ISBN 0 7340 2190 9

Chain Ladder Bias

by

Greg Taylor

The University of Melbourne

RESEARCH PAPER NUMBER 92

January 2002

Centre for Actuarial Studies
Department of Economics
The University of Melbourne
Victoria 3010
Australia

CHAIN LADDER BIAS

Greg Taylor

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

November 2001

Summary. The chain ladder forecast of outstanding losses is known to be unbiased under suitable assumptions. According to these assumptions, claim payments in any cell of a payment triangle are dependent on those from preceding development years of the same accident year. If all cells are assumed stochastically independent, the forecast is no longer unbiased. Section 6 shows that, under very general assumptions, it is biased upward. This result is linked to earlier work on some stochastic versions of the chain ladder.

Keywords. Chain ladder, IBNR.

1. Introduction

The chain ladder (CL) approach to estimation of a loss reserve is well known. It is described, for example, by Taylor (2000).

Its origins are not altogether clear, but it seems likely that it originated as a heuristic device. As such, it may be viewed as a non-parametric estimator. The precise definition is given in Sections 2 and 3.

Kremer (1982) recognised that the CL involved a log-linear cross-classification structure. A number of parametric stochastic versions of the CL developed from this, eg Hertig (1985), Renshaw (1989), Verrall (1989, 1990, 1991).

Mack (1994) pointed out that these stochastic models gave mean estimates of liability that differed from the "classical" CL estimate. While the form of stochastic model underlying the classical CL was speculative, due to the latter's heuristic nature, Mack suggested one. It is distribution free. Details are given in Section 2. Mack also identified the differences between this and the other stochastic models.

Whereas the cross-classified models typically assume stochastic independence of all cells in the data set, the CL (in Mack's formulation) does not. It was shown by Mack (1993) that the algorithm of the classical CL produced unbiased forecasts of liability under its own assumptions.

However, it does not necessarily do so under the alternative assumption of independence between all cells. Some papers have studied the bias in estimates of liability in the parametric cross-classified models mentioned above, but little is known of the bias in the classical CL forecast when all cells are independent.

The purpose of the present paper is to investigate the direction of bias in this case.

2. Framework and notation

Consider a square array X of stochastic quantities $X(i,j) \geq 0$, $i = 0,1,\ldots,I$; $j = 0,1,\ldots,I$.

Denote row sums and column sums as follows:

$$R(i,j) = \sum_{h=0}^{j} X(i,h) \tag{2.1}$$

$$C(i,j) = \sum_{g=0}^{i} X(g,j). \tag{2.2}$$

In addition introduce the following notation for the total sum over a rectangular subset of X:

$$\begin{aligned} T(i,j) &= \sum_{g=0}^{i}\sum_{h=0}^{j} X(g,h) \\ &= \sum_{g=0}^{i} R(g,j) \\ &= \sum_{h=0}^{j} C(i,h). \end{aligned} \tag{2.3}$$

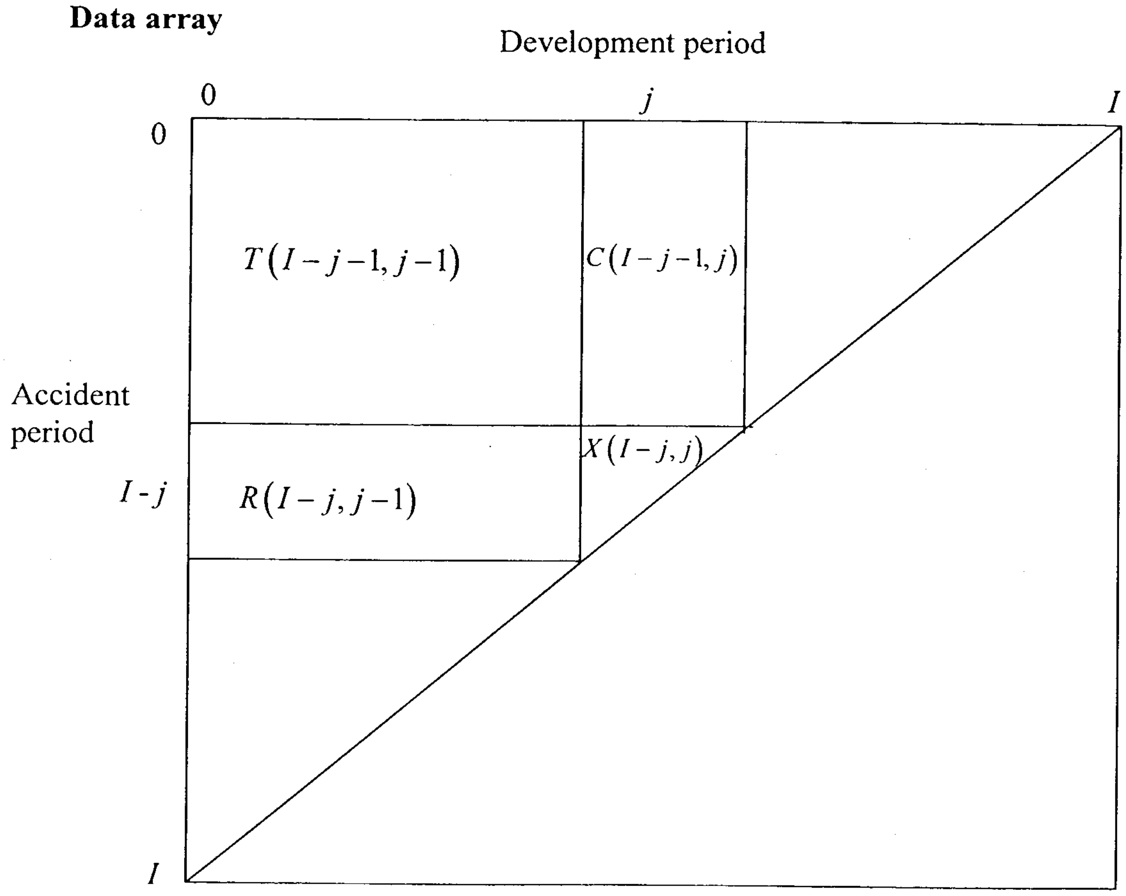
Generally, in the following, any summation of the form $\sum_{a}^{b}$ with $b < a$ will be taken to be zero.

