Estimating Loss Reserves for Dental Claims

Yun Bai
Curt Baragar
Angela Olandese
Nandita Singh
Rahila Sunesara

May 2, 2005

Abstract Insurers are required to estimate a liability at the end of every month for dental services already performed, but for which payment has not been made. These are typically claims that have not yet been reported or are under investigation. The traditional methods of making this estimate are viewed as being inadequate in light of the computing power and accessibility of data. We developed two models for estimating dental claims based on modern techniques and compared their results. We found that one model, based on a time series analysis, performed very well on many data sets. We also found that there appears to be no model that works well on all data sets, suggesting that each client of Delta Dental must be considered individually when estimating loss reserves.

Work done in partial fulfillment of the requirements of Michigan State University MTH 844, advised by Toby Hall, Associate Actuary, Delta Dental Plan of Michigan, and Professor Peiru Wu, MSU mathematics.

Table of Contents

Introduction

Delta Dental Plans Association is a nationwide non-profit organization of independently operated dental health service plans. Delta Dental Plan of Michigan, created in 1957, is now the second largest affiliation within the whole association; approximately 90% of Michigan's dentists are current participants of the dental health service network, providing coverage to more than 3.9 million people in over 2,500 Michigan-based groups.

Delta Dental Plan of Michigan, like all insurers, is required to estimate a liability at the end of every month for dental services already performed, but for which payment has not been made. These are typically claims that have not yet been reported or are under investigation. The traditional methods of making this estimate are viewed as being inadequate in light of the computing power and accessibility of data. They were developed during WWII, when computing devices were less common than they are today. These models thus relied on as little data input as possible, since number crunching was both time-consuming and costly. In modern times, with the availability of computers and much more data, we have the ability to make a more accurate model.

Some of these newer models have already been applied to medical insurance, car insurance, and professional liability insurance. Dental insurance, however, has unique characteristics which require a separate evaluation of the methods. For instance, dental claim amounts are generally small when compared to medical, auto, and liability claims. Also, the *run-out*, which is the duration of time between the service being performed and the payout of the claim, is shorter than that of health insurance, averaging only 2 to 3 months. Another characteristic of dental claims is that there are usually spikes of dental services in late December, early January, and also during the summer months. This is caused by dental plan policies and the fact that many people have more time for dental visits during the summer. Thus, these newer estimation models, which have performed well for insurance in other industries, may not be as accurate when applied to dental insurance.

In this paper, we develop two models for estimating outstanding claims. We use data from General Motors employees provided by Delta Dental in order to compare these methods. The first method considered uses a log-linear regression model, and the second uses a time-series approach. We will show how these approaches were implemented and how well they estimated unpaid dental claims.

Loss Reserves

One of the most important functions performed by an actuary in the property/casualty insurance industry is loss reserving, whereby the actuary determines the present liability associated with future claim payments [1]. These claim reserves are created when an event such as a teeth cleaning has occurred, but payment has not been made as of the *valuation date*, which is the date upon which accounting entries are closed and future claim estimates are established.

For instance, suppose a policyholder goes to her dentist on Monday to have a cavity removed. Since the fee for this service will be paid by her provider, after the procedure her dentist submits a claim to them for reimbursement. Let us assume that this dentist, like many others, submits all claims for the week on Friday, and then within a few days they arrive in the claims office of the dental insurance provider. Suppose however, that Wednesday is the valuation date. Since this dentist is not going to submit his claims until after the valuation date, these claims are referred to as *incurred but not reported*, or IBNR. Once the dentist submits these claims, there will be some lag time before he is reimbursed. This is because of among other things, the time the claims are in the mail delivery system, and processing time at the provider's claims office. During this lag time, these claims are referred to as *reported but not paid*, or RBNP.

The total claim reserve is made up of the two previously mentioned pieces. The much larger piece is the IBNR, and the smaller portion of the total claim reserve is the RBNP. Delta Dental of Michigan must estimate approximately 20 claim reserves each month, which requires many hours of computations.

