CREDIBILITY FOR SEVERITY REVISITED

Vincent Goulet,* Antoni Forgues,† and Jiatao Lu‡

ABSTRACT

It is basic actuarial knowledge that the pure premium of an insurance contract can be written as the product of the expected claim number and the expected claim amount. Actuaries use credibility theory to incorporate the contract's individual experience into this calculation in a statistically optimal way. For many years, however, the use of credibility was limited to the frequency component. Starting with the paper by Hewitt (1971), there have been various suggestions as to how credibility theory also can be applied to the severity component of the pure premium. The latest such suggestion, Frees (2003), revived the interest in the problem.

In this paper, we review four different formulas incorporating frequency and severity into credibility calculations. We then compare by simulation which one is most accurate at predicting a contract's next-year outcome. It is found that the classical formula of Bühlmann (1967) is as good as the other ones in many cases. Alternatives, however, may offer easier analysis of the separate effects of frequency and severity on the premium.

We also show that all the formulas reviewed in this paper stem from the same minimization problem, and we present a general, integrated, solution. At the same time, we complete Gerber (1972) by providing a proof to the main result of this paper and by stating required additional assumptions.

---

1. INTRODUCTION

In casualty insurance, the pure premium of a contract generally can be written as the product of the expected claim number and the expected claim amount. Starting with the seminal papers of Mowbray (1914) and Whitney (1918), actuaries used credibility theory to incorporate a contract's individual past experience into its pure premium. For many years, however, the application of this procedure was limited to the first component of the premium—the claim frequency. Mayerson, Jones, and Bowers (1968) then studied the correct way to include the second component—claim severity—into credibility calculations, but in the limited fluctuations context.

In his now classical paper ''Credibility for Severity,'' Hewitt (1971) tackled the problem of appropriately incorporating both claim frequency and claim severity in greatest accuracy credibility calculations. In two excellent, but somewhat neglected, discussions, Gerber (1972) and Bühlmann (1972) complemented and enhanced Hewitt's paper. Frees (2003), although mainly focused on multivariate credibility, proposed an alternative way to account for frequency and severity in the credibility premium, hence reviving interest in the subject.

This paper compares numerically the formulas proposed by the above authors to one another and to the classical formula of Bühlmann (1967). We try to determine which one is most accurate in predicting

---

* Vincent Goulet, PhD, is an Associate Professor, École d'actuariat, Pavillon Vachon, bureau 1620, Université Laval, Québec G1K 7P4, Canada, vincent.goulet@act.ulaval.ca.

† Antoni Forgues is an analyst, Towers Perrin, 71 High Holborn, London, England WC1V 6TP, antoniforgues@towersperrin.com.

‡ Jiatao Lu is a master's student, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada, richjt@hotmail.com.

the future experience of a contract. From a more theoretical point of view, we furthermore show that the various formulas all belong to the same family of solutions, and we provide a proof to the main result of Gerber (1972).

This contribution thus should be of interest to practicing actuaries hesitating as to the correct way to include severity in their credibility calculations, and to instructors wishing to discuss this subject in their courses. There is little emphasis on technical details, but we give many references to the reader interested in filling in the gaps. Proofs are relegated to the Appendix.

2. THE PROBLEM AND MATHEMATICAL MODEL

To simplify the discussion and ease notation, we start by considering only one insurance contract and no weights (or volumes). We are thus in the framework of the original model of Bühlmann (1967); see Goovaerts et al. (1990) or Klugman, Panjer, and Willmot (1998) for details. Following Bühlmann (1972), we, however, add a slight modification to the classical setting to segregate frequency and severity.

We assume that a contract is characterized by two risk parameters, each represented by a random variable: Λ for the frequency of claims and Θ for the severity of claims. These random variables are independent by assumption.

Conditionally on (Λ, Θ), the contract's total amount of claims follows the assumptions of the collective model of risk theory (Gerber 1979; Daykin, Pentikainen, and Pesonen 1994). That is, the total amount of claims in year $t = 1, \ldots, n$ is

$$S_t = X_{t1} + \cdots + X_{tN_t}, \tag{2.1}$$

where N_t is the number of claims and X_{tu} the amount of claim u. It should be understood here that the individual claim amounts (and hence the frequency) are available to the actuary for ratemaking purposes.

The key quantities needed in the sequel are summarized in Table 1. It is well known that

$$\mu(\Lambda, \Theta) = \mu_N(\Lambda)\mu_X(\Theta), \tag{2.2}$$

$$\sigma^2(\Lambda, \Theta) = \sigma_N^2(\Lambda)\mu_X^2(\Theta) + \mu_N(\Lambda)\sigma_X^2(\Theta). \tag{2.3}$$

