

A Bayesian Approach to Excess of Loss Pure Premium Rating

by Jack Barnett

ABSTRACT

This paper demonstrates a Bayesian approach for estimating loss costs associated with excess of loss reinsurance programs. The main features of this approach are that (1) prior severity distributions are adjusted for historical emergence patterns underlying the experience data, (2) maximum likelihood estimation is used to estimate a ground-up loss ratio for each prior severity distribution, (3) a posterior severity distribution is derived using a Bayesian approach, and (4) a posterior ground-up loss ratio is derived using a Bayesian approach. This paper illustrates a simple implementation of the approach and tests the model by simulating from known frequency and severity distributions and fitting the model to the simulated “data.”

KEYWORDS

Bayesian estimation, maximum likelihood estimation, emergence patterns, exposure curves

1. Introduction

When pricing excess of loss (XOL) programs, many reinsurance and large account pricing actuaries may rely on assumptions that are difficult to support. This is particularly true when the actuary is uncertain of the underlying exposure or of which exposure curve(s) to use. An additional complication arises when the reinsurer does not receive all individual claims in the reinsurance submission, because claims are typically reported only above some truncation point. Finally, maximum likelihood estimation (MLE) and, by extension, Bayesian analysis is generally not amenable to modeling long-tailed lines of business in which claims tend to develop upward or downward over time. This paper attempts to solve all of these problems within a single coherent framework.

Meyers (2005) provides a Bayesian methodology for estimating loss costs associated with high layers. An introduction to the methodology can be found as a solution to the COTOR Challenge. It's been many years in the making, but Meyers' solution to the COTOR Challenge was a primary motivation for much of the work I have done here. Following his solution to the COTOR Challenge, Meyers offers solutions to address what he feels are two serious shortcomings of the methodology: (1) the way in which prior severity models were developed and (2) the effect of settlement time on ultimate value. These are major issues for which this paper offers two more solutions. Finally, the framework presented here includes a Bayesian estimate (the mean of the posterior loss ratio distribution) for the ground-up loss ratio. Actuarial literature is relatively absent any discussion of Bayesian loss ratios for XOL rating.

Gelman (2006) distinguishes three categories of priors: (1) prior distributions giving numerical information that is crucial to estimation of the model—such a prior would be a traditional informative prior, which might come from a literature review or explicitly from an earlier data analysis; (2) prior distributions that are not supplying any controversial information but are strong enough to pull the data away from

inappropriate inferences that are consistent with the likelihood—this type of prior might be called a weakly informative prior; and (3) prior distributions that are uniform, or nearly so, and basically allow the information from the likelihood to be interpreted probabilistically—such priors are noninformative priors or maybe, in some cases, weakly informative. The approach taken here uses a weakly informative prior for the prior severity distributions. The prior loss ratio distribution is implied by the prior severity distributions (as described in section 5).

When significant claim volume is available, accurate severity distributions can be estimated by using empirical methods, nonparametric statistics, or maximum likelihood methods (Klugman, Panjer, and Willmot 2009). When little or no data is available, actuaries may rely on an exposure curve. There are numerous approaches for deciding upon an exposure curve. For example, one can use an external curve acquired from a statistical agency or an internal curve developed from a similar business line. In the limiting case, where there is no data, the Bayesian approach defaults to the average of the prior distributions. In such cases, with little or no data, it becomes useful to calibrate prior severity distributions to agree with the exposure curve selected for the underlying insured. In such a case, we prefer an informative prior. As the volume of claim data grows, it becomes less critical for the average of the priors to agree with the exposure curve selected for the underlying insured. This paper does not assume that an exposure curve underlying the insured risk is available. A Bayesian approach is most useful when we are between the two extremes of no data and a significant volume of data. The posterior severity distribution can be viewed as a nonlinear weighing of data and prior curves.

Prior to the application of the model, it will be assumed that the premium is on level and losses have been properly adjusted for trends and changes in attachment points and limits. Improper adjustments may lead to estimates with greater error.

Section 2 outlines how the severity distribution and the likelihood function can be adjusted for age of claim. This is accomplished through estimating

claim count emergence patterns from various sizes of loss triangles and applying these patterns directly to the likelihood function. Section 3 examines 24 severity distributions used as priors for Bayesian analysis purposes. Using “only” 24 curves is a relatively simple implementation of the approach. The curves are not meant to be prescriptive but are selected for illustrative purposes.

Section 4 describes the assumptions and sample data (one simulation) underlying the model. Section 5 describes the process of estimating an underlying loss ratio for each curve using MLE. Section 6 describes the calculation of the posterior weight for each curve. Section 7 describes a Bayesian approach for estimating the ground-up loss ratio weighting over all MLE estimates, also called the mean of the posterior loss ratio distribution. Section 8 compares Bayesian claim count estimates with experience and MLE estimates for 200 sets of data (simulations). Section 9 summarizes and concludes.

2. Adjusting the likelihood function

The likelihood function is the probability of observing the specified data, given a hypothetical value of the parameter. In the example presented here, the data is the number of reported claims in each layer and the hypothetical parameter is a hypothetical severity distribution. The MLE estimate is the distribution yielding the highest probability of obtaining the observed reported claim counts, weighted over all layers. However, in the case of liability business, data is immature. Ultimate claim counts by layer will generally not be the same as reported claim counts by layer. In general, curves fitted to reported and to ultimate claim counts utilizing MLE will not produce the same parameter estimates.

Reported claims may be available, but can we get the severity distribution for ultimate claims? The form of the likelihood function used in this paper is based on grouped data. Let a range of possible values be partitioned as $c_0 < c_1 < \dots < c_k$, where c_0 is the smallest possible value in the model (i.e., the truncation point) and c_k is the largest possible value (i.e.,

the maximum policy limit). Let $[F(c_j; \theta) - F(c_{j-1}; \theta)]$ represent the proportion of ultimate claims in the band c_j to c_{j-1} , given the distribution F and parameter θ . The grouped likelihood function is given by the expression

$$L(\theta) = \prod_j [F(c_j; \theta) - F(c_{j-1}; \theta)]^{n_j}, \quad (1)$$

where n_j is the number of claims in the band. Note that the proportion of claims for the reported data will differ from ultimate claim counts by the emergence pattern. In other words, the relative proportion of claims reported in each band is the following:

$$\% \text{ Reported}_j \propto [F(c_j; \theta) - F(c_{j-1}; \theta)]. \quad (2)$$

For reported claims, the proportion of claims in each band is

$$\% \text{ Reported}_j \propto [F(c_j; \theta) - F(c_{j-1}; \theta)] / \sum_j \% \text{ Reported}_j \propto [F(c_j; \theta) - F(c_{j-1}; \theta)]. \quad (3)$$

The denominator is simply a normalizing constant so that the probabilities sum to unity.

The *adjusted* likelihood function for reported claims is

$$L(\theta) = \prod_j \{ \% \text{ Reported}_j \propto [F(c_j; \theta) - F(c_{j-1}; \theta)] / \sum_j \% \text{ Reported}_j \propto [F(c_j; \theta) - F(c_{j-1}; \theta)] \}^{n_j}. \quad (4)$$

Exhibit 2.1 illustrates one way we might simulate reported claim counts from a lognormal distribution with an underlying emergence pattern. Only 10 of 30,000 simulations are displayed. Column (2) represents an ultimate claim value. At any point in time, the estimated claim value may be more, or less, than the ultimate value. We are assuming an emergence pattern as shown in column (9). Smaller claims, from \$0 to \$100K, are assumed to be reported more quickly. We assume that 90% of all claims less than \$100K have been reported and that of the largest claims, only 60% have been reported. It is common for reinsurers

Exhibit 2.1. One way to simulate reported claim counts from a lognormal distribution with an underlying emergence pattern

		<div style="border: 1px solid black; padding: 5px; display: inline-block;"> <p style="text-align: center;"><i>Actual</i></p> <p>μ: 10.42</p> <p>σ: 2.12</p> </div>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
Simulation	Rand()	Loss	Emergence Pattern	Random Variable	Reported (1=Yes/0=No)	Reported	C_{j-1}	C_j	Emergence Pattern	
1	0.379	17,451	0.90	0.295	1	17,451	0	100,000	0.90	
2	0.942	930,844	0.70	0.943	0		100,000	250,000	0.80	
3	0.068	1,412	0.90	0.837	1	1,412	250,000	500,000	0.75	
4	0.766	155,793	0.80	0.303	1	155,793	500,000	1,000,000	0.70	
5	0.566	47,743	0.90	0.546	1	47,743	1,000,000	2,000,000	0.65	
6	0.344	14,327	0.90	0.638	1	14,327	2,000,000	4,000,000	0.60	
7	0.226	6,789	0.90	0.397	1	6,789	Maximum Policy Limit			
8	0.247	7,843	0.90	0.381	1	7,843				
9	0.763	152,775	0.80	0.955	0					
10	0.967	1,656,936	0.65	0.939	0					

to examine separately the incurred and paid claim count triangles by size of loss. To convert ultimate counts to reported counts, we first augment the emergence pattern, shown in column (3). A unit-uniform random variable is generated, shown in column (4). If the random variable is less than the emergence pattern [i.e., if column (4) < column (3)], the claim is considered reported [column (5) = 1]. If the claim is reported, it is shown in column (6). At this point, roughly 90% of the claims less than \$100K show up as reported claims in column (6), while only 60% of the claims from \$2M to \$4M show up as reported.

Exhibit 2.2 illustrates the application of equation (4) to the reported claim counts simulated in Exhibit 2.1. In practice, it is typical to minimize the negative log-likelihood function, $-\ln[L(\theta)] = \sum_j n_j \ln[(3)]$. Minimizing column (9) produces the parameters shown in the upper left of the exhibit. We can see that the fit is in close agreement with the underlying severity parameters, or “actual.”

So what does this exercise tell us? It tells us that if we have reported claim amounts (counts) and an estimate of the emergence pattern, we can replicate the parameters of the underlying severity distribu-

tion, and we are well on our way to implementing a Bayesian approach. Why not just use adjusted MLE rather than a Bayesian approach? Because we usually don’t see 25,655 reported claims [the total of column (8) in Exhibit 2.2]. Fewer claims leads to greater parameter uncertainty underlying the MLE estimates. For smaller portfolios, Bayesian estimates utilizing the mean of the posterior distribution are less impacted by random variation than are MLE estimates. Krishnamurthy (2017) provides an illustration of and explanation for this case. However, the adjusted MLE does seem especially well suited for pricing very large portfolios or creating exposure curves for individual lines of business.

