WORKERS COMPENSATION EXCESS RATIOS: AN ALTERNATIVE METHOD OF ESTIMATION

HOWARD C. MAHLER

Abstract

This paper presents an alternative method of calculating excess ratios for workers compensation insurance. While the method shares many similarities with that presented by Gillam [1], there are important differences in approach. The (adjusted) data is relied upon directly for lower limits. For higher limits this is supplemented by a mixed Pareto-exponential distribution fitted to the (adjusted) data.

1. INTRODUCTION

The excess ratio for a limit L is defined as the ratio of losses excess of L to the total ground-up losses. If $f(x)$ is the probability density function for the size of loss distribution, then the excess ratio is defined as:

$$R(L) = \frac{\int_L^\infty (x-L)f(x)\,dx}{\int_0^\infty x f(x)\,dx}.$$

The excess ratio can also be written in terms of the limited expected value $\mathrm{E}[X;L]$ and the mean $\mathrm{E}[X]$: $R(L) = 1 - \mathrm{E}[X;L]/\mathrm{E}[X]$. See for example Hogg and Klugman [2].

The excess ratio is an important statistic with many applications. For example, it can be used to calculate excess loss factors for workers compensation insurance, as discussed in Gillam [1]. Generally, the excess loss factor for a limit is the product of an excess ratio and a permissible loss ratio with the possible addition of a risk load.[1]

[1] The similar excess loss and allocated expense factors (ELAFs) are for use in the ALAE option for retrospective rating. Let $1+a$ be the factor to load losses for allocated loss adjustment expense. Then for an accident limit L, one computes $R(L/(1+a))$ and multi-

Since excess loss factors are typically calculated by hazard group and accident limit, excess ratios need to be estimated by hazard group and by accident limit. This paper will show one method of estimating such excess ratios, with an emphasis on general principles rather than on the important details that may affect the estimate in specific situations.

2. DATA

As always, the first step is to collect the appropriate data. As in Gillam [1], Unit Statistical Plan data is used at third, fourth, and fifth report. All medical only losses are assumed to be below any accident limit. For lost time claims, the Unit Statistical Plan data has the individual claim size[2] for those claims greater than $2,000. (Those claims of size less than $2,000 may be reported on a grouped basis; all of their losses are below any accident limit.)

The reported class codes can be used to divide the data into hazard groups.[3] Using the reported information the proportion of loss dollars excess of any accident limit can be calculated.[4]

In order to illustrate the method in this paper, it will be applied to Massachusetts workers compensation data from composite policy years 1988/1989 at 5th report,[5] 1989/1990 at 4th report and 1990/1991 at 3rd report. In practice it is often appropriate to examine indications using data from several evaluation dates.

---

plies by a permissible loss and allocated LAE ratio, in order to get an ELAF. This method of calculating ELAFs assumes that the expected ALAE ratio is approximately the same for all claim sizes. The effects of variations in this assumption are beyond the scope of this paper. (ALAE data by claim has only recently started to be collected by workers compensation rating bureaus.)

[2] Paid losses plus case reserves, divided between medical and indemnity.

[3] There are four hazard groups, with hazard group 4 having the highest expected claim severity.

[4] Provided the accident limit is greater than $2,000 (times any adjustment factors).

[5] Composite policy year 1988/1989 includes all data from policies with effective dates from July 1, 1988 to June 30, 1989. Fifth report is evaluated 66 months from policy inception.

3. ADJUSTMENTS TO DATA

The claim severity is adjusted from that observed in the data to that expected in the policy effective period at the appropriate report.[6] Each claim is multiplied by an appropriate trend, law amendment and development factor. The product of these adjustment factors for a particular example is shown in Exhibit 1. These adjustment factors should be calculated in a manner consistent with the procedures that produced the rates.[7] In other words, whatever procedure was used to project past losses to ultimate in the effective period in order to estimate rates, should also be applied to the data in order to estimate excess ratios on a consistent basis.

In the example presented in Exhibit 1, the adjustment factors vary by injury kind and between medical and indemnity. To the extent that expected severity trend, development, or law changes differ significantly by claim size within injury kind, the adjustment factors could be varied by claim size interval as well. This refinement is beyond the scope of this paper.

The excess loss factors are used in pricing excess coverage on a per occurrence basis, as discussed in Gillam [1]. Therefore, we are interested in a distribution of loss values for accidents rather than claims. After the above adjustments, all claims are grouped into accidents except medical only claims.[8] A computer program was used which groups data by hazard group, accident date, and policy number, on the assumption that a single policy will not incur two or more accidents on one particular date.[9]

---

[6] The factors in Exhibit 1 include an estimate of average development to ultimate. However, the impact of the dispersion of claim sizes due to development beyond fifth report has not been taken into account.

[7] While the adjustment factors are an important part of the process, they do not represent a difference in the method presented. Therefore, the details are beyond the scope of this paper. Gillam [1] gives an example. See Feldblum [3] for a general discussion of workers compensation ratemaking.

[8] Medical only losses are much smaller than the accident limits purchased, and thus none of them will exceed a relevant loss limit. All medical only losses are assumed to be primary.

[9] Claims without a reported accident date are grouped by hazard group and claim number.

