

USING MULTI-DIMENSIONAL CREDIBILITY TO ESTIMATE CLASS FREQUENCY VECTORS IN WORKERS COMPENSATION

BY

JOSE COURET AND GARY VENTER

ABSTRACT

The US workers compensation system is different from those in many countries, but it is reinsured in the world-wide market and so has international impact. From its origin in the early 20th century it has been a laboratory for actuarial credibility techniques. In recent years deductibles have been increasing, so that fairly high excess coverage is now commonplace. This puts growing emphasis on estimation of the percentage of loss that is excess of high deductibles. A key element of the excess percentage is the frequency of loss by injury type. Fatalities and permanent disabilities cost more than other injury types, so when they have high relative frequency, more of the claims cost arises from large losses. The vector of claim frequency by injury type can be estimated by class of business using multi-dimensional credibility techniques. Historically the fraction of costs excess of various retentions has been calculated for large groups of classes (hazard groups) and not individual classes. We show, by testing a hold-out sample, that credibility estimation by class does produce additional information in comparison to a widely-used seven-hazard-group system.

USING MULTI-DIMENSIONAL CREDIBILITY TO ESTIMATE CLASS FREQUENCY VECTORS IN WORKERS COMPENSATION

The regression framework for credibility started with Hachemeister (1975) and has generated an impressive literature since. We apply a small portion of this to estimate vectors of relative frequency by injury type for classes of US workers compensation insureds. Credibility is non-parametric so it is not usually considered necessary to look at the actual distributions of the data or the parameters, but since it is based on minimizing squared error, it may be inappropriate for heavy-tailed distributions, where relative error is more important. Credibility on the logarithms of the data is often suggested in that case. Here, however, credibility is applied to loss frequencies, which are probably safely within the applicability of squared error. Some performance tests show that credibility does help with this estimation.

US workers compensation classes are based on industry breakdowns and in some cases even particular occupations within industries, to the extent that these have different accident potential. For instance, workers comp rates and excess losses for casting can differ significantly depending on whether you are casting pig iron, pottery, broken bones, dice or actors. Also, the occupation class rates vary by state, but usually not among the regions within a state.

Estimating loss costs by class has been a focus of actuaries since the workers compensation system started in 1912. In fact early versions of credibility theory and even the creation of the Casualty Actuarial Society arose to deal with workers compensation issues. The recognition that large individual claims are a driving force in loss costs also came early, with an excess reinsurance pool formed by 1914. Rating plans that charge large manufacturers their actual small losses plus expected large losses also have a long history, so excess rates are not a new topic. However true excess coverage is newer, and is now a large part of the workers compensation business. Thus there is a growing interest in estimation of excess loss costs.

This is driving an exploration of methodologies that might be useful for excess loss estimation. Insurance companies in the US report class-by-state data to rating bureaus. These bureaus also license data that can be used in studies of loss risk, but the licenses typically do not allow further dissemination of the data. Rating bureaus as well as large insurers and reinsurers are all investigating ways to improve the estimation of excess costs for the class/state cells. Traditionally this estimation used a four-hazard-group breakout for estimation of injury-type¹ frequencies. Hazard groups are collections of classes similar in their excess loss potential, relative to total losses. There are significant claims-cost differences across injury types, so estimating the frequency of claims by type for each hazard group has been the main emphasis of excess rating.

Recent innovations in excess rating include bureaus using more hazard groups – as many as nine, with seven being the industry standard – to get more homogeneity in loss potential; looking at possible differences in claims costs within an injury type across classes or hazard groups; and looking for better ways to combine data from different state systems. Some insurers and reinsurers have tried to produce individual class excess rates, at least for the larger classes. We compare the proposed class credibility approach to the seven-hazard-group frequency-only method, and find that it does provide improved estimation.

There are some variations in the way injury types are defined. We use the following injury types: fatal, permanent total injury, major permanent partial impairment, minor permanent partial, temporary total, and medical-only. We abbreviate these F, PT, major, minor, TT, and med. A typical breakout of relative frequency and severity by injury type, using TT as the base, is:

¹ A categorization of injuries by severity, defined in more detail below.

