SOCIETY OF ACTUARIES

Article from:

Health Section News

April 2003 – Issue No. 45

A Practical Method for Incorpoarting Pended Claims in Medical IBNR Estimates

by Peter K. Reilly

Introduction

Medical incurred but not reported claim reserves (IBNR) are a principal driver of reported financial results of health insurers. These reserves, while not as long tailed as long term care or disability income insurance contracts, can be quite material and the mis-estimating of these reserves can add considerable variability to reported financial results. The vast majority of this misestimating risk comes from the most recent dates of service. Traditional completion or lag factor analysis is often relied upon by valuation actuaries to set IBNR reserves and the most recent months are where these methods are least useful.

In this paper, I present a practical framework for incorporating a full set of available information into the estimation of IBNR reserves that should reduce estimation error. The practical results would be reduced capital requirements supporting the health risk business, via reduced reserve margins, and a better understanding of emerging results allowing one to more quickly take the appropriate steps to manage the business and represent a financial statement that more accurately reflect true date of service results.

Background

IBNR reserves by definition depend on a company's accounting treatment of claims payments. Two common approaches are: 1) recording a claim as paid when the payment is issued and 2) when the draft clears the banking system.

Regardless of which definition is used, the valuation actuary needs to estimate what the company's obligations are for GAAP and statutory accounting purposes. For this paper, I will assume a check issued basis. Estimation methodologies typically rely on past patterns of claim payments and how they have developed. This is done by arranging all known paid claims by date of service and month of payment into a triangle format. This is typically referred to as a lag table. With sufficient history, stable submission and claims processing times and a stable trend environment, past payment patterns can be used to make accurate IBNR estimates.

A few issues arise. First, it should be obvious that because of the definition of paid claims adopted, there are known claims that have been received and pended but not yet adjudicated and potentially processed and held pending release in an account payable (AP) account. However, the claims that have been pended but not adjudicated are not directly translatable into a resultant payment. Some of these claims will be denied or paid at an amount less than submitted. Once adjudicated, some of the claims will have differing payment levels due to contractual terms. Because of these issues, pended claims are often not brought directly into the IBNR estimation process but are instead relied on for anecdotal information only. Depending on payment patterns and whether checks are held in pended status, nearly a month's worth of claims may be on hand but essentially ignored in setting IBNR.

Completion Factor Methods

Completion factor methods rely on the premise that past payment patterns will hold on average in the future. Since it is often the case that the more recent dates of service months may only be 5 percent – 30 percent complete, there is substantial leverage in the volatility of payment patterns into the IBNR reserve. Given this, reserve actuaries will typically choose more conservative estimation methods to ensure the adequacy of the IBNR reserve.

Lets define:

$P_{i,j}$ = Paid claim for DOS_i, paid in period $_j$

Where DOS = Date of Service

Incurred claims for DOS_i

$$IC_i = \sum_{j=0}^{\infty} P_{i,j}$$

Obviously, where $j < \infty$, there exists the potential that claims are still outstanding. When $_j$ gets close to 0 the amount of outstanding claims becomes material. Incurred but not reported for DOS_i, held as of time or duration incurred+$_k$

$$IBNR_{i,k} = \sum_{j=i+k}^{\infty} P_{i,j} \quad \text{where } P_{i,j} \text{ is unknown}$$

Starting with DOS months where it is reasonably certain that few, if any, claims remain to be paid, this formula can be used to work backwards to estimate the balance of the lag table.

$$CF_{i,k} = \frac{\sum_{j=0}^{i+k} P_{i,j}}{IC_i}$$

where k is the payment duration and $0 < CF_{i,j+k} < 1$ ignoring recoveries.

Completion factor methods typically take values over many dates of service for particular payment duration as an estimate or predictor of current payment patterns. A six-month average method might be (^ denoting estimate)

$$\text{e.g. } \hat{CF}_{i,0} = \sum_{i=-1}^{-6} CF_{i,0} / 6$$

This would be used to estimate IC i by

$$\hat{IC}_i = \frac{P_{i,0}}{\hat{CF}_{i,0}}$$

In plain English, this means if on average we believe the prior 6 DOS were 10 percent complete, for example, after one month of payment, we can gross up the one month of payment known for the current DOS, by dividing by .1, to predict the incurred claim. The IBNR reserve would then be

$$\hat{IC} - P_{i,0}$$

Besides ignoring the known information of the pended claims, this process also breaks down when claim payment pattern changes are occurring. For example, claims may be received and processed faster due to electronic claim submission and claim auto-adjudication. Averaging methods always assume the CF will be within a specified historical range. This can be or may be an inappropriate restriction.