Accuracy in estimating these reserves is crucial. In fact, most jurisdictions now require by law that a qualified actuary attest to the adequacy and appropriateness of the insurer's loss reserves [1]. These reserves are by far the largest liabilities carried by property/casualty insurers. Also, by attesting that the insurer has set aside enough money to pay all obligations that already exist, the actuary is providing policyholders with assurance that their benefits will be paid. If the reserves are appropriate, there should be enough funds to guarantee the existing obligations to the policyholders even if the insurer were declared insolvent.

Traditional Methods of Estimating

Most actuaries today develop the bulk of their loss reserve liability by using some variation of the chain-ladder method [2], which utilizes the month-to-month ratios between cumulative payments as a basis for projection. These ratios, known as *development factors*, are then used to estimate the ultimate payments for each month. So the expected ultimate payments minus the amount already paid, represents the reserve requirement. Thus the chain-ladder method has the underlying assumption that future reserves may be sufficiently predicted by those of the past.

Chain-Ladder Method

In order to use this method, data must be summarized by the month incurred versus the number of lag months before payout of the claim. The paid claims must also be compiled cumulatively. The resulting grid, shown in Figure 1, is often referred to as the *claims triangle*. Used with dental claims, the chain-ladder method usually requires at least 24 months of data to develop reasonable development factors. In our abbreviated example, we will use 12 months of data to illustrate the process [3].

	Lag Months											
Incured Month	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Jan	1750	4250	5300	6000	6125	6195	6230	6255	6270	6300	6305	6305
Feb	1500	4250	5800	6525	6625	6700	6730	6740	6765	6790	6795	
Mar	1600	4400	6050	6725	6850	6900	6925	6965	6985	6995		
Apr	1125	3775	5325	6065	6135	6180	6235	6250	6255			
May	1900	4100	5575	6225	6285	6345	6385	6400				
Jun	1500	4100	5450	6100	6200	6260	6290					
Jul	2000	3950	5520	6270	6360	6420						
Aug	1700	4000	5650	6350	6465							
Sep	1400	4300	5740	6465								
Oct	2225	4625	6150									
Nov	1700	3650										
Dec	1575											

Figure 1. Cumulative paid claims.

The goal of the actuary is to now fill in the bottom half of the triangle as accurately as possible. In order to accomplish this, the month-to-month ratios between cumulative payments are used as a basis for projection. These ratios, along with their averages, are shown in Figure 2. These average values are the development factors that are used to fill in the lower half of the claims triangle.

	Ratio of Successive Lag Months										
Incured Month	**1/0**	**2/1**	**3/2**	**4/3**	**5/4**	**6/5**	**7/6**	**8/7**	**9/8**	**10/9**	**11/10**
Jan	2.429	1.247	1.132	1.021	1.011	1.006	1.004	1.002	1.005	1.001	1.000
Feb	2.833	1.365	1.125	1.015	1.011	1.004	1.001	1.004	1.004	1.001	
Mar	2.750	1.375	1.112	1.019	1.007	1.004	1.006	1.003	1.001		
Apr	3.356	1.411	1.139	1.012	1.007	1.009	1.002	1.001			
May	2.158	1.360	1.117	1.010	1.010	1.006	1.002				
Jun	2.733	1.329	1.119	1.016	1.010	1.005					
Jul	1.975	1.397	1.136	1.014	1.009						
Aug	2.353	1.413	1.124	1.018							
Sep	3.071	1.335	1.126								
Oct	2.079	1.330									
Nov	2.147										
average	2.535	1.356	1.126	1.016	1.009	1.006	1.003	1.003	1.003	1.001	1.000

Figure 2. Computing the development factors.

In reality, the average is not often used to determine the development factors. It is often advantageous to use a 3-month or a 6-month average, which more accurately reflect trends, but there are a wide variety of averaging approaches that may be used. Also, actuaries typically employ a range of methods to smooth the month-to-month variations in development factors in order to obtain a final set of projections.