Exhibit 2 shows a comparison of excess ratios computed using ungrouped claim data and data grouped into accidents. For ungrouped data, the excess ratios obtained after adjusting the limit by dividing by a factor of 1.1 were also examined.[10]

The differences between excess ratios calculated from the ungrouped and grouped data were relatively small. At lower limits the 1.1 factor seemed to produce too much of an adjustment, but at higher limits it approximated the effect of the grouping of the data into accidents.

It should be noted that while these results may be interesting, they are far from conclusive. They represent the results for one state for one point in time. At the higher limits random fluctuations are expected to produce differing results over time. Even more importantly, the method used to group claims into accidents is far from perfect. Thus, it is inappropriate to assume the difference represents an error in either method of accounting for multi-claim accidents.

The accident data resulting from the grouping process forms the basis of the analysis.[11] The excess ratios computed from this data are shown in Exhibit 3.

4. CURVE FITTING PROCEDURE

The mean residual life statistic provides a convenient way to examine the tails of loss distributions.[12] The mean residual life at a limit x is defined as $e(x) = (\text{dollars excess of } x)/(\text{number of accidents larger than } x)$. Figure 1 displays the mean residual lives for each of the four hazard groups. As expected, the higher the hazard group the larger the mean residual life. However, as we reach higher limits the data in the two smaller hazard groups,

---

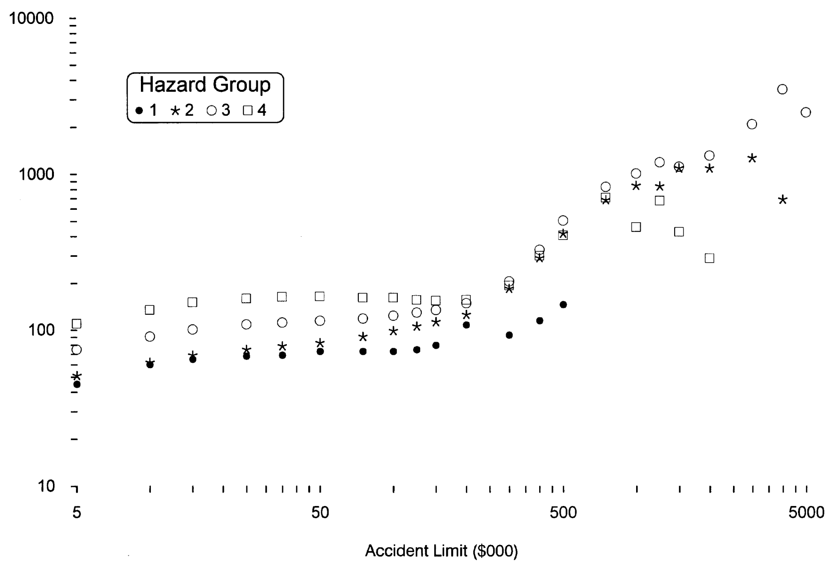
[10] This is similar to the method in Gillam [1].

[11] For the three composite policy years combined there were a total of 157,726 lost-time accidents of which 13,699 had adjusted values greater than $100,000.

[12] See for example Hogg and Klugman [2]. The mean residual life is the average excess cost of a claim that exceeds a given limit.

FIGURE 1

OBSERVED MEAN RESIDUAL LIVES ($000) BY HAZARD GROUP



Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
Based on the following number of lost-time claims: HG1 1391, HG2 97876, HG3 57307, HG4 1114.

1 and 4, becomes sparse. The chance of a very large claim appearing in the data for these hazard groups is too small[13] to get a reliable estimate of the mean residual life at high limits.

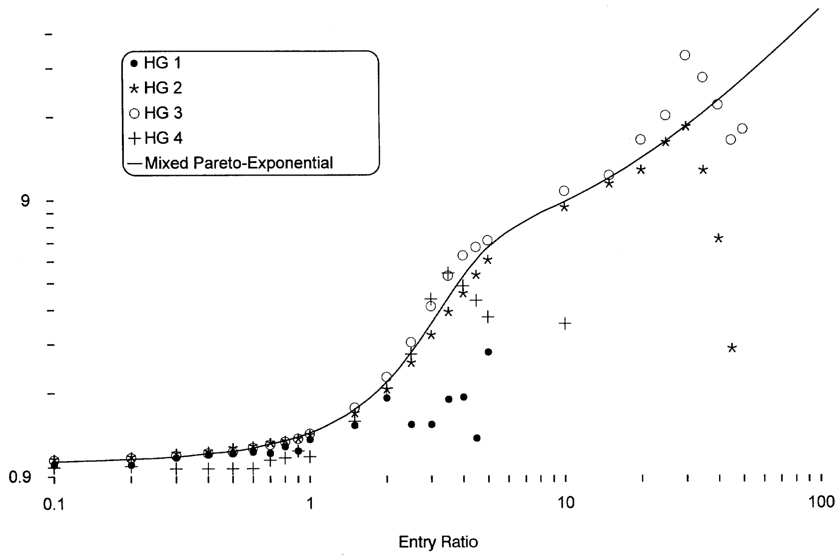
The hazard groups seem to have a similar pattern, with the mean residual life increasing, at least up to an accident limit of several million dollars. A number of adjustments are made to the accident data in order to fit a distribution to it.

