

- Normalize data to entry ratios; pool states and years
- Chart the Mean Residual Lifetime of losses and log of losses (Peak Over Threshold [POT] charts—see Embrechts et al. [2003], pp. 352–370)
- Use POT chart of log-losses to determine the slope parameter m (Hill estimator) and the splice point a (see Embrechts et al. [2003], pp. 330–335)
- One tail for use in all states

Appendix A.2—Useful Formulas

Lognormal

In what follows, $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$ denotes the CDF of the standard normal distribution. Therefore, the CDF of the lognormal distribution with parameters μ and σ is given by:

$$F(r) = \Phi(\mathrm{z}) \text{ where } z = \frac{lnr-\mu}{\sigma}$$

its mean by:

$$\bar{r} = e^{\mu + \frac{\sigma^2}{2}}$$

and its excess ratio function by:

$$R_F(r) = 1 - \Phi(\mathrm{z} - \sigma) - \mathrm{r}\frac{1 - F(r)}{\bar{r}}$$

and its MRL function by:

$$MRL_F(r) = \frac{\bar{r}\big(1 - \Phi(\mathrm{z} - \sigma)\big)}{1 - \Phi(\mathrm{z})} - r\,.$$

Generalized Pareto Distribution (GPD)

For the parameterization used here, which is not standard,[12] the CDF of the GPD is given by:

$$G(b, m; x) = \begin{cases} 1 - \left(\dfrac{b}{mx + b}\right)^{\frac{m+1}{m}} & m \neq 0 \\ 1 - e^{-\frac{x}{b}} & m = 0 \end{cases}.$$

Its mean is just the b parameter. Its excess ratio function is:

$$R_G(x) = \begin{cases} \left(\dfrac{b}{mx + b}\right)^{\frac{1}{m}} & m \neq 0 \\ e^{-\frac{x}{b}} & m = 0 \end{cases}$$

and its mean residual lifetime function is linear:

$$MRL_G(x) = mx + b.$$

The case $m > 0$ is the usual Pareto distribution.

[12] One will often see the CDF of a GPD expressed as $F(x) = 1\text{-}(1+\xi(x/\beta))^{-1/\xi}$ where $\xi = m/(m+1) \neq 0$ has the same sign as m and is sometimes called the *tail index* or *shape parameter* and $\beta = \mu/(m+1)$ the *scale parameter.*

Splicing

Fix a "splice point" a and suppose we have a claim severity distribution with CDF F that we want to modify to a new distribution with CDF $\tilde{F}$ so that for losses greater than a, the new distribution $\tilde{F}$ follows a second distribution with CDF G. The case of interest is when F is a mixture of lognormal distributions and G is a GPD. It is natural to specify this "spliced distribution" in terms of its survival function:

$$S_{\tilde{F}}(x) = \begin{cases} S_F(x) & x \le a \\ S_F(a) S_G(x-a) & x \ge a. \end{cases}$$

Now suppose we have the equation:

$$MRL_F(a) = \mu_G.$$

Then $\mu_{\tilde{F}} = \mu_F$ and one can readily determine the excess ratio function of the spliced distribution $\tilde{F}$ from those of F and G:

$$R_{\tilde{F}}(x) = \begin{cases} R_F(x) & x \le a \\ R_F(a) R_G(x-a) & x \ge a. \end{cases}$$

When G is GPD with parameters m and $b = MRL_F(a)$ we have:

$$R_{\tilde{F}}(x) = \begin{cases} R_F(x) & x \le a \\ R_F(a) \left(\dfrac{b}{m(x-a)+b} \right)^{\frac{1}{m}} & x \ge a. \end{cases}$$

Appendix B.1—Parameters of the Lognormal LDF Dispersion Models: Overview

The kernel density distribution that the individual claim development and dispersion (D&D) model initially assigns to the open claim of size x (reopened claims are treated the same as open) and report t is a lognormal distribution with standard parameters:

$$\ln(x) + \mu(t; x) \text{ and } \sigma^2(t)$$

- Ratemaking LDFs by state, claim grouping, and report are converted to "open-only" LDF factors
- Those open-only factors are used to modify the first parameter by an additive constant by state, claim grouping, and report (flat factor in entry ratio space)
- That adjustment assures that the model has the same expected loss at ultimate as implied by the ratemaking LDFs, by state, claim grouping, and report

Appendix B.2—Parameters of the Lognormal LDF Dispersion Models: Details

As noted in the paper, the development and dispersion (D&D) model adjusts open claims to an ultimate basis in two steps—first to a 10th report and then from a 10th report to ultimate:

- Step 1 (through 10th report)—D&D varies by size of claim based on regression
  - Linear regression considers individual claim development from report t to report 10 and relates it with the claim amount at report t
  - A linear regression model is determined:
    - ❑ For claims open at each of four reports t, for $t = 6,7,8,9$
    - ❑ For each of the five claim groupings
    - ❑ 20 models in total
- Step 2 (beyond 10th report)—D&D does not vary by size of claim[13]
  - Consider the pattern of variability of observed logarithms of annual LDFs from reports t to $t + 1$, for $t = 4, 5, 6, 7, 8, 9$
  - Use exponential regression, tempered with judgment, and project those patterns from report $t > 9$ to claim closure

The D&D steps 1 and 2 are not correlated. Sizes of loss and per-claim development beyond 10th report are not related.

---

[13] We reviewed claims data from NCCI Financial Data Call 31 (Large Loss and Catastrophe) and found that it did not support a relationship between per-claim development beyond 10th report and size of claim.