Once the development factors are determined, simple multiplication allows the bottom half of the claims triangle to be completed. For instance, for the claims incurred in December from Figure 1, the chain-ladder method predicts that $1575 × 2.535 = $3993 in claims will be paid out by the end of January. Similar calculations are used to complete the triangle, which can be seen in Figure 3.

	Lag Months											
Incured Month	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
Jan	1750	4250	5300	6000	6125	6195	6230	6255	6270	6300	6305	6305
Feb	1500	4250	5800	6525	6625	6700	6730	6740	6765	6790	6795	6795
Mar	1600	4400	6050	6725	6850	6900	6925	6965	6985	6995	7002	7002
Apr	1125	3775	5325	6065	6135	6180	6235	6250	6255	6274	6280	6280
May	1900	4100	5575	6225	6285	6345	6385	6400	6419	6438	6445	6445
Jun	1500	4100	5450	6100	6200	6260	6290	6309	6328	6347	6353	6353
Jul	2000	3950	5520	6270	6360	6420	6459	6478	6497	6517	6523	6523
Aug	1700	4000	5650	6350	6465	6523	6562	6582	6602	6622	6628	6628
Sep	1400	4300	5740	6465	6568	6628	6667	6687	6707	6728	6734	6734
Oct	2225	4625	6150	6925	7036	7099	7142	7163	7185	7206	7213	7213
Nov	1700	3650	4949	5573	5662	5713	5747	5765	5782	5799	5805	5805
Dec	1575	3993	5415	6097	6194	6250	6288	6306	6325	6344	6351	6351

Figure 3. The completed claims triangle.

When using the chain-ladder method on General Motors data from January 2001 to December 2002, we get results that are fairly consistent, but not very accurate. The results for the predicted total claims are shown in Table 1. Notice that the error rates increase for later months as more estimations are performed in filling in the bottom half of the claims triangle. For the incurred month of January 2002, for instance, only 12 claims estimates are required, whereas for December 2002, 23 claims estimates are needed to complete the bottom row of the triangle. The error rates are adjusted in the next column to show the results of a full 23 months of predictions.

Table 1. The chain-ladder results are not very accurate.

Incurred Month	Estimated Total Reserve	Actual Total Reserve	Error Rate	Error Rate (Adj)
January 2002	$20,700,651	$22,510,971	-0.08042	-0.15414
February 2002	$17,020,042	$18,592,943	-0.08460	-0.14967
March 2002	$16,603,872	$18,204,862	-0.08794	-0.14448
April 2002	$17,766,278	$19,555,651	-0.09150	-0.14030
May 2002	$17,059,650	$18,852,790	-0.09511	-0.13672
June 2002	$15,249,024	$16,914,862	-0.09848	-0.13324
July 2002	$15,233,693	$16,933,302	-0.10037	-0.12825
August 2002	$15,310,938	$17,077,299	-0.10343	-0.12521
September 2002	$14,510,293	$16,236,981	-0.10634	-0.12229
October 2002	$16,393,411	$18,447,298	-0.11134	-0.12194
November 2002	$14,017,010	$15,931,985	-0.12020	-0.12566
December 2002	$14,274,114	$16,968,877	-0.15881	-0.15881

Modified Chain-Ladder

Delta Dental has been using a modified chain-ladder method to estimate claim reserves. This method is based on the fact that dental claims are incurred more often on certain days of the week. More people visit the dentist on Tuesday than on Monday or Friday, for instance. Also, very few dentists are open on weekends, so that not many claims are incurred on these days. Delta Dental finds a *weighted average* for each day and uses this information to predict the total amount of claims that will be submitted in say, January of 2006, which has five Tuesdays – one more than that of the previous January. This allows Delta Dental to take into account that some months have five weekends, which allows for fewer working days than the more typical four-weekend month.

Delta Dental analyzed their in-house data from the years 2002 and 2003 to establish these weighted averages. They found that Tuesday was the most popular day for dental procedures and thus assigned this day a weight of 1. All other days were weighted accordingly, as shown in Table 2.

Table 2. Day of dental procedure – weighted averages.