In a typical loss reserving framework, i denotes accident period, j development period, and available data will consists of observations on the triangular subset Δ of X:

$$\Delta = \{X(i,j), i = 0,1,\ldots,I; j = 0,1,\ldots,I-i\} \tag{2.4}$$

Figure 2.1 illustrates the situation.

Figure 2.1
Data array



Still in a loss reserving context, Δ would represent some form of claims experience, eg claim counts or claim amounts. The loss reserving problem consists of forecasting the lower triangle in Figure 2.1, conditional on Δ. There is particular interest in forecasting $R(i,I) \mid \Delta, \quad i=1,\dots,I$.

3. Chain ladder forecast

Define

$$\begin{aligned}\hat{v}(j) &= T(I-j-1,j+1)/T(I-j-1,j) \\ &= 1+C(I-j-1,j+1)/T(I-j-1,j)\end{aligned} \tag{3.1}$$

and

$$\hat{R}(i,I) = R(i,I-i)\prod_{k=I-i}^{I-1}\hat{v}(k) \tag{3.2}$$

The value of $\hat{R}(i,I)$ calculated in this way will be referred to as the **chain ladder forecast** (CLF) of $R(i,I)$.

4. Chain ladder models

4.1 Dependent increments

The CLF has been formulated in Section 3 just as an algorithm. No model for the data X has yet been stated.

It is evident that the properties of the CLF will depend on the model. This and the next sub-section consider two alternative models. The first is represented by the following two assumptions.

Assumption 1. $E\left[R(i,j+1) \mid X(i,0), X(i,1),\dots,X(i,j)\right] = v(j)R(i,j)$. (4.1)

Assumption 2. $X(i_1,j_1)$ and $X(i_2,j_2)$ are stochastically independent for $i_1 \neq i_2$.

Remark 1. It follows from Assumptions 1 and 2 that

$$E\left[R(i,j+1) \mid \Delta\right] = v(j)R(i,j) \tag{4.2}$$

for any $j \geq I-i$ (ie future j).

Since $R(i,j+1) = R(i,j) + X(i,j+1)$, one may re-write (4.2) in the form:

$$E\left[X(i,j+1) \mid \Delta\right] = \left[v(j)-1\right]R(i,j). \tag{4.3}$$

Remark 2. It is clear from (4.3) that $R(i,j)$ and $X(i,j+1)$ are **not** independent. Generally, under Assumptions 1 and 2, the $X(i,j)$ for fixed i are **not** independent.

Theorem 1 (Mack). Under Assumptions 1 and 2,

(1) $\hat{\nu}(j)$ is an unbiased estimator of $\nu(j)$ for $j = 0,1,\dots, I-1$; and

(2) the CLF $\hat{R}(i,I)$ is an unbiased estimator of $E[R(i,I)|\Delta]$ for i = 1,2,.., I.

Proof. See Mack (1993).

□

It is also convenient to re-write (4.2) in the form:

$$E[R(i,j+1)/R(i,j)|\Delta] = \nu(j). \tag{4.4}$$

4.2 Independent increments

Replace Assumptions 1 and 2 by 1a, 2a and 3 as follows.

Assumption 1a. $E[R(i,j+1)]/E[R(i,j)] = \eta(j).$ (4.5)

Assumption 2a. $X(i_1,j_1)$ and $X(i_2,j_2)$ are stochastically independent for $(i_1,j_1) \neq (i_2,j_2)$.

Define the set

$$D_i = \{(g,h): g \le I-k-1, h \le k+1, \ \ k = I-i,\dots,I-1\}. \tag{4.6}$$

Assumption 3. $T(g,h) > 0$ for $(g,h) \in D_i$.

Remark 3. It is implicit in Assumption 1a that $E[R(i,j)] \neq 0$. By the assumed non-negativity of the $X(i,j)$, $E[R(i,j)] > 0$ for each i,j.

Remark 4. A comparison of (4.4) and (4.5) indicates that $\nu(j)$ and $\eta(j)$ are different quantities (for fixed j) since

$$E[R(i,j+1)/R(i,j)] \neq E[R(i,j+1)]/E[R(i,j)]. \tag{4.7}$$

This fact was pointed out by Mack (1994).

By Assumption 3, applied to (3.1), all $\hat{\nu}(k)$ appearing in (3.2) are defined and strictly positive.

The conditions of Theorem 1 no longer hold, and so the CLF is not necessarily unbiased.

5. Earlier results

A definitive result under Assumptions 1a and 2a was obtained by Hachemeister and Stanard (1975). This result was not well known for some time and was subsequently re-discovered (Renshaw and Verrall, 1994; Schmidt and Wünsche, 1998).

Theorem 2 (Hachemeister and Stanard). Suppose that

$$X(i,j) \sim \text{Poisson}\left[a(i)b(j)\right] \tag{5.1}$$

for constants $a(i), b(j) > 0,\ i = 0,1,...,I;\ b = 0,1,...,I$. Suppose also that Assumption 2a holds. Then $\hat{\nu}(j)$ is the MLE of $\sum_{h=0}^{j+1} b(h) / \sum_{h=0}^{j} b(h)$ and $\hat{R}(i,I)$ the MLE of $R(i,I)$.

□

Remark 5. The quantity $\sum_{h=0}^{j+1} b(h) / \sum_{h=0}^{j} b(h)$ is equal to $\eta(j)$.

