STANDARD ERRORS OF PREDICTION IN CLAIMS RESERVING: A COMPARISON OF METHODS WORKSHOP

PETER ENGLAND
RICHARD VERRALL

1998 GENERAL INSURANCE CONVENTION
AND
ASTIN COLLOQUIUM

GLASGOW, SCOTLAND: 7-10 OCTOBER 1998

Standard Errors of Prediction in Claims Reserving: A Comparison of Methods

Peter D England
Richard J Verrall

Note: This paper was first presented at Insurance: Mathematics and Economics 1998, Lausanne, Switzerland, and has been submitted to the journal "Insurance: Mathematics and Economics" under the title "Analytic and bootstrap estimates of prediction errors in claims reserving".

Analytic and bootstrap estimates of prediction errors in claims reserving

Peter England[1] and Richard Verrall[2]

[1] CGU plc, St Helen's, 1 Undershaft, London, EC3P 3DQ
[2] Department of Actuarial Science and Statistics, City University, Northampton Square, London, EC1V 0HB

Abstract

Renshaw and Verrall (1994) have suggested a stochastic claims reserving model which reproduces the reserve estimates provided by the standard chain ladder model (subject to constraints on the pattern of negative incremental claims). Their model falls within the generalised linear modelling (GLM) framework and can be fitted easily using standard statistical software packages. Having fitted the model, it is possible to obtain analytic prediction errors of the reserves. The GLM framework suggests appropriate goodness-of-fit measures, and also gives a choice of appropriate residual definitions which can be used for informal diagnostic checks of the fitted model. Residuals can also be used in a bootstrap exercise providing a computationally simple method of obtaining estimates of the reserve prediction errors. In this paper, we consider an appropriate residual definition for this purpose, and show how the bootstrap prediction errors can be computed easily in a spreadsheet, without the need for statistical software packages. The bootstrap prediction errors are compared with their analytic equivalent, and also compared with other methods commonly used, including Mack's distribution free approach (Mack 1993) and methods based on log-linear models.

1. Introduction

In recent years, considerable attention has been given to the relationship between various stochastic models and the chain ladder technique. Stochastic models have been constructed with the aim of producing *exactly* the same reserve estimates as the traditional deterministic chain ladder model. At first sight, this might seem like a futile exercise: why use a complex stochastic method to find reserve estimates when a simple deterministic method will suffice? The answer is that as well as the reserve estimates, there are other aspects of the model which are of importance, such as the underlying distributional assumptions of the model being fitted, estimates of the likely variability in the parameter estimates, and an estimate of the goodness-of-fit of the model. It is also useful to know where the data deviate from the fitted model, and to have a sound framework within which other models can be fitted and compared.

To date, two models have been suggested as stochastic chain ladder models, Mack's distribution free approach (Mack, 1994), and Renshaw and Verrall's approach using generalised linear models (Renshaw and Verrall, 1994). Both models provide reserve estimates which are identical to those provided by the deterministic chain ladder model (under suitable constraints explained in section 2), and allow estimates of reserve variability to be calculated. Other models have been proposed which provide reserve estimates which are usually close to those from the chain ladder model, but not identical.

A key advantage of Renshaw and Verrall's approach is that it is embedded within the generalised linear modelling framework, widely used in statistical modelling. Theory associated with generalised linear models can be used to suggest how parameter estimates can be obtained, and also to suggest appropriate goodness-of-fit measures and residual definitions. The theory can also help in deriving analytic standard errors of prediction (prediction errors) of reserve estimates.

Residuals can also be used in a bootstrap exercise to provide bootstrap standard errors. It is important when bootstrapping to use a residual definition which is appropriate to the model under consideration. Renshaw and Verrall's stochastic chain ladder model suggests a residual definition which is appropriate for bootstrapping chain ladder reserve estimates. This residual definition deviates from the definition used in previous papers on bootstrapping reserve estimates (e.g. Brickman *et al* 1993, Lowe 1994), and overcomes some of the difficulties previously identified.

Analytic prediction errors involve complex formulae which are difficult to evaluate. On the other hand, bootstrap prediction errors are remarkably easy to calculate, and can be computed using a spreadsheet, without recourse to specialised statistical modelling packages.

In the following section, we provide a brief overview of a range of stochastic "chain ladder" type reserving models, not just those mentioned above. Sections 3 and 4 introduce analytic prediction errors and bootstrap prediction errors. An example in which results from the various models are compared is contained in Section 5. An outline of the calculations required for the bootstrap prediction errors appears in the Appendix.

2. Stochastic "Chain Ladder" Type Reserving Models

A significant step in the search for a stochastic chain ladder model was made by Kremer (1982) who focused on the logarithm of incremental claims amounts as the response and regressed on two non-interactive covariates.

