A COMPARISON OF STOCHASTIC CLAIM RESERVING METHODS

by

ERIC M. MANN

B.S., Kansas State University, 2006

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2011

Approved by:

Major Professor
Dr. Haiyan Wang

Abstract

Estimating unpaid liabilities for insurance companies is an extremely important aspect of insurance operations. Consistent underestimation can result in companies requiring more reserves which can lead to lower profits, downgraded credit ratings, and in the worst case scenarios, insurance company insolvency. Consistent overestimation can lead to inefficient capital allocation and a higher overall cost of capital. Due to the importance of these estimates and the variability of these unpaid liabilities, a multitude of methods have been developed to estimate these amounts.

This paper compares several actuarial and statistical methods to determine which are relatively better at producing accurate estimates of unpaid liabilities. To begin, the Chain Ladder Method is introduced for those unfamiliar with it. Then a presentation of several Generalized Linear Model (GLM) methods, various Generalized Additive Model (GAM) methods, the Bornhuetter-Ferguson Method, and a Bayesian method that link the Chain Ladder and Bornhuetter-Ferguson methods together are introduced, with all of these methods being in some way connected to the Chain Ladder Method. Historical data from multiple lines of business compiled by the National Association of Insurance Commissioners is used to compare the methods across different loss functions to gain insight as to which methods produce estimates with the minimum loss and to gain a better understanding of the relative strengths and weaknesses of the methods.

Key terms: Stochastic Claims Reserving, Chain Ladder, Bornhuetter-Ferguson, Generalized Linear Model, Generalized Additive Model, Bayesian, Insurance

Table of Contents

List of Tables

Chapter 1 - Introduction

The business of insurance is unique because unlike many other industries the cost of the product is unknown to the company at the time of sale. This is true for both the losses that occur on policies and the expenses incurred during the process of adjusting and paying for that loss. Due to the importance of these amounts insurance professionals have developed a wide variety of methods to estimate insurance losses. Historically, most of these loss estimation methods consisted of deterministic calculations used to find the point estimate of losses. However, more recent methods have utilized statistical models with a random component meant to provide a reasonable range of loss estimates. This paper will focus on describing and implementing stochastic methods for estimating ultimate losses using real insurance loss data for the purpose of comparing these methods to each other and determining their relative effectiveness.

Estimating ultimate losses is one of the most important aspects of an insurance company. Losses make up the largest portion of the liabilities on the balance sheet, and accurate loss estimation allows investors to fairly determine the value of a company before investing in it. It also provides regulators the information they need to determine whether they should step in to stabilize a company and protect policyholder interests. Additionally, reliable loss estimation enables management to make sound business decisions. It can let them know the value of a book of business that they are considering purchasing, whether a line of business is profitable, what rates are adequate, whether underwriting guidelines are effective, and how to allocate capital in the most strategic and efficient ways (Friedlund 2010). Accurate loss estimation is critical for almost all aspects of an insurance company.

Accurately estimating ultimate losses can also be quite difficult. Property and casualty insurers provide coverage for a variety of risks with distinctive loss patterns. Lines of business like property coverage typically have a lot of claims with small loss amounts that are quickly paid. However, other risks like asbestos liability can go without claims for decades before large, infrequent losses are experienced. These are frequently referred to as short tailed lines and long tailed lines respectively. A variety of other lines exhibit characteristics that are in the middle of these two extremes. Furthermore, many products have unique provisions in deductibles, limits, and exclusions that can also add complexity to the loss estimation process. As a result, accurately estimating insurance losses is not a trivial task. This paper will focus on combinations of longer and shorter tailed lines with varying amounts of data.

Definitions

When an insured has a covered incident they report it to their insurer. The insurer will record details of the claim and classify it as either a reported claim, or also known as an incurred claim. At that time an insurance company representative known as a claims adjuster will, based on their professional judgment, estimate how much that claim will cost the insurance company. This estimate consists of two parts. One part is the case reserve which is the adjuster's estimate of the cost to restore the claimant to their prior event status. The other part that is estimated is the allocated loss adjustment expense reserve (ALAE). This is the estimated expense necessary to investigate, defend, and affect the settlement of a claim. Examples of this expense include the adjuster's time in handling the claim, legal fees to defend the claim, and any other expenses that can be directly linked to the claim (CAS 1988). In most instances, case reserves and ALAE are grouped together and estimated as a whole. However, there are also methods to estimate ALAE individually. Another type of expense is the unallocated loss adjustment expense (ULAE). This includes expenses not directly related to a claim like rent, utilities, salaries of employees not involved with claims (CAS 1988). Estimation of ULAE will not be covered in this paper.

As time progresses and claims evolve case reserves change because a claim is often not paid all at once but over time. Additionally, case reserves may be increased or decreased as more information about the claim becomes available. As a claim is paid the case reserve is converted into another amount called paid loss. Cumulative paid loss is the sum of the paid loss amounts for all of the periods that the claim is open and usually includes ALAE. Another important loss term is incurred loss. Incurred loss is the cumulative paid loss plus the change in the case reserve. This represents the known value of what the claim costs the company at a specific point in time. Paid and incurred losses can also be expressed in incremental amounts that are the difference in the cumulative amounts between two points in time. The process of a claim going from an initial estimate to a final total paid loss is known as loss development.

Table 1.1 provides an example of loss development and displays how case reserves, paid losses, and incurred loss relate to each other.

Table 1.1 An example of the life of a claim payment. Over time the loss transitions from a case reserve to a paid loss.

Time	Case Reserve	Cumulative Paid Loss	Cumulative Incurred Loss	Event
1	100	0	100	A claim is reported.
2	70	30	100	$30 of the claim is paid.
3	150	30	180	Case reserve is increased by $80.
4	110	70	180	$40 of the claim is paid.
5	0	180	180	The last $110 of the claim is paid.

The final loss concept is the total unpaid claim estimate at a point in time. This can be mathematically represented as the difference in the cumulative paid loss at the ultimate payment time and the cumulative incurred loss at a particular point in time. This estimate consists of five parts including outstanding case reserve, a provision for future development on known claims, an estimate for reopened claims, a provision for claims incurred but not reported (IBNR), and a provision for claims in transit which are incurred and reported, but not recorded (Friedlund 2010). The total unpaid claim is the unknown random variable that this paper is attempting to estimate. This contrasts with the cumulative incurred loss which is a fixed and known value. The estimate of the total unpaid claim estimate combined with the cumulative incurred losses equals the ultimate loss that is vital to an insurance company's balance sheet and managerial decisions.

Data Considerations

As can be seen in the prior example the time at which a claim is viewed can affect the amount of the cumulative incurred loss and the estimate of the total unpaid claim. Thus, the time of evaluation will be a crucial component of the analysis. Most data will be evaluated at the end of a period of time like a 3-month quarter or a 12-month year. Also note that for the purposes of this paper it will be assumed that the time an event occurs will also be the same time it is reported and recorded.

Besides the underlying loss process other events like salvage, subrogation, and reinsurance can affect reported cumulative incurred losses and ultimate losses. When a car is destroyed the insurance company takes ownership of the vehicle after paying the claim. It is common for them to sell the scraps of the vehicle as salvage. Subrogation occurs when an insurance company pursues legal restitution against a third party for injuries the third party is liable for against the insured. Finally, reinsurance is insurance for insurance companies. It is common for smaller companies to purchase reinsurance to protect against loss volatility and to meet statutory accounting requirements. These three types of events can reduce the reported cumulative incurred losses so that they are less than what they would be under a theoretical loss generating process. For the purpose of this paper the effects of salvage, subrogation, and reinsurance will not be modeled separately but their effects will be implicitly included in the cumulative incurred losses.

The most common way to show loss data is with a loss triangle. The rows represent the set of losses that occurred during a particular accident year where accident year is defined as the year in which an event occurs and for the purpose of this paper, when the event is reported. The columns represent how much time has passed since the claim occurred. Loss triangles can consist of cumulative or incremental paid or incurred losses, cumulative or incremental paid or incurred claims, or average paid or incurred losses. Tables 1.2 and 1.3 are examples of basic cumulative and incremental loss triangles. Note that in this example as time passes a new diagonal is added to the triangle. Once losses are no longer changing it is assumed that they have arrived at their final ultimate value, and it is the ultimate value that must be estimated for each accident year at different development periods. The difference in the ultimate loss and the recent diagonal of cumulative losses is the unpaid claim amount that this paper is attempting to estimate.

Table 1.2 An example of a cumulative loss triangle. Each cell represents the cumulative losses for a particular accident year after developing for a period of time.