Appendix B.3—Parameters of the Lognormal LDF Dispersion Models: Step 1 Further Details

For claims open at report t and within a claim grouping, a linear regression estimates the log per-claim LDF as a function of a transformed per-claim loss amount at report t

- The regression equation is used to estimate the mean $\mu_1(t; x)$ of the t^{th} to 10th log LDF; this estimate varies with the size of claim x
  - Transforming the size of claim metric improves the fit, especially for the largest claims
  - In place of entry ratio x as explanatory variable, the revised model uses a compressed size of claim metric $\gamma(x)$:

    $\gamma(x) = \ln x$ for $x \geq 1$; $\quad \gamma(x) = x - 1$ for $x \leq 1$
- The variance of the distribution of the residual gives an estimate of the variance $\sigma_1^2(t)$ of the t^{th} to 10th log LDF; this estimate does not vary with the size of claim x
- The proportion $\rho(t)$ of claims open at report t that remain open at 10th is calculated from the claim data used in the regression

Appendix B.4—Parameters of the Lognormal LDF Dispersion Models: Step 2 Further Details

The model assumes that beyond 10th report:

- The mean of the distribution of log annual LDFs is 0 for all claims and all years
- The annual claim closure rate is constant for each claim grouping

The variance σ_2^2 of the second step is estimated for each claim grouping using:

- A constant annual claim closure rate
- Projected variances of the distributions of annual log per-claim LDFs
- Exponential decay model is appropriate (becoming linear with a log vertical scale)
- A judgmentally assigned asymptote as the long-term estimate of the variance of the log annual LDF
- Formulas for the decay model
- Let $y(t) =$ empirical variance of annual per-claim LDF from report t to $t + 1$

Formula to project variance is:

$$y(t) = a + c \cdot e^{bt}$$

in which a is an assumed asymptotic long-term variance[14] and b and c are constants to be estimated. The linear regression model with coefficient vector $= \beta$:

$$\ln(y(t) - a) = \ln c + bt = \beta_0 + \beta_1 t + \epsilon(t)$$

is used to yield the estimates $c = e^{\beta_0}$ and $b = \beta_1$. The formula for the variance of 10th-ultimate log LDF is as follows, letting $1 - s$ be the constant annual claim closure rate and N be the maximum duration to closure after report $t = 10$:

$$\sigma_2^2 \approx \left(\frac{a}{1-s}\right)\left(1 - \frac{Ns^N(1-s)}{1-s^N}\right) + \left(\frac{ce^{10b}}{1-e^b}\right)\left(1 - \left(\frac{e^b(1-s)}{1-s^N}\right)\left(\frac{1-\left(se^b\right)^N}{1-se^b}\right)\right).$$

The maximum additional duration to closure (N) after 10 years by claim group is given in Table 3:

Table 3. Maximum additional duration to close after 10 years

Claim Group	*N* (years)
Fatal	25
Permanent Total	30
Likely to Develop PP/TT	20
Not Likely to Develop PP/TT	15
Medical Only	10

In summary, to each open claim of size x at latest report t, the D&D model assigns a mean $\mu(t; x)$ and variance $\sigma^2(t)$ to the log per-claim LDF distribution; with just two uncorrelated component steps: $\mu(t; x) = \mu_1(t; x)$ and $\sigma^2(t) = \sigma_1^2(t) + \rho(t)\sigma_2^2$

Where:

- $\mu_1(t; x) =$ linear estimate for the mean of the t^{th} to 10th log LDF
- $\sigma_1^2(t) =$ estimated variance of the t^{th} to 10th log LDF
- $\rho(t) =$ proportion of claims still open at 10th report
- $\sigma_2^2 =$ estimated variance of the 10th to ultimate log LDF

The values $\mu(t; x)$ and $\sigma_2(t)$ are the standard parameters for the lognormal density model for the LDFs of an open claim of size x at latest report t (prior to balancing with ratemaking LDFs by state, claim group, and report).

---

[14] A value of 0.015 was selected for a by reviewing PT claims data from unit data and Call 31.

Appendix C.1—Parameters of Countrywide Excess Ratio Curve as a Lognormal Mixture and Pareto Tail

This section describes the ingredients that go into the parametric form for expressing the excess ratio $R(r)$ as a function of entry ratio r. The severity distribution is a mixture of two lognormal distributions with parameters μ_1, μ_2 and σ_1, σ_2, respectively. The weight assigned the first component is denoted ω_1 and the weight assigned the second component is denoted $\omega_2 = (1 - \omega_1)$. A Pareto tail distribution is spliced onto the lognormal mixture at an entry ratio denoted by a and termed the splice point (the splicing preserves an entry ratio mean of 1). The parameters used for the Pareto are denoted b and m, chosen to exploit the characterization of the Pareto distribution as one having a linear mean residual lifetime function of slope m and intercept b, which is also the mean of the distribution.

The CDF of the lognormal components are:

$$F_i(r) = \Phi(z_i) \text{ where } z_i = \frac{\ln r - \mu_i}{\sigma_i}, \text{ where } i = 1, 2.$$

Their means are:

$$\overline{r_i} = e^{\mu_i + \frac{\sigma_i^2}{2}}$$

And in particular:

$$1 = \omega_1 \overline{r_i} + (1 - \omega_1)\overline{r_2}$$

as we are working with entry ratios. The excess ratio functions of the lognormal components are:

$$R_i(r) = 1 - \Phi(z_i - \sigma_i) - \mathrm{r}\frac{1 - F_i(r)}{\overline{r_i}}.$$

The CDF of the lognormal mixture portion is:

$$F(r) = \omega_1 F_1(r) + (1 - \omega_1)F_2(r), r \leq a$$

And the excess ratio function for the lognormal mixture portion is loss weighted average:

$$R(r) = \omega_1 \overline{r_1} R_1(r) + (1 - \omega_1 \overline{r_1})R_2(r), r \leq a.$$