Type:	F	PT	Major	Minor	TT	Med
FREQUENCY RELATIVE TO TT	0.006	0.006	0.085	0.37	1.00	4.3
SEVERITY	60	125	40	4	1.00	0.2

These numbers are illustrative and can vary significantly by state and class, but they show the extent to which the excess costs are driven by the rare but expensive serious injuries. Excess losses above different retentions will be driven by different mixes of injury types, since minor – for example – is much more frequent but less severe than F or PT. Thus getting a good estimate of the class relative frequency vector is a key step in calculating excess loss potential. We will focus on estimating the vector of relativities of F, PT, major, and minor to TT.

1. DEVELOPMENT OF THE CREDIBILITY PROCEDURE

The serious injury types have low frequencies so class claim counts are fairly unstable. What helps however is that they are correlated. The physical circumstances producing fatal, permanent total, and major permanent partial injuries are often quite similar, as slight differences in an accident can yield significantly different outcomes. Thus a class that experiences a lot of major claims is likely to have a higher-than-average propensity to produce PT and fatal claims as well. By applying multi-dimensional credibility theory, the correlations can be used to improve the estimates of each of the elements.

Credibility estimation is similar to regression in that it is a linear model fit by minimizing squared errors. Taking correlation into account is like doing a multiple regression, where the levels of several variables get different weights in estimating the outcomes. However credibility is not exactly the same as regression. Credibility estimates the (unobserved) population mean for a group, and the sample mean is one of the estimators that gets some credibility. In a regression the sample mean would be the dependent variable and the factors (coefficients) would only apply to the other variables. Credibility can be looked at as a regression for which the dependent variable (the population mean) is not observed. Instead a model is postulated about how the observed variables arise stochastically from the unobserved population mean. The data is used to estimate the parameters of that model in order to minimize the expected squared error in the estimate of the population mean. The approach below is basically an application of the methods outlined in Venter (1985).

This is based on a somewhat simplified version of credibility theory, which is outlined first. In standard credibility theory, an individual insured (policyholder or class) is assumed to have n iid observations X_1, \dots, X_n whose distribution is controlled by a parameter θ , where θ is an instance of a random variable Θ with density $\pi(\theta)$. Define $\mu(\theta) = E(X_j | \Theta = \theta)$ and $v(\theta) = \text{Var}(X_j | \Theta = \theta)$.

Often $\mu(\theta)$ is called the hypothetical mean and $v(\theta)$ the process variance. In classical statistics, $\mu(\theta)$ is called the population mean, but Charles Hewitt, a Bayesian actuary, considered $\mu(\theta)$ to be a model construct, not a truly existing entity, and so called it hypothetical, and the terminology has persisted. Let $\mu = E\mu(\Theta)$, $v = Ev(\Theta)$, and $a = \text{Var}[\mu(\Theta)]$. Here v is the expected process variance and a is the variance of the hypothetical means.

The standard Bühlmann approach is to estimate $\mu(\theta)$ linearly as $a_0 + \sum a_j X_j$ by minimizing the expected squared error. The answer is $a_0 = (1 - z)\mu$, $a_j = z/n$ for $j > 0$, where $z = n/(n + k)$, $k = v/a$. Denoting by X^* the average of the X_j , this estimates $\mu(\theta)$ by $zX^* + (1 - z)\mu = \mu + z(X^* - \mu) = EX^* + z(X^* - EX^*)$.

The simplified approach is to start with the estimator $zX^* + (1 - z)\mu$ for $\mu(\theta)$ so only z has to be estimated. This simplification does not give up much because the best linear estimate of the mean from the observations is the sample mean. The assumptions imply $\mu(\theta) = X^* + v(\theta)^{\frac{1}{2}}\varepsilon = \mu + a^{\frac{1}{2}}\eta$, where ε and η are independent mean 0, variance 1 deviations. This can be generalized somewhat to having two estimators X and Y of C with expected squared errors s^2 and t^2 , respectively, where s and t might even be random variables themselves. The problem then becomes to find the z that minimizes $E\{[C - zX + (z - 1)Y]^2\}$.

The assumptions imply that $X = C + s\varepsilon$, $Y = C + t\eta$. To find z , set the derivative of the expected squared error to zero. Then $0 = E\{[C - zX + (z - 1)Y][Y - X]\} = E\{-zs\varepsilon + (z - 1)t\eta\}[t\eta - s\varepsilon] = E[zs^2\varepsilon^2 + (z - 1)t^2\eta^2] = zE[s^2] + (z - 1)E[t^2]$. Thus $z = E(t^2) / [E(s^2) + E(t^2)]$.