Accident Year	Age of Development in Months					
	12	**24**	**36**	**48**	**60**	**Ultimate**
2003	100	120	140	150	155	155
2004	105	122	150	153	153	153
2005	120	140	135	151	160	160
2006	148	160	177	188	188	
2007	160	186	210	215		
2008	185	199	230			
2009	195	220				
2010	205					

Table 1.3 An example of an incremental loss triangle. Each cell represents the incremental losses for a particular accident year after developing for a period of time.

Accident Year	Age of Development in Months					
	12	**24**	**36**	**48**	**60**	**72**
2003	100	20	20	10	5	0
2004	105	17	28	3	0	0
2005	120	20	15	16	9	0
2006	148	12	17	11	0	
2007	160	26	24	5		
2008	185	14	41			
2009	195	25				
2010	205					

Chapter 2 - Model Review

Chain Ladder Technique

The first model to discuss that is used to estimate losses is known as the Chain Ladder method. Since it is heavily used in the actuarial exams, the Chain Ladder method is one of the most widely known methods in the actuarial community. While the point of this paper is not to focus on this method, this method forms a foundation from which other methods are developed. Thus, a brief review of the Chain Ladder technique is being included so that the reader is familiar with it.

Let C_{ij} be the incremental paid or incurred losses for the ith accident year and jth development period. Then the cumulative losses for the ith accident year and jth development period are defined as:

$$D_{ij} = \sum_{k=1}^{j} C_{ik}$$

Then the Chain Ladder development factor estimate for the jth development period is defined as:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}}$$

where is the total number of development periods. Then to determine the estimate for the cumulative claims for the next period or the next several periods compute:

$$\hat{D}_{i,k} = \hat{D}_{i,k-1}\hat{\lambda}_k \text{ for } k = n-i+3, n-i+4, \cdots n$$

Table 2.1 shows estimated unpaid losses for Accident Year 2009 using the Chain Ladder method where values are rounded to the nearest whole number (England et al. 2002). Bolded cells are estimates that have been calculated out at the bottom of the table to demonstrate how to arrive at them.

Table 2.1 A demonstration of the Chain Ladder technique of loss estimation with bolded cells calculated below.

Accident Year	Age of Development in Months					
	12	**24**	**36**	**48**	**60**	**Ultimate**
2003	100	120	140	150	155	155
2004	105	122	150	153	153	153
2005	120	140	135	151	160	160
2006	148	160	177	188	188	
2007	160	186	210	215		
2008	185	199	230			
2009	195	220	**247**	**261**	**267**	**267**
2010	205					
$247=\frac{(140+150+135+177+210+230)}{(120+122+140+160+186+199)}(220)$				$261=\frac{(150+153+151+188+215)}{(140+150+135+177+210)}(247)$		
$267=\frac{(155+153+160+188)}{(150+153+151+188)}(261)$				$267=\frac{(155+153+160)}{(155+153+160)}(267)$		

Note that there are multiple unwritten variations to the Chain Ladder method. These include taking a straight average of the ratio of the columns for each ith row and using that to calculate the column estimate for the next development period instead of the weighted average as calculated above. Or taking a five year straight average of factors and excluding high and low values. Or taking straight averages and applying some unique weighting scheme to the values based on the practitioner's judgment. Finally, it should also be noted that this method can be heavily influenced by unusually large or small development amounts in the historical period or in the $D_{i,n-i+1}$ period from which projections are being made from (Friedlund 2010).

Generalized Linear Model Methods

The first set of stochastic methods to discuss involve Generalized Linear Models that utilize an over-dispersed Poisson distribution, gamma distribution, over-dispersed negative binomial, normal approximation to the negative binomial, log-normal distribution, and Tweedie distribution. These particular GLM methods generate estimates practically identical to the volume weighted Chain Ladder method introduced in the prior section (England et al. 2002). However, their specification of an underlying probability distribution enables calculation of different standard errors, different confidence intervals, and other diagnostics.

The first method assumes that incremental claim amounts are distributed via an over-dispersed Poisson distribution. The model is parameterized as follows:

$$E\left(C_{ij}\right) = m_{ij}$$

$$Var\left(C_{ij}\right) = \phi m_{ij}$$

$$\log\left(m_{ij}\right) = c + \alpha_i + \beta_j$$

Thus, this model says that the log of incremental losses is a function of an accident year (row) effect and a development period (column) effect. Note that ϕ is an over-dispersion parameter that is estimated from the data.

According to Renshaw & Verrall (1998) it should be noted that this model has constraints. The first constraint is that $\alpha_1 = \beta_1 = 0$ which is a result of the model being over parameterized. The second constraint is that losses must be integers. This is not usually difficult to achieve as losses are usually rounded to the dollar on the financial statements anyway. The third restraint is that each $C_{ij} \geq 0$. The third constraint says that losses in each period should be positive. But due to salvage and subrogation described earlier it is possible for losses in a period to be negative. However, quasi-likelihood maximization can be used to estimate negative losses and losses that are not integers, and thus overcome the second and third constraints. If there are negative incremental losses, quasi-likelihood estimation should be used and the Pearson χ^2 statistic should be used instead of deviances for modeling goodness of fit (Renshaw et al.

1998). Finally, the fourth constraint states that $\sum_i C_{ij} \geq 0$, or sum of the incremental losses in a column cannot be negative. This fourth constraint must not be violated for the model to work.

The over-dispersed Poisson model can also be adjusted and turned into a model that utilizes a gamma distribution (Mack 1991). The linear predictor is the same for this model but has different expectations. This model usually produces estimates similar to the Chain Ladder approach but that is not always true.

$$E\left(C_{ij}\right) = m_{ij}$$

$$Var\left(C_{ij}\right) = \phi m_{ij}^2$$

$$\log\left(m_{ij}\right) = c + \alpha_i + \beta_j$$

The next model is the over-dispersed negative binomial model. This is a recursive model and it parameterizes incremental losses in the following way:

$$E\left(C_{ij}\right) = \left(\lambda_j - 1\right) D_{i,j-1}$$

$$Var\left(C_{ij}\right) = \phi \lambda_j \left(\lambda_j - 1\right) D_{i,j-1}$$

λ_j and $D_{i,j-1}$ are defined in the same way as they were for the Chain Ladder Model and ϕ is an over-dispersion parameter that is again estimated from the data. This model is essentially derived from the Poisson model and has the same expected values and predictive distributions (Verrall et al. 2000).

The main difference between the Poisson model and the negative binomial model are their likelihood functions. The Poisson model utilizes an unconditional likelihood function while the negative binomial utilizes a conditional likelihood function that is conditioned on the latest cumulative claims. The conditional model estimates $n-1$ column parameters while the unconditional model estimates $2n-1$ parameters. The implication under the negative binomial model is that the observed cumulative losses and row totals are fixed and the Poisson model treats them as realized values of random variables where their expected values were estimated based upon the observed values (Verrall et al. 2000). The result of these different assumptions is that the negative binomial has a smaller estimation variance while the Poisson model has a larger estimation variance.

As stated earlier in the over-dispersed Poisson model $\sum_i C_{ij} \geq 0$. If this is violated England & Verrall (2002) suggest a normal approximation to the negative binomial distribution can be used. Since the normal distribution has a support that contains negative values this model has the flexibility to handle the violation of that summation assumption. But be aware that the normal approximation requires estimation of more parameters, and thus is less desirable if the assumption of positive incremental losses is not violated. The authors also indicate that more research should be completed to adjust this symmetric distribution to handle insurance losses, which are typically skewed. The normal approximation of incremental claims model is parameterized in the following way:

$$E\left(C_{ij}\right) = \left(\lambda_j - 1\right) D_{i,j-1}$$

$$Var\left(C_{ij}\right) = \phi_j D_{i,j-1}$$

Cumulative claims can also be modeled and is parameterized with the following:

$$E\left(\frac{D_{ij}}{D_{i,j-1}}\right) = \lambda_j = c + y_{j-1}$$

with $y_1 = 0, j \geq 2$

$$Var\left(\frac{D_{ij}}{D_{i,j-1}}\right) = \frac{\phi_j}{D_{i,j-1}}$$

The next model is a lognormal model and has been widely used over time due to its ease of implementation. This model assumes that:

$$\eta_{ij} = \log\left(C_{ij}\right) \overset{iid}{\sim} N\left(m_{ij}, \sigma^2\right)$$

$$\eta_{ij} = c + \alpha_i + \beta_j$$

This model produces estimates similar to the Chain Ladder method but the estimates are not always close. Take note that since this model is looking at the log of the incremental claims, the incremental claims must also be greater than zero.

