CASUALTY ACTUARIAL SOCIETY FORUM

Spring 1994

Volume Two

Including Selected Papers from the
1994 Variability in Reserves Prize Program

CASUALTY ACTUARIAL SOCIETY
ORGANIZED 1914

CASUALTY ACTUARIAL SOCIETY

Date: March 1994

To: CAS Readership

Re: *The Forum*, 1994 Special Edition

This special edition of *The Forum* includes Stephen Philbrick's "Accounting for Risk Margins," a paper that was commissioned by the Committee on Reserves. Mr. Philbrick's paper is preceded by a brief introduction by Paul O'Connell, explaining the committee's charge in funding this paper.

The main body of this issue is devoted to ten papers submitted for the Committee on the Theory of Risk prize on how to measure the variability of loss reserves. Gary Venter gives an introduction and summary of the ten Theory of Risk papers. He ends his introduction by listing outstanding issues that merit further research.

As always, any submissions, question or comments may be directed to me, or anyone on the Committee on *The Forum*.

Very Truly Yours,

Joel Kleinman
Chairperson, The Committee on *The Forum*

The Casualty Actuarial Society *Forum*

The Casualty Actuarial Society *Forum* is a non-refereed journal printed by the Casualty Actuarial Society. The viewpoints express in it do not necessarily reflect the views of the Casualty Actuarial Society.

The *Forum* is edited by the committee for the Casualty Actuarial Society Forum. The committee for the *Forum* invites all members of the CAS to submit papers on topics of interest to the actuarial community. Articles need not be written by a member of the CAS, but should have content of interest to the CAS membership.

The *Forum* is printed on a periodic basis, based on the number of articles submitted. Its goal is to publish two editions during the calendar year.

Comments or questions about the *Forum* may be directed to the committee for the Casualty Actuarial Society Forum.

Stephen W. Philbrick's paper "Accounting for Risk Margins" was edited by the CAS Committee on Reserves.

The Theory of Risk call papers were edited by the CAS Committee on the Theory of Risk

VOLUME TWO OF TWO VOLUMES

Correction

The paper "A Quantification of Snader's Deductible Safety Factor," by John Rollins and Monty J. Washburn, which appeared in the 1994 edition of the *Forum, Including the 1994 Ratemaking Call Papers*, is a copyrighted paper. The copyright notice was inadvertently deleted in printing the *Forum*.

The copyright notice should have read as follows:

> Copyright 1993, National Council on Compensation Insurance. All rights reserved.
>
> The material contained in this paper represents the ideas of the authors and not necessarily those of the National Council on Compensation Insurance.

We regret the error, and apologize to all affected parties.

Spring 1994 CAS Forum

Table of Contents

4. Estimation of Reserves, and Variances of Reserves.

It has been shown that the chain ladder model can be considered as a two-way analysis of variance. This linear model, and other linear models, can be used effectively for analysing claims data and producing estimates of expected total outstanding claims for each year of business. The methods have in common the assumption that the data is lognormally distributed, and the linear models are therefore applied to the logged incremental claims rather than the raw incremental claims data. The problem therefore arises of reversing the log transformation to produce estimates on the original scale. It is this problem which is addressed in this section; in particular the unbiasedness of the estimates is considered. This problem was first addressed in Verrall(1991a), in which the following analysis was given.

4.1. Identically Distributed Data

Before considering the claims run-off triangle, consider n independently, identically distributed observations which are lognormally distributed.

i.e. $Z_1, \ldots, Z_n$ are independent

and $Z_i \sim$ lognormal.

Suppose also that $E(Z_i) = \theta$. (4.1)

The aim is to estimate θ and to find the mean square error (or variance, if the estimate is unbiased) of the estimate. One way of proceeding towards the estimation of θ is to take logs of the data and analyse the resulting sample using normal distribution theory. This is an approach which can be generalised to data which is not identically distributed and so is the most appropriate for claims data.

Let $Y_i = \log Z_i \qquad (i=1,\ldots,n).$ (4.2)

Since Z_i has a lognormal distribution, Y_i has a normal distribution.

Suppose $Y_i \sim N(\mu, \sigma^2)$.

Then $\theta = \exp(\mu + \frac{1}{2}\sigma^2)$. (4.3)

The maximum likelihood estimates of μ and σ^2 are

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{\mu})^2$$

and the maximum likelihood estimate of θ can be obtained by substituting $\hat{\mu}$ and $\hat{\sigma}^2$ into equation (4.3):

$$\hat{\theta} = \exp (\hat{\mu} + \tfrac{1}{2}\hat{\sigma}^2) \qquad (4.4)$$

Finney (1941) showed that the maximum likelihood estimate of θ is biased. In order to correct for the bias, Finney introduced the function $g_m (t)$, where

$$g_m (t) = \sum_{k=0}^{\infty} \frac{m^k (m + 2k)}{m (m + 2) \dots (m + 2k)} \frac{t^k}{k!} \qquad (4.5)$$

and m is the degrees of freedom associated with $\hat{\sigma}^2$. In this case $m = n-1$.

It can be shown that an unbiased estimate of θ is $\tilde{\theta}$ where

$$\tilde{\theta} = \exp (\hat{\mu}) g_m (\tfrac{1}{2}(1-\tfrac{1}{n}) s^2) \qquad (4.6)$$

and $s^2 = \frac{n}{n-1} \hat{\sigma}^2$ is an unbiased estimate of σ^2.

One advantage of the use of linear models is that standard errors of the parameter estimates can be produced. These can be used to find standard errors on the original unlogged scale. The variance of $\tilde{\theta}$ is τ^2, where

$$\tau^2 = E (\tilde{\theta}^2) - (E (\tilde{\theta}))^2.$$

An unbiased estimate of $E (\tilde{\theta}^2)$ is obviously $\tilde{\theta}^2$ (since the expectation of this is $E (\tilde{\theta}^2)$) and

$$(E (\tilde{\theta}))^2 \qquad = (\exp (\mu + \tfrac{1}{2} \sigma^2))^2$$

$$= \exp (2\mu + \sigma^2).$$

By analogy with the unbiased estimation of θ, an unbiased estimate of

$$\exp (2\mu + \sigma^2)$$

is $$\exp (2\bar{\mu}) g_m ((1-\tfrac{2}{n}) s^2).$$

Thus an unbiased estimate of τ^2 is

$$\tilde{\tau}^2 = \exp (2\bar{\mu}) \Big[(g_m (\tfrac{1}{2} (1-\tfrac{1}{n}) s^2))^2 - g_m ((1-\tfrac{2}{n}) s^2) \Big]. \quad (4.7)$$

For comparison purposes, the corresponding maximum likelihood estimates are also found. The maximum likelihood estimate of the variance of the maximum likelihood estimate of θ, $\hat{\theta}$, is

$$\exp (2\hat{\mu} + \tfrac{\hat{\sigma}^2}{n}) \Big[\exp (\tfrac{\hat{\sigma}^2}{n}) \Big[1 - \tfrac{2\hat{\sigma}^2}{n} \Big]^{-\frac{1}{2}(n-1)} - \Big[1 - \tfrac{\hat{\sigma}^2}{n} \Big]^{-(n-1)} \Big]. \quad (4.8)$$

4.2 Unbiased Estimation for Claims Runoff Triangles

A claims runoff triangle consisting of incremental claims (assumed positive) is now considered. It is assumed that the data have been adjusted for inflation and exposure. Z_{ij} is incremental claims in row i, column j.

$$\text{Let } \theta_{ij} = E (Z_{ij}). \quad (4.9)$$

Estimates of θ_{ij} are required along with standard errors of these estimates. In particular, estimates of $\{ \theta_{ij} : i=1,...,t ; j=t-i+2,...,t \}$ are required, as these are the estimates of the expected outstanding claims. The row totals of the estimates also have to be considered, as these are the estimates of the expected total outstanding claims for each year of business.

$\{ Z_{ij} : i=1,...,t;j=1,...,t-i+1 \}$ are assumed to be independently, lognormally distributed.

Let $Y_{ij} = \log Z_{ij}$. (4.10)

Then Y_{ij} are independently normally distributed.

Suppose that $\{ Y_{ij} : i=1,...,t; j=1,...,t-i+1 \}$ are modelled by

$$E(Y_{ij}) = \underline{X}_{ij}\,\underline{\beta} \quad (4.11)$$

$$Var(Y_{ij}) = \sigma^2 \quad (4.12)$$

where $\underline{X}_{ij}$ is a row vector of explanatory variables and $\underline{\beta}$ is a column vector of parameters, both of length p.

The linear model for the whole triangle is

$$E(\underline{Y}) = X\,\underline{\beta} \quad (4.13)$$

where X is an (nxp) matrix whose rows are $\underline{X}_{ij}$

and $\underline{Y}$ is the vector of observations.

n is the number of observations (for a triangular array $n=\frac{1}{2}t(t+1)$), and the errors are assumed to be independently, identically normally distributed.

The expected value of the lognormally distributed data, θ_{ij} , is related to the mean and variance of the normally distributed data by

$$\theta_{ij} = \exp(\underline{X}_{ij}\,\underline{\beta} + \tfrac{1}{2}\sigma^2) \quad (4.14)$$

Thus the maximum likelihood estimate of θ_{ij} is

$$\hat{\theta}_{ij} = \exp(\underline{X}_{ij}\,\hat{\underline{\beta}} + \tfrac{1}{2}\hat{\sigma}^2) \quad (4.15)$$

where $\hat{\underline{\beta}} = (X'X)^{-1}X'\underline{y}$

and $\hat{\sigma}^2 = \frac{1}{n}(\underline{y} - X\hat{\underline{\beta}})'(\underline{y} - X\hat{\underline{\beta}})$

The general theory of estimation from linear models when the data is lognormally distributed was considered by Bradu and Mundlak (1970). It can be shown that an unbiased estimate of

$$\exp (\underline{Z} \, \underline{\beta} + a\sigma^2)$$

for any row vector $\underline{Z}$ of length p and scalar a, is

$$\exp (\underline{Z} \, \hat{\underline{\beta}}) \, g_m \Big[(a - \tfrac{1}{2} \underline{Z} \, (X'X)^{-1} \underline{Z}') \, s^2 \Big] \tag{4.16}$$

where s^2 is an unbiased estimate of σ^2 and m is the number of degrees of freedom associated with s^2.

i.e. $s^2 = \frac{n}{n-p} \hat{\sigma}^2$

$$= \frac{1}{n-p} (\underline{y} - X\hat{\underline{\beta}})'(\underline{y} - X\hat{\underline{\beta}})$$

and $m = n - p$.

Thus an unbiased estimate of θ_{ij} is $\tilde{\theta}_{ij}$, where

$$\tilde{\theta}_{ij} = \exp (\underline{X}_{ij} \, \hat{\underline{\beta}}) \, g_m \Big[\tfrac{1}{2} (1 - \underline{X}_{ij} \, (X'X)^{-1} \underline{X}'_{ij}) \, s^2 \Big]. \tag{4.17}$$

Note that $\mathrm{Var} (\hat{\underline{\beta}}) = (X'X)^{-1} \sigma^2$

and hence $\mathrm{Var} (X_{ij} \, \hat{\underline{\beta}}) = \underline{X}_{ij} \, (X'X)^{-1} \underline{X}'_{ij} \, \sigma^2$.

It can therefore be seen that $\underline{X}_{ij} \, (X'X)^{-1} \underline{X}'_{ij} \, s^2$ is an estimate of $\mathrm{Var} (X_{ij} \, \hat{\underline{\beta}})$.

The variance of the unbiased estimate of θ_{ij}, $\tilde{\theta}_{ij}$, is τ^2_{ij}, where

$$\tau^2_{ij} = \mathrm{Var} (\tilde{\theta}_{ij})$$

$$= E (\tilde{\theta}^2_{ij}) - (E (\tilde{\theta}_{ij}))^2. \tag{4.18}$$

An unbiased estimate of E ($\tilde{\theta}_{ij}^2$) is $\tilde{\theta}_{ij}^2$ and

$$(E (\tilde{\theta}_{ij}))^2 = \theta_{ij}^2$$

$$= \exp (2\underline{X}_{ij} \underline{\beta} + \sigma^2).$$

Hence an unbiased estimate of τ_{ij}^2 is $\tilde{\tau}_{ij}^2$, where

$$\tilde{\tau}_{ij}^2 = \exp(2\underline{X}_{ij}\hat{\underline{\beta}})\Big[(g_m (\tfrac{1}{2} (1-\underline{X}_{ij} (X'X)^{-1} \underline{X}_{ij}')s^2))^2 - g_m((1-2\underline{X}_{ij} (X'X)^{-1} \underline{X}_{ij}')s^2)\Big]. \quad (4.19)$$

4.3. Unbiased Estimates of Total Outstanding Claims

The purpose of the analysis of the claims data is to produce estimates of the expected total outstanding claims, R_i, for each year of business, and the total outstanding claims, R, for the whole triangle.

An unbiased estimate of R_i is $\tilde{R}_i$, where

$$\tilde{R}_i = \sum_{j=t-i+2}^{t} \tilde{\theta}_{ij} . \quad (4.20)$$

The variance of $\tilde{R}_i$ can be calculated as follows:

$$\text{Var} (\tilde{R}_i) = \text{Var}\left[\sum_{j=t-i+2}^{t} \tilde{\theta}_{ij} \right]$$

$$= \sum_{j=t-i+2}^{t} \left[\text{Var} (\tilde{\theta}_{ij}) + 2 \sum_{k=j+1}^{t} \text{Cov} (\tilde{\theta}_{ij} , \tilde{\theta}_{ik}) \right] \quad (4.21)$$

Now

$$\text{Cov} (\tilde{\theta}_{ij} , \tilde{\theta}_{ik}) = E (\tilde{\theta}_{ij} \tilde{\theta}_{ik}) - E (\tilde{\theta}_{ij}) E (\tilde{\theta}_{ik}) ,$$

and an unbiased estimate of this can be obtained using the same method as that which was used to find $\tilde{\tau}_{ij}^2$ in section 4.2.

It can be shown that if

$$\tau_{ijk} \quad = \operatorname{Cov}(\tilde{\theta}_{ij}, \tilde{\theta}_{ik}),$$

an unbiased estimate of τ_{ijk} is $\tilde{\tau}_{ijk}$, where

$$\tilde{\tau}_{ijk} = \exp((\underline{X}_{ij} + \underline{X}_{ik})\hat{\underline{\beta}}) \Big[g_m(\tfrac{1}{2}(1 - \underline{X}_{ij}(X'X)^{-1}\underline{X}'_{ij})s^2) g_m(\tfrac{1}{2}(1 - \underline{X}_{ik}(X'X)^{-1}\underline{X}'_{ik})s^2) - g_m((1 - \tfrac{1}{2}(\underline{X}_{ij} + \underline{X}_{ik})(X'X)^{-1}(\underline{X}_{ij} + \underline{X}_{ik}))s^2) \Big] \tag{4.22}$$

Hence an unbiased estimate of Var ($\tilde{R}_i$) is

$$\sum_{j=t-i+2}^{t} \left[\tilde{\tau}_{ij}^2 + 2 \sum_{k=j+1}^{t} \tilde{\tau}_{ijk} \right] \tag{4.23}$$

By extending the limits of the summations, the total outstanding claims for the whole triangle can also be considered.

4.4 Prediction Intervals

Having found an unbiased estimate of total outstanding claims, it is now possible to produce a prediction interval for total outstanding claims. The purpose of the analysis so far has been to produce an estimate of total outstanding claims and an estimate of the variance of this estimate. It is often desirable to find a 'safe' value which is unlikely to be exceeded by the total actual claims.

Let $\quad$ R = total outstanding claims for the whole triangle

and $\quad \tilde{R}$ be an unbiased estimate of E(R).

Suppose that a 95% upper confidence bound on R is required. i.e. it is required to find a value, k, such that

$$P(R \leq \tilde{R} + k) = 0.95 \tag{4.24}$$

i.e. find k such that

$$P(R - \tilde{R} \leq k) = 0.95. \tag{4.25}$$

Since $\tilde{R}$ is an unbiased estimate of $E(R)$,

$$E(\tilde{R}) = E(R) \tag{4.26}$$

and hence

$$E(R - \tilde{R}) = 0. \tag{4.27}$$

Also, $\tilde{R}$ is based on past data and is thus independent of R under the assumptions of the model. Thus

$$\mathrm{Var}(R - \tilde{R}) = \mathrm{Var}(R) + \mathrm{Var}(\tilde{R}). \tag{4.28}$$

In section 4.3, an unbiased estimate of $\mathrm{Var}(\tilde{R})$ was derived and it is possible to derive an unbiased estimate of $\mathrm{Var}(R)$ using the theory which was used in that section. By independence,

$$\mathrm{Var}(R) = \sum_{i=2}^{t} \sum_{j=t-i+2}^{t} \mathrm{Var}(Z_{ij}) \tag{4.29}$$

and an unbiased estimate of $\mathrm{Var}(Z_{ij})$ is required. This can be derived as follows, using the method of section 4.2.

Z_{ij} has a lognormal distribution, and the variance of this distribution is given by:

$$\mathrm{Var}(Z_{ij}) = \exp(2\underline{X}_{ij}\underline{\beta} + \sigma^2)(\exp(\sigma^2) - 1)$$

$$= \exp(2\underline{X}_{ij}\underline{\beta} + 2\sigma^2) - \exp(2\underline{X}_{ij}\underline{\beta} + \sigma^2) \tag{4.30}$$

Hence, applying equation (4.16), an unbiased estimate of $\mathrm{Var}(Z_{ij})$ is

$$\exp(2\underline{X}_{ij}\hat{\underline{\beta}})\left[g_m\left(2(1-\underline{X}_{ij}(X'X)^{-1}\underline{X}'_{ij})s^2\right) - g_m\left((1-2\underline{X}_{ij}(X'X)^{-1}\underline{X}'_{ij})s^2\right)\right]. \tag{4.31}$$

It is not inappropriate to use a Normal approximation since R and $\tilde{R}$ are, typically, combinations of a reasonably large number of lognormally distributed random variables. Thus a 95% upper bound on total outstanding claims can be found:

$$\tilde{R} + 1.645 \sqrt{Var(R) + Var(\tilde{R})} \qquad (4.32)$$

$\sqrt{Var(R) + Var(\tilde{R})}$ is the root mean square error of prediction.

4.5 Bayesian Estimation for Claims Runoff Triangles

When a method is used which is based on Bayes theory, Bayesian estimators should be used. The Bayesian estimators have a slghtly simpler form than the unbiased estimators and so are sometimes used in their place in a classical analysis. When used in a classical analysis, no prior information is assumed.

Suppose that Z_{kl} has a lognormal distribution with parameters θ and σ, and that the posterior distribution of θ, given D, is normal with mean m and variance τ^2.

i.e. $$\log Z_{kl} \mid \theta \sim N(\theta, \sigma^2)$$

$$\theta \mid D \sim N(m, \tau^2)$$

Suppose also that σ^2 and τ^2 are known. Then

$$E(Z_{kl} \mid D) = e^{m+\frac{1}{2}\sigma^2+\frac{1}{2}\tau^2}$$

and $$Var(Z_{kl} \mid D) = e^{2m+\sigma^2+\tau^2}(e^{\sigma^2+\tau^2} - 1)$$

Similar methods can be used to calculate the covariances, total outstanding claims and the variance of the total outstanding claims.

The Bayes estimate of outstanding claims for year of business i is

$$\sum_{j>n-i+1} E(Z_{ij} \mid D) \qquad (4.33)$$

and the Bayes estimate of the variance is

$$\sum_{j>n-i+1} \left[Var(Z_{ij} \mid D) + 2\sum_{k>j} Cov(Z_{ij}, Z_{ik} \mid D) \right] \qquad (4.33)$$

4.6. Example

This example illustrates and compares the two most basic methods of claims reserving considered in this thesis: the chain ladder method and the two-way analysis of variance. This gives an opportunity to compare the two. For the analysis of variance model, both the unbiased and maximum likelihood estimates of outstanding claims are given. The data used is taken from Taylor and Ashe (1983), and was given in section 2.

The estimates of the parameters in the chain ladder linear model and their standard errors are shown in table 4.1.

Table 4.1

	Estimate	Standard error
Overall mean	6.106	0.165
Row parameters	0.194	0.161
	0.149	0.168
	0.153	0.176
	0.299	0.186
	0.412	0.198
	0.508	0.214
	0.673	0.239
	0.495	0.281
	0.602	0.379
Column parameters	0.911	0.161
	0.939	0.168
	0.965	0.176
	0.383	0.186
	-0.005	0.198
	-0.118	0.214
	-0.439	0.239
	-0.054	0.281
	-1.393	0.379

The standard errors are obtained from the estimates of the variance-covariance matrix of the parameter estimates:

$$(X'X)^{-1} \hat{\sigma}^2$$

where $\hat{\sigma}^2$ is the estimate of the residual variance. For this example, $\hat{\sigma}^2 = 0.116$.

Since the data is in the form of a triangle (there are the same number of rows and columns) and the matrix X is based solely on the design, the standard errors are the same for each row and column parameter. The row parameters are contained within a much smaller range than the column parameters: (0.149, 0.673) compared with (-1.393, 0.965). It can also be seen that there is an indication that the row parameters follow an increasing trend. It is to be expected that the row parameters should be contained within a fairly small range, since the rows are expected to be similar. Any pattern in the row parameters gives an insight into, and depends upon, the particular claims experience. It is thus quite common to observe that the row parameters lie in a small range, but not typical that they follow a trend.

Table 4.2

286170	711785	731359	750301	418911	283724	252756	182559	266237	67948
357848	766940	610542	482940	527326	574398	146342	139950	227229	67948
410587	1021245	1049329	1076506	601040	407078	362646	261930	381987	
352118	884021	933894	1183289	445745	320996	527804	266172	425046	
379337	943516	969461	994572	555294	376094	335044	241994		
290507	1001799	926219	1016654	750816	146923	495992	280405		
339233	843767	866971	889425	496588	336334	299624			
310608	1108250	776189	1562400	272482	352053	206286			
378676	941872	967773	992840	554327	375439				
443160	693190	991983	769488	504851	470639				
389421	968599	995234	1021012	570056					
396132	937085	847498	805037	705960					
420963	1047052	1075844	1103710						
440832	847631	1131398	1063269						
457887	1138894	1170213							
359480	1061648	1443370							
396651	986582								
376686	986608								
344014									
344014									

The fitted values for the analysis of variance model are shown in table 4.2. These are the unbiased estimates and are shown with the actual observations for comparison. In this figure, and in all similar ones in future, the top entries are the estimates and those underneath are the actual observations.

Table 4.3 shows a plot of residuals (fitted value − actual value) against fitted value.

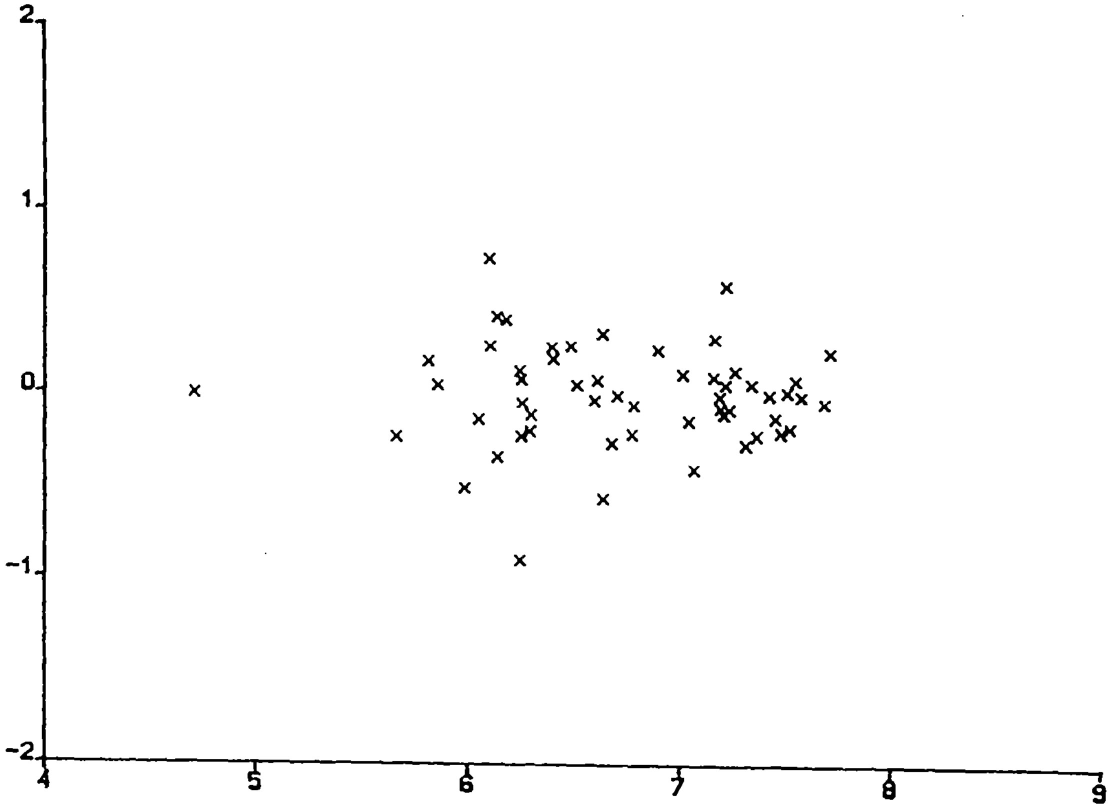


Table 4.3

There is no discernible pattern in the residuals, and they seem to be randomly scattered, so there is no cause to question the model on the basis of this plot. Of course, it is possible to look further into a residual analysis and study the plots of residuals against year of business and delay. This is considered further using the GLIM system, by Renshaw (1989). The main purpose of this paper is to extend the possible range of analyses and to consider rigorous statistical estimation, rather than to find the most appropriate linear model to fit, and so the residual analysis will not be commented on further.

Of most interest to practitioners are the predicted outstanding claims for each year of business, which are the row totals of predicted values. Table 4.4 shows the maximum likelihood predictions of the outstanding claims in the lower triangle, and table 4.5 shows the unbiased predictions. The method does not produce any predictions for the first row, and each row contains one more predicted value.

<u>Table 4.4</u>

								101269
							357398	93599
						217465	319835	83761
					335047	243001	357392	93597
				386433	345088	250283	368102	96402
			617309	418743	373941	271209	398880	104462
		1206369	674243	457364	408430	296223	435668	114097
	1026594	1053911	589034	399564	356813	258787	380610	99678
888831	913640	937951	524224	355600	317554	230313	338732	88710

Table 4.5

								96238
							350362	88841
						215218	313105	79394
					332848	240075	349268	88564
				384305	342028	246696	358900	91006
			613257	415031	369373	266419	387593	98281
		1193906	666126	450811	401216	289387	421005	106752
	1006382	1031734	575643	389575	346716	250077	363813	92248
844677	867203	889047	496032	335695	298762	215487	313486	79483

It can be seen that the maximum likelihood estimates are all higher than the unbiased estimates, as was to be expected.

Table 4.6

	Analysis of Variance		Chain Ladder
Row	Maximum Likelihood	Unbiased	
2	101269	96238	94630
3	450997	439203	464668
4	621061	607717	702101
5	1029037	1010755	965576
6	1446307	1422934	1412202
7	2184544	2149953	2176089
8	3592393	3529202	3897142
9	4164990	4056189	4289473
10	4595556	4339873	4618035

The total predicted outstanding claims for each year of business (the row totals of the predicted outstanding claims) are shown in table 4.6. There are three estimates given, the maximum likelihood and unbiased estimates from the analysis of variance model, and the chain ladder estimate.

It can be seen that the maximum likelihood estimates differ most significantly from the unbiased which is where the number of observations used in the estimation is the greatest. The maximum likelihood estimate is asymptotically unbiased, and the greater the number of observations used to estimate the parameters, the closer are the two. The chain ladder estimates are sometimes higher and sometimes lower than the analysis of variance estimates. There is nothing significant that can be inferred from the differences. This confirms that the crude chain ladder method is a reasonable 'rough-and-ready' method for calculating outstanding claims, although the more proper method, statistically, is the analysis of variance method (using unbiased estimation).

The total predicted outstanding claims are:

Analysis of Variance	Maximum Likelihood	18186154
	Unbiased	17652064
Chain Ladder		18619916

The following table shows the unbiased estimates of the total outstanding claims for each year of business, the standard errors of these estimates and the root mean square error of prediction. This table can be used in setting safe reserves, and gives an idea of the likely variation of outstanding claims.

Table 4.7

Unbiased Estimate	Standard Error	Mean Square Error Of Prediction
96238	35105	47202
439203	108804	163217
607717	127616	182847
1010755	195739	269224
1422934	273082	357593
2149953	429669	538533
3529202	775256	942851
4056189	1052049	1197009
4339873	1534943	1631306

The unbiased estimate of total outstanding claims is 17652064 and the root mean square error of prediction is 2759258. Thus a 95% upper bound on total outstanding claims is

$$17652064 + 1.645 \times 2759258 = 22191043$$

This could be regarded as a "safe" reserve for this triangle according to the chain ladder linear model using unbiased estimation.

5. Estimation of the Development Factors

When considering outstanding claims, it is important to use unbiased estimators. However, when comparing several sets of runoff patterns it is simpler to use maximum likelihood theory since unbiasedness is not critical. There are two sets of parameters whose distributions can usefully be found: the development factors, $\{ \lambda_j : j = 2 , \ldots , t \}$, and the proportions of ultimate claims, $\{ S_j : j = 1 , \ldots , t ; \sum_{j=1}^{t} S_j = 1 \}$. It has already been shown that the following relationship between the proportions of ultimate claims and the development factors holds:

$$S_1 = \frac{1}{\sum_{l=2}^{t} \lambda_l} \qquad (5.1)$$

$$S_j = \frac{\lambda_j - 1}{\sum_{l=j}^{t} \lambda_l} \qquad (j \geq 2) \qquad (5.2)$$

It was also shown by Kremer that the proportions of ultimate claims are related to the column parameters of the linear model as follows

$$S_j = \frac{e^{\beta_j}}{\sum_{l=1}^{t} e^{\beta_l}} \qquad j = 1 , \ldots , t \qquad (5.3)$$

where $\beta_1 = 0$ by definition.

Finally, the relationship between the parameters of the chain ladder and linear models was proved in Verrall (1991b):

$$\lambda_j = 1 + \frac{e^{\beta_j}}{\sum_{l=1}^{j-1} e^{\beta_l}} \tag{5.4}$$

The parameters of the additive model can be estimated using maximum likelihood estimation. The variance-covariance matrix of the parameter estimates can be obtained from the Fisher information matrix by differentiating the log-likelihood a second time. Further details of the theory of maximum likelihood which is used in this section can be found in Cox and Hinkley (1974).

Since maximum likelihood estimates are invariant under parameter transformations, the maximum likelihood estimates of the development factors and the proportions of ultimate claims can be obtained by substituting the estimates of $\{ \beta_j : j = 1 , \dots , t \; ; \beta_1 = 0 \}$ into equations (5.3) and (5.4). In addition to the parameter estimates, it is useful to have standard errors of the parameter estimates which can be obtained by maximum likelihood theory. The particular advantage of using maximum likelihood estimation is that the second moments are relatively straightforward to obtain. Denoting the variance-covariance matrix of $\{ \beta_j : j = 1 , \dots , t \; ; \beta_1 = 0 \}$ by $V(\underline{\beta})$, the variance-covariance matrix of $\{ \lambda_j : j = 2 , \dots , t \}$ and $\{ S_j : j = 1 , \dots , t \; ; \sum_{j=1}^{t} S_j = 1 \}$ are given by

$$V(\underline{\lambda}) = \left(\frac{\partial \lambda}{\partial \underline{\beta}}\right) V(\underline{\beta}) \left(\frac{\partial \lambda}{\partial \underline{\beta}}\right) \tag{5.5}$$

and $$V(\underline{S}) = \left(\frac{\partial \underline{S}}{\partial \underline{\beta}}\right) V(\underline{\beta}) \left(\frac{\partial \underline{S}}{\partial \underline{\beta}}\right) \tag{5.6}$$

It is thus necessary to obtain the matrices of the first derivatives of the respective parameter vectors. The (j,k)th element of $\left(\frac{\partial \lambda}{\partial \underline{\beta}}\right)$ can be obtained from equation (5.4) and is given by:

$$\frac{\partial \lambda_j}{\partial \beta_k} = \begin{cases} 0 & k > j \\ \dfrac{e^{\beta_j}}{\sum_{l=1}^{j-1} e^{\beta_l}} & k = j \\ \dfrac{e^{\beta_j} e^{\beta_k}}{\left(\sum_{l=1}^{j-1} e^{\beta_l}\right)} & k < j \end{cases}$$

$$= \begin{cases} 0 & k > j \\ \lambda_j - 1 & k = j \\ -(\lambda_j - 1)(\lambda_k - 1) & k < j \end{cases} \tag{5.7}$$

Similarly, the (j,k)th element of $\left(\frac{\partial \underline{S}}{\partial \underline{\beta}}\right)$ can be obtained from equation (5.3) and is given by:

$$\frac{\partial S_j}{\partial \beta_k} = \begin{cases} -\dfrac{e^{\beta_j} e^{\beta_k}}{\left(\sum_{l=1}^{t} e^{\beta_l} \right)} & k \neq j \\ \\ \dfrac{e^{\beta_j} \left(\sum_{l=1}^{t} e^{\beta_l} - e^{\beta_j} \right)}{\left(\sum_{l=1}^{t} e^{\beta_l} \right)} & k = j \end{cases}$$

$$= \begin{cases} - S_j \, S_k & k \neq j \\ \\ S_j \, (1 - S_j) & k = j \end{cases} \qquad (5.8)$$

Estimates of the variance-covariance matrices can be obtained by substituting estimates of the parameters into equations (5.7) and (5.8).