3. Prior distributions

For this paper, we implement just 24 severity distributions as priors. Exhibit 3.1 displays the parameters and some statistics for the 24 curves. The mean of the prior distribution is an equal weighted average of the curves, labeled “PRIOR (AVG)” (the excess severity is also weighted by the probability of a claim’s exceeding the truncation point, \$100K).

Exhibit 2.2. Application of equation (4) to reported claim counts simulated in Exhibit 2.1

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fit (Estimate)		$\hat{\mu}$ 10.43								
		$\hat{\sigma}$ 2.12								
c_{j-1}	c_j	Emergence Pattern	$F(c_{j-1};\theta)$ lower	$F(c_j;\theta)$ upper	layer probabilities $F(c_j;\theta) - F(c_{j-1};\theta)$	(4)x(1)	(5)/ constant	$\ln(6)$	Reported Claims n_j	negative-log likelihood $-(7)^*(8)$
0	100,000	0.90	0.000	0.694	0.694	0.625	0.734	-0.309	18831	5823.933
100,000	250,000	0.80	0.694	0.826	0.132	0.106	0.124	-2.087	3187	6651.261
250,000	500,000	0.75	0.826	0.897	0.071	0.053	0.062	-2.773	1593	4416.801
500,000	1,000,000	0.70	0.897	0.944	0.047	0.033	0.039	-3.252	982	3193.341
1,000,000	2,000,000	0.65	0.944	0.973	0.028	0.018	0.021	-3.834	574	2205.323
2,000,000	4,000,000	0.60	0.973	1.000	0.027	0.016	0.019	-3.947	488	1926.150
Maximum Policy Limit					1.000	0.851	1.000		25655	24216.808
						Normalizing Constant				Minimize this

“ACTUAL” denotes the statistics associated with the lognormal curve in section 2. It should be noted that statistics and probabilities for the prior curves are net of any insurer attachment/limit profile. These figures can be calculated easily in most exposure rating models. For the implementation here, I have selected lognormal curves for prior distributions. Although we may be unsure of the underlying stochastic process, I have found that in *most* cases this choice of prior produces a reasonable estimate of the underlying severity distribution even when the underlying process is not lognormal. However, using priors that better match the distributional form generally produces estimates with lower variance and bias. I suggest that readers identify the most appropriate distribution underlying their data and use priors that match that distribution. This exposition uses expressions and calculations for the lognormal distribution. MLE is one possible approach for identifying the distribution. Given the same number of parameters, one could select the distribution that produces the greatest value of the likelihood function.

How might prior severity distributions be constructed? Assume we are pricing a \$5M excess \$5M layer on a per-occurrence basis and receive only

claims above some truncation, say \$100K.¹ To calibrate the curves as shown in Exhibit 2.2, and in general, the following procedure can be utilized. Examine the company’s exposure curves. What are the lowest and highest probabilities indicated by the curves for the highest layer, \$5M × \$5M? The prior severity distributions should span the low and high probabilities indicated by the curves. For the exposure curves available to me, a range from 1% to 9% was sufficient. These figures represent the proportion of claim counts above \$5M, given only claims above \$100K. Next, I’ve judgmentally selected three scale parameters: 9, 10, and 11. With the given scale parameters, I use Excel’s solver to find the shape parameters to match the target probabilities shown in Exhibit 3.1, column (14). In practice, I suggest using more than three scale parameters and extending the range of probabilities beyond the highest and lowest indicated by the exposure curves. Extending the ranges reduces the bias that may result when rating an account that tests the extremes of the prior severity curves.

¹For reinsurance, it is uncommon to receive claims at such a low threshold (though I have personally seen all claims for medical malpractice and legal liability lines). However, for large account pricing, it is common to receive all claims. The truncation here is for illustrative purposes only.

Exhibit 3.1. Prior severity distributions

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Curve	Lognormal Parameters μ σ	Excess Severity	Severity Ground Up	Probability >\$100K	% Loss-T $(3)X(5)/4$	35,000 x 100,000	50,000 x 135,000	65,000 x 185,000	250,000 x 500,000	500,000 x 500,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000
									$F(c_1;\theta) - F(c_2;\theta)$				
1	9.0	393,605	71,166	0.117	64.7%	0.219	0.189	0.146	0.228	0.121	0.058	0.029	0.010
2	9.0	453,173	83,162	0.126	68.9%	0.207	0.182	0.142	0.229	0.128	0.064	0.035	0.014
3	9.0	527,510	98,833	0.137	73.0%	0.195	0.173	0.138	0.229	0.134	0.071	0.041	0.019
4	9.0	621,171	119,601	0.148	77.0%	0.182	0.165	0.133	0.228	0.139	0.079	0.049	0.026
5	9.0	740,297	147,523	0.161	80.7%	0.169	0.155	0.128	0.225	0.144	0.086	0.058	0.035
6	9.0	893,192	185,605	0.175	84.1%	0.156	0.145	0.121	0.220	0.148	0.094	0.067	0.048
7	9.0	1,091,159	238,286	0.190	87.2%	0.142	0.135	0.114	0.213	0.151	0.101	0.078	0.066
8	9.0	1,349,634	312,195	0.208	89.8%	0.129	0.123	0.107	0.204	0.152	0.107	0.089	0.090
9	10.0	1,898	413,781	0.213	69.1%	0.202	0.182	0.145	0.237	0.131	0.063	0.031	0.010
10	10.0	1,977	473,573	0.222	72.6%	0.191	0.174	0.141	0.236	0.137	0.070	0.037	0.014
11	10.0	2,068	547,988	0.232	76.0%	0.180	0.166	0.136	0.235	0.142	0.077	0.044	0.019
12	10.0	2,171	641,529	0.243	79.3%	0.169	0.158	0.131	0.232	0.147	0.084	0.052	0.028
13	10.0	2,290	760,281	0.254	82.5%	0.158	0.149	0.126	0.228	0.151	0.092	0.061	0.035
14	10.0	2,431	912,485	0.267	85.4%	0.146	0.139	0.119	0.222	0.155	0.099	0.071	0.048
15	10.0	2,599	1,109,379	0.280	88.1%	0.134	0.129	0.112	0.215	0.156	0.106	0.082	0.066
16	10.0	2,804	1,366,336	0.295	90.4%	0.121	0.119	0.104	0.205	0.156	0.112	0.093	0.090
17	11.0	1,657	447,754	0.378	74.2%	0.176	0.169	0.142	0.249	0.147	0.073	0.035	0.010
18	11.0	1,729	507,550	0.383	76.8%	0.168	0.162	0.138	0.246	0.152	0.080	0.042	0.014
19	11.0	1,811	581,709	0.389	79.5%	0.159	0.155	0.133	0.243	0.156	0.087	0.049	0.019
20	11.0	1,905	674,658	0.394	82.1%	0.150	0.147	0.128	0.238	0.160	0.094	0.058	0.026
21	11.0	2,014	792,399	0.400	84.6%	0.141	0.139	0.122	0.233	0.163	0.101	0.067	0.035
22	11.0	2,143	943,086	0.405	86.9%	0.131	0.130	0.116	0.225	0.164	0.108	0.077	0.048
23	11.0	2,298	1,137,868	0.412	89.1%	0.121	0.121	0.109	0.217	0.164	0.114	0.088	0.066
24	11.0	2,488	1,392,048	0.418	91.1%	0.111	0.112	0.101	0.206	0.162	0.119	0.099	0.090
PRIOR (AVG):		651,163	261,632	0.289	66.9%	0.159	0.146	0.124	0.227	0.152	0.093	0.063	0.041
ACTUAL:	10.42	2.12	709,952	0.303	81.9%	0.157	0.150	0.127	0.232	0.154	0.091	0.059	0.030

PRIOR (averages) for columns 7 thru 14 are weighted by probabilities in column (5).

(3) = $(EXP(\mu+0.5\sigma^2) * NORMSDIST((LN(\$10M)-\mu)/\sigma-\sigma)+\$10M*(1-NORMSDIST((LN(\$10M)-\mu)/\sigma)))-(EXP(\mu+0.5\sigma^2)*NORMSDIST((LN(\$100K)-\mu)/\sigma-\sigma)+\$100K*(1-NORMSDIST((LN(\$100K)-\mu)/\sigma))))/(5)$

(4) = $EXP(\mu+0.5\sigma^2)*NORMSDIST((LN(\$10M)-m)/\sigma-s)+\$10M*(1-NORMSDIST((LN(\$10M)-m)/\sigma))$

(5) = $1 - NORMSDIST((LN(\$100K) - \mu)/\sigma)$

(7) = $(NORMSDIST((LN(\$135k) - \mu) / \sigma) - NORMSDIST((LN(\$100K) - \mu) / \sigma)) / (5)$

4. Assumptions and “data”

The following discussion assumes an identical severity distribution for each risk. This is an appropriate assumption for testing the efficacy of the model. The underlying ground-up loss ratio for each risk will vary. Since each risk is of identical severity, the impact of the loss ratio is felt through the frequency of claims.

In the example that follows, “data” will be the result of one simulation, to which we fit the model. We simulate an emergence pattern, a ground-up loss ratio, and the size of claim(s). The Bayesian approach will simultaneously fit both the severity distribution and the ground-up loss ratio. We examine 10 years of losses with the following additional assumptions: a \$6M premium is written each year, and the loss ratio is drawn from a lognormal distribution with mean 100% and coefficient of variation 20.2% ($\sigma = .2$).² Size of loss is drawn from a lognormal distribution with parameters as shown in Exhibit 2.1, $\mu = 10.42$ and $\sigma = 2.12$. Maximum loss size (the policy limit) is \$10M. The expected annual number of ground-up claims over all risks is 22.8 (\$6M divided by an expected severity of \$262,713). However, for each risk, the expected frequency varies as a function of the underlying loss ratio. We assume that the reinsurer receives only those claims above \$100K (due to truncation: $T = \$100K$). Given this lognormal assumption, roughly 6.9 claims per year are in excess of \$100K (69 claims over 10 years). Based on simulated emergence patterns, roughly 78% of these claims are reported. The percentage reported declines as the size of loss increases. Frequency is simulated from a negative binomial distribution with a variance-to-mean ratio of $1 + cL$, with $c = .01$ ³ and where L is the expected number of ground-up claims.