In the credibility model $E(s^2)$ is the expected process variance and t^2 is already a constant – the variance of the hypothetical means. The general result can be expressed as $z = [1/E(s^2)] / [1/E(s^2) + 1/E(t^2)]$ so the weight on X is proportional to the reciprocal of its variance, and similarly for Y . This is a standard statistical result.

The multivariate correlated credibility derivation below follows this approach. The class population means for the injury types are estimated as linear functions of the sample means of all the injury types for the class, and the coefficients are estimated by minimizing the expected squared error.

This procedure is applied to ratios of claims counts by injury type to TT claims counts. Denote the observed ratios for the four injury types F, PT, major, and minor as V , W , X , and Y . Assume that each class has parameters that determine its distribution of claim counts by injury type but since the parameters are unobserved they are considered random variables. For the i^{th} class, denote its population (or hypothetical) mean ratios as v_i , w_i , x_i , and y_i . These means are among the parameters that are considered random variables since they are not known.

Focusing on PT, the observed sample claim count ratio to TT, W , for class i at time t , W_{it} , based on m_{it} TT claims, can be formulated:

$$m_{it} W_{it} = \sum_{j=1}^{m_{it}} (w_i + \varepsilon_{jt})$$

or

$$W_{it} = w_i + \left(\sum_{j=1}^{m_{it}} \varepsilon_{jt} \right) / m_{it}$$

where the ε_{jt} are independent mean-zero innovations whose standard deviation σ_{W_i} varies by class but not by time. This looks at each TT claim as an exposure that could produce a PT claim and sometimes does. As one possible example, ε_{jt} might be $1 - w_i$ with probability w_i and $-w_i$ with probability $1 - w_i$. This gives it mean 0 and variance $\sigma_{W_i}^2 = w_i(1 - w_i)$. Then every TT claim produces a random draw of 0 or 1, and the sum of these m_{it} draws is the number of PT claims $m_{it}W_{it}$. Then $\text{Var}(W_{it} | w_i) = \sigma_{W_i}^2 / m_{it}$. This is the statistical import of this formulation. The more TT claims there are, the lower the random fluctuation of the annual observed class ratio W_{it} from its mean is assumed to be.

Denote by W_i the class sample mean ratio over all the time periods (assumed independent):

$$W_i = \sum_{t=1}^N m_{it} W_{it} / \sum_{t=1}^N m_{it}$$

where there are N periods of observation, and similarly for V , X , and Y . Let m_i denote the sum over times t of the m_{it} and m denote the sum over classes i of the m_i . Then the assumption that $\text{Var}(W_{it} | w_i) = \sigma_{W_i}^2 / m_{it}$ implies that $\text{Var}(W_i | w_i) = \sigma_{W_i}^2 / m_i$.

Taking correlation into account in a linear estimate would estimate an injury-type ratio to TT for a class as a linear combination of the observations for all the injury types for that class. The simplified credibility theory uses only those estimators that are linear combinations of sample means, and looks for the linear combination that minimizes the expected squared error of the estimate. Here the expected squared error is taken across all the classes in the hazard group. For PT this can be expressed as minimizing the expected squared difference between the linear estimator and the unobserved conditional mean w_i :

$$(1) \quad \text{minimize } E[(a + bV_i + cW_i + dX_i + eY_i - w_i)^2],$$

The coefficients sought are a , b , c , d , and e . Differentiating (1) wrt a gives:

$$(2) \quad a = -E(bV_i + cW_i + dX_i + eY_i - w_i).$$

Substituting this a back into $a + bV_i + cW_i + dX_i + eY_i$ makes this estimate of w_i

$$Ew_i + b(V_i - EV_i) + c(W_i - EW_i) + d(X_i - EX_i) + e(Y_i - EY_i)$$

The expectation for w_i is unconditional, so it is over all classes. If this procedure is done within a hazard group, it then estimates the hazard group mean ratio of PT to TT. Thus if the class' own data gets zero credibility, the hazard

group ratio will be used. Also the unconditional expected value of W_i is Ew_i (the conditional mean is w_i), so c is the traditional credibility factor Z . In the traditional case, the estimation uses only class i 's own sample mean W_i so the estimator is $Ew_i + c(W_i - EW_i) = cW_i + (1 - c)EW_i$.