A technical note is that the parameter β_1 (which is defined to be zero) has to be included in the matrix of partial derivatives in equation (5.8) since there are n parameters in the vector $\underset{\sim}{S}$. The variance-covariance matrix of the parameters of the additive model which is obtained from a standard least squares analysis has to augmented to include an extra row and column, all of whose entries are zero. This is not necessary for equation (5.7).

5.1 Example

The method described in section 5 is of use when comparing several different sets of data and therefore a different example will be used than in other sections for illustration purposes. The method is applied to six sets of employers' liability data which have been obtained from the DTI returns. The names of the companies to which the data apply have been suppressed, and it should be commented that this mathematical analysis is only one part of the process by which reserves are set. In particular, the DTI data are gross of reinsurance. The results here should therefore be regarded as a statistical analysis which would give further information to the claims reserver who would use the other information available.

We now consider the parameter estimates for each of the three models in turn. Beginning with the additive model the estimates of the column parameters { $\beta_j : j = 2 , \ldots , t$ } and their standard errors are given in the following figure:

COMPANY:

1		2		3		4		5		6	
1.796	0.121	1.748	0.148	2.236	0.249	1.846	0.248	1.941	0.201	2.010	0.082
1.848	0.126	1.857	0.155	2.080	0.261	2.260	0.260	2.248	0.211	2.246	0.086
1.669	0.133	1.654	0.163	1.978	0.273	2.159	0.272	2.204	0.221	2.129	0.091
1.413	0.139	1.400	0.171	1.725	0.287	1.986	0.286	1.981	0.232	1.863	0.095
0.994	0.147	1.200	0.180	1.535	0.303	1.535	0.302	1.514	0.245	1.485	0.100
0.615	0.155	0.705	0.190	1.057	0.320	1.235	0.319	0.788	0.259	1.050	0.106
0.415	0.164	0.339	0.201	0.667	0.338	0.644	0.337	0.227	0.274	0.782	0.112
0.038	0.175	0.025	0.215	-0.099	0.360	0.222	0.359	-0.540	0.291	0.234	0.120
-0.812	0.189	-0.407	0.232	-0.300	0.390	0.047	0.388	-0.993	0.315	0.155	0.129
-0.915	0.212	-1.821	0.260	-0.715	0.437	0.382	0.435	-1.311	0.353	-0.324	0.145
-2.513	0.264	-1.492	0.323	-1.708	0.543	-0.896	0.541	-3.206	0.439	-0.304	0.180

Before going on to the parameters which have a physical interpretation, it should be noticed that it is already possible to see some differences between the companies. In particucar, the standard errors of the parameters are larger for some companies (3 and 4) than for others (6). This will be mirrored in the parameter estimates and standard errors of the other models.

Next, consider the chain ladder model. The estimates of the development factors $\{\lambda_j : j = 2, \ldots, t\}$ and their standard errors are given in following table:

COMPANY:

1		2		3		4		5		6	
7.027	0.727	6.742	0.850	10.36	2.327	7.332	1.569	7.963	1.401	8.466	0.616
1.904	0.101	1.950	0.130	1.773	0.181	2.307	0.300	2.189	0.222	2.117	0.086
1.397	0.041	1.398	0.050	1.394	0.084	1.512	0.109	1.520	0.090	1.469	0.033
1.220	0.022	1.221	0.027	1.219	0.046	1.285	0.059	1.274	0.046	1.245	0.017
1.119	0.012	1.148	0.019	1.149	0.032	1.141	0.030	1.135	0.023	1.135	0.009
1.073	0.008	1.079	0.010	1.080	0.018	1.092	0.020	1.057	0.010	1.077	0.006
1.055	0.006	1.051	0.007	1.050	0.012	1.047	0.011	1.031	0.006	1.055	0.004
1.036	0.005	1.035	0.006	1.022	0.006	1.029	0.008	1.014	0.003	1.030	0.003
1.015	0.002	1.022	0.004	1.018	0.005	1.024	0.007	1.009	0.002	1.027	0.003
1.013	0.002	1.005	0.001	1.012	0.004	1.032	0.011	1.006	0.002	1.016	0.002
1.003	0.001	1.007	0.002	1.004	0.002	1.009	0.004	1.001	0.000	1.016	0.003

Finally, consider the multiplicative model. The estimates of the proportions of ultimate claims in each development year $\{ S_j : j = 1, \ldots, t ; \sum_{j=1}^{t} S_j = 1 \}$ and their standard errors are given in the following table:

COMPANY:

1		2		3		4		5		6	
0.032	0.003	0.032	0.004	0.023	0.005	0.021	0.005	0.023	0.004	0.022	0.002
0.196	0.016	0.184	0.019	0.218	0.036	0.135	0.025	0.162	0.023	0.162	0.010
0.206	0.015	0.205	0.019	0.186	0.029	0.204	0.032	0.220	0.027	0.205	0.011
0.172	0.013	0.167	0.015	0.168	0.026	0.184	0.028	0.211	0.025	0.182	0.009
0.133	0.011	0.130	0.013	0.131	0.021	0.155	0.024	0.169	0.022	0.140	0.007
0.088	0.008	0.106	0.011	0.108	0.019	0.099	0.017	0.106	0.015	0.096	0.006
0.060	0.006	0.065	0.007	0.067	0.013	0.073	0.014	0.051	0.008	0.062	0.004
0.049	0.005	0.045	0.006	0.045	0.010	0.040	0.009	0.029	0.005	0.047	0.003
0.034	0.004	0.033	0.005	0.021	0.005	0.027	0.007	0.014	0.003	0.027	0.002
0.014	0.002	0.021	0.004	0.017	0.005	0.022	0.006	0.009	0.002	0.025	0.002
0.013	0.002	0.005	0.001	0.011	0.004	0.031	0.011	0.006	0.002	0.016	0.002
0.003	0.001	0.007	0.002	0.004	0.002	0.009	0.004	0.001	0.000	0.016	0.002

The runoff patterns of the companies can be compared using the two tables above. For example, 1 and 2 seem quite similar, and some of the companies have more runoff in later development years than others. The standard errors can also be compared, with the same conclusions as above.

6. Bayesian Linear Models and Credibility Theory

Bayes estimates for the linear model were investigated by Lindley and Smith (1972) and also Smith (1973). In the actuarial literature, the recent paper by Klugman (1989) has studied the use of hierarchical linear models in a rating context. It has already been seen that many of the models commonly used to analyse claims runoff triangles can be regarded as linear models, and we now analyse these models from a Bayesian point of view. This analysis has two purposes: firstly the practitioner may have some information, from other data for example, which can be used to specify a prior distribution for the parameters in the model and secondly the Bayesian analysis gives rise in a natural way to estimators which have a credibility theory interpretation.

In the first case the prior distribution is set by the practitioner and the usual prior-posterior analysis can be carried out. The models which we are using assume normal (really log-normal) distributions, and so it is only necessary to specify the mean and variance of the prior distribution (which is also normal). For example, if there is a lot of evidence to suggest that the row parameters are all 0.1, a normal distribution with mean 0.1 and small variance can be used as prior. If there is not much prior information, the prior variance can be set larger. Indeed, in the limit, as the prior variance becomes large, we revert back to ordinary least-squares estimation of the parameters.

In the second case, we will be using empirical priors. Thus the estimation will be empirical Bayes and we will assume that certain of the parameters are exchangeable. The historical requirement that credibility estimators be linear will also be considered and we could claim to have credibility formulae. The situation has some similarities with credibility estimators of risk premiums in that we can regard the rows in a runoff triangle as a set of risks and proceed as Buhlmann (1967) - see Goovaerts and Hoogstad (1987) for a full description of Buhlmann's method. In the case of claims runoff triangles the rows contain different numbers of elements, and there are also the column parameters to contend with. This approach, starting from credibility premiums and working through to a credibility theory for loss runoff triangles was suggested by De Vylder (1982) - again see Goovaerts and Hoogstad (1987) for an exposition of the method. The present method starts from runoff triangles and proceeds to credibility formulae via the linear models. One of the major advantages of the linear model approach is that standard errors of the estimates are also produced.

For consistency, the constraints

$$\alpha_1 = \beta_1 = 0$$

on the first stage distribution have been retained. This also facilitates the comparison with the recursive approaches such as that based on the Kalman filter. It does, however, introduce a slight degree of àssymmetry into the prior distribution and it might be considered more appropriate to use a constraint such as

$$\sum \alpha_i = \sum \beta_j = 0.$$

It is also possible to apply the constraint at the second stage and use the following prior distribution:

$$\alpha_i \sim N(0, \sigma_\alpha^2) \qquad i = 1, \ldots, t.$$

The affect of the exchangeability assumption is similar whichever constraint is used: the estimates are shrunk towards a central value and stability is introduced. The amount of shrinkage is greatest where the number of observations is small.

6.1 Bayes Estimates for the Analysis of Variance Model

In this section the use of two-stage Bayesian linear models which assume that there is some prior information is described. A prior distribution can be written down using the prior knowledge. It was shown earlier that the chain ladder linear model can be written as a linear model in the form:

$$\underline{y} | \underline{\beta} \quad \sim \quad N(X\underline{\beta}, \Sigma)$$

and the prior information is quantified in the prior distribution on $\underline{\beta}$

$$\underline{\beta} | \underline{\theta}_l \quad \sim \quad N(A_l \underline{\theta}_l, C_l)$$

A situation which may occur is that there are similar sets of data available which give information on the individual parameters. In this case A_1 can be taken as an identity matrix, the prior estimates can be put into $\underline{\theta}_1$ and their variances into C_1. In many cases C_1 will be a diagonal matrix of variances, although it is not necessary that the covariances are zero. In this case, the prior distribution becomes:

$$\underline{\beta}|\underline{\theta}_1 \quad \sim \quad N(\underline{\theta}_1, C_1) \tag{6.1}$$

Assuming that the errors are independent, $\Sigma = \sigma^2 I_n$. I_n is a square identity matrix of dimension (nxn).

The Bayes estimate of the parameter vector is the solution, $\tilde{\underline{\beta}}$, of

$$(\sigma^{-2} X'X + C_1^{-1})\,\tilde{\underline{\beta}} = \sigma^{-2} X'X\,\hat{\underline{\beta}} + C_1^{-1}\,\underline{\theta}_1 \tag{6.2}$$

and the variance-covariance matrix of the estimates is

$$\mathrm{Var}(\tilde{\underline{\beta}}) = [\,\sigma^{-2} X'X + C_1^{-1}\,]^{-1} \tag{6.3}$$

The equation for $\tilde{\underline{\beta}}$ (3.4.2) can be written as a credibility formula:

$$\tilde{\underline{\beta}} = z\hat{\underline{\beta}} + (1 - z)\,\underline{\theta}_1 \tag{6.4}$$

where $z = (\sigma^{-2} X'X + C_1^{-1})^{-1} \sigma^{-2} X'X$ is the credibility factor.

It is interesting to note that the credibility factor has been generalised into a credibility matrix, since z is a (pxp) matrix. There will be a subtle dependence of the elements in the Bayes estimator $\tilde{\underline{\beta}}$ on each of the elements in the least squares estimator. It is not possible to write a credibility formula separately for each factor in the form

$$\tilde{\alpha}_i = z\,\hat{\alpha}_i + (1 - z)\theta_i$$

To estimate the variance σ^2, the modal procedure described is used. The estimate of σ^2 is s^2, where

$$s^2 = (\underline{y} - X\tilde{\underline{\beta}})'(\underline{y} - X\tilde{\underline{\beta}})/(n+2) \tag{6.5}$$

Thus the equations which give the Bayes estimates are (6.2), with σ^2 replaced by s^2, and (6.5).

The procedure begins with $s^2 = 0$ and iterates between the solutions of

$$(s^{-2}X'X + C_1^{-1})\tilde{\underline{\beta}} = s^{-2}X'X\hat{\underline{\beta}} + C_1^{-1}\underline{\theta}_1$$

and

$$s^2 = (\underline{y} - X\tilde{\underline{\beta}})'(\underline{y} - X\tilde{\underline{\beta}})/(n+2)$$

6.2 Empirical Bayes Estimates for the Chain Ladder Linear Model

The previous section described the use of a two-stage conventional Bayesian model to analyse claims data. This section uses a three-stage Bayesian model described in Verrall (1990) to derive empirical Bayes estimates for the chain ladder model. This method uses an improper prior distribution at the third stage for the row parameters and improper priors at the second stage for the overall mean and the column parameters. This means that for the overall mean and the column parameters the same assumptions are made as for the maximum likelihood estimators.

The row parameters are assumed to be independent samples from a common distribution - of course, they are unobservable, but this is the underlying assumption. A similar assumption is made in credibility theory. When premiums are calculated using credibility theory, a risk parameter is assigned to each risk and these are assumed to be independently, identically distributed. The set of risks is known as a collective, and the distribution from which the risk parameters is drawn is known as the structure of the collective. The situation in the claims reserving case is similar for the row parameters, but is complicated by the column parameters.

The estimators produced will combine information from each row with information from the triangle as a whole. The prior distribution (i.e. the second stage distribution) is estimated from the data, and hence the estimators have an empirical Bayes interpretation.

The linear model for the chain ladder method is

$$\underline{y}|\underline{\beta} \sim N(X\underline{\beta}, \sigma^2 I) \qquad (6.6)$$

and the constraint $\alpha_1 = \beta_1 = 0$ will be used.

The errors have been assumed to be independently, identically distributed. X is as defined in the first section.

As in credibility theory, a structure is put onto the row parameters $\alpha_2, \alpha_3, \ldots, \alpha_t$: they are assumed to be independent observations from a common distribution. For the overall mean, μ, and the column parameters $\beta_2, \beta_3, \ldots, \beta_t$, the same distributional assumptions as for ordinary maximum likelihood estimation will be used. Thus at the second stage

$$
\underline{\beta} \mid w, \psi, \underline{\xi} \sim N \left(\begin{bmatrix} 1 & 0 & \cdots & & 0 \\ 0 & 1 & 0 & & \vdots \\ \vdots & 1 & \vdots & & \\ \vdots & \vdots & \vdots & & \\ 0 & 1 & 0 & & \\ \vdots & 0 & 1 & & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} w \\ \psi \\ \xi_2 \\ \vdots \\ \xi_t \end{bmatrix}, \begin{bmatrix} \sigma^2_{\mu} & & & & & \\ & \sigma^2_{\alpha} & & & & \\ & & \ddots & & & \\ & & & \sigma^2_{\alpha} & & \\ & & & & \sigma^2_{\beta} & \\ & & & & & \ddots \\ & & & & & & \sigma^2_{\beta} \end{bmatrix} \right)
$$

and take $\sigma_{\mu}^{-2} \to 0$ and $\sigma_{\beta}^{-2} \to 0$. (6.7)

ψ is the mean of the common distribution of the row parameters $\alpha_2, \ldots, \alpha_t$.

Although the assumptions on the estimation of μ and $\beta_2, \ldots, \beta_t$ are the same as for the maximum likelihood estimation, the estimators produced will not be the same because of the treatment of the row parameters.

A vague prior distribution (a third-stage distribution) is used for ψ. Since $\sigma_{\mu}^{-2} \to 0$ and $\sigma_{\beta}^{-2} \to 0$, a third-stage distribution is not needed for w and $\xi_2, \ldots, \xi_t$. Hence a combination of two-stage and three-stage models is used.

The Bayes estimate of $\underline{\beta}$, $\tilde{\underline{\beta}}$, is given by

$$
\tilde{\underline{\beta}} = \left(\sigma^{-2} X'X + \begin{bmatrix} 0 & & & & \\ & \sigma_{\alpha}^{-2} & & & \\ & & \ddots & & \\ & & & \sigma_{\alpha}^{-2} & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \right)^{-1} \left(\sigma^{-2} X'X \hat{\underline{\beta}} + \begin{bmatrix} 0 & & & & \\ & \sigma_{\alpha}^{-2} & & & \\ & & \ddots & & \\ & & & \sigma_{\alpha}^{-2} & \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \tilde{\alpha}. \\ \vdots \\ \tilde{\alpha}. \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right)
$$

(6.8)

where $\tilde{\alpha}. = \frac{1}{t-1} \sum_{i=2}^{t} \tilde{\alpha}_i$

and has a credibility interpretation.

It can be seen that the empirical Bayes estimates of the row parameters are in the general form of credibility estimates: they are the weighted average of the maximum likelihood estimates and the (weighted) average of the estimates from all the rows. The situation is complicated by the fact that $X'X$ is not a diagonal or block-diagonal matrix, so that the estimation of μ, β_2, ... , β_t involves the estimates of α_2, ... ,α_t and vice versa. This is entirely natural since changing the estimates of the row parameters obviously forces changes in the other estimates. However, it can be seen that the form of the estimates is the same as the form of credibility estimates. They are the weighted average of the maximum likelihood estimates and the (weighted) average of the estimates to which the credibility theory type assumptions have been applied. The weights depend on the precision of the estimates.

As before, the variances σ^2 and σ_α^2 are replaced by modal estimates s^2 and s_α^2 , which are given by

$$s^2 \quad = \quad \frac{\nu\lambda+(\underline{y}-X\underline{\tilde{\beta}})'(\underline{y}-X\underline{\tilde{\beta}})}{n+\nu+2} \tag{6.9}$$

$$s_\alpha^2 \quad = \frac{\nu_\alpha\lambda_\alpha+\sum_{i=2}^{t}(\tilde{\alpha}_i-\tilde{\alpha}.)^2}{t+\nu_\alpha+1} \tag{6.10}$$

where ν, λ, ν_α and λ_α are set by the prior distribution of the variances. The derivation of these formulae, and the discusion of the prior parameter values is given in Lindley and Smith (1972).

Again, the estimates are obtained by iterating between (6.8) and (6.9),(6.10). This is illustrated in the example.

The empirical Bayes assumptions could also be applied to the column parameters, although this is of little practical use.

6.3. Example

To illustrate the effect of the assumptions made in the empirical Bayes theory, namely that the row parameters are independent observations from a common distribution, the Taylor and Ashe data is reanalysed in this example.

The estimates of the parameters and their standard errors are shown in table 6.1:

Table 6.1

	Empirical Bayes Estimate	No Prior Estimate	Standard Error Of Bayes Estimate
Overall Mean	6.157	6.106	0.131
Row Parameters	0.225	0.194	0.124
	0.193	0.149	0.129
	0.198	0.153	0.133
	0.300	0.299	0.138
	0.371	0.412	0.144
	0.421	0.508	0.150
	0.493	0.673	0.159
	0.383	0.495	0.170
	0.391	0.602	0.185
Column Parameters	0.893	0.911	0.128
	0.911	0.939	0.133
	0.915	0.965	0.139
	0.319	0.383	0.147
	-0.080	-0.005	0.156
	-0.199	-0.118	0.170
	-0.515	-0.439	0.190
	-0.120	-0.054	0.224
	-1.444	-1.393	0.306

The estimate of the variance of the row parameter distribution is 0.0289.

The empirical Bayes assumptions have been applied to the row parameters only. The effect of these assumptions is that the row parameters have been drawn towards a central point (a weighted average). The lower row parameter estimates have increased, while the higher ones have decreased. This can be seen more clearly from the graph given in section 7.3 which shows a plot of the maximum likelihood and empirical Bayes estimates of the row parameters, togther with the estimates from the dynamic model discussed in section 7.

Table 6.2 shows the row totals and their standard errors. For comparison purposes, the Bayes estimates with no prior assumptions are also given.

Table 6.2

Empirical Bayes Estimates	Bayes No Prior	Empirical Bayes Standard Error
109448	110927	46963
479568	482157	148617
655656	660810	162104
1033109	1090752	220459
1388261	1530532	270730
2002772	2310959	374041
3018896	3806976	572899
3780759	4452396	720836
3811869	5066116	752593

The empirical Bayes estimate of total outstanding claims is 16280338 and the estimate of the standard error of total outstanding claims is 1313997.

The empirical Bayes standard errors are lower than the estimates with no prior information. The estimates of total outstanding claims for the later rows have benn quite considerably reduced, reflecting the reduction in the estimates of the row parameters. The empirical Bayes procedure has thus given less weight to the estimates of the parameters from the later years: it has allowed that the rise in the maximum likelihood parameter estimates from row to row may be due to random variation. As more data becomes available, and there is more evidence in favour of either of these possibilities, this may, or may not, be revised.

7. State Space Models

The previous section described the empirical Bayes framework in which it is assumed that the row parameters have the same prior mean. The advantage of this assumption is the connection made between the accident years. The chain ladder technique suffers from over-parameterisation which is a result of the accident years being regarded as cimpletely separate. The empirical Bayes assumption is one way of overcoming this. Another way of tackling this problem, and in some ways a superior way, is to use a state space approach. This method assumes a recursive connection between the rows, rather than the static assumption made by the empirical Bayes method that all the rows are similar. The state space model assumes that each accident year is similar to the previous one. Just how similar can be governed by the choice of a parameter variance. Section 7.1 describes the state space approach to the chain ladder linear model.

Another problem with the chain ladder technique is, paradoxically, that it makes too much connection between the accident years. It does this by assuming that the shape of the run-off is the same for all accident years: the same development parameters are used. It is also possible to relax this assumption, and details of this are given in section 7.2.

7.1 A state space representation of the chain ladder linear model.

In order to consider the state space model and dynamic estimation methods, it is necessary to set up the two-way analysis of variance model in a recursive form. This takes advantage of the natural causality of the data. The data which makes up the claims runoff triangle are received in the form :

$$Z_{1,1}, \begin{bmatrix} Z_{1,2} \\ Z_{2,1} \end{bmatrix}, \begin{bmatrix} Z_{1,3} \\ Z_{2,2} \\ Z_{3,1} \end{bmatrix}, \cdots, \tag{7.1}$$

and in year t the data which are received are

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t-1} \\ \vdots \\ Z_{t,1} \end{bmatrix} \tag{7.2}$$

The set of data vectors which together make up the whole triangle form a time series:

$Z_1, Z_2, \ldots, Z_t, \ldots$

In this time series, the data vector expands with t: for a triangular set of data, dim (Z_t) = t.

If the data are in the shape of a rhombus, which occurs when the early years of business are fully run off, then Z_t will reach a point when its dimension does not increase.

The analysis can be approached from in the context of multivariate time series. However, the special relationships between the elements of consecutive data vectors mean that it is simpler to generalise the theory of classical and Bayesian time series to two-dimensional processes. For a fuller discussion of the use of classical time series, the reader is referred to Verrall (1989).

There are two methods for calculating the forecast values and their standard errors. The simplest is to use the final parameter estimates and variance-covariance matrix as would be the case in a standard least-squares analysis. The more proper method calculates one-step-ahead , two-step-ahead , ... , (t-1)-steps-ahead forecasts at time t and their variance-covariance matrices. However, since the recursive approaches do not store covariances between, for example, the one-step-ahead and the (t-1)-step-ahead forecasts, the calculation of the variances of the forecasts causes problems. For this reason the first method will be used.

The chain ladder linear model takes the following form when three years' data have been received:

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{13} \\ Y_{22} \\ Y_{31} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{13} \\ e_{22} \\ e_{31} \end{bmatrix}$$

where $Y_{ij} = \log Z_{ij}$.

When the data are handled recursively, the model becomes:

$$Y_{1,1} = \mu + e_{1,1}$$

$$\begin{bmatrix} Y_{1,2} \\ Y_{2,1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} e_{1,2} \\ e_{2,1} \end{bmatrix}$$

$$\begin{bmatrix} Y_{1,3} \\ Y_{2,2} \\ Y_{3,1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \alpha_3 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} e_{1,3} \\ e_{2,2} \\ e_{3,1} \end{bmatrix} \qquad (7.3)$$

etc.

In general, the state vector at time t is defined by:

$$\underline{\theta}_t = \begin{bmatrix} \mu \\ \alpha_2 \\ \beta_2 \\ \vdots \\ \alpha_t \\ \beta_t \end{bmatrix} \qquad (7.4)$$

and equation (7.3) is called the observation equation. The state vector at time t is related to the state vector at time t-1 by the system equation. A recursive version of the chain ladder method is achieved by defining the system equation matrices as

$$\underline{\theta}_{t+1} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 \end{bmatrix} \underline{\theta}_t + \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \underline{u}_t \qquad (7.5)$$

where $\underline{u}_t$ contains the prior distribution of $\begin{bmatrix} \alpha_{t+1} \\ \beta_{t+1} \end{bmatrix}$.

The new parameters at time t+1 are $\begin{bmatrix} \alpha_{t+1} \\ \beta_{t+1} \end{bmatrix}$ and equation (7.5) implies that the existing parameters are unchanged, while the new parameters are treated as stochastic inputs.

If the variance of the errors, e_{ij} , is known and vague priors are used for the parameters, this method gives exactly the same results as ordinary least-squares estimation. It has the advantage that the data can be handled recursively. Also, it gives a method of implementing Bayesian estimation on some or all of the parameters. It has been assumed that the prior estimates of the parameters are uncorrelated: in other words that the stochastic input vector, $\underline{u}_t$, and the state vector, $\underline{\theta}_t$, are independent.

The equations above are an example of a state space system; a more general form is now considered. The models for $\underline{Y}_1$, $\underline{Y}_2$, ... , $\underline{Y}_t$, ... , which together make up the triangle can be written as

$$\begin{aligned} \underline{Y}_1 &= F_1\underline{\theta}_1 + \underline{e}_1 \\ \underline{Y}_2 &= F_2\underline{\theta}_2 + \underline{e}_2 \\ &\vdots \\ \underline{Y}_t &= F_t\underline{\theta}_t + \underline{e}_t \end{aligned} \qquad (7.6)$$

where $\underline{Y}_t = \log \underline{Z}_t$

Equation (7.6) is an observation equation and forms one part of a state system to which the Kalman filter can be applied in order to obtain recursive estimates of the parameters. $\underline{\theta}_t$ is known as the state vector and is related to $\underline{\theta}_{t-1}$ by the system equation. The observation equation and the system equation together make up the state space representation of the analysis of variance model.

The system equation relates $\underline{\theta}_t$ to $\underline{\theta}_{t-1}$ and defines how the state vector evolves with time. Thus, the time evolution of the system is defined on the state vector and the observation vector is then related to the state vector by the observation equation. There are many choices of system equation, the most general being:

$$\underline{\theta}_{t+1} = G_t\,\underline{\theta}_t + H_t\,\underline{u}_t + \underline{w}_t \qquad (7.7)$$

where $\underline{u}_t$ is a stochastic input vector

and $\underline{w}_t$ is a disturbance vector

and the distributions of $\underline{u}_t$ and $\underline{w}_t$ are:

$$\underline{u}_t \sim N(\hat{\underline{u}}_t, U_t)$$

$$\underline{w}_t \sim N(\underline{0}, W_t)$$

The choices of G_t, W_t and the distribution of $\underline{u}_t$ govern the dynamics of the system.

Suppose $\underline{\theta}_t \mid (\underline{Y}_1, \underline{Y}_2, \dots, \underline{Y}_{t-1}) \sim N(\hat{\underline{\theta}}_{t|t-1}, C_t)$. (7.8)

i.e. the distribution of the parameters, based on the data up to time $t-1$ is normal with mean $\hat{\underline{\theta}}_{t|t-1}$ and variance-covariance matrix C_t.

From equations (7.6) and (7.7), the distribution of $\underline{Y}_t$ given information up to time $t-1$ is

$$\hat{\underline{Y}}_{t|t-1} \sim N(F_t \hat{\underline{\theta}}_{t|t-1}, F_t C_t F_t' + V_t) \qquad (7.9)$$

When the observed value of $\underline{Y}_t$ is received, the state estimate can be updated to $\hat{\underline{\theta}}_{t|t}$ and the distribution of the state vector at time t forecast using equation (7.8).

The recursion is given by the following equations, a proof of which can be found in (for example) Davis & Vinter (1985). If the system and observation equations are given by equations (7.6) and (7.7), and the distribution of $\underline{\theta}_t$ given information at time t-1 is given by (7.9), then the distribution of the state vector can be updated when $\underline{Y}_t$ is received using the following recursion:

$$\hat{\underline{\theta}}_{t+1|t} = G_t \hat{\underline{\theta}}_{t|t-1} + H_t \hat{\underline{u}}_t + K_t (\underline{Y}_t - \hat{\underline{Y}}_t) \qquad (7.10)$$

where $$K_t = G_t C_t F_t' (F_t C_t F_t' + V_t)^{-1} \qquad (7.11)$$

$$C_{t+1} = G_t C_t G_t' + H_t U_t H_t' - G_t C_t F_t' (F_t C_t F_t' + V_t)^{-1} F_t C_t G_t' + W_t \qquad (7.12)$$

and $$\hat{\underline{Y}}_t = F_t \hat{\underline{\theta}}_{t|t-1} \qquad (7.13)$$

7.3 Dynamic Estimation of the Row Parameters

A model which applies dynamic estimation to the row parameters has the following system equation:

$$\underline{\theta}_{t+1} = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ 0 & \dots & 1 & 0 & \\ 0 & \dots & \dots & 0 & \end{bmatrix} \underline{\theta}_t + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{bmatrix} u_t + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix} w_t$$

where u_t has the prior distribution of β_{t+1}

and w_t is a disturbance term.

Thus the new row parameter, α_{t+1} , is related to α_t by:

$$\alpha_{t+1} = \alpha_t + w_t \qquad (7.14)$$

and a sophisticated smoothing method is produced.