---

[13]For example, we can estimate that on average we expect about 0.4 accidents greater than $1 million for hazard group 1. This is based on 96 accidents greater than $100,000 in the data set for hazard group 1 and a tail probability of the fitted mixed Pareto-exponential distribution of .0038 at an entry ratio of 12.4. (.0038)(96) $\approx$ .4. In the reported data there were 39 accidents greater than $1 million of which none were in hazard group 1, 11 were in hazard group 2, 27 were in hazard group 3, and 1 was in hazard group 4.

FIGURE 2

MEAN RESIDUAL LIVES BY HAZARD GROUP
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



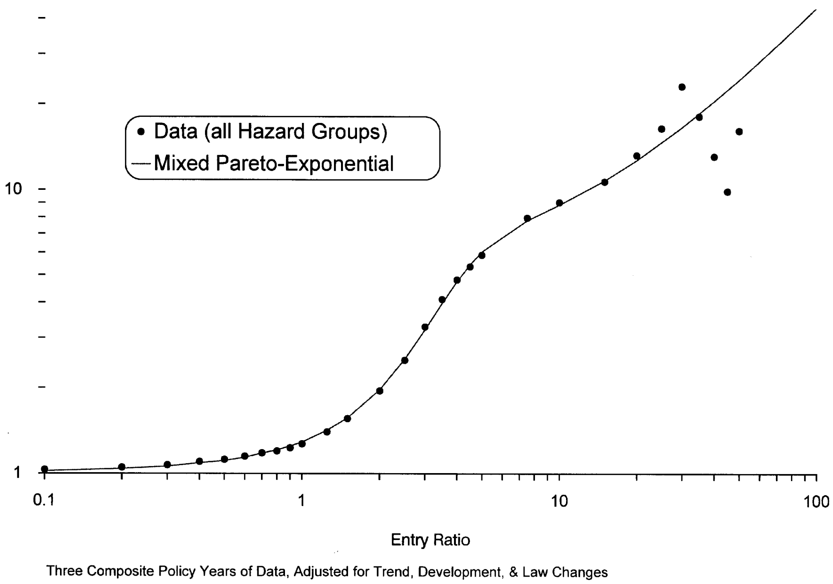
Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
Based on the following number of lost-time claims: HG1 1391, HG2 97876, HG3 57307, HG4 1114.

First, the accident data for third, fourth, and fifth report are combined. Next, for each of the four hazard groups, the data are truncated and shifted at $100,000.[14] Finally, each of these four sets of data is normalized to a mean of unity. Figure 2 shows the mean residual lives for the resulting truncated, shifted and then normalized data by hazard group. Bearing in mind the limited data for hazard groups 1 and 4, it is plausible that the normalized

[14]Accidents with losses less than or equal to $100,000 are eliminated from consideration (for now). Those of size $x > \$100,000$ have $100,000 subtracted from them and appear in the truncated and shifted data as $x - \$100,000$.

FIGURE 3

MEAN RESIDUAL LIVES, OBSERVED VS. FITTED
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



hazard group data might all come from approximately the same distribution. These four sets of normalized data are combined, as displayed in Figure 3.

A mixture[15] of Pareto and exponential distributions is fit to this combined data[16] using the method of maximum likeli-

[15] See for example, Hogg and Klugman [2] for a discussion of the mixture of loss distribution models. The probability density function is $f(x) = pg(x) + (1-p)h(x)$, where g and h are each probability density function.

[16] The data has been combined across reports, injury kinds, and hazard groups, representing over 13,000 accidents over $100,000 in size.

hood.[17] The Pareto and exponential curves are standard size of loss distributions, described in Exhibits 5 and 6. The mixed distribution is (p) (Pareto distribution) + ($1 - p$) (exponential distribution) where p is a fitted parameter with a value between zero and one. Together with the two Pareto parameters (shape and scale) and the single exponential parameter, the mixed distribution has a total of four parameters.

The fitted parameters are displayed in Exhibit 4. Figure 3 compares the mean residual lives for the fitted curve and the observed data. Figure 4 shows the probability density functions for the mixed Pareto-exponential as well as the Pareto and exponential distributions. For small entry ratios the mixed curve behaves as the short-tailed exponential, while for larger entry ratios it behaves as the long-tailed Pareto.

Figure 5 compares the excess ratios for the mixed distribution to that of the exponential and the Pareto. As derived in the Appendix, the excess ratio for the mixed distribution is a weighted average of the excess ratios of the individual distributions, with the weights being (p) (mean of Pareto) and ($1 - p$) (mean of the exponential). In this case, the weights are .2132 and .7868.[18] Thus, for lower entry ratios the excess ratio of the mixed distribution is close to that for the exponential. At higher limits, the excess ratio for the short-tailed exponential is too small to contribute significantly. Therefore, the excess ratio of the mixed distribution for higher entry ratios is approximately 21% of that for the Pareto.