Exhibit 4.1 displays the expected emergence pattern for each of the layers (bands) of loss. The emergence patterns are unknown parameters and are

therefore estimated. This paper suggests an approach to estimating these patterns, but that is not a goal of the model. Losses are reported more quickly for lower layers than for higher layers. The variation around the expected pattern for each age is modeled as lognormal with a sigma of 50%. After emergence patterns are simulated for each age, they are scaled so that the patterns add up to unity for each accident period. Exhibit 4.2 displays the reported claims, ultimate claims, estimated emergence pattern, and actual emergence pattern for the \$3M × \$2M layer. Since each year has identical exposure, the incremental claim count pattern can be calculated simply as the incremental number of claims reported in the lag period (the development period) divided by the years of data for that age. For example, in the triangle shown in Exhibit 4.2, one claim is reported from age 1 to age 2 (over accident periods 2007 through 2015). This results in .111 claims expected to emerge from age 1 to age 2 (1 claim over 9 years). For 10 years of data, 5 claims have been reported to date, 7 ultimate claims have been simulated, and we estimate 2.46 claims incurred but not reported (IBNR). The estimated percentage reported for the \$3M × \$2M layer is 67.1% ($5/[5 + 2.46]$). The 67.1% will be the \$3M × \$2M contribution to adjust the 24 severity distributions to a reported level. Exhibit 4.3 displays the simulated loss ratio, reported claims, IBNR claims, ultimate claims, and estimated emergence pattern (percentage reported) for all the layers. The simulated loss ratio is an unknown parameter, which the Bayesian framework will estimate. The reported claims and the ultimate claims are totals over 10 years. Column (4), “Ultimate,” represents an “experience” or “emergence” estimate. The Bayesian estimate will be compared with this estimate.

The adjustment of the severity distributions to a reported basis from an ultimate basis is shown in Exhibit 4.4. At the top left of the exhibit is the emergence pattern, or percentage reported, in each layer (see Exhibit 4.1). Examining layer probabilities reveals that adjusting the distributions shifts some probability from the higher-attaching layers to the lower-attaching layers.

²Since the underlying loss ratio is 100%, the premium in this example is equal to the pure premium (i.e., the amount necessary to pay losses).

³In my experience, $c = .01$ is a reasonable estimate of “contagion.” The model is robust to the choice of c .

Exhibit 4.1. Expected emergence patterns

Expected Emergence Patterns										
Layer	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
35,000 x 100,000	0.501	0.177	0.096	0.063	0.045	0.034	0.027	0.022	0.019	0.016
50,000 x 135,000	0.454	0.178	0.103	0.070	0.052	0.040	0.033	0.027	0.023	0.020
65,000 x 185,000	0.437	0.178	0.105	0.072	0.054	0.043	0.035	0.029	0.025	0.022
250,000 x 250,000	0.421	0.177	0.107	0.074	0.056	0.045	0.037	0.031	0.027	0.024
500,000 x 500,000	0.341	0.171	0.114	0.085	0.068	0.057	0.049	0.043	0.038	0.034
1,000,000 x 1,000,000	0.266	0.158	0.117	0.094	0.080	0.069	0.062	0.056	0.051	0.047
3,000,000 x 2,000,000	0.199	0.141	0.115	0.100	0.089	0.081	0.075	0.070	0.066	0.063
5,000,000 x 5,000,000	0.134	0.117	0.107	0.101	0.097	0.094	0.091	0.088	0.086	0.084

Exhibit 4.2. Reported claims, ultimate claims, estimated emergence pattern, and actual emergence pattern, \$3M x \$2M layer

cumulative to date: 3,000,000 x 2,000,000												(1)	(2)	(3)	(4)
Year	OLP	age 1	age 2	age 3	age 4	age 5	age 6	age 7	age 8	age 9	age 10	Reported to date	IBNR Claims	Expected Ultimate	Simulated Ultimate
2007	6,000,000	0	0	0	0	0	0	0	0	0	0	0	0.000	0.000	0
2008	6,000,000	0	1	1	1	1	2	2	2	2		2	0.000	2.000	2
2009	6,000,000	0	0	0	0	1	1	1	1			1	0.000	1.000	1
2010	6,000,000	0	0	0	1	1	1	1				1	0.000	1.000	1
2011	6,000,000	0	0	0	0	0	0					0	0.000	0.000	0
2012	6,000,000	0	0	0	0	0						0	0.200	0.200	0
2013	6,000,000	0	0	1	1							1	0.367	1.367	1
2014	6,000,000	0	0	0								0	0.510	0.510	0
2015	6,000,000	0	0									0	0.635	0.635	1
2016	6,000,000	0										0	0.746	0.746	1
Expected Incremental Emergence:		0.111	0.125	0.143	0.167	0.200	0.000	0.000	0.000	0.000		5	2.456	7.456	7
Expected Cumulative Emergence:		0.746	0.635	0.510	0.367	0.200	0.000	0.000	0.000	0.000					
												Estimated Emerged Claims: Total (1)/Total (3)		0.671	
												Actual Emerged Claims: Total (1)/Total (4)		0.714	

Exhibit 4.3. Simulated loss ratio, reported claims, IBNR claims, ultimate claims, and estimated emergence pattern, all layers

Simulated Loss Ratio: 133.4%						
Layer	(1) Reported	(2) IBNR	(3) Ultimate, or "Experience Estimate"	(4) (2)+(3)	(5) Percentage Reported, or "Emergence Pattern"	(6) Underlying Frequency[Parameter] "ACTUAL"
35,000 x 100,000	17	1.98	18.98	18.98	89.6%	14.48
50,000 x 135,000	13	1.76	14.76	14.76	88.1%	13.82
65,000 x 185,000	10	1.57	11.57	11.57	86.4%	11.75
250,000 x 250,000	15	2.79	17.79	17.79	84.3%	21.46
500,000 x 500,000	9	2.08	11.08	11.08	81.2%	14.20
1,000,000 x 1,000,000	11	4.01	15.01	15.01	73.3%	8.45
3,000,000 x 2,000,000	5	2.46	7.46	7.46	67.1%	5.41
5,000,000 x 5,000,000	1	0.67	1.67	1.67	60.0%	2.78
Total:	81	17.32	98.32	98.32	82.4%	92.34

Exhibit 4.4. Ultimate and reported probabilities

		% Reported																
		88.1%	86.4%	84.3%	81.2%	73.3%	67.1%	60.0%										
(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	
	50,000 x 135,000	65,000 x 185,000	250,000 x 250,000	500,000 x 500,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000	Normalizing Constant	35,000 x 100,000	50,000 x 135,000	65,000 x 185,000	250,000 x 250,000	500,000 x 500,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000	Total Probability	
				$[F(c_{j;\theta}) - F(c_{j-1;\theta})]$					% Reported x $[F(c_{j;\theta}) - F(c_{j-1;\theta})]/\text{Normalizing Constant}$									
0.219	0.189	0.146	0.228	0.121	0.058	0.029	0.010	0.847	0.231	0.197	0.149	0.227	0.116	0.050	0.023	0.007	1.000	
0.207	0.182	0.142	0.229	0.128	0.064	0.035	0.014	0.844	0.220	0.190	0.146	0.229	0.123	0.056	0.027	0.010	1.000	
0.195	0.173	0.138	0.229	0.134	0.071	0.041	0.019	0.839	0.208	0.182	0.142	0.230	0.129	0.062	0.033	0.013	1.000	
0.182	0.165	0.133	0.228	0.139	0.079	0.049	0.026	0.834	0.195	0.174	0.138	0.230	0.136	0.069	0.039	0.018	1.000	
0.169	0.155	0.128	0.225	0.144	0.086	0.058	0.035	0.828	0.183	0.165	0.133	0.229	0.142	0.076	0.047	0.025	1.000	
0.156	0.145	0.121	0.220	0.148	0.094	0.067	0.048	0.821	0.170	0.156	0.128	0.226	0.147	0.084	0.055	0.035	1.000	
0.142	0.135	0.114	0.213	0.151	0.101	0.078	0.066	0.813	0.157	0.146	0.122	0.221	0.151	0.091	0.064	0.049	1.000	
0.129	0.123	0.107	0.204	0.152	0.107	0.089	0.090	0.803	0.143	0.135	0.115	0.214	0.153	0.098	0.074	0.067	1.000	
0.202	0.182	0.145	0.237	0.131	0.063	0.031	0.010	0.845	0.214	0.189	0.148	0.236	0.125	0.055	0.025	0.007	1.000	
0.191	0.174	0.141	0.236	0.137	0.070	0.037	0.014	0.841	0.204	0.182	0.145	0.237	0.132	0.061	0.030	0.010	1.000	
0.180	0.166	0.136	0.235	0.142	0.077	0.044	0.019	0.837	0.193	0.175	0.141	0.237	0.138	0.067	0.035	0.013	1.000	
0.169	0.158	0.131	0.232	0.147	0.084	0.052	0.026	0.832	0.182	0.167	0.136	0.235	0.144	0.074	0.042	0.019	1.000	
0.158	0.149	0.126	0.228	0.151	0.092	0.061	0.035	0.826	0.171	0.159	0.132	0.233	0.149	0.081	0.050	0.026	1.000	
0.146	0.139	0.119	0.222	0.155	0.099	0.071	0.048	0.819	0.160	0.150	0.126	0.229	0.153	0.089	0.058	0.035	1.000	
0.134	0.129	0.112	0.215	0.156	0.106	0.082	0.066	0.811	0.148	0.141	0.120	0.223	0.156	0.096	0.068	0.049	1.000	
0.121	0.119	0.104	0.205	0.156	0.112	0.093	0.090	0.801	0.136	0.131	0.113	0.216	0.158	0.102	0.078	0.067	1.000	
0.176	0.169	0.142	0.249	0.147	0.073	0.035	0.010	0.841	0.188	0.177	0.146	0.249	0.142	0.063	0.028	0.007	1.000	
0.168	0.162	0.138	0.246	0.152	0.080	0.042	0.014	0.837	0.179	0.170	0.142	0.248	0.147	0.070	0.033	0.010	1.000	
0.159	0.155	0.133	0.243	0.156	0.087	0.049	0.019	0.833	0.171	0.164	0.138	0.246	0.152	0.076	0.040	0.013	1.000	
0.150	0.147	0.128	0.238	0.160	0.094	0.058	0.026	0.828	0.162	0.156	0.133	0.243	0.157	0.083	0.047	0.019	1.000	
0.141	0.139	0.122	0.233	0.163	0.101	0.067	0.035	0.822	0.153	0.149	0.128	0.239	0.161	0.090	0.055	0.026	1.000	
0.131	0.130	0.116	0.225	0.164	0.108	0.077	0.048	0.815	0.144	0.141	0.123	0.233	0.163	0.097	0.064	0.035	1.000	
0.121	0.121	0.109	0.217	0.164	0.114	0.088	0.066	0.808	0.135	0.132	0.116	0.226	0.165	0.103	0.073	0.049	1.000	
0.111	0.112	0.101	0.206	0.162	0.119	0.099	0.090	0.798	0.125	0.123	0.109	0.217	0.165	0.109	0.083	0.068	1.000	

5. MLE estimate for the ground-up loss ratio

For each curve, we can calculate the probability of a claim's exceeding the truncation ($T = \$100K$), along with the ground-up and excess claim severities. Given a ground-up loss ratio, it is possible to estimate the ground-up claim count and the number of claims in excess of a given truncation point. Recall that we started our example (or analysis) with an estimate of the total on-level premium for all years. Applying a loss ratio produces total loss, total ground-up claim counts, and the number of claims above the truncation.

