MINIMUM BIAS WITH GENERALIZED LINEAR MODELS

ROBERT L. BROWN

Abstract

The paper "Insurance Rates with Minimum Bias" by Robert A. Bailey [3] presents a methodology which is used by a large number of Canadian casualty actuaries to determine class and driving record differentials. In his paper, Bailey outlines four methods (two directly and two by reference to a previous paper by Bailey and Simon). No presentation has ever been made of an analysis of the applicability of these methods on Canadian data. Also, no attempt has been made within the Casualty Actuarial Society literature to augment Bailey's discussion using other statistical approaches now familiar to members of the Society.

This paper analyzes the four Bailey methodologies using Canadian data and then introduces five models using a modern statistical approach. (It should be noted that one of these statistical models turns out to be a reproduction of one of Bailey's models.)

The paper then gives a brief study of generalized linear models followed by an explanation of one possible way of mathematically modeling the specified Canadian data to the given models on the computer using a statistical software package called GLIM (Generalized Linear Interactive Modeling).

1. BACKGROUND

The concept of minimum bias was first introduced to insurance as a means of setting fair rates for groups of exposure units that could be classed in several different ways.

In his paper, "Insurance Rates with Minimum Bias" [3], Robert Bailey expresses the problem most eloquently:

"Although we may get a more reliable indicated adjustment for brick dwellings by combining all brick classes, and a more reliable indicated adjustment for

> small dwellings by combining all small dwelling classes, we cannot be so confident that the adjustment for brick dwellings and the adjustment for small dwellings will combine to produce the proper net adjustment for small brick dwellings. The data for small brick dwellings may be insufficient to be fully reliable but it will always provide some information. So we should look at it and take it into consideration. We should try to use a ratemaking system which, instead of producing each set of adjustments successively one after another, produces all sets of adjustments simultaneously. In this way the adjustments for brick dwellings and for small dwellings will both reflect the indication of small brick dwellings as well as the total for brick dwellings and the total for small dwellings. Such a system will produce a better result than a system which ignores the data in each subdivision."

In this 1963 paper, Bailey was actually expanding on work first presented in his 1960 paper, "Two Studies in Automobile Insurance Ratemaking," coauthored with LeRoy J. Simon [4].

In their 1960 paper, Bailey and Simon laid out four criteria for an acceptable set of relativities:

1. It should reproduce the experience for each class and merit rating (*driving record*) class and also the overall experience, i.e., be *balanced* for each class and in total.
2. It should reflect the relative *credibility* of the various groups involved.
3. It should provide a minimal amount of *departure* from the raw data for the maximum number of people.
4. It should produce a rate for each subgroup of risks which is close enough to the experience so that the differences could reasonably be caused by *chance*.

Using these criteria, the authors introduced four models (two multiplicative and two additive) that have proven very popular with actuaries.

Since the 1960 paper dealt with the same two variables for auto ratemaking (class and driving record) as we analyze in this paper, we will present Bailey's historical formulae as they would exist for two parameters.

Let r_{ij} be the factor from the actual experience that indicates losses for the n_{ij} risks that can be characterized by parameters x_i and y_j. Thus, for example, r_{ij} could be the average loss cost for cell (i,j) corresponding to class i and driving record j.

In the 1963 paper, Bailey introduces the multiplicative model whereby:

$$x_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} y_j} \tag{1.1}$$

and similarly for y_j $\left(\text{i.e., } y_j = \dfrac{\sum_i n_{ij} r_{ij}}{\sum_i n_{ij} x_i}\right)$.

We will refer to this later as Model 1.

He also introduces an additive model whereby:

$$x_i = \sum_j \frac{n_{ij}(r_{ij} - y_j)}{\sum_j n_{ij}} \tag{1.2}$$

and similarly for y_j. We will refer to this as Model 2.

From the 1960 paper, we also have two models, one multiplicative and one additive. For the multiplicative model, which we will refer to as Model 3, we have:

$$x_i = \left[\frac{\sum_j \dfrac{n_{ij} r_{ij}^2}{y_j}}{\sum_j n_{ij} y_j}\right]^{1/2} \tag{1.3}$$

and similarly for y_j.

For the additive model, which we will refer to as Model 4, we have:

$$\Delta x_i = \frac{\sum_j n_{ij}\left(\dfrac{r_{ij}}{x_i + y_j}\right)^2 - \sum_j n_{ij}}{2\sum_j n_{ij}\left(\dfrac{r_{ij}}{x_i + y_j}\right)^2\left(\dfrac{1}{x_i + y_j}\right)}\,. \tag{1.4}$$

Finally, in the 1963 paper, Bailey introduces two tests that can be used to evaluate the appropriateness of the models. They are the chi-squared statistic and the absolute value statistic. For a multiplicative model, the respective formulae are:

$$\text{Absolute Value} = \frac{\sum_{ij} n_{ij} \left| r_{ij} - x_i y_j \right|}{\sum_{ij} n_{ij} r_{ij}} \tag{1.5}$$

$$\text{Chi-Squared} = \sum_{ij} \frac{n_{ij}(r_{ij} - x_i y_j)^2}{x_i y_j} \tag{1.6}$$