A Practical Method for Incorporating Pended Claims Into the IBNR

It should be apparent that if claim submission patterns remain constant, a slow down or speed up of claims adjudication will result in an increase or decrease in claims held in pended status.

Regression methods can be used to expand the CF model to incorporate this data directly. In addition, this method replaces, for better or worse, moving averages as the predictor for the CF with an ordinary least squares estimator.

The proposed models can be stated as:

$$IBNR\ PMPM_{i,k} = a \cdot \sum_{j=0}^{k} CumPaid\ PMPM_{i,j} + b \cdot Pended\ dollars\ PMPM_{i,k} + E_i$$

Where

IBNR PMPM $_{i,k}$ = Restated incurred but not reported reserve for date of service $_i$ after $_k$ months of payment (duration k) divided by exposure (members) at time i.

CumPaid PMPM $_{i,j}$ = paid claims for date of service i paid in duration j divided by exposure (members) at time $_i$.

Pended dollars PMPM $_{i,k}$ = dollars pended in the system payable for date of service $_i$ at time (duration) k divided by exposure (members) at time $_i$.

Each of these variables is a vector of observations whose length will depend on available data. The minimum amount of data required is a function of degrees of freedom necessary to estimate the model parameters and the desired level of statistical precision of the estimated parameters. They should be balanced against the possibility that the parameters may change over time as changes to adjudication speeds occur. It might also be possible that the coefficient b in the pended claim portion could be affected by seasonality (particularly for deductible plans), and additional data would be required to incorporate this effect.

This model jointly estimates completion factors, one for cumulative paid claims and one for pend claims held for DOS i. If b = 0, then the model reduces to the CF model where

(continued on page 32)

$$a = \left(\frac{1}{CF_{i,k}} - 1\right)$$

$$\left(IC_i = \sum_{j=0}^{k} P_{i,j} + IBNR_i\right)$$

The parameters a and b can be estimated using ordinary least square (regression) or OLS methodologies. Note that the model form does not include a constant. All the usual considerations for using OLS with time series such as uncorrelated error terms, should be considered to ensure unbiased, efficient estimation of parameters.

Model Form Variations and Other Considerations

The model described above can be augmented or have model form variations that may improve the ability to fit the data and forecast more accurately.

Natural Logarithms: Experience has shown that the pended dollars PMPM will have considerable noise and scale issues relative to the dependent variable, estimated reserve PMPM. This occurs since pended claims are not adjudicated yet and may turn into paid claims at varying rates due to contractual considerations and denial rates.

Accounts Payable (AP) Pends: If an organization pends adjudicated claims for cash flow purposes, these can be handled in two separate ways. The dependent variable can be transformed by subtracting these claims prior to modeling and the current AP pends can be added back into the predicted reserves later to get the IBNR estimated. This treats AP pends as known claims which lead to the second potential treatment. The AP pends vector for a particular duration can be added as a third predictor variable into the model.

Working Days Variable: While incurred claims typically have seasonality in medical coverages, this seasonality is embedded in both the dependent variable and the predictor variables and therefore a separate variable is usually not necessary in the model. However, there is a separate, more subtle dynamic at work in the process.

The process of generating, submitting and adjudicating claims is a continuous process for the most part. The divvying up of the data into monthly time series is somewhat arbitrary. This decision however injects some variation into the dependent variable in that different months have different lengths. More specifically, they have different numbers of days (working days, mail days, processing days, etc.) where claims are typically generated and processed. This information is NOT embedded in the snapshots of pended claims, as these should be independent of the arbitrary month end cutoffs. Cumulative paid claims will have this embedded. For example, all things held constant more claims will get processed in a longer month than in a shorter month. That may lead to over estimation of reserves, particularly in the earlier durations. The addition of a working days variable which counts the effective numbers of processing days may help adjust this out.

Experience has shown that this is only important in the first few durations as the month-to-month variations in days average out as the exposure period lengthens.