	2002	2003
Sunday	.0047	.0041
Monday	.9433	.9476
Tuesday	1.0000	1.0000
Wednesday	.8367	.8349
Thursday	.9031	.9002
Friday	.4101	.3945
Saturday	.1086	.1054

The next step is to determine how many of each of the days of the week is in each month of the year of interest. For example, how many Sundays, Mondays, etc. are in March of 2003? This, of course, must be calculated for each year. These numbers are then multiplied by their respective weighted average and the results are combined, resulting in the number of weighted days per month, shown in Table 3.

Table 3. Number of weighted days per month.

	2002	2003
January	18.0596	18.0744
February	16.7735	16.7735
March	17.1055	17.8304
April	18.7196	18.4305
May	17.8492	17.1201
June	16.8842	17.7239
July	17.7250	18.3866
August	18.0264	16.9270
September	16.7778	17.7735
October	19.5102	18.9099
November	15.9512	15.5558
December	15.7617	17.0650

Because dental visits are very seasonal, instead of using the previous month as a predictor of the current month (as the standard chain-ladder method does), Delta Dental uses information from the current month of the previous year. Delta Dental figures out the claims per subscriber, per day of one year ago. They do this by dividing the total incurred claims in the current month of the previous year by the number of subscribers at that time. They then divide this number by the number of weighted days for that month to come up with the number of incurred claims per subscriber per day for the current month of one year ago.

There is one other factor that Delta Dental considers, and this one requires a professional actuary. They use a multiplier that is called the *trend* or *adjustment factor* that takes into account such things as inflation, remarkably bad weather (which keeps people indoors and out of the dentist's office), price increases, or the event that the entire office staff at Delta Dental took the last two days of the month off as a special company promotion, thus delaying the processing of some claims until the following month.

Combining these factors leaves us with Delta Dental's method of estimating monthly claims:

$$\text{Monthly Claim Estimate} = I \times T \times S \times D,$$

where I is the number of incurred claims per subscriber per weighted day for the current month of the previous year, T is the trend, S is the number of current subscribers, and D is the number of weighted days for the month.

This method has proven to be fairly accurate, with error rates usually around ± 2½ %, but it is not automated and can take a considerable amount of time to perform.

Designing a Model

The perfect model would be one that was almost entirely automated. The actuary would simply input the client name, the time period in question, and some other details, and the computer would determine the requested loss reserve. In the search for this model, we investigated many ideas from actuarial and statistical literature, and finally decided on two ideas that seemed the most promising.

Log-Linear Regression

For our initial attempt at improving Delta Dental's current process of estimating claims, we implemented a log-linear regression model using S-PLUS statistical software. We started with a basic formula for computing the IBNR claim amount, which determines how much money must be set aside monthly as the loss reserve:

$$\textit{claim dollar amount} = \textit{average cost per procedure} \times \textit{number of procedures}.$$

In computing the average cost of a dental procedure, we used a moving 3-month average by considering the most recent three months and using equal lag time. As an example, suppose we would like to predict the average payment of a dental procedure that incurred in October 2004 and was paid in December 2004. We used the information provided in the Table 4 below to compute this value.

Table 4. Computing the average payment.

Incurred Month	Paid Month	Amount Paid	Est. Number of Procedures	Average Cost per Procedure
July 2004	September 2004	$168,792.00	1721	$98.07786
August 2004	October 2004	$262,189.37	2769	$94.68739
September 2004	November 2004	$204,855.35	2026	$101.11320

Thus the estimated average payment of a dental procedure incurred in October, 2004 and paid in December, 2004 is

$$\frac{\$98.07786 + \$94.68739 + \$101.1132}{3} \approx \$97.96.$$

We used a few different techniques to arrive at the values in Table 4. The number of procedures is estimated from a log-linear regression function. The input variables of this function are the number S of subscribers during the incurred month, the number D of seasonality-adjusted dental service days of the incurred month, the number N_0 of dental procedures incurred and paid during the incurred month, and the number N_C of dental procedures paid in the current month.

