Credibility Prediction using Collateral Information

Edward W. Frees and Peng Shi

April 27, 2016

Abstract

In property-casualty insurance ratemaking, insurers often have access to external information which could be manual rates from a rating bureau or scores from a commercial predictive model. Such collateral information could be valuable because the insurer might either not have sufficient rating information nor the predictive modeling expertise to produce an effective score.

This article shows how to blend collateral information with an insurer's own experience for ratemaking in a predictive modeling framework. Bayesian methods are employed to allow analysts to incorporate their personal knowledge about the precision of the external score. Using conjugate priors, we show that closed-form credibility predictions exist for a variety of distributions including the Tweedie family. A simulation study is performed to demonstrate the prediction with collateral information in a variety of hypothetical scenarios. We further apply the proposed approach to an automobile insurance dataset from Massachusetts. Both the simulation and empirical studies demonstrate situations where combining external information with internal company information provides lift in the prediction of out-of-sample policies.

1 Introduction

Dating back at least to papers by Mowbray (1914) and Whitney (1918), credibility has enjoyed a long history in actuarial science. As seen in the Mowbray and Whitney papers, credibility helps to address two important problems:

- Sharing of information among risk classes. It is common for a distinct risk class to lack information or exposure upon which the insurer can adequately base prices. When developing a rate for a specific risk class, it is naturally desirable to use information from related risk classes.
- Infusing collateral information into resulting rates. At times, an insurer may not have sufficient information from a specific risk class nor from related risk classes and so wishes to incorporate external ("collateral") information.

A natural tool for incorporating collateral information in a disciplined manner is through the use of Bayesian methods (cf., Norberg (1979), page 202, and Jewell (1975)). Bailey (1950a,b) introduced the Bayesian model into credibility theory and showed the equivalence between the Bayesian predictive mean and traditional credibility pricing formulas in specific cases. As noted by these authors, the Bayesian paradigm not only readily permits the sharing of information among risk classes but also allows the analyst to incorporate collateral information.

This result was considerably generalized by Jewell (1974) who showed that linear credibility estimators can be achieved through the use of conjugate priors in linear exponential families. Subsequently, Dannenburg et al. (1996) demonstrated how to incorporate variable weights and Ohlsson and Johansson (2006) extended it to allow parameters to vary by policyholder as one would observe in an insurance portfolio. Moreover, Ohlsson and Johansson (2006) gave specific results for the Tweedie family, a special case of considerable interest in insurance applications and this study. In this work, we also incorporate collateral information in a GLM context and so extend this line of research. In some sense, this paper is a dual application to that of Ohlsson (2008) who also relied on Ohlsson and Johansson (2006) but focused on the risk sharing aspects of credibility.

The rest of the article is structured as follows: Section 2 motivates our modeling framework and Section 3 introduces the closed-form credibility predictors. A simulation study is performed in Section 4. We demonstrate the value added by the collateral information in a variety of hypothetical scenarios such as different types of score, sample size, and data variability. Section 5 applies the proposed approach to an automobile insurance dataset from Massachusetts. Section 6 concludes the paper and technical details are summarized in Section 7.

2 Motivation

In this article, we consider cross-sectional sampling. So, think of each policyholder being observed once and policyholders' claims experience as being unrelated. (In the Appendix Section 7, we more precisely assume that claims are independent, conditional on uncertainties in the collateral information introduced in the following.) For the ith policyholder, use y_i to denote the dependent variable (claim) and $\mathbf{x}_i$ to denote a vector of explanatory (covariate) variables that provide rating information about the policyholder. For this sampling scheme, the information for the model development or training data set is $\{(y_i, \mathbf{x}_i), i = 1, \ldots, n\}$ that we label as Y_{Train}. In a similar way,

the information for the model validation set is $\{(y_i, \mathbf{x}_i), i = n+1, \ldots, n+n_{Valid}\}$ that we label as Y_{Valid}.

Base Case - No Covariate Information

In the base model, the insurer's ith policyholder has claim y_i with mean μ_i. The mean is not observed and may be estimated by policyholder covariates. Before introducing covariates, we assume that the insurer has available an observed score or "manual premium" that is provided by an external agency. Denote the score as $\mu_{\alpha,i}$ and relate it to the mean as a multiplicative effect

$$\mu_i = \alpha \times \mu_{\alpha,i}.$$

The external score is an estimate of the true mean and we use α to denote the corresponding relative error. The variable α may vary by policyholder or risk class. For the moment, we omit the subscript on this term.

From a frequentist perspective, one can think about the term α as a measurement error induced by the score. It is well known in the statistical literature that ignoring this aspect can induce bias in all model coefficients, cf., Carroll et al. (2012). We utilize a Bayesian framework and interpret the distribution of $\{\alpha\}$ as representing the knowledge that the actuary has of the score. Before seeing any data, we assume unbiased scoring and so the mean of the prior distribution is one. This distribution may be subjective and allows the analyst a formal mechanism to inject his or her assessments into the model.

With the training sample Y_{Train} , it will be straightforward to use Bayesian procedures to directly form an estimate of the scoring procedure bias as $\mathrm{E}(\alpha|Y_{Train})$. Then, for a new policy in Y_{Valid}, we are able to form a prediction using $\mathrm{E}(y_i|Y_{Train}) = \mu_{\alpha,i} \times \mathrm{E}(\alpha|Y_{Train})$ for i in $\{n+1, \ldots, n+n_{Valid}\}$.

Of course, it is certainly possible to focus on the random mean μ_i, that is, using Bayesian procedures to update μ_i directly. As will be seen, creation of a new variable α allows us to decompose the random aspect of the uncertainty from other portions, such as covariate information. Moreover, this decomposition will facilitate interpretability as we seek to combine different categorical (factor) random effects.

Introducing Covariates

As a next step, we assume that the insurer has one or more covariates that could be included in the model. For example, thinking of y_i representing the claim on the ith personal automobile policy, the insurer knows whether or not the policyholder also owns a homeowners policy ($x_i = 1$ if yes and $= 0$ otherwise). The insurer could incorporate this information into the collateral score model using the representation

$$\ln \mu_i = \ln \alpha + \ln \mu_{\alpha,i} + x_i \beta,$$

where β is a parameter to be estimated. For this motivation section, we use a logarithmic link function. More generally, the insurer could have a set of covariates that could be included through the representation

$$\ln \mu_i = \ln \alpha + \ln \mu_{\alpha,i} + \mathbf{x}_i' \boldsymbol{\beta}. \tag{1}$$

Here, $\mathbf{x}_i' = (x_{i1}, \ldots, x_{iK})$ represents a set of K explanatory variables and $\boldsymbol{\beta}$ is the corresponding set of parameters. We interpret the term $\mathbf{x}_i'\boldsymbol{\beta}$ as representing the effects of the insurer's portfolio on claims (e.g., insurer underwriting standards).

Similar to the base case, the training sample Y_{Train} is $\{(y_i, \mathbf{x}_i), i = 1, \ldots, n\}$ and Section 3 will show how to use Bayesian procedures to estimate $\mathrm{E}(\alpha|Y_{Train})$. Based on Y_{Train}, we may estimate the regression coefficients $\boldsymbol{\beta}$ (say, $\mathbf{b}$). Then, we will be able to form a prediction using $\mathrm{E}(y_i|Y_{Train}) = \mathrm{E}(\alpha|Y_{Train}) \times \mu_{\alpha,i} \times \exp(\mathbf{x}_i'\mathbf{b})$ for i in $\{n+1, \ldots, n+n_{Valid}\}$.