The probability of surviving to the splice point is:

$$S = 1 - F(a)$$

And since the mean residual lifetime at the splice point must preserve a mean of 1, we have the following equation that can be used to find the value b to assign to the b parameter:

$$R(a) = bS.$$

The CDF of the Pareto tail portion is:

$$F(r) = 1 - S\left(\frac{b}{m(r-a)+b}\right)^{\frac{m+1}{m}}, r \geq a.$$

Finally, from the formula for the excess ratio of a Pareto distribution, we have the formula for the Pareto tail portion as:

$$R(r) = S\left(\frac{b}{m(r-a)+b}\right)^{\frac{1}{m}}, r \geq a.$$

Appendix C.2—Deriving the State Excess Ratio Curves From the Countrywide Curves

The process starts with a credibility weighted relativity r of the standard deviation of state log losses to that of the countrywide. That is,

$$r = Z\left(\frac{\sigma \text{ for logged losses for claim group in state}}{\sigma \text{ for logged losses for claim group countrywide}}\right) + (1 - Z).$$

Here the credibility weight is determined as:

$$Z = \frac{N}{N+k}$$

where k varies by claim group, as shown in Table 4, and N is the expected number of such claims in the claim group for the state.

Table 4. Credibility *k* value by claim group

Claim Group	*k*
Fatal	60
Permanent Total	33
Likely to Develop PP/TT	73
Not Likely to Develop PP/TT	129
Medical Only	373

For the state curve, we use the same w_i as for the countrywide. We replace each of the μ_i and σ_i with $r\mu_i$ and $r\sigma_i$, respectively. This multiplies the standard deviation of the log losses by a factor of r. We then replace the new μ_i with $\mu_i + c$, where c is the constant that produces a mean of 1 for the state curve. Note that adding this constant does not change the standard deviation of the logarithmic entry ratio of the state curve. This yields the lognormal mixture for the state curve.

Now we determine the parameters for the Pareto. We keep the splice point a and the "slope" m the same as for the countrywide, and determine the b parameter for the state curve. The parameter value b

can be determined by matching the Mean Residual Lifetimes of the lognormal mixture and the GPD tail. To this end, the Mean Residual Lifetime of a lognormal component at a is:

$$MRL_i(a) = \frac{e^{\left(\mu_i + \frac{\sigma_i^2}{2}\right)}\left(1 - \Phi\left(\frac{\ln(a) - \mu_i}{\sigma_i} - \sigma_i\right)\right)}{1 - \Phi\left(\frac{\ln(a) - \mu_i}{\sigma_i}\right)} - a\,.$$

The Mean Residual Lifetime of the lognormal mixture at splice point a is, therefore, the sum weighted by the frequency of claims surviving to the splice point:

$$MRL(a) = \frac{\sum_{i=1}^{2} w_i\left(1 - \Phi\left(\frac{\ln(a) - \mu_i}{\sigma_i}\right)\right) MRL_i(a)}{\sum_{i=1}^{2} w_i\left(1 - \Phi\left(\frac{\ln(a) - \mu_i}{\sigma_i}\right)\right)}.$$

Since the overall mean of the GPD is the b parameter, setting $b = MRL\ (a)$ will produce a distribution function $F(r)$, defined as in the previous section, whose Mean Residual Lifetime at splice point a is also b and whose overall mean is therefore also 1. This last b is the state-specific b parameter. This now gives the complete state curve.

We compared the empirical claim experience before and after applying development and dispersion and found that the state relativity of the standard deviation of the logarithmic loss to countrywide was very similar. Accordingly, interim updates to the relativities are based on empirical experience.

Appendix D.1—Claim Count Model

Claim counts are assumed to follow a negative binomial distribution.

The model can be written as:

$$\log(\mu_{ghr}) = \delta_{shr} + \gamma_g + \xi_s + \eta_{hg} + \psi_{sg} + \omega_{sh} + \rho_r$$

Table 5 shows a description of the terms in the model.

Table 5. Description of terms in claim count model

Term	Description
$\boldsymbol{\mu_{ghr}}$	expected number of claims in claim group g, state s, hazard group h, and policy period r
$\boldsymbol{\delta_{shr}}$	log of payroll in state s, hazard group h, and policy period r
$\boldsymbol{\gamma_g}$	factor for claim group g
$\boldsymbol{\xi_s}$	factor for state s
$\boldsymbol{\eta_{hg}}$	factor for hazard group h specific to claim group g
$\boldsymbol{\psi_{sg}}$	factor for interaction between state s and claim group g
$\boldsymbol{\omega_{sh}}$	factor for interaction between state s and hazard group h
$\boldsymbol{\rho_r}$	factor for policy period r

The parameters ψ_{sg} and ω_{sh} are credibility weights (or "shrunk") using multilevel modeling. This is sometimes referred to as *partial pooling*.

Additionally, η_{hg} is assumed to have a structure described later in this section.

The following are notes on each parameter in the equation:

- The log of the payroll (δ_{shr}) is known from data and serves as the exposure base
- The policy period factor (ρ_r) is fixed for all other factors and serves to account for differences between policy periods, such as benefit levels and trend in frequency per payroll
- The state-specific parameters ($\xi_s, \psi_{sg}, \omega_{sh}$) account for the overall state variation separately from the state-to-state variation of individual claim groups and hazard groups, especially considering credibility
    - The state factor (ξ_s) is estimated very accurately, since every claim for all the modeled claim groups for a state contributes to its estimation. Using this factor to account for state differences allows for state variation that interacts with either claim groups or hazard groups to be estimated more accurately. This is analogous to reducing the variance between groups and has the effect of shrinking the state-claim group (ψ_{sg}) and state-hazard group (ω_{sh}) factors toward 1.0.
- The claim group factor (γ_g) accounts for the base frequency per payroll in each claim group

- The hazard group factor (η_{hg}) differs by claim group and has the following structure:

$$\eta_{hg} = \begin{cases} \hat{\eta}_{h1} & if\ g\ is\ Fatal \\ \hat{\eta}_{h1} \cdot \alpha_1 & if\ g\ is\ Permanent\ Total \\ \hat{\eta}_{h2} & if\ g\ is\ Not\ Likely \\ \hat{\eta}_{h2} \cdot \alpha_2 & if\ g\ is\ Likely \end{cases}$$

  This structure reflects the results of an inspection of empirical hazard group relativities by claim group relativities, where we found that while they were more extreme for certain claim groups, the different claim groups varied similarly across hazard groups. For this purpose, we chose to group Fatal and Permanent Total together, and Likely and Not Likely together.