For an additive model, the respective formulae are:

$$\text{Absolute Value} = \frac{\sum_{ij} n_{ij} \left| r_{ij} - x_i - y_j \right|}{\sum_{ij} n_{ij} r_{ij}} \tag{1.7}$$

$$\text{Chi-Squared} = \sum_{ij} \frac{n_{ij}(r_{ij} - x_i - y_j)^2}{x_i + y_j} \tag{1.8}$$

2. INTRODUCTION

The data used in this paper were collected by the Insurers Advisory Organization (IAO) from third party auto liability totals for Canada for the years 1981, 1982, and 1983. The data have been grouped by class and driving record and their differentials have been determined according to class ($x_1, x_2, \ldots, x_{13}$) and driving record ($y_1, y_2, \ldots, y_5$). The differentials satisfy the objective of minimizing the bias in the rates.

Two main types of rate models are examined in the paper:

1) The multiplicative model; and
2) The additive model.

Under the multiplicative model, a driver in class i with driving record j will pay the rate

$$(BR_m) \times x_i y_j.$$

Under the additive model the same driver would pay the rate

$$(BR_a) + x_i + y_j,$$

where BR is a base rate and BR, x_i, and y_j vary by the model applied. Thus it can be seen that an entire rate manual can be constructed from 13+5 numbers. The only constraints placed on these 18 numbers are:

1) $\sum_{j=1}^{5} \sum_{i=1}^{13} n_{ij} f(x_i, y_j) =$ total loss dollars,

   where $f(x_i, y_j)$ is the premium that a class i driver with driving record j would pay; and
2) Each of the 65 premiums must be as "fair" as possible.

It is this second constraint that leads to the idea of minimum bias.

Robert Bailey introduced two different bias functions in his paper, "Insurance Rates with Minimum Bias." Each is a function of the new premium $f(x_i, y_j)$ and the expected unit loss costs which were written as r_{ij}. The two functions he introduced corresponded to the two different ratemaking models in use, the multiplicative and additive models. In each case the differential x_i or y_j was the one which minimized the total average difference in each class and driving record.

The average difference for class i for the multiplicative model is:

$$\frac{\sum_{j=1}^{5} n_{ij}(r_{ij} - f(x_i, y_j))}{\sum_{j=1}^{5} n_{ij} r_{ij}}. \tag{2.1}$$

Setting this equal to zero gives

$$\sum_{j=1}^{5} n_{ij}(r_{ij} - f(x_i, y_j)) = 0, \tag{2.2}$$

which implies

$$\sum_{j=1}^{5} n_{ij}r_{ij} = x_i \sum_{j=1}^{5} n_{ij}y_j. \tag{2.3}$$

This, in turn, gives the recursive method for calculating the x_i's and y_j's that are described in the two aforementioned papers.

Similarly, under the additive model, the average difference for class i is:

$$\frac{\sum_{j=1}^{5} n_{ij}(r_{ij} - (x_i + y_j))}{\sum_{j=1}^{5} n_{ij}r_{ij}}. \tag{2.4}$$

Setting this equal to zero gives

$$\sum_{j=1}^{5} n_{ij}(r_{ij} - y_j) = x_i \sum_{j=1}^{5} n_{ij}. \tag{2.5}$$

From this comes the second of Bailey's recursions.

More generally, one needs only to define a bias function,

$f(r_{ij}, n_{ij}, x_i, y_j)$,

and then minimize

$$\sum_{i=1}^{13} \sum_{j=1}^{5} f(r_{ij}, n_{ij}, x_i, y_j) \tag{2.6}$$

with respect to $(x_1, \ldots, x_{13})$ and $(y_1, \ldots, y_5)$.

Thus it can be seen that Bailey's concept is simply an exercise in statistical modeling.

The second method of ratemaking is actually an exercise in statistically modeling the expected values (the r_{ij}'s) and then solving for the x_i's and y_j's so as to maximize the likelihood of the r_{ij}'s being generated by the model.

Thus if one assumes that the r_{ij}'s are independent observations from a random variable with distribution function

$$f^*(z, f(x_i, y_j)),$$

then the likelihood function becomes

$$\prod_{j=1}^{5} \prod_{i=1}^{13} \{f^*(r_{ij}, f(x_i, y_j))\}^{n_{ij}} = L. \tag{2.7}$$

Thus the model chosen is the one which maximizes L with respect to the x_i's and y_j's with the distribution f^* and the rate form $f(x_i,y_j)$ having already been chosen. More conveniently, one might choose to maximize the log likelihood function

$$\ln \{L\} = \sum_{j=1}^{5} \sum_{i=1}^{13} n_{ij} \ln \{f^*(r_{ij}, f(x_i,y_j))\} \tag{2.8}$$

so that

$$n_{ij} \ln \{f^*(r_{ij}, f(x_i, y_i))\} \tag{2.9}$$

takes the place of

$$f(r_{ij}, n_{ij}, x_i, y_j) \tag{2.10}$$

as the bias function, and we find the maximum value instead of the minimum.