The next Generalized Linear Model approach was developed by Wright (1990) and utilizes two distributions to create a model frequently found in insurance pricing. This model assumes that the number of claims $\left(N_{ij}\right)$ is distributed via a Poisson distribution with expectation and variance:

$$E\left(N_{ij}\right) = e_i p_{ij} = e_i a_j k_i j^{A_i} \exp\left(-b_i j\right)$$

$$Var\left(N_{ij}\right) = E\left(N_{ij}\right)$$

where

$$p_{ij} = a_j k_i j^{A_i} \exp\left(-b_i j\right)$$

In this specification e_i is an exposure base. The definition of exposure for the purpose of this model is different than the typical definition found in most insurance papers. In this model exposure is defined as the expected total claim payments for a particular accident year. Also, note that a_j is a known technical adjustment ranging between $1/8$ and 1, and it depends upon whether data is viewed annually, bi-annually, or quarterly. k, b, and A are unknown parameters to estimate, and j is the development period time. p_{ij} takes this form as the delay from accident to payment is likely to have approximately a gamma distribution (Wright 1990). Wright's reasoning behind this assumption is that payment occurs when several successive processes have been completed where each process is likely to have approximately a negative exponential delay (Wright 1990). As a result, the expected number of claims for a particular period is the total number of claim payments for an accident year multiplied by the probability that a payment is made in the jth development period, or

$$E\left(N_{ij}\right) = e_i p_{ij} .$$

The amount of an individual claim $\left(X_{ij}\right)$ is distributed via a gamma distribution with expectation and variance:

$$E\left(X_{ij}\right) = \exp\left(\delta t\right) kj^{\lambda}$$

$$Var\left(X_{ij}\right) = v\left\{E\left(X_{ij}\right)\right\}^2$$

In this parameterization k, and λ are unknown parameters to estimate, v is a proportionality constant relating the mean and variance of the gamma distribution severity amounts. δt is an optional term representing claims inflation with $t = i + j$, that is, t represents the evaluation date of the data. δ is a constant force of claims inflation.

When combined into an aggregate distribution incremental losses have the following expectation and variance:

$$E\left(C_{ij}\right) = m_{ij} = e_i a_i k_i j^{A_i} e^{-b_i j} e^{\delta t} kj^{\lambda}$$

$$Var\left(C_{ij}\right) = \left(1+v\right) kj^{\lambda} e^{\delta t} E\left(C_{ij}\right)$$

Renshaw derived a simple version of this model with mean and variance (England et al. 2002):

$$E\left(C_{ij}\right) = \exp\left(\mu_{ij} + c + \alpha_i + \beta_i \log\left(j\right) + \gamma_i j + \delta t\right) \qquad (1)$$

and

$$Var\left(C_{ij}\right) = \phi_{ij} E\left(C_{ij}\right)$$

with

$$\phi_{ij} = \left(1+v\right) E\left(X_{ij}\right) = \left(1+v\right) \exp\left(\delta t\right) kj^{\lambda}.$$

From this it can be shown that the link and linear predictor are the following:

$$\log\left(E\left(C_{ij}\right)\right) = \eta_{ij} = \mu_{ij} + c + \alpha_i + \beta_i \log\left(j\right) + \gamma_i j + \delta t$$

In this model, and μ_{ij} is a known offset term, typically exposures with $\mu_{ij} = \ln\left(e_i a_j\right)$. Other variables are then transformed with $c = \ln\left(k\right)$, $a_i = \ln\left(k_i\right)$ with $k_1 = 1$, $\beta_i = \lambda + A_i$, and $\gamma_i = -b_i$. ϕ_{ij} is usually held constant for all i and j, and it is estimated by the deviance divided by the degrees of freedom or with joint modeling (England et al. 2001). Since the expectation is so different from the other GLM methods estimates will not likely be equivalent to the Chain Ladder method of estimating incremental losses.

The final type of GLM model is the Hoerl curve. This model utilizes a different linear predictor for either the lognormal, over-dispersed Poisson, or the gamma models previously mentioned. The linear predictor takes the form:

$$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j \quad (j > 0)$$

$$\exp(\eta_{ij}) = A_i j^{\beta_i} e^{\gamma_i j}$$

where

$$A_i = \exp(c + \alpha_i).$$

Note that when including δt and μ_{ij} this model turns into the same linear predictor developed for the Poisson/gamma distribution discussed earlier. And similarly to the Tweedie linear predictor found in equation (1) the accident year effect is considered categorical and the development period effect is considered continuous. The advantage of using a linear predictor of this form is that it corresponds to a shape more similar to loss run-off patterns. Furthermore, it also allows for the modeling of accident year and development year interaction effects. Another advantage of this model is by treating development time period as a continuous covariate future development can be extrapolated beyond what is currently in the data. However, it is also unlikely that this model will fit well over the development time period being modeled (England et al. 2002). A special case of this model is one in which $\beta_i = \beta$ and $\gamma_i = \gamma$. This restriction then implies that losses run off in a manner independent of accident year.

Bayesian Models

The next section discusses the use of Bayesian methods as a means of creating estimates of unpaid losses. These methods can provide the practitioner a means of estimating losses while still incorporating outside information like industry data or expert opinions. The first part discusses the popular Bornhuetter-Ferguson method. Then a more general framework is presented and it is discussed how the Bornhuetter-Ferguson method is simply a special case of the general Bayesian framework.

Bornhuetter-Ferguson

The earliest Bayesian method is the Bornhuetter-Ferguson technique (Bornhuetter et al. 1972). In a typical Chain Ladder method unpaid losses can be expressed as:

$$U_i^{(CL)} - D_{i,n-i+1} = U_i^{(CL)} \frac{1}{\lambda_{n-i+2}\lambda_{n-i+3}\cdots\lambda_n}\left(\lambda_{n-i+2}\lambda_{n-i+3}\cdots\lambda_n - 1\right) = U_i^{(CL)}\left(1 - \frac{1}{\lambda_{n-i+2}\lambda_{n-i+3}\cdots\lambda_n}\right)$$

In this depiction $U_i^{(CL)}$ is the Chain Ladder estimate of the final ultimate losses. And:

$$U_i^{(CL)} = D_{i,n-i+1}\left(\lambda_{n-i+2}\lambda_{n-i+3}\cdots\lambda_n\right)$$

Under the Bornhuetter-Ferguson technique the unpaid losses are expressed as:

$$U_i^{(BF)}\left(1 - \frac{1}{\lambda_{n-i+2}\lambda_{n-i+3}\cdots\lambda_n}\right)$$

In this characterization $U_i^{(BF)}$ is some other external or prior estimate of the expected ultimate losses. Typically this could be other estimates created by the pricing actuaries or management's expectations of what losses will finally become. However, a more general Bayesian method with greater flexibility can be applied as well.

General Bayesian Framework

Going back to the over-dispersed Poisson model described in the GLM section of this paper, no prior beliefs were incorporated when estimating the row effects. Recall the following linear predictor was utilized:

$$\eta_{ij} = c + \alpha_i + \beta_j \qquad (6)$$

However, it can be assumed that the accident years (rows) do have a prior distribution. One of the most reasonable prior distributions to use is the gamma distribution. Since $\exp(c + \alpha_i)$ in the GLM models can be interpreted as the ultimate aggregate loss for the ith accident year and $\exp\left(\beta_j\right)$ can be interpreted as the proportion of those losses that have been realized, then using a gamma distribution as a prior is reasonable model to use since it is also often used for modeling aggregate losses. Thus, it can be reasonably assumed that each accident year (row) effect α_i has the following prior distribution:

$$\alpha_i \mid \gamma_i, \tau_i \overset{ind}{\sim} Gamma\left(\gamma_i, \tau_i\right)$$

where:

$$E\left(\alpha_i\right) = U_i^{(BF)} = \frac{\gamma_i}{\tau_i}.$$

Recall from the discussion of the over-dispersed Poisson and over-dispersed negative binomial GLMs. These two distributional assumptions were a result of assumptions regarding unconditional and conditional likelihood functions. As a result, in the Bayesian context there are two ways to estimate β_j.

The first way involves assuming either a non-informative prior that assigns equal weight to any particular estimate as the distribution of β_j or using plug in estimates of λ_j from the Chain Ladder method (Verrall 2001). Under the non-informative prior, since no weight is being given to any particular estimate, this results in estimates of β_j implied by what is calculated via the Chain Ladder method, and it is similar to the conditional likelihood method described previously. If using the Chain Ladder factors, it is also important to estimate the development year (column) effect first before estimating accident year (row) effects. To estimate ϕ, maximum likelihood estimates of ϕ, similar to what is used for the frequentist GLM, can be used as plug in estimates for ϕ. Or to create a full Bayesian model a prior distribution of ϕ can be specified and integrated it out as well (Verrall 2001). If plug in values are used to estimate ϕ and Chain Ladder factors are used to estimate β_j then the posterior distribution of C_{ij} is an over-dispersed negative binomial distribution (Verrall 2001). It has mean

$$\left(Z_{ij} D_{i,j-1} + \left(1 - Z_{ij}\right)\left(\frac{\gamma_i}{\tau_i}\right)\frac{1}{\lambda_j \lambda_{j+1} \cdots \lambda_n}\right)\left(\lambda_j - 1\right)$$

where:

$$Z_{ij} = \frac{\dfrac{1}{\lambda_j \lambda_{j+1} \cdots \lambda_n}}{\tau_i \phi + \dfrac{1}{\lambda_j \lambda_{j+1} \cdots \lambda_n}}.$$

It is interesting to note that when all weight is put towards the prior, that is the prior has no variance, then the model reduces to the Bornhuetter-Ferguson method. When no weight is given to the prior, or equivalently when the prior is flat and has infinite variance, the model resembles the frequentist Chain Ladder method of estimation. It is also important to realize that the weight given to either estimate is heavily influenced by the prior parameter τ_i that is chosen. Finally, as the model looks at data that is more developed it can be seen from the weighting formula that more weight is given to the Chain Ladder estimate (Verrall 2001).