Exhibit 1 shows sample values for the calculation of the claim counts for an NCCI state. Note that Medical-Only claim counts are determined directly using reported data, and Permanent Total claim counts are calculated via a separate procedure.

Appendix Exhibit 1. Sample calculation of expected claim counts by claim group and hazard group for an NCCI state

(1) Payroll in $ millions ($e^{\delta_{shr}}$)

	Hazard Group						
Policy Period	A	B	C	D	E	F	G
5/1/11–4/30/12	953	3,388	15,369	3,376	6,203	1,978	597
5/1/10–4/30/11	944	3,290	15,104	3,293	5,940	1,854	627
5/1/09–4/30/10	921	3,245	14,332	3,142	5,695	1,862	593
5/1/08–4/30/09	921	3,217	14,080	3,072	5,450	1,813	570
5/1/07–4/30/08	911	3,288	14,546	3,119	5,678	1,885	568

(2) Factors for Policy Period (ρ_r)

Policy Period	Factors for Policy Period (ρ_r)
5/1/11–4/30/12	1.000
5/1/10–4/30/11	1.051
5/1/09–4/30/10	1.089
5/1/08–4/30/09	1.102
5/1/07–4/30/08	1.205

(3) Adjusted Payroll = (1) x (2)

	Hazard Group						
	A	B	C	D	E	F	G
Adjusted Payroll ($ millions)	5,060	17,885	79,893	17,402	31,495	10,224	3,214

(4) State Factor (ξ_s)

State Factor (ξ_s)
0.900

(5) Factor for Interaction Between State and Hazard Group (ω_{sh})

	Hazard Group						
	A	B	C	D	E	F	G
Relativity	1.293	1.139	1.112	0.984	0.961	0.820	0.787

(6) Claim Group Frequency (γ_g)

Claim Group	Claims per $ Million Payroll (γ_g)
Fatal	0.00032
Likely PP/TT	0.05770
Not Likely PP/TT	0.29446

(7) Factor for Hazard Group Specific to Claim Group (η_{hg})

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	1.000	0.984	0.769	2.459	3.224	10.672	16.531
Likely PP/TT	1.000	0.741	0.371	0.699	0.628	1.390	1.313
Not Likely PP/TT	1.000	0.743	0.374	0.701	0.631	1.386	1.310

(8) Factor for Interaction Between State and Claim Group (ψ_{sg})

Claim Group	Claim Group Factor (ψ_g)
Fatal	1.261
Likely PP/TT	0.900
Not Likely PP/TT	0.881

(9) Expected Number of Claims by Claim Group and Hazard Group = (3) x (4) x (5) x (6) x (7) x (8)

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	2.341	7.172	24.440	15.075	34.933	32.010	14.972
Likely PP/TT	306	705	1,539	560	889	544	155
Not Likely PP/TT	1,528	3,535	7,756	2,806	4,459	2,714	775

Note: Claim counts for the Fatal claim group are shown to three decimal places.

Appendix D.2—Severity Model

The expected severity is assumed to follow a gamma distribution.

The model can be written as:

$$\log(\mu_{ghr}) = \gamma_g + \xi_s + \eta_{hg} + \psi_{sg}$$

Table 6 shows a description of the terms in the model.

Table 6. Description of terms in severity model

Term	Description
γ_g	base severity for claim group g
ξ_s	factor for state s
η_{hg}	factor for hazard group h specific to claim group g
ψ_{sg}	factor for interaction between state s and claim group g

The parameter ψ_{sg} is credibility weighted (or "shrunk") using multilevel modeling. This is sometimes referred to as *partial pooling*.

Additionally, η_{hg} is assumed to have a structure described later in this section.

The developed, on-leveled, and trended empirical average severities at the claim group-hazard group-state level follow a gamma distribution with parameters adjusted to reflect the reduction in variance associated with an increase in number of claims.

The following are notes on each parameter in the equation:

- The base severity (γ_g) can be thought of as an intercept and is analogous to the idea of a base rate in a rating system.
- The state-specific parameters (ξ_s, ψ_{sg}) account for the overall state variation separately from the state-to-state variation of individual claim groups, especially considering credibility. This is similar to the claim count model.
    - The state factor (ξ_s) is estimated very accurately, since every claim for all the modeled claim groups for a state contributes to its estimation. Using this factor to account for state differences allows for state variation that interacts with claim groups to be estimated more accurately. This is analogous to reducing the variance between groups and has the effect of shrinking the state-claim group (ψ_{sg}) factor toward 1.0.
- The hazard group factor (η_{hg}) differs by claim group but is common for all states and has the following structure across the claim groups:

$$\eta_{hg} = \eta_h \cdot \alpha_g$$

This structure reflects the results of an inspection of empirical hazard group relativities by claim group relativities, where we found that while they were more extreme for certain claim groups, the different claim groups varied similarly across hazard groups. This is similar to the structure for claim counts, but all claim groups in the severity model have a common shape (η_h) and vary only by a common multiple (α_g) that reflects the magnitude of the shape.

Exhibit 2 shows sample values for the calculation of the severities for an NCCI state. Note that Medical-Only severities are determined directly using reported data, and Permanent Total severities are calculated via a separate procedure.