Multiple Scores

As a variation, we can imagine a situation where there is more than one set of collateral information. Suppose that we have two sets of collateral scores with their associated uncertainties, $\alpha_1 \times \mu_{1,\alpha,i}$ and $\alpha_2 \times \mu_{2,\alpha,i}$. If we are unsure how to combine them in our claims model, then it would be sensible to use (unknown) scaling factors γ_1 and γ_2 and consider a variation of equation (1),

$$\begin{aligned}
\ln \mu_i &= x_{i1}\beta_1 + \cdots + x_{iK}\beta_K + \gamma_1 \ln(\alpha_1 \mu_{1,\alpha,i}) + \gamma_2 \ln(\alpha_2 \mu_{2,\alpha,i}) \\
&= x_{i1}\beta_1 + \cdots + x_{iK}\beta_K + \gamma_1 \ln \mu_{1,\alpha,i} + \gamma_2 \ln \mu_{2,\alpha,i} + \gamma_1 \ln \alpha_1 + \gamma_2 \ln \alpha_2 \\
&= \tilde{\mathbf{x}}_i'\tilde{\boldsymbol{\beta}} + \ln \tilde{\alpha},
\end{aligned}$$

where $\tilde{\mathbf{x}}_i' = (x_{i1}, \ldots, x_{iK}, \ln \mu_{1,\alpha,i}, \ln \mu_{2,\alpha,i})$ is a set of known covariates, $\tilde{\boldsymbol{\beta}} = (\beta_1, \ldots, \beta_K, \gamma_1, \gamma_2)'$ is a set of variables to be estimated, and $\ln \tilde{\alpha} = \gamma_1 \ln \alpha_1 + \gamma_2 \ln \alpha_2$ is a random source of uncertainty. In a similar way, incorporating multiple sets of collateral scores is simply a special case of equation (1).

If the scaling factors γ_1 and γ_2 are known, then we can again use equation (1) but with uncertainty $\ln \tilde{\alpha} = \gamma_1 \ln \alpha_1 + \gamma_2 \ln \alpha_2$ and offset $\ln \mu_{\alpha,i} = \gamma_1 \ln \mu_{1,\alpha,i} + \gamma_2 \ln \mu_{2,\alpha,i}$.

Introducing Multiple Sources of Collateral Information

The collateral information may also contain multiple sources, each representing a different type of uncertainty. For example, returning to the personal automobile example, one can imagine one set of uncertainties (α_1) for retirees and another set (α_2) for all other drivers. In general, we will assume that there are q sets of uncertainties represented as $\boldsymbol{\alpha}^* = (\ln \alpha_1, \ldots, \ln \alpha_q)'$ that feed into a claims model as

$$\ln \mu_i = \ln \mu_{\alpha,i} + \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{z}_i'\boldsymbol{\alpha}^*. \tag{2}$$

Here, $\mathbf{z}_i' = (z_{i1}, \ldots, z_{iq})$ represents a set of q explanatory variables for a linear allocation of the appropriate sources of collateral information to the ith policyholder.

Factor Random Effects

To further develop intuition, think of the special case where we have split up the rating schedule into q categories, where q may range in the hundreds (for example, in personal auto, one can think of many combinations of age, gender, territory, and so forth). Now, $\boldsymbol{\alpha}^*$ represents a categorical factor so that each z_{ij} is a binary variable assigning the ith observation to the jth level of the factor. Standard mappings (e.g., Frees (2010), Section 4.7) allow one to readily go from regression notation, where we distinguish observations using i, to a one-factor notation, where we distinguish observations using ij. In the one factor notation, there are $i = 1, \ldots, c$ factors and, for the jth

factor, there are $j = 1, \ldots, n_j$ observations, for a total of $n_1 + \cdots + n_c = n$ observations. Henceforth, we use the factor notation.

3 Credibility Prediction

To recapitulate, we assume that the claims distribution is a component of a generalized linear model (GLM). For the ith policyholder that is in the jth level of the factor, we specify a conditional mean

$$\mathrm{E}(y_{ij}|\boldsymbol{\alpha}) = \alpha_j \ \times \ \mu_{\alpha,ij} \ \times \ \exp\left(\mathbf{x}_{ij}'\boldsymbol{\beta}\right) \tag{3}$$

where α_j reflects the uncertainty about the score, $\mu_{\alpha,ij}$ is the (externally) provided score, and $\exp\left(\mathbf{x}_{ij}'\boldsymbol{\beta}\right)$ represents insurer-specific adjustments reflecting covariate effects. Our prior belief is that the scoring procedure is unbiased and so we assume that $\mathrm{E}(\alpha_j) = 1$. Thus, the (unconditional) mean is $\mu_{ij} = \mu_{\alpha,ij} \exp\left(\mathbf{x}_{ij}'\boldsymbol{\beta}\right)$. Although our theory allows μ_{ij} to be a (smooth) nonlinear function of covariates, in practice we often specify a logarithmic link function used in equation (3).

We next specify a prior distribution to reflect the uncertainty of the collateral information summarized in α_j. Fortunately, modern-day computational methods permit a wide scope of alternative choices using, for example, Markov Chain Monte Carlo (MCMC) methods (see, for example, Hartman (2014)). With a prior distribution, it is straightforward to calculate the marginal distribution of y by integrating over the prior and then use maximum likelihood to estimate the parameters of the conditional outcome distribution that include $\boldsymbol{\beta}$.

Although it is possible to calculate posterior means using MCMC techniques for general prior distributions, it is also desirable to specify distributions where closed-form expressions are available. In the Appendix Section 7, we consider several exponential families including the normal, Poisson, gamma, inverse Gaussian, and Tweedie. This section generalizes the work of Ohlsson and Johansson (2006) who focussed on the Tweedie distribution (that includes the Poisson and gamma cases). For each family, we specify a natural conjugate prior density in Appendix equation (9) with mean $\mathrm{E}\ \alpha_j = 1$ and dispersion parameter ϕ_α. For example, in the case of the Tweedie distribution, the dispersion parameter associated with the natural conjugate prior is $\phi_\alpha = \mathrm{Var}(\alpha_j)/\mathrm{E}\ \alpha_j^p$.

In these special cases, we have an explicit expression for the posterior mean of the form

$$\mathrm{E}(\alpha_j|Y_{Train}) = \zeta_j + (1-\zeta_j)\overline{(y/\mu)}_{Wj}, \tag{4}$$

where ζ_j is a credibility factor

$$\zeta_j = \frac{\phi}{\phi + \phi_\alpha W_j}, \tag{5}$$

that is determined by the sum of weights within the jth factor, $W_j = \sum_{i:z_{ij}=1} b_2(\mu_{ij})$, and $\overline{(y/\mu)}_{Wj} = \sum_{i:z_{ij}=1}(y_{ij}/\mu_{ij})b_2(\mu_{ij})/W_j$, a weighted average. Here, z_{ij} is a binary variable that is one if the ith policyholder is in the jth risk category. The parameter ϕ and the function $b_2(\cdot)$ depend on the choice of the outcome (claims) distribution. For example, Appendix Table 9 shows that $b_2(\mu_{ij}) = \mu_{ij}^{2-p}$ for the Tweedie distribution.