By independence of Λ and Θ, it then follows that

$$m = m_N m_X, \tag{2.4}$$

$$s^2 = s_N^2(a_X + m_X^2) + s_X^2 m_N, \tag{2.5}$$

and

$$a = a_N a_X + a_N m_X^2 + a_X m_N^2. \tag{2.6}$$

Table 1

Summary of Frequency–Severity Model

	Aggregate Losses	Frequency	Severity
Random variable	$S_t\|\Lambda, \Theta$	$N_t\|\Lambda$	$X_{tu}\|\Theta$
Risk premium	$\mu(\Lambda, \Theta) = \mathrm{E}[S_t\|\Lambda, \Theta]$	$\mu_N(\Lambda) = \mathrm{E}[N_t\|\Lambda]$	$\mu_X(\Theta) = \mathrm{E}[X_{tu}\|\Theta]$
Conditional variance	$\sigma^2(\Lambda, \Theta) = \mathrm{Var}[S_t\|\Lambda, \Theta]$	$\sigma_N^2(\Lambda) = \mathrm{Var}[N_t\|\Lambda]$	$\sigma_X^2(\Theta)\ \mathrm{Var}[X_{tu}\|\Theta]$
Structure parameters	$m = \mathrm{E}[\mu(\Lambda, \Theta)]$	$m_N = \mathrm{E}[\mu_N(\Lambda)]$	$m_X = \mathrm{E}[\mu_X(\Theta)]$
	$s^2 = \mathrm{E}[\sigma^2(\Lambda, \Theta)]$	$s_N^2 = \mathrm{E}[\sigma_N^2(\Lambda)]$	$s_X^2 = \mathrm{E}[\sigma_X^2(\Theta)]$
	$a = \mathrm{Var}[\mu(\Lambda, \Theta)]$	$a_N = \mathrm{Var}[\mu_N(\Lambda)]$	$a_X = \mathrm{Var}[\mu_X(\Theta)]$

The goal in greatest accuracy credibility consists in finding the closest (in the mean square sense) estimator of the aggregate risk premium (2.2).

3. SOLUTIONS TO THE CREDIBILITY PROBLEM

The classical solution to the credibility problem stated above was proposed by Bühlmann (1967). Discussions by Gerber (1972) and Bühlmann (1972) of Hewitt (1971) then yielded alternative formulas incorporating severity into calculations explicitly. Frees (2003) proposed yet another formula.

These formulas, reviewed below, are all related. At various degrees, the proposed credibility premiums are all formulas of the form

$$\alpha\bar{S} + \beta\bar{N}m_X + \gamma\bar{X}m_N + \eta m_N m_X$$

closest to the aggregate risk premium in the mean square sense. That is, the premiums minimize the expression

$$E[(\mu(\Lambda, \Theta) - \alpha\bar{S} - \beta\bar{N}m_X - \gamma\bar{X}m_N - \eta m_N m_X)^2], \tag{3.1}$$

where

$$\bar{N} = \frac{1}{n}\sum_{t=1}^{n} N_t, \tag{3.2}$$

$$\bar{X} = \frac{1}{N_\Sigma}\sum_{t=1}^{n}\sum_{u=1}^{N_t} X_{tu}, \quad N_\Sigma = \sum_{t=1}^{n} N_t\,, \tag{3.3}$$

and

$$\bar{S} = \frac{1}{n}\sum_{t=1}^{n} S_t = \bar{N}\bar{X}. \tag{3.4}$$

3.1 The Bühlmann Formula

Applied to the situation outlined in Section 2, the original model of Bühlmann (1967) will make use only of the total claim amounts $S_1, \ldots, S_n$ to calculate the credibility premium of the following year.

It is well known by now that setting $\beta = \gamma = 0$ in expression (3.1) and minimizing yields the credibility premium

$$\pi^{B}_{n+1} = z\bar{S} + (1 - z)m, \tag{3.5}$$

with

$$z = \frac{n}{n + K_B}, \quad K_B = \frac{s^2}{a}. \tag{3.6}$$

It should be noted that equation (3.5) does take both frequency and severity into account through structure parameters s^2 and a in equation (3.6). Since these parameters rely on aggregate claim amounts only, it proves impractical, however, to measure separately the impact of frequency and severity on the credibility premium.

3.2 The Bühlmann-Hewitt Formula

Hewitt (1971) is a very enlightening exposition of how claim frequency and claim severity interact in the calculation of the credibility premium. As Bühlmann (1972) showed, however, Hewitt's approach is a simple rewriting of equations (3.5) and (3.6) with s^2 and a replaced by the right-hand sides of equations (2.5) and (2.6).

More specifically, the Bühlmann-Hewitt credibility premium for year $n + 1$, π^{BH}_{n+1}, is

$$\pi^{BH}_{n+1} = z\bar{S} + (1 - z)m, \tag{3.7}$$

with

$$z = \frac{n}{n + K_{BH}}, \tag{3.8}$$

but

$$K_{BH} = \frac{s_N^2(a_X + m_X^2) + s_X^2 m_N}{a_N a_X + a_N m_X^2 + a_X m_N^2}. \tag{3.9}$$

Defining

$$K_N = \frac{s_N^2}{a_N}, \qquad K_X = \frac{s_X^2}{a_X},$$

$$C_N = \frac{m_N^2}{a_N}, \qquad C_X = \frac{m_X^2}{a_X},$$

one can rewrite equation (3.9) as

$$K_{BH} = \frac{K_N(1 + C_N) + \dfrac{K_X}{m_N} C_X}{1 + C_N + C_X}.$$

The aggregate credibility constant K_{BH} then appears as a weighted average of the credibility constant based on claim numbers only, K_N, and the ratio of the credibility constant based on claim amounts only, K_X, and the expected number of claims, m_N.

Obviously, the Bühlmann-Hewitt premium is theoretically equivalent to the Bühlmann premium since $K_B = K_{BH}$. In practice, however, the structure parameters (see Table 1) are estimated from the available data. With Bühlmann's formula, one would estimate parameters m, s^2, and a directly. On the other hand, the Bühlmann-Hewitt formula calls for estimation of the components m_N, m_X, s_N^2, s_X^2, a_N, and a_X. The two approaches are not equivalent.

For example, the traditional estimator of parameter s^2, $\hat{s}^2$, is unbiased. Now, even if estimators $\hat{s}_N^2$, $\hat{s}_X^2$, $\hat{a}_X$, $\hat{m}_N$, and $\hat{m}_X$ are unbiased for their respective parameter,

$$\hat{s}_N^2(\hat{a}_X + \hat{m}_X^2) + \hat{m}_N \hat{s}_X^2 \tag{3.10}$$

will be a biased estimator of s^2. The same discussion holds true for the estimation of a versus the estimation of $a_N a_X + a_N m_X^2 + a_X m_N^2$.