So we have

$$\log(N_C) = a\log(S) + b\log(D) + c\log(N_0),$$

where *a, b*, and *c* are parameters obtained from using the log-linear regression function on the previous 12 months' claim records.

It is very common for the employees' dental insurance plans to also cover their spouses and children . Therefore, we considered the number of subscribers as a weighted sum of the employees, their spouses, and their dependents. The employees' weight is automatically 1 and the weights for their spouses and dependents are determined by the dental care policies. For example, if employees, spouses, and children are all equally insured by their policy, we would assign weights of 1 to all of them. In other cases, the weights are real numbers varying from 0 to 1.

As we discussed earlier, the actual numbers of service days for different months are not equal, which requires careful adjustments for seasonality. Here we considered two factors for seasonality adjustments. First, we adjusted the number of working days for weekdays and holidays. We computed the number of Sundays, Mondays, Tuesdays, etc. of each month for a given year, and then subtracted off the days falling on those holidays when dentists' offices are normally closed. Using an empirical set of weights for the seven days of a week, we then computed a weighted sum of service days for each month as described previously.

We then computed the ratio of claim/subscriber/day for the past 12 months, used the maximum ratio as the denominator, and computed a set of weights for the 12 months of the given year. Before we applied these monthly weights to the weighted numbers of days described earlier, we smoothed them by raising them to a power between 0 and 1. When the smoothing power was chosen as 0, all the monthly weights became 1's and no adjustment was performed. When it was chosen to be 1, the monthly weights had the sharpest impact on the results. We decided to use a smoothing factor of 0.5.

The number of dental claims incurred and paid in the current month is a very important factor in estimating IBNR claims. As can be seen in Figure 4, although patterns of fluctuations of dental services occurring on different months are not obvious, the decreasing pattern run-out patterns of IBNR claims is very consistent. Thus, once we knew the size of dental claims incurred and paid in the current month, it was easier to predict those claims incurring in that month but being paid in later months.

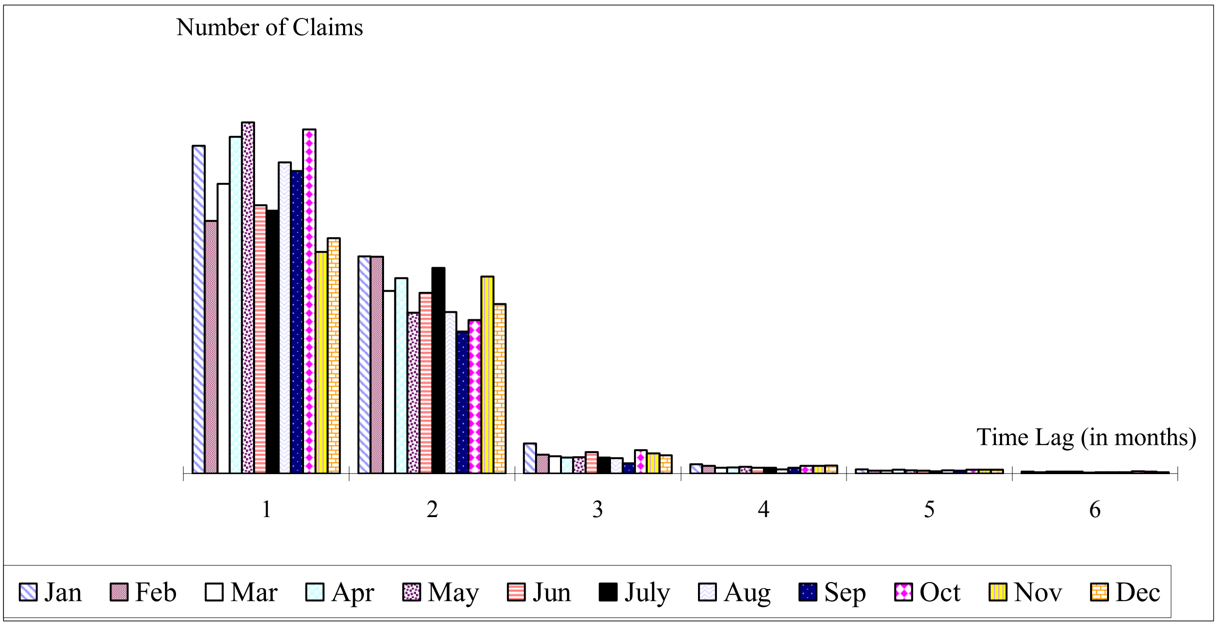

Figure 4. Run-out distribution of GM in 2002.