The second method of estimating β_j is to use some type of improper prior, usually another gamma distribution, to specify β_j and jointly model the accident year (row) and development year (column) effects simultaneously. This method of jointly modeling corresponds to the unconditional likelihood method, and it generates estimates that are different from those obtained when the development years (columns) are first estimated separately (Verrall 2001). This results in an over-dispersed Poisson distribution similar to the frequentist GLM approach. However, this method is more in line with traditional Bayesian ways of estimation since it assumes distributions for all of the parameters of interest.

The use of Bayesian GLMs contains advantages and disadvantages worth discussing. The main advantages of this method is that it allows the user to use outside information, like the prior estimate of ultimate losses, to affect their estimate and compute a weighted value that both reflects the Chain Ladder estimate and the Bornhuetter-Ferguson estimate. The main disadvantage of this model is that it has difficulty dealing with negative incremental losses. The model completely fails when the column totals are negative (Verrall 2001).

Non-Parametric Models

When the practitioner does not feel that a parametric model adequately describes a set of data non-parametric methods can be applied to create estimates of incremental claim payments. This section will focus on Generalized Additive Models as a method of estimating loss development over time.

These non-parametric models can be very closely related to the GLMs formulated in the previous section of this paper. When incremental claims are defined with:

$$E\left(C_{ij}\right) = m_{ij}$$

$$Var\left(C_{ij}\right) = \phi m_{ij}^{\rho}$$

ϕ is a scaling parameter and ρ is an integer that implies what error distribution is specified. For example, $\rho = 0,1,2,$ and 3 give a normal, Poisson, gamma, and inverse Gaussian respectively for the error distribution. A GAM utilizes a linear predictor that is different from a standard GLM. Recall the following GLM predictors:

$$\eta_{ij} = c + \alpha_i + \beta_j$$

or

$$\eta_{ij} = c + \alpha_i + \beta_i \log\left(j\right) + \gamma_i j$$

The GAM linear predictors can be written as

$$\eta_{ij} = c + s_{\theta_i}\left(i\right) + \beta_j$$

or

$$\eta_{ij} = s_{\theta_i}\left(i\right) + s_{\theta_j}\left(\log\left(j\right)\right) + s_{\theta_j}\left(j\right) \qquad (2)$$

where $s_{\theta_i}\left(i\right)$ is a continuous variable on the accident year (row) effect, $s_{\theta_j}\left(j\right)$ is a continuous variable on the development year (column) effect, and θ is a smoothing factor determined by the practitioner.

The GAM linear predictor can be created with locally weighted regression smoothers (loess), cubic smoothing splines, and kernel smoothers. The cubic smoothing spline can be found by minimizing the penalized residual sum of squares

$$\sum_i \left(y_i - s\left(x_i\right)\right)^2 + \theta \int \left(s''\left(t\right)\right)^2 dt .$$

In this model θ acts as a smoothing parameter. Larger values of θ result in smoother models with smaller variance. However, this creates a trade-off as the bias of the model will increase (Verrall 1996). Additionally, when θ is close to zero the function exactly fits each point and becomes more like the Chain Ladder estimate. As θ approaches infinity the function tends to a linear function and becomes more like a GLM (England et al. 2001).

According to Verrall (1996) the loess estimates of S_θ can be found with the following algorithm:

1. Define $N\left(x_0\right)$ to be the set of k nearest neighbors of x_0.
2. Calculate $\Delta\left(x_0\right) = \max_{x_i \in N\left(x_i\right)} \left| x_0 - x_i \right|$.
3. Calculate weights, w_i, for each point in $N\left(x_0\right)$, where

   $$w_i = T\left(\left| x_0 - x_i \right| \big/ \Delta\left(x_0\right) \right),$$

   where T is the tri-cube weight function:

   $$T\left(u\right) = \begin{Bmatrix} \left(1-u^3\right)^3 & \text{for } 0 \le u \le 1 \\ 0 & \text{otherwise} \end{Bmatrix}$$

4. Regress Y on X in $N\left(x_0\right)$ using the weights $\left\{w_1, w_2, \cdots, w_m\right\}$ with the regression being linear, quadratic, or cubic.

The local linear regression has been found to work well enough for the loss development problem.

When errors are believed to be non-normal and from the exponential class a weighted version of the above formula can be applied. Also note that GAMs require the sum of the columns of losses be greater than or equal to zero as is similar to several of the other methods that have been reviewed. If this requirement is violated a constant can be applied to losses to adjust them appropriately (Verrall 1996).

Finally, a combination of parametric and non-parametric functions can be created to form the linear predictor

$$\eta_u = \sum_{v=1}^{p-r} x_{uv}\beta_v + \sum_{v=p-r+1}^{p} s_v\left(x_u\right)$$

where $p =$ number of predictors and $r =$ number of parametric predictors.

The primary advantages of these non-parametric models are that they allow the practitioner the ability to easily and consistently smooth estimates, and using the linear predictor found in equation (2), extrapolate with the same model. It is also robust against small numbers of negative incremental losses but will always produce positive estimates (England et al. 2001). Finally, GAMs can allow for a model that is not over-parameterized as in the case of GLMs (Verrall 1996).

Chapter 3 - Analysis of Models

Research Data

Data used in the comparison of methods consists of publicly available research data compiled and cleaned by the Casualty Actuarial Society (Meyers et al.2011). The CAS received the raw data from Schedule P – Analysis of Losses and Loss Expenses in the National Association of Insurance Commissioners (NAIC) database. Schedule P is part of the statutory accounting statements required for all property and casualty insurance companies in the United States. The data covers years 1988 to 2006 and has a full ten year development of losses for accident years 1988 to 1997. Thus, it consists of a full ten by ten grid of cumulative and incremental losses. This allows researchers to utilize the upper left triangle of this data to build a model to predict the lower right triangle. Data of interest includes cumulative incurred losses and ALAE and cumulative paid losses and ALAE. The data was divided into several types of insurance including personal private passenger auto insurance, commercial private auto insurance, other liability, workers compensation, products liability, and medical malpractice.

There are several advantages and disadvantages of the data set. The primary advantage is that since this is annual statement data it is publicly available, it has been audited by company accountants and actuaries, and it is relatively homogenous within each line of business. However, a disadvantage of the data is that it is net of reinsurance contracts. This means that the historical losses are not what the company actually paid, but what they paid after reinsurance contracts had been implemented. This can create distortions as two companies could have identical losses but different reinsurance contracts and thus, different loss development patterns. It can also create distortions in the loss data if the provisions in the reinsurance contracts changed over the time period being studied. Another disadvantage with the data set is that it combines multiple coverages within a line of business. For example, private passenger auto aggregates losses for coverages with slow development patterns like bodily injury and fast development patterns like collision. This could create distortions if the distribution of coverages offered changes over time. Finally, note that in the data larger insurers that have multiple companies have their multiple companies grouped into one company.

Analysis of Data

An analysis of the Schedule P loss data was conducted to determine the relative effectiveness of the various methods. The objective of the analysis was to compare estimated total unpaid loss estimates with their accompanying actual unpaid loss estimates. Estimates of incremental paid losses for each accident year and development period combination were generated by applying the methods discussed in the prior sections. Then the sums of those estimates were calculated to determine the total unpaid claim estimate as of accident year 1997 for each model and each company. Additionally, actual incremental paid losses were also totaled to provide a base of which to determine the relative effectiveness of each model. Since the data also included the posted reserve for 1997, that is, the amount actually estimated by the companies' actuaries and management, those estimates were also used for comparison purposes.