The row parameters are related recursively and the column parameters are left as they were if their prior distribution is vague (although the estimates change because of the change in the estimation of the row parameters). The state variance is set as 0.0289, in order to compare with the empirical Bayes procedure. The practitioner is free to choose this variance as he sees fit: the larger the variance, the less connection is made between the rows. It is also possible to let this variance depend on t, and thus allow the amount of smoothing to be controlled by the perceived changes in the claims experience. It is also possible to obtain an estimate of this variance from the data, using maximum likelihood estimation. In the case in which the variance is set as 0.0289, the parameter estimates are as follows:

Table 7.1

Parameter	Estimate	Standard Error
μ	6.119	0.163
α_2	0.187	0.151
α_3	0.170	0.148
α_4	0.196	0.152
α_5	0.296	0.158
α_6	0.396	0.164
α_7	0.482	0.171
α_8	0.550	0.183
α_9	0.536	0.202
α_{10}	0.546	0.238
β_2	0.906	0.158
β_3	0.940	0.165
β_4	0.951	0.173
β_5	0.364	0.183
β_6	-0.028	0.195
β_7	-0.145	0.212
β_8	-0.457	0.236
β_9	-0.062	0.278
β_{10}	-1.406	0.378

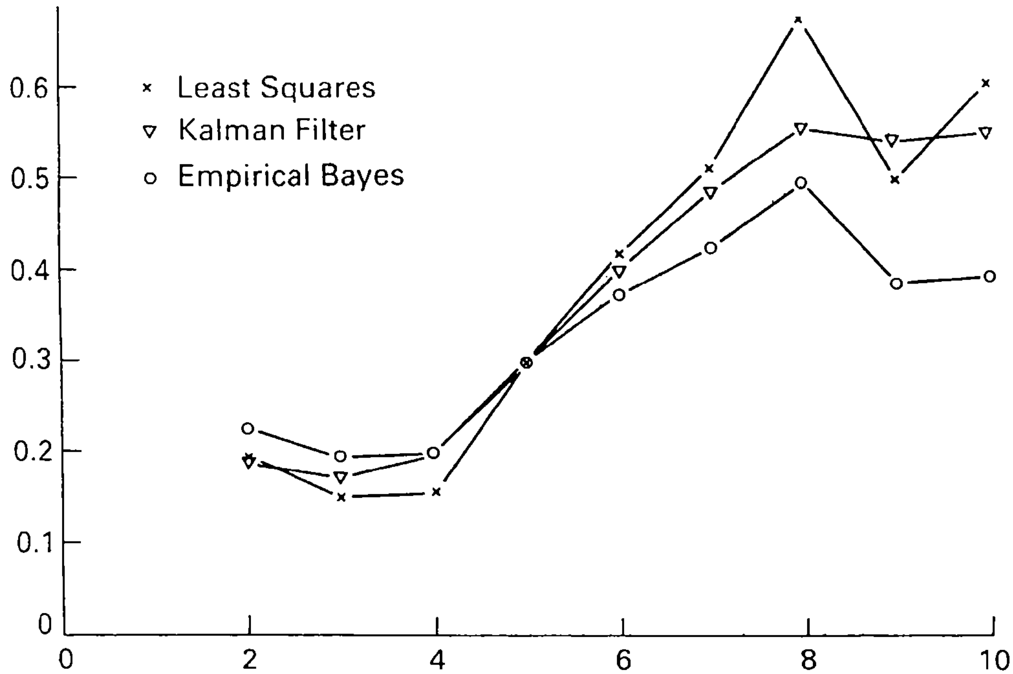
The row totals and their standard errors are given in the following table:

Table 7.2

Row	Predicted Outstanding Claims	Standard Error
2	109955	59278
3	491787	187134
4	686441	206954
5	1076957	277762
6	1486991	347441
7	2217311	491998
8	3309887	744931
9	4545466	1048855
10	4591188	1169469

The predicted overall total outstanding claims is 18515984 and the standard error of this estimate is 2660211. The standard error is lower than that when no prior knowledge is assumed because of the recursive relationship between the parameters. The effect of the Kalman filter on the parameter estimates will be illustrated by a graph, but it is interesting to compare the results with the empirical Bayes approach.

The following graph shows the parameter estimates for three cases: the model with no prior knowledge, the empirical Bayes model and the state space model. It can be seen from the graph that the state space model and empirical Bayes estimates have both smoothed the estimates of the row parameters to a certain degree. The empirical Bayes estimates have been drawn towards the overall estimate, with the amount of change depending on the data through the variation in each row and between the rows. The differences in the estimates of the row parameters has affected the estimates of outstanding claims. The standard errors have been reduced because the estimation has involved more of the data for each parameter. This is a beneficial affect of any of the Bayesian methods.



Least Squares
Kalman Filter
Empirical Bayes
0.6
0.5
0.4
0.3
0.2
0.1
0
0
2
4
6
8
10

7.4 Dynamic estimation of the development factors

It is well-known that the chain ladder technique assumes that the shape of the run-off curve is the same for each accident year, since the same development factors are used. However, it is doubtful whether this is justified in practice. It is likely that there will be a similarity between the run-offs in successive accident years, and it is possible to formulate a state space model to allow this without imposing an identical shape for each year. The basic chain ladder linear model is

$$E(Y_{ij}) = \mu + \alpha_i + \beta_j \qquad (7.15)$$

Allowing the devolpment factors to be completely separate for each accident year would lead to the following model:

$$E(Y_{ij}) = \mu + \alpha_i + \beta_{ij} \qquad (7.16)$$

We would expect the parameters β_{ij} to be similar for succesive values of i and so we impose the model

$$\beta_{i+1,j} = \beta_{ij} + \text{stochastic disturbance} \qquad (7.16)$$

The variance of the stochastic disturbance can be treated in much the same way as for the row parameters in section 7.3. We can now allow the shape of the run-off to vary from accident year to accident year by the choice of the variance of this stochastic disturbance. If it is zero, the run-off pattern is the same in each accident year and as it increases, the connection becomes less significant. We can allow the variance to depend on t and input a large value for one time point if it is believed that there has been a sudden change in the run-off pattern.

To illustrate the effect of this model, we analyse the data given in section 2, with the variance in equation (7.16) taken as 0.01. The main interest in this case is the effect on the run-off pattern, and so table 7.3 gives just the column parameters, β_{ij}.

Table 7.3

Column parameters from model with the same run-off in each row (from table 7.1):

0.906	0.940	0.951	0.364	-0.028	-0.145	-0.457	-0.062	-1.406

Column parameters from model with the same dynamic run-off pattern:

0.925	0.886	0.914	0.383	0.025	-0.175	-0.479	-0.074	-1.413
0.917	0.895	0.945	0.361	-0.035	-0.135	-0.460	-0.063	
0.920	0.907	0.964	0.361	-0.080	-0.130	-0.447		
0.918	0.920	0.980	0.332	-0.050	-0.161			
0.895	0.942	0.951	0.352	-0.026				
0.894	0.960	0.940	0.375					
0.890	0.990	0.944						
0.898	1.014							
0.897								

This illustration shows how changes in the run-off pattern can be observed. For example, the first column parameter is generally decreasing and the second one is increasing.

8. Conclusions

This paper has explored the various models which are available within the framework of the chain ladder linear model. It is envisaged that the practitioner will find all of these of use. The following points are of particular note.

Firstly, any of the Bayesian methods will improve upon the least squares (or uninformative prior) approach on the basis of parameter stability. This is because more information is used in estimating each parameter. For example, in the least squares case, there is only one data point from which to estimate the last row parameter; the Bayesian methods use the data from the other rows as well. To illustrate the affect of this consider a change in the data point in the last row from its present value of 344014 to 544014. The following table shows the predicted outstanding claims for each row from the different models. The first column shows the original results with no prior information.

Table 8.1

Row	Original Results			Revised Results		
	No prior Information	Dynamic Estimation	Empirical Bayes	No prior Information	Dynamic Estimation	Empirical Bayes
2	110927	109955	109448	110927	109958	110094
3	482157	491787	479568	482157	491822	481329
4	660810	686441	655656	660810	686637	657998
5	1090752	1076957	1033109	1090752	1078058	1039692
6	1530532	1486991	1388261	1530532	1491978	1400466
7	2310959	2217311	2002772	2310959	2239482	2024720
8	3806976	3309887	3018896	3806976	3399256	3063229
9	4452396	4545466	3780759	4452396	4847221	3819051
10	5066116	4591188	3811869	8011412	5261069	4411270

The last row prediction using no prior information has changed in proportion with the change in the data point. The other methods have dampened down this change because they use more information in the estimation of the parameter. They therefore exhibit greater predictor stability.

It is important to realise that the results must be used correctly. For example, it is often not necessary to produce a 95% upper confidence bound (a 'safe' reserve) on oustnanding claims for each row, but only for the whole triangle, although the 'safe' reserve for the whole triangle may be allocated among the rows. This is important since it can be seen that the standard errors for each row are, in general, relatively large. The standard error of the overall total is more reasonable. To extend this further, the practitioner may be required to set a 'safe' reserve for the whole company, rather than for each triangle; this would reduce the relative size of the standard error still further.

There are now a number of Bayesian methods which are available to the claims reserver, all of which have particular advantages over the classical estimation method. The chain ladder linear model represents a great step forward from the crude chain ladder technique and has opened the way to more sophisticated estimation methods.

9. References

B. Ajne (1989) *Exponential Run-off* Claims Reserving Manual, Vol.2, Institute of Actuaries.

D.Bradu and Y.Mundlak (1970) Estimation in Lognormal Linear Models. JASA Vol 6?, pp 198-211

H.Buhlmann (1967) *Experience Rating and credibility.* Astin Bulletin, Vol.4. No.3, pp 199-207.

S.Christofides (1990) *Regression Models based on Log-incremental Payments* Claims Reserving Manual, Institute of Actuaries, London.

D.R.Cox and D.V.Hinkley (1974) *Theoretical Statistics.* Chapman and Hall, London.

M.H.A.Davis & R.B.Vinter (1985) *Stochastic Modelling and Control.* Chapman & Hall.

D.J.Finney (1941) On the Distribution of a Variate whose logarithm is Normally Distributed. JRSS Suppl. 7 pp 155-61

Goovaerts and Hoogstad (1987) *Credibility Theory.* Surveys of Actuarial Studies , No.4.

S.Klugman (1989) *Credibility for Classification Ratemaking via the Hierarchical Normal Linear Model.* Proc. of Casualty Actuarial Society, Vol. 74, pp 272-321.

E.Kremer (1982) *IBNR-Claims and the Two-Way Model of ANOVA.* Scand. Act. J., Vol.1, pp 47-55.

D.V.Lindley and A.F.M.Smith (1972) *Bayes Estimates for the Linear Model (with Discussion)* JRSS, Series B, Vol. 34, No.1 pp 1-41.

A.Renshaw (1989) *Chain Ladder and Interactive Modelling* J.I.A. Vol.116, pp 559-587.

A.F.M.Smith (1973) *A General Bayesian Linear Model.* JRSS, Series B, Vol.35, No.1, pp67-75.

G.C.Taylor and F.R.Ashe (1983) *Second Moments of Estimates of Outstanding Claims.* J. of Econometrics, Vol.23, pp 37-61.

R.J.Verrall (1989) *Modelling Claims Runoff Triangles with Two-Dimensional Time Series,* Scand. Act. J, pp 129-138.

R.J.Verrall (1990) *Bayes and Empirical Bayes Estimation for the Chain Ladder Model* Astin Bulletin, Vol.20, No2, pp217-243.

R.J.Verrall (1991a) *On the Unbiased Estimation of Reserves from Loglinear Models* Insurance: Mathematics and Economics, Vol. 10 pp75-80.

R.J.Verrall (1991b) *Chain Ladder and Maximum Likelihood* J.I.A. Vol.118, pp 489-499.

F.de Vylder (1982) *Estimation of IBNR claims by Credibility Theory.* Insurance: Mathematics and Economics 1, pp 35-40.

B.Zehnwirth (1985) Interactive Claims Reserving Forecasting System, Benhar Nominees Pty Ltd., Tunawarra, NSW, Australia.

Probabilistic Development Factor Models with Applications to Loss Reserve Variability, Prediction Intervals, and Risk Based Capital

by Ben Zehnwirth

1.0 INTRODUCTION AND SUMMARY

The present paper aims to present a statistical modelling framework and environment for conducting loss reserving analysis. The modelling framework and approach affords numerous advantages including increased accuracy of estimates and modelling of loss reserve variability. Since the loss reserve is likely to be the largest item in the insurer's balance sheet and is subject to much uncertainty, modelling of loss reserve variability is an integral component of assessing insurer solvency and assessment of risk based capital.

The paper is organised as follows:

Forecasting and some modelling concepts are introduced in Section 2. The salient features of the data that ought to be captured by a model are discussed and arguments in favour of probabilistic models are presented. It is emphasised that the only way to assess loss reserve variability is through probabilistic models. The statistical MODELLING FRAMEWORK is introduced where each model in the framework has four components of interest. The first three involve trends in the three directions, development year, payment/calendar year and accident year and the fourth component is the random fluctuations (distributions) about the trends.

In Section 3 we begin by discussing trend adjustments to a univariate time series and illustrate how analogous adjustments to loss reserving data cannot be handled by graph and ruler, mainly as a consequence of the projection of the payment/calendar year trends onto the development year and accident year directions. Two deterministic models Cape Cod (CC) and Cape Cod with constant inflation (CCI) are discussed. Age-to-age development factors are defined as trend parameters.

A rich class of deterministic development factor models is introduced in Section 4 where each model in the framework contains the three trend components of interest. It is shown how as a result of the projection of calender year (trends), a very simple trend model causes very different development year trends (development factors) for different accident years. Standard actuarial techniques based on age-to-age link ratios of the cumulative payments cannot capture the payment/calendar year trends in the payments.

In Section 5 the class (or family) of deterministic development factor models that only contain trend components in the three directions is extended to include random fluctuations. The resulting models in the rich Development Factor Family (DFF) are probabilistic models that relate the distributions of 'payments' in the various cells in the triangle by trend parameters. It is emphasised that one of the principal uses of regression is the estimation (or fitting) of distributions. Estimation of a model belonging to the DFF involves the fitting of distributions to the cells in the loss development array. Data based on a simple DFF model are generated (simulated) and it is demonstrated how the development year patterns are invariably complex. The trends cannot be determined from the age-to-age link ratios nor from graphs. For readers who are sceptics and may argue "But this is simulated data" should read Section 12 where we analyse real life data involving a line written by a larger insurer for which the age-to-age link ratios on the cumulative payments are relatively smooth. HOWEVER, there are major shifts in payment/calendar year trends in the payments that are quite alarming.

We use regression for a number of purposes:

* Estimation of trends.
* Estimation or fitting of distributions.

In Section 6 we demonstrate how regression can also be employed to adjust data for trends. We state as a THEOREM that the only way to separate payment/calendar year trends from development year trends is by application of regression. Practical applications of regressions involving real life data sets are given in Sections 12 and 13.

In Section 6 we also present a number of tests that we believe any sound loss reserving statistical framework should pass. It is shown that standard actuarial techniques based on age-to-age link ratios fail these minimum tests.

As a result of the dependence of the payment/calendar year direction on the other two directions, many of the models in the DFF that contain many parameters cannot be estimated in a spreadsheet or statistical package and some that can be estimated may contain much parameter uncertainty. This phenomenon, known as multicollinearity, is discussed in Section 7 and motivates the introduction of varying parameter, dynamic or credibility models. Varying parameters or stochastic parameters can also be regarded as proxies for the myriad of variables that affect the complex claims generating process.

In Section 8 we show how the (fixed) parameter regression models may be estimated in a spreadsheet or statistical package and how an estimated model may be employed in producing forecast distributions of (incremental) payments. The forecast (estimated) distributions provide information required for the assessment of risk based capital and solvency.

Additional modelling concepts including parsimony, Akaike Information Criterion and distributional assumptions are discussed in Section 9. Moreover, we describe the importance of the twin concepts of stability and validation analysis and show how data with unstable trends (in the payments) are less predictable (subject to greater uncertainty) than data with stable trend (and some random fluctuations). Parameter *uncertainty (or instability) can reduce predictability much more than process* uncertainty.

Accuracy of forecast distributions is also discussed. We emphasise that the "optimal" statistical model, when trends are unstable, may not be the best for producing forecasts and discuss what assumptions may be appropriate for the future, especially in the light of analysing other data types. Instability in trends in the more recent payment years in the incremental payments requires more actuarial judgment about future trends.

The model building strategy and selection of appropriate assumptions about the future are discussed in Section 10. It is stressed that the model building strategy is necessarily an iterative cycle of model specification, estimation and testing. If trends in the more recent payment/calendar years are unstable, the nature of the instability and possible explanation for the instability is relevant information in deciding on assumptions for the future. This typically may require analysis of other data types employing the advocated modelling framework. We conclude in Section 10 with a discussion of time series models versus explanatory (or casual) models and offer arguments for the superiority of the former over the latter.

Section 11 discusses how prediction intervals may be derived from the forecast distributions and how they are relevant to the assessment of risk based capital and solvency. Prediction intervals computed from the forecast distributions are conditional on the assumptions made about the future remaining true.

The preliminary diagnostic analysis and the model building strategy are illustrated with two real life examples. Project 1 of Section 12 is concerned with real data of a large company. In terms of standard age-to-age link ratio techniques the data and ratios are relatively smooth and it does not appear that there are any problems. HOWEVER, there are major shifts in payment/calendar year trends in the payments that are alarming especially since the new high trend cannot be explained by a corresponding increase in speed of closure of claims. Project 2 of Section 13 also involves real data. Here the link ratios are relatively irregular, yet trends are stable, so that three years earlier estimation of the same model would have forecast the distributions of payments in the cells of the last three payment/calendar years and moreover would have produced the same outstanding reserve estimates.

In Section 14 we remark about an important extension of the DFF MODELLING FRAMEWORK that makes the family of models infinitely richer.

The paper concludes with summary remarks in Section 15.

Throughout the paper we also hope to dispel a number of pervasive loss reserving myths regarding data, age-to-age link ratios, volume, credibility, sources of information, actuarial judgment (when and where required), business knowledge, statistical probabilistic modelling and forecasting.

2.0 STATISTICAL FORECASTING

The best way to suppose what may come, is to remember what is past.

George Savile, Marquis of Halifax.

In this section we discuss a number of fundamental statistical forecasting concepts including which salient features of the data should be "*remember what is past*".

2.1 FORECASTING

Indeed it (forecasting) has been likened to driving a car blindfolded while following directions given by a person looking out the back window. Nevertheless, if this is the best we could do, it is important that it should be done properly, with the appreciation of the potential errors involved. In this way it should at least be possible to negotiate straight stretches of road without a major disaster.

Andrew C. Harvey [9]

In the loss reserving context the '*straight stretches*' are the stable trends in the (incremental) payments. If the trends have been stable in past years, we are confident in supposing the same trends in the future.

2.2 WHY A PROBABILISTIC OR STOCHASTIC MODEL?

There are extremely compelling reasons as to why we should be using probabilistic models to model insurance data, whether for the purpose of loss reserving, rate making or any other purpose.

According to Arthur Bailey's [2] paper ***Sampling Theory in Casualty Insurance***, any insurance data can only be regarded as "*an isolated sample* ...". See top of page 8 of the text book ***Foundations of Casualty Actuarial Science*** [5]. Bailey is basically saying that any insurance data can only be regarded as a sample (path) from perhaps a very complex process.

If a fair coin is tossed 100 times. the mean number of heads is 50, but the probability of observing 50 heads is only 0.08. If a fair dice numbered 1 to 6 is rolled, the mean is 3.5, yet the probability of observing 3.5 is zero. (The variability inherent in coin tossing in known as process uncertainty).

So, the probability of observing the mean in most, if not all, insurance processes is zero. Given, that we do not observe the mean, we need to compute more than just the mean. The mean on its own is not terribly informative. We need to also compute the standard deviation, so that we have some idea of how 'far' our (future) observations will be from the mean. The best, of course, is to compute the whole distribution. From the computed distribution we can derive the moments, percentiles and prediction (confidence (*sic*)) intervals.

Returning to the text book ***Foundations of Casualty Actuarial Science*** [5], the introductory chapter 1, top of page 2, says "*The mention of probabilities reminds us to state the obvious, that probability theory (whether classical or Bayesian) forms the basis of actuarial science. If the actuaries hadn't probability theory, they would have to invent it.*" Indeed, the author also believes that statistical probabilistic methods are essential to actuarial studies, and it is principally by the aid of such methods that these studies may be raised to the rank of sciences.

2.3 MODELLING FRAMEWORK

The models considered in the present paper are relatively simple. They have four components of interest that have a straightforward interpretation.

The first three components are the trends in the three directions, **development year, accident year** and **payment/calendar year**. The fourth component is the random fluctuations about the trends. The random fluctuations component is just as important as the three trend components and is necessarily an integral part of the model. The data or transform thereof are decomposed thus:

DATA = TRENDS + RANDOM FLUCTUATIONS

The concept of **trends** and **random fluctuations** about trends is over two hundred years old. These concepts have been widely used in analysing (and forecasting) any univariate time series such as sales, stock market prices, interest rates, consumption, energy and so on.

The principal aim of analysing a loss development array is to obtain a sensible description of the data. The trends in the past, especially in the payment/calendar year direction, are determined and the random fluctuations about the trends are quantified, so that it can be best judged which assumptions should be used for future trends (and random fluctuations). The models are probabilistic (equivalently, stochastic) since the probability distributions of the random fluctuations 'about' the trends are identified. Probabilistic models are testable and can also be validated. They also afford numerous other advantages including computation of risk margins required for the assessment of **risk based capital**.

IF THE TRENDS ARE STABLE THEN THE MODEL WILL VALIDATE WELL AND BE STABLE. If the trends are unstable then the decision about future trends is no longer straightforward. Instability in trends with little random variation about the trends makes data less predictable then stable trends with much random fluctuation. See Sections 9.6, 10.2 and 10.3. The same principles apply to the modelling of a univariate time series.

The 'best' identified model contains assumptions (equivalently, information). All the assumptions must be tested to ensure they are supported by the data (experience).

As we proceed through the model identification strategy we are extracting information (about trends and stability thereof and the amount of random variation) and we 'hope' that the 'best' identified model tells us that the calendar year trend is stable (especially more recently). If trends are not stable then we may not necessarily use the optimal statistical model for forecasting. See Section 9.6.

None of the numerous models contained in the MODELLING FRAMEWORK actually represent explicitly the underlying claims generating processes. The multitude of variables involved in generating the claims are invariably complex. What we attempt to achieve is the identification of a parsimonious model in terms of the simple components of interest for which all the assumptions inherent in the (probabilistic) model are supported by the data. It is subsequently argued that the experience (data) can be regarded as a sample (path) from the identified probabilistic model. The multitude of variables that are the determinants of the claims processes are proxied by the TRENDS and the (residual) variance of the RANDOM FLUCTUATIONS. Another classical modelling example in insurance where the same kind of modelling concepts are used is when a Pareto distribution, say, is fitted to loss sizes. It is not assumed that the Pareto distribution represents the underlying generating process. Whatever is driving the claims is very complex and depends on many variables. All that is assumed is that the experience (sample) can be regarded as a realisation from the estimated Pareto distribution. Subsequently the estimated Pareto distribution is used to estimate probabilities of very large claims including those exceeding the maximum observed claim in our sample and most importantly it is used to quantify probabilistically the variability in loss sizes.

The principal advantage of an explicit statistical model is that it makes the assumptions clear. Other advantages include improved accuracy and quantification of variability required for assessment of risk based capital and testing of solvency.

3.0 THE GEOMETRY OF TRENDS AND AGE-TO-AGE DEVELOPMENT FACTORS

In this section we show that loss development arrays possess only two independent directions, not three, and define age-to-age development factors as development year trends.

3.1 TREND ADJUSTMENTS TO A UNIVARIATE SERIES

In one dimension, or equivalently for any univariate series, trend concepts are intuitive and natural.

Consider the series log P_t where P_t is the price of gasoline in year t. Figure 3.1.1 below depicts the log P_t series (dark line segments) over a 20 year period.

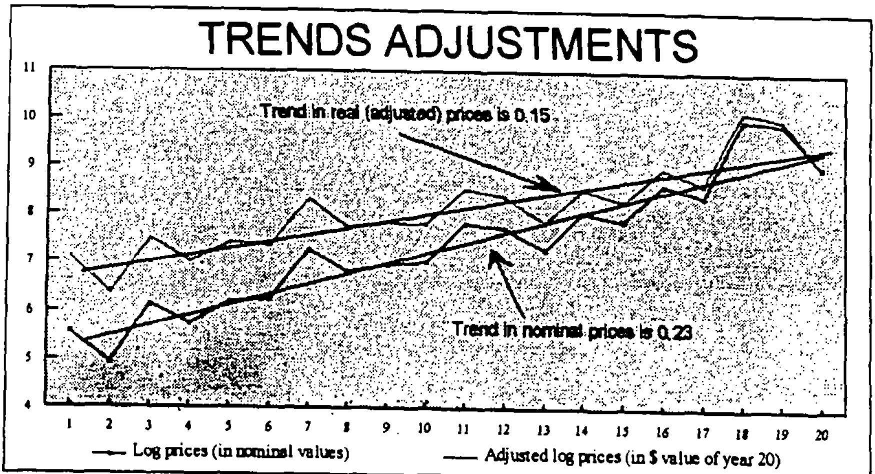


Figure 3.1.1

It appears that there is a constant average trend in the nominal prices. The least squares estimate of the trend is 0.23, say. So prices have been growing at an average rate of 23%. However, 23% is the nominal growth, since there has been economic inflation over the 20 year period. Suppose economic inflation has been 8% continuous rate for the whole 20 year period. The light line segments represent the log prices adjusted to the $ value of year 20.

The trend in the adjusted prices is 23% - 8% = 15%. If instead, one was only given the nominal prices and the adjusted prices (without knowing the adjustment), the 8% adjustment could be determined by estimating the difference in trends in the two series. Trends (on a log scale) are additive.

So, REGRESSION as an approach to estimating trends and adjusting data, immediately suggests itself.

3.2 TREND PROPERTIES OF LOSS DEVELOPMENT ARRAYS

Since a model is suppose to capture the trends in the data, it behoves us to discuss the geometry of trends in the three directions, víz., **development year** (or delay), **accident year** and **payment** (or calendar) **year**.

Development years are denoted by d; d=0,1,2,...,s-1; accident years by w; w=1,2,...,s; and payment years by t; t=1,...,s.

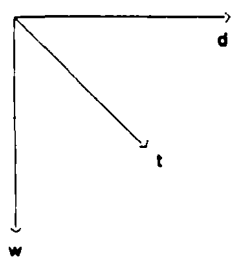


Figure 3.2.1

The payment year variable t can be expressed as t = w + d. This relationship between the three directions implies that there are only two 'independent' directions.

The two directions, delay and accident year, are orthogonal, equivalently, they have zero correlation. That is, trends in either direction are not projected onto the other. The payment year direction t however, is not orthogonal to either the delay or accident year directions. That is, a trend in the payment year direction is also projected onto the delay and accident year directions. Similarly, accident year trends are projected onto payment year trends.

In order to aid the exposition we shall assume, without loss of generality, that the numbers in the loss development array are incremental payments. **It is emphasised that all the arguments and concepts presented apply to all loss development arrays including incurreds, counts, averages and so on.**

We now illustrate the geometric properties of trends of a loss development array with some data.

Consider the following triangle of incremental paid losses:

Triangle One

100	200	150	100	80	60	40	20
100	200	150	100	80	60	40	
100	200	150	100	80	60		
100	200	150	100	80			
100	200	150	100				
100	200	150					
100	200						
100							

This triangle will be said to satisfy the Cape Cod assumptions, viz., homogeneity of age-to-age development factors across accident years and homogeneity of levels across accident years. Each accident year has the same initial starting value, that is, same value in delay 0.

Suppose we subject the payments to a 10% yearly inflation across the payment years. We obtain the next triangle:

Triangle Two

100	220	182	133	117	97	71	39
110	242	200	146	129	106	78	
121	266	220	161	142	117		
133	293	242	177	156			
146	322	266	195				
161	354	292					
177	390						
195							

To obtain the t^{th} diagonal of the second triangle, we multiply each payment in the t^{th} diagonal of triangle one by $(1.1)^{t-1}$.

We observe the following:

1. For triangle two, age-to-age development factors are homogeneous across accident years but are 10% higher than in triangle one.

2. In triangle two there is a 10% accident year trend.

Observations 1 and 2 imply that triangle two could be obtained from one by the two successive (and commutative) operations: subject triangle one to 10% per year trend in accident year direction to obtain:

Triangle Three

100	200	150	100	80	60	40	20
110	220	165	110	88	66	44	
121	242	182	121	97	73		
133	266	200	133	106			
146	293	220	146				
161	322	242					
177	354						
195							

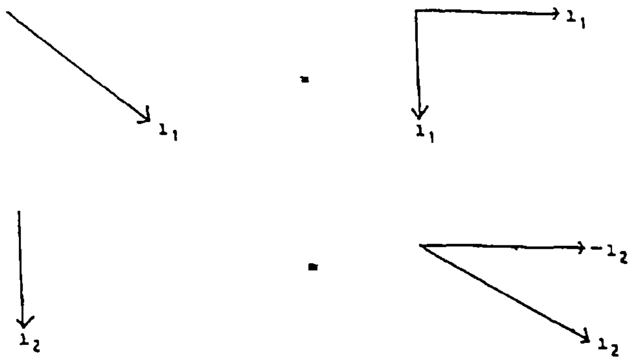
and then subject triangle three to 10% trend in the development year direction to obtain:

Triangle Four

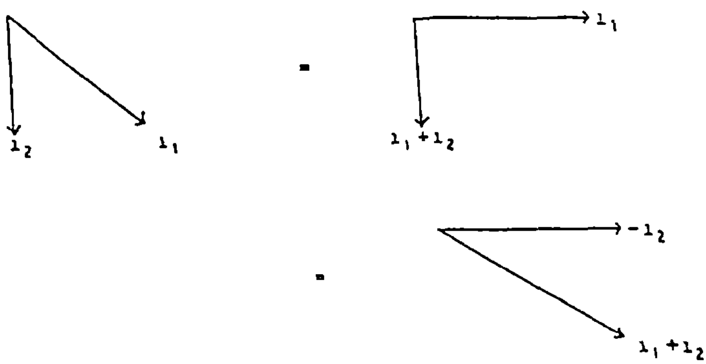
100	220	182	133	117	97	71	39
110	242	200	146	129	106	78	
121	266	220	161	142	117		
133	293	242	177	156			
146	322	266	195				
161	354	292					
177	390						
195							

Triangle four is the same as triangle two. A loss development array depicted by triangle two (or four) is said to satisfy the Cape Cod with constant payment year inflation assumptions.

The following displays demonstrate the equivalence of trends in general.



The above equivalence relations are exemplified by the relationships between the four triangles. We also have,



It is important that the reader understands the relationship and difference between Cape Cod (CC) data and Cape Cod with constant inflation (CCI) data.

CC data have accident years that are completely homogeneous (homogeneity of level or values at development year zero and homogeneity of age-to-age factors). CCI data can be obtained from CC data by subjecting the payment years to a constant trend. If we remove the constant payment year trend from the CCI data we will have CC data.

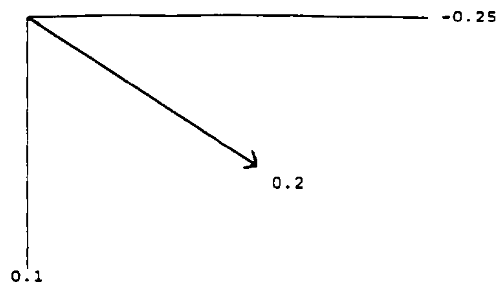
So, the difference between CCI data and CC data is a calendar year trend adjustment. If we did not know how the CCI data were created from the CC data, how would we determine the (simple) difference?

With the univariate series considered in Section 3.1 the difference between the nominal prices and adjusted prices can be determined by estimating the trend, using *eye and ruler*, for each series. Estimating trend using *eye and ruler* can be regarded as a form of crude regression. With the loss reserving data CC and CCI, it also makes sense to estimate the payment year trends and subsequently conclude that the difference in the two loss development arrays resides in the difference in the two trends. But how do we estimate the trends? Given the dimensionality of the data, eye and ruler are not useful. Moreover, given the geometry of trends, we need to separate the trends in the three directions. Equivalently, we need to determine the payment year trends after adjusting for development year trends.

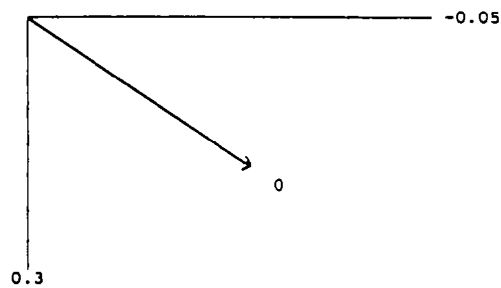
Accordingly, formal regression is suggested as the only way of separating the trends.

A number of words of caution. In actual fact the "true" trends in the three directions are non-identifiable. It is only the resultant trends that are identifiable.

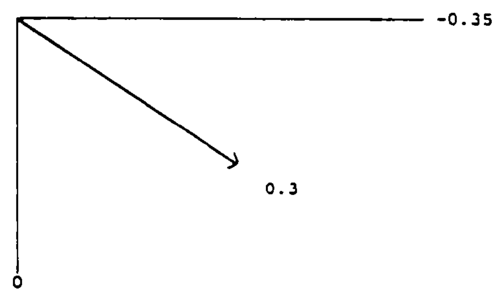
Here is an example. Consider a CC triangle for which the (continuous) trend across development years is constant and is -0.25. Suppose to this CC triangle we introduce a continuous calendar year trend of 0.2 and a continuous accident year trend of 0.1. The adjusted triangle can be represented thus:



Alternatively, it can be represented as:



or,



All three trend triangles are the same and would produce the same projections for the completion of the rectangle. We have three directions (or variables) but only two independent equations.

3.3 DETERMINISTIC AGE-TO-AGE DEVELOPMENT FACTORS

Consider, at first, only one accident year (say, the first) that takes the value p(d) at development year d and let y(d) = log p(d).

Define: $\alpha = \log p(o)$

and

$\gamma = y(j) - y(j-1)$.