Equations (4) and (5) have pleasing interpretations that are common in credibility expressions. On the one hand, the credibility factor ζ_j tends to one as either $\phi_\alpha \to 0$ or $\phi \to \infty$. In either case,

we think of the uncertainty associated with the score being very (increasingly) small relative to the dispersion in the outcome distribution. On the other hand, the credibility factor ζ_j tends to zero as either $\phi \to 0$ or $W_j \to \infty$. Intuitively, the credibility (of the score) is small with high precision data or as the number of observations in the jth level of the factor becomes large, indicating substantial information content in the data. Additional details are in Appendix Section 7.

As yet another special case, suppose that the uncertainty grouping is sufficiently refined so that the covariates are constant within the uncertainty group. In this case, μ_{ij} is a constant over the set $\{i : z_{ij} = 1\}$ and equal to, say, μ_j. Then, the weight is a constant times the number of observations in group j, $W_j = n_j \times b_2(\mu_j)$, the weighted average becomes a simple average $\overline{(y/\mu)}_{Wj} = \sum_{i:z_{ij}=1} y_{ij}/(n_j\mu_j) = \bar{y}_j/\mu_j$, and the credibility factor reduces to

$$\zeta_j = \frac{\phi}{\phi + \phi_\alpha n_j b_2(\mu_j)}. \tag{6}$$

For predicting claims from the validation sample, we can calculate an estimator of $\mathrm{E}(y_{ij}|Y_{Train})$ for i in $n+1, \ldots, n+n_{Valid}$. Using estimates based on Y_{Train}, the predictor for the ith policyholder in the validation sample is

$$\left(\zeta_j + (1-\zeta_j)\overline{(y/\mu)}_{Wj}\right) \ \mu_{\alpha,ij} \ \exp\left(\mathbf{x}_{ij}'\mathbf{b}\right). \tag{7}$$

Illustration

To get a better handle on the credibility factors, assume that you wish to apply the credibility factors in equation (5). How does one think about the distribution of the uncertainty of scores, α_j? We know that the expected uncertainty is one ($\mathrm{E}\ \alpha_j = 1$) and that the dispersion parameter, at least in the Tweedie case, is close to the variance of α_j (specifically, $\phi_\alpha = \mathrm{Var}\ \alpha_j / \mathrm{E}\ \alpha_j^p$).

To give a better sense of the dispersion parameter prior distribution, Figure 1 compares prior distributions over different values of dispersion parameters. This prior distribution corresponds to a Tweedie distribution with shape parameter $p = 1.5$. From the figure, we see that large values of ϕ_α mean that the distribution is fatter tailed and right-skewed. Conversely, a relatively small value of ϕ_α (equal to 0.01) gives a distribution that appears to be approximately normally distributed.

How does the prior distribution affect the credibility factor? To give insights into this question, we turn to the special case in equation (6) where covariates are constant within the uncertainty group. As a benchmark, we consider the parameters of a Tweedie regression model that will be described in detailed in Section 4. Specifically, we assume a Tweedie distribution with parameters $p = 1.5$, $\phi = 1,087.709$ and $\mu = 211.87$.

For these parameter choices and the credibility factor in equation (6), Figure 2 compares credibility factors over several dispersion parameters and group sizes. As anticipated, smaller values of ϕ_α mean that we have more confidence in the (external) score and so the credibility factor is closer to one. Further, larger group sizes mean that we have more confidence in the posterior mean so that the credibility factor is lower. Note that in this paper, the credibility factor measures the amount of belief in the prior mean, not in the data. We could have easily defined the credibility factor in terms of its complement (1-ζ); however, our goal is to emphasize the credibility of the collateral (prior) information, not the data.

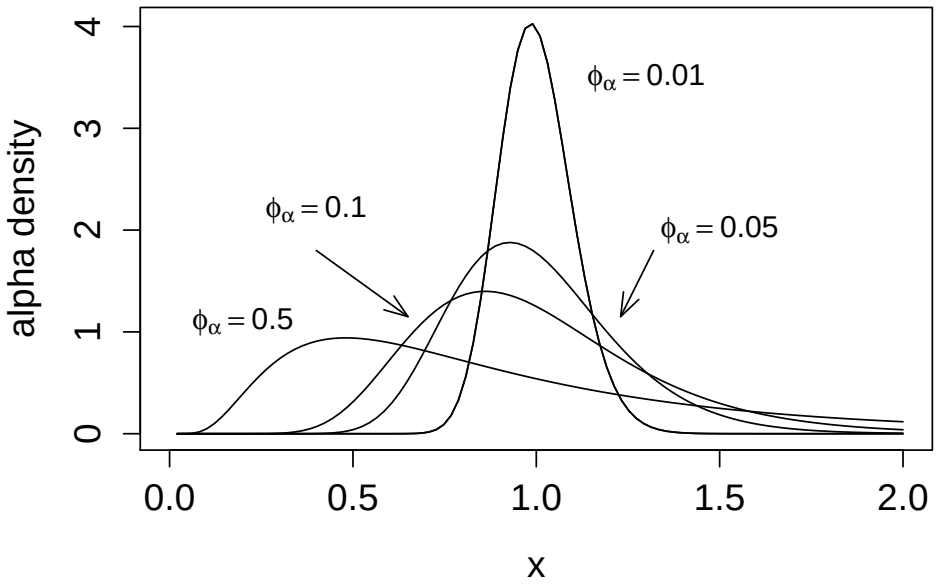

Figure 1: Prior Distributions for Different Dispersion Parameters ϕ_α

4 Simulation Study

4.1 Simulation Design

For the simulation study, we consider a situation where an insurer has a portfolio of policies in a ratemaking development year. The insurer has policyholder characteristics (xs) and claims (ys) upon which rating predictors can be developed. The insurer also contacts an external agency that provides one or more scores based on the characteristics in the insurer's portfolio in the development year. These scores can be used by the insurer to produce modified rating predictors. An insurer then compares the alternative predictors using a new data set, "out-of-sample testing." As part of the testing procedure, the external agency also provides scores based on the characteristics of the new out-of-sample policyholders.

Prior to giving the comparison results, this section describes the data generating process, alternative scores provided, and the rating predictors.

4.1.1 Data Generating Process

We simulated a portfolio of policies and claims experience based on a sample of Massachusetts automobile experience reported in Frees (2014). Table 1 provides the policyholder distribution by two rating factors, an age-based rating group and territory. Thinking of an insurer's experience in a specific state or province, we consider sample sizes of 2,000 and 10,000 policyholders. For

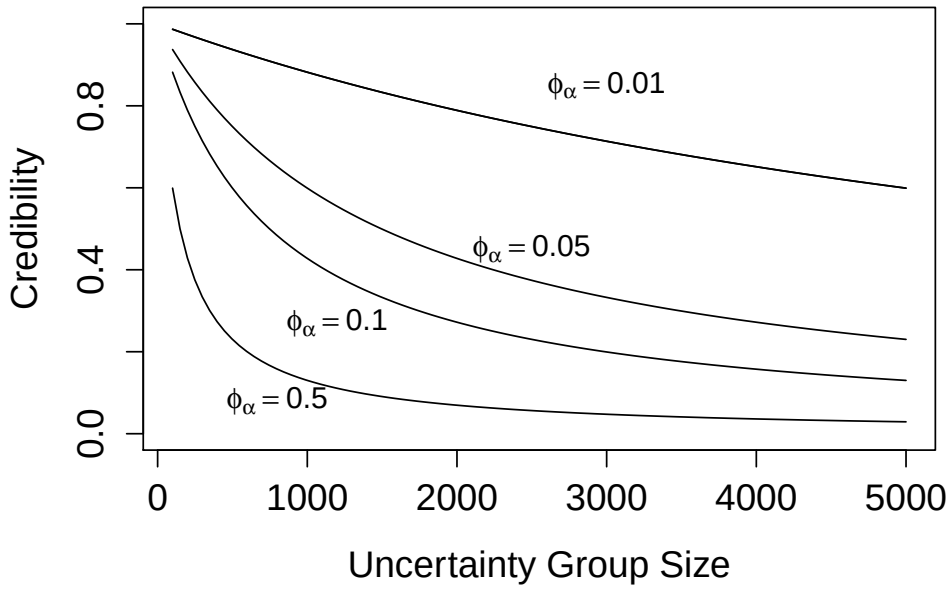

Figure 2: Credibility Factors by Uncertainty Parameters ϕ_α and Group Size

each situation, we generated a portfolio of policyholders using the distribution reported in Table 1 for the model or ratemaking development (in-) sample, and another data set of the same size for out-of-sample testing.