On a side note, it is worth mentioning that estimator (3.10), although biased, has a smaller variance than the unbiased estimator $\hat{s}^2$. This is because of a reduction in the mean square error of the estimator. Recall (Hogg, Craig, and McKean 2005) that the variance of an estimator $\hat{\theta}$ of a quantity θ is

$$\mathrm{Var}[\hat{\theta}] = \mathrm{MSE}(\hat{\theta}) + (\mathrm{bias}(\hat{\theta}))^2.$$

Section 4 discusses estimation of the structure parameters in more detail.

3.3 The Gerber Formula

The alternative formula proposed by Gerber (1972) provides a theoretical justification to an otherwise intuitively sound formula. Since the (aggregate) pure premium is the product of the frequency and severity pure premiums, it seems natural to express the (aggregate) credibility premium as the product of the frequency and severity credibility premiums.

Following Gerber (1972), we show in the Appendix that, with additional assumptions to be discussed there, the premium minimizing expression (3.1) is

$$\begin{aligned}\pi^{G}_{n+1} &= z_N z_X \bar{N}\bar{X} + z_N(1 - z_X)\bar{N}m_X \\ &\quad + (1 - z_N)z_X\bar{X}m_N + (1 - z_N)(1 - z_X)m_N m_X \\ &= (z_N\bar{N} + (1 - z_N)m_N)(z_X\bar{X} + (1 - z_X)m_X), \end{aligned} \tag{3.11}$$

with

$$z_N = \frac{n}{n + K_N}, \quad K_N = \frac{s^2_N}{a_N}, \tag{3.12}$$

$$z_X = \frac{w}{w + K_X}, \quad K_X = \frac{s^2_X}{a_X}, \tag{3.13}$$

and w is the observed (or fixed) total number of claims of a contract: $w = \sum_{t=1}^{n} n_t$.

This result is remarkable: the Gerber premium (3.11) is the product of the Bühlmann premium on claim numbers only, and the Bühlmann premium on claim amounts only.

Note that when working with a portfolio of contracts, the number of claims in any given year are unlikely to be the same for each contract. That is, w will vary from one contract to another. As a result, one actually needs to use the Bühlmann-Straub model (Bühlmann and Straub 1970) to calculate the severity credibility premium.

3.4 The Frees-Jewell Formula

One contribution of Frees (2003) is another credibility formula that considers data from both the claim number and claim amounts processes. The version we present here is for a single contract with no weights.

To emphasize the effect of claim frequency compared to the Bühlmann premium, Frees minimizes expression (3.1) with $\gamma = 0$. According to Bühlmann (1973), W. S. Jewell already proposed a credibility premium of the form

$$\alpha\bar{S} + \beta\bar{N}m_X + \eta m_N m_X$$

at the ASTIN Colloquium in 1973 and in a technical paper not available to the present authors. Consequently, we attach the name of Jewell to the resulting premium:

$$\pi^{FJ}_{n+1} = z_S\bar{S} + \tilde{z}_N\bar{N}m_X + (1 - z_S - \tilde{z}_N)m, \tag{3.14}$$

with

$$\tilde{z}_N = \frac{na_N s^2_X m_N - na_X s^2_N m^2_N}{(na_N + s^2_N)(na_X(a_N + m^2_N) + a_X s^2_N + s^2_X m_N)} \tag{3.15}$$

and

$$z_S = \frac{n}{n + K_{FJ}}, \quad K_{FJ} = \frac{a_X s^2_N + s^2_X m_N}{a_N a_X + a_X m^2_N}. \tag{3.16}$$

We denote equation (3.15) with $\tilde{z}_N$ to distinguish it from equation (3.12). Note that it is not possible to write the latter in the form $n/(n + K)$, where K would be some constant.

4. ESTIMATION OF THE STRUCTURE PARAMETERS

As previously mentioned in Section 3.2, in practice, one has to estimate the structure parameters from a portfolio of similar contracts.

In our simulation study, we wanted to use the same estimators in all four models, so that a difference in premium accuracy does not stem from the estimation of the structure parameters process. We thus favored the usual Bühlmann-Straub model estimators (Goovaerts et al. 1990; Goovaerts and Hoogstad 1987). For example,

$$\hat{m} = \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} \bar{S}_i, \quad z_\Sigma = \sum_{i=1}^{I} z_i,$$

$$\hat{s}^2 = \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} (S_{it} - \bar{S}_i)^2,$$

$$\hat{a} = \frac{1}{I-1} \sum_{i=1}^{I} z_i (\bar{S}_i - \hat{m})^2,$$

where subscript $i = 1, \ldots, I$ identifies the contract. Estimator $\hat{a}$, whose value depends on itself, is called a pseudo-estimator, as De Vylder (1981) originally put it. It is well known that the above estimators are unbiased, but that the resulting credibility premiums are biased.

The estimators of parameters m_N, s_N^2, a_N, m_X, s_X^2, and a_X are similar, with S replaced by N or X as appropriate. Estimation of the severity parameters obviously will require one additional summation and the use of weights.