Let C_{ij} denote the incremental claims amount arising from accident year i paid in development year j. Let $Y_{ij} = \log(C_{ij})$ and consider the log-normal class of models $Y_{ij} = m_{ij} + \varepsilon_{ij}$ with

$$Y_{ij} \sim IN(m_{ij}, \sigma^2) \tag{2.1}$$

$$\varepsilon_{ij} \sim IN(0, \sigma^2) \tag{2.2}$$

$$m_{ij} = \eta_{ij} \tag{2.3}$$

$$\eta_{ij} = c + \alpha_i + \beta_j \qquad \alpha_1 = \beta_1 = 0 \tag{2.4}$$

The normal responses Y_{ij} are assumed to decompose (additively) into a deterministic non-random component with mean $m_{ij} = \eta_{ij}$ and a homoscedastic normally distributed random error component about a zero mean. The use of the logarithmic transform immediately

imposes a limitation on this class of models in that incremental claim amounts must be positive.

Equations 2.1, 2.2, 2.3 and 2.4 define the model introduced by Kremer. Accident year and development year are treated as factors, with a parameter α_i for each accident year i and a parameter β_j for each development year j. It should be noted that this representation implies the same development pattern for all accident years, where that pattern is defined by the parameters β_j.

Parameters in the predictor structure η_{ij} are estimated by maximum likelihood, which in the case of the Normal error structure is equivalent to minimizing the residual sum of squares. The unknown variance σ^2 is estimated by the residual sum of squares divided by the degrees of freedom (the number of observations minus the number of parameters estimated).

Given the parameter estimates, the predicted values on a log scale can be obtained by introducing those estimates back into equation 2.4. Unfortunately, exponentiating to give predicted values on the untransformed scale introduces a bias, which must be corrected. Specific details can be found in Renshaw (1989) and Verrall (1991a). This model usually produces predicted values which are close to those from the simple chain ladder model.

Standard results from statistical theory allow prediction errors to be calculated for reserve estimates, and also allow diagnostic checks of the fitted model to be performed by analysing appropriate residuals.

It should be noted that the model can be extended by considering alternatives to the linear predictor specified in equation 2.4. This log-normal "chain-ladder" model and further generalisations have been considered by Zehnwirth (1989, 1991), Renshaw (1989), Christofides (1990) and Verrall (1991a, 1991b), amongst others.

In 1994, two papers were published, both of which derived stochastic models giving the same reserve estimates as the deterministic chain ladder model. Mack (1994) presented a distribution free approach, whereas Renshaw and Verrall (1994) presented a model in which the distributional properties underlying the model were fully specified. In an earlier paper, Mack (1993) derived reserve standard errors for his distribution free approach. The approach of Renshaw and Verrall (1994) is considered in detail because of the relevance when introducing the bootstrap. In the example in Section 5, the results from Mack (1994) are compared with those obtained by Renshaw and Verrall (1994), and by using a bootstrap approach.

Renshaw and Verrall (1994) proposed modelling the incremental claims C_{ij} directly as the response, with the same linear predictor as Kremer, but linking the mean to the linear predictor through the logarithmic link function, while using an "over-dispersed" Poisson error distribution. Formally,

$$E[C_{ij}] = m_{ij} \quad \text{and} \quad Var[C_{ij}] = \phi E[C_{ij}] = \phi m_{ij} \tag{2.5}$$

$$\log(m_{ij}) = \eta_{ij} \tag{2.6}$$

$$\eta_{ij} = c + \alpha_i + \beta_j \qquad \alpha_1 = \beta_1 = 0 \tag{2.7}$$

Equations 2.5, 2.6 and 2.7 define a generalised linear model in which the response is modelled with a logarithmic link function and the variance is proportional to the mean (hence "over-dispersed" Poisson). The parameter ϕ is an unknown scale parameter estimated as part of the fitting procedure.

Since this model is a generalised linear model, standard statistical software can be used to obtain maximum (quasi) likelihood parameter estimates, fitted and predicted values. Standard statistical theory also suggests goodness-of-fit measures and appropriate residual definitions for diagnostic checks of the fitted model.

Renshaw and Verrall were not the first to notice the link between the chain ladder model and the Poisson distribution (see Appendix A of Mack (1991)), but were the first to implement the model using standard methodolgy in statistical modelling, and to provide a link with the analysis of contingency tables.

It should be noted that the model proposed by Renshaw and Verrall is robust to a small number of negative incremental claims, since the responses are the incremental claims themselves (rather than the logarithm of the incremental claims as in log-normal models). However, because of the way in which the model structure is parameterised and the estimates obtained, it is necessary to impose the restriction that the sum of incremental claims in every row and every column of the data triangle must be positive. Furthermore, because of the logarithmic link function, fitted values are always positive. This usually makes the model unsuitable for use with incurred claims, which often include overestimates of case reserves in the early stages of development leading to a series of negative incremental incurred claims in the later stages of development.

Mack (1991) suggested a further model which is relevant to this paper, although it is not described as a chain ladder model. Mack proposed a multiplicative parametric structure for the mean incremental claims amounts which are modelled as Gamma response variables, and used a rather complex fitting procedure for obtaining maximum likelihood parameter estimates. As Renshaw and Verrall (1994) note, exactly the same model can be fitted using a generalised linear model in which the incremental claim amounts are modelled as independent Gamma response variables, with a logarithmic link function and the same linear predictor as Kremer (1982). Formally,

$$E\left[C_{ij}\right] = m_{ij} \quad \text{and} \quad Var\left[C_{ij}\right] = \phi E\left[C_{ij}\right]^2 = \phi m_{ij}^2 \tag{2.8}$$

$$\log(m_{ij}) = \eta_{ij} \tag{2.9}$$

$$\eta_{ij} = c + \alpha_i + \beta_j \qquad \alpha_1 = \beta_1 = 0 \tag{2.10}$$

The only difference between this model and the stochastic chain ladder model proposed by Renshaw and Verrall (1994) is that the variance is now proportional to the mean squared. The model defined by equations 2.8, 2.9 and 2.10 can be fitted using standard statistical

software capable of fitting GLMs. Like the log-normal models, fitted values from this model are usually close to those from the standard chain ladder model, but not exactly the same.