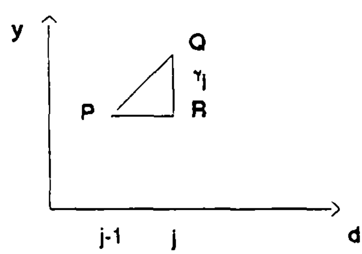


Figure 3.3.1

The parameter α (alpha), denotes the initial value, or intercept, or level whereas the parameter γ_j represents the trend, on a logarithmic scale, from development year j-1 to development year j.

The parameter γ_j is a difference on a log scale and since the length of PR in Figure

3.3.1 is 1, γ_j is the slope of the line PQ, and hence is the trend between development years j-1 and j.

Now,

$$y(d) = y(o) + y(1) - y(o) + \ldots + y(d) - y(d-1)$$

$$= \alpha + \sum_{j=1}^{d} \gamma_j \quad . \qquad (3.3.1)$$

That is, y(d) can be expressed as the initial value plus the sum of the differences to development year d. The differences can also be regarded as trends. Indeed,

$$\gamma_j = y(j) - y(j-1)$$

$$= \log p(j) - \log p(j-1)$$

$$= \log \left[\frac{p(j)}{p(j-1)} \right] \quad . \qquad (3.3.2)$$

One of the principal reasons for taking logarithms of the data is because the difference of two logarithms is equivalent to analysing trends and approximately equivalent to analysing percentage changes.

The trend parameter γ_j is the log of the ratio p(j)/p(j-1). The latter ratio is an age-to-age development factor. So, γ_j can also be interpreted as a log of a development factor. Indeed, in what follows we shall refer to it as a development factor (on a log scale).

Consider the following monotonically increasing series {p(j)} for which the trends are depicted in the Figure 3.3.2 below.

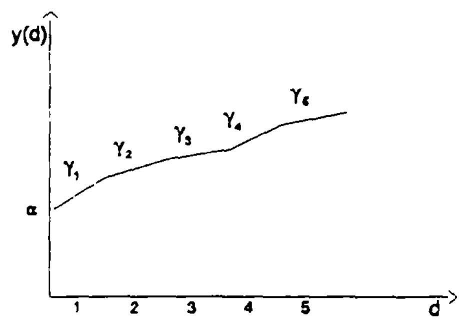


Figure 3.3.2

The γ's represent both the differences in y values and the trends depicted by the straight line segments.

Accordingly, development factors on a log scale form a curve comprising of straight line segments (trends).

4.0 DETERMINISTIC DEVELOPMENT FACTOR MODELS

In this section we develop the mathematical description of the two models corresponding to triangles one and two respectively of Section 3.2.

Let p(w,d) denote the value in the loss development array corresponding to accident year w and development year d and set y(w,d) = logp(w,d).

4.1 CAPE COD (CC)

Consider triangle one of Section 3.2. Each accident year has the same α value, viz., α =log100 and each accident year has the same development factors $\gamma_1, \gamma_2, \ldots, \gamma_6$ (γ_7). For example, γ_3=log(100/150).

So, we can write

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j \qquad (4.1.1)$$

Equation (4.1.1) describes the deterministic CC model.

4.2 CAPE COD WITH CONSTANT INFLATION (CCI)

Consider now triangle two of Section 3.2. It was obtained from triangle one by subjecting it to a constant trend in the payment year direction.

Let's denote the payment year trend on a logarithmic scale by the Greek letter, ι (called iota). For triangle two ι =log 1.1.

The value y(w,d) that lies in payment year w + d is inflated by ι (w + d - 1).

So, for triangle two,

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j + \iota \cdot (w + d - 1) \ . \qquad (4.2.1)$$

The last equation may be re-cast.

$$y(w,d) = \alpha + \iota \cdot w - \iota + \sum_{j=1}^{d} (\gamma_j + \iota) \ . \qquad (4.2.2)$$

The two foregoing equations are identical and represent the CCI deterministic model. The latter equation tells us that the level parameter for accident year w is $\alpha + \iota + w - \iota$, so that there is an ι trend along the accident years and that the development factor from delay j-1 to j is $\gamma_j + \iota$. This is just a mathematical verification that the payment year trend ι projects on the other two directions.

4.3 CC FAMILY AND CCI FAMILY

There are other CC models for which the CC assumptions viz., homogeneity of accident years, apply.

For example, it may be that $\gamma_3 = \gamma_4 = \ldots = \gamma_8$, so that the trends from development year two to eight are constant as depicted below:

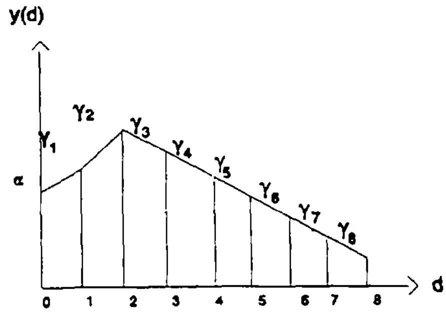


Figure 4.3.1

Another possibility is that all development factors γ_1, γ_2, ..., are equal to γ say, so that we could write:

$$y(w,d) = \alpha + \gamma d \quad . \tag{4.3.1}$$

This model we call the single development factor (SDF) model. It is a straight line curve on a log scale and exponential curve on the $ scale. It is the same curve for each accident year.

So, we can construct many variants of the CC model (4.1.1.). In the sequel, anytime we refer to CC without an added qualification we shall mean model (4.1.1) with distinct γ's.

Similarly, depending on the "relationships" in the γ's in the CCI model, we can construct many variants of the CCI model.

4.4 A CC MODEL WITH THREE INFLATION PARAMETERS

The data in Appendix A1 to Appendix A4 are generated as follows.

First we create payments based on formula:

p(w,d)= exp(alpha- 0.2*d).

So this is deterministic SDF data (where the accident years are homogeneous). See Appendix A1.

On a log scale we introduce a 10% trend from 1978-82, 30% trend from 1982-83 and 15% trend from 1983-91. See Appendix A2.

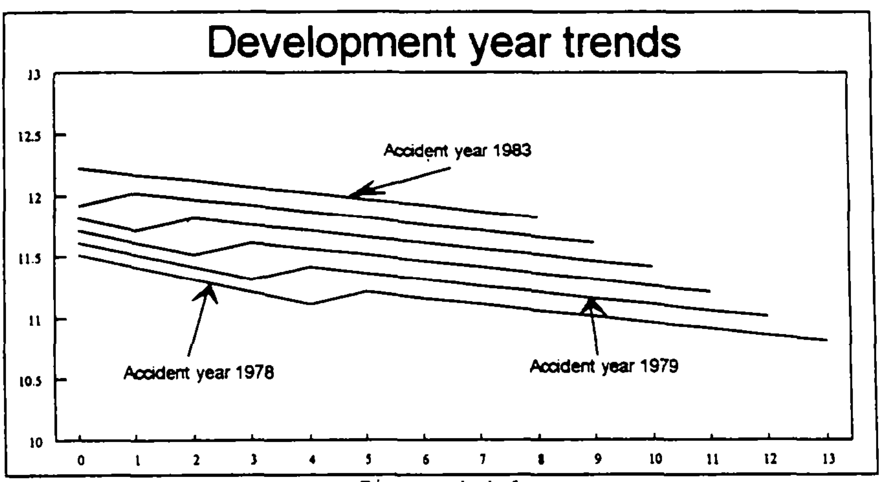


Figure 4.4.1

Figure 4.4.1 displays the graph of the log data versus development year for the first six accident years. The reader can reproduce this graph in a spreadsheet.

Observe how calendar year trends project onto development years and accident years.

Consider the first accident year 1978. The 10% calendar year trend projects onto the development year, so that the resultant trend from development year 0 to development year 4 is -0.2 (the gamma) + .1 (the iota)= -.1. The 30% trend between calendar years 1982 and 1983 also projects onto the development year so that the trend between development year 4 and 5 is +.1=-0.2+.3. Thereafter the trend is -.2+ 15=-.05. Since .15 is larger than 1, the decay in the tail is less rapid (-.05>-.1).

Consider the next accident year 1979. First up to development year 3, this accident year is 10% higher than the previous one since the 10% calendar year trend also projects onto the accident years. The 10% upward trend is one development year earlier than in previous accident year since the 30% trend is a calendar year change.

So, changing calendar year trends can cause some interesting development year patterns. The pattern is different for each accident year. The calendar year trends cannot be determined by the link ratios (Appendix A4).

The patterns became much more complicated in the presence of random fluctuations superimposed on the trends. See Section 5 for a discussion of the current example including random fluctuations.

The model describing the data we have constructed can be represented pictorially thus:

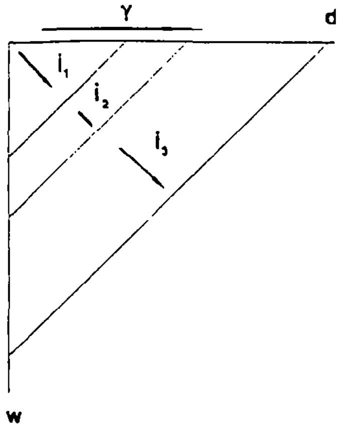


Figure 4.4.2

where $\gamma = -0.2$, $\iota_1 = 0.1$, $\iota_2 = 0.3$ and $\iota_3 = 0.15$.

Writing the equations explicitly is not necessary. Indeed, it is too complicated.

We note that the resultant trend (age-to-age development factor) between development years j-1 and j is the (base) development factor γ between the two development years plus the payment year trend ι (iota) between the two corresponding payment years.

The above model can be described succinctly in terms of the five parameters, α, γ, ι_1, ι_2 and ι_3. We could create a slightly more involved model by adding accident year trends (more α's).

4.5 CHAIN LADDER (CL)

The chain ladder (CL) statistical model is described in Christofides [4]. It is a two-way ANOVA model where accident years and development years are two factors at various levels. The CL statistical model is the direct statistical extension of the standard age-to-age development factor technique. See Christofides [4] for details. It is written (omitting the random fluctuations).

$$y(w,d) = \alpha_w + \sum_{j=1}^{d} \gamma_j \quad . \qquad (4.5.1)$$

The parameter α_w corresponding to accident year w represents the effect of accident year w and the parameter $\gamma_j - \gamma_{j-1}$ (difference in trends) represents the effect of development year j. The number of parameters in the model is 2s-1.

The CC model assumes complete accident year homogeneity, that is, same α and same γ_j's. For the CL model we assume homogeneity of development factors (γ_j's), but heterogeneity of levels (α's).

The principal deficiency of the CL model is that it does not relate the calendar years in terms of trends.

If we do not have an estimate of trends in the past, how do we know what assumptions we can make about the future trends? See comments by George Savile at beginning of Section 2.0 and the discussion in Section 9.6.

HOWEVER, the CL model is an extremely powerful interpretive tool as we shall see in Section 6 and more impressively in an application to a real life example in Section 12.

4.6 THE SEPARATION MODEL (SM)

The separation method separates the base systematic run-off pattern (assumed homogeneous across accident years) from exogenous influences, viz., payment year inflation (or effects). The deterministic model is usually expressed (parametrized) as

$$p(w,d) = e(w)\, b_d\, \lambda_{w+d} \quad ,$$

where the { e (w) } are the exposures, proportional to number of claims incurred. $\{ b_d \}$ are the development factors and the parameter λ_{w+d} expresses the 'effect' of payment year t = w + d.

The *corresponding model in our framework is written* (parametrized) as

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j + \sum_{t=2}^{w+d} \iota_t \quad , \qquad (4.6\ 1)$$

where the parameters $\{ \gamma_j \}$ are the base systematic development factors and ι_t is the force of inflation from payment year t - 1 to payment year t.

The model has 2s - 1 parameters.

Note that this model necessarily assumes that there are significant changes in inflation rates (trends) between every two contiguous payment years and, moreover that *there are significant changes in base development factors between every two* development years.

Refer to the discussion of Section 9.6 where we show that if trends are indeed unstable then the payments are not terribly well predictable.

4.7 DETERMINISTIC DEVELOPMENT FACTOR FAMILY

Let's reconsider the model of Section 4.4. It can be described succinctly as a version of CC (viz., SDF) subjected to three payment year trends. If we remove the three payment year trends, we are back to SDF. On this model we could also superimpose (add) accident year trends.

So, any determistic development factor model (DFF) can be described as some version of CC subject to payment year trends and accident year trends.

Mathematically, the family of development factor models is

$$y(w,d) = \alpha_w + \sum_{j=1}^{d} \gamma_j + \sum_{t=2}^{w+d} \iota_t \quad . \tag{4.7.1}$$

A model has a level parameter α_w for accident year w - it represents the effect or level or exposure of the accident year. Between every two development years, we have a development factor or trend parameter γ_j (the factor from delay j-1 to j) and between every two payment years we have a trend (or inflation) parameter ι_t, the inflation from payment year t-1 to t.

All models considered thus far belong to the development factor family. For example, CC is written as:

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j \quad . \tag{4.7.2}$$

So for CC type model $\alpha_w = \alpha$ (for each w) and $\iota_t = 0$ for each t.

There is no need to memorise the equation representing the family of models. All that needs to be understood is that the parameters of a model comprise (i) trends (development factors) in the development year direction (the γ's), (ii) levels (exposures) for each accident year (the α's) and (ii) trends (inflation) in the payment year direction (ι 's). Furthermore, any payment year trend projects on the other two directions.

5.0 STOCHASTIC DEVELOPMENT FACTOR MODELS

In this section the class of deterministic DFF models (4.7.1) that only contain trend components is extended to include random fluctuations.

Consider one accident year only for which the deterministic model is

$$y(d) = \alpha + \sum_{j=1}^{d} \gamma_j \quad . \tag{5.1}$$

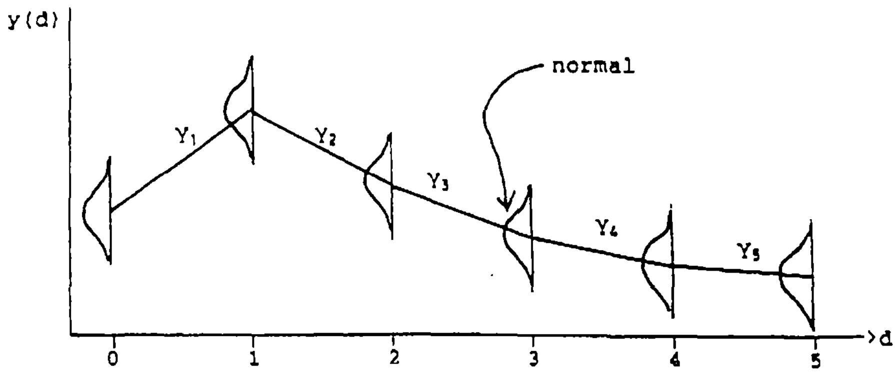
This model says that at delay o we can only observe one (log) value, viz α. Similarly, for the other delays. Between any two delays we can only observe one trend, the trend corresponding to the development factor.

We now assume that around the trends there are random fluctuations. We write

$$y(d) = \alpha + \sum_{j=1}^{d} \gamma_j + \varepsilon \quad , \tag{5.2}$$

where ϵ the error term, has a normal distribution with mean 0 and variance σ^2. In actuarial parlance σ^2 is known as the process uncertainty. Given that the errors are random variables, the dependent variable y is also a random variable.

The probabilistic (stochastic or regression) model is depicted below.



For the stochastic model, α is no longer the value of y observed at delay 0. It is the mean of y(o). Indeed, y(o) has a normal distribution with mean α and variance σ^2.

Similarly, γ_j is not the observed trend between delay j-1 and j, but rather it is the mean trend.

The parameters of the stochastic model represent means of random variables. Indeed, the model (on a log scale) comprises a normal distribution for each development year where the means of the normal distributions are related by the parameter α and the trend parameters $\gamma_1, \gamma_2, \ldots, \gamma_5$.

From equation (5.2) we have

$$y(d) - y(d-1) = \gamma_d + \varepsilon_d - \varepsilon_{d-1} , \tag{5.3}$$

where ε_d is the 'error' at delay d.

Accordingly,

$$E\left[\log\frac{p(d)}{p(d-1)}\right] = \gamma_d \quad . \tag{5.4}$$

That is, the development factor γ_d is the mean of the log of the ratio on the $ scale. A development factor is a parameter.

Based on model (5.2), the random variable p(d) has a lognormal distribution with,

Median = $\exp[\alpha + \sum_{j=1}^{d} \gamma_j]$, (5.5)

Mean = mean $\cdot$ exp $[0.5\ \sigma^2]$, (5 6)

and

Standard

Deviation = mean $\cdot \sqrt{\exp[\sigma^2] - 1}$. (5.7)

Since, y(d) - y(d-1) ~ $N(\gamma_d , 2\sigma^2)$, we have

$$E\left[\frac{p(d)}{p(d-1)}\right] = \exp[\gamma_d + \sigma^2], \tag{5 8}$$

so that the development factor on the $ scale (the mean of a ratio) is given by the last equation.

The stochastic model for p(d) comprises a lognormal distribution for each development year where the medians of the lognormal distributions are related by equation (5.5) and the means are related by equation (5.6). So, in fitting or estimating the model (Section 8) we are essentially fitting a lognormal distribution to each development year. The curve (on a log scale) comprising straight line segments is only one component of the model. The principal component comprises the distributions.

As another example, we consider the stochastic CC model, viz.,

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j + \varepsilon \quad . \tag{5.9}$$

In this model we assume, for example, that y(1,0),...,y(s,0) are observations from a normal distribution with mean α and variance σ^2.

The assumptions contained in the model must be tested to ensure that they are not violated by the data.

The stochastic development factor family **(DFF)** is written as:

$$y(w,d) = \alpha_w + \sum_{j=1}^{d} \gamma_j + \sum_{t=2}^{w+d} \iota_t + \varepsilon \quad . \tag{5.10}$$

Note that the mean trend between cells (w,d-1) and (w,d) is $\gamma_d + \iota_{w+d}$ and the mean trend between cells (w,d) and (w+1,d) is $\alpha_{w+1} - \alpha_w + \iota_{w+1+d}$.

A model belonging to the DFF of (stochastic) models relates the lognormal distributions of the cells in the triangle. On a log scale the distribution for each cell is normal where the means of the normal distributions are related by the "trends" equation belonging to the family (4.7.1).

Another deficiency of the CL probabilistic model is that it contains the explicit assumption that the errors for the youngest accident year and the last development year are both zero. The chance of that, is zero!

We now return to the deterministic development factor model of Section 4.4.

To all the log "payments" in the triangle we add random numbers from a normal distribution with mean zero. Equivalently, to the trends depicted in Figure 4.4.1, we add random numbers from a normal distribution displayed in Appendix A5. The sum of trends (Appendix A2) plus random fluctuations (Appendix A5) is displayed in Appendix A6.

The graph of the first six accident years of the data in Appendix A6 is given in the Figure 5.2 below.

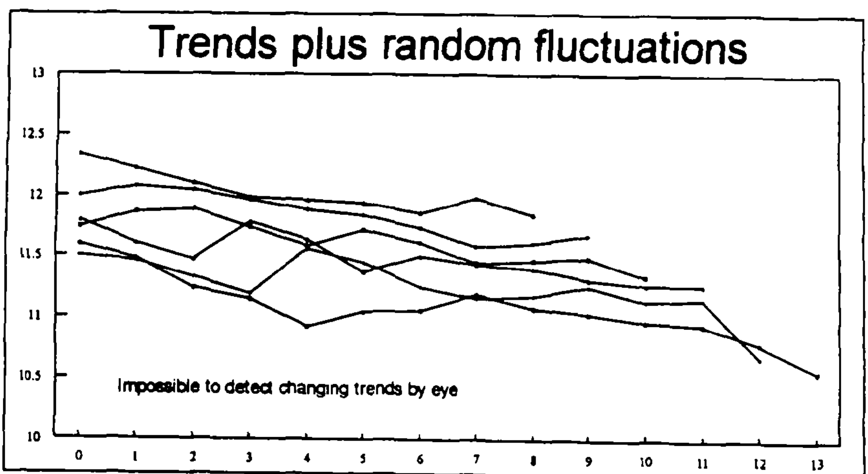


Figure 5.2

NOTE that it is impossible to determine the trends and/or change in trends by eye or from the age-to-age link ratios of the cumulative payments (Appendix A9). See Appendices A7 - A9. THE TRENDS CAN ONLY BE DETERMINED BY USING REGRESSION.

Notwithstanding the fact that the DFF modelling framework can be applied to any loss development array, much of the remainder of the discussion will involve analysis of the incremental payments for the following reasons:

* the geometry of trends;
* simplicity and parsimony;
* *distributions of future payments is relevant information for financial* statements.

Other reasons are given in Sections 10.3 and 10.4.

6.0 REGRESSION AS A FORM OF ADJUSTMENT AND MINIMUM TESTS

Hitherto we have applied regression for two related purposes. Estimation of trends in the 'payments' and estimation of the distribution of payments in each cell. The estimated trends relate the means of the distributions on a log scale.

For example, if the CC model is an appropriate model, then the 'payments' come from lognormal distributions and the means of the log 'payments' lie on the surface:

$$y(w,d) = \alpha + \sum_{j=1}^{d} \gamma_j \quad .$$

6.1 REGRESSION AS A FORM OF ADJUSTMENT

Regression is also a very powerful approach to adjusting data, especially in the loss reserving context.

In view of the fact that payment/calendar year trends project onto the other two directions, a graph of the data in one direction gives no indication of the trends. See for example, the simulated data with three payment year trends discussed in Section 5, and in particular, Figure 5.2.

We define a residual by

$$\hat{\varepsilon} = y - \hat{y} \quad .$$

That is, a residual is an observed value minus its fitted value.

Residuals can be interpreted as the data adjusted for what has been fitted. Let's consider a number of examples.

Suppose we simulate (generate) a triangle based on a CC model. The model generating the data can be written

CC DATA = CC TRENDS + ERROR (randomness)

If to the data we estimate the CC model, then the residual is

residual = CCDATA - FITTED CC TRENDS

= estimate of error,

that is, the residuals represent the data after we take away (subtract) what we fitted, alternatively, the residuals represent the data adjusted for what we fit. Here we subtract the estimates of the trends we used to create the data, so residuals should represent what is left, which is "randomness" in the three directions. "Random" residuals versus payment years are depicted in Figure 6.1.1.

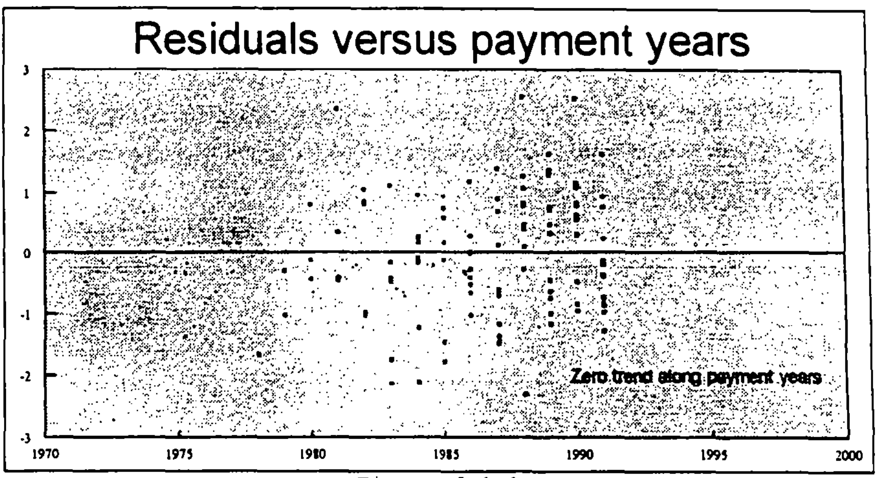


Figure 6.1.1

Suppose we now generate,

DATA = CC data + 10% calendar year trend

If we fit the CC model to this data the residual is

residual = DATA - fitted CC TRENDS

= estimate of error + 10% calendar year trend

So here residuals versus payment/calendar years will exhibit a straight upward trend (+ randomness) as depicted in Figure 6.1.2. After removing the CC trends from the data, there still remains the 10% calendar year trend plus the random fluctuation.

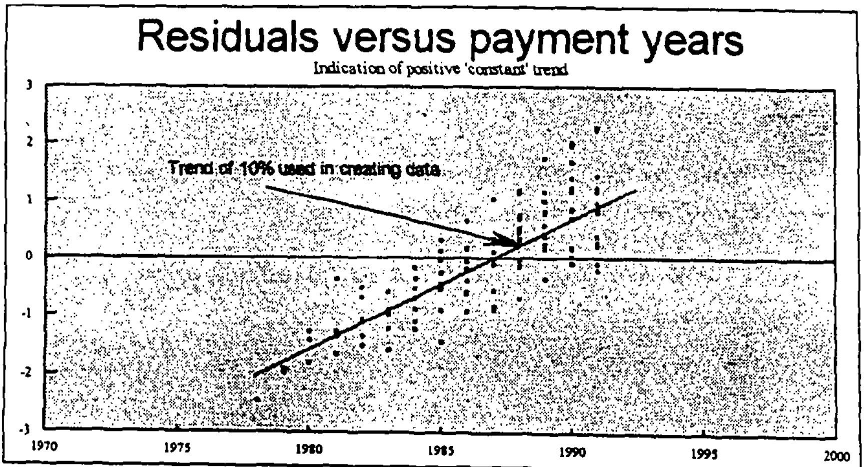


Figure 6.1.2

If you estimate the average trend in these residuals in a spreadsheet you would obtain an estimate of approximately 10% (the trend introduced into the data).

If we estimate the CCI model to the data, we are essentially estimating a trend parameter through the payment year residuals (Figure 6.1.2) of the previous CC model.

Now the residuals versus payment years should be random as we have removed (subtracted) all the (estimated) trends we introduced into the data.

Consider now data created as follows

DATA = CC data + 10% trend (calendar years 1978-85)
+ 20% trend (calendar years 1985-91)

If we fit the CC model to this data the residual is

residual = DATA - fitted CC TRENDS

= estimate of error + 10% (78-85)+15% (85-91)

The residuals versus payment/calendar years exhibit two trends, one from 1978-85 and sharper trend from 1985-91. See Figure 6.1.3 below.

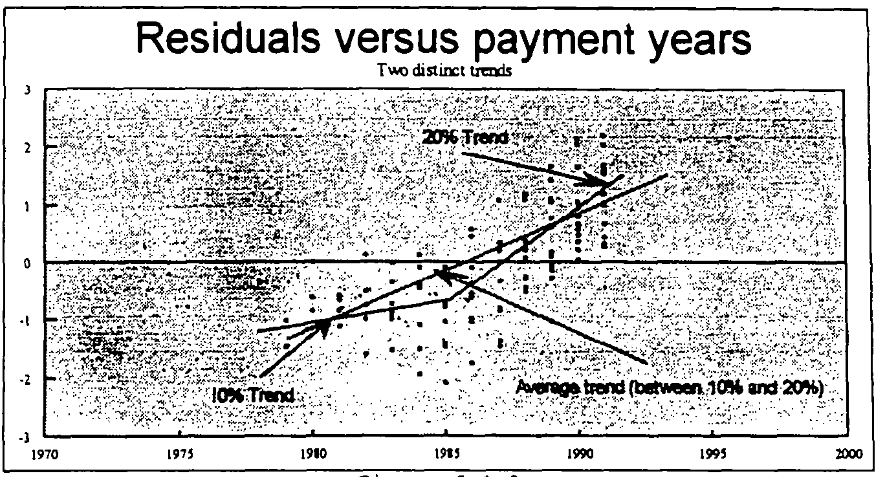


Figure 6.1.3

In now estimating the CCI model to the data, we are essentially estimating a trend parameter through the payment year residuals of Figure 6.1.3. The average trend is between 10% and 20%. The residuals versus payment years are now 'v-shaped'. See Figure 6.1.4 below.

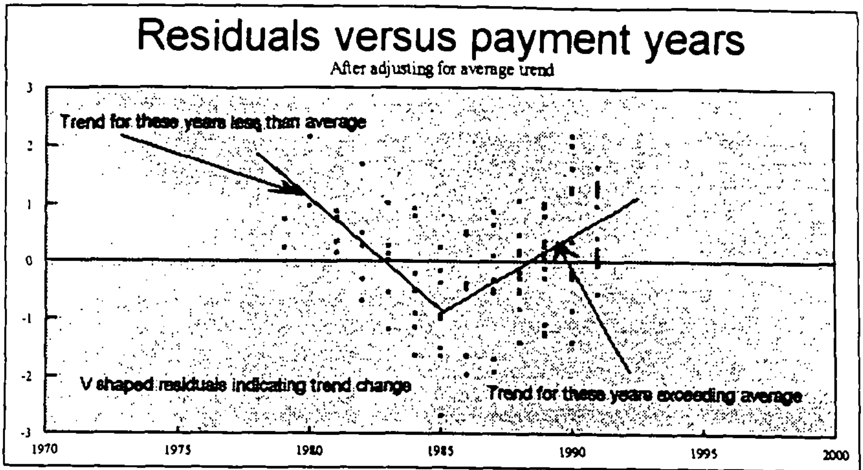


Figure 6.1.4

We are now led to estimate the two trends.

In view of the fact that calendar year trends project onto the other two directions, we can only obtain an indication of payment year trends, after we first remove the development year trends from the data (and vice versa).

REGRESSION IS A VERY POWERFUL TECHNIQUE FOR SEPARATING THE TRENDS IN THE THREE DIRECTIONS FROM RANDOM FLUCTUATIONS

In Section 12 we analyse a real life example that possesses relatively smooth age-to-age link ratios, yet there are major shifts in calendar year trends that are quite alarming.

6.2 MINIMUM TESTS

The author believes that a sound loss reserving statistical modelling framework should pass a number of very simple basic fundamental tests.

Turning to the univariate (log price) series of Section 3.1, if the (average) trend in the nominal prices is zero, that is, the prices are random about a zero trend then this feature in the data could be determined informally by examining the graph with eye and ruler and formally in a spreadsheet by estimating the trend, showing that it is insignificant and testing the residuals for randomness. Hence,

Test 1: If the (incremental) payments in a loss development array are random observations (from a lognormal distribution), and accordingly there are no trends in each of the three directions, then a sound loss reserving methodology should determine this.

We illustrate with an example. Appendix B1 contains incremental payments drawn at random from the same lognormal distribution. Note the variability. The mean forecast or fitted value for each cell is the same. Indeed, estimation of the CC model, for example, to the data would yield insignificant γ 's, as they should be. Application of the DFF modelling framework will allow us to identify the salient features of the data extremely fast.

The age-to-age link ratios are displayed in Appendix B3 and do not appear to convey much relevant information. (Compare with age-to-age link ratios in Appendix B5. What can you tell?)

For those readers who feel that random data (no trends) represents a pathological case, should analyse a number of Lloyd's Syndicates data.

Returning to the univariate series of Section 3.1, it is rather straightforward to identify both informally and formally the difference between the nominal prices and the adjusted prices. A second loss reserving test is suggested.

Test 2: Consider any real life incremental paid loss development array. Create from this array a second array by subjecting it to a number of trends, for example, a 10% trend (say) in the first five calendar years (say), and a 15% trend (say) in the subsequent calendar years, then a sound loss reserving methodology will allow for a quick determination of the simple difference between the two loss development arrays.

The DFF modelling framework passes Test 2 with flying colors. The reader will find that by applying Test 2 to standard age-to-age link ratio techniques they fail it. That is because standard techniques do not satisfy the necessary and simple property of additivity of trends.

In order to dispel the myth that smooth age-to-age link ratios imply stability of trends we analyse in Section 12 a real life array with smooth factors and find major trend instability that is quite alarming and in order to dispel the converse myth that rough age-to-age link ratios imply trend instability, we analyse in Section 13 a real life array with rough ratios and find stability so that had we used the same model estimated three years earlier, it would have accurately predicted the distributions for the last three calendar years and would have given the 'same' outstanding estimates.

To further illustrate the impact of randomness of payments on age-to-age link ratios, Appendix B4 contains an array generated by an SDF probabilistic model with constant 20% calendar year trend. The link ratios are presented in Appendix B5 and appear relatively rough. Yet, the same model estimated four years earlier would have predicted the distributions of the payments of the last four years and would have produced the 'same' completion of the rectangle!

It is interesting to also observe that even though the data in Appendix B4 has a 20% calendar year (and accident year) trend, as you step down a column (development year), sometimes the numbers decrease rather than increase (by 20%).

For example, (1989, 1) to (1991, 1) the payment reduces from 767664 to 350789. This is explained by the random fluctuations component of the model. Examine now Figure 3.1.1 and note that even though the mean trend in nominal prices is 23%, prices from one year to the next do not necessarily increase. This is due to the random fluctuations. So, the same phenomenon applies to loss reserving data.