5. SIMULATION MODELS

We used a total of seven different models for simulation of claim data:

1. Basic model: for claim numbers, we let

$$N_t | \Lambda \sim \text{Poisson}(\Lambda),$$

$$\Lambda \sim \text{Gamma}(\alpha, \gamma).$$

Following Hewitt (1971), we retained a lognormal-normal combination for claim amounts:

$$X_{tu} | \Theta \sim \text{Lognormal}(\Theta, \sigma_1^2),$$

$$\Theta \sim \text{Normal}(\mu, \sigma_2^2).$$

For illustration purposes, Table 2 presents the formulas of the structure parameters for this model. We implemented the simulation with $\alpha = 2$, $\gamma = 1$, $\mu = \ln 1{,}500 - 1$, and $\sigma_1^2 = \sigma_2^2 = 1$.

2. Low-frequency model: Same distribution as in the basic model, but with $\gamma = 8$. The other parameters remain unchanged.
3. High-frequency model: Same distribution as in the basic model, but with $\gamma = 0.1$. The other parameters remain unchanged.
4. Linear severity model: The frequency distributions are the same as in the basic model, but the severity distributions are replaced by

Table 2

Formulas of the Structure Parameters in the Basic Simulation Model

Parameter	Aggregate Losses	Frequency	Severity
m	$\frac{\alpha}{\gamma} e^{\mu + \frac{\sigma_1^2 + \sigma_2^2}{2}}$	$\frac{\alpha}{\gamma}$	$e^{\mu + \frac{\sigma_1^2 + \sigma_2^2}{2}}$
s^2	$\frac{\alpha}{\gamma} e^{2(\mu + \sigma_1^2 + \sigma_2^2)}$	$\frac{\alpha}{\gamma}$	$e^{2\mu + \sigma_1^2 + 2\sigma_2^2}(e^{\sigma_1^2} - 1)$
a	$\frac{\alpha}{\gamma^2} e^{2\mu + \sigma_1^2 + \sigma_2^2}((\alpha + 1)e^{\sigma_2^2} - \alpha)$	$\frac{\alpha}{\gamma^2}$	$e^{2\mu + \sigma_1^2 + \sigma_2^2}(e^{\sigma_2^2} - 1)$

Table 3

Values of the Structure Parameters in the Simulation Models

Model	Parameter	Aggregate Losses	Frequency	Severity
Basic	m	3.000×10^3	2	1.500×10^3
	s^2	3.325×10^7	2	1.051×10^7
	a	2.770×10^7	2	3.866×10^6
	K	1.201×10^0	1	2.718×10^0
Low frequency	m	3.750×10^2	0.25	1.500×10^3
	s^2	4.156×10^6	0.25	1.051×10^7
	a	4.328×10^5	0.03125	3.866×10^6
	K	9.604×10^0	8	2.718×10^0
High frequency	m	3.000×10^4	20	1.500×10^3
	s^2	3.325×10^8	20	1.051×10^7
	a	2.770×10^9	200	3.866×10^6
	K	1.201×10^{-1}	0.1	2.718×10^0
Linear severity	m	3.000×10^3	2	1.500×10^3
	s^2	1.500×10^7	2	3.750×10^6
	a	1.350×10^7	2	1.500×10^6
	K	1.111×10^0	1	2.500×10^0
Heavy tail for Λ	m	3.000×10^3	2	1.500×10^3
	s^2	3.325×10^7	2	1.051×10^7
	a	1.718×10^8	25.556	3.866×10^6
	K	1.936×10^{-1}	0.078	2.718×10^0
Heavy tail for Θ	m	3.000×10^3	2	1.500×10^3
	s^2	2.457×10^8	2	7.765×10^7
	a	2.622×10^8	2	4.294×10^7
	K	9.372×10^{-1}	1	1.808×10^0
Heavy tail for Λ and Θ	m	3.000×10^3	2	1.500×10^3
	s^2	1.815×10^9	2	8.625×10^8
	a	1.327×10^9	25.556	4.294×10^7
	K	1.368×10^0	0.078	2.009×10^1

$$X_{tu}|\Theta \sim \text{Exponential}(\Theta),$$

$$\Theta \sim \text{Gamma}(3.5,\ 3{,}750),$$

a combination yielding a linear severity credibility premium.

5. Heavy tail for Λ: The distribution of Λ in the basic model is replaced by $\Lambda \sim \text{Lognormal}(\ln 2 - 1, 2)$. The other distributions and their parameters remain unchanged.
6. Heavy tails for Θ: The distribution of Θ in the basic model is replaced by $\Theta \sim \text{Normal}(\ln 1{,}500 - 2, 3)$. The other distributions and their parameters remain unchanged.
7. Heavy tails for Λ and Θ: Frequency model as in model 5 and the following severity model:

$$X_{tu}|\Theta \sim \text{Lognormal}(\Theta, 3)$$

$$\Theta \sim \text{Normal}(\ln 1{,}500 - 3, 3).$$

The resulting values of the structure parameters for all seven models above are found in Table 3.

6. RESULTS

To compare the accuracy of the four credibility premiums, we simulated six years of experience for portfolios of 100 contracts. We calculated the credibility premium π_6 using the first five years of experience and compared it with the actual outcome, S_6. For each simulation, the accuracy of the various formulas was measured by the mean square error:

$$\text{MSE} = \frac{1}{100} \sum_{i=1}^{100} (\pi_{i,6} - S_{i,6})^2,$$

Table 4

Results of 10,000 Simulations with True Values of the Structure Parameters

Model	Measure	Bühlmann	Gerber	Frees-Jewell
Basic	MSE	1.00027	1.00000	1.00015
	RMSE	2.96	3.29	3.32
	No. smallest MSE	3399	4872	1729
	No. smallest RMSE	8283	1523	194
Low frequency	MSE	1.00014	1.00000	1.00005
	RMSE	25.8	28.0	28.0
	No. smallest MSE	4032	4081	1887
	No. smallest RMSE	6867	2199	934
High frequency	MSE	1.00006	1.00000	1.00006
	RMSE	0.30	0.30	0.31
	No. smallest MSE	3479	4845	1676
	No. smallest RMSE	6911	2287	802
Linear severity	MSE	1.00010	1.00003	1.00000
	RMSE	2.14	2.23	2.27
	No. smallest MSE	3314	4648	2038
	No. smallest RMSE	7657	2310	33

where $i = 1, \ldots, 100$ identifies the contract. To keep numbers low and highlight the differences, we also considered the relative mean square error:

$$\text{RMSE} = \frac{1}{100} \sum_{i=1}^{100} \left(\frac{\pi_{i,6} - S_{i,6}}{\pi_{i,6}} \right)^2 .$$

It should be noted, however, that this measurement is not minimized by the credibility formulas.