In fact, the results of Schmidt and Wünsche establish equivalence between CLF and MLE for more general distributional assumptions than (5.1). However, these more general cases do not fall within the current area of interest in which Assumption 2a holds.

Note that, because of (5.1), Theorem 2 deals with one of the log-linear cross-classification structure mentioned in Section 1. The theorem provides a case in which the CLF is ML.

Bias in models of this type was investigated by Verrall (1991) and Doray (1996). Suppose that

$$\log X(i,j) = \log a(i) + \log b(j) + \varepsilon(i,j) \tag{5.2}$$

where

$$\varepsilon(i,j) \sim N\left(0, \sigma^2\right) \tag{5.3}$$

and the $\varepsilon(i,j)$ are stochastically independent.

In this structure, Assumptions 1a and 2a hold with $\eta(j)$ as in Remark 5.

Suppose the data triangle Δ is available and let $\hat{\theta}$ denote an estimate of IBNR claims:

$$\hat{\theta} = \sum_{i=1}^{I} \sum_{j=I-i+1}^{I} \hat{\theta}(i,j) \tag{5.4}$$

where $\hat{\theta}(i,j)$ is an estimator of $X(i,j)$.

The MLE's of the parameters in (5.2) and (5.3) are found by regression, because of the normal error structure. The MLE of σ^2 takes the form

$$\hat{\sigma}^2 = \text{RSS}/n \tag{5.5}$$

where RSS is the residual sum of squares and n the number of data points $\left(=(I+1)(I+2)/2\right)$.

This last estimator is biased, and its unbiased form is:

$$\tilde{\sigma}^2 = \text{RSS}/(n-p) \tag{5.6}$$

where p is the number of parameters (other than σ^2) in (5.2).

Doray considers a number of estimators $\hat{\theta}$, including:

- $\hat{\theta}_U$, the unique uniformly minimum variance unbiased estimator;
- $\hat{\theta}_V$, a modification of the MLE in which $\hat{\sigma}^2$ is replaced by $\tilde{\sigma}^2$.

Theorem 3 (Doray). Suppose that $X(i,j)$ are log normally distributed according to (5.2) and (5.3) with the $\varepsilon(i,j)$ stochastically independent (so that Assumptions 1a and 2a hold). Then

$$\hat{\theta}_U < \hat{\theta}_V \tag{5.7}$$

$$E\left[\hat{\theta}_U\right] < E\left[\hat{\theta}_V\right]. \tag{5.8}$$

□

This shows that, for the log normal case, $\hat{\theta}_V$, an approximation to MLE is biased upward.

There is no particular connection between Theorems 2 and 3, one relating to Poisson variates and the other to log normal. Nonetheless, the results obtained are that:

- the CLF is MLE in one case;
- an MLE approximation is biased upward in the other case.

This raises a question as to whether the CLF is generally biased upward. Section 6 considers this question.

6. Chain ladder bias

Theorem 4. Define

$$Y=\prod_{k=I-i}^{I-1}\hat{v}(k). \tag{6.1}$$

Then

$$\frac{\partial^2 Y}{\partial X^2(g,h)}=0 \text{ for } (g,h)\notin D_i; \tag{6.2}$$

for $(g,h)\in D_i$ and $h\leq I-i$,

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)}=2\sum_{k=I-i}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}\text{x}\left[\frac{1}{T(i-1,I-i)}+\sum_{l=I-i+1}^{k}\frac{R(I-l,l)}{T(I-l-1,l)T(I-l,l)}\right] \tag{6.3}$$

for $(g,h)\in D_i$ and $h>I-i$,

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)}=2\sum_{k=h}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}\sum_{l=h}^{k}\frac{R(I-l,l)}{T(I-l-1,l)T(I-l,l)}. \tag{6.4}$$

These results do **not** depend on the Assumptions 1a, 2a and 3.

Proof. See appendix. □

Theorem 5. Under Assumptions 1a, 2a and 3, and if $X(g,h)$ is not degenerate for at least one $(g,h)\in D_i$, the CLF $\hat{R}(i,I)$ is biased upward as an estimate of $E\left[R(i,I)\right]$ in the sense that

$$E\left[\hat{R}(i,I)\right]>E\left[R(i,I)\right]. \tag{6.5}$$

Proof. See appendix. □

Remark 6. An alternative form of (6.5) is:

$$E_{R(i,I-i)}E\left[\hat{R}(i,I)\,|\,R(i,I-i)\right] > E\left[R(i,I)\right]. \tag{6.6}$$

Note that Theorem 5 does **not** state that

$$E\left[\hat{R}(i,I)|\,R(i,I-i)\right] > E\left[R(i,I)\right].$$

Indeed, it follows from (3.2), (6.1) and the fact that Y is independent of $R(i, I-i)$ that

$$E\left[\hat{R}(i,I)|\,R(i,I-i)\right] = R(i,I-i)E\left[Y\right] > E\left[R(i,I)\right]\frac{R(i,I-i)}{E\left[R(i,I-i)\right]}, \tag{6.7}$$

the last step following from (A.23).

It is evident that (6.6) (equivalently (6.5)) follows from (6.7).

7. Conclusion

Theorem 5 shows that, under very general distribution free conditions, **the CLF is biased upward**. A simulation test of prediction bias in the chain ladder and other models was carried out by Stanard (1985). One of his experiments dealt with the case in which the total number of claims in an accident year is a Poisson variate and is multinomially distributed over development years. It may be shown that distinct cells in a row of the claim count triangle are then stochastically independent, and so Theorem 5 applies.

Stanard's simulations did in fact find upward bias in the CLF.

8. Acknowledgements

Dr Alois Gisler read earlier drafts of this paper and provided a number of valuable comments and insights. The paper possibly would not have occurred without his assistance.