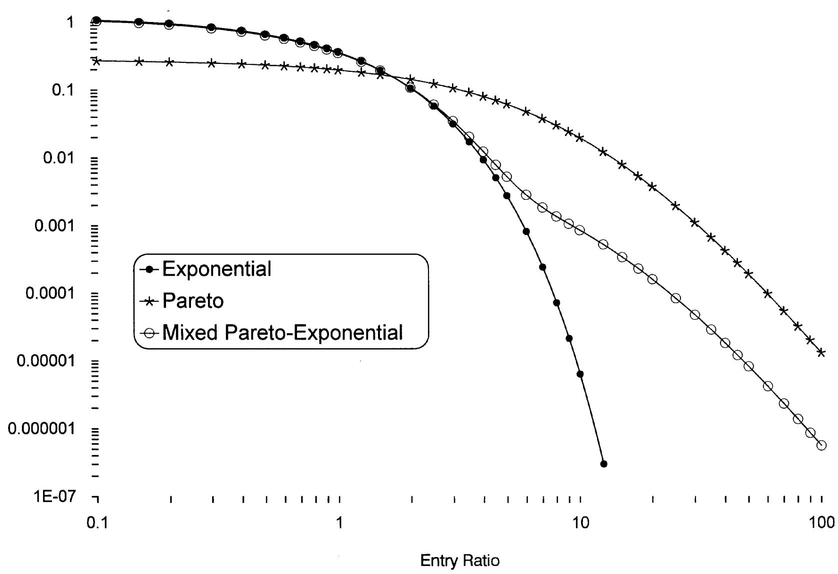
---

[17] The result of the maximum likelihood method has a mean slightly different from unity, so the scale parameters of the Pareto and exponential have been adjusted so as to have the desired mean of unity. The method of maximum likelihood is a commonly used method for fitting size of loss distributions to either grouped or ungrouped data, as discussed in Hogg and Klugman [2]. In this case, the method was applied to the individual data points rather than data grouped into intervals.

[18] This is based on $p = .04294$, a mean of the Pareto of $12.83704/(3.58490 - 1) = 4.9662$ and a mean of the exponential of .82205, as shown in Exhibit 4. $.2132 = (.04294)(4.9662) / \{(.04294)(4.9662) + (.95706)(.82205)\}$.

FIGURE 4

PROBABILITY DENSITY FUNCTIONS
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



Pareto parameters: 3.58490 and 12.83704. Exponential parameter: .82205. Weight to the Pareto: .04294.

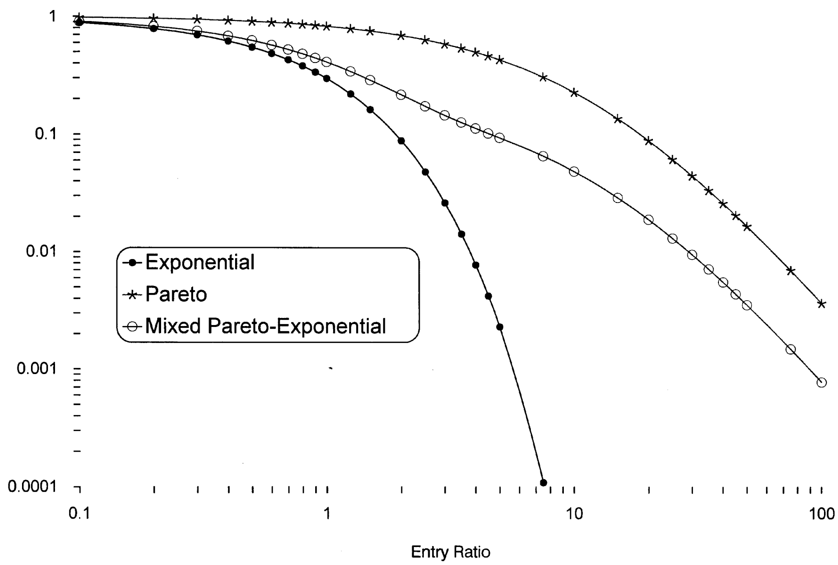
5. ESTIMATION OF EXCESS RATIOS

For each hazard group this fitted curve, scaled to the observed mean, is used in Exhibit 7 to estimate the excess ratios for the data truncated and shifted at $100,000.

The excess ratios for accident limits less than or equal to $100,000 are determined directly from the data. For accident limits L above $100,000, the excess ratio is estimated from the product of (empirical excess ratio at $100,000) $\times$ (excess ratio estimated from mixed Pareto-exponential curve for $L - \$100,000$). See the Appendix. The former is shown in Exhibit 2, the latter in

FIGURE 5

EXCESS RATIOS
Data Truncated and Shifted at $100,000 and then Normalized to a Mean of Unity



Pareto parameters: 3.58490 and 12.83704. Exponential parameter: .82205. Weight to the Pareto: .04294.

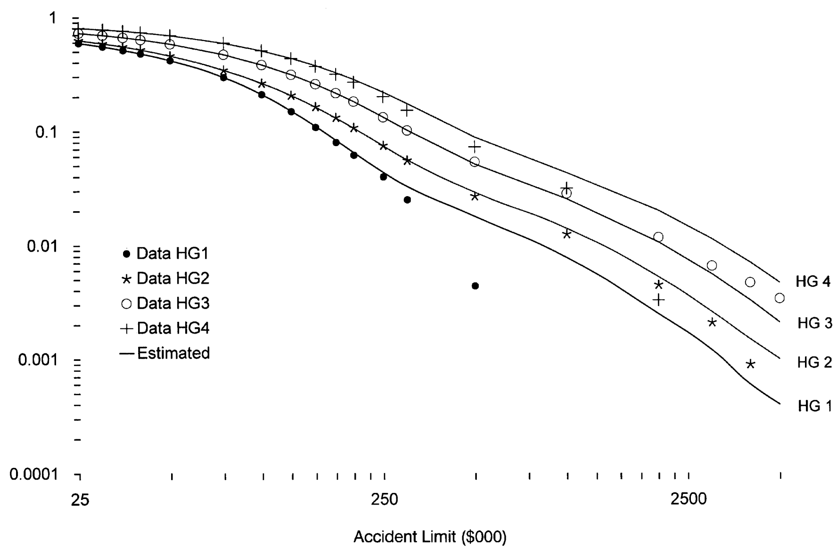
Exhibit 7, while the product is in Exhibit 8.[19] Figure 6 compares the estimated and observed excess ratios.