These errors were then averaged over 10,000 simulations. We also recorded the number of times each formula had the smallest MSE and RMSE.

The calculations were coded in R (R Development Core Team 2005). The functions used to simulate the data and compute the structure parameters estimators are part of the R package `actuar`.

6.1 True Values of the Structure Parameters

Table 4 first presents the results of 10,000 simulations when the true values of the structure parameters were used in the calculations. In that way, estimation of the parameters does not have any impact on the relative accuracy of the credibility formulas. Note that the Bühlmann and Bühlmann-Hewitt formulas are equivalent in this case, so results for the latter are not shown in the table. Only the first four simulation models were used.

The mean square errors were scaled to 1 to ease reading. To give an idea of the scale, the mean square error with the Bühlmann and Bühlmann-Hewitt premiums using the basic model was 3.75×10^7.

It can be seen from these results that the Gerber premium is more accurate using the mean square error measure, and the Bühlmann (and Bühlmann-Hewitt) premiums fare better using the relative mean square error. However, the differences between the various premiums are actually quite small.

6.2 Estimated Structure Parameters

Of more practical interest, Table 5 presents the results when structure parameters were estimated using the estimators of Section 4 in all four approaches and for each of the seven simulation models.

In general, the Bühlmann and Bühlmann-Hewitt formulas are the ones performing best. The former has a slight advantage, most likely because of the fact that fewer structure parameters have to be estimated. The Gerber formula follows closely, while the Frees-Jewell formula lags behind. The larger errors of this formula are generally due to negative premiums, something that usually happens when $\bar{S} = 0$ and $z_S + \tilde{z}_N > 1$.

Table 5

Results of 10,000 Simulations with Estimated Structure Parameters

Model	Measure	Bühlmann	Bühlman-Hewitt	Gerber	Frees-Jewell
Basic	MSE	1.0000	1.0049	1.0057	1.0058
	RMSE	2.99	3.04	3.44	8.38
	No. smallest MSE	4244	1811	2492	1453
	No. smallest RMSE	4947	3232	1032	789
Low frequency	MSE	1.0065	1.0000	1.0157	1.0020
	RMSE	32.9	31.1	33.2	63.7
	No. smallest MSE	4666	2399	1539	1047
	No. smallest RMSE	4071	3246	1412	1098
High frequency	MSE	1.0005	1.0003	1.0000	1.0002
	RMSE	0.30	0.30	0.32	0.32
	No. smallest MSE	4154	1782	2694	1370
	No. smallest RMSE	5176	2788	1170	866
Linear severity	MSE	1.0056	1.0000	1.0017	1.0011
	RMSE	2.16	2.09	2.26	2.30
	No. smallest MSE	3860	1888	2517	1735
	No. smallest RMSE	5490	3026	1042	442
Heavy tail for Λ	MSE	1.0000	1.0407	1.0108	1.0779
	RMSE	46	73	373	378
	No. smallest MSE	2646	1211	4747	1396
	No. smallest RMSE	7140	2768	92	0
Heavy tail for Θ	MSE	1.0188	1.0058	1.0000	1.0058
	RMSE	10	12	8	139
	No. smallest MSE	4427	1382	2932	1259
	No. smallest RMSE	4018	2227	1907	1848
Heavy tail for Λ and Θ	MSE	1.0000	1.1793	1.1892	1.2936
	RMSE	91	692	5601	5630
	No. smallest MSE	4105	1667	3530	619
	No. smallest RMSE	6367	3436	175	21

The Gerber formula gains accuracy when the frequency of claims is high. Rankings are less constant when reducing the credibility factors, but then all formulas tend toward the collective premium m as the credibility factors tend toward zero.

We also conducted simulations (results not shown here) where there was no claim heterogeneity ($a_X = 0$). The Gerber and Frees-Jewell formulas are equivalent in this case, and they were consistently more accurate than the other two. In turn, the Bühlmann-Hewitt formula was more accurate than the Bühlmann formula. This should be expected because $a_X = 0$ can be ‘‘hard coded’’ into the Gerber and Bühlmann-Hewitt formulas. However, the claim heterogeneity still will be estimated with the Bühlmann formula.

Using 500 contracts instead of 100 did not change the orderings found in Table 5. Increasing the number of contracts improves the quality of the structure parameter estimators and, essentially, had the effect of improving the accuracy of the Frees-Jewell formula and making all premiums closer to each other. The Bühlmann formula generally remained the most accurate.

It should be noted that some rankings may change from one simulation study to another. We believe, however, that the sufficiently large number of repetitions (10,000) and number of contracts (100 and 500) make our results indicative of the relative accuracy of the four formulas studied in this paper.