From this analogy a system of 18 equations can be derived, all of which must be satisfied by the x's and y's.

$$\frac{\partial}{\partial x_i} \ln \{L\} = \sum_{j=1}^{5} \sum_{i=1}^{13} \frac{n_{ij} f_1(x_i,y_i)\, f_2^* \,(r_{ij}, f(x_i,y_i))}{f^*(r_{ij}, f(x_i,y_i))}, \tag{2.11}$$

and similarly for y_i.

Thus the maximum of $\ln \{L\}$ gives

$$\frac{\partial}{\partial x_i} \ln \{L\} = 0 , \quad i=1, 2, \ldots, 13, \text{ and} \tag{2.12}$$

$$\frac{\partial}{\partial y_i} \ln \{L\} = 0 , \quad j=1, 2, \ldots, 5. \tag{2.13}$$

From the nature of the expression of $\frac{\partial}{\partial x_i} \ln\{L\}$, it is easy to deduce that solving the 18 equations analytically is probably impossible and that some iterative method must be employed.

This paper explains how generalized linear models can be used to solve this problem.

3. THE LOSS COST APPROACH

The loss cost is defined as the incurred losses divided by the exposure units. There is a loss cost for each class and driving record combination which produces a matrix of loss costs. One's class is defined according to age, sex, marital status, and use of car (see Appendix A). Driving record is defined as the number of years of claim-free experience. For example, driving record 5 is defined as 5 years of claim-free experience. Statistics are available for driving record 0, 1, 2, 3, and 5.

By law, premium and claim statistics are collected by the Insurance Bureau of Canada in the seven provinces not operating under a government monopoly (the latter three being British Columbia, Saskatchewan, and Manitoba). The Province of Quebec has a government-administered no-fault system for bodily injury liability claims, so that the statistics published for class and driving record are usually analyzed for the six remaining provinces in total and for the last three policy years in total. This allows enough data for credibility. Many Canadian actuaries then derive pricing differentials for the two parameters, class and driving record, using a methodology consistent with that presented in Bailey's paper. The formulae from Bailey's 1963 *PCAS* paper have been reduced to two variables, x and y, representing class and driving record, the two parameters of interest. Note that the order of the variables is irrelevant.

The 13×5 matrix of loss costs becomes the parameter r_{ij} in all the formulae. The variable n_{ij} is the number of cars or exposure units. The computer model is then solved for the x_i and y_j differentials using the matrix of calculated loss costs as the r_{ij} and the exposure units as the n_{ij}.

Class 02 and driving record 3 define the base class and base driving records respectively. For multiplicative models, their differentials are each set to 1. In additive models, the base differentials are each set to

0. All differentials are normalized with the base differential. (See Appendix B for non-normalized and normalized differentials.)

The net premiums are calculated next, such that the total net premium income would equal the total dollars of loss. The premium in cell 023 (Class 02, Driving Record 3) is the base rate (BR). Using the formulae previously derived in the introduction,

$$BR_m \sum_{j=1}^{5} \sum_{i=1}^{13} n_{ij} x_i y_j = \text{incurred loss, for multiplicative model,} \tag{3.1}$$

and

$$\sum_{j=1}^{5} \sum_{i=1}^{13} (BR_a + x_i + y_j) n_{ij} = \text{incurred loss, for additive model,} \tag{3.2}$$

are the respective net premiums.

Bailey's 1963 paper introduces the following two models:

Model 1: Bailey's Minimum Bias Multiplicative Model

$$x_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} y_j} \tag{3.3}$$

and similarly for y_j. (Note that this model can be derived using maximum likelihood estimation for a Poisson distribution within a loglinear model, as shown later in the Statistical Approach section of the paper.)

Model 2: Bailey's Minimum Bias Additive Model

$$x_i = \frac{\sum_j n_{ij}(r_{ij} - y_j)}{\sum_j n_{ij}} \tag{3.4}$$

and similarly for y_j.

The 1960 paper by Bailey and Simon introduced two other methods, namely:

Model 3: Bailey and Simon—Multiplicative

$$x_i = \left[\frac{\sum_j \dfrac{n_{ij} r_{ij}^2}{y_j}}{\sum_j n_{ij} y_j} \right]^{1/2} \tag{3.5}$$

and similarly for y_j.

Model 4: Bailey and Simon—Additive

$$\Delta x_i = \frac{\sum_j n_{ij} \left(\dfrac{r_{ij}}{x_i + y_j} \right)^2 - \sum_j n_{ij}}{2 \sum_j n_{ij} \left(\dfrac{r_{ij}}{x_i + y_j} \right)^2 \left(\dfrac{1}{x_i + y_j} \right)} . \tag{3.6}$$

and similarly for y_j.

For Model 4, different starting values of x converge to different non-normalized class and driving record relativities, but $x_i + y_j$ and the normalized class and driving record differentials are independent of the starting values of x and y.