Appendix Exhibit 2. Sample calculation of expected severity by claim group and hazard group for an NCCI state

(1) Base Severity for Claim Group (γ_g)

Claim Group	Base Severity for Claim Group (γ_g)
Fatal	285,559
Likely PP/TT	91,090
Not Likely PP/TT	25,518

(2) State Factor (ξ_s)

State Factor (ξ_s)
0.946

(3) Factor for Hazard Group Specific to Claim Group (η_{hg})

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	1.000	1.097	1.122	1.209	1.281	1.382	1.436
Likely PP/TT	1.000	1.259	1.334	1.604	1.852	2.235	2.454
Not Likely PP/TT	1.000	1.215	1.276	1.491	1.683	1.973	2.136

(4) Factor for Interaction Between State and Claim Group (ψ_{sg})

Claim Group	Claim Group Factor (ψ_g)
Fatal	0.700
Likely PP/TT	1.366
Not Likely PP/TT	1.046

(5) Expected Severity by Claim Group and Hazard Group = (1) x (2) x (3) x (4)

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	189,207	207,468	212,336	228,690	242,346	261,457	271,611
Likely PP/TT	117,736	148,227	157,043	188,869	218,061	263,128	288,954
Not Likely PP/TT	25,262	30,691	32,226	37,664	42,528	49,845	53,949

Appendix E—Treatment of Permanent Total Claims

Exhibits 3 and 4 show sample values for the calculations for the claim counts and severities for the PT claim group for an NCCI state.

Appendix Exhibit 3. Sample calculation of expected claim counts by hazard group for the PT claim group for an NCCI state

(1) State Claim Count: Base Period 5/1/2000 to 4/30/2005. These values are calculated using the same process as shown in Appendix Exhibit 1, but these values include the PT claim group and use an older time period.

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	3.710	14.647	36.657	21.374	50.941	48.294	29.459
Permanent Total	5.013	19.045	53.456	22.731	50.082	33.133	16.509
Likely PP/TT	447	1,125	2,232	850	1,401	919	283
Not Likely PP/TT	2,013	5,036	9,873	3,810	6,263	4,176	1,287
Total Non-PT Lost-Time	2,463	6,176	12,141	4,681	7,715	5,143	1,599

Note: Claim counts for the Fatal and PT claim groups are shown to three decimal places.

(2) Initial Proportion of PT Claim Count to Total Non-PT Lost-Time $= \frac{(1)_{PT}}{(1)_{Total\ Non-PT\ Lost-Time}}$

	Hazard Group						
	A	B	C	D	E	F	G
Proportion	0.00203	0.00308	0.00440	0.00486	0.00649	0.00644	0.01032

(3) Fitted State Claim Counts. These values are taken from the final claim counts shown in Appendix Exhibit 1. The Total Non-PT Lost-Time claim count is calculated.

	Hazard Group						
Claim Group	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Fatal	2.341	7.172	24.440	15.075	34.933	32.010	14.972
Likely PP/TT	306	705	1,539	560	889	544	155
Not Likely PP/TT	1,528	3,535	7,756	2,806	4,459	2,714	775
Total Non-PT Lost-Time	1,836	4,247	9,319	3,381	5,383	3,290	945

(4) Estimated PT Claim Count = $(2) \times (3)_{Total\ Non-PT\ Lost-Time}$

	Hazard Group						
	A	**B**	**C**	**D**	**E**	**F**	**G**
PT Claim Count	3.737	13.098	41.032	16.417	34.942	21.197	9.755

Appendix Exhibit 4. Sample calculation of expected severities by hazard group for the PT claim group for an NCCI state

(1) State PT Severity: Base Period 5/1/2000 to 4/30/2005. These values are calculated using the same process as shown in Appendix Exhibit 2, but these values include the PT claim group and use an older time period. Only the values for the PT claim group are shown here, although data for all lost-time claim groups are used in the model to produce these PT values.

	Hazard Group						
Claim Group	**A**	**B**	**C**	**D**	**E**	**F**	**G**
PT	590,710	817,530	896,879	1,043,220	1,257,983	1,526,118	1,780,806

(2) Calculation of Trend Factors

Trend Stage	Annual Indemnity Trend	Annual Medical Trend	Trend Period Start Date	Trend Period End Date	Number of Years	Indemnity Trend Factor	Medical Trend Factor
First Stage	1.050	1.067	5/15/2003	5/15/2010	7.005	1.407	1.575
Second Stage	1.020	1.030	5/15/2010	4/1/2017	6.885	1.146	1.226

(3) Combined Trend Factors = First Stage Trend x Second Stage Trend

	Indemnity Trend Factor	Medical Trend Factor
Combined Trend Factor	1.613	1.931

(4) Selected On-Level Factor. The on-level factor reflects changes in PT benefit levels between the base period and the effective time period.

	Indemnity	Medical
On-Level Factor	1.101	1.127

(5) PT Indemnity/Medical Split. The PT Indemnity/Medical Split is calculated using developed PT loss dollars in the base period.

	Indemnity	Medical
PT Loss Weight	0.231	0.769

(6) Combined Trend and On-Level Factors

$$Total = (3)_{Indemnity} \times (4)_{Indemnity} \times (5)_{Indemnity} + (3)_{Medical} \times (4)_{Medical} \times (5)_{Medical}$$

	Indemnity	Medical	Total
Combined Trend and On-Level Factors	1.776	2.175	2.083

(7) Estimated PT Severity $= (1) \times (6)_{Total}$

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
PT	1,230,525	1,703,019	1,868,314	2,173,161	2,620,538	3,179,099	3,709,645

Appendix F—Excess Ratio Curve for Losses Including ALAE

Let s denote the ratio of the loss-only severity to the loss-including-ALAE severity, for the claim group and hazard group in a given state. For a fixed loss limit L, let r be the corresponding entry ratio when the limit is viewed as a pure loss and $\hat{r} = sr$ be the entry ratio when that limit is viewed as applying to a loss that includes ALAE. For each claim group i, let E_i be the excess ratio function for the state on a pure loss basis and let $\hat{E}_i$ be the excess ratio function for the state on a loss-with-ALAE basis, with both $\hat{E}_i$ and E_i being functions of the applicable entry ratio. $\hat{E}_i$ is calculated using the same formula structure as E_i but with different parameter values. Let $ALAE_{State}$ and $ALAE_{CW}$ denote state and countrywide ALAE percentages, respectively. Then the formula for the claim group component excess ratio for loss with ALAE is:

$$E_i^{ALAE}(\hat{r}) = Min\left(Max\left(E_i(\hat{r}) + \left(\frac{ALAE_{State}-1}{ALAE_{CW}-1}\right)\left(\hat{E}_i(\hat{r}) - E_i(\hat{r})\right), sE_i(r)\right), 1 - s + sE_i(r)\right).$$

The appropriate value to use for $ALAE_{CW}$ is 1.127. The value to use for $ALAE_{State}$ is the ALAE factor appropriate for the state and time period.