Also for each curve, given a loss ratio, we can estimate the expected number of claims in each layer. Applying the emergence pattern (Exhibit 4.3) to these counts gives us the expected number of reported claims in each layer. We can assume a frequency distribution for each layer and solve for the loss ratio that maximizes the likelihood of observing the reported claim counts in each layer, given the expected number of reported claims in each layer. Under this framework, we can view the loss ratio as a parameter of the frequency distribution. For the ground-up claim counts, we assume a negative binomial distribution with a variance-to-mean (VTM) ratio of $1 + cL$, with $c = .01$ and where L is the number of ground-up claims as a function of the MLE. The model is robust to the choice of c . The maximum likelihood function is the product of eight frequency distributions normalized for probability in each band; the probability in each band can be found in Exhibit 4.3, columns (16) through (23). Normalizing assures us that the total probability summed over the eight frequency distributions is 100%. Normalizing doesn't affect the MLE for each curve, but it does affect the relative likelihood between curves, which in turn influences the ground-up loss ratio (i.e., the mean of the posterior loss ratio distribution) weighted over all curves (more in section 7).

The VTM ratio generally declines as we move up into successively higher layers. Patrik and Mashitz

(1990) describe expressions that modify the ground-up VTM ratio as we move into higher layers. We can view the dispersion in the ground-up claim count as the result of independent Poisson random variables, each with a different underlying mean. We assume the mean is gamma distributed, with parameters A and B calibrated to produce the observed ground-up claim count and VTM ratio. Expressions are as follows:

$$\text{Ground-up claim counts} = A/B$$

$$\text{VTM ratio} = (1 + B)/B$$

$$\text{Ground-up parameters are } B = 1/(\text{VTM} - 1)$$

$$A = (\text{Ground-up counts}) \times (\text{VTM} - 1)$$

As we move into higher layers, Patrik and Mashitz proceed to show that the scale parameter $B \rightarrow B/$ (probability of attaching). The expression for the VTM above gets updated as B gets transformed. The probability of attaching is calculated directly from the prior severity distribution.

Exhibit 5.1 takes us through the calculations for curve #1 (see Exhibit 2.2 for parameters). The MLE estimate shown on line (2), 97.8%, minimizes the negative log-likelihood function on line (22). The negative binomial density function on line (19) evaluates the probability of observing the actual number of reported claims, given the number of expected reported claims and the VTM ratio. The expected number of claims is a function of the ground-up loss ratio. An identical calculation is performed for each curve.

6. Posterior weights for the severity distribution

The severity likelihood function depends on two expressions. The first expression is equation (4) above, the adjusted likelihood function. The second expression will be described below.

It is quite common that only claims in excess of some given truncation are submitted to a reinsurer. In

Exhibit 5.1. MLE estimate of the ground-up loss ratio for curve #1

Estimated on-level premium ALL YEARS	60,000,000	(1)																
Ground up Loss Ratio (MLE):	97.8%	(2)																
Excess Loss Percentage:	64.7%	(3)																
Excess Loss Rate :	63.3%	(4)	(2)x(3)															
Excess Loss:	37,958,445	(5)	(4)x(1)															
Excess Severity:	393,605	(6)																
Expected Excess Claims:	96.44	(7)	(5)/(6)															
Probability of attaching:	11.7%	(8)																
Expected Ground Up Claims (over 10 year period):	824.6	(9)	(7)/(8)															
Assumed Contagion:	1.0%	(10)																
Assumed VTM Ground Up:	1.825	(11)	1+(10)x(9)/10															
B (scale parameter gamma distribution):	1.213	(12)	1/[1(1)-1]															
A (scale parameter gamma distribution):	100	(13)	(9)x(12)/10															
Curve #1_Probability of attaching [q]:																		
B' = B/q	10.37	(14)																
VTM:	1.096	(15)																
Layer Probability:	0.219	(16)																
Emergence Pattern [% Reported]:	89.6%	(17)																
Expected Reported:	18.9	(18)																
Actual Reported Claims (simulated):	17.0	(19)																
Negative Binomial pdf:	0.0843	(20)																
Normalizing constants (adjusted layer probs):	0.231	(21)																
log likelihood :	-3.94	(22)																
Negative loglikelihood:	43.1159	(23)																

many instances, a severity distribution that appears to be a good candidate (fit) from an excess perspective may not be a good candidate from a ground-up perspective. Why does this matter? Because in many instances, curves that seem to fit well just above the truncation point do not adequately capture the tail of the underlying severity distribution. A curve that fits well on an excess basis might imply a loss ratio significantly or implausibly different than expected. Therefore, we examine the MLE for the ground-up loss ratio. If the MLE is significantly different than the expected loss ratio (ELR), the curve is not a good candidate for the underlying severity distribution. The ELR can be the result of a detailed analysis for the individual risk, or it can be the result of an industry study. Given estimates of the ELR, uncertainty surrounding the ELR, and the process risk underlying the frequency and severity assumptions, we can calculate the likelihood of observing the MLE. This is the second expression mentioned above.

For our case study, we assume that ground-up triangles are not available, and we utilize a mean industry loss ratio of 100%, with individual risks varying around this mean loss ratio. We will assume a log-normal distribution with $\sigma = 40\%$. Our estimate of the dispersion around the ELR is considerably wider than the underlying 20%. Again, we will assume the MLE is lognormally distributed with mean 100% and $\sigma \approx (.4^2 + \text{process variance})^{.5}$, where the process variance is calculated assuming the underlying ELR of 100%. For the likelihood, we calculate the density of the MLE using the parameters above. For this calculation, we use Excel's built-in function LOGNORM.DIST. Exhibit 6.1 takes us through the calculation for curve #1. The calculation assumes that the premium charged for excess losses mirrors the ELR. A discussion of each term used to calculate the process variance is beyond the scope of this paper. However, analytical expressions are defined in Exhibit 6.1. We use simulation to validate the process variance.

Exhibit 6.2 displays the posterior weight for each severity distribution. Columns (26) and (27)

display ground-up weights and excess weights. Equation (5) states the posterior weight as proportional to the product of the probability of the data given the model, also called the *likelihood*, and the prior probability. Since prior weights are equal, posterior weights are proportional to the likelihood function.

$$\text{Posterior}\{\text{model}|\text{data}\} \propto \text{Probability}\{\text{data}|\text{model}\} \\ \times \text{Prior}\{\text{model}\} \quad (5)$$

Ground-up weights are necessary to calculate ground-up statistics such as ground-up severity, while excess weights are used to calculate excess statistics. For informational purposes, the MLE for each curve is shown in column (28). One can see that the likelihood of the MLE, and thus the weights, declines as the MLE deviates from the ELR. The likelihood function in column (24) may potentially create numerical underflow problems. One possible way of handling this is by taking logarithms of each term in the likelihood function, adding the absolute value of the smallest log-likelihood value to each log-likelihood value, exponentiating the results, and then normalizing.

Exhibit 6.3 displays the posterior severity distribution. To create Exhibit 6.3, we append columns (26) and (27) from Exhibit 6.2 to Exhibit 2.2. The last row of columns (7) through (14) of Exhibit 6.3 represents the posterior severity distribution, conditioned on claims greater than the truncation. It is important to note that given the statistics in the last row of columns (3) through (6), and an estimate for the ground-up loss ratio, claim count estimates for the eight bands can be produced. We will refer to these claim count estimates as Bayesian claim count estimates, or Bayesian estimates. It is also instructive to note that curve #12 produces the greatest likelihood (posterior weight) of generating the observed data. Looking back at Exhibit 2.2, we can see that compared with the other curves, the parameters and statistics for curve #12 are relatively close to the underlying, or ACTUAL.