These four classic methodologies were tested on the Canadian data split by rural and urban territories. As explained earlier, Bailey introduced two tests in his 1963 paper that can be used to evaluate the appropriateness of a model: the chi-squared statistic and the absolute value statistic. The chi-squared statistic and absolute value statistic for the first four models are as follows:

	Model			
	1	2	3	4
Urban Territories:				
Chi-Squared	6,684,350	56,886,610	6,552,692	10,854,933
Absolute Value	.05145	.05773	.05178	.06226
Rural Territories:				
Chi-Squared	7,101,723	115,079,807	6,459,712	8,309,002
Absolute Value	.06621	.07042	.07651	.08372

Note: In Canada, geographic territories are split depending on whether they are predominantly urban or predominantly rural in nature, and separate class and driving record relativities are accordingly derived.

4. THE LOSS RATIO APPROACH

In the loss cost approach described on the previous pages, it is assumed implicitly that the distribution of all other parameters is completely homogeneous across class and driving record. One approach to correct or adjust for heterogeneity in the distribution of any parameters not being directly analyzed would be to use the loss ratio method in defining the r_{ij} matrix for the minimum bias calculation.

The minimum bias analysis can be done using a loss ratio approach as follows. Given the 13 class differentials and the 5 driving record differentials provided by the IAO (see Appendix B), a 13 $\times$ 5 matrix of "existing differentials" is calculated for all 65 cells (note that cell 023 will equal 1.00). The loss ratios for these 65 cells are then calculated (incurred losses/earned premiums). Each of these entries is then divided by the loss ratio calculated for cell 023. This matrix multiplied by the "existing differential" matrix gives the "indicated differential" matrix, which is used as the r_{ij} in the minimum bias calculation. Before using the generalized linear modeling technique, the first iteration must be performed manually to convert the matrix of differentials into a matrix of rates.

This calculation can be done as follows:
Let BR = the base rate, where

$$BR \sum_{ij} n_{ij} r_{ij} = \text{incurred losses.} \tag{4.1}$$

Then,

$$BR \cdot r_{ij} = \text{the new } r_{ij} \text{ matrix.} \tag{4.2}$$

Generalized linear modeling can now be used in exactly the same manner as was used for the loss cost approach.

Using the loss ratio approach to calculate the r_{ij}'s, the results using the criteria outlined in Bailey's paper are as follows:

	Model			
	1	2	3	4
Urban Territories:				
Chi-Squared	6,689,226	108,368,525	6,553,952	10,828,905
Absolute Value	.05029	.05623	.05060	.06188
Rural Territories:				
Chi-Squared	5,741,837	129,531,557	5,224,229	6,754,593
Absolute Value	.06587	.07018	.07600	.08197

Using Bailey's criteria, the loss ratio method appears to provide slightly better results than the loss cost method in general.

5. A STATISTICAL APPROACH

The following is a summary of the results obtained from the minimum bias analysis for several possible models that can be derived using modern statistical formulae.

a) Maximum Likelihood Methods
If the losses for cell (i,j) are modeled by $L_{ij} = n_{ij}r_{ij}$, then $E(L_{ij}) = n_{ij}E(r_{ij})$.

Model 5: Exponential—Multiplicative
$L_{ij} \sim$ exponential and

$$E(L_{ij}) = n_{ij}E(r_{ij}) = n_{ij}x_iy_j \tag{5.1}$$

$$f(L_{ij}) = f(n_{ij}, r_{ij}) = \frac{1}{n_{ij}x_iy_j} \exp\left\{-\left(\frac{r_{ij}n_{ij}}{n_{ij}x_iy_j}\right)\right\}$$

$$= \frac{1}{n_{ij}x_iy_j} \exp\left\{-\left(\frac{r_{ij}}{x_iy_j}\right)\right\}. \tag{5.2}$$

The log likelihood function is

$$l = -\sum_i \sum_j \left(\ln n_{ij} + \ln x_i + \ln y_i + \frac{r_{ij}}{x_i y_j} \right) \tag{5.3}$$

$$\frac{\partial l}{\partial x_k} = 0 \Rightarrow \frac{1}{x_k} \sum_j 1 - \frac{1}{x_k^2} \sum_j \frac{r_{kj}}{y_j} = 0 \tag{5.4}$$

$$\Rightarrow x_i = \frac{\sum_j \frac{r_{ij}}{y_j}}{\sum_j 1} \tag{5.5}$$

and similarly for y_j.

Model 6: Normal—Multiplicative

$$L_{ij} \sim N(\mu_{ij}, \sigma^2) \tag{5.6}$$

$$\mu_{ij} = n_{ij} x_i y_j \tag{5.7}$$

$$f(L_{ij}) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2} (L_{ij} - \mu_{ij})^2 \right\}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2\sigma^2} (r_{ij} n_{ij} - x_i y_j n_{ij})^2 \right\}. \tag{5.8}$$

The log likelihood function is

$$l = \sum_i \sum_j \{ -\ln \sigma (\sqrt{2\pi}) \} - \frac{1}{2\sigma^2} \sum_i \sum_j n_{ij}^2 (r_{ij} - x_i y_j)^2 \tag{5.9}$$

$$\frac{\partial l}{\partial x_k} = 0 \Rightarrow \sum_j n_{kj}^2 \, y_j (r_{kj} - x_k y_j) = 0 \tag{5.10}$$

$$\Rightarrow x_i = \frac{\sum_j n_{ij}^2 r_{ij} y_j}{\sum_j n_{ij}^2 y_j^2} \tag{5.11}$$

and similarly for y_j.