We noticed that the log-linear regression function of the past 12 months' claim records usually had a very high degree of linearity. Table 5 shows several multiple R-square coefficients of the regression functions for the GM data for 2002.

Table 5. A high degree of linearity can be seen.

Incurred Month	Paid Month	coefficient of determination, r^2
January 2002	February 2002	0.9999
February 2002	March 2002	1
March 2002	April 2002	1
April 2002	May 2002	1
May 2002	June 2002	1
June 2002	July 2002	1
July 2002	August 2002	1
August 2002	September 2002	1
September 2002	October 2002	1
October 2002	November 2002	1
November 2002	December 2002	1

This is a very interesting discovery because it shows that the function fits the data perfectly. However, it doesn't produce a similar high degree of accuracy when used with the data for predicting claims.

We applied the log-linear regression model on data from various corporate plans, and compared the predicted values with the actual claims. The results for GM 2002 data are shown in Table 6.

Table 6. Log-linear error rates for GM 2002 data.

Incurred Month	Paid Month	Estimated Reserve	Actual Reserve	Error Rate
January 2002	February 2002	$8,027,954	$8,857,533	-0.09366
February 2002	March 2002	$6,820,038	$8,505,161	-0.19813
March 2002	April 2002	$6,688,719	$7,250,155	-0.07744
April 2002	May 2002	$6,706,167	$7,516,271	-0.10778
May 2002	June 2002	$6,919,037	$6,328,022	0.09340
June 2002	July 2002	$6,740,350	$6,547,831	0.02940
July 2002	August 2002	$7,711,536	$7,548,472	0.02160
August 2002	September 2002	$6,874,790	$5,871,233	0.17093
September 2002	October 2002	$5,630,507	$5,491,918	0.02524
October 2002	November 2002	$7,112,756	$5,469,690	0.30039
November 2002	December 2002	$7,269,077	$7,113,821	0.02182

The error rates typically fall within a range of ±10%, except for some extremely "bad" months where the error rates reach as high as 30%. We also tested the model on other data sets and found the same problem. The variation in the error rate is very high and it is difficult to tell in advance if the prediction is going to be accurate.

Time Series – The Holt-Winters Seasonal Algorithm

In light of the problems with consistency we had with the log-linear regression model, we investigated a method that seemed better able to handle the fluctuations in the claims data. We will start with an introduction to the Holt-Winters Seasonal Algorithm [4], which is designed to handle both trend and seasonal variation in the data.

Suppose we have a series of measurements $Y_1, Y_2, \ldots, Y_n$ which contains both trend and seasonality with period d. Then the *forecast function* that takes this into account is

$$F_{n+h} = (\hat{a}_n + \hat{b}_n h)\,\hat{c}_{n-d+h}, \qquad h = 1, 2, \ldots, \tag{1}$$

where F_{n+h} is called the *forecast value*, and $\hat{a}_n$, $\hat{b}_n$, and $\hat{c}_n$ can be thought of as estimates of the "trend level" $\hat{a}_n$, "trend slope" $\hat{b}_n$, and "seasonal component" $\hat{c}_n$ at time n. If k is the smallest integer such that $n + h - kd \le n$, then we set

$$\hat{c}_{n+h} = \hat{c}_{n+h-kd}, \qquad h = 1, 2, \ldots,$$

while the values of $\hat{a}_i, \hat{b}_i$, and $\hat{c}_i$, $i = d+2, \ldots, n$, are found from the recursions