For each policyholder, we simulated claims using the Tweedie distribution with parameters reported in Table 2. We used a logarithmic link function with scale parameter $p = 1.5$ and initial choice of a dispersion parameter $\phi = 250$. For example, suppose that we wish to estimate expected claims for an Adult driver in Territory 6. Because these are the reference levels, the estimate is $\exp(5.356) = 211.87$. The corresponding probability of zero claims is $\exp\{-\mu_{ij}^{2-p}/(\phi(2-p))\} = 89.0\%$. Use $\boldsymbol{\beta}_0$ to denote the vector of parameters in Table 2.

Table 1: Proportion of Policies by Rating Group and Territory

Rating Group	Proportion	Territory	Proportion
A – Adult	0.76616	1	0.18410
B – Business	0.01269	2	0.19360
I – Youthful with less	0.03453	3	0.11245
than 3 years Experience		4	0.20300
M – Youthful with	0.04190	5	0.18921
3-6 years Experience		6	0.11764
S – Senior Citizens	0.14472		

Table 2: Tweedie GLM Coefficients

Rating Group	Estimate	Territory	Estimate
B	0.340	1	-0.743
I	1.283	2	-0.782
M	0.474	3	-0.552
S	-0.033	4	-0.480
		5	-0.269

Intercept is 5.356
Reference levels are "A" for Rating Group and "6" for Territory

4.1.2 Alternative Scores

A score is provided by an external agency that is based on policyholder characteristics. The best (although unattainable in practice) score is the mean, $\exp(\mathbf{x}'\boldsymbol{\beta}_0)$, that we label `ScoreTrue`.

To derive alternative scores, we assume that the external agency has other data that follows the same distribution as in Section 4.1.1. Ideally, the analyst for the external agency (i) works with a large data set, (ii) employs an extensive set of covariates, and (iii) uses modern (appropriate) statistical methods. To assess these alternatives, we provide eight alternative scores that vary by:

- sample size, either a relatively large sample size (`LS`) 100,000 or a small sample size (`SS`) 10,000,
- number of covariates, either including both age and territory (`Full`) or a reduced set, only age, (`Red`), and
- statistical methods, either a GLM using a Tweedie distribution (`GLM`) or a linear model (`LM`).

Thus, for example, `LS_Full_GLM` denotes a score that the analyst derives using a sample size of 100,000, both age and territory covariates, and a GLM representation. As another example, `SS_Red_LM` denotes a score derived using a sample size of 10,000, only the age covariate, and a linear model.

Scores are calibrated from data and so are subject to estimation error. To generate the scores, we used $q = 6$ risk categories corresponding to different territories. We use the conjugate prior distribution described in Section 7 with parameter $\phi_\alpha = 0.01$. This corresponds to a standard deviation of α_j of approximately 0.1.

4.1.3 Rating Predictors

The analyst for the insurer has many choices of rating predictors. First, one option is to use only company experience, ignoring any scores provided by an external agency. We will assume that the insurer analyst is only using GLM representations but, like the external analyst, may be working with a limited set of covariates. Specifically, we distinguish between the cases when the analyst has a full set of covariates, including both age and territory (`Full`), and a reduced set, only age, (`Red`).

Second, another option is to use only the score provided by the external agency, ignoring company information. As described in Section 4.1.2, there are eight such scores available, in addition to the baseline true score.

A third option is for the analyst to use the externally provided score as an offset in a GLM model and then incorporate additional company covariates as available. We also considered including the externally provided score as a variable in a GLM model together with company covariates. No real insights were garnered from this alternative option and so we do not describe results here.

Fourth, the analyst may incorporate the externally provided score as an offset, use company covariates, and modify the predictors based on the insurer's belief in the scores using the predictors described in Section 3. For our work, we allow the belief parameter to vary over $\phi_\alpha = 0.5, 0.1, 0.01, 0$.

4.1.4 Out-of-Sample Summary Measures

We choose seven criteria to measure how each model performs in terms of out-of-sample prediction. The first three statistics are standard out-of-sample validation statistics, e.g. Frees (2010); they measure how far away the predicted values deviate from the observed values in the hold-out sample. Thus, the smaller the numbers, the better are the predictions. The mean absolute (percentage) error computes the average of the (percentage) absolute error between the prediction and the observed value; the root mean square error is the square root of the average squared distance between the prediction and the observed values.

The next three statistics measure the correlation between predicted values and observed values in the hold-out sample. The larger the numbers, the better are the predictions. Recall that the Pearson correlation is obtained by dividing the covariance of two variables by their standard deviations and the Spearman correlation is defined as the Pearson correlation coefficient of the ranked variables. That is, the original values need to be converted to ranks, and Spearman correlation is less sensitive than Pearson correlation to outliers in the tails of both samples. The Gini coefficient is a newer measure developed in Frees et al. (2013). It measures the correlation between the rank of predictor and the corresponding outcome from a hold-out sample. The seventh statistic is a Gini index that corresponds closely to the correlation coefficient. This Gini index is twice the average covariance between the the outcome in the hold-out-sample and the rank of the predictor. All correlations are reported on a percentage scale, that is, multiplied by 100.

4.2 Simulation Results

Table 3 summarizes the out-of-sample statistics for our credibility rating predictors. For ease of comparison, Panel A presents the results in the case $\phi_\alpha = 0.0$ so there is no presumed uncertainty associated with the score. Even in Panel A we see that the mean absolute error and the mean absolute percentage error give non-intuitive results. In each case, going from the "Full" set of covariates to the "Red" (reduced), they actually increase, indicating a poorer ability to predict. The root mean square error statistic fares better although still does not provide the type of separation that one would like to see. Although useful for many applications, because our outcome claims variable contains many zeros and, when positive, has a skewed distribution, these traditional measures are less useful for rating analyses.

Panels B, C, and D of Table 3 present results as the analyst changes his or her belief about the score's precision. As one increases the value of ϕ_α, one places less credibility on the score and more on the data. Interestingly, when even acknowledging a small imprecision in the score, the case $\phi_\alpha = 0.01$, the Gini correlation increases from approximately 5.22 to 7.48 even for scores that use only a reduced set of covariates. This is because we selected the risk classes to correspond to the set of information, territory, that is "missing" in both the company's covariates and external

agency covariates. By averaging over these risk classes in one period, the insurer has a very useful nonparametric predictor of claims in the next period. Results for the traditional Pearson and Spearmen coefficients are consistent with the newer Gini correlations. Because of the literature cited above, we focus henceforth on the newer Gini measure.