3. Analytic Estimates of Reserve Prediction Errors

One of the principle advantages of stochastic reserving models is the availability of estimates of reserve variability. Commonly used in prediction problems (as we have here) is the standard error of prediction, also known as the prediction error, or root mean square error of prediction. Consider accident year i and claim payments in development year j (yet to be observed). The mean square error of prediction is given by

$$E\left[(C_{ij} - \hat{C}_{ij})^2\right] \cong Var\left[C_{ij}\right] + Var\left[\hat{C}_{ij}\right] \tag{3.1}$$

For a detailed justification of equation 3.1, see Renshaw (1994). Equation 3.1 is valid for the log-normal reserving models, the over-dispersed Poisson model, and the Gamma model. Note that the mean square error of prediction can be considered as the sum of two components, variability in the data (process variance) and variability due to estimation (estimation variance). The precise form of the two components of the variance is dictated by the specification of the model fitted. For the log-normal model defined by equations 2.1, 2.2, 2.3 and 2.4, the precise form of the two components of variance can be found in Renshaw (1989) or Verrall (1991a).

A general form for the process variance can be derived for the over-dispersed Poisson and Gamma models. From equations 2.5 and 2.8, it can be seen that

$$Var\left[C_{ij}\right] = \phi m_{ij}^{\rho} \tag{3.2}$$

where $\rho = 1$ for the over-dispersed Poisson model and $\rho = 2$ for the Gamma model.

For the estimation variance, we note that for the over-dispersed Poisson and Gamma models

$$\hat{C}_{ij} = m_{ij} = \exp(\eta_{ij})$$

Then using the *delta method*,

$$Var\left[\hat{C}_{ij}\right] \cong \left|\frac{\partial m_{ij}}{\partial \eta_{ij}}\right|^2 Var\left[\eta_{ij}\right]$$

giving

$$E\left[(C_{ij} - \hat{C}_{ij})^2\right] \cong \phi m_{ij}^{\rho} + m_{ij}^2 Var\left[\eta_{ij}\right] \tag{3.3}$$

The final component of equation 3.3, the variance of the linear predictor, is usually available directly from statistical software packages, enabling the mean square error to be calculated without difficulty. The standard error of prediction is the square root of the mean square error.

The standard error of prediction for origin year reserve estimates and the total reserve estimates can also be calculated. Denoting the triangle of predicted claims contributing to the reserve estimates by Δ, then the reserve estimate in origin year i is given by summing the predicted values in row i of Δ, that is

$$C_{i+} = \sum_{j \in \Delta_i} C_{ij}$$

From Renshaw (1994), the mean square error of prediction of the origin year reserve is given by

$$E\left[(C_{i+} - \hat{C}_{i+})^2\right] \cong \sum_{j \in \Delta_i} \phi m_{ij}^{\rho} + \sum_{j \in \Delta_i} m_{ij}^2 Var\left[\eta_{ij}\right] + 2 \sum_{\substack{j_1, j_2 \in \Delta_i \\ j_2 > j_1}} m_{ij_1} m_{ij_2} Cov\left[\eta_{ij_1} \eta_{ij_2}\right] \qquad (3.4)$$

The total reserve estimate is given by

$$C_{++} = \sum_{i,j \in \Delta} C_{ij}$$

and the mean square error of prediction of the total reserve is given by

$$E\left[(C_{++} - \hat{C}_{++})^2\right] \cong \sum_{i,j \in \Delta} \phi m_{ij}^{\rho} + \sum_{i,j \in \Delta} m_{ij}^2 Var\left[\eta_{ij}\right] + 2 \sum_{\substack{i_1 j_1 \in \Delta \\ i_2 j_2 \notin \Delta \\ i_1 j_1 \neq i_2 j_2}} m_{i_1 j_1} m_{i_2 j_2} Cov\left[\eta_{i_1 j_1} \eta_{i_2 j_2}\right] \qquad (3.5)$$

Equations 3.4 and 3.5 require considerable care when summing the appropriate elements. The covariance terms are not readily available from statistical software packages. However, provided the *design matrix* and *variance-covariance matrix* of the parameter estimates can be extracted from the statistical software package used, a full matrix of the covariance terms can be calculated. Indeed, the variances of the linear predictors are simply the diagonal of such a matrix.

Note that the first term in the accident year and overall prediction errors is simply the appropriate sum of the process variances. The remaining terms relate to the estimation variance.