This method provides a smooth transition from relying on data for lower accident limits to relying on a fitted curve to provide some of the information at higher accident limits. It is important to note that even at higher accident limits an important contribution to the excess ratio is $R(100{,}000)$ which is calculated directly from the data.

[19]It should be noted that for a limit of $100,000 the two methods automatically give the same answer since the excess ratio estimated from the curve at 0 is always unity.

FIGURE 6

EXCESS RATIOS BY HAZARD GROUP
Observed versus Estimated



Three Composite Policy Years of Data, Adjusted for Trend, Development, & Law Changes.
For accident limits of $100,000 or less, the estimated excess ratio is equal to the observed excess ratio.

6. SELECTION OF A TRUNCATION POINT

The $100,000 truncation point was selected to permit the maximum reliance on reported data while still retaining enough data above the truncation point to permit the reasonable fitting of a loss distribution. For this technique and data set, a truncation point around $100,000 achieves the desired balance. Other values such as $50,000 or $150,000 could also have been used without substantially altering the estimated excess ratios.

In general, the truncation point should be a round number prior to the "thinning out" of the data. In this data set there are over 13,000 accidents with values greater than $100,000, with the two smallest hazard groups having about 100 or 200 accidents.[20] For the two larger hazard groups, a higher truncation point could have been selected, but for hazard groups 1 and 4 a higher truncation point would make it difficult to get a reliable average value to use to normalize the data.[21]

7. FEATURES OF THE PROCEDURE

This procedure allows us to rely on the actual data for the lower layers where there is a larger volume of data less subject to random fluctuation. The task of fitting curves to the smaller accidents is avoided totally.

Fitting curves to the combined data regardless of injury kind allows claims to be grouped into accidents.[22] It also avoids relying on the sometimes arbitrary or judgmental division of claims between injury kind.[23] The mixed Pareto-exponential distribution fit to the truncated and shifted data assigns the preponderance of weight to the short-tailed exponential distribution.[24] The long-tailed Pareto distribution models the behavior of the extreme tail of the accident distribution and has a very large effect on the estimated excess ratios for limits over $500,000.

Thus, the estimation procedure can be viewed in terms of three layers. The layer of losses below $100,000 is estimated without curve fitting. The layer from $100,000 to about $500,000 is

[20]There are 96 accidents from hazard group 1 and 228 accidents from hazard group 4.

[21]These average values are used in Exhibit 7 in order to calculate excess ratios by hazard group.

[22]An accident may consist of claims of several different injury kinds. For the calculation of the effect of accident limits it is not inherently necessary to divide dollars between injury kinds.

[23]Note, however, that prior to grouping by accident, claims of differing injury kinds have somewhat different adjustment factors applied to them, as shown in Exhibit 1.

[24]As is common in the use of mixed distribution, a mixture of a longer and shorter tailed distribution was selected. Originally, the short-tailed distribution was a Weibull. However, the fitted Weibull portions of the mixed distribution were very close to an exponential. Therefore, the one parameter exponential was substituted for the two parameter Weibull of which the exponential is a special case.

modeled largely by the exponential distribution. The layer above about $500,000 is modeled largely by the Pareto distribution.[25]

8. CONCLUSION

Actuaries should be familiar with the Pareto distribution, the exponential distribution, and truncated and shifted data. These basic concepts have been employed together in a procedure with a powerful ability to fit the observed data. This procedure of estimating excess ratios is likely to be useful in various practical applications.

[25]The parameters of the fitted Pareto-exponential determine the approximate layers above $100,000. Although it may be conceptually useful to think of it that way, there is no actual division into layers above $100,000.

REFERENCES

[1] Gillam, William R., "Retrospective Rating: Excess Loss Factors," *PCAS* LXXVIII, 1991, pp. 1–40.

[2] Hogg, Robert V. and Stuart A. Klugman, *Loss Distributions*, New York, Wiley, 1984.

[3] Feldblum, Sholom, "Workers' Compensation Ratemaking," *Casualty Actuarial Society Forum*, Special Edition including 1993 Ratemaking Call Papers.

[4] Venter, Gary G., "Scale Adjustments to Excess Expected Losses," *PCAS* LXIX, 1982, pp. 1–14.

EXHIBIT 1

COMBINED TREND, LAW, AND DEVELOPMENT FACTORS

INDEMNITY

Composite Pol. Yr.	Injury Kind 1	Injury Kind 2	Injury Kind 3	Injury Kind 4	Injury Kind 5
88/89	1.79	1.82	1.41	1.28	1.04
89/90	1.53	1.87	1.38	1.26	0.95
90/91	1.42	1.56	1.44	1.31	0.91

MEDICAL

Composite Pol. Yr.	Injury Kind 1	Injury Kind 2	Injury Kind 3	Injury Kind 4	Injury Kind 5	Injury Kind 6
88/89	2.29	2.29	2.29	1.85	1.85	1.85
89/90	3.93	2.06	2.15	1.74	1.61	1.67
90/91	3.30	1.96	2.00	1.62	1.38	1.50

Notes: Product of separate factors calculated to bring all losses to ultimate and a common level, consistent with a 10/1/96 effective date. Injury Kind 1 = Fatal, Injury Kind 2 = Permanent Total, Injury Kind 3 = Major Permanent Partial, Injury Kind 4 = Minor Permanent Partial, Injury Kind 5 = Temporary Total, Injury Kind 6 = Medical Only.