Tables 4 and 5 present out-of-sample summary statistics for small sample and small dispersion cases. For these alternative cases, we focus on the Gini correlations. Table 4 provides similar information compared to Table 3 except we now assume that the insurer has 2,000 policyholders for in-sample analysis and out-of-sample testing. The scoring procedures by the external agency remain the same. The table shows that the results for this smaller sample size are consistent with those in Table 3 except that now the belief parameter ϕ_α must be larger, placing more emphasis on the data, in order to overcome a poor score.

Table 5 returns to the scenario of 10,000 policyholders available in- and out-of-sample yet now consider the case where the outcome dispersion parameter ϕ reduces from 250 to 100. Because of this reduction in dispersion, all correlations are larger than the corresponding elements in Table 3 yet the conclusions remain essentially the same.

Table 3: Out-of-Sample Statistics for Credibility Predictors

$n = 10,000, \quad \phi = 250$

With Company Experience Adjustment, Reduced Covariates

	Mean Absolute Error	Mean Absolute Perc Error	Root Mean Square Error	Correlations*			Simple Gini
				Pearson	Spearman	Gini	
Panel A. $\phi_\alpha = 0.0$							
LS_Full_GLM	278.920	184.216	721.825	11.031	5.400	7.837	32.680
LS_Red_GLM	279.620	184.127	723.020	9.427	3.368	5.217	21.688
LS_Full_LM	279.301	184.569	722.058	10.680	5.425	7.867	32.802
LS_Red_LM	279.690	184.200	723.084	9.317	3.378	5.224	21.715
SS_Full_GLM	279.427	184.638	722.003	10.865	5.241	7.648	31.890
SS_Red_GLM	279.620	184.127	723.020	9.427	3.368	5.217	21.688
SS_Full_LM	280.002	287.171	722.278	10.485	5.253	7.660	31.937
SS_Red_LM	279.716	216.308	723.105	9.292	3.379	5.226	21.725
Panel B. $\phi_\alpha = 0.01$							
LS_Full_GLM	288.857	182.296	723.061	10.786	5.431	7.861	32.776
LS_Red_GLM	283.796	180.882	722.262	10.541	5.104	7.483	31.206
LS_Full_LM	289.411	182.707	722.882	10.587	5.469	7.909	32.978
LS_Red_LM	283.846	180.969	722.254	10.506	5.115	7.496	31.258
SS_Full_GLM	289.465	182.720	723.334	10.687	5.349	7.763	32.368
SS_Red_GLM	283.796	180.882	722.262	10.541	5.104	7.483	31.206
SS_Full_LM	290.391	253.368	723.280	10.438	5.370	7.795	32.499
SS_Red_LM	283.895	212.954	722.275	10.483	5.113	7.492	31.241
Panel C. $\phi_\alpha = 0.1$							
LS_Full_GLM	298.500	183.709	726.078	10.416	5.367	7.744	32.286
LS_Red_GLM	287.745	181.372	722.473	10.749	5.297	7.704	32.121
LS_Full_LM	299.213	184.182	725.374	10.221	5.414	7.813	32.573
LS_Red_LM	287.779	181.471	722.389	10.738	5.318	7.734	32.248
SS_Full_GLM	299.200	184.156	726.475	10.351	5.324	7.695	32.082
SS_Red_GLM	287.745	181.372	722.473	10.749	5.297	7.704	32.121
SS_Full_LM	300.478	244.978	726.083	10.108	5.355	7.739	32.264
SS_Red_LM	287.851	215.296	722.413	10.720	5.312	7.729	32.224
Panel D. $\phi_\alpha = 0.5$							
LS_Full_GLM	300.363	184.403	726.854	10.349	5.352	7.716	32.170
LS_Red_GLM	288.482	181.870	722.622	10.735	5.305	7.708	32.136
LS_Full_LM	301.107	184.891	726.036	10.148	5.400	7.787	32.467
LS_Red_LM	288.514	181.972	722.521	10.724	5.327	7.739	32.266
SS_Full_GLM	301.080	184.858	727.278	10.288	5.311	7.671	31.979
SS_Red_GLM	288.482	181.870	722.622	10.735	5.305	7.708	32.136
SS_Full_LM	302.428	244.757	726.824	10.039	5.344	7.718	32.178
SS_Red_LM	288.590	216.335	722.547	10.706	5.325	7.737	32.258

Legend: LS means large sample, SS means small sample

Full means full set of covariates, Red means reduce set of covariates

GLM means generalized linear model, LM means linear model

*All correlations are reported on a percentage scale, that is, multiplied by 100.

Table 4: Out-of-Sample Gini Correlations for Credibility Predictors - Small Sample
$n = 2,000, \quad \phi = 250$
With Company Experience Adjustment, Reduced Covariates

	Gini Correlation		Gini Correlation
$\phi_\alpha = 0.0$		$\phi_\alpha = 0.01$	
ScoreTrue	7.307		
LS_Full_GLM	7.224	LS_Full_GLM	7.049
LS_Red_GLM	4.345	LS_Red_GLM	5.423
LS_Full_LM	7.276	LS_Full_LM	7.095
LS_Red_LM	4.395	LS_Red_LM	5.467
SS_Full_GLM	7.036	SS_Full_GLM	6.889
SS_Red_GLM	4.345	SS_Red_GLM	5.423
SS_Full_LM	7.091	SS_Full_LM	6.930
SS_Red_LM	4.397	SS_Red_LM	5.468
$\phi_\alpha = 0.10$		$\phi_\alpha = 0.50$	
LS_Full_GLM	6.889	LS_Full_GLM	6.759
LS_Red_GLM	6.194	LS_Red_GLM	6.239
LS_Full_LM	6.959	LS_Full_LM	6.834
LS_Red_LM	6.229	LS_Red_LM	6.279
SS_Full_GLM	6.807	SS_Full_GLM	6.676
SS_Red_GLM	6.194	SS_Red_GLM	6.239
SS_Full_LM	6.843	SS_Full_LM	6.743
SS_Red_LM	6.230	SS_Red_LM	6.276

Legend: LS means large sample, SS means small sample
Full means full set of covariates, Red means reduce set of covariates
GLM means generalized linear model, LM means linear model

Table 5: Out-of-Sample Gini Correlations for Credibility Predictors - Small Dispersion
$n = 10,000, \quad \phi = 100$
With Company Experience Adjustment, Reduced Covariates

	Gini Correlation		Gini Correlation
$\phi_\alpha = 0.0$		$\phi_\alpha = 0.01$	
ScoreTrue	12.650		
LS_Full_GLM	12.486	LS_Full_GLM	12.413
LS_Red_GLM	7.964	LS_Red_GLM	12.196
LS_Full_LM	12.512	LS_Full_LM	12.521
LS_Red_LM	7.974	LS_Red_LM	12.223
SS_Full_GLM	12.382	SS_Full_GLM	12.370
SS_Red_GLM	7.964	SS_Red_GLM	12.196
SS_Full_LM	12.403	SS_Full_LM	12.474
SS_Red_LM	7.973	SS_Red_LM	12.219
$\phi_\alpha = 0.10$		$\phi_\alpha = 0.50$	
LS_Full_GLM	12.290	LS_Full_GLM	12.271
LS_Red_GLM	12.318	LS_Red_GLM	12.321
LS_Full_LM	12.401	LS_Full_LM	12.384
LS_Red_LM	12.361	LS_Red_LM	12.368
SS_Full_GLM	12.252	SS_Full_GLM	12.236
SS_Red_GLM	12.318	SS_Red_GLM	12.321
SS_Full_LM	12.366	SS_Full_LM	12.351
SS_Red_LM	12.358	SS_Red_LM	12.364

Legend: LS means large sample, SS means small sample
Full means full set of covariates, Red means reduce set of covariates
GLM means generalized linear model, LM means linear model

5 Massachusetts Automobile Claims

In this example, we consider a database of personal automobile claims from the Commonwealth Automobile Reinsurers (CAR) in Massachusetts, described in Ferreira Jr and Minikel (2010). The CAR is a statistical agent for motor vehicle insurance in the Commonwealth of Massachusetts and collects insurance data for both private passengers and commercial automobiles in the state. In Massachusetts, individuals who drive a car must purchase third party liability (property damage and bodily injury) and personal injury protection (PIP) coverage for their personal vehicle.