Model 7: Normal—Additive

$$L_{ij} \sim N(\mu_{ij}, \sigma^2) \tag{5.12}$$

$$\mu_{ij} = (x_i + y_j) n_{ij} \tag{5.13}$$

The log likelihood function is

$$l = \sum_i \sum_j \{- \ln (\sigma\sqrt{2\pi})\} - \frac{1}{2\sigma^2} \sum_i \sum_j n_{ij}^2 (r_{ij} - x_i - y_j)^2 \tag{5.14}$$

$$\frac{\partial\, l}{\partial\, x_k} = 0 \Rightarrow \sum_j n_{kj}^2 \,(r_{kj} - x_k - y_j) = 0 \tag{5.15}$$

$$\Rightarrow x_i = \frac{\sum_j n_{ij}^2 (r_{ij} - y_j)}{\sum_j n_{ij}^2} \tag{5.16}$$

and similarly for y_j.

Model 8: Poisson—Multiplicative (Bailey's Model 1)

We will now show that by using maximum likelihood estimation, a Poisson—Multiplicative model will reproduce Bailey's Model 1.

$$f(L_{ij}) = \frac{e^{-x_i y_j n_{ij}} (x_i y_j n_{ij})^{n_{ij} r_{ij}}}{(n_{ij} r_{ij})!} \, . \tag{5.17}$$

The log likelihood function is

$$l = \sum_i \sum_j \{(n_{ij} r_{ij} [\ln x_i + \ln y_j + \ln n_{ij}] - x_i y_j n_{ij} - \ln(n_{ij} r_{ij})!\} \tag{5.18}$$

$$\frac{\partial l}{\partial x_k} = 0 \Rightarrow \sum_j \left\{ \frac{n_{kj} r_{kj}}{x_k} - n_{kj} y_j \right\} = 0 \tag{5.19}$$

$$\Rightarrow x_i = \frac{\sum_j n_{ij} r_{ij}}{\sum_j n_{ij} y_j} \tag{5.20}$$

which is Bailey's Model 1.

b) Least Squares Estimate Methods
Model 9 LSE—Multiplicative

$$SS = \sum_i \sum_j n_{ij}(r_{ij} - x_i y_j)^2 \tag{5.21}$$

$$\frac{\partial SS}{\partial x_k} = 0 \Rightarrow \sum_j n_{kj} y_j (r_{kj} - x_k y_j) = 0 \tag{5.22}$$

$$\Rightarrow x_i = \frac{\sum_j n_{ij} r_{ij} y_j}{\sum_j n_{ij} y_j^2} \tag{5.23}$$

and similarly for y_j.

These models were applied to the Canadian data and the resulting chi-squared statistic and absolute value statistic are as follows:

	Model			
	5	6	7	9
Urban Territories:				
Chi-Squared	13,059,115	7,023,572	14,040,792	7,009,249
Absolute Value	.12810	.04175	.07145	.05621
Rural Territories:				
Chi-Squared	11,877,604	9,210,338	9,776,232	7,623,831
Absolute Value	.18830	.05155	.07633	.07757

6. GENERALIZED LINEAR MODELS

A generalized linear model is a probability distribution for an observed random variable vector Y given a set of explanatory vectors x_1, $x_2, \ldots, x_p$ which satisfy the following three conditions:

1) The distribution of each y_i of the vector Y ($i = 1, 2, \ldots, n$), given $x_{i1}, \ldots, x_{ip}$, belongs to an exponential family. The probability density function (pdf) for each y is of the form

$$\exp\left[\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi)\right] \tag{6.1}$$

where θ, known as the *canonical* parameter, is a function of $x_{i1}, \ldots, x_{ip}$ that involves known parameters; and ϕ, known as the *dispersion* parameter, is constant for all i.

It can be shown that

$$\mu_i = E(y_i) = b'(\theta_i) \tag{6.2}$$

and

$$\mathrm{var}(Y_i) = b''(\theta_i)a(\phi). \tag{6.3}$$

2) The explanatory variables enter only as a linear sum of their effects, the linear predictor, η; hence, for each i,

$$\eta_i = \sum_{j=1}^{p} x_{ij}\beta_j = X\beta$$

where the β_j effects are the linear parameters to be estimated.

3) The expected value of each observation can be expressed as some function of its linear predictor, $\eta_i = g(\mu_i)$, where $g(.)$ is a monotonic and differentiable function known as the link function.

The link function is a transformation between the linear function and the mean. Those readers not familiar with generalized linear models are encouraged to read references [8] and [13].

7. GLIM

With the previous references, it would be possible for an actuary to program the technique known as generalized linear models with the information already provided. However, as one might expect, packages do exist to do this type of analysis. One of the best known and most complete and flexible is called the Generalized Linear Interactive Modeling (GLIM) system. It was designed by the Royal Statistical Society and is available at the address given at the end of the Bibliography. From this point on, GLIM will refer to the Royal Statistical Society program.