Consider now the unusual large value of 1317425 corresponding to (1985,6). It is not unusual. It comes from the tail of the lognormal distribution. Given that the lognormal is skewed to the right, values greater than the median tend to be 'far' from the median, whereas values less than the median tend to be relatively close to the median.

7.0 VARYING PARAMETER, DYNAMIC OR CREDIBILITY MODELS

7.1 MULTICOLLINEARITY

Many of the models within the family (5.10) cannot be estimated in a spreadsheet or any statistical package. Models that contain "many" iotas, alphas and gammas suffer from a problem known as multicollinearity. This problem is explained as follows.

To estimate the Ordinary Least Squares line for the simple linear regression:

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad , \qquad (7.1.1)$$

we estimate the intercept α and slope β by minimising the error sum of squares,

$$SS = \sum (y_i - \alpha - \beta x_i)^2 \quad .$$

Taking partial derivatives of the last equation with respect to α and β, and setting them to zero we obtain:

$$-2\sum (y_i - \alpha - \beta x_i) \quad = \quad 0 \qquad (7.1.2)$$

and

$$-2\sum x_i (y_i - \alpha - \beta x_i) \quad = \quad 0 \qquad (7.1.3)$$

Equivalently,

$$\bar{y} - \alpha - \beta \bar{x} \quad = \quad 0 \qquad (7.1.4)$$

and

$$\sum x_i\, y_i - n\alpha\bar{x} - \beta\sum x_i^2 \quad = \quad 0 \; . \qquad (7.1.5)$$

The two linear homogeneous equations are known as the normal equations and their solution yields the least squares estimates of α and β.

For a model having P parameters in the DFF family, a spreadsheet (or a statistical package) sets up P linear homogeneous equations in order to solve for P unknowns. However, as a result of the non-orthogonality of the payment year direction with the other two directions, some of the equations in the normal equations are redundant, e.g.,

$$\alpha - \beta \quad = \quad 2$$

and

$$2\alpha - 2\beta \quad = \quad 4$$

So, there is no unique solution.

If there are some equations that are almost redundant, e.g.,

$$\alpha - \beta \quad = \quad 2$$

and

$$2\alpha - 2\beta \quad = \quad 4.00001,$$

then the estimates will have high standard errors, so that the resulting model will be unstable.

7.2 OVERCOMING MULTICOLLINEARITY

The phenomenon of multicollinearity associated with fixed parameter models can be interpreted in terms of information. There is not sufficient information in one loss development array to estimate many payment year parameters and accident year parameters (especially, for more recent accident years). Another interpretation is that the independent variables in the regression are not really independent. We showed in Section 3 that calendar year trends are related to development year trends and accident year trends.

If we include another α parameter for the last accident in our model we are using one single datum to estimate that parameter. That is, we assign full credibility to the last accident year's datum and zero credibility to previous years in respect of the estimation of the additional α parameter. A better approach may be to assign some credibility to the previous years data and less than full credibility to the last year's datum.

We are motivated to introduce exponential smoothing/varying parameter/credibility models, as a result of multicollinearity. Multicollinearity can lead to fixed parameter regression models that (i) are unstable and (ii) have large prediction errors.

The technique of exponential smoothing has received widespread use in the context of forecasting a time series. It originated more than 40 years ago without any reference to an underlying model that makes the technique optimal.

We first present heuristic arguments for exponential smoothing and varying parameter models. The following illustrations and arguments may be viewed from two different perspectives. The data may be regarded as either

(1) sales data over time, or

(2) incremental paid losses for delay 0 across accident years.

(i) Constant mean level (one parameter)

Suppose we have a sequence of time series observations $y_1, y_2, \ldots y_n$ such that

$$y_t = \alpha + \varepsilon_t \quad , t = 1, \ldots, n$$

where α is a constant mean level and ε_t is a sequence of uncorellated errors with constant variance. Figure 7.2.1 below depicts such a series.

The model describing the data is the simplest regression model.

Our model has only one parameter. so that the years are completely homogeneous (stable!).

If α is known, the best forecast of a future observation $y_{(n)+1}$, based on information up to time n, is

$$\hat{y}_{(n)+1} = a \; .$$

If the parameter α is unknown, we estimate it from the past data $(y_1 \ldots y_n)$ by its ordinary least squares estimate,

$$\hat{\alpha} = \Sigma y_i/n,$$

so that the one-step-ahead forecast of $y_{(n)+1}$ is now

$$\hat{y}_{(n)+1} = \bar{y} .$$

We can now write,

$$\hat{y}_{(n+1)+1} = \hat{y}_{(n)+1} + \frac{(y_{n+1} - \hat{y}_{(n)+1})}{n+1}$$

The last equation indicates how a forecast from time origin n+1 can be expressed as a linear combination of the forecast from time origin n and the most recent observation. This is the simplest credibility formula, due to Gauss [8], used when updating sample averages. Since the mean level α is assumed constant, each observation contributes equally to the forecast.

The above formula for updating sample averages is an experience rating (credibility) formula in the context of adjusting a premium, assuming the risk (parameter) does not change from year to year.

In computing $\hat{\alpha}$ $(=\bar{y})$ we assign the same weight to each observation. From the loss reserving perspective, we are assuming that the accident years are completely homogeneous. In order to estimate the next years premium, we use all the accident years' data!

We now turn to another example.

(ii) Unstable mean level (each year its own parameter)

Here,

$$y_t = \alpha_t + \varepsilon_t$$

where the mean level α_t changes dramatically in successive time periods. Each year t has its own parameter α_t. Figure 7.2.2 depicts a series of y_t values that may be generated by this model.

Figure 7.2.2

Here, the best we could do, is forecast $y_{(n)+1}$ by

$$\hat{y}_{(n)+1} = y_n .$$

We are assigning zero weight to the past and full weight to the current observation. From the loss reserving perspective, accident years are completely heterogeneous, so that each accident year's individual parameter is estimated by that year's individual experience.

(iii) Locally constant mean level exponential smoothing and credibility

Often situations present themselves where the mean is approximately constant locally. Assigning equal weights to the past would be too restrictive and assigning zero weight would result in **loss of information**. It would be more reasonable to choose weights that decrease (geometrically) with the age of the observations.

We could have

$$\hat{y}_{(n)+1} = Ky_n + K(1-K)y_{n-1} + K(1-K)^2 y_{n-2} + \ldots$$

For n sufficiently large this may be written

$$\hat{y}_{(n)+1} = \hat{y}_{(n-1)+1} + K(y_n - \hat{y}_{(n-1)+1})$$

$$= (1-K)\hat{y}_{(n-1)+1} + Ky_n . \qquad (7.2.1)$$

This is also a credibility formula.

Muth [12] showed that the exponential smoothing formula (7.2.1) is an optimal forecast for the following model:

$$y_t = \alpha_t + \varepsilon_t : Var[\varepsilon_t] = \sigma^2_\varepsilon$$

$$\alpha_t = \alpha_{t-1} + \eta_t : Var[\eta_t] = \sigma^2_\eta \qquad (7.2.2)$$

Here the mean level α_t process is a random walk. If $\sigma^2_\eta = 0$, then we have the constant mean level situation (i) and if σ_η^2 is large we have the unstable mean level situation (ii). The parameter σ_η^2 should be chosen as small as possible at the same time ensuring that the trend in the data is captured.

Choosing σ_η^2 (relative to σ_ε^2) that minimises the SSPE yields the maximum likelihood estimates of σ_η^2.

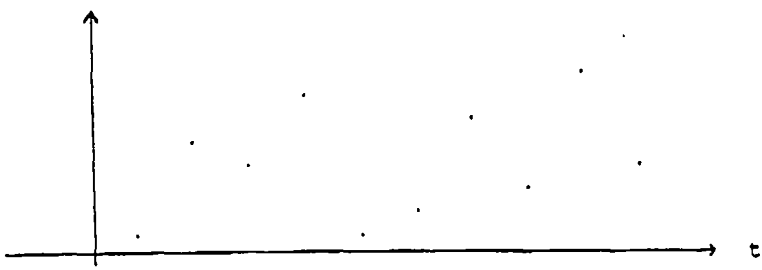


Figure 7.2.3

The exponential smoothing formula (7.2.1) formally credibility weights all the observations. It is an experience rating formula for a risk (parameter) that changes. If in the situation depicted in Figure 7.2.3, one were to assign zero weight to the past in place of using formula (7.2.1), then much information would be potentially lost.

We illustrate the methodology of formula (7.2.1) in the loss reserving context.

Suppose, for the sake of argument, there are only two accident years (but more than three development years), and the γ and ι parameters are zero.

We have,

$$y(1,d) = \alpha_1 + \varepsilon(1,d);\ d=0,1,2,\ldots,n_1-1 \text{(say)} \qquad (7.2\ 3)$$

and

$$y(2,d) = \alpha_2 + \varepsilon(2,d);\ d=0,1,2,\ldots,n_2-1 \text{(say)} \qquad (7.2.4)$$

The first accident year has n_1 observations and the second n_2 observations. Denote the sigma-squared assigned to observations by σ^2. Accordingly, Var[ε (1,d)] = Var[ε (2,d)] = σ^2.

The relation between α_2 and α_1 is given by

$$\alpha_2 = \alpha_1 + \eta : \text{Variance}(\ \eta\ ) = \sigma^2_\eta\ . \qquad (7.2.5)$$

Substituting equation (4.4) for α_1 into (4.3) yields:

$$y(2,d) = \alpha_1 + \eta + \varepsilon(2,d)\ . \qquad (7.2.6)$$

Combining the last equation with (4.2) we have,

$$y(1,d) = \alpha_1 + \varepsilon\ (1,d)$$

with (7.2.7)

$$y(2,d) = \alpha_1 + \eta + \varepsilon\ (2,d)$$

Since, conditional on α_1 the observations y(2,0), y(2,1), ... are correlated, we reduce by sufficiency to obtain:

$$\bar{y}_1 = \alpha_1 + \varepsilon_1$$

and

$$\bar{y}_2 = \alpha_1 + \varepsilon_2$$

where $$Var[\varepsilon_1] = \sigma^2/n_1 \quad , \quad Var[\varepsilon_2] = \sigma^2/n_2 + \sigma^2_\eta$$

and $$\bar{y}_1 = \sum_{d=0}^{n_1-1} y(1,d)/n_1 \quad , \quad \bar{y}_2 = \sum_{d=0}^{n_2-1} y(2,d)/n_2 \ .$$

The estimate of α_1 minimises the weighted error sum of squares

$$w_1(\bar{y}_1 - \alpha_2)^2 + w_2(\bar{y}_2 - \alpha_1)^2 \ ,$$

where

$$w_1^{-1} = Var[\varepsilon_1] = \sigma^2/n_1 \ ,$$

and

$$w_2^{-1} = Var[\varepsilon_2] = \sigma^2/n_2 - \sigma^2{}_\eta \ .$$

/

Similarly, the estimate of α_2 is obtained by minimising,

$$w_1(\bar{y}_2 - \alpha_2)^2 + w_2(\bar{y}_1 - \alpha_2)^2 \quad ,$$

where now $\quad w_1^{-1} = \sigma^2/n_2 \quad$ and $\quad w_2^{-1} = \sigma^2/n_1 + \sigma_\eta^2$

The estimates of $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are given by respectively,

$$\hat{\alpha}_1 = (1 - z_1)\bar{y}_2 + z_1\bar{y}_1$$

and

$$\hat{\alpha}_2 = (1 - z_2)\bar{y}_1 + z_2\bar{y}_2$$

where,

$$z_1 = \frac{\dfrac{n_1}{\sigma^2}}{\dfrac{n_1}{\sigma^2} + \dfrac{n_2}{\sigma^2 + n_2\,\sigma_\eta^2}} \quad \text{, and} \quad z_2 = \frac{\dfrac{n_2}{\sigma^2}}{\dfrac{n_2}{\sigma^2} + \dfrac{n_1}{\sigma^2 + n_1\sigma_\eta^2}}$$

Both $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are credibility estimators.

The smaller σ_η^2 is (relative to σ^2), the more information is being pooled across the two years in estimating α_1 and α_2. We are credibility weighting the two years' data.

For a description of general recursive credibility formulae, see Zehnwirth [14].

We conclude this section by remarking that even in the absence of multicollinearity, varying parameter models are more stable and validate better than the 'corresponding' fixed parameter regression models. Moreover, according to A.C. Harvey's [9] modern book on forecasting, explanatory variables are "proxied by a stochastic trend".

8.0 PARAMETER ESTIMATION AND FORECASTING OF DISTRIBUTIONS

In the present section we describe how the (fixed parameter) regression models may be set up in a spreadsheet (or a statistical package) for the twofold purpose of estimating the model parameters and forecasting the distributions of future (incremental) payments.

A practical illustration of this procedure for the chain ladder statistical model is given by Christofies [4] in the second volume of the Institute of Actuaries Loss Reserving Manual [11].

8.1 ESTIMATION

In order to estimate a regression model in a spreadsheet we need to create, corresponding to each dependant observation y, the values of the (row) design vector containing the values of the independent variables.

Let y(w,d) = log p(w,d) and let β' be a row vector holding the parameters of the model, that is,

$$\beta' = (\alpha_1, \alpha_2, \ldots, \alpha_k, \quad \gamma_1, \ldots, \gamma_p \quad \iota_1, \ldots, \iota_m) \quad .$$

The model has (i) k distinct α parameters where α_1 represents the level for accident years 1,2. ., w_1 (say); α_2 represents the level of accident years w_1 + 1, ... w_2 (say), and so on, (ii) l distinct γ parameters where γ_1 is trend along development years 0, 1, ..., d_1; γ_2 is trend along development years d_1, $d_1 + 1$, ... d_2 and so on and (iii) m distinct iota parameters where ι_1 represents the trend along payment years 0,1,2,...;t_1; ι_2 represents the trend along payment years t_1, ..., t_2, and so on.

The arguments k, l and m may take the value 0.

The corresponding design vector is

$$\underline{x}'(w,d) = (\delta_{11}, \delta_{12}, .., \delta_{1k}, \quad \delta_{21}, ..., \delta_{2l}, \delta_{31}, ... \delta_{3m})$$

where each δ is a variable defined as follows

$$\delta_{1j} = 1 \text{ if } w_{j-1} + 1 \leq w \leq w_j \; (w_0 = 1)$$

$$= 0, \quad \text{otherwise} \quad ;$$

$$\delta_{21} = 1$$

and $\delta_{2j} = d - d_{j-1}$, If $d \geq d_{j-1} + 1$ ($j \geq 2$)

$$= 0 \quad , \quad \text{otherwise} \quad ;$$

and

$$\delta_{3j} = w + d - t_{j-1}, \text{ if } w + d \geq t_{j-1}$$

$$= 0 \quad , \quad \text{otherwise.}$$

We now stack the y observations to form a column vector

$$\underline{y} = (y(1,0), \ldots\ y(1,s-1),\ y(2,1),\ \ldots,\ y\ (2,s-2),\ \ldots,\ \ldots\ y(s,0)) \quad .$$

and corresponding design vectors to form a design matrix,

$$X = (\underline{x}'(1,0),\ldots\ldots,\underline{x}'(s,0)) \quad .$$

The observation equation can now be written

$$\underline{y} = X\underline{\beta} + \underline{\varepsilon} \quad ,$$

where $\underline{\varepsilon}$ contains independent errors from a normal distribution with mean zero and variance σ^2.

To estimate a DFF model in a spreadsheet, one needs to specify the column vector $\underline{y}$ and the columns of X as the independent variables.

The spreadsheet will create $\hat{\underline{\beta}}$, the ordinary least squares estimator of $\underline{\beta}$, and some other statistics including R^2, S^2 and standard errors of parameters.

The estimate of the variance - covariance matrix of $\hat{\underline{\beta}}$ is given by

$$V(\hat{\underline{\beta}}) = \hat{\sigma}^2(X'X)^{-1} \quad .$$

Some statistical packages such as MINITAB will produce the variance - covariance matrix as explicit output. Residuals and standardised residuals are straightforward to compute.

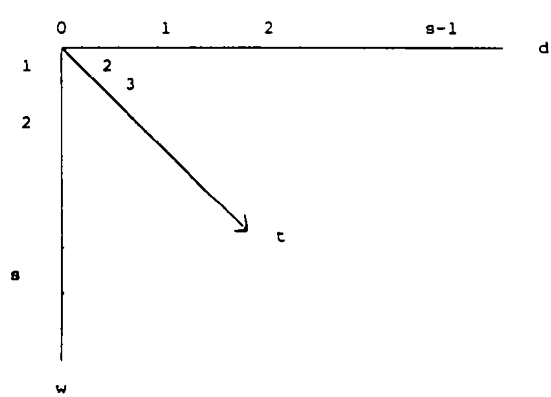
A lucid exposition of multivariate regression theory is given in Chatterjee and Price [3].

8.2 FORECASTING (PREDICTION) OF DISTRIBUTIONS

We have stressed repeatedly that a regression model is a probabilistic model and that the models contained in our rich DFF framework relate the normal distributions of the log payments of the cells in the loss development array by (trend) parameters.

We now would like to obtain estimates of normal distributions for payment years exceeding s.

That is, for calendar years beyond the evaluation year.



Consider a cell (w,d) for which $w+d>s$ and $d \leq s-1$.

Suppose we assume that the mean trend along payment years $\geq$ s is $\hat{\iota}_s$, the estimate of trend from payment year s-1 to s. (If ι_s is not a parameter in the model then $\hat{\iota}_s = 0$). We also assume that the standard deviation of the trend is se ($\hat{\iota}_s$), the standard deviation of the estimate. We stress emphatically that the larger se ($\hat{\iota}_s$) is, the mean trend $\hat{\iota}_s$ being the same, the larger the (mean) payments.

The vector of parameter estimates now contains the $\hat{\alpha}'$s, $\hat{\gamma}'$s but only one iota estimate, viz, $\hat{\iota}_s$.

The (design) independent value in the design vector $\underline{x}'(w,d)$ corresponding to $\hat{\iota}_s$ is now (w+d-s) = number of payment years from s to w+d. The other parameters contain the same design elements as in the estimation stage. The forecast $\hat{y}$ of y corresponding to cell (w,d) is given by:

$$\hat{y}\,(w,d) \quad = \quad \underline{x}'\,(w,d)\;\hat{\underline{\beta}} \quad .$$

We can now stack all forecasts $\hat{y}$ into a vector $\hat{\underline{Y}}$ and design vectors $\underline{x}'$ into a matrix X.

The estimate of the variance - covariance matrix of $\hat{\underline{Y}}$ is

$$V\,(\hat{\underline{Y}}) \qquad = X'\;V\,(\hat{\underline{\beta}})\;X + \hat{\sigma}^2\;I \quad ,$$

where I is the identity matrix.

The quantity $\hat{\sigma}^2$ is the estimate of the process variance (uncertainty), whereas

$$X' \, V(\hat{\underline{\beta}}) X$$

is a function of the variance of $\hat{\underline{\beta}}$, representing the parameter uncertainty.

Since $V(\hat{\underline{\beta}})$ is a function of $\hat{\sigma}^2$, the estimates of parameter uncertainty and process uncertainty are related. Quite often the smaller σ^2 is (relatively speaking), *the smaller the parameter uncertainty.*

Using Fisher's fiducial approach we can argue that our forecast for the distribution of y(w,d) is normal with mean $\hat{y}(w,d)$ and variance $V(\hat{y}(w,d))$, the diagonal element of $V(\hat{\underline{y}})$ corresponding to y(w,d).

Indeed, $\underline{y}$ has a multivariate normal distribution with mean $\hat{\underline{y}}$ and variance covariance $V(\hat{\underline{y}})$.

So, by applying standard regression theory we can compute our estimate of the multivariate normal distribution of the y values in the lower right of the rectangle.

Each estimate $\hat{y}$ of the corresponding y variable is best in the sense that it minimises the mean square error.

$$E[\,(y - f(\mathbf{y}))^2\,],$$

over all statistics f(.), where f(.) is a function of the data $\mathbf{y}$.

In order to obtain the distributions (multivariate) of the (incremental) payments and accident year and payment year sums, we employ the relationship between the multivariate lognormal and the multivariate normal distributions and standard statistical theory involving variances of sums. The means of the lognormal distributions are best estimates of the corresponding incremental payments.

We remark that our forecast distributions can also be argued for from a Bayesian viewpoint. The forecasts are Bayes with respect to a noninformative prior.

The reader will appreciate that to write a macro in a spreadsheet for a particular model in the modelling framework would be extremely prohibitive in terms of time, let alone writing a macro for each model

For readers that are interested, the author can make available a Lotus worksheet containing some of the models discussed in the real life study of Section 13.

9.0 MODELLING CONCEPTS

9.1 INTRODUCTION

The mechanisms by which claim severities, frequencies and delays are generated are invariably complex. When a model is constructed, it is not intended to be an accurate description of every aspect of the claims processes. The aim is to simplify the underlying processes in such a way that the essential features are brought out. According to Milton Friedman [7]: *'A hypothesis is important if it 'explains' much by little...'*. Similar views are expressed by Popper [13]: *'Simple statements... are to be prized more highly than less simple ones because they tell us more; because their empirical content is greater, and because they are better testable.'*

The "essential features" of the data in the loss reserving context are the trends and the random fluctuations about the trends. We decompose the data thus:

Log 'payments' = Trends + Random Fluctuations

Another way of thinking of this statistical model is to regard the Trends as a mathematical description of the main features of the data and the Random Fluctuations (or error or noise component) as all those characteristics not 'explained' by the Trends. All the complex mechanisms involved in generating the data are implicitly included in the model as creating the Trends plus the residual variance in the Random Fluctuations. See also Section 7 on varying parameter models.

The final identified model that 'explains' the data does not represent explicitly the underlying generating process. The model has probabilistic properties for which the data may be regarded as a sample (path) from it. Another classical modelling example in insurance where the same kind of modelling concepts are used is when we fit a Pareto distribution, say, to loss sizes. We do not assume that the Pareto distribution represents the underlying generating process. Whatever is driving the claims is very complex and depends on many factors. All we are saying is that our experience (sample) can be regarded as a random sample from the estimated Pareto distribution. The estimated Pareto distribution describes the variability in the loss sizes.

By way of summary, in order to take account of variables (or factors) not included in the Trends, we consider probabilistic models. See also Section 7 on varying parameter models.

There are a number of criteria for a good model with high predictive power:

* Ockham's Razor - parsimony;
* goodness of 'fit';
* validation and stability.

9.2 OCKHAM'S RAZOR - PARSIMONY

Ockham's razor, also known as the principle of parsimony, says that in a choice among competing models, other things being equal, the simplest is preferable. Accordingly, a parsimonious model that provides a description of the salient features of the data may be preferable to a complicated one for which the residual variance in the error is smaller (and so R-squared is larger). See also Section 10.4.

We stress R-squared (or adjusted R-squared) does not measure the predictive power of a model.

Consider two data generating models, Model 1 is,

$$y_t = \mu + \varepsilon_t \quad , \qquad (9.2.1)$$

where $\varepsilon_t \sim N(o, \sigma^2)$ and the signal to noise ratio μ/σ^2 is large. Here, R-squared = 0 and since σ^2 is "small" predictions based on samples from this model will be relatively accurate.

For Model 2,

$$y_t = \alpha + \beta t + \varepsilon_t \quad , \tag{9.2.2}$$

where $\epsilon_t \sim N(0, \sigma^2)$. Suppose σ^2 is relatively large and R-squared is 85%. Predictions based on samples from this model will have larger errors than predictions in the first model. The forecasting errors are not a function of R-squared.

The consequences of adopting an inappropriate model will depend on its relationship to the 'true' model.

Underparametrisation - it imposes invalid constraints on the 'true' model.

Overparametrisation - the model is more general than is necessary.

Overparametrisation has different consequences to underparametrisation. Overparametrisation leads to high errors of prediction. The forecasts are extremely sensitive to the random component (in contrast to the trends) in the observations. Indeed, overfitting can be disastrous in certain circumstances. Overfitting a model is equivalent to including randomness as part of the (systematic) trend (component). Underparametrisation, on the other hand, tends to lead to bias rather than instability.

The dangers of overparametrisation are illustrated with a simple example. Imagine we have some yearly sales figures, as depicted below in Figure 9.2.1, and generated by

$$Y_t = 1 + 2t + 3t^2 + \varepsilon_t \ ,$$

say, where the ε_t's are random from $N(0,\sigma^2)$, and Y_t represents the number of sales in year t.

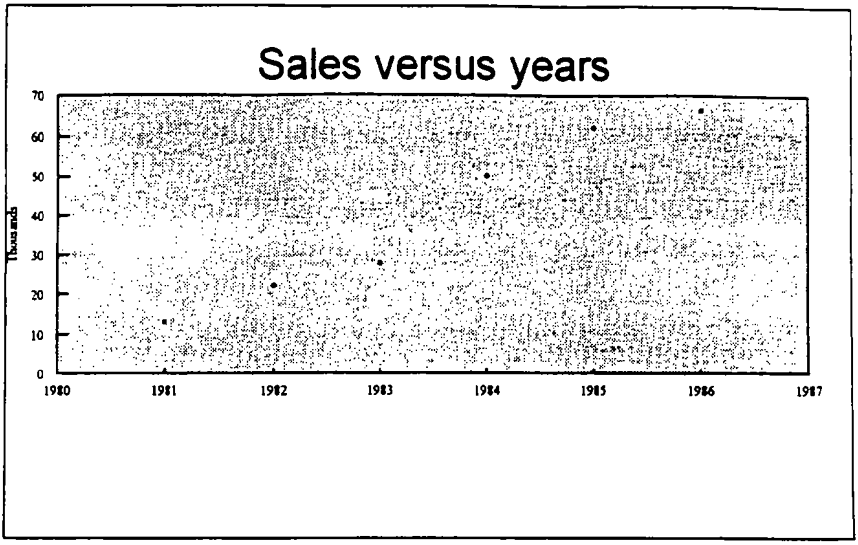


Figure 9.2.1

We wish to forecast sales for 1987. We could estimate a straight line model:

$$Y_t \quad = \quad \beta_0 + \beta_1 * t + \varepsilon_t \quad . \qquad (9.2.3)$$

This model produces residuals that are not random and is therefore rejected. The quadratic model,

$$Y_t \quad = \quad \beta_0 + \beta_1 * t + \beta_2 * t^2 + \varepsilon_t \quad , \qquad (9.2.4)$$

on the other hand, produces residuals that appear random. Moreover, R-squared is higher and parameters are significant.

We could try a fifth degree polynomial, viz.,

$$Y_t \quad = \quad \beta_0 + \beta_1 * t + \beta_2 * t^2 + \ldots . + \beta_5 * t^5 + \varepsilon_t \quad . \qquad (9.2.5)$$

This model will produce zero residuals, that is, it will go through every data point and the R^2 = 100%. However, it is <u>useless</u> from the point of view of forecasting. Why? If we change only one data point marginally, the forecast will change to a very large degree. Moreover, if we use the model at year end 1986 to forecast sales for 1988, re-estimate the model at year end 1987 to update our forecast for 1988, the two forecasts would be completely different. The data are **NOT** unstable. **IT IS THE MODEL THAT IS UNSTABLE**. The model is incredibly sensitive to the random component in the data. It should only be sensitive to the systematic trend. Incidentally, standard techniques based on calculation of age-to-age link ratios suffer from the same defect.

9.3 AKAIKE INFORMATION CRITERION AND INFORMATION

It has been emphasised that in comparing the goodness of 'fit' of various models, an appropriate allowance should be made for parsimony. This has a good deal of appeal, especially where the model may be based primarily on pragmatic considerations.

Akaike Information Criterion (AIC) is both a function of S^2 and the number of parameters in the model. It is an information theoretic criterion that can be used for discriminating between any two models, even if they are non-nested. It originated with the work of Akaike.

In general the AIC is given by

$$AIC = -2\log(\text{likelihood}) + 2P$$

For DFF models it reduces to

$$AIC = N\log[2\prod S^2(MLE)] - N + 2P,$$

where

(i) N = Number of observations,

(ii) s^2 (MLE) is the maximum likelihood estimator of σ^2 ,

and (iii) P denotes the number of parameters.

The aim is to select a model with a minimum (relative) AIC. Note that the AIC can be used to discriminate between any two models, irrespective of whether they have any parameters in common.

9.4 RECURSIVE RESIDUALS AND SSPE

Consider a time series $z_1, z_2, \ldots, z_n$ where $\hat{z}_{t+1}(t)$ denotes a forecast of z_{t+1} based on the data $z_1, z_2, \ldots, z_t$. That is, the forecast is based on the information up to time t only. The one-step-ahead forecast (prediction) error is given by

$$\hat{\varepsilon}_t(1) = Z_{t+1} - \hat{Z}_{t+1}(t) \quad .$$

The notation $\hat{\varepsilon}_t(1)$ expresses the fact that it is the one-step-ahead prediction error that is calculated from past data up to and including time t. The estimates of the parameters of the model are only based on the data $Z_1, Z_2, \ldots, Z_t$.

In order to compute the errors $\{\hat{\varepsilon}_t(1)\}$ the model has to be estimated many times.

The sum of the squared one-step-ahead prediction errors, denoted by SSPE, is given by

$$\text{SSPE} = \sum_{t=t_0}^{n} \hat{\varepsilon}^2_t(1) \quad .$$

The time t_0 is chosen so that it exceeds the maximum number of parameters amongst the models being considered; by at least one.

Computation of the SSPE may take much time even with a good spreadsheet program, as the model has to be estimated for sub-samples, $\{Z_1,..., Z_t\}$; $t=t_0$, $t_0+1,...,n-1$.

Readers familiar with exponential smoothing will note that the optimal smoothing constant of exponential smoothing is determined by minimising the SSPE. See Abraham and Ledholter [1] for a lucid exposition of exponential smoothing.

By way of summary of the quality of 'fit' statistics, consider the quadratic polynomial example of Section 9.2, and suppose there are at least twenty data points. The relative magnitudes of R^2, AIC and SSPE as we fit polynomials of order one to six (say) are:

. R^2 increases with more parameters;

. AIC decreases from polynomial of order one to polynomial of order two, subsequently increasing as degree of polynomial increases (for most samples);

. SSPE behaves in much the same way as AIC.

Accordingly, a polynomial of degree exceeding two would have performed worse in a forecasting context than a polynomial of degree two, had we used them each year.

A relatively 'low' SSPE is preferable to a high SSPE. Naturally, there are other aspects of testing, including significance of parameters, distributional assumptions, residual displays and the number of parameters.

The 'tests' should be seen as complementary rather than competitive.

/

9.5 OUTLIERS, SYMMETRIC DISTRIBUTIONS AND NORMALITY

Outliers are data points with large standardised residuals. Observations classified as outliers have residuals that are large relative to the residuals for the remainder of the observations.

Estimates of parameters and supporting summary statistics may be sensitive to outliers. Residual displays provide information on outliers. Moreover, if omission of outliers from the regression affects the output, then that provides more evidence that the omitted observations are in fact outliers.

An outlier may be a result of a coding error, in which case it should be assigned zero weight, or it may be a genuine observation that is unusual and accordingly has a large influence on the estimates, unless it is assigned reduced weight.

To detect outliers routinely, we need a rule of thumb that can be used to identify them. A **Box plot** is a schematic plot devised by J.W. Tukey. The following steps summarise the general procedure for constructing a Box plot.

. Order the data.

. Find the median (M), lower quartile (LQ), upper quartile (UQ) and mid-spread (MS), where MS = UQ - LQ. [1]

. Find the upper and lower boundaries defined by

LB = LQ - 1.5*MS

UB = LQ + 1.5*MS.

---

[1] Footnote: LQ and UQ are actually the lower and upper hinges. They are only approximately the quartiles.

List all outliers. An outlier is defined as any observation above the upper boundary or below the lower boundary.

Construct a Box plot as follows:

(a) Draw a horizontal scale;

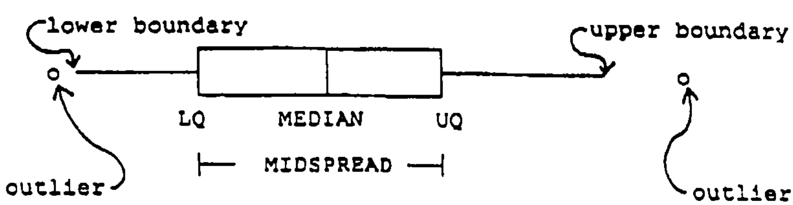
(b) Mark the position of the median using " | ";

(c) Draw a rectangular box around the median, with the right side of the box corresponding to the UQ and the left side corresponding to the LQ. The length of the box is equal to the MS. The median divides the box into two boxes;

(d) Find the largest and smallest observations between the boundaries and draw straight horizontal lines from the UQ to the largest observation below the upper boundary and from the LQ to the smallest observation above the lower boundary;

(e) Mark all observations (outliers) outside the boundaries with hollow circles (o). If an outlier is repeated, mark the number of times it is repeated.