4. Bootstrap Estimates of Reserve Prediction Errors

Where a standard error is difficult or impossible to estimate analytically, it is common to adopt the bootstrap. In claims reserving, we are interested in the prediction error of the sum of random variables, and the bootstrap technique is a natural candidate for this. In regression type problems, it is common to bootstrap the residuals, rather than bootstrap the data themselves (see Efron and Tibshirani 1993). However, it is important to use an appropriate residual definition for the problem at hand. For linear regression models with Normal errors, the residuals are simply the observed values less the fitted values. For generalised linear models, an extended definition of residuals is required which have (approximately) the usual properties of Normal theory residuals (see McCullagh and Nelder 1989). The most commonly used residuals in generalised linear models are the Deviance residuals and the

Pearson residuals. A third residual, less commonly used, is the Anscombe residual. The precise form of the residual definitions is dictated by the error distribution. For the stochastic chain ladder model defined by equations 2.5, 2.6 and 2.7, we use the form of residuals suitable for Poisson GLMs, which are:

Unscaled Deviance residual: $$r_D = sign(C-m)\sqrt{2(C\log(C/m)-C+m)} \quad (4.1)$$

Unscaled Pearson residual: $$r_P = \frac{C-m}{\sqrt{m}} \quad (4.2)$$

Unscaled Anscombe Residuals: $$r_A = \frac{\frac{3}{2}(C^{2/3}-m^{2/3})}{m^{1/6}}$$

The bootstrap process involves resampling, with replacement, from the residuals. A bootstrap data sample is then created by inverting the formula for the residuals using the resampled residuals, together with the fitted values. Given r and m, it can be seen that equation 4.1 cannot be solved analytically for the observed incremental claims, C, making deviance residuals less suitable for bootstrapping. However, it is easy to solve equation 4.2 for C. Given a resampled Pearson residual r_P^* together with the fitted value m, the associated bootstrap incremental claims amount, C^*, is given by

$$C^* = r_P^*\sqrt{m} + m \quad (4.3)$$

It is also possible to solve the Anscombe residuals for C, but they are not considered here any further because they are less commonly used and because it is desirable to use a residual definition when bootstrapping which is consistent with the estimation of the scale parameter (see below).

Having obtained the bootstrap sample, the model is refitted and the statistic of interest calculated. The process is repeated a large number of times, each time providing a new bootstrap sample and statistic of interest. The bootstrap standard error is the standard deviation of the bootstrap statistics.

In the context of stochastic claims reserving, resampling the residuals (with replacement) gives rise to a new triangle of claims payments. Strictly, we ought to fit the over-dispersed Poisson GLM to the bootstrap sample to obtain the bootstrap reserve estimates. However, we can obtain identical reserve estimates using standard chain ladder methodology. It is at this point that the usefulness of the bootstrap process becomes apparent: we do not need sophisticated software to fit the model, a spreadsheet will suffice. To obtain the bootstrap standard errors of the reserve estimates, it is necessary to repeat the process a large number of times (say, N), each time creating a new bootstrap sample, and obtaining chain ladder reserve estimates. The bootstrap standard errors are the standard deviations of the N bootstrap reserve estimates. Once set up, the process is very quick, taking only a few seconds on a standard desktop computer.

The bootstrap standard error is an estimate of the square root of the *estimation* variance. However, it cannot be compared directly with the analytic equivalent since the bootstrap

standard error does not take account of the number of parameters used in fitting the model: the bootstrap process simply uses the residuals with no regard as to how they are obtained. The analytic estimates of the estimation variance do allow for the number of parameters estimated since they involve variance and covariance terms which implicitly involve the scale parameter ϕ in their calculation. The scale parameter is estimated as either the model deviance divided by the degrees of freedom, or the Pearson chi-squared statistic divided by the degrees of freedom, the choice usually making little difference. The deviance and Pearson chi-squared statistics are obtained as the sum of the squares of the corresponding residuals. The degrees of freedom are defined as the number of data points (in the original data sample) less the number of parameters used in fitting the model. Therefore, the Deviance scale parameter is given by

$$\phi_D = \frac{\sum r_D^2}{n-p}$$

and the Pearson scale parameter is given by

$$\phi_P = \frac{\sum r_P^2}{n-p} \tag{4.4}$$

where n is the number of data points in the sample, p is the number of parameters estimated and the summation is over the number (n) of residuals. It can be seen that an increased number of parameters used in fitting the model introduces a penalty (*ceteris paribus*).

For consistency, we use the Pearson scale parameter in the analytic estimation variance, and the Pearson residuals in the bootstrap process. The bootstrap estimation variance is analogous to the analytic estimation variance without adjusting for the number of parameters (as though the scale parameter had been calculated by dividing by n not $n-p$). To enable a proper comparison between the estimation variances given by the two procedures, it is necessary to make an adjustment to the bootstrap estimation variance to take account of the number of parameters used in fitting the model. The appropriate adjustment is to multiply the bootstrap estimation variance by $n/(n-p)$.

To obtain the bootstrap prediction error, it is necessary to add an estimate of the process variance, which is simply the scale parameter multiplied by the reserve estimates (see equations 3.4 and 3.5 when $\rho = 1$). The reserve estimates are given by the initial chain ladder projection, and the scale parameter is calculated by summing the squares of the residuals used in the bootstrap exercise. The process variance can also be computed in a spreadsheet. The bootstrap prediction error is then given by

$$PE_{bs}(R) = \sqrt{\phi_P R + \frac{n}{n-p}\left(SE_{bs}(R)\right)^2}$$

where R is an accident year or total reserve, and $SE_{bs}(R)$ is the bootstrap standard error of the reserve estimate.

