JIA 120 (1993) 171-183

J.I.A. **120**, I, 171–183

NEGATIVE INCREMENTAL CLAIMS: CHAIN LADDER AND LINEAR MODELS

BY R. J. VERRALL, M.A., M.Sc., Ph.D. AND Z. LI, M.Sc.

(of The City University, London)

ABSTRACT

This paper considers the application of loglinear models to claims run-off triangles which contain negative incremental claims. Maximum likelihood estimation is applied using the three parameter lognormal distribution. The method can be used in conjunction with any model which can be expressed in lognormal form. In particular the chain ladder technique is considered. An example is given and the results compared with the basic actuarial method.

KEYWORDS

Chain Ladder; Linear Models; Lognormal Distribution

1. INTRODUCTION

In recent years, a statistical framework for the analysis of claims run-off triangles has been built up. This applies loglinear models to incremental claims which are assumed to be positive. A summary of the theory of loglinear models in claims reserving is given by Verrall (1990).

The standard method of dealing with data sets containing negative incremental claims has been to add a suitably large constant before taking logarithms, and subtracting the constant after forecasts have been made.

This paper examines the method of choosing the constant, the sensitivity of the results to the choice of the constant, and compares the method with the 'standard' actuarial technique, 'chain ladder'.

Kremer (1982) showed that the chain ladder technique is equivalent to applying a two-way analysis of variance model to the logged incremental claims. Direct comparison of the two approaches is, therefore, possible. The chain ladder technique uses the cumulative claims in its calculations, and would, therefore, appear, prima facie, to be untroubled by negative incremental claims. In fact, it emerges from a deeper examination of the chain ladder technique, in the light of the paper by Kremer, that the issue is more subtle. The statistical approach implies that the 'best' estimates (in maximum likelihood sense) involve geometric means (i.e. roots of products) of incremental claims rather than the arithmetic means of incremental claims in the chain ladder technique. The estimates of the development factors in the chain ladder technique can be reformulated to involve incremental claims rather than cumulative claims, and the substitution of arithmetic means for geometric means then appears to be a device for handling negative incremental claims. This can also be done in the statistical model, but it begs the question of whether it is a sensible procedure to follow.

Thus, this paper gives a maximum likelihood estimate of the 'threshold parameter' (which is added to the data before taking logarithms) and compares the results with those using the chain ladder technique.

The paper is set out as follows. Section 2 gives a summary of the use of linear models in claims reserving, and, in particular, the 'chain ladder linear model'. It should be noted that the chain ladder linear model is only one example of linear models which have been applied to claims run-off triangles. Section 3 derives a maximum likelihood estimate of the threshold parameter. Section 4 gives some illustrations of the method for various data sets, and compares the results with the chain ladder technique. Some general issues in the fitting of the models (including the model underlying the chain ladder technique) are discussed. It should be noted that this paper is concerned entirely with the statistical aspects of the problem, the practitioner will use the statistical results in combination with skill and judgement to set the actual reserves.

2. Loglinear Models

A fuller discussion of loglinear models in the context of claims reserving is given elsewhere (see Verrall, 1991a, 1991b). We will assume that the data consist of a triangle of claims, as shown below. There is no loss of generality, as the models can be fitted to other shapes of data.

$$\begin{array}{lllll}
Z_{1,1} & Z_{1,2} & \dots & \dots & Z_{1,n} \\
Z_{2,1} & Z_{2,2} & \dots & Z_{2,n-1} & \\
\vdots & & \ddots & & \\
Z_{n,1} & & & &
\end{array}$$

Z_{ij} = incremental claims in year of business i, development year j.

In some cases, exposure and inflation factors are given which can be used to standardise the data before analysis.

Define $Y_{ij} = \log(Z_{ij})$. A linear model takes the form:

$$Y_{ij} = X_{ij}\boldsymbol{\beta} + e_{ij} \tag{1}$$

where $\boldsymbol{\beta}$ is a vector of parameters, X_{ij} is a row from a design matrix, and e_{ij} is an error with mean zero.

It is usually assumed that e_{ij} are independently and identically distributed, usually with a normal distribution with variance σ^2. Both of these assumptions can be relaxed.

