

CHAIN LADDER BIAS

Taylor, Greg

Taylor Fry Consulting Actuaries
Level 4, 5 Elizabeth Street
Sydney NSW 2000
Australia

Professorial Associate, Centre for Actuarial Studies
Faculty of Economics and Commerce
University of Melbourne
Parkville VIC 3052
Australia

Phone: 61 2 9223 5268

Fax: 61 2 9223 6851

greg@taylorfry.com.au

March 2001

Summary

The chain ladder (CL) is known to provide unbiased estimates of loss reserve under certain conditions. However, these conditions may not always be realistic. For example, they do **not** include the case in which all cells of a run-off triangle are stochastically independent. It is shown that, under this condition, the CL estimate tends to estimate the median loss reserve, and to be downward biased relative to the mean. The approximate bias correction factor is calculated.

Keywords: chain ladder, loss reserving

Table of Contents

1. Introduction	2
2. Chain ladder model	3
3. Model bias	4
4. Bias correction	8
5. Numerical example.....	9
6. Conclusions	9

Appendices

- A Proofs of propositions**
- B Data for numerical example**

1. Introduction

The chain ladder (CL) model is without doubt the most widely known and applied loss reserving technique. Yet comparatively little is known of its stochastic properties.

It has its origins in deterministic loss reserving, but some stochastic developments have been made in recent years. For example, Hertig (1985) developed a stochastic version, as did Renshaw (1989) and Verrall (1989, 1990, 1991). The Renshaw and Verrall (1991) references used Generalised Linear Models. For other references, see Taylor (2000).

As pointed out by Mack (1994), some of these stochastic extensions actually depart from the “classical” CL structure. This structure will be described in Section 2.1, and is the sole focus of the present paper.

Within the classical structure, there have been two main developments.

First, Hachemeister and Stanard (1975) showed that the CL algorithm is maximum likelihood for Poisson claim counts. This result, not well known for some time, was subsequently re-discovered (Verrall, 1991; Schmidt and Wünsche, 1998).

Second, Mack (1993) developed a formula for the standard error of a CL reserve estimate.

Despite these developments, the literature has little to offer on the question of **bias** in CL estimates. Even the more formal of the frameworks mentioned above, such as maximum likelihood estimation, are uninformative on this subject.

It is possible to base the CL model on assumptions that render its estimates unbiased. This is discussed in Section 2. But this may sometimes seem contrived, and it is advisable to examine the bias introduced when these assumptions are breached.

An empirical study of this question has recently been made by Schiegl (2001). Two of the findings were that CL estimates of loss reserves:

- exhibited a **downward bias**
- were very close to unbiased estimators of the median (rather than the mean) loss reserve.

The purpose of the present paper is to provide some theoretical exploration of these results.

2. Chain ladder model

Consider a run-off triangle Δ consisting of quantities $C(i, j)$, $i = 0, 1, \dots, I$; $j = 0, 1, \dots, I - i$. The $C(i, j)$ may denote claim counts or amounts.

Define

$$A(i, j) = \sum_{m=0}^j C(i, m), \quad i = 0, 1, \dots, I; \quad j = 0, 1, \dots, I - i. \quad (2.1)$$

Make the following assumptions.

Assumption 1. $E[A(i, j+1) | A(i, j)] = v(j) A(i, j)$, where the $v(j)$, $j = 0, \dots, I-1$ are constants dependent on j but not i .

Assumption 2. $V[A(i, j+1) | A(i, j)] = \sigma^2(j) A(i, j)$, where the $\sigma^2(j)$, $j = 0, \dots, I-1$ are also constants dependent on j but not i .

Assumption 3. $C(i_1, j_1)$ and $C(i_2, j_2)$ are stochastically independent for $i_1 \neq i_2$.

Now consider the **future** observations associated with Δ , ie $C(i, j)$, $i = 1, 2, \dots, I$; $j = I - i + 1, I - i + 2, \dots, I$ and suppose these also subject to Assumptions 1 and 3. The **loss reserve** associated with Δ is

$$L(\Delta) = \sum_{i=1}^I \sum_{j=I-i+1}^I C(i, j). \quad (2.2)$$

The CL estimation of $L(\Delta)$ proceeds as follows (see eg Taylor, 2000, Section 2.2.4). Estimate $v(j)$ by

$$\hat{v}(j) = \sum_{i=0}^{I-j-1} A(i, j+1) / \sum_{i=0}^{I-j-1} A(i, j). \quad (2.3)$$

By Assumption 1, $\hat{v}(j)$ is unbiased. Indeed, each ratio $A(i, j+1) / A(i, j)$, $i = 0, 1, \dots, I - j - 1$, is an unbiased estimator of $v(j)$.