The derivative of (1) wrt b gives:

$$aEV_i + E[V_i(bV_i + cW_i + dX_i + eY_i - w_i)] = 0$$

Plugging in (2) for a then yields:

$$(-E(bV_i + cW_i + dX_i + eY_i - w_i)) EV_i + E[V_i(bV_i + cW_i + dX_i + eY_i - w_i)] = 0.$$

This can be rearranged to give:

$$(3) \quad \text{Cov}(V_i, w_i) = b\text{Var}(V_i) + c\text{Cov}(V_i, w_i) + d\text{Cov}(V_i, X_i) + e\text{Cov}(V_i, Y_i).$$

Doing the same thing for c , d , and e will yield three more equations that look like (3), but with the variance moving over one position each time. This will end up with four equations that can be written as a single matrix equation:

$$(4) \quad \begin{pmatrix} \text{Cov}(V_i, w_i) \\ \text{Cov}(W_i, w_i) \\ \text{Cov}(X_i, w_i) \\ \text{Cov}(Y_i, w_i) \end{pmatrix} = C \cdot \begin{pmatrix} b \\ c \\ d \\ e \end{pmatrix},$$

where C is the covariance matrix of the class by injury-type sample means, which show up in (3), for instance. Thus (4) is the requirement needed for b , c , d , and e to be the optimal weights for estimation of w_i .

The problem becomes how to estimate the covariances needed to solve (4). The variance of V_i conditional on the true class mean v_i is $\sigma_{V_i}^2/m_i$ where $\sigma_{V_i}^2$ is called the process variance for V_i . The unconditional variance of V_i is $EPV_V/m_i + VHM_V$, where the latter term, the variance of the hypothetical means for V , is the variance across the classes of the unobserved V means v_i . EPV_V is $E(\sigma_{V_i}^2)$, which does not depend on i as the expectation is across all classes.

This uses the well-known formula $\text{Var}(V_i) = E[\text{Var}(V_i|\lambda)] + \text{Var}[E(V_i|\lambda)]$. Assuming the conditioning is on all the unobserved parameters for the class, the analogous formula for covariance is $\text{Cov}(V_i, W_i) = E[\text{Cov}(V_i, W_i|\lambda)] + \text{Cov}[E(V_i|\lambda), E(W_i|\lambda)] = E[\text{Cov}(V_i, W_i|\lambda)] + \text{Cov}(v_i, w_i)$. Since W is w_i plus a random innovation independent of w_i , $\text{Cov}(W_i, w_i)$ is $\text{Cov}(w_i, w_i)$ which is VHM_W . Similarly, $\text{Cov}(V_i, w_i) = \text{Cov}(v_i, w_i)$, etc.

This model assumes that the ratios observed for any year for each type of injury by class are the class-injury means plus independent random draws. This assumption makes $E[\text{Cov}(V_i, W_i|\lambda)]$ zero. This can be seen by expanding

$E(V_i W_i)$ and $E(V_i)E(W_i)$ in terms of means and independent innovations. Thus $\text{Cov}(V_i, W_i) = \text{Cov}(v_i, w_i)$ which can be estimated by the sample covariance.

Thus the off-diagonal elements of C can be estimated by the sample covariances, $\text{Cov}(W_i, w_i)$ is VHM_W and the covariances of $V_i, X_i,$ and Y_i with w_i are their covariances with W_i from the sample covariance. A reasonable calculation of the sample covariance, say of V and W , would seem to be

$$\Sigma(V_i - V)(W_i - W)m_i/m = \Sigma V_i W_i m_i/m - VW.$$

The estimates of VHM_W and EPV_W , etc. needed for this are standard in credibility theory. One reference is Dean (2005) who gives the following formula for VHM_V , converted to our notation:

$$(5) \quad V\hat{H}M_V = \left(\sum_{i=1}^R m_i (V_i - V)^2 - (R - 1) E\hat{P}V_V \right) / \left(m - \left(\frac{1}{m} \right) \sum_{i=1}^R m_i^2 \right).$$