A variety of models were applied to evaluate the data. Methods used included parametric, non-parametric methods, and deterministic calculations. For the parametric methods Generalized Linear Models with an over-dispersed Poisson distribution, negative binomial distribution, gamma distribution, lognormal distribution, and Tweedie distribution were constructed. Additionally, within this category, the Bayesian model resulting in a negative binomial with a mean equal to a weighting of the Chain Ladder and Bornhuetter-Ferguson methods was utilized. For non-parametric methods, Generalized Additive Models were implemented including Locally Weighted Regression Smoothers and a Cubic Spline. Linear predictors for the models included

$$\eta_{ij} = c + \alpha_i + \beta_j$$

and

$$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$$

for the GLMs and

$$\eta_{ij} = c + s_{\theta_i}(i) + \beta_j$$

and

$$\eta_{ij} = s_{\theta_i}(i) + s_{\theta_j}(\log(j)) + s_{\theta_j}(j)$$

for the GAMs. As before, these linear predictors are for the ith accident year and jth development period. The analysis was performed using the R programming software (R Development Core Team 2010) and several library packages were also included in addition to the base packages. The libraries employed include ‘gam’ (Hastie 2010), ‘tweedie’ (Dunn 2010), ‘MASS’ (Venebles et al 2002) and ‘statmod’ (Smyth et al 2010). The code be requested from the author.

As is often the case when working with real world data, not all of the data exhibited ideal characteristics for a loss development analysis. This is not so say that the data was incorrect, as it has been heavily audited as part of the financial statements of the companies. Problems were related to sparseness of the data and unusual activities of companies. Examples include situations where losses did not go back a full ten years, but only some of the years, or when losses occurred in years in the middle of the triangle, implying that the company may have only written business for a few years and then stopped. Another issue was too many negative incremental paid losses or situations when the entire total unpaid losses ended up actually being negative. As noted in the descriptions of the various models, many methods do not perform optimally when incremental losses are negative and completely fail when total development period incremental losses are negative. As a result, some data sets were adjusted or discarded in an attempt to produce viable estimates and compute relevant statistics to compare the models. Adjustments to data included setting incremental losses less than or equal to zero to values of one. Other adjustments included discarding data sets where actual total unpaid losses equaled a value less than or equal to zero. Finally, a few other company data sets were also removed as a result of several models failing to produce estimates or producing unrealistic estimates. Table 3.1 lists the line of business, the number of companies originally in the data set for that line of business, and the number of companies utilized after poor data sets were culled.

In addition to alterations to the data, some models required alterations or assumptions in order to provide estimates for the data sets. For the Chain Ladder model the estimated development factors, or $\hat{\lambda}_j$, were checked for error statements, negative values, or blanks and were given values of one when an error did occur. For the over-dispersed Poisson model quasi-likelihood maximization was applied and the over-dispersion parameter, ϕ, was estimated from the data. Similarly for the negative binomial model quasi-likelihood maximization was used and the parameter, ϕ, was also estimated from the data. For the Tweedie model it should be noted that two parameters were required to be input for the model to run. One parameter was to specify the link and was set equal to zero indicating a log link. This was selected since in many insurance applications, particularly pricing, a log link is the usual link used when modeling insurance losses. The other parameter that required specification was the power of the variance. A value of 1.667 was selected indicating a compound-Poisson distribution with non-zero mass at value of zero. This was also selected since typical insurance applications contain variance powers greater than 1.5 (Dunn et al. 2007). For the cubic smoothing spline the smoothing parameter, θ, was left as .5 which is the value automatically specified in the GAM R program (Hastie 2010).

Table 3.1 The number of companies in the NAIC data set and the number of companies analyzed using the methods organized by line of business.

Line of Business	Initial Number of Companies	Number of Companies Analyzed
Other Liability	295	234
Commercial Auto Private	202	181
Personal Passenger Auto	187	173
Workers Compensation	166	134
Products Liability	94	66
Medical Malpractice	41	36

The Bayesian negative binomial model required the most number of assumptions since it relies heavily on information outside of the insurance data being analyzed. The prior ultimate loss, γ_i/τ_i, was not estimated from the data since there were two parameters to estimate but there was only one full accident year that could be used to estimate it with, and there were no previous years of data that could be used. Instead the prior ultimate loss was derived by determining the 1988 industry wide loss ratio and multiplying it by the net premium for the ith accident year to create a prior loss estimate specific to that company for each accident year. The industry loss ratio was generated by taking the total cumulative paid losses divided by the total premium, net of reinsurance contracts. Net premiums were chosen since losses were net of reinsurance contracts. This process assumes that each company sets their rates with the goal of targeting a loss ratio similar to the industry average and that these loss ratios do not change with time. Note that this method for determining the prior ultimate loss is consistent with what was done in the original Bornhuetter-Ferguson method (Bornhuetter et al. 1972) and is a generally accepted actuarial practice. When net premiums were absent from the data, the mean formula was automatically adjusted to provide no weight to the prior estimate and all weight to the Chain Ladder estimate. As a result, companies' with premium data issues have Bayesian estimates very similar to Chain Ladder estimates.

Another assumption for the Bayesian negative binomial model involved the credibility. Recall that the credibility formula, z, is a function of the Chain Ladder development factors and $\tau_i\phi$ where ϕ is the over-dispersion parameter found in the over-dispersed Poisson model. Since τ_i was not estimated $\tau_i\phi$ was arbitrarily selected to be set equal to .5. This assumes a constant relationship between τ_i and ϕ that may or may not be valid. However, credibility estimates produced by this assumption appeared to be reasonable and ranged from values of .2 to .8 depending on what development period was being estimated.

Many of the GLM and GAM methods rely on convergence of a maximization algorithm to generate estimates. Even with perfect data, these iterative procedures can fail to converge due to the particular shape of the likelihood function or other various factors. When this occurred 'NA' values were assigned to those total unpaid loss estimates to denote failure of the model.

While many of the assumptions and adjustments made to the data and models may not be ideal, it is important to realize that this comparison process needed to be set up so that each company analysis was automated. This contrasts with a professional analysis that would check more assumptions and determine individual solutions for each company that was being analyzed. However, because that kind of time commitment is unrealistic for this volume of data, these compromises needed to be made. It should also be noted that as a result of these simplifications, some methods may not perform as well in this context as they would had they been given more specialized treatment.

Table 3.2 summarizes the various models, their linear predictors, and parameters of interest.

Table 3.2 A summary of the models used to analyze the data.

Model Type	Model	Linear Predictor 1	Linear Predictor 2	Parameters
GLM	Over-Dispersed Poisson	$\eta_{ij} = c + \alpha_i + \beta_j$	$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$	ϕ estimated from data
GLM	Negative Binomial	$\eta_{ij} = c + \alpha_i + \beta_j$	$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$	ϕ estimated from data
GLM	Gamma	$\eta_{ij} = c + \alpha_i + \beta_j$	$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$	None
GLM	Lognormal	$\eta_{ij} = c + \alpha_i + \beta_j$	$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$	None
GLM	Tweedie	None	$\eta_{ij} = c + \alpha_i + \beta_i \log(j) + \gamma_i j$	Compound Poisson Variance
Bayesian GLM	Bayesian Negative Binomial	$\eta_{ij} = c + \alpha_i + \beta_j$	None	$\gamma_i / \tau_i = (LR)(Net\,EP)$ $\beta_i \phi = .5$
GAM	LOESS	$\eta_{ij} = c + s_{\theta_i}(i) + \beta_j$	$\eta_{ij} = s_{\theta_i}(i) + s_{\theta_j}(\log(j)) + s_{\theta_j}(j)$	$\theta = .5$
GAM	Spline	$\eta_{ij} = c + s_{\theta_i}(i) + \beta_j$	$\eta_{ij} = s_{\theta_i}(i) + s_{\theta_j}(\log(j)) + s_{\theta_j}(j)$	$\theta = .5$
Deterministic	Chain Ladder	$\eta_{ij} = c + \alpha_i + \beta_j$	None	None

Analysis of Methods

Once total unpaid loss estimates were calculated for each method and company by line of business these values were then compared to the actual unpaid losses during that time period to form three statistics with which to evaluate the relative performance of the different methods.