GLIM is a computerized statistical package which mathematically models a random quantity (dependent or response variables) and takes into account any related or covariate information (independent or explanatory variables). The model that is produced is the one that maximizes the log likelihood function over the given data set. This paper treats the pure premiums as the dependent variables and attempts to relate the corresponding class and driving differentials to these pure premiums.

8. BASIC COMPONENTS OF A GLIM PROGRAM

The general format of any GLIM program to be run in a batch environment using GLIM consists of a data definition section at the beginning, followed by the actual process or the body of the program, which performs the model fitting. The data definition directives describe the structure of the initial input data matrix, establish the variable labels for the input data, and read the data into the program work area. The GLIM commands which make up the body of the program provide the instructions for GLIM to do the statistical modeling and analysis. (Further details are available from the author.)

9. SUMMARY OF RESULTS

For this paper, a total of 12 models were fit using GLIM with a loss cost approach. Of these 12 models, 6 were run using urban data and 6 with rural data. Within each of the 2 sets of data, the output was further divided to include 3 probability distributions: Poisson, Gamma, and Normal. Finally, each distribution was run using a multiplicative approach and an additive approach.

The interesting statistics from the output are:

1) Deviance: This is the residual variance, or the variability not explained by the model. In particular it is twice the drop in log likelihood

between a model which fits the data perfectly (one parameter per observation) and the model actually fit. (This is similar to Bailey's chi-squared statistic.)
2) Degrees of Freedom
3) Fitted Values: This column gives the new premium rates as fitted by the model.
4) Estimate and Standard Error: These can be used to generate the differentials.

The deviances produced by GLIM for the 12 models are summarized below. Note that for two of the models (Poisson and Exponential—Additive using rural data) the deviances have been entered as "***" indicating that the fitted mean is out of range for the error distribution (i.e., an inappropriate model).

Deviance Under Loss Cost Approach

Distribution	Link	Model	Urban	Rural
Poisson	log	Multiplicative	6,596,200	5,295,126
Gamma*	log	Multiplicative	18,373	32,614
Normal	log	Multiplicative	3,413,386,183	1,518,522,878
Poisson	identity	Additive	10,422,477	***
Gamma*	identity	Additive	37,425	***
Normal	identity	Additive	4,084,117,310	1,902,075,827

*It should be noted that a Gamma distribution is of the form:

$$f(x;\alpha,\lambda) = \frac{\lambda^{\alpha} x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

and an Exponential distribution is of the form:

$$f(x;\lambda) = \lambda e^{-\lambda x}.$$

Thus an Exponential distribution is a Gamma distribution with the parameter $\alpha = 1$.

10. ADVANTAGES/DISADVANTAGES OF GLIM

Advantages

Once the user is familiar with the GLIM package, only a minimal amount of computer knowledge is necessary to utilize GLIM as an automobile ratemaking tool.

Once the GLIM system is in working order, it is easy to test different statistical models, as only minor changes must be made to the GLIM command file.

GLIM can reproduce results obtained by Bailey's multiplicative model because a GLIM Poisson log-linear model (under maximum likelihood estimation) is equivalent to Bailey's multiplicative model.

Disadvantages

To this point we have not been able to use GLIM to reproduce results obtained by Bailey's additive model, the Bailey and Simon additive model, or the Bailey and Simon multiplicative model. However, there should not be any need to produce premium rates using these models, since Bailey's multiplicative model (GLIM Poisson log-linear model) produces premium rates which provide a better fit to the data.

GLIM is a difficult package with which to become familiar. However, becoming familiar with GLIM should not be a problem, as this report answers most of the questions new GLIM users might ask.

Conclusion

In conclusion, it is safe to say that the advantages of GLIM outweigh the disadvantages, and, therefore, the GLIM statistical package could be used to determine net premium rates for automobile insurance. Further, once one is familiar with GLIM, many other property-casualty applications become apparent (e.g., loss reserving models).

REFERENCES

[1] Ajne, "A Note on the Multiplicative Rate-Making Model," *ASTIN Bulletin*, 8, 1974, p. 144.

[2] Andersen, "Multiplicative Poisson Models with Unequal Cell Rates," *Scandanavian Journal of Statistics*, 4, 1977, p. 153.

[3] R. A. Bailey, "Insurance Rates with Minimum Bias," *PCAS* L, 1963, p. 4.

[4] R. A. Bailey and L. Simon, "Two Studies in Automobile Insurance Ratemaking," *ASTIN Bulletin*, 1, 1960, p. 192.

[5] L. A. Baxter, S. M. Coutts, and G. A. F. Ross, "Applications of Linear Models in Motor Insurance," *Transactions of the 21st International Congress of Actuaries*, 2, 1979, p. 11.

[6] Peter Berg, "On the Loglinear Poisson and Gamma Model," *ASTIN Bulletin*, 11, 1980, p. 35.

[7] S. M. Coutts, "Motor Insurance Rating, An Actuarial Approach," *Journal of the Institute of Actuaries*, 111, 1984, p. 87.

[8] A. Dobson, *An Introduction to Statistical Modeling*, Chapman and Hall, 1983.

[9] J. van Eeghen, E. K. Greup, and J. A. Nijssen, "Rate Making," *Surveys of Actuarial Studies*, 2, Nationale-Nederlanden N.V., 1983.