Here V is the weighted average of the V_i 's using weights m_i and EPV is estimated from the sample variances for each class. This estimate can be negative, in which case it is set to 0, as then EPV accounts for all the observed variation and there is no support for individual risk differences. Also Dean (2005) gives:

$$(6) \quad E\hat{P}V_V = \frac{\sum_{i=1}^R \sum_{t=1}^N m_{it} (V_{it} - V_i)^2}{R(N - 1)}.$$

Note that this is not divided by the sum of the m 's as it is the process variance per unit of exposure, which here is the TT count. The diagonal elements of C for the i^{th} class are then $EPV_V/m_i + VHM_V, EPV_W/m_i + VHM_W$, etc.

2. PERFORMANCE TESTING

We had seven years of class data, at various levels of maturity, for the estimation, but for testing we discarded the quite immature first report. As a test, we calculated injury-type relativities from the even reports (2, 4, and 6) and used these to predict the hold out sample of odd reports (3, 5, and 7). Holding out the last two years and predicting by the older years gave comparable results. For testing credibility itself, the current approach seems to make more sense, as it is more neutral with respect to trend and loss development. Table A summarizes the sum of squared prediction errors by hazard group for the ratio of major permanent partial to temporary total claims. The squared errors are calculated for each combination of state and class – about 16,000 observations overall. Table B lists totals for all injury types.

TABLE A
SUM OF SQUARED PREDICTION ERRORS – MAJOR PERMANENT PARTIAL CLAIMS

(1) Hazard Group	(2) Prediction Based on HG	(3) Prediction Based on Raw Even	(4) Prediction Based on Cred. Proc.
A	33.4	52.9	33.0
B	206.9	363.2	205.1
C	250.8	479.3	248.0
D	89.9	129.9	89.3
E	333.2	484.0	330.8
F	247.1	306.3	242.6
G	263.7	386.2	256.9
Total	1,425.0	2,201.7	1,405.6

TABLE B
SUM OF SQUARED PREDICTION ERRORS BY INJURY TYPE

(1) Injury Type	(2) Prediction Based on HG	(3) Prediction Based on Raw Even	(4) Prediction Based on Cred. Proc.
Fatal	43.6	65.2	43.5
PT	34.7	83.6	34.6
Major PP	1,425.0	2,201.7	1,405.6
Minor PP	6,756.9	10,360.3	6,558.0
Med. Only	417,260.8	434,837.9	351,270.8

The credibility procedure yields a modest reduction in the sum of squared prediction errors. The credibility procedure is designed to minimize the expected deviation between the true class mean and its sample estimator *over the same time period* (the even years). In performance testing, however, a three-year holdout sample of odd years is used as a proxy for the even year's true mean. That is, the estimator derived from even-year data is compared to the ratio for the holdout sample from odd years. Thus, there is some disconnect between the expectation being minimized and the statistic being used to gauge performance.

Relative incidence ratios are impacted by unknown covariates – with levels varying between odd and even years, indeed over all years. Such unknown effects contribute to the error of the predictions. A class code represents the experience of a dynamic portfolio of individual insurance policies whose composition varies over time. In addition, the individual class-state ratios are quite

volatile. Improving the estimates of the means might produce only a small improvement in the total sum of squared errors.

Interestingly, researchers trying to test the Capital Asset Pricing Model (CAPM) faced a similar measurement problem. Under CAPM, individual excess stock returns are hypothetically proportional to beta. The Security Market Line (SML) is a graphical representation of this relationship. Early tests of CAPM which focused on empirical validation of the SML ran into statistical problems. Bodie, Kane and Marcus (BKM) write:

“The next wave of tests was designed to overcome the measurement error problem that led to biased estimates of the SML. The innovation in these tests, pioneered by Black, Jensen, and Scholes (BJS), was to use portfolios rather than individual securities. Combining into portfolios diversifies away most of the firm-specific part of the returns, thereby enhancing the precision of estimates of beta and the expected rate of return of the portfolio of securities. This mitigates the statistical problems that arise from measurement error in the beta estimates.”