Exhibit 6.1. Likelihood of the MLE, given the ELR and total variance for curve #1

Estimated Annual on-level premium:	6,000,000	(1)
Expected Loss ratio:	100.0%	(2) See Exhibit 2, Column (6)
Excess Loss Percentage:	64.7%	(3); (1)x(2)x(3)
Excess Loss E[S]:	3,881,160	(4); (1)x(2)x(3)
Excess Severity E[X]:	393,605	(5); See Exhibit 2, Column (3)
Expected Excess Claim Counts E[N]:	9.861	(6); (4)/(5)
Var[N]:	10.833	(7); = (6) x [1 + c x(6)], c=.01
E[X X>T]:	493,605	(8); (5)+T
E[X ²]:	1.34E+11	(9); = (EXP(2*μ+0.5*(2*σ) ²)*NORMSDIST((LN(PL)-μ)/(σ-2*σ)))/(NORMSDIST((LN(PL)-μ)/(σ-2*σ))*(1-LOGNORM.DIST(PL,μ,σ,1)))
E[X X>T]²:	2.44E+11	(10); (8) ²
F ₂ (T) [second moment distribution function]:	5.93E+08	(11); = EXP(2*m+(4*σ ² /2)*NORMSDIST((LN(T)-μ)/(σ-2*σ))
Probability of attaching:	11.7%	(12); See Exhibit 2, Column (5)
E[X ² X>T]:	1.14E+12	(13); [(9)-(11)]/(12)
Var(X):	8.95E+11	(14); (13)-(10)
Process variance, Var(S)	1.15E+13	(15); [(6)x(14)+(7)x(10)] ⁵
Experience period in years:	10	(16)
Normalizing constant:	0.847	(17); See Exhibit 5, Column (15)
Standard Deviation of the Excess Loss Ratio:	30.0%	(18); SQRT((15)/[(16)x(17)]) / [(4)/(2)]
σ ₁ :	29.3%	(19); SQRT(LN((18)/2*(1 ² +1))
Parameter Uncertainty, σ ₂ :	40.0%	(20)
Total Variance:	24.6%	(21); (19) ² + (20) ²
Standard Deviation, σ _T :	49.6%	(22); SQRT(21)
μ:	-12.3%	(23); LN(1)-0.5*(22)/2
MLE:	0.978	(23)
Likelihood MLE _{Expected LR and total variance} :	0.806	(24); LOGNORM.DIST(MLE,μ,σ _T ,0)

$$Var(S) = E[N]Var(X) + Var(N)E[X]^2$$

- X: Random variable representing size of loss
- N: Random variable representing number of excess claims.
- S: Random variable representing total loss
- T: Truncation, \$100K
- PL: Maximum Policy Limit, \$10M

Exhibit 6.2. Posterior weights for severity distributions

		See Exhibit 4B											
		17	13	10	11	9	5	1					
Reported Claims		21.0%	16.0%	12.3%	11.1%	13.6%	6.2%	1.2%					
% Distribution													
See Exh 2		See Exhibit 5											
(5)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)
Probability	% Reported x [F(C_i;θ) - F(C_{i-1};θ)]/Normalizing Constant												
>\$100K	35,000 x	50,000 x	65,000 x	250,000 x	500,000 x	1,000,000 x	3,000,000 x	5,000,000 x	Group Likelihood Function*	Likelihood of MLE given ELR & Tot Var	Ground Up Posterior Wgts (24)x(25)/(5) #	Excess Posterior Wgts (24)x(25) #	MLE
0.117	100,000	135,000	185,000	250,000	500,000	1,000,000	2,000,000	5,000,000	9.57E-72	0.806	0.0081	0.0041	97.8%
0.126	0.231	0.197	0.149	0.227	0.116	0.050	0.023	0.007	4.56E-71	0.702	0.0312	0.0170	106.0%
0.137	0.220	0.190	0.146	0.229	0.123	0.056	0.027	0.010	1.57E-70	0.576	0.0813	0.0480	116.8%
0.148	0.208	0.182	0.142	0.230	0.129	0.062	0.033	0.013	3.68E-70	0.437	0.1338	0.0856	131.0%
0.161	0.195	0.174	0.138	0.230	0.136	0.069	0.039	0.018	5.54E-70	0.300	0.1274	0.0885	149.6%
0.175	0.183	0.165	0.133	0.229	0.142	0.076	0.047	0.025	4.89E-70	0.183	0.0629	0.0475	174.2%
0.190	0.170	0.156	0.128	0.226	0.147	0.084	0.055	0.035	2.25E-70	0.096	0.0140	0.0115	206.9%
0.208	0.157	0.146	0.122	0.221	0.151	0.091	0.064	0.049	4.58E-71	0.043	0.0012	0.0010	250.5%
0.213	0.143	0.135	0.115	0.214	0.153	0.098	0.074	0.067	2.31E-71	0.835	0.0112	0.0103	96.4%
0.222	0.214	0.189	0.148	0.236	0.125	0.055	0.025	0.007	8.65E-71	0.720	0.0345	0.0331	105.3%
0.232	0.204	0.182	0.145	0.237	0.132	0.061	0.030	0.010	2.36E-70	0.583	0.0731	0.0733	116.8%
0.243	0.193	0.175	0.141	0.237	0.138	0.067	0.035	0.013	4.49E-70	0.436	0.0992	0.1040	131.5%
0.254	0.182	0.167	0.136	0.235	0.144	0.074	0.042	0.019	5.54E-70	0.295	0.0792	0.0870	150.7%
0.267	0.171	0.159	0.132	0.233	0.149	0.081	0.050	0.026	4.09E-70	0.177	0.0335	0.0386	175.6%
0.280	0.160	0.150	0.126	0.229	0.153	0.089	0.058	0.035	1.61E-70	0.093	0.0065	0.0079	208.6%
0.295	0.148	0.141	0.120	0.223	0.156	0.096	0.068	0.049	2.87E-71	0.041	0.0005	0.0006	252.4%
0.378	0.136	0.131	0.113	0.216	0.158	0.102	0.078	0.067	5.45E-71	0.838	0.0149	0.0243	97.4%
0.383	0.188	0.177	0.146	0.249	0.142	0.063	0.028	0.007	1.45E-70	0.712	0.0333	0.0550	106.9%
0.389	0.179	0.170	0.142	0.248	0.147	0.070	0.033	0.010	2.90E-70	0.568	0.0523	0.0877	118.9%
0.394	0.171	0.164	0.138	0.246	0.152	0.076	0.040	0.013	4.13E-70	0.418	0.0541	0.0920	134.1%
0.400	0.162	0.156	0.133	0.243	0.157	0.083	0.047	0.019	3.95E-70	0.279	0.0340	0.0586	153.6%
0.405	0.153	0.149	0.128	0.239	0.161	0.090	0.055	0.026	2.32E-70	0.166	0.0117	0.0205	178.8%
0.412	0.144	0.141	0.123	0.233	0.163	0.097	0.064	0.035	7.57E-71	0.086	0.0020	0.0035	211.9%
0.418	0.135	0.132	0.116	0.226	0.165	0.103	0.073	0.049	1.17E-71	0.038	0.0001	0.0002	252.7%
0.418	0.125	0.123	0.109	0.217	0.165	0.109	0.083	0.068	Total:				1.0000
									1.0000				1.0000

* Curve #1: 9.57E-72 = (.231^17)(.197^13)(.149^10)(.227^15)(.116^9)(.050^11)(.023^5)(.007^1)

#: Weights are normalized to sum to unity.

Exhibit 6.3. Posterior severity distribution

	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(26)	(27)
Curve	Excess Severity	Severity Ground Up	Probability >\$100K	% Loss>T [3]X[5][4]	35,000 x 100,000	50,000 x 135,000	65,000 x 185,000	250,000 x 250,000	500,000 x 500,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000	Ground Up Posterior Wgts	Excess Posterior Wgts
					F(C _i ;θ) - F(C _{i+1} ;θ)									
1	393,605	71,166	0.117	64.7%	0.219	0.189	0.146	0.228	0.121	0.058	0.029	0.010	0.0081	0.0041
2	453,173	83,162	0.126	68.9%	0.207	0.182	0.142	0.229	0.128	0.064	0.035	0.014	0.0312	0.0170
3	527,510	98,883	0.137	73.0%	0.195	0.173	0.138	0.229	0.134	0.071	0.041	0.019	0.0813	0.0480
4	621,171	119,601	0.148	77.0%	0.182	0.165	0.133	0.228	0.139	0.079	0.049	0.026	0.1338	0.0856
5	740,297	147,523	0.161	80.7%	0.169	0.155	0.128	0.225	0.144	0.086	0.058	0.035	0.1274	0.0885
6	893,192	185,605	0.175	84.1%	0.156	0.145	0.121	0.220	0.148	0.094	0.067	0.048	0.0629	0.0475
7	1,091,159	238,266	0.190	87.2%	0.142	0.135	0.114	0.213	0.151	0.101	0.078	0.066	0.0140	0.0115
8	1,349,634	312,195	0.208	89.8%	0.129	0.123	0.107	0.204	0.152	0.107	0.089	0.090	0.0012	0.0010
9	413,781	127,332	0.213	69.1%	0.202	0.182	0.145	0.237	0.131	0.063	0.031	0.010	0.0112	0.0103
10	473,573	144,931	0.222	72.6%	0.191	0.174	0.141	0.236	0.137	0.070	0.037	0.014	0.0345	0.0331
11	547,988	167,404	0.232	76.0%	0.180	0.166	0.136	0.235	0.142	0.077	0.044	0.019	0.0731	0.0733
12	641,529	196,470	0.243	79.3%	0.169	0.158	0.131	0.232	0.147	0.084	0.052	0.026	0.0992	0.1040
13	760,281	234,553	0.254	82.5%	0.158	0.149	0.126	0.228	0.151	0.092	0.061	0.035	0.0792	0.0870
14	912,485	285,091	0.267	85.4%	0.146	0.139	0.119	0.222	0.155	0.099	0.071	0.048	0.0335	0.0386
15	1,109,379	353,015	0.280	88.1%	0.134	0.129	0.112	0.215	0.166	0.106	0.082	0.066	0.0065	0.0079
16	1,366,336	445,459	0.295	90.4%	0.121	0.119	0.104	0.205	0.156	0.112	0.093	0.090	0.0005	0.0006
17	447,754	228,304	0.378	74.2%	0.176	0.169	0.142	0.249	0.147	0.073	0.035	0.010	0.0149	0.0243
18	507,550	253,208	0.383	76.8%	0.168	0.162	0.138	0.246	0.152	0.080	0.042	0.014	0.0333	0.0550
19	581,709	284,379	0.389	79.5%	0.159	0.155	0.133	0.243	0.156	0.087	0.049	0.019	0.0523	0.0877
20	674,658	323,842	0.394	82.1%	0.150	0.147	0.128	0.238	0.160	0.094	0.058	0.026	0.0541	0.0920
21	792,399	374,379	0.400	84.6%	0.141	0.139	0.122	0.233	0.163	0.101	0.067	0.035	0.0340	0.0586
22	943,086	439,834	0.405	86.9%	0.131	0.130	0.116	0.225	0.164	0.108	0.077	0.048	0.0117	0.0205
23	1,137,868	525,558	0.412	89.1%	0.121	0.121	0.109	0.217	0.164	0.114	0.088	0.066	0.0020	0.0035
24	1,392,048	639,044	0.418	91.1%	0.111	0.112	0.101	0.206	0.162	0.119	0.099	0.090	0.0001	0.0002
PRIOR:	651,163	261,632	0.269	66.9%	0.154	0.146	0.124	0.227	0.152	0.093	0.063	0.041		
ACTUAL:	709,952	262,713	0.303	81.9%	0.157	0.150	0.127	0.232	0.154	0.091	0.059	0.030		
POSTERIOR:	689,547	193,398	0.232	80.2%	0.166	0.156	0.130	0.232	0.149	0.086	0.054	0.028		

PRIOR (averages) for columns 7 thru 14 are weighted by probabilities in column (5).
 POSTERIOR (averages) for columns 7 thru 14 are weighted by column (27).