The choice of X_{ij} governs the model which is applied to the data, and several possible models (including the chain ladder model) are given below.

If the triangle of data $\{Y_{ij};\ i=1, \ldots, n;\ j=1, \ldots, n-i+1\}$ is expressed as a vector, the model can be written in the following form:

$$Y = X\beta + e. \tag{2}$$

The parameter estimates can be obtained by maximum likelihood or least squares methods, and are the solution of:

$$X^T X\hat{\beta} = X^T y. \tag{3}$$

An estimate of the error variance, σ^2, can also be obtained.

It is usually assumed that e_{ij} has a normal distribution. In this case the maximum likelihood estimate of the expected value of Y_{ij}, θ_{ij}, can be obtained by direct substitution:

$$\theta_{ij} = e^{X_{ij}\hat{\beta} + \frac{1}{2}\hat{\sigma}^2}. \tag{4}$$

Of the models which can be cast in this form, the most widely-used is the chain ladder which has the form:

$$Y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \tag{5}$$

where μ is the overall mean, α_i is a business year effect, and β_j is a development year effect. (For technical reasons, $\alpha_1 = \beta_1 = 0$.)

The relationship between this loglinear model and the chain ladder technique was first pointed out by Kremer (1982).

As an example, consider a 3×3 triangle of incremental claims:

$$\begin{array}{ccc} Z_{11} & Z_{12} & Z_{13} \\ Z_{21} & Z_{22} & \\ Z_{31} & & \end{array}$$

The model for this triangle (after taking logs of the data) is:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{13} \\ y_{22} \\ y_{31} \end{bmatrix} = \begin{bmatrix} 1\,0\,0\,0\,0 \\ 1\,0\,0\,1\,0 \\ 1\,1\,0\,0\,0 \\ 1\,0\,0\,0\,1 \\ 1\,1\,0\,1\,0 \\ 1\,0\,1\,0\,0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \alpha_3 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{13} \\ e_{22} \\ e_{31} \end{bmatrix} \tag{6}$$

The estimation of the parameters can be performed effectively in a statistical package such as GLIM (see Renshaw, 1989), or in a spreadsheet package such as SuperCalc5 (see Christofides, 1990). The reader is referred to one of these papers, or Verrall (1990), for an example of the application of the model to data consisting of exclusively positive incremental claims.

Other models which can be cast in the loglinear form include the Gamma curve run-off suggested by Zehnwirth (1985), and the exponential tail suggested by

Ajne (1989). Examples of the application of these models are given in the cited references.

When the data contain negative values the following procedure is usually adopted in order to avoid problems with taking logarithms of negative values.

Choose a suitably large constant τ.

Add τ to all incremental claims (so that they are all positive).

Apply the linear model to $\log(Z_{ij}+\tau)$.

Estimate outstanding claims.

Subtract τ from all estimates and forecasts of claims.

It will be seen, in the next section, that this is equivalent to using a 'three parameter lognormal distribution'. The choice of the constant τ (which is to be regarded as a third parameter), can be performed by maximum likelihood methods. At present, it is usually chosen arbitrarily.

3. The Three Parameter Lognormal Distribution

Consider first the standard two parameters distribution with density:

$$f(z) = \frac{1}{z\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} (\log z - \mu)^2 \right\}. \tag{7}$$

Figure 1 shows the density function of a typical two parameter lognormal distribution.

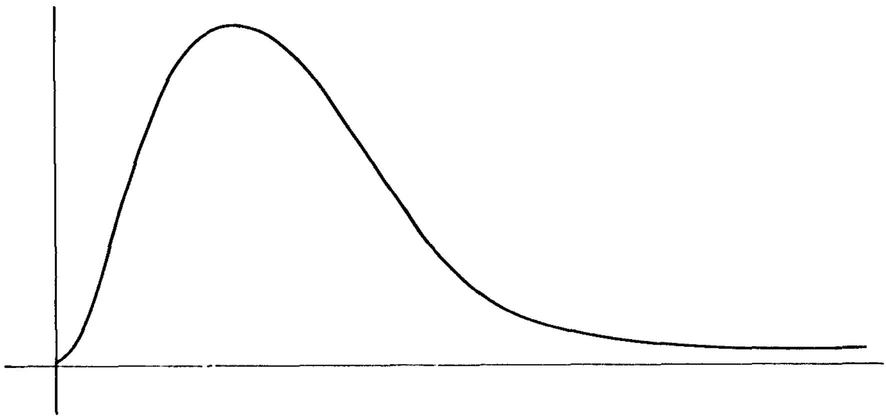

Figure 1. Two Parameter Lognormal Density.