Note: The excess ratio curve for losses with ALAE can be viewed as the weighted sum of two excess ratio functions:

$$E_i(\hat{r}) + \left(\frac{ALAE_{State}-1}{ALAE_{CW}-1}\right)\left(\hat{E}_i(\hat{r}) - E_i(\hat{r})\right) = \left(1 - \frac{ALAE_{State}-1}{ALAE_{CW}-1}\right) E_i(\hat{r}) + \frac{ALAE_{State}-1}{ALAE_{CW}-1}\hat{E}_i(\hat{r})$$

that is then subject to a lower bound of $sE_i(r)$ that corresponds to the case where the additional ALAE has no contribution to the excess and is also subject to an upper bound of $1 - s + sE_i(r)$ that corresponds to the case where the additional ALAE has full contribution to the excess.

Because the excess ratio function for losses with ALAE is determined by this formulaic adjustment of excess ratios, this construction does not provide a parametric claim severity distribution function.

Finally, for the loss limit L, the overall excess ratio for loss with ALAE is the loss-weighted average:

$$E^{ALAE}(L) = \sum_i \widehat{\omega}_i \cdot E_i^{ALAE}(\hat{r})$$

where the $\widehat{\omega}_i$ denote the weights that correspond to itemizing the losses including ALAE for the state and hazard group into claim groups.

Appendix G—Calculation of Severities Including ALAE

Appendix Exhibit 5. Sample calculation of expected severities including ALAE by claim group and hazard group for an NCCI state

(1) Expected loss on a loss-only basis, calculated by multiplying expected claim counts by expected severities. (Fatal, Likely PP/TT, and Not Likely PP/TT from Appendix Exhibit 1. PT values from Appendix Exhibit 3. Medical-Only values calculated as reported from unit data.)

	Hazard Group						
Claim Group	A	B	C	D	E	F	G
Fatal	443,014	1,487,960	5,189,480	3,447,497	8,465,864	8,369,227	4,066,563
PT	4,598,375	22,306,722	76,660,356	35,676,057	91,568,062	67,386,530	36,186,006
Likely PP/TT	36,027,211	104,500,008	241,689,517	105,766,864	193,855,977	143,141,729	44,787,898
Not Likely PP/TT	38,600,488	108,491,174	249,946,987	105,685,608	189,630,436	135,278,756	41,810,279
Medical Only	10,506,947	27,324,096	56,680,683	19,513,664	30,287,788	16,593,726	4,609,704

(2) State ALAE Percentage (from state rate or loss cost filing)

State ALAE Percentage
0.116

(3) Countrywide ALAE Relativities by Claim Group

Claim Group	Countrywide ALAE Relativity
Fatal	0.0590
PT	0.0782
Likely PP/TT	0.1188
Not Likely PP/TT	0.1132
Medical Only	0.1320
Total	0.1067

(4) Ratio of State ALAE to Countrywide ALAE $= \frac{(2)}{(3)_{Total}}$

Ratio of State ALAE to Countrywide ALAE
1.087

(5) State ALAE Percentage by Claim Group= $(3) \times (4)$

Claim Group	Countrywide ALAE Relativity

Fatal	0.0641
PT	0.0850
Likely PP/TT	0.1292
Not Likely PP/TT	0.1231
Medical Only	0.1435

(6) Expected Loss Including ALAE by Claim Group and Hazard Group $= (1) \times [1 + (5)]$

	Hazard Group						
Claim Group	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Fatal	471,430	1,583,402	5,522,346	3,668,627	9,008,886	8,906,050	4,327,403
PT	4,989,311	24,203,149	83,177,708	38,709,090	99,352,806	73,115,458	39,262,393
Likely PP/TT	40,680,292	117,996,668	272,904,839	119,427,146	218,893,376	161,629,148	50,572,463
Not Likely PP/TT	43,350,917	121,842,808	280,707,098	118,691,970	212,967,598	151,927,045	46,955,725
Medical Only	12,014,749	31,245,245	64,814,654	22,313,975	34,634,242	18,975,011	5,271,220

(7) Expected Number of Claims from Appendix Exhibits 1 and 3. Medical-Only claim counts are determined directly using reported unit data.

	Hazard Group						
Claim Group	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Fatal	2.341	7.172	24.440	15.075	34.933	32.010	14.972
PT	3.737	13.098	41.032	16.417	34.942	21.197	9.755
Likely PP/TT	306	705	1,539	560	889	544	155
Not Likely PP/TT	1,528	3,535	7,756	2,806	4,459	2,714	775
Medical Only	8,756	19,492	40,696	12,956	18,431	9,125	2,401

(8) Expected Severity including ALAE by Claim Group and Hazard Group $= \frac{(6)}{(7)}$

	Hazard Group						
Claim Group	**A**	**B**	**C**	**D**	**E**	**F**	**G**
Fatal	201,344	220,775	225,955	243,358	257,890	278,227	289,033
PT	1,335,139	1,847,802	2,027,150	2,357,914	2,843,326	3,449,373	4,025,024
Likely PP/TT	132,942	167,371	177,326	213,263	246,224	297,112	326,274
Not Likely PP/TT	28,371	34,468	36,192	42,299	47,761	55,979	60,588
Medical Only	1,372	1,603	1,593	1,722	1,879	2,079	2,195