Estimate $A(i, I)$ by

$$\hat{A}(i, I) = A(i, I-i) \hat{v}(I-i) \hat{v}(I-i+1) \dots \hat{v}(I-1). \quad (2.4)$$

Then $L(\Delta)$ is estimated by

$$\hat{L}(\Delta) = \hat{L}_i(\Delta)$$

with

$$\begin{aligned} \hat{L}_i(\Delta) &= \hat{A}(i, I) - A(i, I-i) \\ &= \sum_{i=1}^I A(i, I-i) [\hat{v}(I-i)\hat{v}(I-i+1)\dots\hat{v}(I-1)-1]. \end{aligned} \quad (2.5)$$

Again by Assumption 1, $\hat{L}(\Delta)$ is unbiased. Indeed, Assumption 1 is crucial to the unbiasedness. Consider, for example, an alternative assumption.

Assumption 1a. All $C(i, j)$ are stochastically independent.

In this case,

$$A(i, j+1)/A(i, j) = 1 + C(i, j+1)/A(i, j) \quad (2.6)$$

with $C(i, j+1)$ and $A(i, j)$ stochastically independent. Then $A(i, j+1)/A(i, j)$, and hence $\hat{v}(j)$ defined by (2.3), are no longer necessarily unbiased estimators of $v(j)$.

3. Model bias

3.1 Theoretical

Now consider the CL estimator (2.5) when Assumption 1a holds instead of Assumption 1. Assumption 2 will **not** be made.

Initially, the quantity of interest will be the CL ratio

$$X(i, j) = A(i, j+1)/A(i, j) \quad (3.1)$$

for fixed but arbitrary i, j and its logged version

$$Y(i, j) = \log X(i, j) = \log A(i, j+1) - \log A(i, j). \quad (3.2)$$

Definition. A random n -tuple (X_1, \dots, X_n) will be said to be **symmetrically distributed** about (μ_1, \dots, μ_n) if

$$\begin{aligned} \text{Prob}[X_i \leq \mu_i - w_i, i = 1, 2, \dots, n] = \\ \text{Prob}[X_i \geq \mu_i + w_i, i = 1, 2, \dots, n] \text{ for all } w_1, \dots, w_n \geq 0. \end{aligned} \quad (3.3)$$

Only the cases $n = 1, 2$ will be required below.

Definition. Random variables X and Y with joint d.f. F will be said to be **exchangeable** if $F(x, y) = F(y, x)$.

Proposition 3.1. Let $A(i, j+1)$ and $kA(i, j)$ be exchangeable for some constant k . Then $Y(i, j)$ is symmetrically distributed about $\log k$.

□

The proof of this appears, with other proofs in Appendix A.

Remark. There is **no assumption** here about stochastic independence between $A(i, j)$ and $A(i, j+1)$.

Proposition 3.2. Let the ordered pair $[\log A(i, j+1), \log A(i, j)]$ be symmetrically distributed. Then $Y(i, j)$ is symmetrically distributed.

□

Proposition 3.3. For given j , let each pair $A(i, j+1)$ and $k_j A(i, j)$, $i = 0, 1, \dots, I - j - 1$, be exchangeable for some constant k_j . Then each $Y(i, j)$ is symmetrically distributed. Moreover, each $Y(i, j)$, $X(i, j)$ and $\hat{v}(j)$ is an unbiased estimate of its median. If the hypotheses hold for each $j = 0, 1, \dots, I-1$, then each $\hat{A}(i, m)$, $i = 1, 2, \dots, I$; $m = I - i + 1, \dots, I$ is an unbiased estimate of its median.

Remark. The conditions of the proposition require only pairwise exchangeability of the $A(i, j+1)$ and $k_j A(i, j)$. For different i , the distributions involved may be quite different.