Suppose $Z_1, \ldots, Z_n$ is a sample of independently, identically distributed lognormal random variables with density $f(z)$. The maximum likelihood estimates of the parameters are given by equations (8) and (9):

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} \log z_i \tag{8}$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (\log z_i - \hat{\mu})^2. \tag{9}$$

It can be seen that:

$$\hat{\mu} = \log \left(\prod_{i=1}^{n} z_i \right)^{\frac{1}{n}}. \tag{10}$$

The lognormal distribution is only defined for positive value of Z. If any negative values occur in the sample, it can create problems calculating:

$$\left(\prod_{i=1}^{n} z_i \right)^{\frac{1}{n}}.$$

Two possible ways of dealing with this are as follows.

It is possible to replace the geometric mean:

$$\left(\prod_{i=1}^{n} z_i \right)^{\frac{1}{n}}.$$

by the arithmetic mean:

$$\frac{1}{n} \sum_{i=1}^{n} z_i.$$

This will give a fairly similar result, and is the procedure adopted by the chain ladder technique. It could also be done in the context of the loglinear model.

Alternatively, it is possible to add a constant to all values in the sample, so that none is less than zero. This is equivalent to shifting the lognormal distribution so that it becomes what is known as a three parameter lognormal distribution. Its density is:

$$f(z) = \frac{1}{(z + \tau)\sqrt{(2\pi\sigma^2)}} \exp\left\{ -\frac{1}{2\sigma^2} [\log(z + \tau) - \mu]^2 \right\} \tag{11}$$

and a typical example is illustrated in Figure 2.

The parameter τ is known as the 'threshold parameter'.

Consider now the data in the claims run-off triangle as given at the beginning of Section 2, and the loglinear model as given by equation (2). Incorporating the threshold parameter, τ, the density of Z_{ij} becomes:

$$f(z_{ij}) = \frac{1}{(z_{ij} + \tau)\sqrt{(2\pi\sigma^2)}} \exp\left\{ -\frac{1}{2\sigma^2} [\log(z_{ij} + \tau) - X_{ij}\boldsymbol{\beta}]^2 \right\}. \tag{12}$$

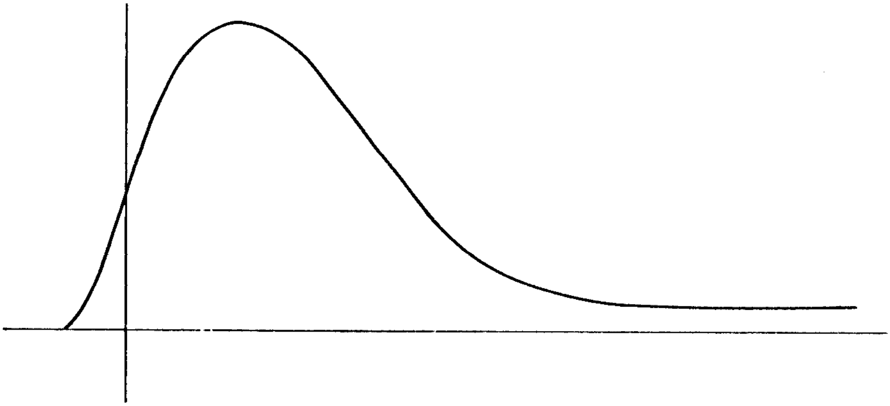

Figure 2. Three Parameter Lognormal Density.

Redefining y_{ij} as:

$$y_{ij} = \log\,(z_{ij} + \tau) \tag{13}$$

the likelihood of the triangle is:

$$\frac{1}{(2\pi\sigma^2)^{N/2}\prod\limits_{i,j}(z_{ij} + \tau)} \exp\left\{-\frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta)\right\} \tag{14}$$

where $N = n(n+1)/2$ is the number of observations in the triangle.