Appendix H—Comparing Per-Claim and Per-Occurrence Excess Ratios

Table 7 itemizes multiple claim occurrences according to the number of claims in the occurrence:

Table 7. Probability of multiclaim occurrences containing different numbers of claims

Claim Count per Non-Singleton Occurrence	Probability
2	73.3%
3	14.3%
4	5.1%
5	2.4%
6	1.2%
7	0.8%
8	0.6%
9	0.4%
10	0.2%
More than 10	1.7%

Table 8 gives the overall breakdown of losses and claims according to whether they belong to a singleton or a multiple claim occurrence:

Table 8. Losses and claim counts by type of occurrence

Type of Occurrence	Losses	Claim Counts
Singleton	88.8%	98%
Non-Singleton	11.2%	2%
Combined	100.0%	100%

Table 9 itemizes losses and claim counts by claim group, according to whether the claim belongs to a singleton or a multiple claim occurrence:

Table 9. Losses and claim counts by type of occurrence and claim group

	Singleton		Non-Singleton	
Claim Group	Losses	Claim Counts	Losses	Claim Counts
Fatal	0.5%	0.1%	12.4%	2.8%
Permanent Total	7.6%	0.1%	18.5%	0.5%
PP and TT Likely	34.8%	4.2%	58.9%	22.4%
PP and TT Not Likely	49.5%	19.2%	8.9%	9.7%
Medical Only	7.6%	76.4%	1.3%	64.6%

Table 10 gives the severity relativities within each claim group, according to whether the claim belongs to a singleton or a multiple claim occurrence:

Table 10. Severity relativities by type of occurrence and claim group

Claim Group	Singleton	Non-Singleton	Combined
Fatal	0.947	1.017	1.000
Permanent Total	0.861	1.452	1.000
PP and TT Likely	0.896	1.485	1.000
PP and TT Not Likely	0.974	2.134	1.000
Medical Only	0.995	1.221	1.000

Assume that for claims in multiple occurrences:

- Per-claim severity distribution F_m does not vary by size of occurrence
- The correlation ρ between claim sizes within an occurrence also does not vary by the number of claims in the occurrence (estimated from the data to be about 0.25)

Let p_i be the probability that a claim is in an occurrence with exactly i claims. Then we have a formula to estimate the excess ratio $R_{G_m}(x)$ of the distribution G_m of multiple occurrences:

$$R_{G_m}(x) \approx (1-\omega)R_{F_m}(x) + \omega \sum_{i=2}^{\infty} p_i\, R_{F_m}\left(\frac{x}{i}\right)$$

where $\omega = (1 + \rho CV_{F_m}^{2})/(1 + CV_{F_m}^{2})$ and $R_{F_m}(x)$ is the excess ratio function of F_m.

Let F be the overall per-claim severity distribution and G the per-occurrence severity distribution. The relativity of means and CVs between claims from singleton and multiple occurrences are used to derive per-claim severity distributions F_s and F_m for singleton and multiple claims, respectively. The loss weight α for singletons is then:

$$\alpha = \frac{p_1\mu_{F_s}}{p_1\mu_{F_s} + \left(1 - p_1\mu_{F_m}\right)} = \frac{p_1\mu_{F_s}}{\mu_F}.$$

A model that simulates grouping claims into occurrences suggests a value of about 0.98 for the probability p_1 that a claim is a singleton.

Finally, putting the pieces together provides an estimate for the excess ratio function of the distribution G of occurrences:

$$R_G(x) = \alpha R_{F_s}(x) + (1-\alpha)R_{G_m}(x)$$

$$\approx \alpha R_{F_s}(x) + (1-\alpha)\left((1-\omega)R_{F_m}(x) + \omega \sum_{i=2}^{N} p_i\, R_{F_m}\left(\frac{x}{i}\right)\right).$$

Table 11 gives the per-occurrence excess ratios that correspond to certain per-claim excess ratios. Linear interpolation is used on the per-claim excess ratios to get the per-occurrence excess ratio that corresponds to the exact desired per-claim excess ratio.

Table 11. Per-claim excess ratio to per-occurrence excess ratio conversion

Excess Ratios	
Per Claim	Per Occ
1.000000	1.000000
0.990000	0.990032
0.980000	0.980062
0.970000	0.970092
0.960000	0.960123
0.950000	0.950155
0.940000	0.940189
0.930000	0.930226
0.920000	0.920264
0.910000	0.910305
0.900000	0.900349
0.890000	0.890395
0.880000	0.880443
0.870000	0.870494
0.860000	0.860546
0.850000	0.850600
0.840000	0.840656
0.830000	0.830714
0.820000	0.820773
0.810000	0.810835
0.800000	0.800898
0.790000	0.790962
0.780000	0.781027
0.770000	0.771095
0.760000	0.761163
0.750000	0.751234
0.740000	0.741306
0.730000	0.731379
0.720000	0.721453
0.710000	0.711530
0.700000	0.701607
0.690000	0.691686
0.680000	0.681767
0.670000	0.671849
0.660000	0.661933
0.650000	0.652019
0.640000	0.642106
0.630000	0.632194
0.620000	0.622285
0.610000	0.612377
0.600000	0.602471
0.590000	0.592566
0.580000	0.582664
0.570000	0.572763
0.560000	0.562864
0.550000	0.552967
0.540000	0.543071
0.530000	0.533177
0.520000	0.523285
0.510000	0.513395
0.500000	0.503507
0.490000	0.493620
0.480000	0.483735
0.470000	0.473851
0.460000	0.463970
0.450000	0.454089
0.440000	0.444210
0.430000	0.434332
0.420000	0.424456
0.410000	0.414580
0.400000	0.404706
0.390000	0.394832
0.380000	0.384958
0.370000	0.375085
0.360000	0.365212
0.350000	0.355338
0.340000	0.345464
0.330000	0.335588
0.320000	0.325711
0.310000	0.315832
0.300000	0.305951
0.290000	0.296066
0.280000	0.286178
0.270000	0.276286
0.260000	0.266388
0.250000	0.256485
0.240000	0.246574
0.230000	0.236656
0.220000	0.226730
0.210000	0.216794
0.200000	0.206847
0.190000	0.196889
0.180000	0.186917
0.170000	0.176933
0.160000	0.166933
0.150000	0.156917
0.140000	0.146884
0.130000	0.136833
0.120000	0.126763
0.110000	0.116673
0.100000	0.106561
0.090000	0.096426
0.080000	0.086265
0.070000	0.076073
0.060000	0.065843
0.050000	0.055563
0.040000	0.045208
0.030000	0.034737
0.020000	0.024062
0.010000	0.012971
0.005000	0.007075
0.001000	0.001831
0.000500	0.001051
0.000100	0.000305
0.000050	0.000181
0.000010	0.000053
0.000000	0.000000