$$\hat{a}_n = \alpha(Y_n / \hat{c}_{n-d}) + (1-\alpha)(\hat{a}_{n-1} + \hat{b}_{n-1}), \quad (2)$$

$$\hat{b}_n = \beta(\hat{a}_n - \hat{a}_{n-1}) + (1-\beta)\hat{b}_{n-1}, \quad (3)$$

and

$$\hat{c}_n = \gamma(Y_n / \hat{a}_n) + (1-\gamma)\hat{c}_{n-d}, \quad (4)$$

with initial conditions

$$\hat{a}_i = (Y_i + Y_{i+1} + \ldots + Y_{d+i-1}) / d,$$

$$\hat{b}_i = (Y_{d+1} - Y_i) / d,$$

and

$$\hat{c}_i = Y_i / a_i, \qquad i = 1, 2, \ldots, d.$$

Then (2), (3), and (4) can be solved successively for $\hat{a}_i, \hat{b}_i$, and $\hat{c}_i$, $i = d+1, \ldots, n$, and for the forecast value F_{n+h} found in (1).

These forecasts depend, of course, on the parameters α, β, and γ. These can either be chosen arbitrarily (with values between 0 and 1) or chosen in a more systematic way to minimize the sum of squares of the one-step errors $\sum_{i=d+2}^{n}(Y_i - P_{1-i}Y_i)^2$, obtained when the algorithm is applied to the already observed data.

We wrote a routine in SAS and used this to apply this technique to data from both GM and DaimlerChrysler. Table 7 shows the results of the predicted values for GM, which were very encouraging. These values were computed using a 95% confidence interval. None of the months had error rates worse than 5% and in fact, 9 of the 11 months had error rates better that 2%.

Table 7. Time series error rates for GM data.

Paid Month	Estimated Reserve	Actual Reserve	Error Rate
January 2004	$22,615,735	$22,853,981	-0.01042
February 2004	$19,511,385	$19,870,188	-0.01806
March 2004	$22,152,027	$22,320,962	-0.00757
April 2004	$19,733,071	$19,681,708	0.00261
May 2004	$18,951,527	$18,973,871	-0.00118
June 2004	$19,888,948	$19,836,048	0.00267
July 2004	$17,421,289	$16,869,593	0.03270
August 2004	$19,091,069	$19,080,265	0.00057
September 2004	$17,942,679	$17,645,437	0.01685
October 2004	$17,735,690	$17,663,004	0.00412
November 2004	$18,724,324	$17,995,974	0.04047

The same technique was less impressive when applied to the DaimlerChrysler data, however. We see in Table 8 that the error rates in some months have grown to unacceptable levels. In fact, 3 of the 11 months have error rates worse than 10%, and in November the error rate jumps to 63%.

Table 8. Time series error rates for DaimlerChrysler data.

Paid Month	Estimated Reserve	Actual Reserve	Error Rate
January 2004	$7,086,258	$6,254,770	0.13294
February 2004	$6,437,994	$7,772,106	-0.17165
March 2004	$7,702,400	$8,377,119	-0.08054
April 2004	$6,679,047	$6,783,321	-0.01537
May 2004	$6,208,121	$6,296,853	-0.01409
June 2004	$6,081,385	$6,750,757	-0.09916
July 2004	$5,694,408	$5,651,267	0.00763
August 2004	$6,258,938	$6,133,777	0.02041
September 2004	$5,744,592	$5,832,109	-0.01501
October 2004	$5,798,602	$6,186,072	-0.06264
November 2004	$9,615,509	$5,888,723	0.63287

Conclusion

We found many techniques that did not estimate dental claims with consistent accuracy. Some techniques worked well for certain data sets and performed poorly on others. We noticed that some of the models we investigated such as the chain-ladder and the time series required 24 to 36 months of historical data, while the log-linear model required only 13. So for new clients, modeling with time series may not be possible. We conclude that there seems to be no model that will be both accurate and consistent for all data sets. This suggests that each client of Delta Dental must be considered individually when selecting a model for estimating loss reserves. Further analysis may reveal patterns that allow all clients to be grouped according to policy terms, claim lags, seasonal variations, and trends. This may allow loss reserves to be accurately calculated by implementing a few models such as the two mentioned in this paper.