The database summarizes experience of over three million policyholders in year 2006. For each policy, we observe the number of claims, the type of claim for each accident, as well as the total payments associated with each type during the year. Besides the claim information, the data also contain basic risk classification variables. Because the dataset represents experience from several insurance carriers, we only have access to a limited number of common rating variables reported to the bureau.

We take a random sample of 100,000 policyholders for this study. The first 50,000 observations are used as training data to develop the model, and the rest are reserved as hold-out for validation. Table 6 displays the description and summary statistics of the rating factors for the training data.

As described in Ferreira Jr and Minikel (2010), *Rating Group* indicates policyholder characteristics and *Territory Group* indicates the risk level of the driving area defined by the garage town. *Rating Group* is constructed from finer-grained driver classes. *Territory Group* is constructed based the Automobile Insurance Bureau's relativities estimated for the 351 Massachusetts towns and the 10 state-defined regions within the city of Boston. These geographical units are ranked by the estimated risk and then grouped into six territories. Table 6 shows that 77% policyholders are adult drivers and 11% are from the most risky driving territory. The last two columns present the average liability and PIP claims.

Table 6: Description and Summary Statistics of Basic Rating Information[†]

		Pure Premium	
	Mean	Liability	PIP
Rating Group			
A - Adult	0.772	165.371	15.929
B - Business	0.013	173.446	13.924
I - Youth with <3 years experience	0.036	344.686	32.542
M - Youth with 3-6 years experience	0.039	358.413	43.913
S - Senior citizens	0.140	169.500	8.944
Territory Group			
1 - the least risky territory group	0.192	125.979	7.812
2	0.197	163.083	5.226
3	0.114	209.265	11.855
4	0.206	156.951	17.279
5	0.180	202.764	23.189
6 - the most risky territory group	0.110	279.889	45.145

[†] The summary statistics are adjusted by exposure.

To illustrate the value added by the collateral information from external sources, we consider scores produced by ISO Risk Analyzer - a commercial predictive model from Verisk Analytics. Specifically, two sets of scores are used in the credibility prediction, the relativities from the vehicle module for liability and from the environmental module. The former is based on vehicle characteristics and is the focus of this example. The latter captures the effects of granular environmental factors instead of the crude location of garage town indicated by the territory group.

We use the same out-of-sample statistics as introduced in the simulation section. The results for the liability and PIP coverage are reported in Tables 7 and 8, respectively. We compare the credibility predictions with (Panels B-E) and without (Panel A) collateral information - the vehicle module liability score.

In Panel A, the insurer only uses rating variables as covariates in the prediction. We consider three base scenarios representing a range of complexity of predictive models employed by the insurer.

1. A naive insurer, relying on the principal of parsimony, could use a reduced set of covariates - rating group only.

2. A more knowledgable insurer might consider a full set of covariates, including both rating and territory groups in the prediction.

3. In addition to the rough territory information, a sophisticated insurer might incorporate the

more detailed address-specific risk factors, which could be the relativities from the environmental module.

Table 7: Out-of-Sample Statistics for Credibility Predictors of Liability Coverage

	Pearson	Spearman	Gini	Simple Gini
Panel A : Insurer Information				
Rating Group	3.389	3.047	1.931	17.361
Rating Group + Territory Group	4.833	4.883	3.423	41.328
Rating Group + Territory Group + Environmental	4.973	5.535	3.831	46.586
Panel B : $\phi_\alpha = 0$				
Rating Group	3.916	4.343	2.615	31.791
Rating Group + Territory Group	5.276	5.850	3.873	47.101
Rating Group + Territory Group + Environmental	5.355	6.318	4.072	49.518
Panel C : $\phi_\alpha = 0.01$				
Rating Group	4.909	5.702	3.768	45.828
Rating Group + Territory Group	5.223	5.839	3.899	47.414
Rating Group + Territory Group + Environmental	5.299	6.301	4.089	49.725
Panel D : $\phi_\alpha = 0.1$				
Rating Group	5.084	5.702	3.858	46.921
Rating Group + Territory Group	5.146	5.783	3.880	47.181
Rating Group + Territory Group + Environmental	5.224	6.243	4.061	49.390
Panel E : $\phi_\alpha = 0.5$				
Rating Group	5.091	5.707	3.860	46.940
Rating Group + Territory Group	5.132	5.774	3.878	47.162
Rating Group + Territory Group + Environmental	5.210	6.230	4.054	49.302

In Panels B-E, the insurer combines the vehicle liability score with each of the three basic models introduced in Panel A. Consistent with the simulation study, we include the vehicle module relativity as an offset in the Tweedie GLM. We further assume that the measurement error in this external score varies by territory group, i.e. credibility predictions are calculated according to equation (4) where the subscript j refers to territory. The confidence in the score is reflected by parameter ϕ_α with a larger value indicating higher uncertainty. Predictions for $\phi_\alpha = 0, 0.01, 0.1,$ and 0.5 are reported.

We observe similar patterns in both tables and, in general, the out-of-sample statistics suggest that collateral information could improve prediction. First, Panel A suggests that geographic information is an important predictor for this data set. For example, if the insurer is knowledgable enough to use territory group in the prediction, the Gini index increases from 17.36 to 41.33. By further incorporating granular information from the environmental module, one could improve the Gini index to 46.59. Consistent results are also observed in Panels B-E.

Second, the credibility prediction in Panels B-E for the naive insurer reinforces the results observed in Table 3 in the simulation study. Specifically, when even allowing for a small imprecision ($\phi_\alpha = 0.01$) in the score, the Gini index increases from 31.79 to 45.83. This is because the territory group, as suggested by Panel A, is an important risk class indicator for the portfolio of policyholders. Although the naive insurer is not knowledgable to use territory as a covariate, the territory information is brought into the prediction through the factor random effect specification. However, this difference between Panel B and Panels C-E is less prominent for more sophisticated

insurers, because including territory group as a predictor reduces the effect of averaging over these risk classes.

Third, comparing Panels B-E with Panel A, one finds that regardless of the complexity of the predictive model used by the insurer, using the vehicle liability score further improves the prediction. One notices that higher lift is provided by the external score for less sophisticated insurers. For example, the Gini index increases approximately from 41 to 47 for a knowledgable insurer, and from 46 to 49 for a sophisticated insurer. This observation is anticipated because the score is likely to contain more relevant information that is missing in the reduced set of covariates.