It should be noted that no allowance has been made for a tail factor in the bootstrap calculations. It is not obvious how uncertainty in predicted values beyond the range of data observed should be taken into account. A fixed tail factor should not be included as this will increase the reserve estimates but leave the estimation variance unchanged, thus reducing the prediction error as a percentage of the reserve estimate. Extrapolating can only increase the uncertainty, not reduce it.

An example showing the computations required by the bootstrap can be found in the Appendix.

5. Example

To enable a comparison with previously published methods, we use the data from Taylor and Ashe (1983) which was also used by Verrall (1991), Mack (1993) and Renshaw (1989, 1994). The data are shown here in incremental form.

357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
352118	884021	933894	1183289	445745	320996	527804	266172	425046	
290507	1001799	926219	1016654	750816	146923	495992	280405		
310608	1108250	776189	1562400	272482	352053	206286			
443160	693190	991983	769488	504851	470639				
396132	937085	847498	805037	705960					
440832	847631	1131398	1063269						
359480	1061648	1443370							
376686	986608								
344014									

Reserve estimates provided by the deterministic chain ladder, the over-dispersed Poisson model, the Gamma models using the GLM implementation outlined in this paper and the Mack 1991 implementation, and three methods using log-normal models are shown in Table 5.1. Equivalent prediction errors are shown in Table 5.2, with the inclusion of the bootstrap approach and Mack's distribution free approach.

Table 5.1 : Estimated Reserves (000's)

	Chain Ladder	Poisson GLM	Gamma GLM	Mack (1991)	Verrall (1991)	Renshaw/ Christofides	Zehnwirth
i=2	95	95	93	93	96	111	109
i=3	470	470	447	447	439	482	473
i=4	710	710	611	611	608	661	648
i=5	985	985	992	992	1011	1091	1069
i=6	1419	1419	1453	1453	1423	1531	1500
i=7	2178	2178	2186	2186	2150	2311	2265
i=8	3920	3920	3665	3665	3529	3807	3731
i=9	4279	4279	4122	4122	4056	4452	4364
i=10	4626	4626	4516	4516	4340	5066	4965
Total	18681	18681	18085	18085	17652	19512	19124

Table 5.2 : Prediction Errors as % of Reserve Estimate

	Mack's Distn. Free	Poisson GLM Analytic	Bootstrap chain ladder	Gamma GLM	Mack (1991)	Verrall (1991)	Renshaw/ Christofides	Zehnwirth
i=2	80	116	117	48	40 (49)	49	54	49
i=3	26	46	46	36	30 (37)	37	39	35
i=4	19	37	36	29	24 (30)	30	32	29
i=5	27	31	31	26	21 (26)	27	28	25
i=6	29	26	26	24	20 (25)	25	26	24
i=7	26	23	23	24	20 (25)	25	26	24
i=8	22	20	20	26	21 (26)	27	28	26
i=9	23	24	24	29	24 (30)	30	31	30
i=10	29	43	43	37	31 (38)	38	40	39
Total	13	16	16	15	-	15	16	16

The results for Mack (1991), Verrall (1991), Renshaw/Christofides and Zehnwirth have been taken from Mack (1993). The three log-normal models (Verrall 1991, Renshaw/Christofides and Zehnwirth) are all using essentially the same model structure, as defined by equations 2.1, 2.2, 2.3 and 2.4. The differences in the reserve estimates and the prediction errors for the log-normal models are due to alternative methods for implementing the necessary bias correction or in the calculation of σ^2. The prediction error using Mack's distribution free approach has been taken from Mack (1993). Renshaw (1994) used the same data to compare results from the log-normal, Poisson and Gamma chain ladder type models (using a deviance scale parameter), but did not compare his results with Mack's distribution free and Gamma models, and did not consider the bootstrap.

It can be seen that the stochastic chain ladder model of Renshaw and Verrall (the overdispersed Poisson GLM) gives exactly the same reserve estimates as the deterministic chain ladder model (and hence Mack's distribution free stochastic model). The Gamma model implemented as a generalised linear model gives exactly the same reserve estimates as the Gamma model implemented by Mack (1991), which is comforting rather than surprising. It can be seen that these reserve estimates of the Gamma models are close to the chain ladder estimates. The loglinear model implemented by Verrall (1991) gives reserve estimates which are close to those given by the Gamma models, and again they are close to those given by the chain ladder model on the whole. The reserve estimates given by Renshaw/Christofides and Zehnwirth are very close to each other, the difference being due to the calculation of σ^2.