Instead of assigning stocks to portfolios randomly, BJS sort by descending beta and then group by quantile. BKM, in their heuristic explanation of BJS, write:

“.. we need to construct portfolios with the largest possible dispersion of beta coefficients. Other things being equal, a sample yields more accurate regression estimates the more widely spaced are the observations of the independent variables.”

“Rather than allocate 20 stocks to each portfolio randomly, we can rank portfolios by betas. Portfolio 1 will include the 20 highest-beta stocks and Portfolio 5 the 20 lowest beta stocks. In that case a set of portfolios with small nonsystematic components, e_p , and widely spaced betas will yield reasonably powerful tests of the SML.”

CAPM is a theory about the relationship between beta and *expected* excess return. Even an average of returns for an individual stock is volatile. Additionally, there are systematic effects that vary over time, which means that a firm's beta is not constant over time. The BJM approach of using portfolios diversifies away the systematic, firm specific part of variation in returns.

In our case, performance testing based on a sum of squared errors criteria only produces a slight improvement; however, testing based on *ranked portfolios* of state-class combinations demonstrates a significant improvement. Our approach is based on the “quintiles test”, a technique developed by NCCI for testing their experience rating plan.

Dorweiler wrote in 1934:

“A necessary condition for proper credibility is that the credit risks and debit risks equally reproduce the permissible loss ratio. Also, if the proper credibility has been attained, each [random] subgroup of the credit and debit risks, provided it has adequate volume, should give the permissible loss ratio.”

Gillam (1992) explains that the “quintiles test” is a logical extension of Dorweiler's criteria. (Dorweiler's groups were defined by modification range, but varied by size.) In our analysis, class relativities to the hazard group average are analogous to the experience rating modifications of interest to Gillam.

TABLE C
ILLUSTRATION OF QUINTILES TEST – HAZARD GROUP D, PERMANENT TOTAL CLAIMS

(1) Quintile	(2) Odd Relativity	(3) Prediction Based on HG	(4) Prediction Based on Raw Even	(5) Prediction Based on Cred. Procedure
1	0.4951	1.0000	0.3065	0.5648
2	0.8634	1.0000	0.4260	0.8732
3	0.9861	1.0000	0.7513	1.0000
4	1.1269	1.0000	1.3473	1.1038
5	1.5215	1.0000	2.1547	1.4519
Mean	1.0000	1.0000	1.0000	1.0000
SSE		0.5618	0.7315	0.0105

We illustrate the procedure using W relativities for classes in the fairly large hazard group D, but the procedure for the other injury types is the same. This is illustrated in Table C.

For the quintiles test, the classes within hazard group D were first grouped by credibility weighted relativity from the even years' data into five groups of roughly equal size. (The boundaries were selected to make the number of TT claims in each quintile about equal.) The lowest 20% of the credibility weighted class relativities belong to the risks in the first quintile; the next 20% to the second; and so on. Column (2) represents the average relative frequency W for the classes in the quintile *divided by* the corresponding estimate for all of hazard group D. For example, the relative frequency of PT claims (as a ratio to TT) for the classes in the first quintile is about half of the hazard group average; the fifth quintile average exceeds the hazard group average by 52%. Upwardly sloping relativities are desirable, indicating that the credibility procedure used to determine quintiles tends, on average, to identify class differences in W .

The goal is to predict the column (2) frequency relativity for each quintile. Column (3) is a prediction based on the hazard group average. All its entries are equal to unity – since by assumption every quintile has the hazard group D relative frequency for PT. The predictions in Column 4 are based on raw class relativities observed for the even years. The aggregation of the classes in the fifth quintile, for example, has an *even year* relative PT frequency that is 215% of the hazard group average. Column (5) predictions were derived using the multi-dimensional credibility procedure described in this paper.

The exhibit is based on five groups with about 1400 PT claims in the even years. As expected, the ratios in column (2) increase across the quintiles; classes for which credibility predicts high relativities using even year data tend to generate high relativities during the odd years. The predictions in column (3), are too high for the first three quintiles and too low for the last two; the sum of

squared prediction errors based on column (3) predictions is 0.5618. Similarly, the sum of squared prediction errors based on column (4) predictions is 0.7315. The upward slope in column (4) is too steep; the odd years exhibit regression toward the mean. The credibility procedure predictions in column (5), have a sum of squared prediction errors of 0.0105. To the extent the class relativity procedure performs, the ratios in column (5) will track those in column (2); producing a lower sum of squared deviations.