Thus the loglikelihood is:

$$L = -\frac{N}{2}\log\,(2\pi\sigma^2) - \sum_{i,j}\log\,(z_{ij} + \tau) - \frac{1}{2\sigma^2}(y - X\beta)^T(y - X\beta). \tag{15}$$

Differentiating L with respect to β and σ^2 gives the maximum likelihood estimates ostensibly in the same form as before:

$$X^TX\hat{\beta} = X^Ty \tag{16}$$

$$\hat{\sigma}^2 = \frac{1}{N}(y - X\hat{\beta})^T(y - X\hat{\beta}). \tag{17}$$

Note that $y_{ij} = \log(z_{ij} + \tau)$ here.

Also, differentiating with respect to τ gives the following likelihood equation for τ:

$$\hat{\sigma}^2\sum_{i,j}\frac{1}{(z_{ij} + \hat{\tau})} - \frac{1}{\hat{\sigma}^2}\sum_{i,j}\frac{y_{ij} - X_{ij}\hat{\beta}}{(z_{ij} + \hat{\tau})} = 0. \tag{18}$$

Equations (16) and (17) can be solved iteratively with equation (18). No problems have been encountered with the convergence of this procedure, and the results do not appear to be sensitive to the starting values.

The maximum likelihood estimate of θ_{ij} is:

$$\hat{\theta}_{ij} = \exp(X_{ij}\hat{\boldsymbol{\beta}} + \tfrac{1}{2}\hat{\sigma}^2) - \hat{\tau}. \tag{19}$$

The next section illustrates this method, examines the threshold parameter τ, and compares the results with those from the chain ladder technique.

4. Numerical Illustration

As has been previously stated, the linear modelling approach encompasses many different models in common usage, including the chain ladder technique. The illustration in this section will concentrate on the chain ladder technique, but the methods can be used with any of the other loglinear methods. It should also be noted that this section contains only a statistical illustration of the methods. This does not imply that the models applied are necessarily the optimum ones.

We consider the triangle shown in Table 1. The rows and columns are numbered from 1 to 12 for ease of reference.

There are no adjustments to exposure or inflation made to these data.

These data have been obtained from the London Market and, as can be seen,

Table 1. Observed Data

	1	2	3	4	5	6	7	8	9	10	11	12
1	290089	266666	314364	468721	264735	269916	125922	540684	120757	58963	50837	151645
2	401574	648101	673897	656985	458421	373010	31541	279066	98551	177200	−422178	
3	251430	373741	1827086	−429298	801041	746157	109788	212418	101225	−3883		
4	48924	213108	644118	248680	1202333	311357	1067149	697658	650711			
5	62782	278404	880618	611843	243380	335226	205508	164632				
6	10684	109837	189684	581492	69177	323129	207976					
7	271613	290244	587769	660187	681626	413425						
8	151219	183554	485830	431524	427587							
9	97658	141952	369009	450971								
10	51843	119089	530706									
11	145703	421333										
12	21019											

Table 2. Maximum Likelihood Estimates

2	3	4	5	6	7	8	9	10	11	12	1
										193306	2
									−194506	182332	3
								248243	−65584	349210	4
							169469	59075	−220291	148955	5
						189633	50136	−52244	−311331	31111	6
					364805	441258	280667	162806	−135457	258766	7
				338639	225681	296350	147907	38961	−236741	127662	8
			393200	289707	179798	248560	104123	−1883	−270144	84425	9
		261207	359955	258304	150351	217889	76023	−28096	−291581	56675	10
	719916	453479	563166	450255	330343	405363	247781	132128	−160546	226290	11
137773	461643	226566	323342	223720	117922	184112	45078	−56963	−315190	26116	12

contain negative values. The reason for these negative values is not questioned here, although the results suggest that they should be investigated further.

The maximum likelihood estimates of the parameters are as follows:

Threshold parameter = 1474450.

Overall mean = 14·307.

	Row parameters	Column parameters
2	0·017	0·067
3	0·010	0·250
4	0·106	0·120
5	−0·010	0·176
6	−0·085	0·119
7	0·056	0·054
8	−0·023	0·095
9	−0·050	0·008
10	−0·068	−0·062
11	0·037	−0·263
12	−0·089	−0·005

As has been noted in the statistical literature, the likelihood becomes very flat around the maximum likelihood value of the threshold parameter, indicating that values of τ within a large range should be examined. This is considered further in the next section.