An Example

This example is based on actual company data. For confidentiality purposes, the data has been transformed. The model relationships are invariant to the transformation.

In this example, I present a process where only the most recent durations are set using this modeling approach. Later durations are set first using traditional completion factor approaches. The two most recent durations are set iteratively. The IBNR estimate for DOS one month prior at duration is set first. This last data point of course sets the restated IBNR for duration 0. The IBNR for the current month DOS at duration 0 is the set using the model prediction.

Another important issue that was previously referred to is prominent in this example, the time varying parameter problem. In the model form presented earlier, the coefficients (effectively the completion factors) are assumed to be invariant with respect to time. In other words, processing pattern changes over time are averaged out. Traditional completion factor methods attempt to deal with this problem by shortening the averaging length used in selecting completion factors in hopes of limiting the prediction error.

There are two simple ways of dealing with this issue within the modeling framework presented in this paper. The first is to limit the data used to a time period that contains roughly stable processing patterns. The second approach is similar—maintain a longer time span of data, but segment the dependent variables into one or more sub segments that will have the parameters independently estimated. This allows for statistical tests on the hypothesis that the parameters have changed and for the ability to search for optimal points of segmentation.

In this example there are three predictor variables: cumulative paid claims, pended dollars and accounts payable pended dollars. All variables are stated on a per member per month basis. The predictor variables are not logarithmically transformed and I have broken the pended claims and cumulative paid claims variables into two pieces. The coefficients for the most recent time period are the ones used in predicting reserve levels.

The model statistics are indicated in the following tables and graphs.

Duration 1

SUMMARY OUTPUT

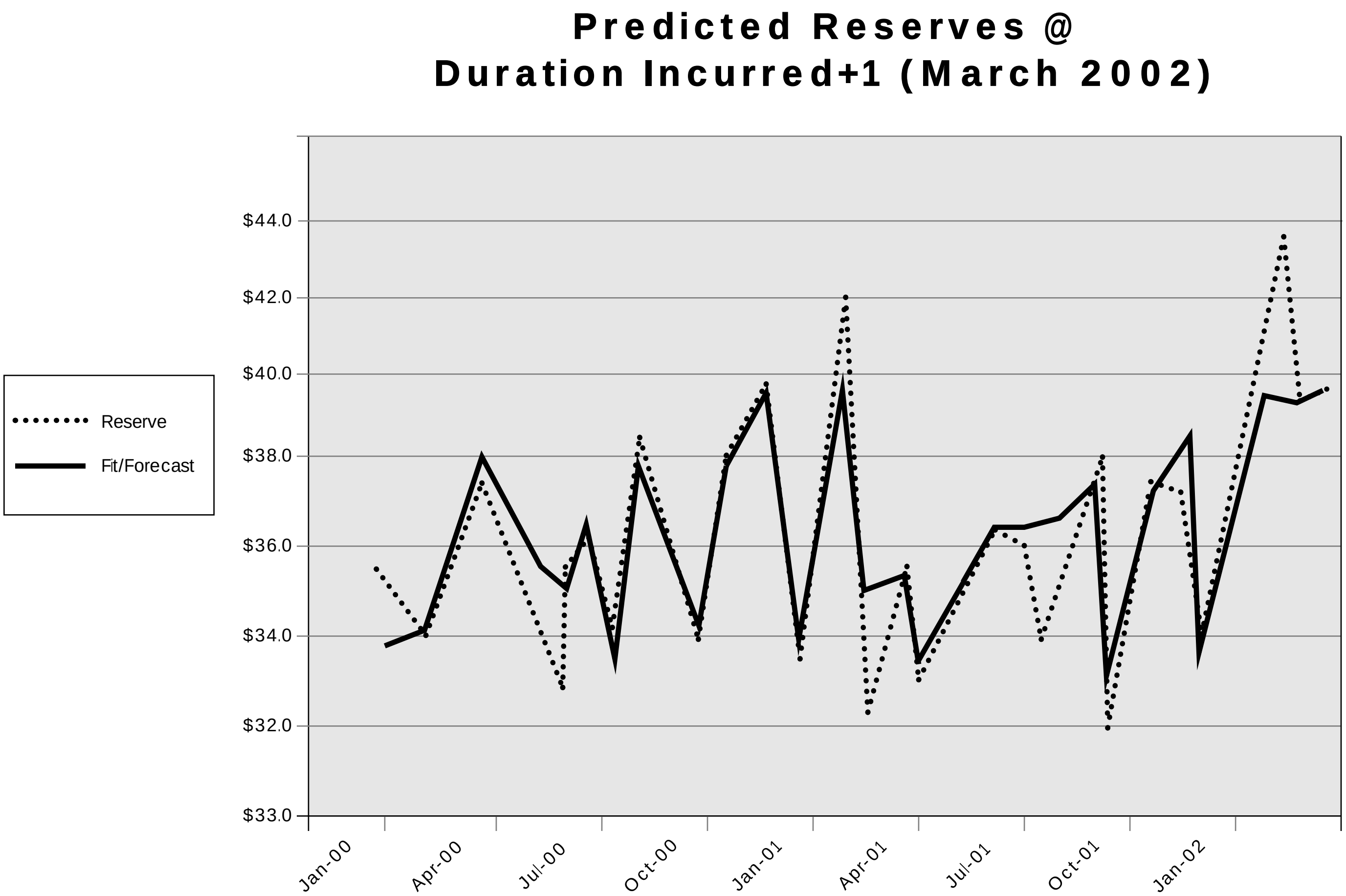
Regression Statistics	
Multiple R	0.854389376
R Square	0.729981205
Adjusted R Square	0.630930006
Standard Error	1.616321986
Observations	26