As indicated by the sum of squared prediction errors in Table D, the results are similar for most other combinations of injury type and hazard group, the exception being hazard group A for all three injury types. The quintiles test suggests that hazard group A is homogeneous and the credibility procedure is not improving the estimation there, although the sum of individual squared errors is slightly better for the credibility estimates. For major the raw class data provides surprisingly good estimates for the quintiles as a whole once these have been defined by the credibility procedure. These estimates are quite good

TABLE D
 QUINTILES TEST SUM OF SQUARED PREDICTION ERRORS BY HAZARD GROUP AND INJURY TYPE

HG	Injury Type	Prediction Based on HG	Prediction Based on Raw Even	Prediction Based on Cred. Proc.
A	Fatal	0.13227	0.94431	0.18952
B	Fatal	0.32630	1.79940	0.05637
C	Fatal	0.83604	1.39413	0.03376
D	Fatal	0.97498	0.87260	0.12111
E	Fatal	0.49691	1.44023	0.05096
F	Fatal	0.39060	1.35362	0.07280
G	Fatal	0.55650	1.23015	0.06035
A	PT	0.03941	1.94151	0.57993
B	PT	0.38273	1.34145	0.11044
C	PT	0.56175	0.55609	0.01180
D	PT	0.56183	0.73151	0.01053
E	PT	0.73195	0.82350	0.07050
F	PT	0.56872	0.53817	0.01812
G	PT	1.09139	0.52326	0.07946
A	Major	0.58481	0.01988	0.05079
B	Major	0.33888	0.03729	0.00870
C	Major	0.38001	0.04108	0.00738
D	Major	0.18900	0.03928	0.01850
E	Major	0.28775	0.07476	0.01030
F	Major	0.32418	0.04703	0.01781
G	Major	0.58518	0.14046	0.00538

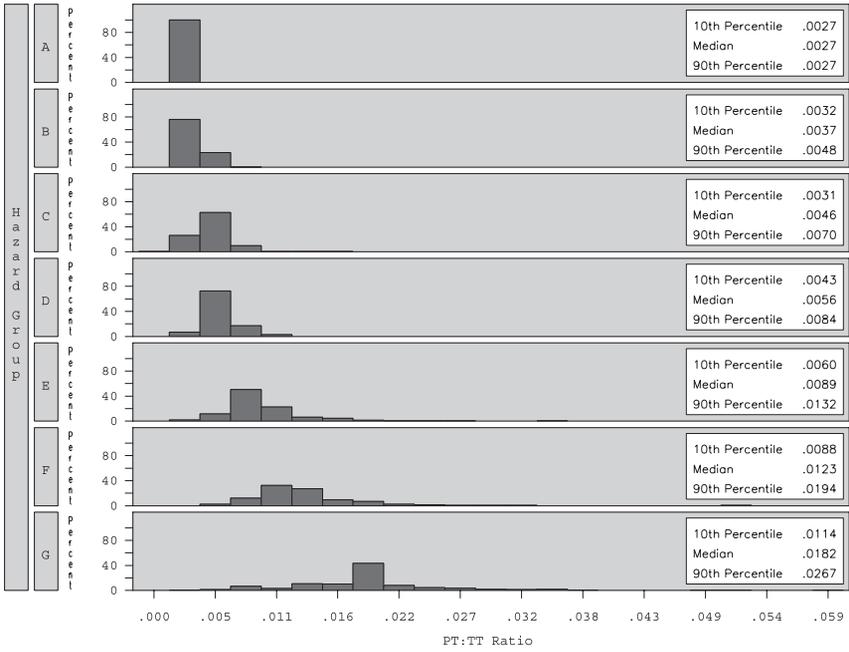


FIGURE 1: Distribution of Predicted PT:TT Ratios by Hazard Group.