[10] T. Grimes, "Claim Frequency Analysis in Motor Insurance," *Journal of the Institute of Actuaries' Students' Society*, 19, 1971, p. 147.

[11] P. D. Johnson and G. B. Hey, "Statistical Studies in Motor Insurance," *Journal of the Institute of Actuaries*, 97, 1971, p. 199.

[12] J. Jung, "On Automobile Insurance Ratemaking," *ASTIN Bulletin*, 5, 1968, p. 41.

[13] P. McCullagh and J. Nelder, *Generalized Linear Models*, Chapman and Hall, 1983.

[14] D. H. Reid, "The Limits of Resolution of a Statistical Rating System," *Transactions of the 22nd International Congress of Actuaries*, 4, 1984, p. 281.

Information regarding GLIM can be obtained from:

The Numerical Algorithms Group Ltd. (NAG)
Mayfield House
256 Bonbury Road
Oxford OX2 7DE
United Kingdom

APPENDIX A

PRIVATE PASSENGER AUTOMOBILE CLASSIFICATION

PLEASURE—NO MALES UNDER 25, NO FEMALE PRINCIPAL OPERATORS UNDER 25:

Class 01: No driving to work; annual mileage of 10,000 or less; 2 or fewer operators per automobile who have held valid operators' licenses for at least the past 3 years.

Class 02: Driving to work 10 miles or less one way permitted; unlimited annual mileage; 2 or fewer operators per automobile.

Class 03: Driving to work over 10 miles permitted; unmarried female occasional drivers under 25 may drive.

PLEASURE OR BUSINESS:

Class 06: Occasional male driver use—male under 25. (Note—the principal operator insures the automobile for use by all other drivers under Classes 01, 02, 03, or 07.)

Class 07: Business primarily; no male drivers under age 25.

PRINCIPAL OPERATORS UNDER 25 YEARS OF AGE:

MARRIED MALE:

Class 08: Ages 20 and under.

Class 09: Ages 21, 22, 23, or 24.

UNMARRIED MALE:

Class 10: Ages 18 and under.

Class 11: Ages 19 and 20.

Class 12: Ages 21 and 22.

Class 13: Ages 23 and 24.

FEMALES—MARRIED OR UNMARRIED:

Class 18: Ages 20 and under.

Class 19: Ages 21, 22, 23, or 24.

APPENDIX B
PAGE 1

EXPOSURES
URBAN

Class	Driving Record 5	3	2	1	0
1	1,032,596	69,952	7,176	6,531	7,531
2	908,551	92,324	12,630	11,138	8,376
3	171,145	22,770	2,333	2,275	2,115
6	22,509	67,929	7,527	8,865	4,315
7	101,962	13,586	1,177	1,214	3,025
8	238	1,471	118	119	57
9	22,395	7,768	890	682	397
10	439	6,876	1,448	1,096	516
11	2,406	17,515	1,421	1,112	874
12	25,362	16,827	1,756	1,420	950
13	37,145	11,345	1,201	981	648
18	2,374	17,957	2,447	1,738	900
19	50,032	18,679	2,212	1,669	905

Rural

Class	Driving Record 5	3	2	1	0
1	588,554	34,156	3,137	2,674	2,853
2	390,669	32,182	4,398	3,768	2,520
3	72,173	9,898	764	732	651
6	6,489	31,307	5,587	6,441	1,902
7	30,164	3,073	231	220	434
8	125	1,239	133	95	45
9	15,172	5,554	578	412	290
10	104	3,473	1,028	700	240
11	552	9,296	853	647	428
12	10,957	6,982	771	589	380
13	14,504	3,922	482	370	233
18	722	9,028	1,447	1,077	400
19	20,085	7,739	979	753	355

APPENDIX B
PAGE 2

INCURRED LOSSES
URBAN

	Driving Record				
Class	5	3	2	1	0
1	$160,542,268	$18,776,760	$1,631,254	$1,497,881	$4,006,765
2	155,275,831	28,026,927	4,960,356	5,161,099	5,966,125
3	31,800,758	6,380,889	746,837	753,405	1,687,178
6	2,265,541	6,965,839	1,106,344	1,747,742	722,654
7	20,952,908	6,983,349	523,905	363,916	1,029,192
8	27,306	1,198,937	52,608	192,758	24,939
9	5,457,632	2,521,186	487,892	171,022	599,265
10	177,911	5,743,938	1,123,514	830,349	361,252
11	650,084	9,317,804	748,966	1,376,913	1,236,611
12	8,128,211	8,259,076	1,948,990	562,474	732,563
13	8,691,528	3,987,223	456,099	602,538	441,043
18	613,659	5,158,436	771,092	649,862	476,125
19	9,869,507	4,620,205	1,104,166	494,677	322,707