7. CONCLUSION

This paper reviewed four different means to incorporate the effect of the severity of claims into credibility premiums calculations. We then compared the relative accuracy of the four formulas by simulation. In light of the results presented in Section 6, the classical formula of Bühlmann (1967) still remains an optimal choice in most situations. The Bühlmann-Hewitt and Gerber approaches are useful mainly to study the impact of frequency and severity separately, or when there is no claim heterogeneity.

If claim frequency is "large," the Gerber formula appears to be slightly more accurate. As Gerber (1972) pointed out himself, though, a nicer feature of this formula is its ability to distinguish a contract with, say, 10 claims of 1,000 from one with a single claim of 10,000. Where the Bühlmann and Bühlmann-Hewitt formulas will yield identical premiums for the two contracts, the Gerber formula will appropriately charge more to the first contract. The Frees-Jewell formula has the same ability.

Another goal of this paper is to show that all the formulas reviewed are related to the same minimization problem. This is formalized in the Appendix.

APPENDIX

GENERAL SOLUTION TO THE MINIMIZATION PROBLEM

In this Appendix, we work out the problem of minimizing

$$\mathrm{E}[(\mu(\Lambda, \Theta) - \alpha\bar{S} - \beta\bar{N}m_X - \gamma m_N\bar{X} - \eta m_N m_X)^2] \tag{A.1}$$

with respect to α, β, γ, and η and obtain general formulas. We see in Sections A.1 and A.2 that these general formulas easily reduce to the Bühlmann and Frees-Jewell cases. The Gerber formula necessitates additional assumptions that are introduced in Section A.3.

Before we proceed, any minimization problem in credibility theory requires that we know the covariance relations between the involved random variables. These are given in the following theorem.

Theorem 1

Under the assumptions of Section 2, the following variance and covariance relations hold:

$$\mathrm{Var}[\bar{S}] = a + \frac{s^2}{n}, \tag{A.2}$$

$$\mathrm{Var}[\bar{N}] = a_N + \frac{s_N^2}{n} = \frac{a_N}{z_N}, \tag{A.3}$$

$$\mathrm{Var}[\bar{X}] = a_X + s_X^2\mathrm{E}\left[\frac{1}{N_\Sigma}\right], \tag{A.4}$$

$$\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S}) = a, \tag{A.5}$$

$$\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N}) = a_N m_X, \tag{A.6}$$

$$\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{X}) = a_X m_N, \tag{A.7}$$

$$\mathrm{Cov}(\bar{S}, \bar{N}) = a_N m_X + \frac{s_N^2}{n} m_X = \mathrm{Var}[\bar{N}]\ m_X, \tag{A.8}$$

$$\mathrm{Cov}(\bar{S}, \bar{X}) = a_X m_N + \frac{s_X^2}{n}, \tag{A.9}$$

$$\mathrm{Cov}(\bar{X}, \bar{N}) = 0. \tag{A.10}$$

PROOF

Recall that claim numbers are conditionally independent given Λ, claim amounts are conditionally independent given Θ, and total (or aggregate) claim amounts are conditionally independent given (Λ, Θ). In addition, we assume that Λ and Θ are independent.

Results of the theorem are thus always derived in a similar fashion by first conditioning on the appropriate random variables. For example,

$$\begin{aligned}\mathrm{Var}[\bar{S}] &= \mathrm{Var}[\mathrm{E}[\bar{S}|\Lambda, \Theta]] + \mathrm{E}[\mathrm{Var}[\bar{S}|\Lambda, \Theta]] \\ &= \mathrm{Var}[\mu(\Lambda, \Theta)] + \mathrm{E}\left[\frac{\sigma^2(\Lambda, \Theta)}{n}\right] \\ &= a + \frac{s^2}{n}\end{aligned}$$

and

$$\begin{aligned}\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N}) &= \mathrm{Cov}(\mathrm{E}[\mu(\Lambda, \Theta)|\Lambda, \Theta], \mathrm{E}[\bar{N}|\Lambda, \Theta]) \\ &\quad + \mathrm{E}[\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N}|\Lambda, \Theta)] \\ &= \mathrm{Cov}(\mu(\Lambda, \Theta), \mu_N(\Lambda)) + 0 \\ &= \mathrm{Cov}(\mu_N(\Lambda)\mu_X(\Theta), \mu_N(\Lambda)) \\ &= \mathrm{Var}[\mu_N(\Lambda)]\mathrm{E}[\mu_X(\Theta)] \\ &= a_N m_X.\end{aligned}$$

More intricate, however, are the expressions involving $\bar{X}$. By definition, $\bar{X}$ is a function not only of the claim amounts $\{X_{tu};\ t = 1, \ldots, n;\ u = 1, \ldots, N_t\}$, but also of the claim numbers $\{N_t;\ t = 1, \ldots, n\}$. Therefore, we have

$$\begin{aligned}\mathrm{Var}[\bar{X}] &= \mathrm{Var}[\mathrm{E}[\bar{X}|\Theta]] + \mathrm{E}[\mathrm{Var}[\bar{X}|\Theta]] \\ &= \mathrm{Var}[\mu_X(\Theta)] + \mathrm{E}\left[\frac{\mu_X(\Theta)}{N_\Sigma}\right] \\ &= a_X + s_X^2\mathrm{E}\left[\frac{1}{N_\Sigma}\right],\end{aligned}$$

where $N_\Sigma = \Sigma_{t=1}^{n} N_t$, as defined in equation (3.3). Moreover,

$$\begin{aligned}\mathrm{Cov}(\bar{S}, \bar{X}) &= \mathrm{Cov}(\mu(\Lambda, \Theta), \mu_X(\Theta)) + \mathrm{E}[\mathrm{Cov}(\bar{S}, \bar{X}|\Lambda, \Theta)] \\ &= \mathrm{Var}[\mu_X(\Theta)]\mathrm{E}[\mu_N(\Lambda)] + \mathrm{E}\left[\frac{\sigma_X^2(\Theta)}{n}\right] \\ &= a_X m_N + \frac{s_X^2}{n}.\end{aligned}$$

The last term in the second equality above is best obtained by further conditioning on $\{N_t;\ t = 1, \ldots, n\}$.