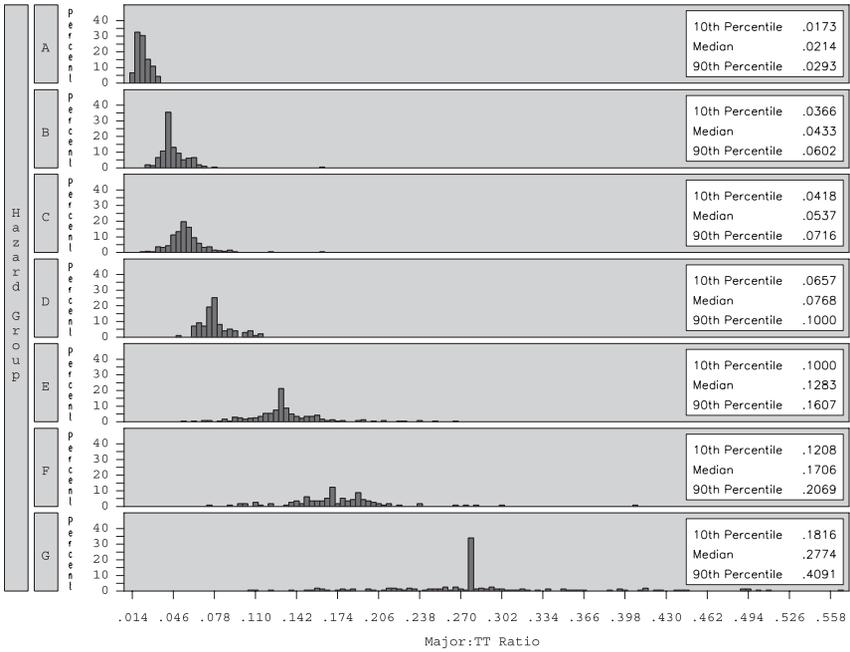


FIGURE 2: Distribution of Predicted Major:TT Ratios by Hazard Group.

for minor as well (not shown). However Table B shows that raw class data does poorly in predicting individual class ratios for all injury types.

Thinking in R^2 terms, the class relativities could be said to “explain” 98% of the “between quintiles” variance. This is not actually a regression, but the statistic is still impressive by real-life actuarial standards. The use of class relativities significantly improves the class frequency by injury type estimation.

The predicted mean ratios by class are distributed around the hazard-group ratios. In some cases these distributions are fairly tight, and in other cases fairly dispersed. Figure 1 illustrates this for PT. Figure 2 is the comparable graph for Major.

In both cases the later hazard groups have higher means and greater dispersion of classes about the mean.

3. CONCLUSIONS

Individual class experience contains information relevant to future large loss relative frequency. A correlated credibility approach using the relationships among injury-type frequencies within each class can utilize this information.

REFERENCES

- BLACK, F., JENSEN, M. and SCHOLES, M. (1972) “The Capital Asset Pricing Model: Some Empirical Tests.” In *Studies in the Theory of Capital Markets*, ed. Michael C. Jensen. New York: Praeger.
- BODIE, Z., KANE, A. and MARCUS, A. (1999), *Investments, 4th Edition*, Irwin/McGraw-Hill, (pages 376-378).
- DEAN, C. (2005) “Topics in Credibility Theory”, Education and Examination Committee of the Society of Actuaries, Construction and Evaluation of Actuarial Models Study Note.
- DORWEILER, P. (1934) “A Survey of Risk Credibility in Experience Rating, Presidential Address at Twentieth Anniversary”, *Proceedings of the Casualty Actuarial Society XXI(1)*.
- GILLAM, W. (1992) “Parametrizing the Workers Compensation Experience Rating Plan”, *Proceedings of the Casualty Actuarial Society, LXXIX*, 21-56.
- HACHEMEISTER, C. (1975) “Credibility for Regression Models with Application to Trend”. In: Kahn, P. (Ed.), *Credibility, Theory and Applications*, Academic Press, New York.
- HARWAYNE, F. (1966) “Insurance Cost of Automobile Basic Protection Plan in Relation to Automobile Bodily Injury Costs”, *Proceedings of the Casualty Actuarial Society, LIII*, 122-158.
- VENTER, G. (1985) “Structured Credibility in Applications – Hierarchical, Multidimensional, and Multivariate Models”, *Actuarial Research Clearing House*, Vol. 2.

JOSE COURET

Guy Carpenter, LLC,

One Madison Avenue,

New York, NY 10010

E-Mail: Jose.R.Couret@guycarp.com

GARY G. VENTER

E-Mail: Gary.G.Venter@guycarp.com