Table 2 shows the maximum likelihood estimates of the outstanding claims.

The penultimate row shows the effect of the large negative value of −422178. This is a significant value in the column, and results in all the predicted values being negative. The effect of the other large negative value of −429298 is diluted by the large positive values in its column.

Turning to the chain ladder technique, the estimates of the development factors are shown below.

Column	Estimates of development factor
2	2·7079
3	2·5256
4	1·3658
5	1·3270
6	1·1829
7	1·1164
8	1·1240
9	1·0675
10	1·0226
11	0·9430
12	1·0547

It can be seen that the penultimate development factor is less than 1, and the predicted outstanding claims on the corresponding columns will also be negative. Table 3 shows the chain ladder estimates of outstanding claims.

Table 3. Chain Ladder Estimates

2	3	4	5	6	7	8	9	10	11	12	
											1
										184599	2
									−227259	205719	3
								114915	−296140	268071	4
							187804	67135	−173011	156613	5
						185012	113192	40464	−104276	94393	6
					338009	402130	246028	87949	−226649	205166	7
				307278	231206	275066	168288	60159	−155033	140338	8
			346499	257223	193542	230258	140874	50359	−129778	117477	9
		256639	313368	232628	175037	208242	127405	45544	−117369	106245	10
	865096	523833	639624	474823	357272	425048	260049	92961	−239565	216859	11
35898	86835	52580	64203	47661	35862	42665	26103	9331	−24047	21767	12

The row totals of estimated outstanding claims for each method, together with the total estimated outstanding claims are shown below.

Row	Chain Ladder	Maximum Likelihood
2	184599	193306
3	−21540	−12174
4	86846	531868
5	238541	157208
6	328784	−926694
7	1052634	1372845
8	1027303	938459
9	1206454	1027787
10	1347738	1060727
11	3615999	3368175
12	398858	1374120
The overall total outstanding claims	9466216	9919627

It can be seen that for most rows the results are very similar. Row 6 shows a large difference and should be examined further. The exposure in that year appears to be lower than the other years, which could distort the results. Overall the results are similar, and it can be concluded that if there are not too many negative values, the method of adding a constant can be satisfactory. In this case there would be a lot of sense in examining the data to decide whether the negative values were due to accounting practices. If these were so, they could be adjusted and many difficulties avoided. Therefore, we now also examine the fit of the methods to the observed data.

Table 4 shows the observed values together with the fitted values using maximum likelihood estimation and the chain ladder technique.

Table 4.

Actual Value
Maximum Likelihood Estimate
Chain Ladder Estimate

	1	2	3	4	5	6	7	8	9	10	11	12
1	290089	266666	314364	468721	264735	269916	125922	540684	120757	58963	50837	151545
	159599	272535	623477	368750	473616	365667	251025	322748	172092	61522	-218290	151545
	146335	249922	604548	366066	446984	331817	249670	297033	181728	64923	-167414	151546
2	401574	648101	673897	656985	458421	373010	31541	279066	98551	177200	-422178	
	187641	302515	659480	400381	507047	397245	280636	353589	200348	87881	-196733	
	178252	304432	736404	445907	544473	404189	304124	361816	221364	79132	-203928	
3	251430	373741	1827086	-429298	801041	746157	109788	212418	101225	-3883		
	176704	290823	645439	388045	494008	384929	269087	341561	189328	77601		
	198646	339263	820658	496925	606768	450433	338920	403214	246691	88186		
4	48924	213108	644118	248680	1202333	311357	1067149	697658	650711			
	343015	468629	858963	575643	692280	572214	444704	524477	356911			
	258854	442091	1069393	647539	790675	586956	441644	525425	321461			
5	62782	278404	880618	611843	243380	335226	205508	164632				
	143441	252260	602732	350524	454353	347471	233963	304976				
	151228	258278	624761	378305	461928	342911	258017	306964				
6	10684	109837	189684	581492	69177	323129	207976					
	25997	129700	451949	218048	314340	215217	109948					
	91147	155668	376553	228010	278411	206678	155511					
7	271613	290244	587769	660187	681626	413425						
	252878	372262	743237	473969	584821	470710						
	198112	338351	818452	495589	605137	449222						
8	151219	183554	485830	431524	427587							
	122220	232573	575488	326587	429054							
	135513	231440	559840	338994	413927							
9	97658	141952	369009	450971								
	79129	186504	520164	277981								
	113438	193738	468642	283772								
10	51843	119089	530706									
	51474	156938	484658									
	102592	175214	423833									
11	145703	421333										
	220513	337660										
	209402	357634										
12	21019											
	21019											
	21019											