The first statistic utilized was the Structural Similarity Index (SSIM) statistic. This is a statistic often associated with imaging, and it measures the similarities between two images. In this case, the images are the vectors of estimated and observed losses. The statistic is calculated by pairing together the total unpaid losses estimated from a particular method with the actual total unpaid losses. The pairs are then ordered by the estimated values and partitioned into subsets with fifteen observations in each subset. Fifteen observations were used so ensure that each $SSIM_i$ was large enough to estimate well but not so large that there weren't many $SSIM_i$'s to average across. For each ith subset the following value is calculated:

$$SSIM_i = \left(\frac{2\hat{\mu}_X\hat{\mu}_Y + c}{\hat{\mu}_X^2 + \hat{\mu}_Y^2 + c}\right)\left(\frac{2\hat{\sigma}_{XY} + c}{\hat{\sigma}_X^2 + \hat{\sigma}_Y^2 + c}\right)$$

$$= \left(\frac{2\hat{\mu}_X\hat{\mu}_Y + c}{\hat{\mu}_X^2 + \hat{\mu}_Y^2 + c}\right)\left(\frac{2\hat{\sigma}_X\hat{\sigma}_Y}{2\hat{\sigma}_X\hat{\sigma}_Y}\right)\left(\frac{2\hat{\sigma}_{XY} + c}{\hat{\sigma}_X^2 + \hat{\sigma}_Y^2 + c}\right)$$

$$= \left(\frac{2\hat{\mu}_X\hat{\mu}_Y + c}{\hat{\mu}_X^2 + \hat{\mu}_Y^2 + c}\right)\left(\frac{2\hat{\sigma}_X\hat{\sigma}_Y + c}{\hat{\sigma}_X^2 + \hat{\sigma}_Y^2 + c}\right)\left(\frac{\hat{\sigma}_{XY}}{\hat{\sigma}_X\hat{\sigma}_Y}\right) \tag{3}$$

In this calculation X is the vector of estimated total unpaid losses and Y is the vector of actual total unpaid losses, and c = .001 is a stabilization factor included to prevent the denominators from equaling zero. The final SSIM statistic is then just average of all of the SSIM statistics in the data set. It will consist of values between zero and one with values closer to one indicating a stronger trend and values closer to zero indicating a weaker trend.

From formula (3) it can be seen that the SSIM statistic is composed of three parts. The first part compares the means between X and Y. The second part compares the variation between X and Y, while the third part determines the correlation between X and Y. Since these three components are being multiplied together and they are all values between zero and one, a strong SSIM implies that the two models produce similar means, variances, and are positively correlated.

The other two statistics employed to evaluate the methods were the Mean Squared Error (MSE) and Mean Absolute Error (MAE).

$$\sum_{i=1}^{n} \frac{\left(E_i - A_i\right)^2}{n}$$

$$\sum_{i=1}^{n} \frac{\left|E_i - A_i\right|}{n}$$

Where E_i and A_i are the estimated and actual total unpaid losses for the ith company. It is important to note that the MSE penalizes models that have a few larger errors versus models with many smaller errors. It should also be understood that each of these two statistics penalizes models that produce errors when estimating larger companies that have more losses as those companies will have larger values of E_i and A_i than smaller companies. This also has practical meaning as a large insurance company should have a greater volume of data making parameters easier to estimate. Furthermore, miss-estimating reserves for a large company would adversely impact more policyholders and shareholders than a smaller company, and thus the social and economic cost of this miss-estimation would be worse relative to a smaller company.

Results

After the three sets of statistics were calculated for each method and line of business combination the methods were ranked across each statistic. The relative best method was judgmentally selected based on the combined rankings of the models, the magnitude of the fit statistics, and whether a model was able to generate all three fit statistics. Ranks were not the only criteria used as some models may rank well but have very poor statistics for a particular method of judging goodness. Since some models should theoretically produce the same estimates then their fit statistics should also be identical. However, due to the failure of some models to produce estimates for particular companies, these statistics actually end up being different. Thus, models were also implicitly evaluated on whether they were able to consistently generate reliable estimates. Since the SSIM statistic measures the similarity between two images and MSE/MAE measures distance, methods that did well in both categories were viewed favorably. It should also be noted that the posted reserve, that is, the reserve determined by the companies' actuaries and management, was also included in the ranking process.

Spearman correlations were calculated between the SSIM and MSE ranks, SSIM and MAE ranks, and MSE and MAE ranks and were utilized in helping to determine the most effective model. As was expected, the MSE and MAE ranks were highly correlated for each model. However, the level of correlation between the SSIM and MSE, and SSIM and MAE varied much more. The Spearman correlations can be found in Appendix A.

For the Personal Private Auto line of business the Chain Ladder and Poisson model with the second linear predictor appeared to be the best models. The negative binomial with the second predictor also appeared to be a strong model but its result may be distorted due to the absence of the SSIM statistic for that method. The lognormal model and Bayesian negative binomial models performed the worst. SSIM statistics for these methods mostly ranged from values of .10 to .46. Since these values are low it is important to remember that claims regarding which method is good are relative to the other methods and not a statement in an absolute sense. The relevant Spearman correlations were .318 and .188 for the SSIM/MSE and SSIM/MAE respectively. Since these showed some positive correlation, then the goodness of fit statistics were giving similar results. This further supports that the Chain Ladder and Poisson models are relatively superior.

For the Commercial Private Auto line of business the Chain Ladder method appeared to be the overall best method. However, some of the other methods also had good but conflicting results. The gamma for the first linear predictor failed to estimate the SSIM but did well for the MSE and MAE. The lognormal models with both linear predictors performed well with the MSE but not as well under MAE and poorly for SSIM. The LOESS with the second linear predictor and spline with the first linear predictor performed well based on the SSIM statistic but poorly with regards to MSE and MAE. Since SSIM values mainly ranged between .13 and .28

the MSE and MAE criteria may be a better measure to rely upon, and thus, the gamma or lognormal models may be superior for this line of business. This argument is validated since the Spearman correlations were strongly negative. The Poisson models with both linear predictors performed the worst.

For the Other Liability line of business the Tweedie model appeared to be the top overall model. The SSIM values ranged from .07 to .14, indicating a very poor fit. Other models that performed well were the Poisson with the first linear predictor and the negative binomial with the second linear predictor. The spline and gamma models performed the worst. The Spearman correlations where positively correlated for the SSIM and MSE, and SSIM and MAE, indicating that the goodness of fit measures were giving somewhat similar results.

For the Medical Malpractice line of business the LOESS models with both linear predictors were the decisive leaders. The SSIM statistics were also generally higher with these sets of models with values ranging from .10 to .66. This indicates that not only is the LOESS model a good model relative to the other models, it may also be a good model in an absolute way. The Poisson model with the first linear predictor appeared to also be a plausible model. And the lognormal and Bayesian negative binomial models performed the worst. The Spearman correlations were the highest of the business lines. This further confirmed the strength of the LOESS models at making predictions.

For the Product Liability line of business the gamma model with the first linear predictor and the Tweedie model performed the best by a wide margin. SSIM values mostly ranged from .17 to .35. The lognormal with the second linear predictor performed poorly with respect to SSIM but very well with respect to MSE and MAE. Thus, it is a model that should be given some thought as a possibly good model. What is interesting about the results for this line of business is that the LOESS method with the second linear predictor and posted reserves were the worst methods. While the Spearman correlations were low, the gamma and Tweedie models produced ranks that were all some of the highest. This strengthens the argument for these models as relatively good models since the low correlations imply that the goodness of fit statistics are measuring different things.

Finally, for the Workers Compensation line of business the negative binomial with the first linear predictor and gamma model with the second linear predictor appeared to perform the best relative to the other models. SSIM values mainly ranged from .17 to .39. It should also be noted that the Tweedie, and both lognormal models performed poorly under the SSIM but very well under the MSE and MAE criteria. Thus, the Tweedie and lognormal models are also ones that could be under consideration as decent models. The negative binomial and spline models performed the worst. Again, the Spearman correlations were relatively high for the SSIM and MSE, and SSIM and MAE. This adds support to the selection of the models.

One comment should be made when considering the posted reserve compared with the other models. When providing reserve estimates it is common for actuaries to try to consistently produce estimates that are biased towards being too large. This is done because if reserve estimates are too low a company may be more likely to become insolvent or face a credit rating downgrade. This can damage an actuary's professional reputation and even lead to claims of malpractice and lawsuits. Thus, while posted reserves never seemed to perform remarkably well relative to other models, they usually didn't perform poorly either. Some of the low performance is likely a result of the incentive to bias estimates.

Overall, no single model had the best ranks in all situations. For the auto lines it appeared that the Chain Ladder produced the best ranks. For the liability only lines like other liability and product liability the Tweedie appears the best. For medical malpractice, the loess models seemed superior. And finally, for workers compensation the gamma or negative binomial appeared to do best. It should also be noted that at this time there is no similar analysis to compare these results to as this analysis was performed shortly after the data was published. Table 3.3 displays a summary of the results. A more in-depth look at the statistics and rankings can be found in Appendix A.

Table 3.3 A summary of the model results by line of business.