7. The ground-up loss ratio (mean of the posterior loss ratio distribution)

As described in section 5, an MLE estimate for the ground-up loss ratio is derived for each curve. We can view these estimates as a set of priors, all initially receiving equal weight. As with the severity distribution, the likelihood of each MLE estimate depends on two expressions. The first expression can be found on line (21) of Exhibit 5.1. However, instead of transforming the likelihood to a negative log-likelihood, we will rely directly upon the likelihood function. To convert the 43.1159 negative log-likelihood found in Exhibit 5.1, we change sign and exponentiate: $e^{-43.1159} = 1.884E - 19$. The second expression is described below.

The eight frequency distributions, one for each layer, are a function of the MLE. As with a severity distribution, a frequency distribution that may be a good candidate for the excess frequency may not be a good candidate for the underlying ground-up frequency. Again, if the MLE is significantly different than the ELR, the MLE is not a good candidate for the underlying frequency distribution. Whereas process risk should be recognized when identifying which severity curves could potentially produce the observed data, the expected process risk underlying each curve should not be reflected when calculating the likelihood of the MLE given the ELR and the surrounding uncertainty (i.e., the parameter uncertainty). Our second expression reflects only this parameter uncertainty. As described above, we assume an industry loss ratio of 100% with individual risks lognormally distributed and $\sigma = 40\%$.

Exhibit 7.1 displays the calculation of the ground-up loss ratio, which is a weighted average over all MLEs. In this example, the underlying, or ACTUAL, loss ratio is 133.4%, while the posterior loss ratio is 146.1%.

Exhibit 7.2 ties together the posterior severity distribution in Exhibit 6.3 with the mean of the posterior loss ratio distribution in Exhibit 7.1 to arrive at Bayesian claim count estimates. We can see that the Bayesian count better mirrors ACTUAL (the under-

lying frequency/parameter) than does the experience estimate. An experience approach is more responsive to the reported claims in the layer and is thus more responsive to random variation. Exhibit 4.3 shows that only 1 claim has been reported at the 5×5 layer, while the true underlying frequency is 2.78 claims. Thus, experience rating is too responsive to the reported claim counts. Where reported claims are higher than the underlying expected reported claims, experience rating tends to produce an estimate that is too high. Where reported claims are lower than the underlying expected reported claims, experience rating tends to produce an estimate that is too low. A Bayesian approach takes into account experience in all layers and does not overreact to random variation to the degree that experience rating does.

8. Bayesian estimates versus MLE and experience estimates

Exhibit 8.1 displays an extract of Bayesian, MLE, and experience fits (estimates) for 200 sets of simulated data; 200 is enough simulations to be an effective illustration of expected results. The MLE approach utilizes parameters as estimated by the methods discussed above: specifically, the adjusted likelihood function, as in equation (3), and the MLE estimate for the ground-up loss ratio, as shown in section 5. The calculations for simulation 1 can be found in appendix A. We use estimated emergence patterns for the experience estimates.⁴ For ease of exposition, only the top three layers for each approach are shown. Columns (1) through (3) are Bayesian claim count estimates, columns (4) through (6) are MLE estimates, and columns (6) through (9) are experience estimates. Each set of claim count estimates relies on

⁴For layers without any claims reported over the entire experience period, we use the emergence pattern of the layer immediately below. However, for experience rating purposes, this method would overstate the expected claim counts in higher layers. In these instances, the ratio of probability in the higher layer to that in the lower layer is applied to the expected claim counts in the lower layer. We use the probabilities associated with the posterior severity distribution.

Exhibit 7.1. Mean of the posterior loss ratio distribution

Curve	(1) MLE	(2) Frequency Likelihood Function	(3) Likelihood of MLE given ELR & Parameter Risk	(4) Posterior weights (2)x(3)/Σ(2)x(3)
1	97.8%	1.884E-19	1.009	0.001
2	106.0%	1.552E-18	0.886	0.006
3	116.8%	8.996E-18	0.718	0.027
4	131.0%	3.473E-17	0.519	0.074
5	149.6%	8.316E-17	0.321	0.110
6	174.2%	1.128E-16	0.162	0.075
7	206.9%	7.672E-17	0.063	0.020
8	250.5%	2.216E-17	0.018	0.002
9	96.4%	5.328E-19	1.029	0.002
10	105.3%	3.381E-18	0.897	0.012
11	116.8%	1.533E-17	0.718	0.045
12	131.5%	4.705E-17	0.512	0.099
13	150.7%	9.126E-17	0.313	0.117
14	175.6%	1.024E-16	0.156	0.066
15	208.6%	5.895E-17	0.060	0.015
16	252.4%	1.481E-17	0.017	0.001
17	97.4%	1.579E-18	1.015	0.007
18	106.9%	6.957E-18	0.872	0.025
19	118.9%	2.250E-17	0.687	0.063
20	134.1%	5.070E-17	0.481	0.100
21	153.6%	7.455E-17	0.289	0.089
22	178.8%	6.565E-17	0.142	0.038
23	211.9%	3.088E-17	0.054	0.007
24	255.7%	6.623E-18	0.015	0.000
POSTERIOR LR:	146.1%	=SUMPRODUCT{(1),(4)}		
		*: Likelihood=LOGNORM.DIST(MLE, μ, σ, 0) where: μ = -.08, σ = .40%		

the same estimated emergence pattern. We also show the underlying, Bayesian, and MLE loss ratios. The underlying loss ratio is the simulated loss ratio, while the Bayesian and MLE loss ratios are estimates. The last column, (13), is the underlying frequency for the \$5M × \$5M layer for each simulated set of data. This is the parameter we are attempting to estimate.

The average of the Bayesian loss ratios shown in Exhibit 8.1 is 105.8%, while the underlying parameter for these 200 simulations is 98.6%. Some of this loss ratio bias is due to the choice of prior curves; however, the bias may also “force” expected counts to more closely mirror reported counts. It is difficult to come up with general rules for constructing prior severity distributions. Not only can the choice

of curves reduce or increase bias, but it also affects variance. This trade-off should be tested in the development of prior curves. MLE loss ratios exhibit less bias than Bayesian loss ratios, but their mean square error is significantly greater.

The “actual” number of claims for each layer is shown in column (13). “Actual” represents the underlying parameter, not the actual simulated counts. Simulation 1 represents estimates for the sample data. For each simulation, any deviation from “actual” is due to process risk. The goal of any model is to produce estimates that are both unbiased and have minimum variance. Exhibit 8.1 shows that the experience and MLE estimates tend to have slightly lower bias but much higher variance than the Bayesian estimates.

Exhibit 7.2. Projected claim counts

		See Exhibit 9			See Exhibit 9						See Exh 10				
		(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Excess Severity	709,952	262,713	0.303	Probability >\$100K	% Loss>T $\frac{(3) \times (5)}{(4)}$	# Claims>T $\frac{\$60M \times (15) \times (5)}{(2)}$	35,000 x	50,000 x	65,000 x	250,000 x	500,000 x	1,000,000 x	3,000,000 x	5,000,000 x	
Severity	669,547	193,398	0.232		81.9%	92.3	100,000	135,000	185,000	250,000	500,000	1,000,000	2,000,000	5,000,000	Loss Ratio
ACTUAL:						92.3	0.157	0.150	0.127	0.232	0.154	0.091	0.059	0.030	133.4%
POSTERIOR:					80.2%	105.0	0.166	0.156	0.130	0.232	0.149	0.086	0.054	0.028	146.1%
Claim Counts															
ACTUAL:						92.3	35,000 x	50,000 x	65,000 x	250,000 x	500,000 x	1,000,000 x	3,000,000 x	5,000,000 x	
POSTERIOR/Bayesian:					105.0	105.0	100,000	135,000	185,000	250,000	500,000	1,000,000	2,000,000	5,000,000	
"Experience" Estimate:							14.48	13.82	11.75	21.46	14.20	8.45	5.41	2.78	
							17.42	16.34	13.67	24.37	15.60	9.02	5.64	2.94	
							18.98	14.76	11.57	17.79	11.08	15.01	7.46	1.67	

Exhibit 8.1. Bayesian versus MLE and experience claim count estimates (over 10 years)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Avg of 200 simulations:	6.46	4.05	2.17	6.12	3.71	1.98	6.05	4.22	2.07	98.6%	105.8%	97.0%	2.05
Actual (@100%LR):	6.33	4.06	2.08	6.33	4.06	2.08	6.33	4.06	2.08	100.0%	100.0%	100.0%	2.08
Bias(@sim LR):	0.22	0.05	0.12	-0.12	-0.29	-0.07	-0.19	0.22	0.01	0.0%	7.2%	-1.6%	0.00
Variance:	0.80	0.62	0.57	1.91	1.55	1.91	8.07	4.11	2.81	0.0%	3.0%	8.3%	0.00
MSE (Var+Bias ²):	0.85	0.63	0.58	1.92	1.63	1.91	8.10	4.16	2.81	0.0%	3.5%	8.4%	0.00
Simulation	Bayesian Estimates			MLE Estimates			Experience Estimates			Underlying Loss Ratio	Bayesian Loss Ratio	MLE Loss Ratio	Underlying 5,000,000 x 5,000,000
	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000	1,000,000 x 1,000,000	3,000,000 x 2,000,000	5,000,000 x 5,000,000				
1	9.02	5.64	2.94	8.75	5.93	3.65	15.01	7.46	1.67	133.4%	146.1%	154.0%	2.78
2	5.05	3.35	2.00	4.13	2.93	2.15	3.79	4.35	1.67	81.9%	86.4%	81.7%	1.71
3	8.44	5.82	3.69	8.73	6.38	4.32	8.14	5.61	5.00	114.0%	146.2%	155.3%	2.37
4	7.54	4.46	2.02	7.94	4.02	1.09	10.33	5.52	0.69	103.9%	115.4%	94.2%	2.16
5	4.81	3.44	2.36	5.29	3.49	1.56	2.50	5.22	1.43	63.7%	86.6%	72.0%	1.33
6	6.01	3.94	2.24	6.47	3.51	1.03	8.93	1.25	1.67	88.8%	99.1%	75.0%	1.85
7	5.66	3.26	1.41	5.07	1.66	0.19	5.28	1.43	0.19	88.2%	86.1%	54.3%	1.84
8	4.08	2.56	1.37	3.01	1.63	0.59	1.43	4.76	0.95	67.0%	67.4%	43.4%	1.40
9	6.70	4.02	1.86	7.24	2.66	0.35	6.47	4.29	0.60	82.4%	102.0%	71.0%	1.72
10	8.40	5.36	2.89	8.40	5.65	3.23	9.47	5.87	2.78	126.9%	137.0%	138.7%	2.64
11	7.42	4.71	2.61	6.48	4.71	4.05	8.96	2.78	7.54	139.4%	125.4%	147.3%	2.90
12	7.25	4.95	3.10	6.71	5.05	4.19	9.43	2.86	4.68	99.6%	125.9%	139.9%	2.07
13	7.09	4.09	1.76	6.52	2.23	0.27	7.40	2.00	0.43	93.4%	107.1%	69.0%	1.94
14	5.34	3.03	1.30	4.56	2.13	0.55	7.04	1.43	0.18	90.6%	83.4%	60.1%	1.89
15	7.21	4.67	2.62	7.03	4.75	2.72	9.44	7.19	1.67	97.6%	119.3%	115.5%	2.03