Proposition 3.4. Suppose that each ordered pair $[\log A(i, j+1), \log A(i, j)]$, $i = 0, 1, \dots, I$; $j = 0, 1, \dots, I - 1$ (that is both past and future) is symmetrically distributed. Then each $Y(i, j)$, $X(i, j)$ and $\hat{v}(j)$ is an unbiased estimator of its median.

□

Remark. Again the distributions of the $A(i, j)$ may differ for different i, j .

Note that, by the symmetry of the distribution of $Y(i, j)$ in Propositions 3.3 and 3.4, it is an unbiased estimator of both its median and its mean. However, because of the non-linearity of the relation between $Y(i, j)$ and $X(i, j)$, the latter is **not** necessarily an unbiased estimator of its mean. Similarly for $\hat{v}(j)$ and $\hat{A}(i, m)$.

In practice, the distributions of the $A(i, j)$ will typically be skewed to the right. The $\hat{v}(j)$ and $\hat{A}(i, m)$ will also have this property. Hence their means exceed their medians. Since the CL estimators of these quantities are unbiased for the medians under the conditions of Proposition 3.3 and 3.4 (and under Assumption 1a), they are **biased downward relative to the means**. This is precisely the result found empirically by Schiegl (2000).

Examples

Example 1. Consider the case in which

$$A \sim N(\mu_A, \sigma_A^2), \quad C \sim N(\mu_C, \sigma_C^2)$$

and A, C are stochastically independent. Here, for brevity A and C denote $A(i, j)$ and $C(i, j+1)$. Define

$$k = (\mu_A + \mu_C) / \mu_A.$$

Then

$$\begin{aligned} kA &\sim N(\mu_A + \mu_C, k^2 \sigma_A^2) \\ A + C &\sim N(\mu_A + \mu_C, \sigma_A^2 + \sigma_C^2). \end{aligned}$$

If $k = [(\sigma_A^2 + \sigma_C^2) / \sigma_A^2]^{1/2}$, then

$$\begin{aligned} kA &\sim N(\mu_A + \mu_C, \sigma_A^2 + \sigma_C^2) \\ A + C &\sim N(\mu_A + \mu_C, \sigma_A^2 + \sigma_C^2) \end{aligned}$$

and $kA, A + C$ are exchangeable, and Proposition 3.3 applies.

The assumption of stochastic independence made here is compatible with Assumption 1a, but not 1. Consider the case, consistent with Assumption 1, in which A is as above and $C = ZA$ with Z, A stochastically independent. Then

$$\begin{aligned} \log(A + C) &= \log A + \log(1 + Z) \\ \log kA &= \log A + \log k. \end{aligned}$$

These two variables will have different variances, and therefore cannot be exchangeable, unless Z has a point distribution. Thus, except for this degenerate case, $A + C$ and kA cannot be exchangeable, and Proposition 3.3 does **not** apply.

Example 2. Consider the case in which A and $A + C$ are jointly log normally distributed, but with A and C stochastically independent, as in Assumption 1a. Then Proposition 3.4 applies.

3.2 Practical

It is possible to produce examples that satisfy the conditions of Proposition 3.3. For example,

$$C(i, j) \sim N(\mu_j, \sigma_j^2) \quad (3.4)$$

with

$$\mu_{j+1} = (k_j - 1) \sum_{m=0}^j \mu_m \quad (3.5)$$

$$\sigma_{j+1}^2 = (k_j^2 - 1) \sum_{m=0}^j \sigma_m^2 \quad (3.6)$$

and subject to Assumption 1a.

This example (which is in fact a re-statement of Example 1 above) is rather contrived, however. Equations (3.5) and (3.6) are restrictive in the relations they allow between the μ_j and σ_j^2 .

In practice, the conditions of Proposition 3.3 are unlikely to hold precisely. It is likely, however, that they will hold approximately, in which case its conclusion will hold approximately.

Similar remarks apply to Proposition 3.4. In this case it is difficult even to produce theoretical examples. This has to do with the fact that the quantities involved in the proposition are logged sums of random variables. Families of variables that are closed under addition typically do not yield tractable log forms.

Once again, however, the proposition may apply approximately. Note particularly the comment towards the end of Section 3.1 that the $A(i, j)$ tend to be right skewed. The log transformation is right tail reducing, and so $\log A(i, j)$ will be less skewed to the right, possibly approximately symmetrical.