Finally, equation (A.10) follows from independence between frequency and severity. □

We now have in hand everything needed to derive general formulas for the minimum of expression (A.1). The procedure below was inspired by Bühlmann (1973). First, let

$$\begin{aligned}V &= \mathrm{E}[(\mu(\Lambda, \Theta) - \alpha\bar{S} - \beta\bar{N}m_X - \gamma m_N\bar{X} - \eta m_N m_X)^2] \\ &= \mathrm{Var}[\mu(\Lambda, \Theta)] + \alpha^2\mathrm{Var}[\bar{S}] + \beta^2\mathrm{Var}[\bar{N}]m_X^2 + \gamma^2\mathrm{Var}[\bar{X}]m_N^2 \\ &\quad - 2\alpha\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S}) - 2\beta\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N})m_X - 2\gamma\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{X})m_N \\ &\quad + 2\alpha\beta\mathrm{Cov}(\bar{S}, \bar{N})m_X + 2\alpha\gamma\mathrm{Cov}(\bar{S}, \bar{X})m_N + 2\beta\gamma\mathrm{Cov}(\bar{N}, \bar{X})m_X m_N \\ &\quad + (1 - \alpha - \beta - \gamma - \eta)^2 m_N^2 m_X^2.\end{aligned}$$

Then, setting equal to 0 the partial derivatives of V with respect to α, β, γ, and η—respectively—yields the general solution

$$\alpha = \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S}) - \beta \mathrm{Cov}(\bar{S}, \bar{N}) m_X - \gamma \mathrm{Cov}(\bar{S}, \bar{X}) m_N}{\mathrm{Var}[\bar{S}]}, \tag{A.11}$$

$$\beta = \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N}) - \alpha \mathrm{Cov}(\bar{S}, \bar{N}) - \gamma \mathrm{Cov}(\bar{N}, \bar{X}) m_N}{\mathrm{Var}[\bar{N}] m_X}, \tag{A.12}$$

$$\gamma = \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{X}) - \alpha \mathrm{Cov}(\bar{S}, \bar{X}) - \beta \mathrm{Cov}(\bar{N}, \bar{X}) m_X}{\mathrm{Var}[\bar{X}] m_N}, \tag{A.13}$$

$$\eta = 1 - \alpha - \beta - \gamma. \tag{A.14}$$

A.1 BÜHLMANN FORMULA

As pointed out in Section 3.1, the Bühlmann—and Bühlmann-Hewitt—premiums follow easily from the above general formulas by setting $\beta = \gamma = 0$. Then,

$$\begin{aligned} \alpha &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S})}{\mathrm{Var}[\bar{S}]} \\ &= \frac{a}{a + \dfrac{s^2}{n}} = z \end{aligned}$$

and $\eta = 1 - z$.

A.2 FREES-JEWELL FORMULA

The Frees-Jewell formula is obtained by setting $\gamma = 0$ in the general formulas. Solving the resulting system of linear equations yields

$$\begin{aligned} \alpha &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S}) - \mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N}) m_X}{\mathrm{Var}[\bar{S}] - \mathrm{Var}[\bar{X}] m_X^2}, \\ \beta &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N})}{\mathrm{Var}[\bar{N}] m_X} - \alpha, \\ \eta &= 1 - \alpha - \beta, \end{aligned}$$

or, after some simplifications,

$$\begin{aligned} \alpha &= \frac{n(a_N a_X + a_X m_N^2)}{n(a_N a_X + a_X m_N^2) + a_X s_N^2 + s_X^2 m_N} = z_S, \\ \beta &= \frac{n a_N s_X^2 m_N - n a_X s_N^2 m_N^2}{(n a_N + s_N^2)(n a_X (a_N + m_N^2) + a_X s_N^2 + s_X^2 m_N)} = \tilde{z}_N, \\ \eta &= 1 - z_S - \tilde{z}_N. \end{aligned}$$

A.3 GERBER FORMULA

In his 1972 discussion of Hewitt (1971), Gerber stated without proof that the premium (3.11) minimizes expression (A.1). Now, replacing the variance and covariances of Theorem 1 in equations (A.11)–

(A.14) quickly leads to a dead end for two reasons: the variance of $\bar{X}$ is not tractable, and $\mathrm{Cov}(\bar{S}, \bar{X})$ is not ''symmetrical'' to the other covariance relations.

To obtain any tractable and meaningful results, one further needs to assume that (1) $\bar{N}$ and $\bar{X}$ are independent, and (2) the number of claims in the denominator of $\bar{X}$ is a fixed constant w as defined in Section 3.3. The first assumption is fairly standard in actuarial science, whereas the second is incorrect but useful from a mathematical point of view. Neither assumption is met in the simulation since one needs to know the number of claims to be able to compute their average amount.