Line of Business	Best Models	Models to Consider	Worst Models
Personal Auto	Chain Ladder Poisson - LP2**	Neg. Binomial - LP2	Lognormal – LP1, LP2 Bayesian N.B.
Commercial Auto	Chain Ladder	Gamma – LP1*, LP2 Lognormal – LP1, LP2 LOESS – LP2	Poisson – LP1, LP2
Other Liability	Tweedie	Poisson – LP1 Neg Binomial – LP2	Spline – LP1, LP2 Gamma – LP1, LP2
Medical Malpractice	LOESS – LP1, LP2	Poisson – LP1	Lognormal – LP1, LP2 Bayesian N.B.
Product Liability	Gamma – LP1 Tweedie	Poisson – LP1 Lognormal – LP2	LOESS – LP2 Posted Reserve
Workers Compensation	Neg. Binomial – LP1 Gamma – LP2	Lognormal – LP1, LP2	Neg Binomial – LP2 Spline – LP1, LP2

***** LP1 refers to the first linear predictor

** LP2 refers to the second linear predictor

Estimate Anomalies

After total unpaid paid loss estimates were calculated for each data set in the different lines of business several anomalies were noted. As stated in prior sections the over-dispersed Poisson, negative binomial, gamma, and lognormal GLMs with the first linear predictor should produce estimates equal to the Chain Ladder method. However, as the companies became smaller in the volume of their data the estimates would often begin to diverge from each other. This was likely caused by adjustments made to the Chain Ladder development factor calculations or irregularities in the likelihood functions GLMs. Either way, low data volume appeared to be the root of the problem as larger data sets produced equivalent estimates.

The other irregularity was the different level of performance that several methods displayed between the personal private auto data set and the commercial private auto data set. This is unusual because personal auto and commercial auto would be expected to be very similar since the coverage, exposures, and types of risks are nearly identical. After viewing the absolute errors produced by the methods between the two data sets it was noted that several methods within the commercial auto data set produced at least one error that was 1,000 to 1,000,000 times larger than any of the other errors for that particular company. Thus, these few excessively poor estimates are likely causing the methods to appear different in performance between these two lines of business.

The final anomaly was that the SSIM statistic failed to calculate for several methods across several lines of business. This appeared to be a result of many reserve estimates that failed to converge due to data issues. When too many estimates failed the corresponding SSIM for that method also failed. The result of this was those methods were not ranked and their rank is only an average of the MSE and MAE. This scenario occurred most frequently with the negative binomial and gamma models with the second linear predictor and on occasion with the gamma model with the first linear predictor. As a result of these anomalies, some methods didn't produce reliable ranks.

Future Work

While this paper has explored various methods of predicting aggregate unpaid loss estimates there is still more work that can be done to explore this topic. One of the first areas would be to explore the relative effectiveness of individual accident year by development period cell estimates. Under the current analysis one method can estimate total unpaid liabilities well simply by consistently over-estimating some cells and underestimating other cells. An analysis by cell would provide a deeper understanding of the effectiveness of the models and provide more data points that can be used to estimate relative effectiveness. It would also be useful in a practical sense as it would allow a better understanding of how these models predict each calendar year which is expressed as a diagonal in the loss triangle. Calendar year estimates are important because they are often used by management for planning purposes. One of the primary reasons that the author decided to look at total unpaid losses was that the NAIC data set contained posted reserves on an aggregate level and not by cell. Since the author was interested in including posted reserves in the analysis this forced the use of total unpaid losses.

Another area of future work would be to evaluate the behavior and performance of the MSE, MAE, and SSIM statistics and ranking procedures. This could accomplished by simulating insurance losses using various loss distributions, using the modeling methods to predict unpaid losses, and using the three statistics to compare the estimated with the actual. The three statistics could be evaluated to determine the ideal ways in which they behave and related to each other. The ranking procedure could also be reviewed and the Spearman correlation coefficient could be utilized to determine how the ranks are ideally related to each other. Looking at this component would facilitate a better selection of the best relative model.

Additionally, other future work would be to explore other Bayesian models that were described earlier in the paper. The Bayesian negative binomial was only explored. However, the Bayesian over-dispersed Poisson, which utilizes an improper prior as the distribution of y_j instead of the non-informative prior employed in the Bayesian negative binomial, could also be reviewed. Additionally, priors could also be selected for the over-dispersion parameter, ϕ. Looking at this additional method would provide another model to evaluate.

Finally, another avenue of opportunity for this analysis would be to loosen some of the constraints placed on the incremental paid losses. As was discussed negative values or values of zero were converted into values of one. This was done to assist in accommodating the likelihood functions for several of the models, most notably the lognormal model. However, since the lognormal never performed remarkably well, and the data represents the entire industry over a 20 year period, it could be excluded from future analysis and thus, assumptions could be relaxed. Likely, this would allow for values to remain negative or at least only be converted to zeros. Loosening constraints would make the modeled data more similar to real data and make the results more closely related to what they would be in a real analysis.

Chapter 4 - Conclusion

Estimating unpaid liabilities for insurance companies is of extreme importance due to the financial and legal ramifications associated with these numbers. This paper has reviewed, utilized, and compared several stochastic and deterministic methods for estimating total unpaid liabilities. It was discovered that the relative effectiveness of a method is largely dependent upon the line of business under review and the characteristics of that unique line of business. However, this analysis provides insight and knowledge to a practitioner to better allow them to estimate this highly important financial information.

Bibliography

Bornhuetter, R.L. & Ferguson, R.E. (1972). The Actuary and IBNR. *Proceedings of the Casualty Actuarial Society*, LIX, 181-195

Casualty Actuarial Society Committee on Ratemaking Principles. (1988). *Statement of Principles Regarding Property and Casualty Loss and Loss Adjustment Expense Reserves*.

Dunn, P.K. (2010). Tweedie: Tweedie Exponential Family Models. R Package Version 2.0.7.

Dunn, P.K. & Smyth, G.K. (2007). Evaluation of Tweedie Exponential Dispersion Model Densities by Fourier Inversion. *Statistics and Computing*, 18

England, P.D. & Verrall, R.J. (2002). Stochastic Claims Reserving in General Insurance. *Institute of Actuaries and Faculty of Actuaries*, 28.

Friedlund, J.F. (2010). *Estimating Unpaid Claims Using Basic Techniques*. Casualty Actuarial Society, Third Version.

Hastie, T. (2010). GAM: Generalized Additive Models. R Package Version 1.03. http://CRAN.R-project.org/package=gam

Mack, T. (1991). A Simple Parametric Model for Rating Automobile Insurance or Estimating IBNR Claims Reserves. ASTIN Bulletin, 22, 1, 93-109.

Meyers, G.G. & Shi, P. (2011). Casualty Actuarial Society. http://www.casact.org/research/index.cfm?fa=loss_reserves_data

Renshaw, A.E. & Verrall, R.J. (1998). A Stochastic Model Underlying the Chain-Ladder Technique. *British Actuarial Journal*, 4, 903-923.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Verrall, R.J. (1996). Claims Reserving and Generalized Additive Models. *Insurance: Mathematics and Economics*, 19, 31-43

Verrall, R.J. (2000). An Investigation into Stochastic Claims Reserving Models and the Chain-Ladder Technique. *Insurance: Mathematics and Economics*, 26, 91-99

Verrall, R.J. (2001). A Bayesian Generalized Linear Model for the Bornhuetter-Ferguson Method of Claims Reserving. *North American Actuarial Journal*

Verrall, R.J. & England, P.D. (2000). Comments on: "A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates, by Mack and Venter. *Insurance: Mathematics and Economics*, 26, 109-111

R Development Core Team (2010). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Smyth, Gordon; Hu, Yifang; Dunn, Peter; and Phipson, Belinda (2010). statmod: Statistical Modeling. R package version 1.4.6. http://CRAN.R-project.org/package=statmod

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Wright, T.S. (1990). A Stochastic Method For Claims Reserving in General Insurance. *Journal of the Institute of Actuaries*. 117, 677-731

Appendix A - Model Summary Statistics

Personal Private Passenger Auto

Model	LP	Statistics: SSIM	MSE	MAE	Ranks: SSIM	MSE	MAE	Mean
Poisson	LP 2	0.407	4.208E+10	39,945	3	1	2	2.00
Chain Ladder		0.465	4.450E+10	40,882	1	3	3	2.33
Neg Binomial	LP 2	NA	4.540E+10	38,413	NA	4	1	2.50
Tweedie		0.334	4.357E+10	42,681	7	2	4	4.33
Loess	LP 2	0.368	4.541E+10	43,643	4	5	5	4.67
Loess	LP 1	0.358	4.589E+10	44,139	5	6	7	6.00
Neg Binomial	LP 1	0.328	4.801E+10	46,518	8	9	9	8.67
Gamma	LP 2	0.280	4.625E+10	43,664	13	7	6	8.67
Poisson	LP 1	0.335	4.822E+10	47,483	6	10	11	9.00
Posted Reserve		0.411	1.034E+11	62,402	2	13	13	9.33
Lognormal	LP 2	0.277	4.690E+10	44,560	14	8	8	10.00
Gamma	LP 1	0.292	4.898E+10	46,717	11	11	10	10.67
Lognormal	LP 1	0.283	5.250E+10	51,630	12	12	12	12.00
Spline	LP 2	0.326	1.039E+12	177,612	10	15	15	13.33
Spline	LP 1	0.328	2.641E+19	391,355,818	9	16	16	13.67
BFCL		0.102	4.571E+11	132,529	15	14	14	14.33