The prediction errors as a percentage of the equivalent reserve estimates of the three log-normal models are very close to each other in total and across accident years. For the Gamma models, at first sight it appears that the prediction errors are quite different. However, Mack (1991) did not make an adjustment for the degrees of freedom used in fitting his model, the appropriate adjustment being division by $n-p$ instead of n when calculating the scale parameter ϕ, where n is the number of data points (55) and p is the number of parameters estimated (19). The adjustment affects both the estimation variance and process

variance. To enable a proper comparison, it is necessary to adjust Mack's prediction errors by a factor f,

where $$f = \sqrt{\frac{n}{n-p}} = \sqrt{\frac{55}{36}} = 1.236$$

The numbers in parentheses show Mack's prediction errors including this adjustment. It can be seen that these are now very close to those given by the Gamma GLM, the differences being due to rounding errors. Mack did not provide a prediction error for the overall reserve. It is perhaps surprising that the prediction errors given by the Gamma models are very close to those given by the log-normal models, particularly the models of Verrall (1991) and Zehnwirth. It should be noted, however, that for both the log-normal and gamma distributions, the variance is proportional to the mean squared.

Although the prediction error for the total reserve given by the Poisson GLM is almost identical to that given by the Gamma and log-normal models, there are some large differences when looking across accident years. The biggest difference is clearly when i=2, where the Poisson model gives a large prediction error of 116%. It should be noted, however, that the denominator (the reserve estimate) is very low, and a large prediction error is not unexpected.

The bootstrap prediction errors (based on 1000 simulations) are extremely close to the analytic prediction errors of the Poisson model, both in total and across accident years. This is remarkable given the radically different methods used in obtaining the estimation variance.

Like the Poisson GLM and bootstrap approaches, Mack's distribution free approach gives a high prediction error when i=2. Prediction errors using Mack's distribution free approach for the other accident years are systematically neither higher nor lower than those given by the other methods. The prediction error for the total reserve, at 13%, is slightly lower than the equivalent figures from the other methods. Again, there is no adjustment for the number of parameters used in fitting the model. It is interesting to note that using the same adjustment factor, f, as for Mack's Gamma model gives 16% for the prediction error of the total reserve, bringing it into line with the other models. Unlike the Gamma model, however, it is not clear that such an adjustment is justified.

6. Conclusions

With the exception of Mack's distribution free approach, all of the stochastic claims reserving models shown in this paper use exactly the same linear predictor structure, that is, the structure introduced by Kremer. The models differ in the error distribution assumed, the choice being between the log-normal, the (over-dispersed) Poisson and the Gamma distributions. The Poisson model is interesting since the reserve estimates given by the model are identical to those given by the standard deterministic chain ladder model (under suitable constraints), and as such, it can be called a stochastic chain ladder model. Mack's distribution free approach is included because it also provides reserve estimates which are identical to those given by the deterministic chain ladder model.

Perhaps more interesting than the reserve estimates themselves are the prediction errors given by the various models. For models in which the distributional assumptions have been specified, it is possible to use an analytic or bootstrap approach. The bootstrap approach has been outlined for the Poisson model only, since it is easy to implement in a spreadsheet environment. Since residuals can be defined for the log-normal and Gamma models, it is also possible to obtain bootstrap prediction errors for these models, but model fitting is more complex.

It has been shown that when comparing prediction errors given by different methods, it is important to ensure that both the estimation variance and process variance have been included, and that they have been calculated in a consistent manner, including adjustment for the number of parameters used in fitting the model.

A comparison of the prediction errors reveals that the Gamma and log-normal models provide very similar results when viewing the prediction errors as a percentage of reserve estimates. The bootstrap prediction errors are remarkably similar to their analytic equivalent, justifying their use with the standard chain ladder model when applied correctly. The bootstrap procedure is practically expedient and does not require the summation of a large collection of terms, unlike the analytic and distribution free approaches.

It is interesting to note that the prediction errors of the reserve totals given by the various methods are reassuringly close in the example in section 5. Although this is often the case, unfortunately it is not always, and care must be taken in making inferences from the results. Further work is needed to justify the use of a particular error distribution in stochastic claims reserving models. In particular, the accuracy and interpretation of accident year prediction errors needs careful consideration. Clearly, it is not appropriate to consider approximate 95% prediction intervals as the reserve estimate $\pm$ twice the prediction error when the prediction error is a large percentage of the reserve estimate. It is best to use the accident year prediction errors as a crude means of assessing confidence in the reserve estimates.

Although we have used the Pearson residuals in our treatment of the bootstrap, Moulton and Zeger (1991) discuss an adjusted Pearson residual which may perform better. The adjustment is difficult to accommodate in a spreadsheet environment, and consequently has been ignored since any outperformance is outweighed by difficulty of implementation.

Acknowledgements

The authors are grateful for the support, encouragement and enthusiasm shown by Stavros Christofides while carrying out this work, and also for the useful discussions with Dr Arthur Renshaw.

References

Brickman, S, Barlow, C, Boulter, A, English, A, Furber, L, Ibeson, D, Lowe, L, Pater, R, and Tomlinson, D (1993) *Variance in claims reserving*. Proceedings of the 1993 General Insurance Convention, Institute of Actuaries and Faculty of Actuaries.

Christofides, S (1990) *Regression models based on log-incremental payments*. Claims Reserving Manual, Vol 2, Institute of Actuaries, London.

Efron, B and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap*. Chapman and Hall.

Kremer, E (1982) *IBNR claims and the two way model of ANOVA*. Scandinavian Actuarial Journal, pp47-55.

Lowe, J (1994) *A practical guide to measuring reserve variability using: Bootstrapping, Operational Time and a distribution free approach*. Proceedings of the 1994 General Insurance Convention, Institute of Actuaries and Faculty of Actuaries.