The author also acknowledges helpful comment by Dr Thomas Mack.

Appendix
Proof of theorems

Lemma 1. Suppose that

$$y = \prod_{i=1}^{n} f_i(x) \tag{A.1}$$

with $x = (x_1, ..., x_m)^T$ and $f_i(x) > 0$ for each i. Then

$$\frac{1}{y}\frac{\partial^2 y}{\partial x_k^2} = \sum_{i=1}^{n} \frac{1}{f_i}\frac{\partial^2 f_i}{\partial x_k^2} + \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{1}{f_i f_j}\frac{\partial f_i}{\partial x_k}\frac{\partial f_j}{\partial x_k}.$$

Proof. Differentiate $\log y$ twice. □

Lemma 2 (multivariate Jensen inequality). Let $X = (X_1, ..., X_m)^T$ where the X_k are pairwise stochastically independent random variables. Let $f : R^m \to R$ be twice differentiable in all its arguments, and suppose that

$$\partial^2 f(X) / \partial X_k^2 \geq 0 \text{ for all } X \text{ and for } k = 1, 2, \ldots, m. \tag{A.2}$$

Then

$$E[f(X)] \geq f(E[X]). \tag{A.3}$$

If strict inequality holds in (A.2) for at least one k, and X_k is not degenerate, then strict inequality holds in (A.3).

Proof. Expand $f(X)$ as the Taylor series:

$$\begin{aligned} f(X) &= f(\mu) + (\partial f(\mu)/\partial X)(X-\mu) \\ &\quad + \tfrac{1}{2}(X-\mu)^T \partial^2 f(\xi)(X-\mu) \end{aligned} \tag{A.4}$$

where $\mu = (\mu_1, ..., \mu_m)^T = E[X]$ and $\xi = \mu + \theta^T(X - \mu)$ for some m-vector θ. Then

$$\begin{aligned} E[f(X)] &= f(\mu) + \tfrac{1}{2}\sum_{k,l=1}^{m} \left[\partial^2 f(\xi)/\partial X_k \partial X_l\right] \mathrm{Cov}[X_k, X_l] \\ &= f(\mu) + \tfrac{1}{2}\sum_{k=1}^{m} \left[\partial^2 f(\xi)/\partial X_k^2\right] \mathrm{Var}[X_k] \end{aligned} \tag{A.5}$$

$$\geq f(\mu), \tag{A.6}$$

where (A.5) follows from the pairwise independence of the X_k, and (A.6) follows from (A.2).

The result (A.6) is the same as (A.3). It is evident that if one of the derivatives in (A.5) is strictly positive and X_k is not degenerate for that k, then the inequality in (A.6) is strict.

□

Proof of Theorem 4. Consider Y defined by (6.1), with $\hat{v}(k)$ defined by (3.1) and (2.3). The observations $X(g,h)$ involved in the $\hat{v}(k)$ constituting Y are just those in D_i. This justifies (6.2).

Now consider $(g,h) \in D_i$. Note that Lemma 1 is applicable to Y because of (6.1). Hence

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)} = \sum_{k=I-i}^{I-1} \frac{1}{\hat{v}(k)}\frac{\partial^2 \hat{v}(k)}{\partial X^2(g,h)} + \sum_{\substack{k,l=I-i \\ k \neq l}}^{I-1} \frac{1}{\hat{v}(k)\hat{v}(l)}\frac{\partial \hat{v}(k)}{\partial X(g,h)}\frac{\partial \hat{v}(l)}{\partial X(g,h)}. \quad \text{(A.7)}$$

There are two cases to be considered in the evaluation of (A.7), according to whether $h \le I-i$ or $h > I-i$.

Case I: $h \le I-i$.

By (3.1),

$$\hat{v}(k) = 1 + C(I-k-1,k+1)/T(I-k-1,k), \quad k = I-i,\ldots,I-1. \quad \text{(A.8)}$$

Thus $X(g,h)$ will not appear in the numerator of the fraction here. It will appear as a summand in the dominator, provided that $g \le I-k-1,\ h \le k$. Then, for $g \le I-k-1,\ h \le k$,

$$\partial \hat{v}(k)/\partial X(g,h) = -C(I-k-1,k+1)/T^2(I-k-1,k) \quad \text{(A.9)}$$

$$\partial^2 \hat{v}(k)/\partial X^2(g,h) = 2C(I-k-1,k+1)/T^3(I-k-1,k). \quad \text{(A.10)}$$

Otherwise, $\partial \hat{v}(k)/\partial X(g,h) = \partial^2 \hat{v}(k)/\partial X^2(g,h) = 0$.

Substitution of (3.1), (A.9) and (A.10) into (A.7) yields

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)} = 2\sum_{k=I-i}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T^2(I-k-1,k)}$$
$$+2\sum_{k=I-i}^{I-g-1}\sum_{l=I-i}^{k-1}\frac{C(I-k-1,k+1)C(I-l-1,l+1)}{T(I-k-1,k+1)T(I-k-1,k)T(I-l-1,l+1)T(I-l-1,l)}$$
$$=2\sum_{k=I-i}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}$$
$$\text{x}\left[\frac{1}{T(I-k-1,k)}+\sum_{l=I-i}^{k-1}\frac{C(I-l-1,l+1)}{T(I-l-1,l+1)T(I-l-1,l)}\right] \qquad \text{(A.11)}$$

Note that, by (2.2), the second member within the square bracket may be expanded as follows:

$$\frac{C(I-l-1,l+1)}{T(I-l-1,l+1)T(I-l-1,l)}=\frac{1}{T(I-l-1,l)}-\frac{1}{T(I-l-1,l+1)} \qquad \text{(A.12)}$$

Substitute (A.12) into the square bracket in (A.11) to obtain

$$\frac{1}{T(I-k-1,k)}+\sum_{l=I-i}^{k-1}\frac{1}{T(I-l-1,l)}-\sum_{l=I-i+1}^{k}\frac{1}{T(I-l,l)}$$
$$=\frac{1}{T(i-1,I-i)}+\sum_{l=I-i+1}^{k}\left[\frac{1}{T(I-l-1,l)}-\frac{1}{T(I-l,l)}\right]$$
$$=\frac{1}{T(i-1,I-i)}+\sum_{l=I-i+1}^{k}\frac{R(I-l,l)}{T(I-l-1,l)T(I-l,l)}, \qquad \text{(A.13)}$$

by (2.1).