Box Plot



We can also conclude (diagnostically) that a distribution is symmetric if the median is approximately half way between the LQ and the UQ.

A DFF model assumes that the weighted standardised residuals come from a normal distribution. Accordingly a normal probability plot should appear approximately linear. That is, the plot of weighted residuals against normal scores should have points that fall close to a straight line. This means that the correlation should be close to unity.

9.6 VALIDATION AND STABILITY

The important question is whether the estimated model can predict outside the sample. It is therefore important to retain a subset (the most recent one or two payment years) of observations for post-sample predictive testing. This post-sample prediction testing is called **VALIDATION**.

VALIDATION of the last payment year, or any payment year, is also related to the concept of **STABILITY**. If we don't use the last payment years' data to estimate the model, the ultimate losses should not differ from that obtained by using the last years' data by more than one standard error. We would like to identify a model that delivers **STABILITY** of reserves from year to year (only if trends are stable).

9.6.1 VALIDATION

Consider the triangle of incremental paid losses depicted below.

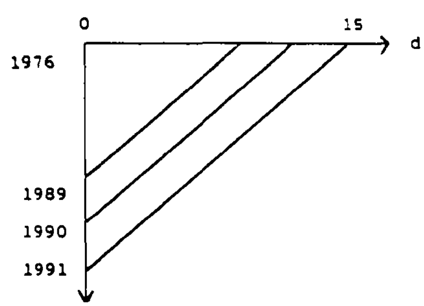


Figure 9.6.1.1

We have model that has been identified and estimated using all the data, up to 1991.

If the same model were estimated at year end 1988, would it predict accurately the incremental payments for 1989, 1990 and 1991? And what do we mean by 'predict accurately'?

Let's illustrate with a fair coin. If a fair coin is to be tossed 100 times we can 'predict accurately' the distribution of the number of heads. The exact distribution is Binomial (100, 0.5). The distribution details the probabilities of all the possible outcomes. If instead, we had a mutilated coin and we required a future prediction based on a sample data then our predicted distribution is Binomial (100, $\hat{p}$) where $\hat{p}$ is an optimal estimate of the true probability p of a head occurring, based on the sample.

We now return to our triangle. At year end 1988, we would estimate the parameters of the same model using the smaller sample and we would predict a distribution for each of the log 'payments' in 1989, 1990 and 1991. See Section 8.2 on forecasting of distributions.

So, one of the most important validation tests is to determine whether the observed log 'payments' in 1989, 1990 and 1991 can be regarded as a sample from the predicted distributions.

More specifically, let $\hat{y}$ be a prediction of a log 'payment' y for a cell in payment year 1989, 1990 or 1991. We call,

$$\tilde{\varepsilon} = y - \hat{y} \quad ,$$

the validated residual or the prediction error.

We test the validated residuals for (i) randomness in the three directions **delay**, **accident year** and **payment year**; (ii) randomness versus predicted values $\hat{y}$ and (iii) most importantly, normality.

9.6.2 STABILITY

Returning to our example of the foregoing section, we ask the question whether at year end 1988 our completion of the rectangle should be materially different to our completion at year end 1991. The answer is in the negative if trends (especially in the payment year direction) are stable.

We illustrate with four examples. (There are numerous others that occur in practice.)

Example 1: Suppose payment year trends (after adjusting for trends in the other two directions) are as depicted in Figure 9.6.2.1 below. The trend is stable and suppose its estimate is 10% ± 2%. How do we know that the trend is stable? Well, as we remove the more recent payment years from the estimation, the estimates of trends do not change (significantly). For example, after removing 1990 and 1991, the estimate of trend is 9.5% ± 2.1%, say. Alternatively, we could estimate a new trend parameter from 1989-1991 and examine whether the trend has changed significantly.

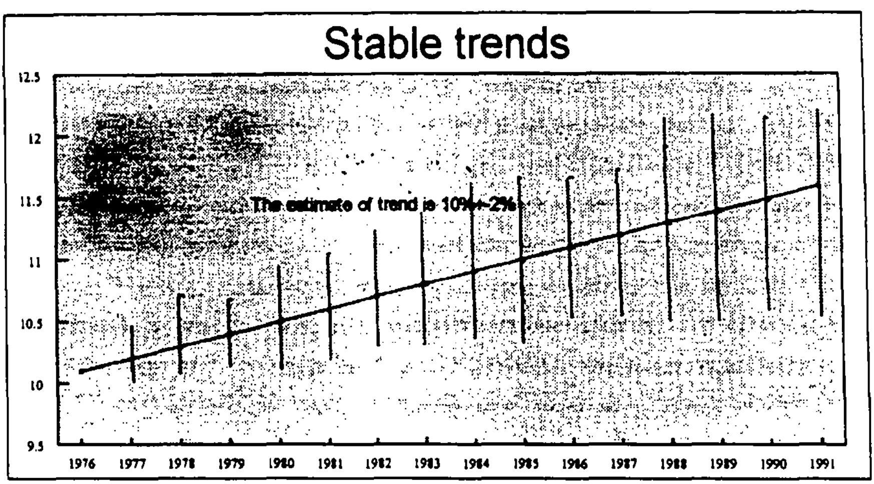


Figure 9.6.2.1

Typically, if the payment/calendar year trend is stable, the model will also validate well. Here our estimates of outstanding payments do not change significantly as we omit recent years.

Example 2: Consider the payment year trends depicted in Figure 9.6.2.2 below.

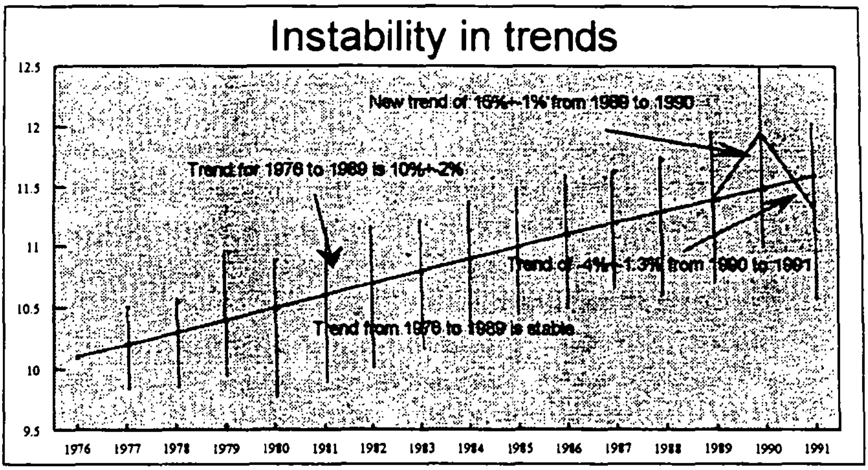


Figure 9.6.2.2

The trend in the years 1976 to 1989 is relatively stable. Its estimate is 10% ± 2%, say. However, the trend from 1989 to 1990 is higher at 15% (± 1%) and from 1990 to 1991 it is -4% (± 1.3%), say. This information is extracted from the "optimal" statistical model. The shifts in trends is a property of the data (determined through the model). A question now emerges as to which trend assumption do we make for the future, first in the absence of any other information. It would be foolhardy to assume the estimate between the last two years of -4% ± 1.3%. The most reasonable assumption (for the future) is a mean trend of 10% with a standard deviation of 2%, that which was estimated for the years 1976-1989.

Suppose we also have access to another data type, the number of closed claims development array. See Sections 10.2 and 10.3. We find utilising our DFF modelling framework that the additional 5% above the 10% between 1989 and 1990 can be explained by a corresponding increase in speed of closures of claims and the -15% from 1990 to 1991 below the 10% from 1976-1989 can be explained by a corresponding decrease in the speed of closures of claims. What assumption about *future trends* in payments should we adopt then? I would still recommend 10% ± 2% for the future. That's a decision based on my judgement and experience. The instability in trends in the last few years means that the model will not validate well. At year end 1990, we would not have forecast the distributions for 1991, for example.

Example 3: It is possible to have a transient change in trend. Consider Figure 9.6.2.3. The business has been moving along 10% ± 2% but between the last two calendar years 1990 and 1991 the trend increases to 20% ± 3%. What do we assume for the future? Well, that depends on the explanation for the increase in trend. Suppose its a "transient" change that can be explained by a new level of benefits that apply retrospectively. Then it is reasonable to assume 10% ± 2% for the future. Suppose instead that subsequent to analysis of claims closed triangle, the trend change is explained by increase in severities. That's a problem, because this means that it is now more likely that the new trend will continue.

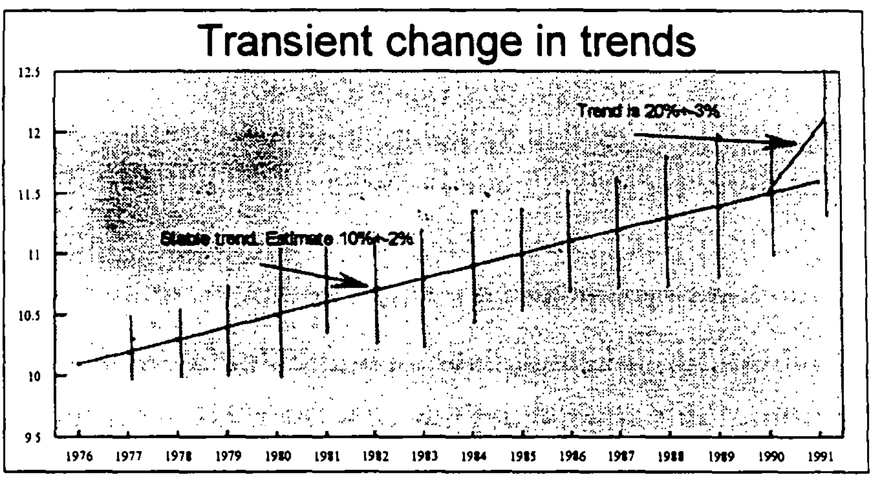


Figure 9.6.2.3

So the decision making process about the future becomes more complicated when trends are unstable. We are talking about trends in the (incremental) payments not age-to-age link ratios.

The last example illustrates an 'unpredictable' loss development array.

Example 4: The payment year trends are depicted in Figure 9.6.2.4 below.

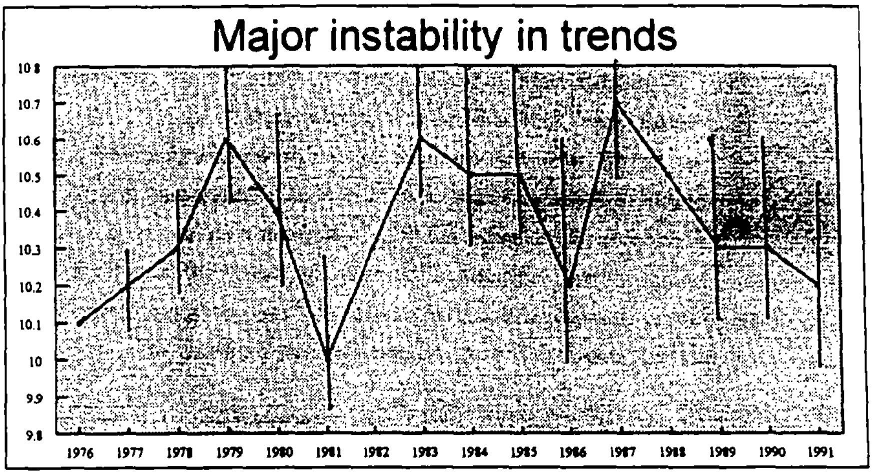


Figure 9.6.2.4

Note instability in trends. At year end 1989, would anyone be able to predict a flat trend for the next year and a downward trend for the following year?

Here, maybe, one could calculate a $\hat{\iota}$, a weighted average of trends estimated in the past with a weighted variance $\hat{\sigma}^2$ and assume for the future a mean trend of $\hat{\iota}$ with standard deviation of trend $\hat{\sigma}$. Since $\hat{\sigma}$ will be relatively large, mean forecasts will be well above the median forecasts and the standard deviation of the distributions relatively large. See Section 8.2

It is instructive to relate the foregoing discussion with the quote from A.C. Harvey [9] given at the beginning of Section 2.1.

9.7 POST-SAMPLE PREDICTIVE TESTING AND MODEL MAINTENANCE

Once a model has been identified for year end 1991, and assumptions about the future are made, the model is stored.

One year later, in 1992, on receipt of additional information (diagonal), there is no need to analyse the (augmented) triangle from the start. We already have a model for which we now conduct post-sample predictive testing and model updating and maintenance.

Has the model at year end 1991, predicted the distributions for 1992? This question is answered by restoring the model, assigning zero weight to "payments" in 1992 and validating the year. We also test for stability of parameters. If the model estimated at year end 1991 does not predict 1992 accurately, we know which parameter is the culprit and accordingly may have to amend the model (slightly).

For example, consider Example 2 of the preceding sub-section. If the 1992 data do not lie on the 10% ± 2% trend, then we have more evidence of changes in trends and our assumption of 10% ± 2% becomes pretty suspect.

Typically, once a model is identified for an incremental paid loss development array, the same model (with occasional minor amendments) is used in every subsequent year.

There is no way that a statistical method can automatically determine the "best" model and assumptions to be adopted for the future Rather, this decision is based on the model identification strategy (that may include analysis of other data types) and considerable judgment, especially if trends in the incremental payments are unstable.

Of course, any information about the nature of the business (especially change in business) may be critical in determining the assumptions for the future.

For example, in a number of loss development arrays of Lloyd's Syndicates analysed by the author, asbestos and pollution claims are not covered by policies written after 1978, say. This means that the calendar year effects of asbestos and pollution claims only apply to accident years prior to 1978. So, the iota estimates applying to accident years prior to 1978, do not apply to accident years post 1978.

For loss development arrays where the forecast uncertainties are relatively large, analysis of "similar" arrays within the company or analysis of industry wide arrays, for the purpose of formally credibility adjusting the parameter (estimates) could prove very useful. Incidentally, credibility is not just a function of volume. It is a myth that if claim numbers are "small" or incremental paids are small, or the triangle dimensions are small, then random fluctuations necessarily swamp the pattern (trends). The noise to signal ratio, equivalently, process uncertainty, may be very small even with small volume. Of course large volume and little process uncertainty does not mean that standard actuarial techniques will pick up the changing trend. See Section 12 for a study of a real life example involving (very) large volume and alarming calendar year shifts that cannot be detected using standard actuarial techniques.

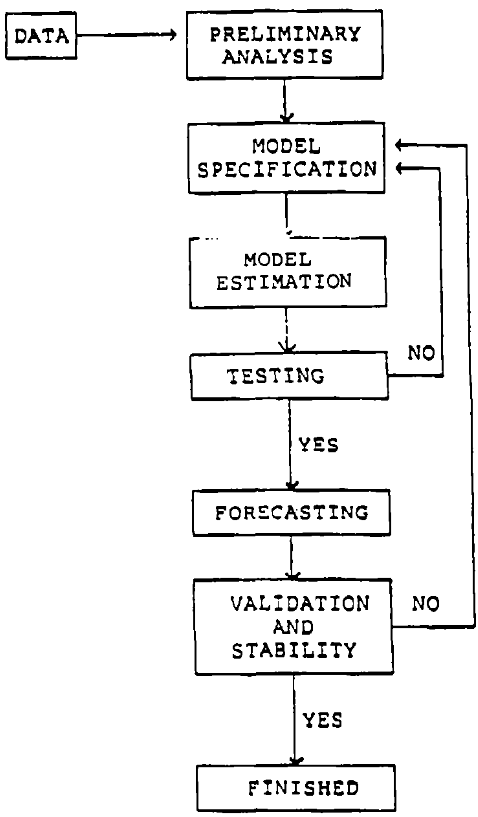
On every subsequent evaluation date post-sample predictive testing is conducted and the model is updated. Since data are recorded sequentially over time, updating procedures that can be applied routinely and that avoid re-analysis of the history are very desirable. See Section 9.6.2.

criterion is not satisfied, the model may have to be re-specified and the identification cycle repeated.

Step 6: Assumptions about the future based on Step 5 involving possibly analysis of other data types (Sections 10.2 and 10.3), are decided and forecasts and standard errors are produced. The final model is stored.

Step 7: Finished.

STEPS IN MODELLING



10. MODEL IDENTIFICATION AND ASSUMPTIONS ABOUT THE FUTURE

The aim is to identify a parsimonious model that separates the (systematic) trends from the random fluctuations and moreover determine whether the trend in the **payment/calendar year** direction is stable.

Recall that models contain information and accordingly the 'best' identified model *conveys information about the loss development array being analysed.*

For example, CCI (with constant development in the tail) indicates that the calendar year trend has been stable. This model should validate well and produce 'stable' outstanding estimates as recent calendar years are added or removed from the estimation. See preceding Sections 9.6.1 and 9.6.2.

10.1 MODEL IDENTIFICATION

The identification of the 'optimal' statistical model involves a number of iterative steps.

Step 1: Preliminary analysis facilitates the diagnostic identification of the heterogeneity in the data. The types of heterogeneity are also diagnostically identified.

Step 2: Based on step 1 a (preliminary) model is specified.

Step 3: The specified model is estimated.

Step 4: The model is checked to ensure that all assumptions contained in the model are satisfied by the data. If the model is inadequate, it has to be re-specified (step 2), and the iterative cycle of model specification - estimation - checking must be repeated.

Step 5: The best identified model is validated and tested for stability. If either

10.2 ASSUMPTIONS ABOUT THE FUTURE

We demonstrated in Section 9.6 that if payment/calendar year trend has been stable, especially in the more recent years, then the assumption about the future trend is relatively straightforward. For example, if the estimate in the last five years has been $\hat{\iota}$ $\pm$ s.e. ($\hat{\iota}$), then we assume for the future a mean trend of $\hat{\iota}$ with a standard deviation of trend of s.e. ($\hat{\iota}$). We do not assume that trend in the future is constant. Our model does include the variability (uncertainty) in trend in the future.

If on the other hand, payment/calendar year trend has been unstable as is illustrated in examples 2 and 3 of Section 9.6, assumptions about future trends are not so obvious and may depend on analysis of other data types.

In Section 10.1 we also cited a practical example where special knowledge about the business is a contributory factor in making decisions about the future. But, that special knowledge is combined with what we found in the past experience.

10.3 OTHER DATA TYPES AND METHODS

Hitherto much emphasis has been placed on the importance of analysing and predicting distributions for (incremental) paid loss development arrays. Reasons given include:

- the geometry of trends;
- simplicity and parsimony;
- distributions of future payments is relevant information for financial statements.

We now discuss other data types and methods.

10.3.1 PAYMENTS PER CLAIM CLOSED

Let the "series" $\{p_t\}$ denote the payments loss development array and the series $\{n_t\}$ denote the closed claims development array.

We shall say that $\{n_t\}$ causes $\{p_t\}$, if taking account of past values of n_t leads to improved predictions of future values of p_t. (This is know as Granger causality.)

Typically, an actuary analyses $z_t = p_t/n_t$ and obtains predictions $\hat{z}_t$ of future values of z_t. The analysis of $\{n_t\}$ leads to predictions $\hat{n}_t$ of future values of n_t.

The future values of p_t are then predicted by $\hat{p}_t = \hat{n}_t \hat{z}_t$.

So, is the forecast $\hat{p}_t$ better than the forecast $\tilde{p}_t$ that only depends on past values of p_t. A forecast is better if its mean square error is less. That is, $\hat{p}_t$ is better than $\tilde{p}_t$ if

$$E[(\hat{p}_t - p_t)^2] < E[(\tilde{p}_t - p_t)^2] \quad .$$

The author believes that $\tilde{p}_t$ is better than $\hat{p}_t$. That is, there is no reduction in forecast error with respect to the given information set { $\hat{z}_t$, $\hat{n}_t$, $\hat{p}_t$ }. However, this does not rule out the possibility that when there is an instability in calendar year trends in $\{p_t\}$ as described in Section 9.6, analysis of $\{n_t\}$ will not lead to improved accuracy of predicting future values of $\{p_t\}$. The information extracted from the analysis of $\{n_t\}$ <u>may</u> improve the actuary's judgment in respect of which assumptions to use for future trends of p_t.

10.3.2 INCURRED LOSSES AND CASE RESERVES

Analysis of incurred losses (paid to date plus case reserves) does not provide information about what is still to be paid. We have given sufficient reasons why any analysis of cumulative data is unsound. And adding case reserves to cumulative paids reduces the information (not increases the information).

Incremental paid losses and case reserves should be analysed separately. That is the best way to determine the information contained in each data type and any relationships that may exist between the two data types.

For example, if there is a trend shift in the incremental paids between calendar years 1984 and 1985 and a corresponding shift in the case reserves one year later, between 1985 and 1986, then we know that the case reserves are lagging the payments.

If instead we found that case reserves are leading the payments then a change in trend in the case reserves between the last two calendar years, for example, may suggest an increase in trend in payments one year later (in the future). See Sections 10.1 and 10.2.

For a small dimensional triangle of a long tail line, case reserves for the early accident years will be helpful in determining the development year trend (γ) in the future.

There are ways of determining whether case reserves have been "accurate" in forecasting subsequent payments. See the paper by Fisher and Lange [6].

Perhaps we should also remark that case reserves vary between and within claims personnel and due to changing reserving philosophy of the company.

10.4 TIME SERIES MODELS VERSUS EXPLANATORY (OR CAUSAL) MODELS

The rich modelling framework advocated by the author contains essentially time series models. The only "causal" variable is time, equivalently payment year, accident year and development year. The past values of the incremental payments are used to forecast future values of the payments.

There is an alternative approach to forecasting in statistics called explanatory or causal models. These models make an attempt to discover the factors (or variables) affecting the behaviour of the claims process.

There are many reasons for preferring time series models to explanatory models.

- Causality based on the definition given in Section 10.3.1 is hard to prove, especially since the causal variables need to also be forecast.

- Simplicity and parsimony discussed in Sections 9.1 and 9.2.

- The claims process is complex and is unlikely to be understood and even if it were understood, it may be extremely difficult to determine the relationships that govern the behaviour of claims. Moreover, its likely the relationship changes with time. This last reason is part motivation for varying parameter models. (See Section 7).

- Explanatory models are difficult to validate and test for stability and when they don't work it may be hard to determine the reason.

By way of summary, we advocate the use of the DFF of models applied primarily to the incremental payments and applied to "related" data types, especially for the case in which calendar year trend instabilities are found in the incremental payments.

11.0 PREDICTION INTERVALS, RISK BASED CAPITAL AND RELATED ISSUES

11.1 INTRODUCTION

Loss reserves often constitute the largest single item in an insurer's balance sheet. An upward or downward 10% movement of loss reserves could change the whole financial picture of the company.

The current paper is not meant to focus on risk based capital and solvency issues, but mainly to stress that these are necessarily probabilistic concepts. The paper's principal intention is to show how the distributions (or variability) of loss reserves may be derived from sample data. It is the variability or uncertainty of loss reserves that is relevant to risk based capital and solvency considerations.

11.2 PREDICTION INTERVALS

We have given persuasive arguments for the use of probabilistic models, especially in assessing the variability or uncertainty inherent in loss reserves. The probability that the loss reserve, carried in the balance sheet, will be realised in the future, is necessarily zero, even if the loss reserve is the best estimate. See Sections 8.0 and 10.3 for definition of best.

Future (incremental) paids may be regarded as a sample path from the forecast (estimated) lognormal distributions. The estimated distributions include both process risk and parameter risk.

The forecast distributions are accurate <u>provided</u> the assumptions made about the future will remain true. For example, if it is assumed that future calendar year trend (inflation) has a mean of 10% and a standard deviation of 2%, and in two years time it turns out that inflation is 20%, then the forecast distributions are far from accurate.

Accordingly, any prediction interval computed from the forecast distributions is <u>conditional</u> on the assumptions about the future remaining true.

Suppose $\hat{p}$ is a mean of a forecast lognormal distribution corresponding to payment p. Both $\hat{p}$ and p are random variables.

Let u = log p, μ = E [u] and σ^2 = Var [u]. A 100 (1-α)% prediction interval for u (a random variable) is given by

$$\mu \pm \sigma Z\ (\alpha/2) \quad ,$$

where Z (α/2) is the 1-α/2 percentage point of the standard unit normal distribution.

A 100 (1-α)% prediction interval for p (=log u) is

$$\exp\ [\mu \pm \sigma\ Z\ (\alpha/2)\ ].$$

The latter interval is non-symmetric about $\hat{p}$ since the lognormal distribution is skewed (to the right). The parameters μ and σ are computed from the mean and standard deviation of p, and the relationship between the lognormal and normal distributions.

The limits of the interval can be interpreted as follows. Suppose repeated samples of the rectangle are taken (from the estimated probabilistic model), then the proportion of times the observed p value will lie in the observed interval (in the long run) is 1-α. Bear in mind that p is a random variable.

The distribution of sums, for example, accident year outstanding payments, is the distribution of a sum of lognormal variables that are correlated. The exact distribution of the sum can be obtained by generating (simulating) samples from the estimated multivariate lognormal distributions. Alternatively, one can approximate the distribution of the sum by a lognormal. Indeed, the lognormal would be the riskiest.

If there are 'many' components in the sum, then the Central Limit Theorem could be invoked, especially if the lognormal distributions of the paids are not terribly skewed. See Section 13 for a real life example.

Insurer's risk can be defined in many different ways. Most definitions are related to the standard deviation of the risk, in particular a multiple of the standard deviation.

If an insurer writes more than one long tail line and aims for a $100(1-\alpha)\%$ security level on all the lines combined, then the risk margin per line decreases the more lines the company writes. This is always true, even if there exists some dependence (correlation) between the various lines.

Consider a company that writes n independent long tail lines. Suppose that the standard error of loss reserve L(j) of line j is se(j). That is, se(j) is the standard error of the loss reserve variable L(j). The standard error for the combined lines L(1)+...+L(n) is

$$se(Total) = [se^2(1)+...+se^2(n)]^{0.5}$$

If the risk margin for all lines combined is k*se(Total), where k is determined by the level of security required, then the risk margin for line j is

$$k*se(Total)*se(j)/[se(1)+...+se(n)]$$

$$<kse(j).$$

The last inequality is true even when se(Total) is not given by the above expression.

If as a result of analysing each line using the DFF modelling framework we find that for some lines trends change in same years and the changes are of the same order of magnitudes, then the paid losses are not independent. (There may also be some probabilistic model, derived from the company's experience, that describes the particular line for that company. In the hundreds of arrays that the author has analysed, no one model described more than one loss development array.

The approach the author is advocating allows the actuary to determine the relationships within and between companies experiences and their relationships to the industry in terms of simple well understood features of the data.

In establishing the loss reserve, recognition is often given to the time value of money by discounting. The absence of discounting implies that the (median) estimate contains an implicit risk margin. But this implicit margin may bear no relationship to the security margin sought. The risk should be computed before discounting (at a zero rate of return).

correlations between the residuals).

In that situation, line i and j are correlated, say, then one should use se(i)+se(j) as the upper bound of the standard error of L(i)+L(j).

We now return to an important modelling concept or 'law of payments'.

Suppose we assume for the future payment/calendar years a <u>mean</u> trend of ($\hat{\iota}$) with a standard deviation (standard error) se ($\hat{\iota}$). Specifically we are saying that the trend ι, a random variable, has a normal distribution with mean $\hat{\iota}$ and standard deviation se ($\hat{\iota}$). Recognition of the relationship between the lognormal and normal distributions tells us that the mean payment increases as se ($\hat{\iota}$) increases (and $\hat{\iota}$ remains constant). The greater the uncertainty in a parameter (the mean remaining constant), the more money is paid out.

The foregoing arguments apply to each parameter in the model.

11.3 RISK BASED CAPITAL

The author understands that the NAIC is drafting regulations where part of the risk based capital requirements will be based on loss reserves. In the article by Laurenzano [10], page 50, the loss reserve component of the risk based capital formula "*selects the worst reserve development* ...".

The approach advocated by the NAIC is flawed for many reasons including:

* The uncertainty in loss reserves (for the future) should be based on a probabilistic model (for the future) that may bear no relationship to reserves <u>carried</u> in the past;

* The uncertainty for each line for each company should be based on a

12.0 ANALYSIS OF PROJECT 1

12.1 INTRODUCTION AND SUMMARY

The principal objectives of the analysis of real life data in this section are to demonstrate that:

1. Age-to-age link ratios based on the cumulative paid losses give no indication about the trends and random fluctuations in the (incremental) payments.

2. Smooth data may have major shifts in calendar year trends.

3. Regression as an approach to adjusting data and determining trends and changes thereof is very powerful.

4. Large company's run-off payments are not necessarily stable in respect of calendar year trends, even though the payments may be extremely smooth (with very little random fluctuations about the trends).

12.2 DATA AND AGE-TO-AGE LINK RATIOS

The data (save a multiplicate factor in order to preserve confidentiality) come from a large insurer and are given in Appendix C1. Accident year exposures, (from memory), represent earned premium (relativities). As we shall see in the next section, the exposures are not that important.

The age-to-age link ratios presented in Appendix C2 are relatively smooth. For the early development years they tend to decrease slightly in the middle accident years and then increase in the latter payment years.

12.3 ANALYSES

We define a normalised payment as the (incremental) paid divided by the corresponding accident year exposure and apply the MODELLING FRAMEWORK to the normalised payments.

If p(w,d) is the incremental payment corresponding to accident year w and development year d, and e(w) is the accident year exposure, then the normalised payment is p(w,d)/e(w) and we define,

$$y(w,d) = \log [p(w,d)/e(w)]$$

Figure C3 (in Appendix C3) represents a graph of the normalised payments versus delay for the first two accident years in the triangle. Observe that the run-off development for both years is remarkably smooth.

The chain ladder (CL) statistical model is given by,

$$y(w,d) = \alpha_w + \sum_{j=1}^{d} \gamma_j + \varepsilon$$

Since the exposures e(w) are absorbed into the parameters α_w, the estimates of the development trends γ_j do not depend on the exposure base used. Indeed, there are other statistics that are invariant (for CL) with respect to exposure base including, AIC, residuals, S-squared, normality testing and forecasts. The chain ladder model adjusts for the different levels (α's) of each accident year.

The estimates of the CL parameters and associated regression table are presented in Appendix C4. R-squared is high and S-squared is small. Hence, the random fluctuations are small. Now, the CL model adjusts the data for development year trends and accident year trends (or levels). Many parameter differences are insignificant but that is not important since we are not trying to identify a parsimonious model here but rather show how some of the models in the FRAMEWORK may be used for fast identification of payment/calendar year trend changes.

So, the residuals represent the data adjusted (after removing) for the average development year trends and the average accident year trends.

Residuals versus development years (Figure C5.1) and accident years (Figure C5.2) are the "best" we can obtain since we have removed the trends in these two directions. In Figure C5.1, the sum of residuals for any one development year is zero and in Figure C5.2, the sum of residuals for any one accident year is zero. HOWEVER, residuals versus payment years (Figure C5.3) exhibit a very strong V shape AND THIS IS FOR SMOOTH DATA OF A LARGE COMPANY. So, after removing accident year and development year trends from the data we observe major shifts in calendar year trends. (Compare this with the simulated data of Sections 4.4 and 5). There appears to be a change in trend in 1984 and definitely a change in trend in 1985.

We now estimate the CC model. It adjusts the data for the average development year trends. Appendix C6 presents the regression output and Figure C7 is a graph of residuals versus payment years that indicates an upward trend (positive inflation). It is hard to tell from this graph whether there is a major shift in trends.

In order to estimate a trend parameter through the residuals of Figure C7, we estimate the CCI model to the data. The regression output is presented in Appendix C8 and the residuals versus payment years are displayed in Figure C9. The average payment year trend is 12.1% ($\pm$ 0.53%). The V shape in residuals is distinct, suggesting very strongly the change in trends.

Our final model introduces another two payment year trend parameters. One from

1984-1985 and one from 1985-1987. The regression output is given in Appendix C10. Note shift in trend from 9.85% to 19.52%. This is quite alarming, especially if it cannot be explained by an increase in speed of finalisations of claims. See Section 10.2, for a discussion of assumptions to be applied for the future.

We now graph in Appendix C11 the lognormalised payments versus delay for the first two accident years. Since 19.52% is much higher than 9.85%, observe that the trend in the tail increases for both accident years, and for accident year 1978 the change is one development year earlier than in accident year 1977. That is because the trend change is a calender year change.