The practical conclusion of this discussion is as follows.

Conclusion. If Assumptions 1 to 3 hold, then the CL estimate of loss reserve is unbiased with respect to the mean. If Assumption 1 is replaced by 1a, and if the conditions of Proposition 3.3 or 3.4 hold approximately, then the CL estimate is likely to be biased downward.

□

4. Bias correction

It is shown in the proofs of Propositions 3.3 and 3.4 that $\log \hat{A}(i, m)$ is symmetrically distributed. Hence it is an unbiased estimator of its mean. Consider the bias of

$$\hat{A}(i, m) = \exp \log \hat{A}(i, m). \quad (4.1)$$

First note the result that, for any symmetrically distributed random variable Z , with mean μ and variance σ^2 , and any entire function $f(\square)$,

$$\begin{aligned} E[f(Z)] &= E[f(\mu + (Z - \mu))] \\ &= E\left[\sum_{k=0}^{\infty} f^{(k)}(\mu)(Z - \mu)^k / k!\right] \\ &\quad \text{[by Taylor's theorem]} \\ &= f(\mu) + \frac{1}{2}\sigma^2 f''(\mu) + 4^{\text{th}} \text{ and higher order terms.} \end{aligned} \quad (4.2)$$

With $f(\square) = \exp(\square)$, (4.2) gives the following third order approximation:

$$E[\exp Z] = (\exp \mu) \left(1 + \frac{1}{2}\sigma^2\right). \quad (4.3)$$

Taking $Z = \log \hat{A}(i, m)$, one finds the following third order approximation:

$$\text{bias correction} \left[\hat{A}(i, m) \right] = \hat{A}(i, m) \times \frac{1}{2} V \left[\log \hat{A}(i, m) \right]. \quad (4.4)$$

An estimate of the variance term in (4.4) is given by Hertig (1985), also summarised in Section 7.3 of Taylor (2000). Hertig adopts the assumption that the $Y(i, j)$ (as defined by (3.2)) are normally distributed, stochastically independent and that

$$V[Y(i, j)] = \sigma_j^2, \text{ independent of } i. \quad (4.5)$$

Then (4.4) becomes (see eg Taylor, 2000, equation (7.20)):

$$\text{bias correction} \left[\hat{A}(i, m) \right] = \hat{A}(i, m) \times \frac{1}{2} \sum_{h=I-i}^{m-1} \hat{\sigma}_h^2 \left[1 + 1/(I-j) \right] \quad (4.6)$$

where

$$\hat{\sigma}_h^2 = \sum_{i=0}^{I-j-1} \left[Y(i, j) - \hat{\xi}_j \right]^2 / (I-j-1) \quad (4.7)$$

$$\hat{\xi}_h = \sum_{i=0}^{I-j-1} Y(i, j) / (I - j). \quad (4.8)$$

Note that $\hat{\xi}_j$ is an estimator of $\log v(h)$ but differs from $\log \hat{v}(h)$. The variance term appearing in (4.6) is in fact a mean square error of prediction (MSEP).

Actually, the bias correction calculated by Hertig is:

$$\text{bias correction} [\hat{A}(i, m)] = \hat{A}(i, m) \left\{ \exp \frac{1}{2} \sum_{h=I-i}^{m-1} \hat{\sigma}_h^2 [1 + 1/(I - j)] - 1 \right\} \quad (4.6a)$$

which is the same as (4.6) to first order in $\hat{\sigma}_h^2$ and is exact for normal $Y(i, j)$, ie log normal $A(i, j)$.

With $m = I$, (4.6) or (4.6a) gives the bias correction for the loss reserve $\hat{L}_i(\Delta)$.

5. Numerical example

Taylor (2000, pp. 204-209) applies Hertig's model to the run-off triangle set out in Appendix B. Bias correction is included in accordance with (4.6a).

The magnitude of this correction can be ascertained by re-calculating the $\hat{A}(i, I)$ with the $\hat{\sigma}_h^2$ set to zero.

This produces a reduction of about \$2.7M, or 1.7%, in the loss reserve reported in Taylor's Table 7.6.

This is less than some of the biases measured by Schiegl (2001). On the other hand, the triangle in Appendix B is comparatively well behaved, whereas Schiegl's examples include some extreme cases, such as Pareto close to infinite variance.