The most satisfactory residual analysis is to use the percentage errors. These are defined as:

$$\frac{Z_{ij}-\hat{Z}_{ij}}{\hat{Z}_{ij}}$$

where Z_{ij} is the actual incremental claim, and $\hat{Z}_{ij}$ is the fitted value.

Figures 3 and 4 show the percentage errors plotted against fitted values for the maximum likelihood and chain ladder methods.

Overall, the fits appear to be very similar. It is clear from the plots that there are some outlying and unusual observations. The negative fitted values on the left are clearly separated from the rest of the data and should be investigated further. Also there is evidence that some observations are outliers. The negative value in

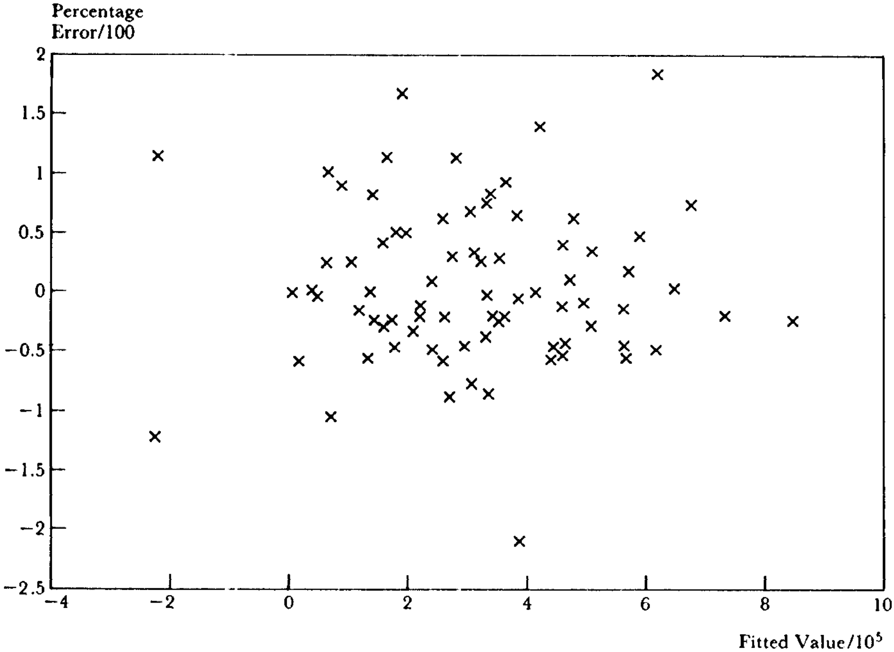

Figure 3. Percentage Error v Fitted Value (Maximum Likelihood Estimates).

row 3, column 4 has a percentage error of over −200% in each model. This is clearly unsatisfactory. There are also some observations which have a positive percentage error of more than 150%.

The largest positive percentage error on the maximum likelihood estimate corresponds to the large value one development year earlier than the negative value in row 3. Clearly these are connected and should be investigated and adjusted before the analysis is performed.

5. Conclusions

The maximum likelihood method can be, as in the illustration given in this paper, an effective method of estimating the threshold parameter. The data and the fitted values should be examined carefully after the model has been fitted. In particular, the effect of outlying observations should be observed, and these should be adjusted if necessary and appropriate. Further research is being carried out on the methods of dealing with outlying observations, including the use of robust estimation methods.