RURAL

Class	5	3	2	1	0
1	$72,818,071	$6,569,811	$580,796	$671,249	$693,731
2	46,799,691	8,000,268	692,669	1,444,989	725,973
3	10,979,651	2,609,918	184,710	127,998	121,386
6	1,439,679	2,052,548	392,274	822,024	83,894
7	3,944,980	847,588	205,711	15,873	348,641
8	65,162	331,414	14,595	77,585	43,629
9	2,313,951	1,407,547	151,523	459,051	38,430
10	23,868	1,886,902	1,005,682	895,735	280,150
11	119,059	3,733,793	280,868	263,630	168,825
12	3,084,451	2,820,342	331,082	174,409	514,735
13	3,232,096	1,191,885	196,718	146,137	48,247
18	87,123	2,876,800	312,692	201,297	264,561
19	2,588,858	1,490,676	214,363	248,212	59,574

APPENDIX B
PAGE 3

INDICATED LOSS COSTS (r_{ij} FOR LOSS COST METHODS)*
URBAN

Class	Driving Record 5	3	2	1	0
1	$155.474	$268.423	$227.321	$229.349	$532.036
2	170.905	303.571	392.744	463.378	712.288
3	185.812	280.232	320.119	331.167	797.720
6	100.650	102.546	146.983	197.151	167.475
7	205.497	514.011	445.119	299.766	340.229
8	114.731	815.049	445.831	1,619.815	437.526
9	243.699	324.561	548.193	250.765	1,509.484
10	405.264	835.360	775.907	757.618	700.101
11	270.193	531.990	527.070	1,238.231	1,414.887
12	320.488	490.823	1,109.903	396.108	771.119
13	233.989	351.452	379.755	614.208	680.622
18	258.492	287.266	315.117	373.914	529.028
19	197.264	247.348	499.171	296.391	356.582

Rural

Class	Driving Record 5	3	2	1	0
1	$123.724	$192.347	$185.144	$251.028	$243.158
2	119.794	248.594	157.496	383.490	288.085
3	152.130	263.681	241.767	174.861	186.461
6	221.865	65.562	70.212	127.624	44.108
7	130.784	275.818	890.524	72.150	803.320
8	521.296	267.485	109.737	816.684	969.533
9	152.515	253.429	262.151	1,114.201	132.517
10	229.500	543.306	978.290	1,279.621	1,167.292
11	215.687	401.656	329.271	407.465	394.451
12	281.505	403.945	429.419	296.110	1,354.566
13	222.842	303.897	408.129	394.965	207.069
18	120.669	318.653	216.097	186.905	661.402
19	128.895	192.619	218.961	329.631	167.814

*(Indicated Loss Costs)$_{ij}$ = (Incurred Losses)$_{ij}$ / (Exposures)$_{ij}$

APPENDIX B
PAGE 4

EARNED PREMIUMS
URBAN

	Driving Record				
Class	5	3	2	1	0
1	$235,547,350	$27,864,747	$3,400,585	$3,481,338	$5,389,670
2	243,651,153	43,214,401	7,044,434	7,003,389	7,036,261
3	46,094,341	10,671,357	1,309,985	1,435,101	1,780,991
6	3,067,070	15,693,178	2,082,685	2,752,201	1,780,886
7	31,782,258	7,443,208	765,162	882,174	3,025,876
8	87,819	905,370	87,173	95,278	60,002
9	7,085,253	4,253,265	579,097	501,015	391,990
10	275,491	7,173,186	1,803,896	1,523,661	971,610
11	1,313,681	16,004,319	1,541,004	1,354,574	1,417,867
12	9,971,261	11,404,388	1,403,404	1,279,771	1,153,945
13	14,047,048	7,442,591	929,577	853,985	758,385
18	647,226	8,212,282	1,337,674	1,070,350	740,322
19	13,409,777	8,580,997	1,215,449	1,027,302	753,915

RURAL

Class	Driving Record 5	3	2	1	0
1	\$101,062,451	\$9,321,598	\$1,043,233	\$1,033,029	\$1,279,751
2	72,046,127	9,420,894	1,566,185	1,560,793	1,211,134
3	13,679,991	2,995,335	279,122	316,566	324,724
6	658,589	5,064,497	1,085,059	1,460,394	500,790
7	6,972,259	1,146,006	104,627	115,386	265,936
8	33,728	532,462	68,025	56,717	31,244
9	3,519,880	2,032,396	255,435	214,240	172,875
10	59,135	3,111,955	1,111,589	873,427	344,456
11	261,452	7,042,735	776,994	688,728	526,422
12	4,200,804	4,204,779	560,158	495,889	369,578
13	4,442,756	1,886,737	281,409	249,126	181,789
18	151,636	3,041,185	581,410	500,843	215,521
19	3,836,080	2,333,954	355,640	317,760	173,441

APPENDIX B
PAGE 5

CLASS AND DRIVING RECORD RELATIVITIES
UNDERLYING THE IAO PREMIUMS

Class	Urban	Rural
1	.86	.94
2	1.00	1.00
3	1.00	1.05
6	.50	.55
7	1.12	1.24
8	1.37	1.48
9	1.20	1.26
10	2.31	3.13
11	2.02	2.65
12	1.48	2.09
13	1.42	1.67
18	1.00	1.16
19	1.00	1.04

Driving Record	Urban	Rural
5	.58	.63
3	1.00	1.00
2	1.20	1.22
1	1.35	1.42
0	1.80	1.63