Substitute (A.13) for the square bracket in (A.11) to obtain

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)}=2\sum_{k=I-i}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)(I-k-1,k)}\text{x}$$
$$\left[\frac{1}{T(i-1,I-i)}+\sum_{l=I-i+1}^{k}\frac{R(I-l,l)}{T(I-l-1,l)T(I-l,l)}\right]. \qquad \text{(A.14)}$$

This proves (6.3).

Case II: $h > I - i$.

To evaluate the derivatives in (A.7) note that, according to (3.1), the $\hat{v}(k)$ involve $X(g,h)$ as follows:

- in the numerator, but not the denominator, for $k = h - 1$
- in the denominator, but not the numerator, for $k > h - 1,\ g \le I - k - 1$.

Note that the restriction $g \le I - k - 1$ is not required in the case $k = h - 1$. This is because in this case $I - k - 1 = I - h \ge g$ for $(g,h) \in D_i$, so the restriction is automatically satisfied.

Hence, (3.1) yields:

$$\frac{\partial \hat{v}(h-1)}{\partial X(g,h)} = \frac{1}{T(I-h,h-1)} \tag{A.15}$$

$$\frac{\partial^2 \hat{v}(h-1)}{\partial X^2(g,h)} = 0 \tag{A.16}$$

$$\frac{\partial \hat{v}(k)}{\partial X(g,h)} = -\frac{C(I-k-1,k+1)}{T^2(I-k-1,k)}, k \ge h,\ g \le I-k-1 \tag{A.17}$$

$$\frac{\partial^2 \hat{v}(k)}{\partial X^2(g,h)} = \frac{2C(I-k-1,k+1)}{T^3(I-k-1,k)}, k \ge h,\ g \le I-k-1. \tag{A.18}$$

Substitute (A.15) – (A.18) into (A.7) to obtain

$$\begin{aligned}
\frac{1}{Y}\frac{\partial^2 Y}{\partial^2 X(g,h)} &= \frac{1}{\hat{v}(h-1)}\frac{\partial^2 \hat{v}(h-1)}{\partial X^2(g,h)} + \sum_{k=h}^{I-g-1} \frac{1}{\hat{v}(k)}\frac{\partial^2 \hat{v}(k)}{\partial X^2(g,h)} \\
&+ \sum_{\substack{k,l=h \\ k\ne l}}^{I-g-1} \frac{1}{\hat{v}(k)\hat{v}(l)}\frac{\partial \hat{v}(k)}{\partial X(g,h)}\frac{\partial \hat{v}(l)}{\partial X(g,h)} \\
&+ \frac{2}{\hat{v}(h-1)}\frac{\partial \hat{v}(h-1)}{\partial X(g,h)}\sum_{k=h}^{I-g-1}\frac{1}{\hat{v}(k)}\frac{\partial \hat{v}(k)}{\partial X(g,h)} \\
&= 2\sum_{k=h}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T^2(I-k-1,k)} \\
&+ 2\sum_{k=h}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}\sum_{l=h}^{k-1}\frac{C(I-l-1,l+1)}{T(I-l-1,l+1)T(I-l-1,l)} \\
&- \frac{2}{T(I-h,h)}\sum_{k=h}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}.
\end{aligned} \tag{A.19}$$

Now apply (A.12) and use the same mode of calculation as led from (A.11) to (A.14). Then (A.19) becomes:

$$\frac{1}{Y}\frac{\partial^2 Y}{\partial X^2(g,h)} = 2\sum_{k=h}^{I-g-1}\frac{C(I-k-1,k+1)}{T(I-k-1,k+1)T(I-k-1,k)}\sum_{l=h}^{k}\frac{R(I-l,l)}{T(I-l-1,l)T(I-l,l)}. \quad \text{(A.20)}$$

This proves (6.4). □

Proof of Theorem 5. Consider Y defined by (6.1). By Theorem 4,

$\partial^2 Y / \partial X^2(g,h) \geq 0$ for all (g,h)

with strict inequality for some (g,h), namely those in D_i. It follows from a multivariate form of Jensen's inequality (see Lemma 2) that

$$E[Y] > \bar{Y} = \prod_{k=I-i}^{I-1} \bar{v}(k) \quad \text{(A.21)}$$

where $\bar{Y}$ is the value obtained by replacing each $X(g,h)$ in Y by its expectation, and $\bar{v}(k)$ is similarly defined.

By (3.1),

$$\begin{aligned}\bar{v}(k) &= E[T(I-k-1,k+1)]/E[T(I-k-1,k)] \\ &= \sum_{g=0}^{I-k-1} E[R(g,k+1)] / \sum_{g=0}^{I-k-1} E[R(g,k)] \qquad [\text{by}(2.3)] \\ &= \eta(k), \text{ by } (4.5).\end{aligned} \quad \text{(A.22)}$$

Substitute (A.22) in (A.21):

$$E[Y] > \prod_{k=I-i}^{I-1} \eta(k) = E[R(i,I)]/E[R(i,I-i)] \quad \text{(A.23)}$$

by (4.5).