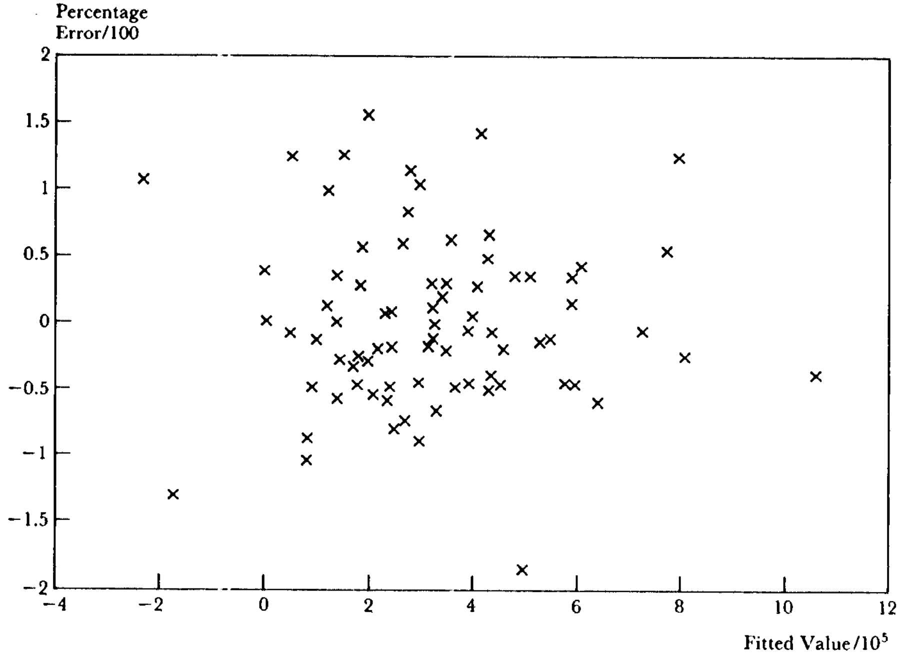

Figure 4. Percentage Error v Fitted Value (Chain Ladder Estimates).

The sensitivity of the results to the threshold parameter should also be examined. For the example given in the previous section, the total forecast outstanding claims were also calculated for other values of the threshold parameter as follows:

τ	Total forecast claims
450000	13455204
1000000	10116739
1474450	9919627
2000000	9785020
5000000	9447599
10000000	9276280
99999999	9077167

where 1474450 is maximum likelihood estimate of the threshold parameter.

In this case the results are not too greatly affected by the choice of the threshold parameter over a very wide range. In other data sets, this is not the case: see Li (1990) for further examples.

The chain ladder technique produces a very similar fit to the loglinear model. There are outlying observations in both cases. The example given in this paper has investigated the similarity between the results from the loglinear model and the chain ladder technique. It should be noted that the chain ladder technique does not necessarily produce 'correct' results, or results with good statistical properties.

ACKNOWLEDGEMENT

Thanks are due to an anonymous supplier of data.

REFERENCES

AJNE, B. (1989). Exponential Runoff. *Claims Reserving Manual*, **2.** Institute of Actuaries.

CHRISTOFIDES, S. (1990). Regression models based on log-incremental claims. *Claims Reserving Manual*, **2.** Institute of Actuaries.

KREMER, E. (1982). IBNR-Claims and the Two-way Model of ANOVA. *Scand. Act. J.* **1,** 47–55.

LI, Z. (1990). Maximum Likelihood Methods in the Estimation of Outstanding Claims. M.Sc. Dissertation, Department of Actuarial Science and Statistics, City University.

RENSHAW, A. E. (1989). Chain Ladder and Interactive Modelling. *J.I.A.* **116,** 559–587.

VERRALL, R. J. (1990). Statistical Aspects of Outstanding Claims Reserving. Presented to a joint meeting of the Royal Statistical Society, General Applications Section, and the Staple Inn Actuarial Society.

VERRALL, R. J. (1991a). Chain Ladder and Maximum Likelihood. *J.I.A.* **118,** 489–499.

VERRALL, R. J. (1991b). On the Unbiased Estimation of Reserves from Loglinear Models. *Insurance; Mathematics and Economics*, **10,** 75–80.

ZEHNWIRTH, B. (1985). *Interactive Claims Reserving Forecasting System*. Benhar Nominees Pty Ltd, Tunuwurra, NSW, Australia.