EXHIBIT 2

OBSERVED EXCESS RATIOS FOR UNADJUSTED DATA[1]

Limit ($000)	Accident Data[2]	Claim Data: Using Limit	Claim Data: Using Limit ÷ 1.1
Hazard Group 2			
25	.5230	.5130	.5373
100	.1417	.1342	.1553
500	.0167	.0157	.0171
1,000	.0087	.0078	.0087
2,000	.0042	.0039	.0043
Hazard Group 3			
25	.6335	.6259	.6465
100	.2369	.2276	.2560
500	.0311	.0295	.0324
1,000	.0128	.0118	.0138
2,000	.0042	.0041	.0047

[1]The data for three separate reports, 88/89 at 3rd, 87/88 at 4th, 86/87 at 5th have been combined and then an excess ratio has been calculated. The data have *not* been adjusted for trend, law amendments, or development.

[2]Claims with the same hazard group, accident date, and policy number are grouped into the same accident.

EXHIBIT 3

EXCESS RATIOS BASED ON ADJUSTED DATA[1]

Accident Limit ($000)	Hazard Group 1[2]	Hazard Group 2	Hazard Group 3	Hazard Group 4[2]
25	0.5950	0.6288	0.7283	0.8064
30	0.5530	0.5888	0.6960	0.7817
35	0.5142	0.5521	0.6655	0.7581
40	0.4791	0.5184	0.6366	0.7353
50	0.4177	0.4586	0.5831	0.6918
75	0.2974	0.3441	0.4709	0.5935
100	0.2106	0.2643	0.3832	0.5098
125	0.1494	0.2072	0.3146	0.4353
150	0.1086	0.1647	0.2604	0.3715
175	0.0804	0.1327	0.2171	0.3165
200	0.0622	0.1081	0.1827	0.2699
250	0.0400	0.0754	0.1333	0.2011
300	0.0252	0.0559	0.1021	0.1526
500	0.0044	0.0271	0.0541	0.0730
1,000	0.0000	0.0126	0.0286	0.0317
2,000	0.0000	0.0045	0.0118	0.0033
3,000	0.0000	0.0021	0.0066	0.0000
4,000	0.0000	0.0009	0.0047	0.0000
5,000	0.0000	0.0000	0.0034	0.0000

[1]Massachusetts Workers Compensation, Composite Policy Years 88/89 at 5th, 89/90 at 4th, 90/91 at 3rd.

[2]Note there is relatively little data for hazard groups 1 and 4 since they each represent only between 1 and 2 percent of total premiums. In this reported data there were no accidents greater than $1,000,000 for hazard group 1 and only one for hazard group 4. Thus the empirical excess ratios at higher limits for these hazard groups are poor estimates of future expected excess ratios.

EXHIBIT 4

MIXED PARETO-EXPONENTIAL DISTRIBUTION

Parameters:

Pareto Shape = s	3.58490
Pareto Scale = b	12.83704
Exponential Scale = c	0.82205
Weight to Pareto = p	0.04294
Mean = 1	Coef. of Var. = 1.94
Variance = 3.75	Skewness = 30

Excess Ratios

Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio
0.1	0.9057	11	0.0431
0.2	0.8217	12	0.0387
0.3	0.7470	13	0.0350
0.4	0.6806	14	0.0317
0.5	0.6214	15	0.0288
0.6	0.5687	20	0.0188
0.7	0.5217	25	0.0131
0.8	0.4797	30	0.0095
0.9	0.4422	35	0.0071
1.0	0.4088	40	0.0055
1.25	0.3397	45	0.0044
1.5	0.2872	50	0.0035
1.75	0.2469	55	0.0029
2.0	0.2157	60	0.0024
2.5	0.1722	65	0.0020
3.0	0.1444	70	0.0017
3.5	0.1255	75	0.0015
4.0	0.1118	80	0.0013
4.5	0.1014	85	0.0011
5.0	0.0929	90	0.0010
6.0	0.0797	95	0.0009
7.0	0.0694	100	0.0008
8.0	0.0610		
9.0	0.0540		
10.0	0.0481		

Note: See the Appendix for a sample calculation of an excess ratio.

EXHIBIT 5

PARETO DISTRIBUTION

$$F(x;s,b) = 1 - \left(1 + \frac{x}{b}\right)^{-s}$$

$$f(x;s,b) = \frac{s}{b}\left(1 + \frac{x}{b}\right)^{-(s+1)}$$

$$\mathrm{E}(X^y) = \frac{b^y \Gamma(y+1)\Gamma(s-y)}{\Gamma(s)}, \qquad -1 < y < s$$

If y is an integer N,

$$\mathrm{E}(X^N) = \frac{b^N N!}{\prod_{i=1}^{N}(s-i)} \qquad N < s$$

$$\text{Mean} = \frac{b}{s-1} \qquad \text{Variance} = \frac{b^2 s}{(s-1)^2(s-2)}$$

$$\text{Coefficient of Variation} = \sqrt{\frac{s}{s-2}} \qquad s > 2$$

$$\text{Skewness} = \frac{2(s+1)}{s-3}\sqrt{\frac{s-2}{s}} \qquad s > 3$$

$$\text{Excess Ratio} = R(x) = \left(1 + \frac{x}{b}\right)^{1-s}$$

$$\text{Mean Residual Life} = e(x) = \frac{b+x}{s-1}$$

Note: s is the shape parameter, b is the scale parameter.