Now take expectations on both sides of (3.2):

$$E[\hat{R}(i,I)] = E[R(i,I-i)]E[Y] > E[R(i,I)], \text{ by (A.23)}. \quad \text{(A.24)}$$

The first step leading to (A.24) is justified by the fact that the $X(g,h)$ involved in Y are those in the set D_i (see the start of the proof of Theorem 4), and this excludes row i of the array X (see (4.6)) on which $R(i,I-i)$ depends. Thus, $R(i,I-i)$ and Y are stochastically independent. □

References

Doray, L.G. (1996). UMVUE of the IBNR reserve in a lognormal linear regression model. **Insurance: mathematics and economics**, 18, 43-57.

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.

Kremer, E. (1982). IBNR claims and the two-way model of ANOVA. **Scandinavian Actuarial Journal**, 47-55.

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.

Renshaw, A.E. and Verrall, R.J. (1994). The stochastic model underlying the chain-ladder technique. Actuarial Research Paper No 63, City University, London.

Schmidt, K.D. and Wünsche A. (1998). Chain ladder, marginal sum and maximum likelihood estimation. **Blätter der Deutsche Gesellshaft Versicherungs-mathematiker**, 23, 267-277.

Stanard, J.N. (1985). A simulation test of prediction errors of loss reserve estimation techniques. **Proceedings of the Casualty Actuarial Society**, 72, 124-148.

Taylor, G.C. (1986). **Loss reserving in non-life insurance**. North-Holland, Amsterdam, Netherlands.

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). On the estimation of reserves from loglinear models. **Insurance: mathematics and economics**, 10, 75-80.

RESEARCH PAPER SERIES

No.	Date	Subject	Author
1	MAR 1993	AUSTRALIAN SUPERANNUATION: THE FACTS, THE FICTION, THE FUTURE	David M Knox
2	APR 1993	AN EXPONENTIAL BOUND FOR RUIN PROBABILITIES	David C M Dickson
3	APR 1993	SOME COMMENTS ON THE COMPOUND BINOMIAL MODEL	David C M Dickson
4	AUG 1993	RUIN PROBLEMS AND DUAL EVENTS	David C M Dickson Alfredo D Egídio dos Reis
5	SEP 1993	CONTEMPORARY ISSUES IN AUSTRALIAN SUPERANNUATION – A CONFERENCE SUMMARY	David M Knox John Piggott
6	SEP 1993	AN ANALYSIS OF THE EQUITY INVESTMENTS OF AUSTRALIAN SUPERANNUATION FUNDS	David M Knox
7	OCT 1993	A CRITIQUE OF DEFINED CONTRIBUTION USING A SIMULATION APPROACH	David M Knox
8	JAN 1994	REINSURANCE AND RUIN	David C M Dickson Howard R Waters
9	MAR 1994	LIFETIME INSURANCE, TAXATION, EXPENDITURE AND SUPERANNUATION (LITES): A LIFE-CYCLE SIMULATION MODEL	Margaret E Atkinson John Creedy David M Knox
10	FEB 1994	SUPERANNUATION FUNDS AND THE PROVISION OF DEVELOPMENT/VENTURE CAPITAL: THE PERFECT MATCH? YES OR NO	David M Knox
11	JUNE 1994	RUIN PROBLEMS: SIMULATION OR CALCULATION?	David C M Dickson Howard R Waters
12	JUNE 1994	THE RELATIONSHIP BETWEEN THE AGE PENSION AND SUPERANNUATION BENEFITS, PARTICULARLY FOR WOMEN	David M Knox
13	JUNE 1994	THE COST AND EQUITY IMPLICATIONS OF THE INSTITUTE OF ACTUARIES OF AUSTRALIA PROPOSED RETIREMENT INCOMES SRATEGY	Margaret E Atkinson John Creedy David M Knox Chris Haberecht
14	SEPT 1994	PROBLEMS AND PROSPECTS FOR THE LIFE INSURANCE AND PENSIONS SECTOR IN INDONESIA	Catherine Prime David M Knox

No.	Date	Subject	Author
15	OCT 1994	PRESENT PROBLEMS AND PROSPECTIVE PRESSURES IN AUSTRALIA'S SUPERANNUATION SYSTEM	David M Knox
16	DEC 1994	PLANNING RETIREMENT INCOME IN AUSTRALIA: ROUTES THROUGH THE MAZE	Margaret E Atkinson John Creedy David M Knox
17	JAN 1995	ON THE DISTRIBUTION OF THE DURATION OF NEGATIVE SURPLUS	David C M Dickson Alfredo D Egídio dos Reis
18	FEB 1995	OUTSTANDING CLAIM LIABILITIES: ARE THEY PREDICTABLE?	Ben Zehnwirth
19	MAY 1995	SOME STABLE ALGORITHMS IN RUIN THEORY AND THEIR APPLICATIONS	David C M Dickson Alfredo D Egídio dos Reis Howard R Waters
20	JUNE 1995	SOME FINANCIAL CONSEQUENCES OF THE SIZE OF AUSTRALIA'S SUPERANNUATION INDUSTRY IN THE NEXT THREE DECADES	David M Knox
21	JUNE 1995	MODELLING OPTIMAL RETIREMENT IN DECISIONS IN AUSTRALIA	Margaret E Atkinson John Creedy
22	JUNE 1995	AN EQUITY ANALYSIS OF SOME RADICAL SUGGESTIONS FOR AUSTRALIA'S RETIREMENT INCOME SYSTEM	Margaret E Atkinson John Creedy David M Knox
23	SEP 1995	EARLY RETIREMENT AND THE OPTIMAL RETIREMENT AGE	Angela Ryan
24	OCT 1995	APPROXIMATE CALCULATIONS OF MOMENTS OF RUIN RELATED DISTRIBUTIONS	David C M Dickson
25	DEC 1995	CONTEMPORARY ISSUES IN THE ONGOING REFORM OF THE AUSTRALIAN RETIREMENT INCOME SYSTEM	David M Knox
26	FEB 1996	THE CHOICE OF EARLY RETIREMENT AGE AND THE AUSTRALIAN SUPERANNUATION SYSTEM	Margaret E Atkinson John Creedy
27	FEB 1996	PREDICTIVE AGGREGATE CLAIMS DISTRIBUTIONS	David C M Dickson Ben Zehnwirth
28	FEB 1996	THE AUSTRALIAN GOVERNMENT SUPERANNUATION CO-CONTRIBUTIONS: ANALYSIS AND COMPARISON	Margaret E Atkinson
29	MAR 1996	A SURVEY OF VALUATION ASSUMPTIONS AND FUNDING METHODS USED BY AUSTRALIAN ACTUARIES IN DEFINED BENEFIT SUPERANNUATION FUND VALUATIONS	Des Welch Shauna Ferris