Mack , T (1991) *A simple parametric model for rating automobile insurance or estimating IBNR claims reserves*. ASTIN Bulletin 22, 1, pp93-109.

Mack, T (1993) *Distribution free calculation of the standard error of chain ladder reserve estimates*. ASTIN Bulletin 23, 2, pp213-225.

Mack, T (1994) *Which stochastic model is underlying the chain ladder model?* Insurance: Mathematics and Economics 15, pp133-138.

McCullagh, P. and Nelder, J. (1989) *Generalised Linear Models, 2nd Edition*. Chapman and Hall.

Moulton, L.H. and Zeger, S.L. (1991) *Bootstrapping generalized linear models*. Computational Statistics and Data Analysis 11, pp53-63.

Renshaw, A.E. (1989) *Chain Ladder and Interactive Modelling (Claims Reserving and GLIM)*. Journal of the Institute of Actuaries 116, III, pp559-587.

Renshaw, A.E. (1994) *On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models*. Actuarial Research Paper No. 65, Department of Actuarial Science and Statistics, City University, London, EC1V 0HB.

Renshaw, A.E. and Verrall, R.J. (1994) *A stochastic model underlying the chain ladder technique*. Proceedings XXV ASTIN Colloquium, Cannes.

Taylor, G.C. and Ashe, F.R. (1983) *Second moments of estimates of outstanding claims*. Journal of Econometrics 23, pp37-61.

Verrall, R.J. (1991a) *On the estimation of reserves from loglinear models*. Insurance: Mathematics and Economics 10 pp75-80.

Verrall, R.J. (1991b) *Chain ladder and maximum likelihood*. Journal of the Institute of Actuaries 18, III, pp489-499.

Zehnwirth, B (1989) *The chain ladder technique – A stochastic model*. Claims Reserving Manual, Vol 2, Institute of Actuaries, London.

Zehnwirth, B (1991) Interactive Claims Reserving Forecasting System (ICRFS), Insureware P/L, E. St Kilda, Victoria 3183, Australia.

Appendix – Calculations required by the Bootstrap

Triangle 1 below shows the cumulative paid claims from the Example, together with the traditional chain ladder development factors.

Triangle 1 – Observed Cumulative Data

357848	1124788	1735330	2218270	2745596	3319994	3466336	3606286	3833515	3901463
352118	1236139	2170033	3353322	3799067	4120063	4647867	4914039	5339085	
290507	1292306	2218525	3235179	3985995	4132918	4628910	4909315		
310608	1418858	2195047	3757447	4029929	4381982	4588268			
443160	1136350	2128333	2897821	3402672	3873311				
396132	1333217	2180715	2985752	3691712					
440832	1288463	2419861	3483130						
359480	1421128	2864498							
376686	1363294								
344014									

Development Factors

3.4906	1.7473	1.4574	1.1739	1.1038	1.0863	1.0539	1.0766	1.0177	1.0000

The first stage is to obtain the cumulative fitted values, given the development factors. The fitted cumulative paid to date equals the actual cumulative paid to date, so we can transfer the final diagonal of the actual cumulative triangle to the fitted cumulative triangle. The remaining cumulative fitted values are obtained backwards by recursively dividing the fitted cumulative value at time t by the development factor at time $t-1$. The results of this operation are shown in Triangle 2.

The incremental fitted values, obtained by differencing in the usual way, are shown in Triangle 3.

Triangle 2 – Cumulative Fitted Values

270061	942678	1647172	2400610	2817960	3110531	3378874	3560909	3833515	3901463
376125	1312904	2294081	3343423	3924682	4332157	4705889	4959416	5339085	
372325	1299641	2270905	3309647	3885035	4288393	4658349	4909315		
366724	1280089	2236741	3259856	3826587	4223877	4588268			
336287	1173846	2051100	2989300	3508995	3873311				
353798	1234970	2157903	3144956	3691712					
391842	1367765	2389941	3483130						
469648	1639355	2864498							
390561	1363294								
344014									

Triangle 3 – Incremental Fitted Values

270061	672617	704494	753438	417350	292571	268344	182035	272606	67948
376125	936779	981178	1049342	581260	407474	373732	253527	379669	
372325	927316	971264	1038741	575388	403358	369957	250966		
366724	913365	956652	1023114	566731	397290	364391			
336287	837559	877254	938200	519695	364316				
353798	881172	922933	987053	546756					
391842	975923	1022175	1093189						
469648	1169707	1225143							
390561	972733								
344014									

The unscaled Pearson residuals, shown in Triangle 4, can be obtained using equation 4.2, together with the observed and fitted incremental data.

Triangle 4 – Unscaled Pearson Residuals

168.93	115.01	-111.94	-311.63	170.23	521.04	-235.52	-98.64	-86.91	0.00
-39.14	-54.51	-47.73	130.76	-177.75	-135.47	252.02	25.11	73.64	
-134.09	77.35	-45.71	-21.67	231.27	-403.77	207.21	58.77		
-92.67	203.92	-184.51	533.16	-390.87	-71.77	-261.92			
184.29	-157.75	122.49	-174.18	-20.59	176.15				
71.17	59.56	-78.52	-183.21	215.31					
78.26	-129.87	108.03	-28.62						
-160.76	-99.91	197.16							
-22.20	14.07								
0.00									

A crucial step in performing the bootstrap is resampling the residuals, with replacement. One such sample is shown in Triangle 5. Notice that residuals may appear more than once when resampled with replacement (e.g. 59.56 appears four times). Care must be taken to ensure that **all** residuals have an equal chance of being selected.