EXHIBIT 6

Exponential Distribution

$$F(x;c) = 1 - e^{-x/c}$$

$$f(x;c) = \frac{1}{c}e^{-x/c}$$

$$\mathrm{E}(X^y) = c^y\Gamma(1+y) \qquad y > -1$$

If y is an integer N,

$$\mathrm{E}(X^N) = c^N N! \qquad N > -1$$

$$\text{Mean} = c$$

$$\text{Variance} = c^2$$

$$\text{Coefficient of Variation} = \text{Standard Deviation} \div \text{Mean} = 1$$

$$\text{Skewness} = 2$$

$$\text{Excess Ratio} = R(x) = e^{-x/c}$$

$$\text{Mean Residual Life} = e(x) = c$$

Note: c is the scale parameter.

EXHIBIT 7

EXCESS RATIOS TRUNCATED AND SHIFTED TO $100,000

	Hazard Group 1		Hazard Group 2		Hazard Group 3		Hazard Group 4	
Accident Limit ($000)	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio	Entry Ratio	Excess Ratio
125	0.3446	0.7165	0.2531	0.7810	0.2020	0.8201	0.1541	0.8590
150	0.6891	0.5265	0.5062	0.6180	0.4041	0.6781	0.3083	0.7413
175	1.0337	0.3983	0.7592	0.4962	0.6061	0.5657	0.4624	0.6429
200	1.3783	0.3110	1.0123	0.4049	0.8081	0.4765	0.6165	0.5606
250	2.0674	0.2086	1.5185	0.2838	1.2122	0.3490	0.9248	0.4335
300	2.7566	0.1565	2.0247	0.2131	1.6162	0.2671	1.2331	0.3438
500	5.5131	0.0856	4.0493	0.1107	3.2325	0.1348	2.4662	0.1746
1,000	12.4046	0.0371	9.1110	0.0533	7.2731	0.0669	5.5489	0.0852
2,000	26.1874	0.0120	19.2343	0.0200	15.3543	0.0279	11.7144	0.0399
3,000	39.9702	0.0055	29.3576	0.0098	23.4355	0.0145	17.8798	0.0223
4,000	53.7531	0.0030	39.4809	0.0057	31.5167	0.0087	24.0453	0.0139
5,000	67.5359	0.0019	49.6042	0.0036	39.5979	0.0056	30.2107	0.0093

Entry Ratio = (LIMIT – $100,000 Truncation Point)/Average Size for Data Truncated and Shifted to $100,000.

Average Size of Loss Data Truncated and Shifted to $100,000 by Hazard Group:

HG 1	72,554
HG 2	98,782
HG 3	123,744
HG 4	162,194

Excess ratio is computed for a Pareto-exponential distribution with parameters:

Pareto Shape	Pareto Scale	Exponential	Weight to Pareto
3.58490	12.83704	0.82205	0.04294

EXHIBIT 8

ESTIMATED EXCESS RATIOS BASED ON ADJUSTED DATA AND CURVES FIT TO DATA TRUNCATED AND SHIFTED AT $100,000

Accident Limit ($000)	Hazard Group 1	Hazard Group 2	Hazard Group 3	Hazard Group 4
25	0.5950	0.6288	0.7283	0.8064
30	0.5530	0.5888	0.6960	0.7817
35	0.5142	0.5521	0.6655	0.7581
40	0.4791	0.5184	0.6366	0.7353
50	0.4177	0.4586	0.5831	0.6918
75	0.2974	0.3441	0.4709	0.5935
100	0.2106	0.2643	0.3832	0.5098
125	0.1509	0.2064	0.3143	0.4379
150	0.1109	0.1633	0.2598	0.3779
175	0.0839	0.1311	0.2168	0.3278
200	0.0655	0.1070	0.1826	0.2857
250	0.0439	0.0750	0.1337	0.2210
300	0.0330	0.0563	0.1024	0.1753
500	0.0180	0.0293	0.0517	0.0890
1,000	0.0078	0.0141	0.0256	0.0434
2,000	0.0025	0.0053	0.0107	0.0203
3,000	0.0012	0.0026	0.0056	0.0114
4,000	0.0006	0.0015	0.0033	0.0071
5,000	0.0004	0.0010	0.0021	0.0047

Note: For accident limits of $100,000 or less the excess ratio is taken directly from Exhibit 3. For accident limits larger than $100,000, the excess ratio is a product of that for $100,000 in Exhibit 3 and the excess ratio shown in Exhibit 7. For example, for hazard group 3 at a limit of $1 million, (.3832)(.0669) = .0256.