Table 8: Out-of-Sample Statistics for Credibility Predictors of PIP Coverage

	Pearson	Spearman	Gini	Simple Gini
Panel A : Inusurer Information				
Rating Group	1.738	2.637	1.707	2.741
Rating Group + Territory Group	4.250	4.439	2.990	6.450
Rating Group + Territory Group + Environmental	4.797	4.475	3.075	6.680
Panel B : $\phi_\alpha = 0$				
Rating Group	1.812	2.872	1.975	4.291
Rating Group + Territory Group	4.217	4.524	3.040	6.605
Rating Group + Territory Group + Environmental	4.667	4.593	3.137	6.817
Panel C : $\phi_\alpha = 0.01$				
Rating Group	2.624	3.935	2.730	5.932
Rating Group + Territory Group	4.233	4.529	3.045	6.617
Rating Group + Territory Group + Environmental	4.677	4.598	3.141	6.824
Panel D : $\phi_\alpha = 0.1$				
Rating Group	4.071	4.691	3.148	6.840
Rating Group + Territory Group	4.280	4.543	3.059	6.646
Rating Group + Territory Group + Environmental	4.708	4.615	3.153	6.851
Panel E : $\phi_\alpha = 0.5$				
Rating Group	4.332	4.722	3.145	6.835
Rating Group + Territory Group	4.302	4.547	3.065	6.661
Rating Group + Territory Group + Environmental	4.723	4.625	3.160	6.867

6 Conclusion

This study was motivated by a practical problem in property casualty insurance ratemaking: How can individual insurers blend external information with their own rating variables to improve prediction? We answered this question by developing credibility predictions in a GLM context using Bayesian methods and hence contributed to the credibility literature on incorporating collateral information.

Collateral information can take different forms. We think of them as a single or multiple scores obtained from external agencies, be it a rating bureau or a proprietary entity. Such a score is conceptually attractive because it could incorporate rating information not available to the insurer or it could be generated through advanced predictive models that the insurer could not afford in house. To adapt the external score to the insurer's own rating scheme, we introduced the uncertainty in the score that is allowed to vary across different risk classes and made statistical

inference for the bias from the data. We employed a Bayesian approach. The advantage of this approach is that it provides a mechanism for the analyst to incorporate his or her prior belief about the uncertainty of the score into the prediction.

Using conjugate priors, we derived close-form credibility predictors for a variety of distributions in the exponential family, with a focus on the Tweedie family in the simulation study and the application of Massachusetts automobile insurance. The Tweedie GLM is commonly used in modeling pure premiums in property-casualty insurance and thus is a natural choice for incorporating external scores such as manual rates or ISO Risk Analyzer relativities. For validation, we noted that the traditional out-of-sample statistics are less useful and emphasized recently developed Gini statistics for measuring the predictive performance.

Acknowledgement

BLINDED

References

Bailey, Arthur L. (1950a). "Credibility procedures, LaPlace's generalization of Bayes' rule and the combination of collateral knowledge with observed data," *Proceedings of the Casualty Actuarial Society*, Vol. 37, pp. 7–23.

——— (1950b). "Credibility procedures, LaPlace's generalization of Bayes' rule and the combination of collateral knowledge with observed data: Discussion," *Proceedings of the Casualty Actuarial Society*, Vol. 37, p. 94:115.

Carroll, Raymond J, David Ruppert, Leonard A Stefanski, and Ciprian M Crainiceanu (2012). *Measurement error in nonlinear models: a modern perspective*. CRC press.

Dannenburg, D. R., Rob Kaas, and Marc J. Goovaerts (1996). *Practical Actuarial Credibility Models*. Amsterdam: Institute of Actuarial Science and Economics, University of Amsterdam.

Ferreira Jr, Joseph and Eric Minikel (2010). "Pay-as-you-drive auto insurance in Massachusetts: a risk assessment and report on consumer, industry and environmental benefits," *Conservation Law Foundation & Environmental Insurance Agency*, p. http://web.mit.edu/jf/www/payd/PAYD_CLF_Study_Nov2010.pdf.

Frees, Edward W. (2010). *Regression modeling with actuarial and financial applications*. Cambridge University Press.

——— (2014). "Frequency and severity models," in Edward W. Frees, Glenn Meyers, and Richard A. Derrig eds. *Predictive Modeling Applications in Actuarial Science*, Cambridge. Cambridge University Press.

Frees, Edward W., Glenn Meyers, and A. David Cummings (2013). "Insurance ratemaking and a Gini index," *Journal of Risk and Insurance*.

Hartman, Brian (2014). "Bayesian computational methods," in Edward W. Frees, Glenn Meyers, and Richard A. Derrig eds. *Predictive Modeling Applications in Actuarial Science*, Cambridge. Cambridge University Press.

Jewell, William S. (1974). "Credible means are exact Bayesian for exponential families," *Astin Bulletin*, Vol. 8, pp. 77–90.

——— (1975). "The use of collateral data in credibility theory: a hierarchical model," *Giornale dell'Istituto Italianao degli Attuari*, Vol. 38, pp. 1–16.

Kaas, Rob, D. Dannenburg, and Marc Goovaerts (1997). "Exact credibility for weighted observations," *ASTIN Bulletin*, Vol. 27, pp. 287–295.

Mowbray, Albert H. (1914). "How extensive a payroll exposure is necessary to give a dependable pure premium?" *Proceedings of the Casualty Actuarial Society*, Vol. I, pp. 25–30.

Norberg, Ragnar (1979). "The credibility approach to experience rating," *Scandinavian Actuarial Journal*, Vol. 1979, pp. 181–221.

Ohlsson, Esbjörn (2008). "Combining generalized linear models and credibility models in practice," *Scandinavian Actuarial Journal*, Vol. 2008, pp. 301–314.

Ohlsson, Esbjörn and Björn Johansson (2006). "Exact credibility and Tweedie models," *Astin Bulletin*, Vol. 36, p. 121.

Whitney, Albert W. (1918). "The theory of experience rating," *Proceedings of the Casualty Actuarial Society*, Vol. IV, pp. 275–293.

7 Appendix: Details of Bayesian Inference for the Generalized Linear Model

To establish notation, we begin with an linear exponential family of the form

$$\mathrm{p}(y,\theta,\phi) = \exp\left(\frac{y\theta - b(\theta)}{\phi} + S(y,\phi)\right), \tag{8}$$

with moments $\mu = \mathrm{E}\ y = b'(\theta)$ and $\mathrm{Var}\ y = \phi b''(\theta)$. Using the link function $\mathrm{g}(\cdot)$, introduce a systematic component $\eta = \mathrm{g}(\mu)$ and $\mu = b'(\theta)$. Further, use $h(z) = (b')^{-1}(z)$ so that $\theta = h(\mathrm{g}^{-1}(\eta))$.

Conjugate Prior Distribution

Now think of θ as a random variable. The natural conjugate prior distribution of θ is

$$f(\theta;\delta,\phi_\theta) = \frac{1}{c(\delta,\phi_\theta)} \exp\left(\frac{\theta\delta - b(\theta)}{\phi_\theta}\right). \tag{9}$$

Some easy calculations show (see Ohlsson and Johansson (2006), Lemma 2.1) that $\delta = \mathrm{E}\ b'(\theta)$. Further, the mode satisfies $\delta = b'(\theta_{mode})$.

We can use the Bayes machinery to update the prior. Using equations (8) and (9), we have

$$\begin{aligned}
\mathrm{f}(\theta|y) &\propto \mathrm{p}(y|\theta) f(\theta) \\
&\propto \exp\left(\frac{y\theta - b(\theta)}{\phi} + \frac{\theta\delta - b(\theta)}{\phi_\theta}\right) = \exp\left(\theta(\frac{y}{\phi} + \frac{\delta}{\phi_\theta}) - b(\theta)(\frac{1}{\phi} + \frac{1}{\phi_\theta})\right) \\
&\propto f(\theta;\delta^* = \frac{\delta\phi + y\phi_\theta}{\phi + \phi_\theta}, \phi_\theta^* = \frac{\phi\phi_\theta}{\phi + \phi_\theta}).
\end{aligned} \tag{10}$$

This interesting and useful result is originally due to Jewell (1974), subsequently extended to include weights by Kaas et al. (1997). Greater focus on the Tweedie distribution was provided in Ohlsson and Johansson (2006).