No.	Date	Subject	Author
30	MAR 1996	THE EFFECT OF INTEREST ON NEGATIVE SURPLUS	David C M Dickson Alfredo D Egídio dos Reis
31	MAR 1996	RESERVING CONSECUTIVE LAYERS OF INWARDS EXCESS-OFF-LOSS REINSURANCE	Greg Taylor
32	AUG 1996	EFFECTIVE AND ETHICAL INSTITUTIONAL INVESTMENT	Anthony Asher
33	AUG 1996	STOCHASTIC INVESTMENT MODELS: UNIT ROOTS, COINTEGRATION, STATE SPACE AND GARCH MODELS FOR AUSTRALIA	Michael Sherris Leanna Tedesco Ben Zehnwirth
34	AUG 1996	THREE POWERFUL DIAGNOSTIC MODELS FOR LOSS RESERVING	Ben Zehnwirth
35	SEPT 1996	KALMAN FILTERS WITH APPLICATIONS TO LOSS RESERVING	Ben Zehnwirth
36	OCT 1996	RELATIVE REINSURANCE RETENTION LEVELS	David C M Dickson Howard R Waters
37	OCT 1996	SMOOTHNESS CRITERIA FOR MULTI-DIMENSIONAL WHITTAKER GRADUATION	Greg Taylor
38	OCT 1996	GEOGRAPHIC PREMIUM RATING BY WHITTAKER SPATIAL SMOOTHING	Greg Taylor
39	OCT 1996	RISK, CAPITAL AND PROFIT IN INSURANCE	Greg Taylor
40	OCT 1996	SETTING A BONUS-MALUS SCALE IN THE PRESENCE OF OTHER RATING FACTORS	Greg Taylor
41	NOV 1996	CALCULATIONS AND DIAGNOSTICS FOR LINK RATION TECHNIQUES	Ben Zehnwirth Glen Barnett
42	DEC 1996	VIDEO-CONFERENCING IN ACTUARIAL STUDIES – A THREE YEAR CASE STUDY	David M Knox
43	DEC 1996	ALTERNATIVE RETIREMENT INCOME ARRANGEMENTS AND LIFETIME INCOME INEQUALITY: LESSONS FROM AUSTRALIA	Margaret E Atkinson John Creedy David M Knox
44	JAN 1997	AN ANALYSIS OF PENSIONER MORTALITY BY PRE-RETIREMENT INCOME	David M Knox Andrew Tomlin
45	JUL 1997	TECHNICAL ASPECTS OF DOMESTIC LINES PRICING	Greg Taylor
46	AUG 1997	RUIN PROBABILITIES WITH COMPOUNDING ASSETS	David C M Dickson Howard R Waters
47	NOV 1997	ON NUMERICAL EVALUATION OF FINITE TIME RUIN PROBABILITIES	David C M Dickson

No.	Date	Subject	Author
48	NOV 1997	ON THE MOMENTS OF RUIN AND RECOVERY TIMES	Alfredo G Egídio dos Reis
49	JAN 1998	A DECOMPOSITION OF ACTUARIAL SURPLUS AND APPLICATIONS	Daniel Dufresne
50	JAN 1998	PARTICIPATION PROFILES OF AUSTRALIAN WOMEN	M. E. Atkinson Roslyn Cornish
51	MAR 1998	PRICING THE STOCHASTIC VOLATILITY PUT OPTION OF BANKS' CREDIT LINE COMMITMENTS	J.P. Chateau Daniel Dufresne
52	MAR 1998	ON ROBUST ESTIMATION IN BÜHLMANN STRAUB'S CREDIBILITY MODEL	José Garrido Georgios Pitselis
53	MAR 1998	AN ANALYSIS OF THE EQUITY IMPLICATIONS OF RECENT TAXATION CHANGES TO AUSTRALIAN SUPERANNUATION	David M Knox M. E. Atkinson Susan Donath
54	APR 1998	TAX REFORM AND SUPERANNUATION – AN OPPORTUNITY TO BE GRASPED.	David M Knox
55	APR 1998	SUPER BENEFITS? ESTIMATES OF THE RETIREMENT INCOMES THAT AUSTRALIAN WOMEN WILL RECEIVE FROM SUPERANNUATION	Susan Donath
56	APR 1998	A UNIFIED APPROACH TO THE STUDY OF TAIL PROBABILITIES OF COMPOUND DISTRIBUTIONS	Jun Cai José Garrido
57	MAY 1998	THE DE PRIL TRANSFORM OF A COMPOUND R_k DISTRIBUTION	Bjørn Sundt Okechukwu Ekuma
58	MAY 1998	ON MULTIVARIATE PANJER RECURSIONS	Bjørn Sundt
59	MAY 1998	THE MULTIVARIATE DE PRIL TRANSFORM	Bjørn Sundt
60	JUNE 1998	ON ERROR BOUNDS FOR MULTIVARIATE DISTRIBUTIONS	Bjørn Sundt
61	JUNE 1998	THE EQUITY IMPLICATIONS OF CHANGING THE TAX BASIS FOR PENSION FUNDS	M E Atkinson John Creedy David Knox
62	JUNE 1998	ACCELERATED SIMULATION FOR PRICING ASIAN OPTIONS	Felisa J Vázquez-Abad Daniel Dufresne
63	JUNE 1998	AN AFFINE PROPERTY OF THE RECIPROCAL ASIAN OPTION PROCESS	Daniel Dufresne
64	AUG 1998	RUIN PROBLEMS FOR PHASE-TYPE(2) RISK PROCESSES	David C M Dickson Christian Hipp
65	AUG 1998	COMPARISON OF METHODS FOR EVALUATION OF THE n-FOLD CONVOLUTION OF AN ARITHMETIC DISTRIBUTION	Bjørn Sundt David C M Dickson