Triangle 5 – Example set of resampled residuals

-157.75	207.21	-261.92	115.01	-22.20	14.07	25.11	168.93	-78.52	59.56
-135.47	-135.47	115.01	184.29	-45.71	176.15	-92.67	115.01	-235.52	
215.31	-71.77	0.00	521.04	78.26	-21.67	59.56	-160.76		
-390.87	-183.21	-86.91	-157.75	-235.52	59.56	184.29			
115.01	77.35	-21.67	-45.71	533.16	0.00				
14.07	533.16	-157.75	203.92	-235.52					
-111.94	-183.21	521.04	-98.64						
-403.77	252.02	-86.91							
203.92	-28.62								
59.56									

Using the resampled residuals in Triangle 5, together with the original incremental fitted values in Triangle 3, a bootstrap data sample can be calculated by using equation 4.3. The bootstrap sample associated with the resampled residuals in Triangle 5 is shown in Triangle 6. The associated cumulative sample is shown in Triangle 7, together with development factors obtained by applying the standard chain ladder to the bootstrap data. The bootstrap reserve estimate is obtained from the development factors and cumulative bootstrap sample in the usual way.

Triangle 6 – Incremental Bootstrap data sample

188083	842558	484657	853267	403007	300180	281353	254108	231609	83474
293040	805657	1095099	1238128	546413	519919	317083	311436	234551	
503702	858204	971264	1569774	634753	389594	406186	170432		
130025	738275	871647	863552	389432	434833	475640			
402982	908346	856956	893928	904049	364316				
362166	1381652	771385	1189647	372609					
321773	794937	1548957	990057						
192942	1442279	1128946							
517999	944509								
378950									

Triangle 7 – Cumulative Bootstrap data sample together with development factors

188083	1030642	1515299	2368566	2771573	3071753	3353106	3607214	3838823	3922297
293040	1098697	2193796	3431924	3978337	4498255	4815338	5126774	5361324	
503702	1361906	2333170	3902945	4537698	4927293	5333479	5503911		
130025	868300	1739947	2603500	2992931	3427765	3903405			
402982	1311328	2168284	3062211	3966260	4330576				
362166	1743818	2515203	3704849	4077458					
321773	1116710	2665667	3655724						
192942	1635221	2764167							
517999	1462508								
378950									

Resampled Development Factors

3.992	1.760	1.502	1.170	1.110	1.092	1.054	1.053	1.021

Bootstrap Reserve Estimates

$i = 2$	116580
$i = 3$	419829
$i = 4$	526745
$i = 5$	1041244
$i = 6$	1537217
$i = 7$	2236020
$i = 8$	3927752
$i = 9$	4769853
$i = 10$	6068470
Total	20643712

The process is completed by repeatedly resampling from the residuals N times, where N is large (e.g. $N = 1000$), each time creating a new bootstrap sample and new bootstrap reserve estimates. The bootstrap standard errors of the reserve estimates are simply the standard deviations of the N bootstrap reserve estimates.

It is important to note that the bootstrap standard error so derived is an estimate of the square root of the estimation variance, with no adjustment for the degrees of freedom. To enable a comparison with the analytic estimation variance it is necessary to make the appropriate adjustment. Furthermore, to obtain the prediction error, it is necessary to add the process variance, which in the case of the chain ladder model is the scale parameter multiplied by the original chain ladder reserve estimate. The scale parameter is calculated as the Pearson chi-squared statistic divided by the degrees of freedom, where the Pearson chi-squared statistic is the sum of the (unscaled) Pearson residuals squared (see equation 4.4).

The various components contributing to the prediction error are shown in Table A.1

Table A.1

	Actual Reserve	Bootstrap SD	Variability Parameter	Variability Data	Prediction Error	Prediction Error %
$i = 2$	94634	68556	84737	70554	110265	117%
$i = 3$	469511	122608	151548	157153	218320	46%
$i = 4$	709638	139107	171941	193204	258634	36%
$i = 5$	984889	165159	204142	227610	305745	31%
$i = 6$	1419459	206556	255310	273250	373964	26%
$i = 7$	2177641	291556	360373	338448	494384	23%
$i = 8$	3920301	517972	640230	454107	784925	20%
$i = 9$	4278972	734598	907987	474426	1024461	24%
$i = 10$	4625811	1577154	1949415	493279	2010856	43%
Total	18680856	2298953	2841582	991281	3009523	16%

The bootstrap standard deviation is the standard deviation of 1000 bootstrap reserve estimates. Parameter variability is the bootstrap standard deviation multiplied by $\sqrt{55/36}$, the degrees of freedom adjustment. Data variability is the square root of the product of the scale parameter and the reserve estimates, where the scale parameter is 52,601. The bootstrap prediction error is the square root of the sum of the squares of parameter variability and data variability.