Multiplicative Random Effects

Assume that the means vary by subject and so use E $y_i = \mu_i$. Assume also that a multiplicative random effect is common to a set of observations and so use $\mathrm{E}(y_i|\alpha) = \alpha\mu_i$. In the GLM notation, we specify $\eta_i = \mathrm{g}(\mathrm{E}(y_i|\tilde{\alpha}))$ and so $\mathrm{g}^{(-1)}(\eta_i) = \alpha\mu_i$. Here, α is a random variable (effect).

To relate this to parameters of linear exponential family, recall that $b'(\theta_i) = \mathrm{g}^{(-1)}(\eta_i) = \alpha\mu_i$. Thus, it useful to define $h(z) = (b')^{(-1)}(z)$ and so write $\theta_i = h(\alpha\mu_i)$. Note that because α is a random variable, so is θ_i.

With these choices, we can express the conditional distribution as

$$\mathrm{f}(y_i; \tilde{\alpha}, \phi, \mu_i) \quad = \quad \exp\left(\frac{y_i(h(\alpha\mu_i)) - b(h(\alpha\mu_i))}{\phi} + S(y_i, \phi)\right). \tag{11}$$

To simplify the expression in equation (11), we would like to be able to write

$$\begin{aligned} h(\alpha\mu) &= h(\alpha)b_2(\mu)/\mu + h_3(\mu) \\ b(h(\alpha\mu)) &= b(h(\alpha))b_2(\mu) + b_3(\mu) \end{aligned} \tag{12}$$

Table 9 provides examples of several distributions where display (12) holds. We note that these are not unique decompositions for each distribution.

Table 9: Exponential Family Distributions Satisfying Equation (12)

Distribution	$b(z)$	$b'(z)$	$h(z)$	$b(h(z))$	$h_3(z)$	$b_2(z)$	$b_3(z)$
Normal	$\frac{z^2}{2}$	z	z	$\frac{z^2}{2}$	0	z^2	0
Poisson	e^z	e^z	$\ln z$	z	$\ln z$	z	0
Gamma	$-\ln z$	$\frac{-1}{z}$	$\frac{-1}{z}$	$-\ln z$	0	1	$-\ln z$
Inverse Gaussian	$-(-2z)^{1/2}$	$(-2z)^{-1/2}$	$\frac{-1}{2z^2}$	$\frac{1}{4z}$	0	$\frac{1}{z}$	0
Tweedie*	$k_1 z^{k_2}$	$k_1 k_2 z^{k_2-1}$	$\left(\frac{z}{k_1 k_2}\right)^{\frac{1}{k_2-1}}$	$k_1\left(\frac{z}{k_1 k_2}\right)^{\frac{k_2}{k_2-1}}$	0	$z^{\frac{k_2}{k_2-1}}$	0

*For Tweedie, one uses $k_2 = (p-2)/(p-1)$ and $k_1 = (1-p)^{k_2}/(2-p)$, for $1 < p < 2, p > 2$

For the prior distribution, define the transformed random effect $\tilde{\alpha} = h(\alpha)$. Assume that $\tilde{\alpha}$ has a (conjugate) density corresponding to (9) with parameters δ and ϕ_α.

With this, $\tilde{\alpha} = h(\alpha)$, equations (11) and (9), we have

$$\begin{aligned}
\mathrm{f}(\tilde{\alpha}|y_i) &\propto \mathrm{p}(y_i|\tilde{\alpha}) f(\tilde{\alpha};\delta,\phi_\alpha) \\
&\propto \exp\left(\frac{y_i h(\alpha\mu_i) - b(h(\alpha\mu_i))}{\phi} + \frac{\tilde{\alpha}\delta - b(\tilde{\alpha})}{\phi_\alpha}\right) \\
&= \exp\left(\frac{y_i(h(\alpha)b_2(\mu_i)/\mu_i + h_3(\mu_i)) - (b(h(\alpha))b_2(\mu_i) + b_3(\mu_i))}{\phi} + \frac{\tilde{\alpha}\delta - b(\tilde{\alpha})}{\phi_\alpha}\right) \\
&\propto \exp\left(\frac{(y_i/\mu_i)\tilde{\alpha}b_2(\mu_i) - b(\tilde{\alpha})b_2(\mu_i)}{\phi} + \frac{\tilde{\alpha}\delta - b(\tilde{\alpha})}{\phi_\alpha}\right) \\
&\propto \exp\left(\tilde{\alpha}\left\{\frac{y_i b_2(\mu_i)}{\mu_i\phi} + \frac{\delta}{\phi_\alpha}\right\} - b(\tilde{\alpha})\left\{\frac{b_2(\mu_i)}{\phi} + \frac{1}{\phi_\alpha}\right\}\right).
\end{aligned} \tag{13}$$

This has the same form as equation (9).

Cross-Sectional Sample

We now combine the conditional outcome distribution over several observations with the prior parameter distribution. We assume that there are q uncertainties and that $\{\alpha_j\}$ are i.i.d. Recall that z_{ij} is a binary variable assigning the ith observation to the jth level of the factor. For the random factor model, observations from different levels of the factor are independent. Thus, we restrict our updating to observations from the same, say, jth level. To this end, consider a set of n_j observations, independent conditional on $\tilde{\alpha}_j = h(\alpha_j)$, with

$$\begin{aligned}
\mathrm{f}(\tilde{\alpha}|\mathbf{y}) &\propto \left(\prod_{i:z_{ij}=1} \mathrm{p}(y_i|\tilde{\alpha})\right) f(\tilde{\alpha}) \\
&\propto \exp\left(\tilde{\alpha}\left\{\frac{1}{\phi}\sum_{i:z_{ij}=1}\frac{y_i}{\mu_i}b_2(\mu_i) + \frac{\delta}{\phi_\alpha}\right\} - b(\tilde{\alpha})\left\{\frac{1}{\phi}\sum_{i:z_{ij}=1} b_2(\mu_i) + \frac{1}{\phi_\alpha}\right\}\right).
\end{aligned} \tag{14}$$

This has the same form as the density in equation (9) with the new parameters

$$\delta^* = \frac{\phi\delta + \phi_\alpha \sum_{i:z_{ij}=1}(y_i/\mu_i)b_2(\mu_i)}{\phi + \phi_\alpha \sum_{i:z_{ij}=1} b_2(\mu_i)} \tag{15}$$

$$\phi_\alpha^* = \frac{\phi\phi_\alpha}{\phi + \phi_\alpha \sum_{i:z_{ij}=1} b_2(\mu_i)} \tag{16}$$

In particular, define the weight $W_j = \sum_{i:z_{ij}=1} b_2(\mu_i)$ and the credibility factor

$$\zeta_j = \frac{\phi}{\phi + \phi_\alpha W_j}. \tag{17}$$

With this, we have $\phi_\alpha^* = \phi_\alpha\zeta_j$ and

$$\delta^* = \zeta_j\delta + (1-\zeta_j)\overline{(y/\mu)}_{Wj}, \tag{18}$$

where $\overline{(y/\mu)}_{Wj} = \sum_{i:z_{ij}=1}(y_i/\mu_i)b_2(\mu_i)/W_j$, a weighted average.