Commercial Private Passenger Auto

Model	LP	Statistics: SSIM	MSE	MAE	Ranks: SSIM	MSE	MAE	Mean
Gamma	LP 2	NA	7.757E+08	7,045	NA	3	1	2.00
Chain Ladder		0.282	8.743E+08	7,272	1	6	5	4.00
Tweedie		0.225	7.907E+08	7,195	6	4	4	4.67
Gamma	LP 1	0.196	8.391E+08	7,072	9	5	2	5.33
Lognormal	LP 1	0.156	6.873E+08	7,177	12	2	3	5.67
Lognormal	LP 2	0.132	5.872E+08	7,375	13	1	6	6.67
Posted Reserve		0.188	1.098E+09	8,798	11	7	7	8.33
Neg Binomial	LP 1	0.192	1.615E+09	9,314	10	8	8	8.67
Spline	LP 1	0.237	4.281E+13	579,994	3	11	12	8.67
Neg Binomial	LP 2	NA	2.136E+09	9,659	NA	9	9	9.00
Spline	LP 2	0.234	4.666E+13	556,388	4	12	11	9.00
Loess	LP 2	0.237	6.364E+13	820,168	2	13	13	9.33
BFCL		0.131	1.006E+10	23,084	14	10	10	11.33
Loess	LP 1	0.232	6.559E+14	2,680,069	5	15	15	11.67
Poisson	LP 1	0.212	1.224E+14	833,585	8	14	14	12.00
Poisson	LP 2	0.224	7.185E+18	206,388,905	7	16	16	13.00

Other Liability

		Statistics			Ranks			
Model	***LP***	***SSIM***	***MSE***	***MAE***	***SSIM***	***MSE***	***MAE***	***Mean***
Tweedie		0.136	6.766E+09	14,578	2	3	1	2.00
Neg Binomial	***LP 2***	NA	6.527E+09	15,302	NA	2	3	2.50
Poisson	***LP 1***	0.122	6.436E+09	14,701	5	1	2	2.67
Loess	***LP 2***	0.139	3.160E+10	25,053	1	8	6	5.00
Loess	***LP 1***	0.134	2.918E+10	25,451	3	7	7	5.67
Chain Ladder		0.112	1.842E+10	24,786	11	5	5	7.00
BFCL		0.070	9.736E+09	23,672	13	4	4	7.00
Lognormal	***LP 1***	0.123	6.781E+10	42,127	4	10	10	8.00
Lognormal	***LP 2***	0.121	4.151E+10	31,176	7	9	8	8.00
Posted Reserve		0.113	2.471E+10	31,239	10	6	9	8.33
Poisson	***LP 2***	0.117	1.071E+13	306,073	8	11	12	10.33
Neg Binomial	***LP 1***	0.116	1.221E+13	240,869	9	12	11	10.67
Spline	***LP 2***	0.122	8.949E+18	195,574,773	6	13	13	10.67
Spline	***LP 1***	0.103	2.617E+19	334,471,696	12	14	14	13.33
Gamma	***LP 1***	NA	2.270E+39	2.965E+18	NA	15	15	15.00
Gamma	***LP 2***	NA	5.040E+119	4.333E+58	NA	16	16	16.00

Medical Malpractice

		Statistics			Ranks			
Model	***LP***	***SSIM***	***MSE***	***MAE***	***SSIM***	***MSE***	***MAE***	***Mean***
Loess	***LP 1***	0.600	2.512E+09	17,871	2	1	1	1.33
Loess	***LP 2***	0.588	2.526E+09	17,989	3	2	2	2.33
Poisson	***LP 1***	0.565	3.530E+09	19,605	5	4	4	4.33
Spline	***LP 2***	0.535	2.753E+09	19,081	10	3	3	5.33
Chain Ladder		0.557	3.570E+09	20,327	7	5	5	5.67
Posted Reserve		0.666	4.335E+09	30,070	1	7	13	7.00
Spline	***LP 1***	0.542	3.728E+09	23,759	8	6	8	7.33
Gamma	***LP 1***	0.566	9.235E+09	23,539	4	13	7	8.00
Tweedie		0.538	5.837E+09	23,019	9	10	6	8.33
Poisson	***LP 2***	0.490	5.219E+09	27,604	12	8	11	10.33
Neg Binomial	***LP 1***	0.492	9.555E+09	25,558	11	14	9	11.33
Lognormal	***LP 1***	0.432	9.070E+09	29,237	14	11	12	12.33
Neg Binomial	***LP 2***	0.477	1.040E+10	26,934	13	15	10	12.67
Gamma	***LP 2***	0.565	5.007E+23	1.129E+11	6	16	16	12.67
BFCL		0.108	5.792E+09	34,061	16	9	15	13.33
Lognormal	***LP 2***	0.380	9.128E+09	31,614	15	12	14	13.67

Product Liability

Model	LP	Statistics SSIM	Statistics MSE	Statistics MAE	Ranks SSIM	Ranks MSE	Ranks MAE	Ranks Mean
Gamma	***LP 1***	0.346	1.003E+09	7,742	2	2	1	1.67
Tweedie		0.345	1.089E+09	9,549	3	3	3	3.00
Neg Binomial	***LP 2***	0.313	1.237E+09	10,488	9	4	4	5.67
Poisson	***LP 1***	0.351	3.023E+09	13,295	1	10	7	6.00
Lognormal	***LP 2***	0.139	9.795E+08	8,344	16	1	2	6.33
Gamma	***LP 2***	0.317	1.693E+09	10,887	8	6	5	6.33
Poisson	***LP 2***	0.330	2.718E+09	14,970	4	9	11	8.00
Lognormal	***LP 1***	0.219	1.585E+09	11,482	13	5	6	8.00
Neg Binomial	***LP 1***	0.295	2.111E+09	14,181	11	8	9	9.33
Spline	***LP 2***	0.327	6.916E+09	17,183	5	12	12	9.67
Chain Ladder		0.267	3.108E+09	13,934	12	11	8	10.33
BFCL		0.188	1.910E+09	14,322	14	7	10	10.33
Loess	***LP 1***	0.324	9.230E+09	18,891	6	14	13	11.00
Spline	***LP 1***	0.319	7.128E+09	18,987	7	13	14	11.33
Loess	***LP 2***	0.308	9.945E+09	19,607	10	15	15	13.33
Posted Reserve		0.177	2.430E+10	41,783	15	16	16	15.67

Workers Compensation

Model	LP	Statistics SSIM	Statistics MSE	Statistics MAE	Ranks SSIM	Ranks MSE	Ranks MAE	Ranks Mean
Gamma	***LP 2***	0.320	1.693E+09	13,251	3	1	1	1.67
Lognormal	***LP 1***	0.275	2.052E+09	15,816	7	2	4	4.33
Neg Binomial	***LP 1***	0.335	4.731E+09	17,383	2	5	8	5.00
Gamma	***LP 1***	0.393	4.934E+09	17,935	1	6	9	5.33
Chain Ladder		0.292	7.090E+09	16,004	4	8	5	5.67
Tweedie		0.246	3.605E+09	14,970	12	4	2	6.00
Lognormal	***LP 2***	0.193	2.808E+09	15,659	14	3	3	6.67
Loess	***LP 2***	0.257	7.599E+09	16,675	8	9	6	7.67
Loess	***LP 1***	0.250	6.846E+09	16,692	10	7	7	8.00
Poisson	***LP 1***	0.280	7.638E+09	20,095	6	10	10	8.67
Poisson	***LP 2***	0.282	1.011E+11	42,552	5	13	12	10.00
BFCL		0.177	2.742E+10	40,252	15	11	11	12.33
Posted Reserve		0.241	3.942E+10	47,056	13	12	13	12.67
Spline	***LP 1***	0.247	1.051E+17	30,699,921	11	14	14	13.00
Spline	***LP 2***	0.257	2.437E+17	42,736,376	9	15	15	13.00
Neg Binomial	***LP 2***	NA	1.696E+82	1.038E+40	NA	16	16	16.00

Spearman Correlations

	SSIM/MSE	SSIM/MAE	MSE/MAE
Personal Auto	0.318	0.188	0.974
Commercial Auto	-0.447	-0.394	0.938
Other Liability	0.309	0.412	0.968
Medical Malpractice	0.450	0.503	0.718
Product Liability	0.032	0.156	0.944
Workers Compensation	0.388	0.303	0.932