No.	Date	Subject	Author
66	NOV 1998	COMPARISON OF METHODS FOR EVALUATION OF THE CONVOLUTION OF TWO COMPOUND R_1 DISTRIBUTIONS	David C M Dickson Bjørn Sundt
67	NOV 1998	PENSION FUNDING WITH MOVING AVERAGE RATES OF RETURN	Diane Bédard Daniel Dufresne
68	DEC 1998	MULTI-PERIOD AGGREGATE LOSS DISTRIBUTIONS FOR A LIFE PORTFOLIO	David C M Dickson Howard R Waters
69	FEB 1999	LAGUERRE SERIES FOR ASIAN AND OTHER OPTIONS	Daniel Dufresne
70	MAR 1999	THE DEVELOPMENT OF SOME CHARACTERISTICS FOR EQUITABLE NATIONAL RETIREMENT INCOME SYSTEMS	David Knox Roslyn Cornish
71	APR 1999	A PROPOSAL FOR INTEGRATING AUSTRALIA'S RETIREMENT INCOME POLICY	David Knox
72	NOV 1999	THE STATISTICAL DISTRIBUTION OF INCURRED LOSSES AND ITS EVOLUTION OVER TIME I: NON-PARAMETRIC MODELS	Greg Taylor
73	NOV 1999	THE STATISTICAL DISTRIBUTION OF INCURRED LOSSES AND ITS EVOLUTION OVER TIME II: PARAMETRIC MODELS	Greg Taylor
74	DEC 1999	ON THE VANDERMONDE MATRIX AND ITS ROLE IN MATHEMATICAL FINANCE	Ragnar Norberg
75	DEC 1999	A MARKOV CHAIN FINANCIAL MARKET	Ragnar Norberg
76	MAR 2000	STOCHASTIC PROCESSES: LEARNING THE LANGUAGE	A J G Cairns D C M Dickson A S Macdonald H R Waters M Willder
77	MAR 2000	ON THE TIME TO RUIN FOR ERLANG(2) RISK PROCESSES	David C M Dickson
78	JULY 2000	RISK AND DISCOUNTED LOSS RESERVES	Greg Taylor
79	JULY 2000	STOCHASTIC CONTROL OF FUNDING SYSTEMS	Greg Taylor
80	NOV 2000	MEASURING THE EFFECTS OF REINSURANCE BY THE ADJUSTMENT COEFFICIENT IN THE SPARRE ANDERSON MODEL	Maria de Lourdes Centeno
81	NOV 2000	THE STATISTICAL DISTRIBUTION OF INCURRED LOSSES AND ITS EVOLUTION OVER TIME III: DYNAMIC MODELS	Greg Taylor

No.	Date	Subject	Author
82	DEC 2000	OPTIMAL INVESTMENT FOR INVESTORS WITH STATE DEPENDENT INCOME, AND FOR INSURERS	Christian Hipp
83	DEC 2000	HEDGING IN INCOMPLETE MARKETS AND OPTIMAL CONTROL	Christian Hipp Michael Taksar
84	FEB 2001	DISCRETE TIME RISK MODELS UNDER STOCHASTIC FORCES OF INTEREST	Jun Cai
85	FEB 2001	MODERN LANDMARKS IN ACTUARIAL SCIENCE Inaugural Professorial Address	David C M Dickson
86	JUNE 2001	LUNDBERG INEQUALITIES FOR RENEWAL EQUATIONS	Gordon E Willmot Jun Cai X Sheldon Lin
87	SEPTEMBER 2001	VOLATILITY, BETA AND RETURN WAS THERE EVER A MEANINGFUL RELATIONSHIP?	Richard Fitzherbert
88	NOVEMBER 2001	EXPLICIT, FINITE TIME RUIN PROBABILITIES FOR DISCRETE, DEPENDENT CLAIMS	Zvetan G Ignatov Vladimir K Kaishev Rossen S Krachunov
89	NOVEMBER 2001	ON THE DISTRIBUTION OF THE DEFICIT AT RUIN WHEN CLAIMS ARE PHASE-TYPE	Steve Drekic David C M Dickson David A Stanford Gordon E Willmot
90	NOVEMBER 2001	THE INTEGRATED SQUARE-ROOT PROCESS	Daniel Dufresne
91	NOVEMBER 2001	ON THE EXPECTED DISCOUNTED PENALTY FUNCTION AT RUIN OF A SURPLUS PROCESS WITH INTEREST	Jun Cai David C M Dickson
92	JANUARY 2002	CHAIN LADDER BIAS	Greg Taylor
93	JANUARY 2002	FURTHER OBSERVATIONS ON CHAIN LADDER BIAS	Greg Taylor
94	JANUARY 2002	A GENERAL CLASS OF RISK MODELS	Daniel Dufresne
95	JANUARY 2002	THE DISTRIBUTION OF THE TIME TO RUIN IN THE CLASSICAL RISK MODEL	David C M Dickson Howard R Waters