6. Conclusions

The main empirical conclusions of Schiegl (2001) under Assumption 1a were as follows.

S1. The CL estimates the median loss reserve without bias.

S2. The CL produces a downward bias in estimating the mean loss reserve.

S3. When each cell of the run-off triangle is a compound Poisson realisation, the size of the bias reduces as the number of claims involved in the run-off triangle increased, though it does not approach zero in the limit.

S4. When each cell of the run-off triangle is a compound Poisson realisation, increasing the variance of the individual claim size severity increase the CL bias.

Conclusions S1 and S2 are the same as derived here in Section 3.

Conclusion S4 follows directly from (4.6) or (4.6a). It is exemplified by a comparison of the numerical example of Section 5, with some of the more extreme examples of Schiegl (see the comment in Section 5).

Conclusion S3 can also be derived, at least in part, from (4.6) or (4.6a) when a log normal approximation is reasonable for individual cells. In this case,

$$\sigma_h^2 = \log \left\{ 1 + [\text{coefficient of variation (c.v.)}]^2 \right\} \quad (6.1)$$

As the expected number of claims in a cell increases, the c.v. of the compound Poisson variate reduces, and so the CL bias reduces. In the limit for large numbers of claims, the c.v. would approach zero, and therefore so would σ_h^2 , and hence the bias. This suggests that the “saturation effect” observed by Schiegl might have been eliminated if larger numbers of claims had been considered.

CL estimates of loss reserve are unbiased with respect to the mean under Assumption 1, but biased under Assumption 1a. Each application of the CL therefore calls for a judgement as to which assumption is the more accurate.

Appendix A Proofs of propositions

Proof of Proposition 3.1. It follows from the definition of exchangeability that, for exchangeable X and Y , $X-Y$ is symmetrically distributed about zero. Apply this result to the case $X = \log A(i, j+1)$, $Y = \log[kA(i, j)]$. This gives $Y(i, j) - \log k$ as symmetrically distributed about zero. The result follows. \square

Proof of Proposition 3.2. Suppose initially that (X_1, X_2) is symmetrically distributed about $(0,0)$. Let $Y = X_1 - X_2$. Then

$$\text{Prob}[Y \leq -y] = \int_S d \text{Prob}[X_1 \leq x_1, X_2 \leq x_2] \quad (\text{A.1})$$

where

$$S = \{(x_1, x_2) : -x_1 - (-x_2) \leq y\} = \{(x_1, x_2) : x_1 - x_2 \geq y\}. \quad (\text{A.2})$$

By (A.1) and the symmetry of (X_1, X_2) ,

$$\text{Prob}[Y \leq -y] = \int_S d \text{Prob}[X_1 \geq x_1, X_2 \geq x_2] = \text{Prob}[Y \geq y]$$

by (A.2).

Thus, Y is symmetrically distributed about zero. A simple modification adapts this proof to the case where (X_1, X_2) are symmetrically distributed about a point other than $(0,0)$. \square

Proof of Proposition 3.3. The result for $Y(i, j)$ follows immediately from the symmetry of its distribution, established by Proposition 3.1. By (3.2), $X(i, j)$ is related one-one to $Y(i, j)$, and so the result follows for $X(i, j)$.

By (2.3),

$$\log \hat{v}(j) - \log k_j = \log \sum_{i=0}^{l-j-1} A(i, j+1) - \log \sum_{i=0}^{l-j-1} k_j A(i, j). \quad (\text{A.3})$$

By hypothesis, and by stochastic independence with respect to i , the two members on the right side of (A.3) are exchangeable. Hence $\log \hat{v}(j)$ is symmetrically distributed about $\log k_j$, the proof of this parallel to that of Proposition 3.1. Then $\log \hat{v}(j)$ is an unbiased estimate of its median, and hence similarly for $\hat{v}(j)$.

By (2.9),

$$\log \hat{A}(i, m) = \log A(i, I-i) + \log \hat{v}(I-i) + \dots + \log \hat{v}(m-1). \quad (\text{A.4})$$

Since the $\log \hat{v}(j)$ have just been shown symmetrically distributed, so must be $\log \hat{A}(i, m)$, and so $\log \hat{A}(i, m)$ is an unbiased estimator of its median. Hence, so is $\hat{A}(i, m)$. \square

Proof of Proposition 3.4. The result for $Y(i, j)$ follows immediately from the symmetry of its distribution, established by Proposition 3.2. The result for $X(i, j)$ then follows just as in Proposition 3.3.