It is worth noting that in some instances, actuaries will price higher layers by applying increased limit factors (ILFs) to lower layers, where there is more experience. In this case, Exhibit 8.1 suggests that pricing would be improved by applying the ILF to the Bayesian estimate for the lower layer. However, deciding which ILF to utilize still requires judgment, while a Bayesian framework is objective in nature.

Exhibits 8.2 and 8.3 are pricing exercises of the Bayesian framework contrasted against the experience and MLE approaches. We assume that an insured will select the lowest of the prices offered in the market. For the 200 simulations, we assume 0% expense and profit loads. For calculating pure premium, the experience approach utilizes experience frequency and the actual underlying severity. For the Bayesian approach, severity is a posterior weighted average of the individual severities for each curve. For the MLE approach, severity can be calculated directly. For contrasting the Bayesian and experience approaches, using actual severity for the experience rate is conservative. The simulations have been sorted in descending order of loss ratio and separated into five quintiles, allowing us to do some testing to determine whether the Bayesian model is picking up differences in loss ratio.

Based on the 200 simulations found in Exhibits 8.2 and 8.3 (showing only the first 10 simulations), if a competitor were to use rates indicated by the experience and MLE approaches, it would write less business at a significantly higher loss ratio. We can also see that the Bayesian approach does a reasonably good job of reflecting loss ratio differences. Recall that we assume an ELR of 100% with a parameter coefficient of variation of 40%. For risks with the highest loss ratios, the top 20%, the model estimate is 130.2%, while the actual is 128.9%. This should alleviate some concern regarding adverse selection. There are 40 simulations in each quintile. The profit or loss is the sum of all simulations in the quintile. The other statistics are averages for the quintile. We can see that the Bayesian loss ratio tracks the underlying loss ratio reasonably well, and that a Bayesian underwriting profit is produced for each quintile.

Exhibit 8.4 is a head-to-head comparison/competition of all three approaches.

9. Summary and conclusions

XOL rating generally requires an exposure curve, a loss ratio, and a credibility procedure to weight experience and exposure estimates. In many instances, at least one and sometimes all three of these parameters are judgmentally selected. Just as experience or MLE estimates can overreact/overfit to random variation in the data, so can actuarial judgment. When we see higher-than-expected loss, we tend to pick conservatively. When we see lower-than-expected loss, we are optimistic. The author believes that relying on judgment may produce results no more accurate, and possibly even less so, than the experience and MLE approaches presented in this paper. Another drawback of approaches that utilize too much judgment is that they are difficult to monitor or recalibrate.

This paper provides a Bayesian framework for XOL rating. There are many possible variations on the approach provided here, each providing possible advantages and/or disadvantages. These variations include the family of prior distributions, how the curves are calibrated, the number of layers utilized, and the introduction of credibility into the procedure. The approach taken here can be viewed as a starting point, or “Barnett approach.”

The author believes that a single set of prior severity distributions allows for the development of a model that could be monitored for performance and recalibrated over time (a technical price can be constructed and compared with actual results to potentially recalibrate the model). Based on the volume of losses simulated in the exhibits above, in my opinion, it is preferable to develop a single set of prior severity distributions that fit well to a wide range of potential losses (severity distributions) and produce results similar to those above.⁵

⁵ In my private work, I have done so. Promulgating prior severity distributions here may be considered anticompetitive.

Exhibit 8.2. Pricing exercise: Bayesian approach versus experience approach

	CV:	0.373	0.816	0.000		Bayesian	Experience	Underlying	
	mean:	2.17	2.07	2.05		3,059,054	3,110,715	3,110,715	
	Variance:	0.571	2.807	0.000		0.217	0.207	0.205	
	MSE:	0.585	2.808	0.000		663,824	642,957	638,392	
						+4.0%	+0.7%	+0.0%	
						5,000,000 xs	5,000,000 xs		
						5,000,000	5,000,000		
								Charged Pure Premium: 61,462,409	
								Expected Profit/(Loss): 11,644,222	
								81.1%	
								+178.8%	
LR Rank	5x5 Bayesian Estimate	5x5 Experience Estimate	5x5 Underlying Parameter	Underlying Loss Ratio	Bayesian Loss Ratio	Severity Bayesian Estimate	Severity Experience Estimate	Profit/(Loss) Bayesian	Profit/(Loss) Experience
1	1.97	3.10	3.36	161.2%	137.9%	2,906,837	3,110,715	(472,481)	0
2	2.25	0.74	3.30	158.5%	141.6%	2,875,399	3,110,715	0	(797,476)
3	3.34	2.00	3.04	146.0%	156.0%	3,112,195	3,110,715	0	(323,681)
4	2.36	1.51	2.97	142.6%	133.5%	2,959,653	3,110,715	0	(455,207)
5	2.18	1.43	2.95	141.8%	123.8%	2,957,507	3,110,715	0	(474,367)
6	2.61	7.54	2.90	139.4%	125.4%	3,122,474	3,110,715	(89,247)	0
7	3.83	6.11	2.90	139.2%	166.2%	3,160,806	3,110,715	308,915	0
8	3.85	3.10	2.87	137.7%	169.8%	3,138,546	3,110,715	0	71,114
9	1.87	1.67	2.85	136.8%	127.1%	2,873,735	3,110,715	0	(367,922)
10	2.16	1.16	2.80	134.3%	133.9%	2,926,714	3,110,715	0	(508,789)
Top 20%	2.57	2.83	2.68	128.9%	130.2%	3,035,309	3,110,715	2,130,374	(9,174,484)
20%-40%	2.36	2.22	2.24	107.7%	114.5%	3,056,768	3,110,715	2,564,464	(7,052,011)
40%-60%	2.13	2.04	2.00	96.1%	103.8%	3,052,608	3,110,715	2,929,837	(6,948,876)
60%-80%	2.02	1.85	1.79	86.2%	97.6%	3,065,362	3,110,715	2,223,023	(5,526,306)
80%-100%	1.76	1.39	1.54	73.8%	82.9%	3,085,222	3,110,715	1,796,524	(5,608,328)

* Combined Ratio assumes 0% expense & profit loads.

Exhibit 8.3. Pricing exercise: Bayesian approach versus MLE approach

	CV: 0.373	0.674	0.000		Bayesian	MLE	Underlying		
mean:	2.17	1.98	2.05	Average Severity:	3,059,054	2,931,609	3,110,715		
				Average Frequency:	0.217	0.198	0.205		
Variance:	0.571	1.906	0.000	Average Pure Premium:	663,824	581,150	638,392		
				[Average PPI]/underlying - 1:	+4.0%	(9.0%)	+0.0%		
MSE:	0.585	1.911	0.000	5,000,000 xs	5,000,000 xs	5,000,000 xs			
					5,000,000	5,000,000			
				Charged Pure Premium:			62,571,875		
				Expected Profit/(Loss):			11,924,739		
				Combined Ratio*:			80.9%		
							43,207,923		
							(33,823,357)		
							+178.3%		
LR Rank	5x5 Bayesian Estimate	5x5 MLE Estimate	5x5 Underlying Parameter	Underlying Loss Ratio	Bayesian Loss Ratio	Severity Bayesian Estimate	Severity MLE Estimate	Profit/(Loss) Bayesian	Profit/(Loss) MLE
1	1.97	3.03	3.36	161.2%	137.9%	2,906,837	3,369,389	(472,481)	0
2	2.25	1.28	3.30	158.5%	141.6%	2,875,399	2,468,898	0	(710,592)
3	3.34	4.34	3.04	146.0%	156.0%	3,112,195	3,350,007	92,238	0
4	2.36	1.90	2.97	142.6%	133.5%	2,959,653	2,820,947	0	(986,575)
5	2.18	1.62	2.95	141.8%	123.8%	2,957,507	2,755,134	0	(472,285)
6	2.61	4.05	2.90	139.4%	125.4%	3,122,474	3,617,545	(89,247)	0
7	3.83	4.69	2.90	139.2%	166.2%	3,160,806	3,306,273	308,915	0
8	3.85	3.92	2.87	137.7%	169.8%	3,138,546	3,091,722	315,747	0
9	1.87	1.36	2.85	136.8%	127.1%	2,873,735	2,724,815	0	(516,372)
10	2.16	2.09	2.80	134.3%	133.9%	2,926,714	2,947,557	0	(252,609)
Top 20%	2.57	2.83	2.68	128.9%	130.2%	3,035,309	3,062,937	1,647,591	(6,962,645)
20%-40%	2.36	2.22	2.24	107.7%	114.5%	3,056,768	2,932,109	2,974,442	(7,573,018)
40%-60%	2.13	2.04	2.00	96.1%	103.8%	3,052,608	2,875,094	3,595,699	(8,299,088)
60%-80%	2.02	1.85	1.79	86.2%	97.6%	3,065,362	2,915,393	1,716,561	(5,147,457)
80%-100%	1.76	1.39	1.54	73.8%	82.9%	3,085,222	2,872,512	1,990,446	(5,841,148)

* Combined Ratio assumes 0% expense & profit loads.