So there is overwhelming statistical evidence of a major shift in calendar year trends in the last two calendar years. What assumptions do we make about the future trends? We could analyse the number of claims closed development array and determine whether the substantial increase in trend in the payments is due to a corresponding increase in trend in the number of closed claims. If the answer is in the negative, then the trend increase must be due to increase in severities which would then be a major concern for the company. See Section 10.2.

In this section we have not identified a parsimonious model for the data. Instead the objective was to demonstrate how some of the models in the MODELLING FRAMEWORK may be used for quick determination of major calendar year shifts (in data that are relatively smooth and do not appear problematic if we are to employ the standard actuarial approaches based on link ratios).

The reader will appreciate that our modelling approach is interactive and terribly computer intensive. In order to identify the calendar year trend changes we have had to estimate four models. To set up each model in a spreadsheet is extremely time consuming. See Section 8.

13.0 ANALYSIS OF PROJECT 2

13.1 INTRODUCTION AND SUMMARY

In the present section we analyse a real life loss development array for which the age-to-age link ratios of the cumulative paids are relatively unstable, yet the trends in the paids are stable.

The "best" identified model is essentially a version of CC with two additional iotas (payment year trend parameters) that are used to adjust for "low" payments in one payment year. The model (and so the trend in the data) is stable and validates very well. Had the model been employed three years earlier, it would have yielded the "same" outstanding payments and would have forecast the distributions of (incremental) payments for the last three years extremely accurately.

13.2 DATA AND PRELIMINARY ANALYSIS

The incremental paid loss development array and accident year exposures are displayed in Appendix D1. The exposures are estimates of the number of ultimate claims incurred in each accident year. We define a normalised payment as the paid divided by the corresponding accident year exposure and identify a DFF model for the normalised payments.

The first step in the preliminary (diagnostic) analysis is to graph the data. Figure D2.1 displays a graph of normalised payments versus development year for all accident years combined. It exhibits a band whose width (variability) increases as the normalised payments get larger.

On the other hand, the graph of the lognormalised payments depicted in Figure D2.2 exhibits a band whose width is relatively constant. That is, % variability is constant with development year suggesting a lognormal distributions for the normalised payments.

The graph in Figure D2.2 also gives us a preliminary idea of a parsimonious number of γ's (development year trend parameters) that may be required in the model.

It appears we require one γ from delay 0-1, one from delay 1-2 (that turns out to be insignificant to zero), one from delay 2-4 and one from delay 4-8.

13.3 MODEL IDENTIFICATION

In this sub-section we implement the model selection strategy discussed in Section 10.

Model 0 and 1: Estimate a CC model with the four γ parameters suggested by the preliminary diagnostic analysis. It turns out that the parameter γ_2 is insignificant from zero, as was anticipated from the graphs. Set γ_2 to zero and re-estimate the model. Regression tables and residual displays are given in Appendix D3 and Appendix D4, respectively.

Residuals versus delay and accident years suggest that the trends in these two directions have been captured well. This diagnostic test can be formalised by adding more parameters and testing for significance of parameters and their differences.

Since we have estimated a CC model, the residuals may be interpreted as the data adjusted for the development year trends.

Residuals versus payment years (Figure D4.3) suggest (i) zero trend from 1975-1979, (ii) low payments in 1974 and (iii) perhaps zero trend from 1969-1973. So we next estimate.

Model 2: This model is the previous CC model with four iota parameters. The first iota represents the trend from 1969-1973, the second iota the trend from 1973-1974 the third iota represents the trend from 1974-1975 and the fourth iota represents the trend from 1975-1979. We find that both the first and fourth iota are insignificant, and the first being less significant than the fourth.

Model 3: Previous model with first iota set to zero. We find that fourth iota is still insignificant.

Model 4: *Previous model with fourth iota set to zero. We find all parameters and* their differences are significant. Moreover, SSPE and AIC are the lowest amongst the four models. Outlier analysis indicates that the observation in accident year 1972, delay 7 is an outlier.

So our final identified model (before conducting validation and stability analysis) has three gammas (0-1, 2-4 and 4-8), two iotas (1973-1974 and 1974-1975) and one alpha, and it also assigns zero weight to (1972,7).

The regression tables and various statistical displays are given in Appendices D5 to D7.

Figure D7.5 of Appendix D7 displays a normal probability plot where r^2 (correlation squared) between the normal scores and ordered residuals is 0.993. The P-value is in excess of 0.5.

So we have shown that the log incremental payments in the cells of the loss development array can be regarded as observations from normal distributions whose means are related by the (trend) parameters given in Appendix D5.

Forecasts, standard errors and % errors based on the model are presented in Appendices D8 and D9, respectively.

Appendix D8

This appendix presents:

(i) each observed payment (OBS);

(ii) each expected model payment (EXP), that is a mean of a lognormal distribution;

(iii) forecasts for each accident year subdivided according to development year (right side of stair-case corresponding to EXP row);

(iv) standard errors of each individual forecast (below each forecast);

(v) total forecast (outstanding) for each accident year and associated standard error (right hand column);

(vi) total forecast (payment) to be made in each future payment year in respect of all the accident years and associated standard errors (bottom row). This is the future liability stream with corresponding uncertainties that may prove useful for asset/liability matching;

(vii) total outstanding with associated standard error (bottom right hand corner).

Expected values and forecasts are estimates of means of lognormal distributions. Standard errors are estimates of standard deviations of lognormal distributions.

Appendix D9

Here we present a quality of fit table comparing the original observed payments with the model expected payments. For each accident year and for each payment year we compute the ratio of the difference in total observed and total expected to the total observed. The quality of fit is high.

13.4 VALIDATION AND STABILITY ANALYSIS

We now re-estimate the same model and assign zero weight to the last three calender years (1979, 1978 and 1977). We aim to determine (i) whether the model estimated at year end 1976, would have forecast the distributions of payments in years 1977-1979 and (ii) are the parameter estimates of the model and the forecasts based on the model stable.

Appendix D10 presents the parameter estimates as of year end 1976. Compare these estimates with those obtained at year end 1979 (Appendix D5). Note that none of the parameter estimates have changed significantly. The estimate of the tail. -0.5544 ($\pm$ 0.0753) at year end 1976, is slightly higher than the estimate -0.6749 ($\pm$ 0.0390), at year end 1979, hence the higher forecasts in the tail. The estimates of iotas 1973-1974 and 1974-1975 are very close (and so stable).

Appendix D11 represents "All" residuals displays. All residuals include those corresponding to observations used in the estimation (1969-1976), and the validated residuals (1977-1979) corresponding to observations not included in the estimation. All displays are great.

In particular, Figure D11.3 shows the distribution of the validated residuals (prediction errors) for 1977-1979 relative to residuals corresponding to years used in the estimation.

Appendix D12 presents displays of the validated residuals (only those corresponding to years 1977-1979). All displays are in good shape.

Most importantly, Appendix D12.4 presents a test whether the lognormalised payments in 1977-1979 come from the forecasted distributions as at year end 1976. The squared correlation between normal scores and validated residuals is 0.959 with a P-value of 0.313.

By way of summary, there is very strong statistical evidence that the model at year end 1976 would have predicted accurately the <u>distributions</u> of 'payments' for 1977-1979.

Let's now compare the forecasts, Appendix D13 (validation model) with Appendix D8.

Total outstanding beyond 1979, based on estimated model at year end 1976 is 12,620,833 $\pm$ 1,072,089 compared with estimated model at year end 1979 of 12,948,473 $\pm$ 1,030,808. No difference.

So, we could have obtained the same answers three years ago (that is, without the last three years information). All other forecasts compare extremely favourably.

Note that in Appendix D13 the expected values corresponding to payment years 1977-1979 actually represent mean forecasts based on estimated model at year end 1976.

From Appendix D14 we see that had we reserved mean forecasts at year end 1976 (for years 1977-1979) we would have underforecast 1977 and 1978 by 13% and 1% respectively, and overforecast 1979 by 5%.

Our findings using probabilistic models have shown that:

* calendar year trends are essentially stable, save the dip in the year 1974;
* the model used three years earlier would have predicted accurately the distributions of payments for the last three years;

and

* rough (irregular) age-to-age link ratios, especially in the early development years, give no indication of stability of trends.

The author has analysed numerous data sets with rough (or irregular) age-to-age link ratios for which the payment/calendar year trends are stable. Conversely, smooth age-to-age link ratios does not mean stability of trends.

We conclude this section by showing how to compute a prediction interval for the total outstanding payments, using the discussion of Section 11.2.

From Appendix D8, the mean outstanding is given by

$$m \quad = \quad \text{mean} = 12{,}948{,}473$$

and the standard deviation (or standard error) by

$$sd = 1,030,808$$

We assume that the total reserve (or liability) L is a random variable with mean m and standard deviation sd and moreover the distribution of L is lognormal.

Put y = log L, then y has a normal distribution with mean μ and standard deviation σ, say.

Therefore,

$$m = \exp[\mu + 0.5\ \sigma^2]$$

and

$$sd = m\ [\exp(\sigma^2) - 1]^{0.5}$$

Solving the last two equations for μ and σ we obtain,

$$\mu = 16.37332$$

and

$$\sigma = 0.079482$$

Employing Section 11.2, a 100 $(1-\alpha)$% prediction interval for the random variable L is given by

$$\exp[16.37648 \pm 0.079482Z\ (\alpha/2)]$$

where Z (α /2) is the 1 - α /2 percentage point of the standard unit normal distribution

The median of the distribution of L is exp [μ] = 12,907,636 which is very close to the mean of 12,948,473. Since σ^2 is small the lognormal distribution is not terribly skewed, so that were we to assume that the distribution of L is normal (rather than lognormal), the prediction intervals would be almost the same.

14. EXTENSION OF THE DFF MODELLING FRAMEWORK

We observed that a fruitful extension of the DFF modelling framework was the introduction of varying parameter (dynamic) models in Section 7.

Another important extension is related to the distributional assumption of normality. Hitherto, we have assumed that the variances of the y values, denoted by σ^2, are identical (constant)

The variance on a log scale can be interpreted as % variability. So constant σ^2 implies constant % variability. For many loss development arrays this assumption is not valid. For some arrays, % variability increases in the tail, for some others, % variability is higher in the early development years. When σ^2 is not constant and varies with development years we need to also model the σ^2's. That is, we introduce a secondary equation.

This is outside the scope of the present paper.

15. CONCLUSIONS

We have argued that the four components of interest regarding a loss development array are the trends in the three directions and the distributions (random fluctuations) about the trends.

A MODELLING FRAMEWORK was introduced where each model contained therein possesses the four components of interest. The modelling approach offers the actuary a way of fitting (estimating) distributions to the cells in a loss development array and predicting (forecasting) distributions for future years that affords numerous advantages including:

- simplicity;
- clarity of assumptions;
- testing of assumptions;
- assessment of loss reserve variability;
- asset/liability matching;
- model maintenance and updating.

We showed how the identified optimal statistical model for the (incremental) payments conveys information about the loss experience to date. In applying the model to predicting distributions of future payments the actuary may (need to) adjust some of the parameters to reflect knowledge about the business and to incorporate his view of the future. View of the future may be based on analysis of other data types, especially if there are instabilities in the payments in the recent calendar years.

A prediction interval computed from the forecast distributions is conditional on the assumptions made about the future remaining true.

In passing we have debunked a number of pervasive loss reserving perceptions concerning data types, age-to-age link ratios, stability, forecasting and regression.

Methods based on age-to-age link ratios do not (and cannot) separate trends from random fluctuations and moreover do not satisfy the basic fundamental property of additivity of trends.

REFERENCES

[1] Abraham, B. and Ledholter. J. (1983). Statistical Methods for Forecasting. John Wiley, New York.

[2] Bailey R.A., "Sampling theory in casualty insurance", PCAS, 1942. pp.29-50.

[3] Catterjee, S. and Price, B. (1977). Regression Analysis by Example. John Wiley, New York.

[4] Christofidies, S. (1980). "Regression models based on log-Incremental Payments". Institute of Actuaries Loss Reserving Manual, Volume 2., Chapter 2.

[5] Foundations of Casualty Actuarial Science, 1989, Casualty Actuarial Science, New York.

[6] Fisher, W.H. and Lange, J.T. (1973). "Loss reserve testing: a report year approach", PCAS, 60, pp.189-207.

[7] Friedman, M. (1953). "The methodology of positive economics", in Essays in Positive Economics, University of Chicago Press, pp.3-43.

[8] Gauss, K.F. (1809). Theoria Motus corpurum coelstium, Werke 7, Hamburg.

[9] Harvey, A.C. (1990). Forecasting, structural time series models and the Kalman filter. Cambridge University Press, New York.

[10] Laurenzano, V. (1992). "Balancing capital with risks", Best's Review, June, 1992.

[11] Loss Reserving Manual, 1990, Institute of Actuaries, London.

[12] Muth, J.F. (1960). "Optimal properties of exponentially weighted forecasts". J. Amer, Statist. Assoc., 55, 299-306.

[13] Popper, K.F. (1959). *The Logic of Scientific Discovery*. Hutchinson.

[14] Zehnwirth, B. (1985). "Linear Filtering and Recursive Credibility Estimation", ASTIN Bulletin, Vol. 15, No. 1, 19-35.

Appendix A1

Model is p = exp(alpha-.2d) with no error or randomness
alpha = 11.51293

Year/delay

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1978	100000	81873	67032	54881	44933	36788	30119	24660	20190	16530	13534	11080	9072	7427
1979	100000	81873	67032	54881	44933	36788	30119	24660	20190	16530	13534	11080	9072	
1980	100000	81873	67032	54881	44933	36788	30119	24660	20190	16530	13534	11080		
1981	100000	81873	67032	54881	44933	36788	30119	24660	20190	16530	13534			
1982	100000	81873	67032	54881	44933	36788	30119	24660	20190	16530				
1983	100000	81873	67032	54881	44933	36788	30119	24660	20190					
1984	100000	81873	67032	54881	44933	36788	30119	24660						
1985	100000	81873	67032	54881	44933	36788	30119							
1986	100000	81873	67032	54881	44933	36788								
1987	100000	81873	67032	54881	44933									
1988	100000	81873	67032	54881										
1989	100000	81873	67032											
1990	100000	81873												
1991	100000													

Appendix A2

y=log(p) plus .1 inf. from 1978-82, .3 inf. from 1982-83 and .15 inf. from 1983-91

Year\delay

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1978	11.5129	11.4129	11.3129	11.2129	11.1129	11.2129	11.1629	11.1129	11.0629	11.0129	10.9629	10.9129	10.8629	10.8129
1979	11.6129	11.5129	11.4129	11.3129	11.4129	11.3629	11.3129	11.2629	11.2129	11.1629	11.1129	11.0629	11.0129	
1980	11.7129	11.6129	11.5129	11.6129	11.5629	11.5129	11.4629	11.4129	11.3629	11.3129	11.2629	11.2129		
1981	11.8129	11.7129	11.8129	11.7629	11.7129	11.6629	11.6129	11.5629	11.5129	11.4629	11.4129			
1982	11.9129	12.0129	11.9629	11.9129	11.8629	11.8129	11.7629	11.7129	11.6629	11.6129				
1983	12.2129	12.1629	12.1129	12.0629	12.0129	11.9629	11.9129	11.8629	11.8129					
1984	12.3629	12.3129	12.2629	12.2129	12.1629	12.1129	12.0629	12.0129						
1985	12.5129	12.4629	12.4129	12.3629	12.3129	12.2629	12.2129							
1986	12.6629	12.6129	12.5629	12.5129	12.4629	12.4129								
1987	12.8129	12.7629	12.7129	12.6629	12.6129									
1988	12.9629	12.9129	12.8629	12.8129										
1989	13.1129	13.0629	13.0129											
1990	13.2629	13.2129												
1991	13.4129													

Appendix A3

Cumulative data (on a $ scale) derived from Appendix A2

100000	190484	272357	346439	413471	487552	558021	625053	688816	749469	807164	862045	914250	963908
110517	210517	301001	382874	473358	559428	641302	719182	793263	863732	930764	994527	055180	
122140	232657	332657	443174	548302	648302	743425	833908	919979	001852	1079732	1153814		
134986	257126	392112	520515	642655	758838	869355	974482	1074482	1169605	1260089			
149182	314055	470886	620068	761975	896961	1025363	1147504	1263687	1374204				
201375	392929	575141	748467	913339	1070170	1219352	1361259	1496245					
233965	456519	668219	869594	1061148	1243360	1416685	1581557						
271828	530399	776359	1010324	1232878	1444578	1645954							
315819	616236	902001	1173829	1432400	1678360								
366930	715964	1047976	1363795	1664212									
426311	831831	1217574	1584504										
495303	966450	1414619											
575460	1122855												
668589													

Appendix A4

Age-to- age link ratios of the cumulative losses of Appendix A3

1.904837	1.429816	1.272002	1.193488	1.179170	1.144535	1.120124	1.102011	1.088054	1.076981	1.067992	1.060558	1.054316
1.904837	1.429816	1.272002	1.236327	1.181830	1.146351	1.121440	1.103008	1.088834	1.077607	1.068505	1.060986	
1.904837	1.429816	1.332224	1.237213	1.182381	1.146726	1.121712	1.103213	1.088994	1.077736	1.068611		
1.904837	1.524979	1.327463	1.234652	1.180786	1.145639	1.120925	1.102618	1.088529	1.077362			
2.105170	1.499375	1.316812	1.228856	1.177152	1.143152	1.119119	1.101248	1.087456				
1.951229	1.463726	1.301361	1.220279	1.171712	1.139400	1.116378	1.099162					
1.951229	1.463726	1.301361	1.220279	1.171712	1.139400	1.116378						
1.951229	1.463726	1.301361	1.220279	1.171712	1.139400							
1.951229	1.463726	1.301361	1.220279	1.171712								
1.951229	1.463726	1.301361	1.220279									
1.951229	1.463726	1.301361										
1.951229	1.463726											
1.951229												

Appendix A5

Random error random from Normal with mean 0

Year\delay

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1978	0.083	0.075	-0.076	-0.065	-0.188	-0.164	-0.101	0.078	0.021	0.029	0.005	0.03	-0.073	-0.241
1979	-0.113	-0.049	-0.086	-0.123	0.148	0.09	-0.06	-0.099	-0.032	0.096	0.028	0.1	-0.331	
1980	0.086	-0.007	-0.037	0.17	0.071	-0.138	0.047	0.022	0.036	0.003	0.004	0.058		
1981	-0.071	0.147	0.067	-0.028	-0.132	0.049	0	-0.117	-0.042	0.026	-0.078			
1982	0.081	0.059	0.073	0.048	0.025	0.029	-0.023	-0.133	-0.044	0.066				
1983	0.117	0.059	-0.017	-0.081	-0.051	-0.024	-0.048	0.124	0.033					
1984	-0.024	-0.026	0.134	0.214	0.071	0.193	-0.022	0.012						
1985	0.022	0.015	0.076	-0.028	-0.004	0.155	0.032							
1986	-0.043	0.181	0.184	-0.192	-0.16	-0.048								
1987	0.07	0.106	0.144	0.032	-0.102									
1988	0.056	-0.195	0.032	0.041										
1989	0.145	0.187	-0.159											
1990	0.001	-0.153												
1991	-0.142													

Appendix A6

Sum of data in Appendices A2 and A5 to produce trends + randomness

Year\delay

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
1978	11.5959	11.4879	11.2369	11.1479	10.9249	11.0489	11.0619	11.1909	11.0839	11.0419	10.9679	10.9429	10.7899	10.5719
1979	11.4999	11.4639	11.3269	11.1899	11.5609	11.4529	11.2529	11.1639	11.1809	11.2589	11.1409	11.1629	10.6819	
1980	11.7989	11.6059	11.4759	11.7829	11.6339	11.3749	11.5099	11.4349	11.3989	11.3159	11.2669	11.2709		
1981	11.7419	11.8599	11.8799	11.7349	11.5809	11.7119	11.6129	11.4459	11.4709	11.4889	11.3349			
1982	11.9939	12.0719	12.0359	11.9609	11.8879	11.8419	11.7399	11.5799	11.6189	11.6789				
1983	12.3299	12.2219	12.0959	11.9819	11.9619	11.9389	11.8649	11.9869	11.8459					
1984	12.3389	12.2869	12.3969	12.4269	12.2339	12.3059	12.0409	12.0249						
1985	12.5349	12.4779	12.4889	12.3349	12.3089	12.4179	12.2449							
1986	12.6199	12.7939	12.7469	12.3209	12.3029	12.3649								
1987	12.8829	12.8689	12.8569	12.6949	12.5109									
1988	13.0189	12.7179	12.8949	12.8539										
1989	13.2579	13.2499	12.8539											
1990	13.2639	13.0599												
1991	13.2709													

Appendix A7

Incremental paids derived from Appendix A6

1978	108651	97529	75879	69418	55542	62875	63697	72468	65114	62436	57983	56551	48528	39023
1979	98706	95216	83025	72396	104914	94174	77103	70538	71747	77567	68934	70467	43560	
1980	133106	109743	96365	130993	112860	87108	99698	92494	89224	82117	78190	78504		
1981	125731	141478	144336	124854	107034	122015	110514	93517	95885	97626	83692			
1982	161765	174888	168704	156514	145495	138954	125480	106927	111179	118054				
1983	226364	203191	179136	159835	156670	153108	142187	160637	139511					
1984	228411	216837	242050	249422	205644	220996	169549	166858						
1985	277868	262472	265375	227499	221660	247187	207918							
1986	302519	360015	343485	224336	220334	234427								
1987	393525	388054	383425	326081	271278									
1988	450855	333667	398276	382277										
1989	572576	568013	382277											
1990	576021	469724												
1991	580068													

Appendix A8

Cumulative paids from Appendix A7

1978	108651	206180	282059	351477	407019	469894	533591	606059	671173	733609	791592	848143	896671	935694
1979	98706	193922	2769473	49343	454257	548431	625534	696072	767819	845386	914320	984787	1028347	
1980	133106	242849	339214	470207	583067	670175	769873	862367	951591	1033708	1111898	1190402		
1981	125731	267209	411545	536399	643433	765448	875962	969479	1065364	1162990	1246682			
1982	161765	336653	505357	661871	807366	946320	1071800	1178727	1289906	1407960				
1983	226364	429555	608691	768526	925196	1078304	1220491	1381128	1520639					
1984	228411	445248	687298	936720	1142364	1363360	1532909	1699767						
1985	277868	540340	805715	1033214	1254874	1502061	1709979							
1986	302519	662534	1006019	1230355	1450689	1685116								
1987	393525	781579	1165004	1491085	1762363									
1988	450855	784522	1182798	1565075										
1989	572576	1140589	1522866											
1990	576021	1045745												
1991	580068													

Appendix A9

Age-to-age factors (link ratios) of the cumulative payments

1978	1.897635	1.368023	1.246111	1.158024	1.154476	1.135556	1.135811	1.107438	1.093025	1.079038	1.071439	1.057216	1.043519
1979	1.964642	1.428136	1.261407	1.300318	1.207314	1.140588	1.112764	1.103074	1.101022	1.081541	1.077070	1.044232	
1980	1.824478	1.396810	1.386166	1.240021	1.149396	1.148764	1.120141	1.103464	1.086294	1.075640	1.070603		
1981	2.125243	1.540161	1.303378	1.199541	1.189631	1.144378	1.106759	1.098903	1.091636	1.071962			
1982	2.081123	1.501121	1.309709	1.219823	1.172107	1.132597	1.099763	1.094321	1.091521				
1983	1.897629	1.417026	1.262588	1.203857	1.165487	1.131861	1.131616	1.101012					
1984	1.949328	1.543629	1.362902	1.219536	1.193454	1.124361	1.108850						
1985	1.944592	1.491125	1.282356	1.214534	1.196981	1.138421							
1986	2.190057	1.518441	1.222993	1.179081	1.161597								
1987	1.986097	1.490577	1.279896	1.181933									
1988	1.740076	1.507667	1.323197										
1989	1.992030	1.335157											
1990	1.815463												
1991													

One cannot determine changing calendar year trends from the age-to-age link ratios.

APPENDIX B1

Random incremental paids from (same) lognormal distribution

DELAY

ACC. YEAR	0	1	2	3	4	5	6	7	8	9
1976	10266	3419	3724	9606	8152	8175	3958	3030	1733	351
1977	1767	2454	6580	2819	1957	2150	3677	4751	2832	
1978	6232	5143	2667	4278	2289	6215	6273	4905		
1979	4597	3591	5909	5156	4013	3557	1961			
1980	2483	3805	3995	6315	3480	3486				
1981	1643	2077	5101	1907	3274					
1982	3270	7230	1853	4158						
1983	3161	2065	5890							
1984	5305	6078								
1985	6127									

APPENDIX B2

Cumulative payments

ACC. YEAR	DELAY 0	1	2	3	4	5	6	7	8	9
1976	10266	13685	17409	27015	35167	43342	47300	50330	52063	55574
1977	1767	4221	10801	13620	15577	17727	21404	26155	28987	
1978	6232	11375	14042	18320	20609	26824	33097	38002		
1979	4597	8188	14097	19253	23266	26823	28784			
1980	4248	8053	12048	18363	21843	25329				
1981	1643	3720	8821	10728	14002					
1982	3270	10500	12353	16511						
1983	3161	5226	11116							
1984	5305	11383								
1985	6127									

APPENDIX B3

Age-to-Age Link Ratios

DELAY

ACC. YEAR	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9
1976	1.333041	1.272122	1.551783	1.301758	1.232462	1.091320	1.064059	1.034432	1.067437
1977	2.388794	2.558872	1.260994	1.143685	1.138024	1.207423	1.221967	1.108277	
1978	1.825256	1.234461	1.304657	1.124945	1.301567	1.233857	1.148200		
1979	1.781161	1.721665	1.365751	1.208435	1.152884	1.073108			
1980	1.895715	1.496088	1.524153	1.189511	1.159593				
1981	2.264150	2.371238	1.216188	1.305182					
1982	3.211009	1.176476	1.336598						
1983	1.653274	2.127057							
1984	2.145711								
1985									

APPENDIX B4

Incremental paids generated by SDF model with 20% calendar year trend

DELAY

YEAR	0	1	2	3	4	5	6	7	8	9	10	11	1
1978	53275	66971	121278	292065	86300	79271	240147	86269	73645	225638	218708	72438	861
1979	31912	85884	42106	150200	88290	82798	230017	346594	169950	113715	48703	82441	1689
1980	24964	96951	208159	697227	213581	251802	489886	387322	524382	133462	206570	76440	
1981	82867	117837	279958	469997	577054	378084	438640	556884	338201	173980	161958		
1982	41268	252181	101806	219303	283631	352082	748704	727854	147742	299994			
1983	32190	491133	239252	228226	375903	494626	323417	482001	157137				
1984	231651	401780	626068	496230	388360	395640	653268	535755					
1985	31273	409563	433997	831822	572787	468844	1317425						
1986	92728	342040	246087	530327	837381	694392							
1987	147772	208578	389162	602683	743423								
1988	146151	209864	1827396	1391050									
1989	81526	767664	1042474										
1990	206885	350789											
1991	559279												

APPENDIX B5

Age-to-age link ratios

DELAY

	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12
1978	2.26	2.01	2.21	1.16	1.13	1.34	1.09	1.07	1.21	1.17	1.05	1.05
1979	3.69	1.36	1.94	1.28	1.21	1.48	1.49	1.16	1.09	1.04	1.06	1.11
1980	4.88	2.71	3.11	1.21	1.20	1.33	1.20	1.22	1.05	1.07	1.02	
1981	2.42	2.39	1.98	1.61	1.25	1.23	1.24	1.12	1.05	1.05		
1982	7.11	1.35	1.55	1.46	1.39	1.60	1.36	1.05	1.10			
1983	16.26	1.46	1.30	1.38	1.36	1.17	1.22	1.06				
1984	2.73	1.99	1.39	1.22	1.18	1.26	1.17					
1985	14.10	1.98	1.95	1.34	1.21	1.48						
1986	4.69	1.57	1.78	1.69	1.34							
1987	2.41	2.09	1.81	1.55								
1988	2.44	6.13	1.64									
1989	10.42	2.23										
1990	2.70											
1991												

Note that link ratios do not tell us that we have a constant stable calendar year trend

APPENDIX C1

ACC. YEAR	DELAY 0	1	2	3	4	5	6	7	8	9	10
1977	153638	188412	134534	87456	60348	42404	31238	21252	16622	14440	12200
1978	176536	226412	158894	104686	71448	47990	35576	24818	22662	18000	
1979	210172	259168	188388	123074	83380	56086	38496	33768	27400		
1980	11448	253482	183370	131040	78994	60232	45568	38000			
1981	219810	266304	194650	120098	87582	62750	51000				
1982	205654	252746	177506	12952	96786	82400					
1983	197716	255408	194648	142328	105600						
1984	239784	329242	264802	190400							
1985	326304	471744	375400								
1986	420778	590400									
1987	496200										

ACCI EXPOSURES

YEAR	
1977	2.20
1978	2.40
1979	2.20
1980	2.00
1981	1.90
1982	1.60
1983	1.60
1984	1.80
1985	2.20
1986	2.50
1987	2.60

APPENDIX C2

AGE LINK RATIOS OF CUMULATIVE PAYMENTS

DELAYS

	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10
1977	2.226337	1.393316	1.183505	1.106992	1.067912	1.046848	1.030445	1.023109	1.019622	1.016259
1978	2.268158	1.392381	1.185665	1.106873	1.064853	1.045149	1.030135	1.026712	1.020665	
1979	2.233123	1.401389	1.187119	1.106787	1.064900	1.041831	1.035220	1.027606		
1980	2.198791	1.394403	1.202128	1.101360	1.070173	1.049607	1.039413			
1981	2.211519	1.400420	1.176416	1.109359	1.070629	1.053616				
1982	2.228986	1.387229	1.203681	1.126446	1.095567					
1983	1.291792	1.429568	1.219719	1.133653						
1984	2.373077	1.465360	1.228344							
1985	2.445719	1.470397								
1986	2.403115									

APPENDIX C3

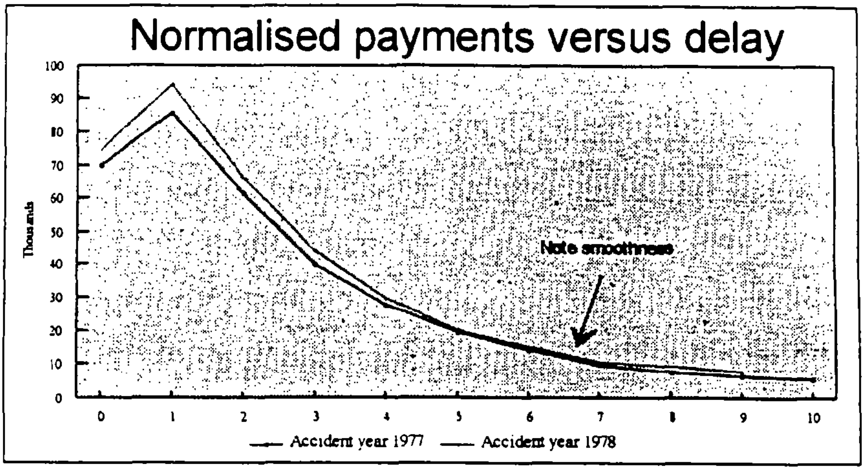


Figure C3

APPENDIX C4 - (Statistical Chain Ladder)

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	0.2511	0.0370	6.79			
2	-0.3069	0.0385	-7.97	-0.5580	0.0650	-8.59
3	-0.3928	0.0406	-9.68	-0.0859	0.0682	-1.26
4	-0.3803	0.0432	-8.81	0.0124	0.0723	0.17
5	-0.3402	0.0464	-7.34	0.0401	0.0773	0.52
6	-0.3384	0.0505	-6.71	0.0018	0.0835	0.02
7	-0.2908	0.0559	-5.20	0.0476	0.0917	0.52
8	-0.2248	0.0637	-3.53	0.0660	0.1030	0.64
9	-0.2152	0.0763	-2.82	0.0095	0.1202	0.08
0	-0.1893	0.1030	-1.84	0.0259	0.1526	0.17

NOT ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1977	11.0484	0.0380	290.75			
1978	11.1402	0.0380	293.17	0.0918	0.0370	2.48
1979	11.3935	0.0385	295.97	0.2533	0.0385	6.58
1980	11.5218	0.0393	293.10	0.1283	0.0406	3.16
1981	11.6001	0.0405	286.71	0.0783	0.0432	1.81
1982	11.7939	0.0420	280.55	0.1938	0.0464	4.18
1983	11.7979	0.0442	266.67	0.0040	0.0505	0.08
1984	11.9095	0.0474	251.04	0.1115	0.0559	1.99
1985	12.0116	0.0524	229.08	0.1022	0.0637	1.60
1986	12.0774	0.0613	196.88	0.0657	0.0763	0.86
1987	12.1592	0.0827	147.00	0.0818	0.1030	0.79