APPENDIX

Excess Ratios, Truncated and Shifted Data

Let $f(x)$ be the size of loss probability density function. Then the excess ratio for limit L is given by:

$$R(L) = \frac{\int_L^\infty (x-L)f(x)\,dx}{\int_0^\infty xf(x)\,dx}$$
$$= \frac{\text{average dollars of loss excess of } L}{\text{average size of loss}}$$
$$= \frac{\text{total dollars of loss excess of } L}{\text{total dollars of loss}}.$$

Assume we have a truncation point of T. Assume we look at the size of loss distribution for the data truncated and shifted at T. So for a loss $x > T$, we instead look at $x - T$. Then the excess ratio for the truncated and shifted data for ground up limit $L > T$ can be written as

$$\hat{R}(L-T).$$

Assume we were computing the (observed) excess ratio for a \$500,000 accident limit, for hazard group 3 data[26]

$$R(\$500{,}000) = \frac{\text{HG3 Losses Excess of \$500,000}}{\text{Total HG3 Losses (including Medical Only)}}.$$

We can also express this in terms of the data truncated and shifted at \$100,000 as follows:

$$R(\$500{,}000) = \frac{\text{HG3 Losses Excess of \$500,000}}{\text{HG3 Losses Excess of \$100,000}}$$
$$\times \frac{\text{HG3 Losses Excess of \$100,000}}{\text{Total HG3 Losses (including Medical Only)}}.$$

[26]For 3rd, 4th, and 5th report combined, adjusted for trend, law changes, and development.

However, the second term is the excess ratio at \$100,000, $R(\$100{,}000)$, while the first term is $\hat{R}(\$400{,}000)$ = excess ratio at \$400,000 for the data truncated and shifted at \$100,000. Thus

$$R(\$500{,}000) = \hat{R}(\$400{,}000) \times R(\$100{,}000).$$

In general, for limits $L > \$100{,}000$

$$R(L) = \hat{R}(L - \$100{,}000) \times R(100{,}000).$$

In the methodology used here, $\hat{R}(L - \$100{,}000)$ is estimated via a curve fit to the data truncated and shifted at \$100,000, while $R(100{,}000)$ is estimated from the data.

Excess Ratios, Mixed Distributions

Let a (mixed) distribution be a weighted average of two other distributions:

$$f(x) = pg(x) + (1-p)h(x).$$

Then the mean is a weighted average of the two means:

$$\begin{aligned} m_f &= \int_0^\infty xf(x)\,dx = \int_0^\infty x\{pg(x) + (1-p)h(x)\}\,dx \\ &= p\int_0^\infty xg(x)\,dx + (1-p)\int_0^\infty xh(x)\,dx \\ &= pm_g + (1-p)m_h. \end{aligned}$$

The excess ratio for limit L is given by:

$$\begin{aligned} R_f(L) &= \frac{\int_L^\infty (x-L)f(x)\,dx}{\int_0^\infty xf(x)\,dx} \\ &= \frac{p\int_L^\infty (x-L)g(x)\,dx + (1-p)\int_L^\infty (x-L)h(x)\,dx}{pm_g + (1-p)m_h} \\ &= \frac{pm_gR_g(L) + (1-p)m_hR_h(L)}{pm_g + (1-p)m_h}. \end{aligned}$$

So the excess ratio for a mixed distribution is a weighted average of the excess ratios for the individual distributions, with weights equal to the product of the mean of each distribution times the weight in the mixture of each distribution.[27]

For example, for the mixed Pareto-exponential distribution with parameters: 3.5849, 12.83704, .82205, .04294, at an entry ratio of 2, the excess ratio is computed as follows:

excess ratio for Pareto (3.5849, 12.83704) at entry ratio 2 (of the mixed distribution)

$$= \left(1 + \frac{2}{12.83704}\right)^{1-3.5849} = .6878$$

excess ratio for exponential (.82205) at entry ratio 2 (of the mixed distribution)

$$= e^{-2/.82205} = .0878$$

mean for Pareto (3.5849, 12, 83704)

$$= \frac{12.83704}{3.5849 - 1} = 4.9662$$

mean for exponential (.82205) = .82205

excess ratio for Pareto-exponential at entry ratio 2

$$= \frac{(.04294)(4.9662)(.6878) + (1 - .04294)(.82205)(.0878)}{(.04294)(4.9662) + (1 - .04294)(.82205)}$$

$$= .2157/1 = .2157.$$

This matches the value shown on Exhibit 4.

Moments of Mixed Models

Moments of a mixed model are a weighted average of the moments of the individual distributions. For example, for the mixed Pareto-exponential distribution with parameters: 3.5849,

[27]This is closely related to the similar result for increased limit factors discussed in Venter [4].

12.83704, .82205, .04294, the moments are a weighted average using weights of .04294 applied to the moments of the Pareto and 1 – .04294 = .95706 applied to the moments of the exponential.

Moment	Pareto	Exponential	Pareto-exponential
1	4.9662	.82205	1
2	80.4478	1.35153	4.7479
3	5296.86	3.3331	230.64

Then the variance of the Pareto-exponential is $4.7479 - 1^2 = 3.7479$. Note that the variance of the mixed distribution is not the weighted average of the individual variances. The skewness of the Pareto-exponential is

$$\{230.64 - (3)(4.7479)(1^2) + 2(1^3)\}/3.7479^{1.5} = 30.1.$$

The coefficient of variation is $(\sqrt{3.7479})/1 = 1.94$. These match the values shown on Exhibit 4.