Exhibit 8.4. Pricing exercise: Bayesian approach versus MLE and experience approaches

CV:	0.373	0.674	0.816	0.000
mean:	2.17	1.98	2.067	2.05
Variance:	0.571	1.906	2.807	0.000
MSE:	0.585	1.911	2.808	0.000

	Bayesian	MLE	Experience	Underlying
Average Severity:	3,059,054	2,931,609	3,110,715	3,110,715
Average Frequency:	0.217	0.198	0.207	0.205
Average Pure Premium:	663,824	581,150	642,957	638,392
[Average PP]/underlying - 1:	+4.0%	(9.0%)	+0.7%	+0.0%

	Severity Bayesian Estimate	Severity MLE Estimate	Severity Experience Estimate	Severity Underlying Estimate
5,000,000 xs 5,000,000	5,000,000	5,000,000	5,000,000	5,000,000
5,000,000 xs 5,000,000 xs	5,000,000	5,000,000	5,000,000	5,000,000

	Charged Pure Premium:	Expected Profit/(Loss):	Combined Ratio:
	46,908,110	23,415,587	34,058,646
	11,995,532	(20,421,530)	(21,503,255)
	74.4%	+187.2%	+163.1%

LR Rank	5x5 Bayesian Estimate	5x5 MLE Estimate	5x5 Experience Estimate	5x5 Underlying Parameter	Underlying Loss Ratio	Bayesian Loss Ratio	MLE Loss Ratio
1	1.97	3.03	3.10	3.36	161.2%	137.9%	162.9%
2	2.25	1.28	0.74	3.30	158.5%	141.6%	121.3%
3	3.34	4.34	2.00	3.04	146.0%	156.0%	169.3%
4	2.36	1.90	1.51	2.97	142.6%	133.5%	120.1%
5	2.18	1.62	1.43	2.95	141.8%	123.8%	109.8%
6	2.61	4.05	7.54	2.90	139.4%	125.4%	147.3%
7	3.83	4.69	6.11	2.90	139.2%	166.2%	179.3%
8	3.85	3.92	3.10	2.87	137.7%	169.8%	171.8%
9	1.87	1.36	1.67	2.85	136.8%	127.1%	109.9%
10	2.16	2.09	1.16	2.80	134.3%	133.9%	127.1%

	Severity Bayesian Estimate	Severity MLE Estimate	Severity Experience Estimate	Severity Underlying Estimate	Profit/(Loss) Bayesian	Profit/(Loss) MLE	Profit/(Loss) Experience
Top 20%	3,035,309	3,062,937	3,110,715	3,110,715	2,130,374	(3,024,344)	(6,919,676)
20%-40%	3,056,768	2,932,109	3,110,715	3,110,715	3,133,648	(4,131,957)	(5,199,278)
40%-60%	3,052,608	2,875,094	3,110,715	3,110,715	3,649,243	(5,580,274)	(3,651,314)
60%-80%	3,065,362	2,915,393	3,110,715	3,110,715	1,609,679	(4,081,084)	(2,678,406)
80%-100%	3,085,222	2,872,512	3,110,715	3,110,715	1,472,589	(3,603,870)	(3,054,582)

* Combined Ratio assumes 0% expense & profit loads.

Readers may be tempted to develop separate sets of curves for different risks, or lines of business, so I offer a word of caution: try to develop an objective set of standards that clearly define under what circumstances each risk will be assigned to a set of curves. Otherwise, a technical price is difficult to monitor and recalibration becomes problematic. Another issue with separate sets of curves is bias. Presumably, separate sets are developed to capture differences in severity potential. A risk that is not properly assigned could potentially lead to greater bias. However, if risks are properly assigned, we generally find an estimate with lower mean square error.

It should be noted that biased low estimates can still be profitable relative to traditional pricing techniques, suggesting that narrow priors with some bias are acceptable. However, this is a function of the relative variance. If a competitor can produce unbiased estimates with comparable variance, it is likely that the biased low portfolio will be unprofitable.

Even if the risk exposure is well known and an exposure curve is available, a Bayesian approach may be preferable if the exposure curve is not properly constructed. Improperly constructed curves may lead to bias in all business priced. A Bayesian approach could be preferable even when a properly constructed exposure curve is available. Recall that an exposure curve may be accurate for all risks in aggregate, but it is unlikely that all risks share the same underlying severity propensity.

The error statistics presented in this paper are a function of known underlying parameters. In practice, we do not know the underlying severity and loss ratio distributions. With the emergence of Bayesian

Markov chain Monte Carlo (MCMC) models, actuaries can estimate the uncertainty associated with expected outcomes. It is the author's hope that readers can embed the Bayesian approach presented here within an MCMC framework.

On a final note, the technical price should be considered a lower bound. If a competitor's price is significantly above the technical price, one need only price below the competitor's price to write the business. In a head-to-head competition with experience rating, Bayesian profitability would be even more pronounced than shown in Exhibit 8.1.

Acknowledgments

The author wishes to thank friends, family, and colleagues.

References

- Gelman, A., "Prior Distributions for Variance Parameters in Hierarchical Models," *Bayesian Analysis* 1:3, 2006, pp. 515–533.
- Gelman, A., "Informative and Noninformative Priors," *Statistical Modeling, Causal Inference, and Social Science* (blog), July 18, 2007, andrewgelman.com/2007/07/18/informative_and/.
- Klugman, S., H. Panjer, and G. Willmot, *Loss Models: From Data to Decisions* (3rd ed.), Hoboken, NJ: Wiley, 2009.
- Krishnamurthy, H., "Why Bayesian Formulations Better than Maximum Likelihood Estimates?" Oct. 9, 2017, colaberry.com/why-bayesian-formulations-serve-as-better-estimates-than-maximum-likelihood-estimates/.
- Meyers, G., "On Predictive Modeling for Claim Severity," *Casualty Actuarial Society Forum*, Spring 2005, pp. 215–253.
- Patrik, G., and I. Mashitz, "Credibility for Treaty Reinsurance Excess Pricing," *Casualty Actuarial Society Discussion Paper Program*, 1990, pp. 323–374.

Appendix A. MLE estimates

Exhibit A.1. MLE estimate: Lognormal parameters

	Fit	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\hat{\mu}$	9.328									
$\hat{\sigma}$	2.484									
		Emergence	F(c _{j-1} ;θ)	F(c _j ;θ)	layer probabilities		(5)/		Reported	negative-
c _{j-1}	c _j	Pattern	lower	upper	[F(c _j ;θ) - F(c _{j-1} ;θ)]/ [1-F(100,000;θ)]	(4)x(1)	constant	ln(6)	Claims	log likelihood
100,000	135,000	0.896	0.810	0.841	0.1634	0.146	0.177	-1.731	17	29.421
135,000	185,000	0.881	0.841	0.870	0.1516	0.133	0.162	-1.823	13	23.696
185,000	250,000	0.864	0.870	0.894	0.1260	0.109	0.132	-2.026	10	20.260
250,000	500,000	0.843	0.894	0.937	0.2248	0.190	0.229	-1.472	15	22.082
500,000	1,000,000	0.812	0.937	0.965	0.1472	0.120	0.145	-1.933	9	17.398
1,000,000	2,000,000	0.733	0.965	0.981	0.0892	0.065	0.079	-2.537	11	27.902
2,000,000	5,000,000	0.671	0.981	0.993	0.0605	0.041	0.049	-3.015	5	15.073
5,000,000	10,000,000	0.600	0.993	1.000	0.0373	0.022	0.027	-3.610	1	3.610
Totals:					1.0000	0.826	1.000		81	159.442
						constant				

Exhibit A.2. MLE estimate: Ground-up loss ratio

Estimated on-level premium ALL YEARS	60,000,000	(1)	Layer1	Layer2	Layer3	Layer4	Layer5	Layer6	Layer7	Layer8
Ground up Loss Ratio (MLE):	154.0%	(2)	19.0%	15.9%	13.0%	10.6%	6.3%	3.5%	1.9%	0.7%
Excess Loss Percentage:	81.9%	(3); See Exhibit 2 (2), Column (6)	10.20	12.19	14.89	18.25	30.52	54.56	104.39	273.78
Excess Loss Rate :	126.2%	(4); (2)x(3)	1.934	1.082	1.067	1.055	1.033	1.018	1.010	1.004
Excess Loss:	75,697,369	(5); (4)x(1)	100	13.0	10.7	18.6	11.7	6.4	4.0	2.2
Excess Severity:	772,129	(6); See Exhibit 2 (2), Column (3)	0.163	0.152	0.126	0.225	0.147	0.089	0.060	0.037
Expected Excess Claims:	98.04	(7); (5)/(6)	89.6%	88.1%	86.4%	84.3%	81.2%	73.3%	67.1%	60.0%
Probability of attaching:	19.0%	(8); See Exhibit 2 (2), Column (5)	14.4	13.1	10.7	18.6	11.7	6.4	4.0	2.2
Expected Ground Up Claims (over 10 year period):	517.1	(9); (7)/(8)	17.0	13.0	10.0	15.0	9.0	11.0	5.0	1.0
Assumed Contagion:	1.0%	(10)	0.0739	0.1057	0.1189	0.0705	0.0932	0.0314	0.1545	0.2450
Assumed VTM Ground Up:	1.517	(11); $1+(10)\times(9)/10$	0.177	0.162	0.132	0.229	0.145	0.079	0.049	0.027
B (scale parameter gamma distribution):	1.934	(12); $1/[(11)-1]$	-4.34	-4.07	-4.16	-4.12	-4.31	-6.00	-4.88	-5.02
A (scale parameter gamma distribution):	100	(13); $(9)\times(12)/10$	36.888							
Curve #1_Probability of attaching [q]:										
B' = B/q										
VTM:										
Layer Probability:										
Emergence Pattern [% Reported]:										
Expected Reported:										
Actual Reported Claims (simulated):										
Negative Binomial pdf:										
Normalizing constants (adjusted layer probs):										
log likelihood :										
Negative loglikelihood:										
										(24); -sum[(23)]