By (2.3),

$$\log \hat{v}(j) = \log \sum_{i=0}^{I-j-1} A(i, j+1) - \log \sum_{i=0}^{I-j-1} A(i, j). \quad (\text{A.5})$$

By the symmetry hypothesis, and stochastic independence with respect to i , the two members of the right side of (A.5) for a symmetrically distributed ordered pair, $\log \hat{v}(j)$ is symmetrically distributed, the proof of this parallel to that of Proposition 3.2.

The remainder of the proof follows that of Proposition 3.3. \square

Appendix B Data for numerical example

	Incurred losses																	
Period of origin	Incurred losses to end of development year n=																	
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000	\$000
1978	9,268	18,263	20,182	22,383	22,782	26,348	26,172	26,184	25,455	25,740	25,711	25,452	25,460	25,422	25,386	25,520	25,646	25,469
1979	9,848	16,123	17,099	18,544	20,534	21,554	23,219	22,381	21,584	21,408	20,857	21,163	20,482	19,971	19,958	19,947	19,991	
1980	13,990	22,484	24,950	33,255	33,295	34,308	34,022	34,023	33,842	33,933	33,570	31,881	32,203	32,345	32,250	32,168		
1981	16,550	28,056	39,995	42,459	42,797	42,755	42,435	42,302	42,095	41,606	40,440	40,432	40,326	40,337	40,096			
1982	11,100	31,620	40,852	38,831	39,516	39,870	40,358	40,355	40,116	39,888	39,898	40,147	39,827	40,200				
1983	15,677	33,074	35,592	35,721	38,652	39,418	39,223	39,696	37,769	37,894	37,369	37,345	37,075					
1984	20,375	33,555	41,756	45,125	47,284	51,710	52,147	51,187	51,950	50,967	51,461	51,382						
1985	9,800	24,663	36,061	37,927	40,042	40,562	40,362	40,884	40,597	41,304	42,378							
1986	11,380	26,843	34,931	37,805	41,277	44,901	45,867	45,404	45,347	44,383								
1987	10,226	20,511	26,882	32,326	35,257	40,557	43,753	44,609	44,196									
1988	8,170	18,567	26,472	33,002	36,321	37,047	39,675	40,398										
1989	10,433	19,484	32,103	38,936	45,851	45,133	45,501											
1990	9,661	23,808	32,966	42,907	46,930	49,300												
1991	14,275	25,551	33,754	38,674	41,132													
1992	13,245	29,206	36,987	44,075														
1993	14,711	27,082	34,230															
1994	12,476	23,126																
1995	9,715																	

References

- Hachemeister, C. A and Stanard, J.N. (1975). IBNR claims count estimation with static lag functions. Paper presented to the XIIth ASTIN Colloquium, Portimão, Portugal.
- Hertig, J. (1985). A statistical approach to the IBNR-reserves in marine reinsurance. **ASTIN Bulletin**, 15, 171-183.
- Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. **ASTIN Bulletin**, 23, 213-221.
- Mack, T. (1994). Which stochastic model is underlying the chain ladder method? **Insurance: mathematics and economics**, 15, 133-138.
- Renshaw, A.E. (1989). Chain ladder and interactive modelling (claims reserving and GLIM). **Journal of the Institute of Actuaries**, 116, 559-587.
- Schiegl, M. (2001). On the safety loading for chain ladder estimates: a Monte Carlo simulation study. Unpublished.
- Schmidt, K.D. and Wünsche A. (1998). Chain ladder, marginal sum and maximum likelihood estimation. **Blätter der Deutsche Gesellschaft Versicherungsmathematiker**, 23, 267-277.
- Taylor, G. (2000). **Loss reserving: an actuarial perspective**. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Verrall, R.J. (1989). A state space representation of the chain ladder linear model. **Journal of the Institute of Actuaries**, 116, 589-609.
- Verrall, R.J. (1990). Bayes and empirical bayes estimation for the chain ladder model. **Astin Bulletin**, 20, 217-243.
- Verrall, R.J. (1991). Chain ladder and maximum likelihood. **Journal of the Institute of Actuaries**, 118, 489-499.