Recommendations

The Holt-Winters time series model seems to have the greatest potential for both accuracy and consistency. We found that as a predictor of dental reserves, this model is typically accurate to within plus or minus 3%, although it is much less accurate when used with certain data sets, such as those of DaimlerChrysler. Further investigations of this method may reveal techniques for improving this problem. Based on our research and data analysis, we recommend these investigations be undertaken. We believe that a relatively small amount of effort in this area has the potential to produce a consistently accurate model for predicting dental claims.

Acknowledgments

Hall, Toby. Mr. Hall is an associate actuary at Delta Dental Plan of Michigan. He explained the problem of estimating claims and was a never-ending source of suggestions and ideas.

Wu, Peiru. Dr. Wu is a professor of mathematics at Michigan State University and acted as our faculty manager on this project. She kept the team on track and mindful of our goals throughout the project.

References

[1] Brown, R., and Gottlieb, L., *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*, 2nd ed., ACTEX publications, Boston, 2001, pp. 109.

[2] Lynch, R., "Calculation of IBNR Reserves with Low Variance," *Health Section News*, August 2003, pp. 20.

[3] Lloyd, J., *Health Reserves*, Society of Actuaries, USA, 2000, pp. 22.

[4] Brockwell, P., and Davis, R., *Introduction to Time Series and Forecasting*, 2nd ed., Springer, New York, 2002, pp. 326.

Appendix I – SAS Code for Holt-Winters Algorithm

```
data GM;
   input date inc @@;

   title 'GM Analysis';
   datalines;
;



proc gplot data=GM;
   /*format date year4.;*/
   plot inc*date;
run;


proc forecast data=GM
              interval=month
              trend=2         /* fit a linear trend model          */
              method=winters     /* use winters smoothing method  */
              out=GMout1
              lead=12
                    seasons = month
                    outest = est1
                    outlimit
                    out1step
                    alpha = .05
              outstd;
   var inc;
   id date;
run;
data est1b;
  set est1(keep=_type_ inc rename=(inc=value));
title 'parameter estimates for winters trend=2, with season';
proc print data=est1b;



proc print data=GMout1;
run;
data est1b;
  set est1(keep=_type_ inc rename=(inc=value));
title 'parameter estimates for winters trend=2, no season';
proc print data=est1b;
```

Appendix II – S-Plus Code for Estimating Claims

```
"ddreserve.month"<-
function(service.file, subscriber.file, month,year, smooth=0.5, lags=2,
wt.sp=0.5, wt.kids=0.5){

# Functionality: Generate the estimation of loss reserve of a given
month and year
# Inputs Description:
# service.file: the filename of the previous claim records
# subscriber.file: the filename of the subscribers
# month: the paid month of IBNR claims
# year: the paid year of IBNR claims
# smooth: smoothing factor for the log-linear regression
# lags: the difference between the earliest incurred month and the
current month
# wt.sp: the weight given to the spouse
# wt.kids: the weigt given to the dependants

        new.chart<-create.serv.chart(service.file)
        sub.chart<-create.sub.chart(subscriber.file)
        year00<-weight.day(2000,7)
    year01<-weight.day(2001,1)
    year02<-weight.day(2002,2)
    year03<-weight.day(2003,3)
    year04<-weight.day(2004,4)
    year05<-weight.day(2005,7)
        day.chart<-rbind(year00, year01, year02, year03, year04, year05)
    estimate<-0
        for(i in 1:lags){
                N<-claim.number(month, year, new.chart, sub.chart,
day.chart, smooth, i,wt.sp, wt.kids)
                P<-claim.amount(month, year, new.chart, i)
                print(paste("Number of claims:",as.character(N)))
                estimate<-estimate+P*N
                }
        error.rate<-calc.errate(new.chart, year, month, lags, estimate)
        return(Month=month, Year=year, Estimate.Amount=estimate,
Error=error.rate)
}
```