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	148.3176173	29.66352	11.35447	2.23E-05
Residual	21	54.862432	2.612497	-	-
Total	26	203.1800493	-	-	-

	Coefficients	*Standard Error*	*t Stat*	*P-Value*	*Lower 95%*	*Upper 95%*
Intercept	0	#NA	#nA	#NA	#NA	#NA
Pended Claims Pre 2001	12.73431218	1.913895284	6.65361	1.38E-06	8.754148	16.71448
Pended Claims Post 2000	12.84985547	2.194842844	5.854567	8.23E-06	8.285429	17.41428
Cum Paid Claims Pre Feb-01	0.055433252	0.101366983	0.546857	0.590241	-0.155371	0.266237
Cum Paid Claims Post Jan-02	0.013223526	0.083951033	0.157515	0.876344	-0.161362	0.187809
AP Pends	-0.478832562	0.683930742	-0.700118	0.491537	-1.901145	0.94348

(continued on page 34)



Duration 2

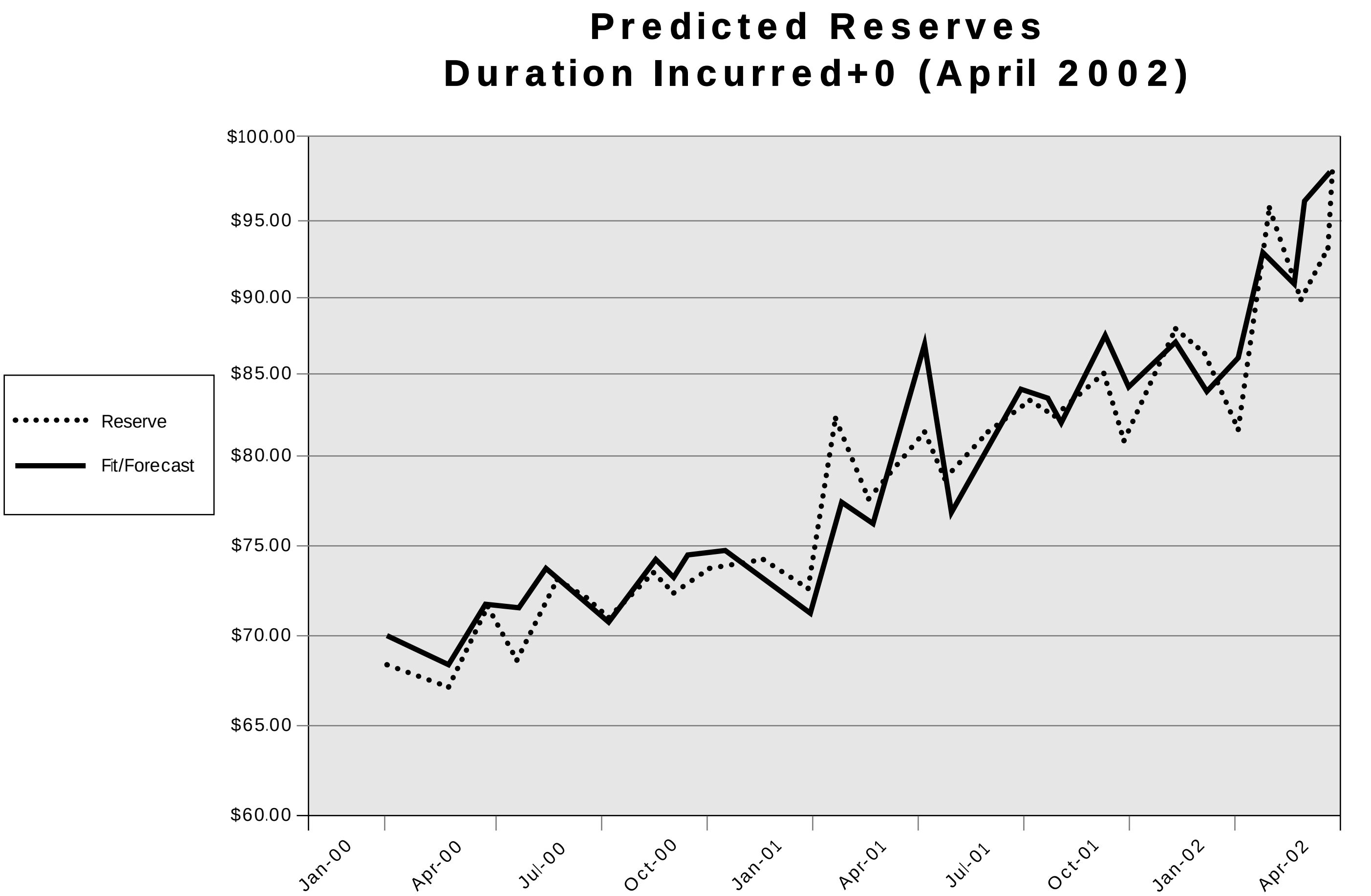
SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.956480244
R Square	0.914854457
Adjusted R Square	0.853918904
Standard Error	2.411341347
Observations	27

ANOVA

	df	SS	MS	F	Significance F
Regression	5	1374.45381	274.8908	47.27622	1.01E-10
Residual	22	127.920476	5.814567	-	-
Total	27	1502.374307	-	-	-

	Coefficients	Standard Error	t Stat	P-Value	Lower 95%	Upper 95%
Intercept	0	#NA	#nA	#NA	#NA	#NA
Pended Claims Pre 2001	20.30555462	0.926769978	21.910003	1.96E-14	18.38355	22.22756
Pended Claims Post 2000	20.67095749	1.147051825	18.02094	1.16E-14	18.29212	23.0498
Cum Paid Claims Pre Feb-01	0.536370206	0.205172893	2.614235	0.0153838	0.110867	0.961873
Cum Paid Claims Post Jan-02	0.511675694	0.15794425	3.239597	0.003764	0.184119	0.839232
AP Pends	0.705568172	0.358653302	1.967271	0.061888	-0.038234	1.44937



The summary statistics indicate fairly high R2's, indicating that a large portion of the historical variance is explained in the model. There is little evidence that the relationship between reserves and pended claims has changed over time. Interestingly, in the duration incurred+1 model cumulative paid claims is not statistically significant. The bulk of the reserve prediction is coming from pended dollars PMPM. AP pends are not statistically significant at duration incurred+1 but are significant at duration incurred+0.

Conclusions

The incorporation of additional information not traditionally incorporated formally in the reserve process has the potential to reduce errors in setting IBNR reserve. Additionally, a statistical approach can facilitate setting confidence limits around reserve estimates and the assessment of the probability of adequate recorded reserves.

The downside to this approach is the requirements of familiarity and skill with certain statistical techniques, potential difficulty in communicating the process to non-technical audiences, potential distrust of the process until its efficacy can be demonstrated and the difficulty in identifying and dealing with time varying parameters.

Peter K. Reilly, FSA, MAAA, is an actuary at Aetna U.S. Healthcare in Blue Bell, PA. He can be reached at ReillyPK@aetna.com.