Appendix I—Illustration of Calculation of Per-Occurrence Excess Ratios Not Including ALAE (Selected Loss Limits)

Appendix Exhibit 6. Sample calculation of per-occurrence excess ratios not including ALAE by loss limit for Hazard Group A for an NCCI state

(1) Severities, not including ALAE, Hazard Group A, calculated from Appendix Exhibits 2 and 4:

Claim Group	Severity
Fatal	189,207
PT	1,230,525
Likely PP/TT	117,736
Not Likely PP/TT	25,262
Medical Only	1,200

(2) Loss weights, not including ALAE, Hazard Group A, calculated from Appendix Exhibit 5:

Claim Group	Loss Weight
Fatal	0.005
PT	0.051
Likely PP/TT	0.400
Not Likely PP/TT	0.428
Medical Only	0.117

(3) Entry Ratios by Loss Limit and Claim Group = $\frac{Loss\ Limit}{(1)}$:

Loss Limit	Fatal	PT	Likely PP/TT	Not Likely PP/TT	Medical Only
$10,000	0.05	0.01	0.08	0.40	8.33
$100,000	0.53	0.08	0.85	3.96	83.34
$500,000	2.64	0.41	4.25	19.79	416.68
$1,000,000	5.29	0.81	8.49	39.58	833.35
$5,000,000	26.43	4.06	42.47	197.92	4166.77

(4) Excess Ratio Curve Parameters by Claim Group (as described in Appendix C.1):

Parameter	Fatal	PT	Likely PP/TT	Not Likely PP/TT	Medical Only
μ_1	–0.145	–0.490	–0.279	–1.619	–0.899
μ_2	–2.209	–1.677	–1.229	–0.222	–1.180

σ_1	0.801	1.127	0.783	1.774	1.269
σ_2	1.727	1.269	1.564	0.920	2.457
ω_1	0.727	0.789	0.152	0.836	0.983
ω_2	0.273	0.211	0.848	0.164	0.017
a	5.85	6.47	56.20	125.00	626.00
b	3.660	4.121	36.530	90.485	1068.114
m	0.67	0.72	0.59	0.47	0.96

(5) Excess Ratios by Claim Group (as calculated in Appendix C.1):

Loss Limit	Fatal	PT	Likely PP/TT	Not Likely PP/TT	Medical Only
$10,000	0.950	0.992	0.923	0.758	0.127
$100,000	0.597	0.921	0.564	0.291	0.044
$500,000	0.120	0.686	0.219	0.087	0.022
$1,000,000	0.039	0.508	0.122	0.043	0.014
$5,000,000	0.003	0.120	0.018	0.005	0.004

(6) Per-Claim Excess Ratio= $\sum_{Claim\ Group}(2)\cdot(5)$:

Loss Limit	Per-Claim Excess Ratio
$10,000	0.763
$100,000	0.405
$500,000	0.163
$1,000,000	0.095
$5,000,000	0.016

(7) Per-Occurrence Excess Ratio (using linear interpolation on Table 11 from Appendix H):

Loss Limit	Per-Occ Excess Ratio
$10,000	0.764
$100,000	0.410
$500,000	0.170
$1,000,000	0.102
$5,000,000	0.020

References

Corro, D. and G. Engl, "The 2004 NCCI Excess Loss Factors," Casualty Actuarial Society *Forum*, Fall 2006, pp. 513–571, www.casact.org/pubs/forum/06fforum/517.pdf.

Daley, T.V., "Class Ratemaking for Workers Compensation: New Developments in Loss Development," *Variance*, Volume 6, Issue 2, 2012, pp. 196–230.

Embrechts, P., C. Klüpperberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer-Verlag, 2003.

Evans, J., "Workers Compensation Excess Loss Development," NCCI, 2011, https://www.ncci.com/Articles/Documents/II_Excess_Loss_Research_Brief.pdf.

Gillam, W. R., "Retrospective Rating: Excess Loss Factors," *Proceedings of the Casualty Actuarial Society* 78, 1991, pp. 1–40, www.casact.org/pubs/proceed/proceed91/91001.pdf.

Gillam, W.R., and J. Couret, "Retrospective Rating: 1997 Excess Loss Factors," *Proceedings of the Casualty Actuarial Society* 84, 1997, pp. 450–481.

Mahler, H., "Retrospective Rating: 1997 Excess Loss Factors [Discussion], " *Proceedings of the Casualty Actuarial Society* 85, 1998, pp. 316–344.

McNeil, A., "Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory," *ASTIN Bulletin* 27, 1997, pp. 117–137.

Robertson, J.P., "NCCI's 2007 Hazard Group Mapping," *Variance*, Volume 3, Issue 2, 2009, pp. 194–213.