With these assumptions, we have

$$\mathrm{Var}[\bar{X}] = a_X + \frac{s_X^2}{w} = \frac{a_X}{z_X}, \tag{A.15}$$

$$\begin{aligned}
\mathrm{Var}[\bar{S}] &= \mathrm{Var}[\bar{N}\bar{X}] \\
&= \mathrm{E}[\bar{N}^2]\mathrm{E}[\bar{X}^2] - \mathrm{E}[\bar{N}]^2\mathrm{E}[\bar{X}]^2 \\
&= (\mathrm{Var}[\bar{N}] + m_N^2)(\mathrm{Var}[\bar{X}] + m_X^2) - m_N^2 m_X^2 \\
&= \mathrm{Var}[\bar{N}]\mathrm{Var}[\bar{X}] + \mathrm{Var}[\bar{N}]m_N^2 + \mathrm{Var}[\bar{X}]m_X^2
\end{aligned} \tag{A.16}$$

and

$$\begin{aligned}
\mathrm{Cov}(\bar{S}, \bar{X}) &= \mathrm{Cov}(\bar{N}\bar{X}, \bar{X}) \\
&= \mathrm{E}[\bar{N}\bar{X}^2] - \mathrm{E}[\bar{N}]\mathrm{E}[\bar{X}]^2 \\
&= \mathrm{Var}[\bar{X}]m_N.
\end{aligned} \tag{A.17}$$

The relations involving $\bar{X}$ are now in every respect symmetrical to those involving $\bar{N}$. Moreover, the product $\mathrm{Var}[\bar{N}]\mathrm{Var}[\bar{X}]$ in equation (A.17) proves essential to obtain the product $z_N z_X$ below. Solving equations (A.11)–(A.14) then easily yields

$$\begin{aligned}
\alpha &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{S}) - \beta\mathrm{Var}[\bar{N}]m_X^2 - \gamma\mathrm{Var}[\bar{X}]m_N^2}{\mathrm{Var}[\bar{N}]\mathrm{Var}[\bar{X}] + \mathrm{Var}[\bar{N}]m_N^2 + \mathrm{Var}[\bar{X}]m_X^2}, \\
\beta &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{N})}{\mathrm{Var}[\bar{N}]m_X} - \alpha, \\
\gamma &= \frac{\mathrm{Cov}(\mu(\Lambda, \Theta), \bar{X})}{\mathrm{Var}[\bar{X}]m_N} - \alpha, \\
\eta &= 1 - \alpha - \beta - \gamma,
\end{aligned}$$

that is,

$$\begin{aligned}
\alpha &= z_N z_X, \\
\beta &= z_N(1 - z_X), \\
\gamma &= (1 - z_N)z_X, \\
\eta &= (1 - z_N)(1 - z_X),
\end{aligned}$$

as needed. This completes the proofs.

ACKNOWLEDGMENTS

By a combination of circumstances, the current affiliations of the authors fail to show that this research was conducted almost completely at Concordia University, Montréal, Canada. The first and third authors also acknowledge financial support from the Natural Sciences and Engineering Research Council of Canada. Finally, the authors wish to thank two anonymous referees for suggestions that yielded improvements to this paper.

REFERENCES

BÜHLMANN, HANS. 1967. Experience Rating and Credibility. *ASTIN Bulletin* 4: 199–207.

———. 1972. Discussion of Hewitt (1971). *Proceedings of the Casualty Actuarial Society* 58: 27–31.

———. 1973. A Comparison of Three Credibility Formulae Using Multidimensional Techniques. *ASTIN Bulletin* 7: 203–7.

BÜHLMANN, HANS, AND E. STRAUB. 1970. Glaubgewürdigkeit für Schadensätze. *Bulletin of the Swiss Association of Actuaries* 70: 111–33. English translation by C. E. Brooks.

DAYKIN, C., T. PENTIKAINEN, AND M. PESONEN. 1994. *Practical Risk Theory for Actuaries.* London: Chapman & Hall.

DE VYLDER, F. ETIENNE. 1981. Practical Credibility Theory with Emphasis on Parameter Estimation. *ASTIN Bulletin* 12: 115–31.

FREES, EDWARD W. 2003. Multivariate Credibility for Aggregate Loss Models. *North American Actuarial Journal* 7: 13–37.

GERBER, HANS U. 1972. Discussion of Hewitt (1971). *Proceedings of the Casualty Actuarial Society* 58: 25–27.

———. 1979. *An Introduction to Mathematical Risk Theory.* Philadelphia: Huebner Foundation.

GOOVAERTS, MARC J., AND W. J. HOOGSTAD. 1987. *Credibility Theory, Surveys of Actuarial Studies, no. 4.* Nationale-Nederlanden N.V., the Netherlands.

GOOVAERTS, MARC J., R. KAAS, A. E. VAN HEERWAARDEN, AND T. BAUWELINCKX. 1990. *Effective Actuarial Methods.* Amsterdam: North-Holland.

HEWITT, C. C. 1971. Credibility for Severity. *Proceedings of the Casualty Actuarial Society* 57: 148–71.

HOGG, ROBERT V., ALLEN T. CRAIG, AND JOSEPH W. MCKEAN. 2005. *Introduction to Mathematical Statistics.* 6th ed. New York: Prentice Hall.

KLUGMAN, STUART A., HARRY H. PANJER, AND GORDON WILLMOT. 1998. *Loss Models: From Data to Decisions.* New York: Wiley.

MAYERSON, A. L., DONALD A. JONES, AND NORMAN L. BOWERS. 1968. On the Credibility of the Pure Premium. *Proceedings of the Casualty Actuarial Society* 55: 175–85.

MOWBRAY, A. H. 1914. How Extensive a Payroll Exposure Is Necessary to Give a Dependable Pure Premium? *Proceedings of the Casualty Actuarial Society* 1: 25–30.

R DEVELOPMENT CORE TEAM. 2005. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

WHITNEY, A. W. 1918. The Theory of Experience Rating. *Proceedings of the Casualty Actuarial Society* 4: 275–93.

Discussions on this paper can be submitted until July 1, 2006. The authors reserve the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.