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX C4

(REGRESSION OUTPUT CONTINUED)

S = 0.0827 S-SQUARED = 0.0068 S-SQUARED(SCI) = 0.0449

S(B) = 0.0827 S(B)-SQUARED = 0.0068 DELTA = 0.0000

R-SQUARED = 99.5 PERCENT N = 66 P = 21.0

SSPE = 0.948 WSSPE = 0.948 AIC = -124.97 AIC(SCI) = -52.18

APPENDIX C5

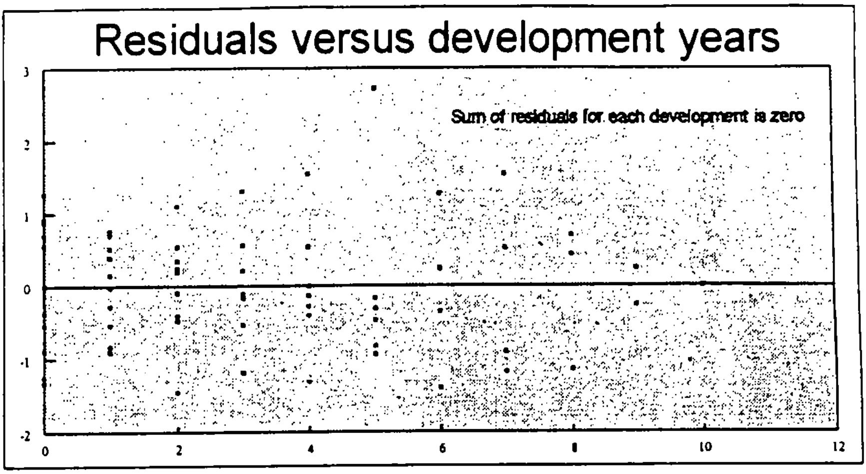


Figure C5.1

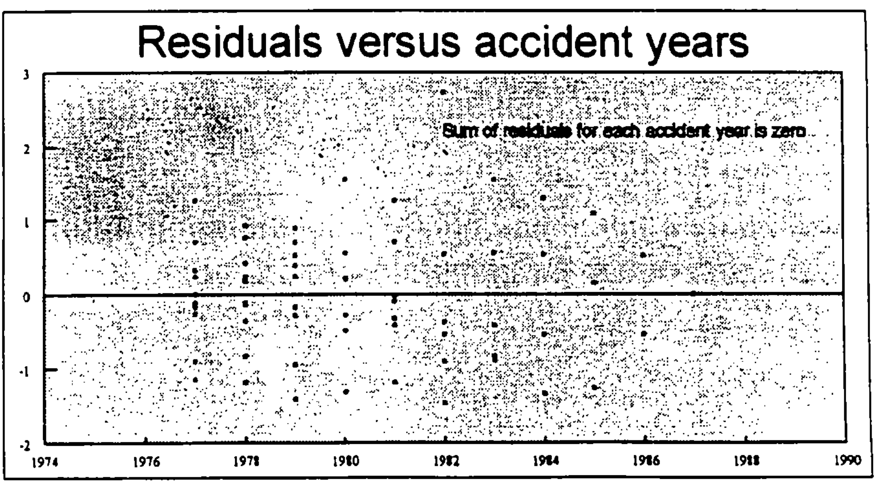


Figure C5.2

APPENDIX C5

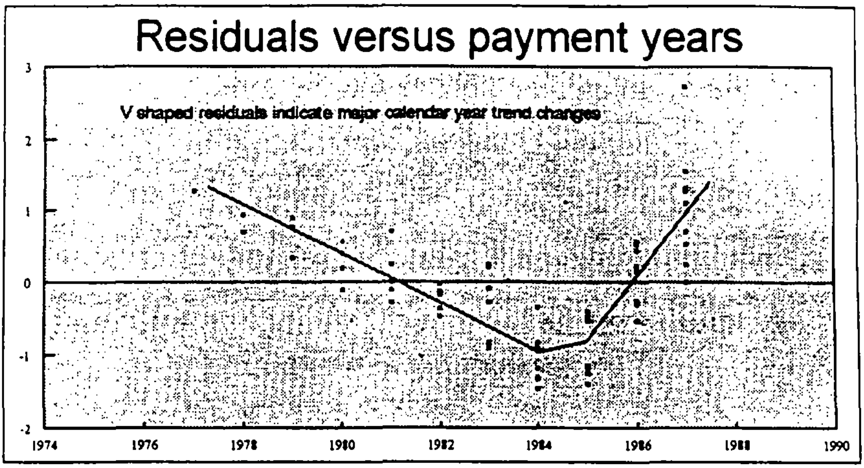


Figure C5.3

APPENDIX C6 - Cape Cod

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	0.2029	0.1416	1.43			
2	-0.3567	0.1489	-2.40	-0.5596	0.2514	-2.23
3	-0.4468	0.1574	-2.84	-0.0901	0.2651	-0.34
4	-0.4352	0.1677	-2.59	0.0116	0.2814	0.04
5	-0.3947	0.1803	-2.19	0.0404	0.3010	0.13
6	-0.4139	0.1962	-2.11	-0.0192	0.3256	-0.06
7	-0.3556	0.2174	-1.64	0.0583	0.3574	0.16
8	-0.3067	0.2475	-1.24	.0.0489	0.4012	0.12
9	-0.3150	0.2958	-1.06	-0.0083	0.4677	-0.02
10	-0.2352	0.3968	-0.59	0.0797	0.5916	0.13

NOT ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1977	11.6776	0.0977	119.53			
1978	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1979	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1980	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1981	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1982	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1983	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1984	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1985	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1986	11.6776	0.0977	119.53	0.0000	0.0000	0.00
1987	11.6776	0.0977	119.53	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX C6

(REGRESSION OUTPUT CONTINUED)

S = 0.3240 S-SQUARED = 0.1050 S-SQUARED(SCI) = 0.0449

S(B) = 0.3240 S(B)-SQUARED = 0.1050 DELTA = 0.0000

R-SQUARED = 91.1 PERCENT N = 66 P = 11.0

SSPE = 7.433 WSSPE = 7.433 AIC = 48.51 AIC(SCI) = -52.18

APPENDIX C7

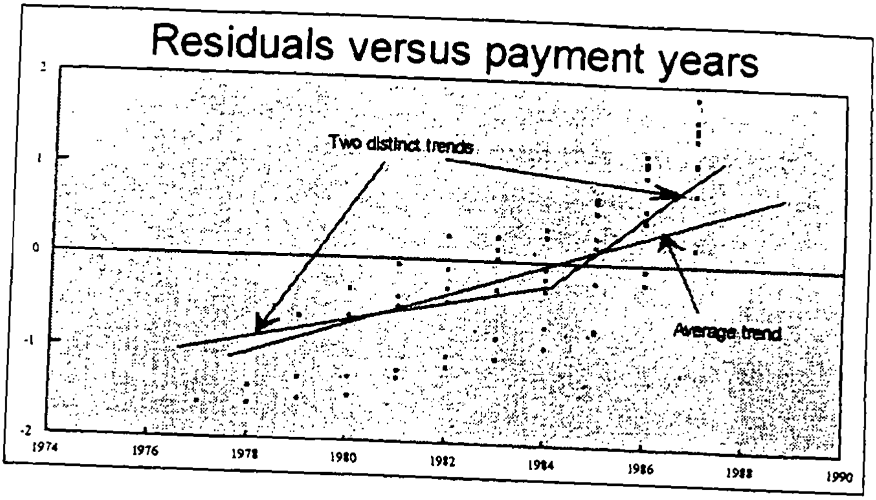


Figure C7

APPENDIX C8 - Cape Cod with constant inflation

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	-RATIO
1	0.1424	0.0439	3.24			
2	-0.4172	0.0462	-9.03	-0.5596	0.0779	-7.19
3	-0.5072	0.0488	-10.39	-0.0901	0.0821	-1.10
4	-0.4956	0.0520	-9.53	0.0116	0.0871	0.13
5	-0.4552	0.0559	-8.14	0.0404	0.0932	0 43
6	-0.4744	0.0608	-7.80	-0.0192	0.1008	-0.19
7	-0.4161	0.0674	-6.18	0.0583	0.1107	0.53
8	-0.3672	0.0767	-4.79	0.0489	0.1243	0.39
9	-0.3754	0.0917	-4.10	-0.0083	0.1449	-0.06
10	-0.2957	0.1230	-2.41	0.0797	0.1832	0.44

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1977	11.0728	0.0403	275.09			
1978	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1979	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1980	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1981	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1982	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1983	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1984	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1985	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1986	11.0728	0.0403	275.09	0.0000	0.0000	0.00
1987	11.0728	0.0403	275.09	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

PMNT YEAR	IOTA	S.E.	T-RATIO	DIFFERENCE IN IOTA	S.E.	T-RATIO
1978	0.1210	0.0053	22.79			
1979	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1980	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1981	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1982	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1983	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1984	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1985	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1986	0.1210	0.0053	22.79	0.0000	0.0000	0.00
1987	0.1210	0.0053	22.79	0.0000	0.0000	0 00

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX C8

(REGRESSION OUTPUT CONTINUED)

S = 0.1004 S-SQUARED = 0.0101 S-SQUARED(SCI) = 0.0449

S(B) = 0.1004 S(B)-SQUARED = 0.0101 DELTA = 0.0000

R-SQUARED = 99.2 PERCENT N = 66 P = 12.0

SSPE = 1.176 WSSPE = 1.176 AIC = -105.40 AIC(SCI) = -52.18

APPENDIX C9

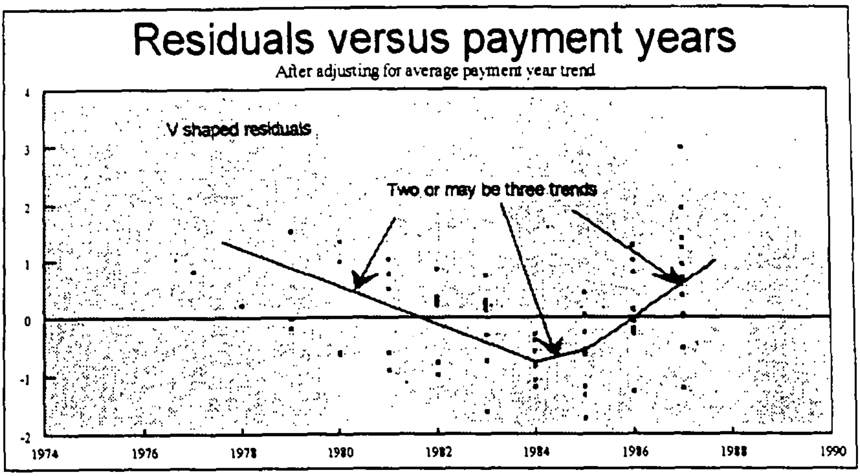


Figure C9

APPENDIX C10 - Cape Cod with three payment year parameters
(1977-84, 1984-1985 and 1985-1987)

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	-0.1505	0.0371	4.05			
2	-0.4098	0.0390	-10.50	-0.5603	0.0657	-8.52
3	-0.5008	0.0413	-12.14	-0.0910	0.0693	-1.31
4	-0.4906	0.0439	-11.17	0.0102	0.0736	0.14
5	-0.4522	0.0472	-9.58	0.0384	0.0787	0.49
6	-0.4748	0.0514	-9.24	0.0225	0.0851	-0.26
7	-0.4222	0.0569	-7.41	0.0526	0.0935	0.56
8	-0.3849	0.0651	-5.91	0.0373	0.1050	0.36
9	-0.4126	0.0780	-5.29	-0.0277	0.1229	-0.23
10	-0.3329	0.1042	-3.19	0.0797	0.1547	0.52

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1977	11.1536	0.0400	278.91			
1978	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1979	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1980	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1981	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1982	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1983	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1984	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1985	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1986	11.1536	0.0400	278.91	0.0000	0.0000	0.00
1987	11.1536	0.0400	278.91	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

PMNT YEAR	IOTA	S.E.	T-RATIO	DIFFERENCE IN IOTA	S.E.	T-RATIO
1978	0.0985	0.0077	12.74			
1979	0.0985	0.0077	12.74	0.0000	0.0000	0.00
1980	0.0985	0.0077	12.74	0.0000	0.0000	0.00
1981	0.0985	0.0077	12.74	0.0000	0.0000	0 00
1982	0.0985	0.0077	12.74	0.0000	0.0000	0.00
1983	0.0985	0.0077	12.74	0.0000	0.0000	0.00
1984	0.0985	0.0077	12.74	0.0000	0.0000	0.00
1985	0.1174	0.0343	3.42	0.0189	0.0385	0.49
1986	0.1952	0.0197	9.91	0.0778	0.0484	1.61
1987	0.1952	0.0197	9.91	0.0000	0.0000	0.00

APPENDIX C10

(REGRESSION OUTPUT CONTINUED)

S = 0.0847 S-SQUARED = 0.0072 S-SQUARED(SCI) = 0.0449

S(B) = 0.0847 S(B)-SQUARED = 0.0072 DELTA = 0.0000

R-SQUARED = 99.4 PERCENT N = 66 P = 14.0

SSPE = 1.000 WSSPE = 1.000 AIC = -126.26 AIC(SCI) = -52.18

APPENDIX C11

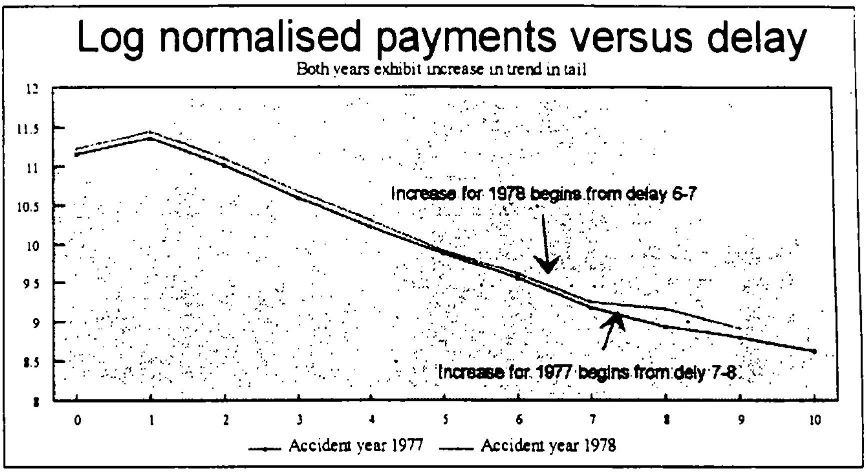


Figure C11

APPENDIX D1

INCREMENTAL PAID LOSSES

DELAY

ACCI. YR	0	1	2	3	4	5	6	7	8
1969	193013	1584331	1151882	778980	475203	143352	128612	70845	25077
1970	376473	1541950	1719509	1032570	289305	382508	270087	108354	23133
1971	568891	1579158	1277822	734670	680369	217221	147800	57099	64829
1972	428753	970640	955898	1095771	510072	491853	242995	299845	
1973	458252	989072	1417606	953222	881133	278778	197156		
1974	355229	948807	1292900	748003	547288	274367			
1975	282419	688332	1158793	903450	629983				
1976	267600	1044790	1216437	527644					
1977	560307	940002	1185899						
1978	360171	1011773							
1979	445545								

ACCI YR	EXPOSURES
1969	523.00
1970	643.00
1971	676.00
1972	673.00
1973	809.00
1974	669.00
1975	513.00
1976	543.00
1977	622.00
1978	703.00
1979	743.00

APPENDIX D2

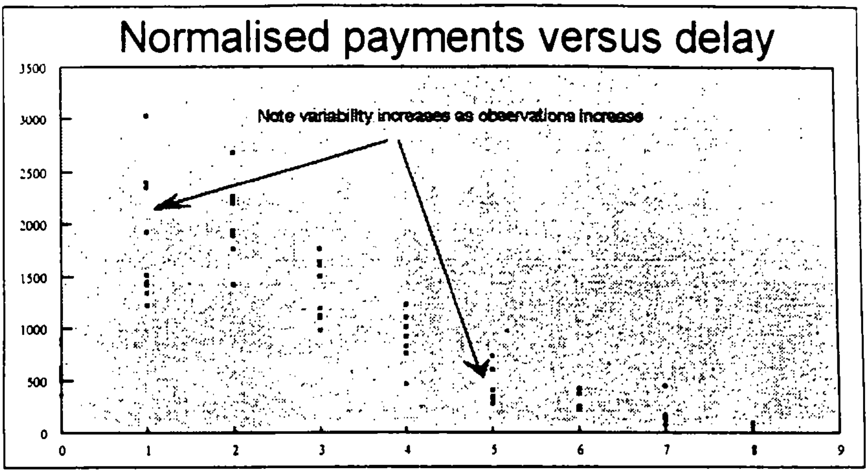


Figure D2.1

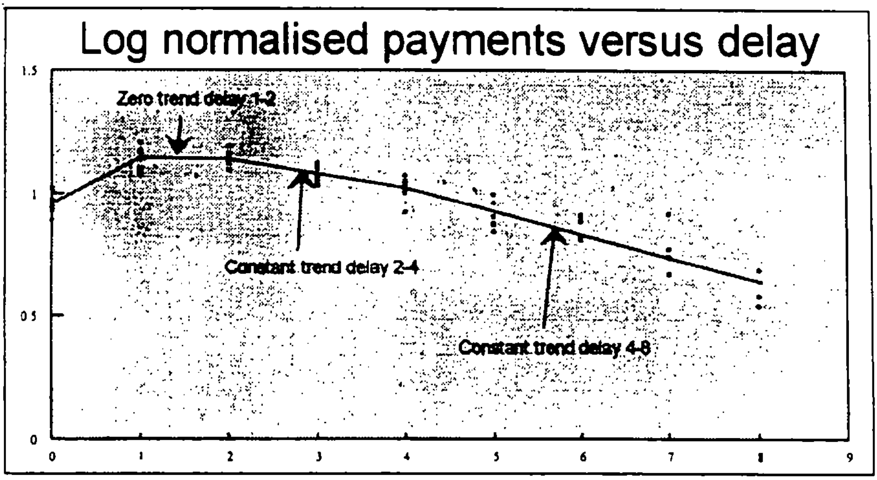


Figure D2.2

APPENDIX D3

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	1.1647	0.1234	9.44			
2	0.0000	0.0000	0.00	-1.1647	0.1234	-9.44
3	-0.3769	0.0631	-5.98	-0.3769	0.0631	-5.98
4	-0.3769	0.0631	-5.98	0.0000	0.0000	0.00
5	-0.6226	0.0466	-13.35	-0.2457	0.0985	-2.49
6	-0.6226	0.0466	-13.35	0.0000	0.0000	0.00
7	-0.6226	0.0466	-13.35	0.0000	0.0000	0.00
8	-0.6226	0.0466	-13.35	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1969	6.3672	0.0997	63.84			
1970	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1971	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1972	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1973	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1974	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1975	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1976	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1977	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1978	6.3672	0.0997	63.84	0.0000	0.0000	0.00
1979	6.3672	0.0997	63.84	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

PMNT YEAR	IOTA	S.E.	T-RATIO	DIFFERENCE IN IOTA	S.E.	T-RATIO
1970	0.0000	0.0000	0.00			
1971	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1972	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1973	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1974	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1975	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1976	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1977	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1978	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1979	0.0000	0.0000	0.00	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX D4

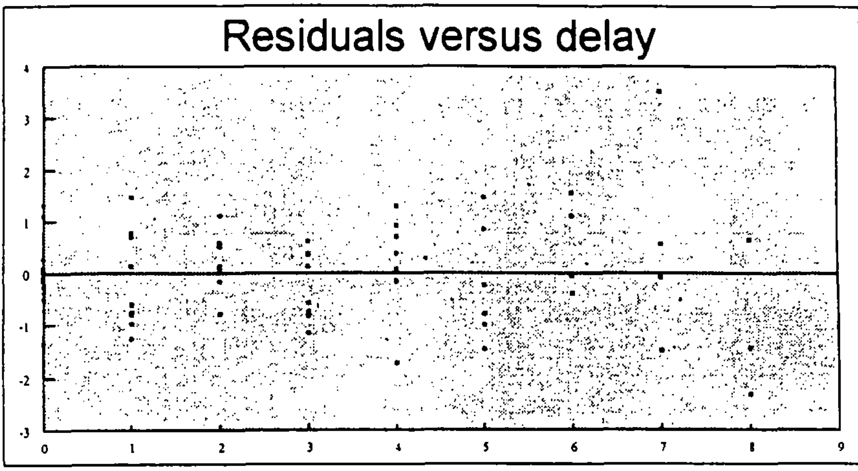


Figure D4.1

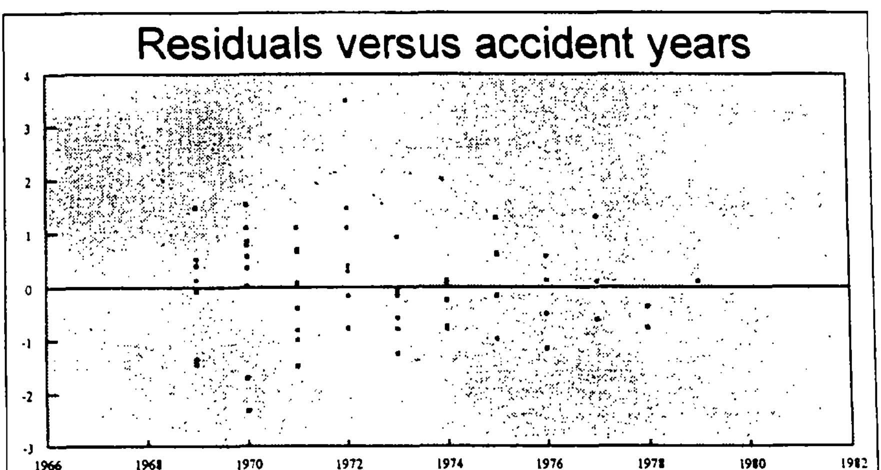


Figure D4.2

APPENDIX D4

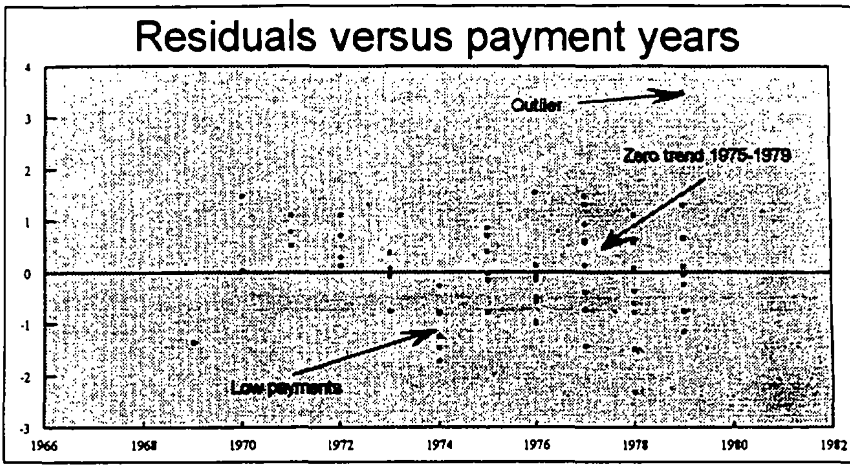


Figure D4.3

APPENDIX D5

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	1.1777	0.0993	11.86			
2	0.0000	0.0000	0.00	-1.1777	0.0993	-11.86
3	-0.3478	0.0519	-6.70	-0.3478	0.0519	-6.70
4	-0.3478	0.0519	-6.70	0.0000	0.0000	0.00
5	-0.6749	0.0390	-17.32	-0.3270	0.0803	-4.07
6	-0.6749	0.0390	-17.32	0.0000	0.0000	0.00
7	-0.6749	0.0390	-17.32	0.0000	0.0000	0.00
8	-0.6749	0.0390	-17.32	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1969	6.4594	0.0927	69.68			
1970	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1971	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1972	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1973	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1974	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1975	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1976	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1977	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1978	6.4594	0.0927	69.68	0.0000	0.0000	0.00
1979	6.4594	0.0927	69.68	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

PMNT YEAR	IOTA	S.E.	T-RATIO	DIFFERENCE IN IOTA	S.E.	T-RATIO
1970	0.0000	0.0000	0.00			
1971	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1972	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1973	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1974	-0.4792	0.1306	-3.67	-0.4792	0.1306	-3.67
1975	0.3723	0.1182	3.15	0.8515	0.2330	3.65
1976	0.0000	0.0000	0.00	-0.3723	0.1182	-3.15
1977	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1978	0.0000	0.0000	0.00	0.0000	0.0000	0 00
1979	0.0000	0.0000	0.00	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX D6

(REGRESSION OUTPUT CONTINUED)

S = 0.2654 S-SQUARED = 0.0704 S-SQUARED(SCI) = 0.5469

S(B) = 0.2654 S(B)-SQUARED = 0.0704 DELTA = 0.0000

R-SQUARED = 93.5 PERCENT N = 62 P = 6.0

SSPE = 7.360 WSSPE = 7.360 AIC = 17.13 AIC(SCI) = 43.81

APPENDIX D7

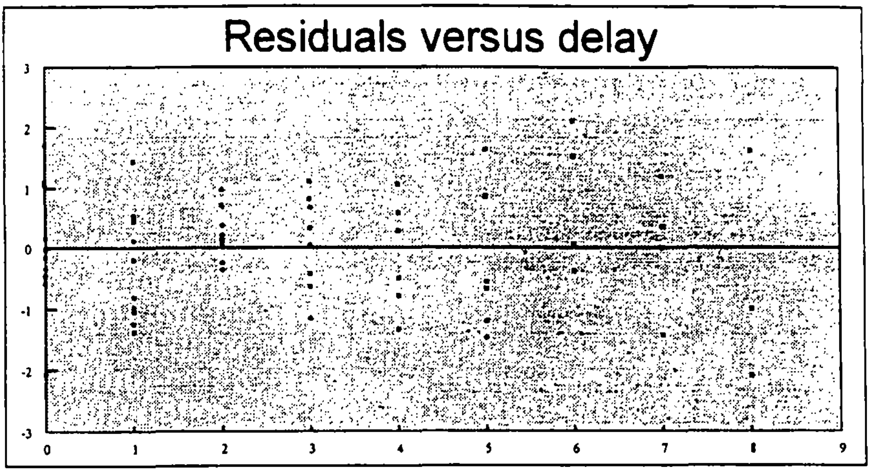


Figure D7.1

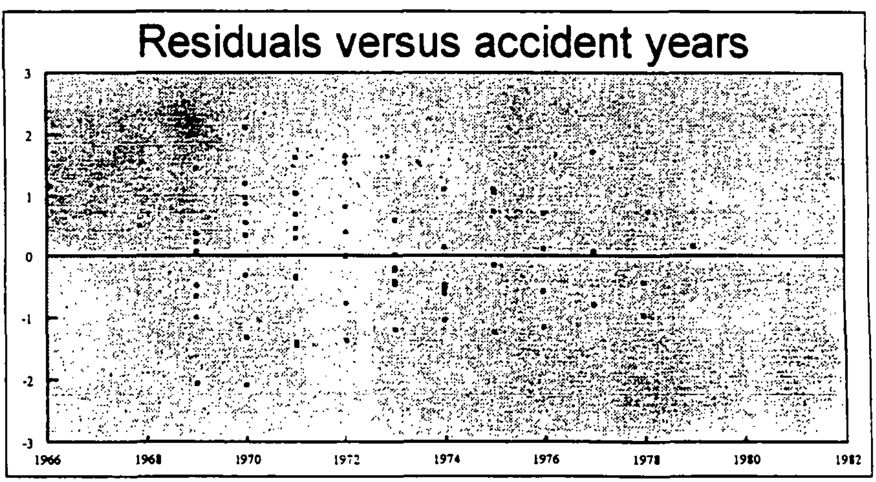


Figure D7.2

APPENDIX D7

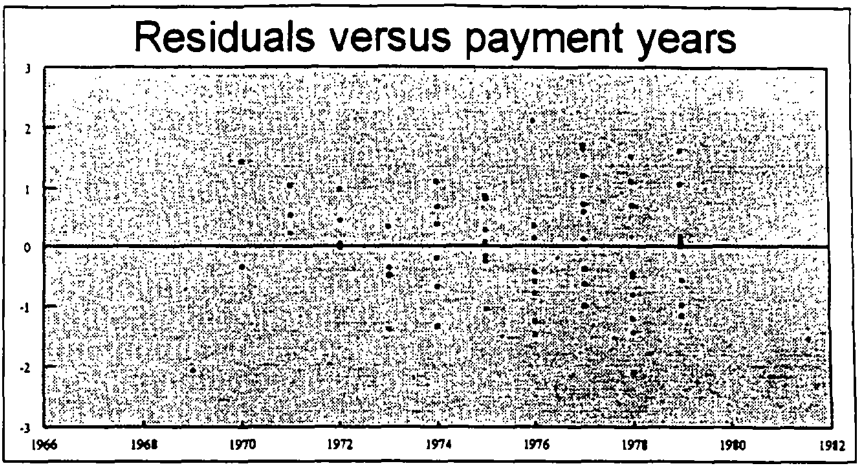


Figure D7.3

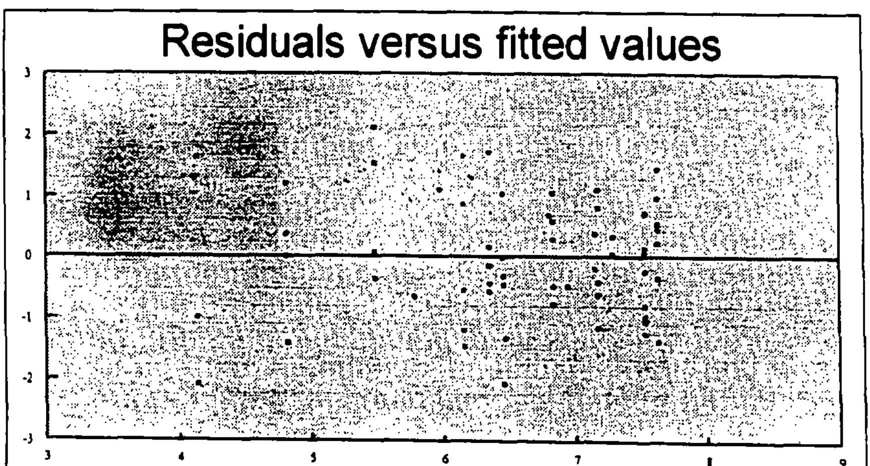


Figure D7. 4

APPENDIX D7

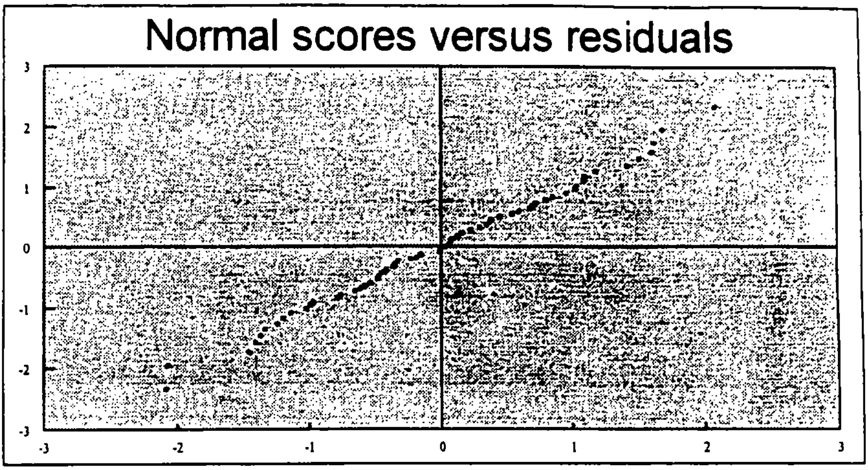


Figure D7.5

FORECASTING OUTPUT

ASSUMED FUTURE INFLATION = 0.0000
STANDARD ERROR = 0.0000

EXPECTED PAYMENTS/OBSERVED PAYMENTS +--------+ **FORECAST MEAN PAYMENTS/STANDARD ERRORS**

(PAYMENTS IN $1 S)

EXP:	346295	1123112	1123112	793172	561671	177321	130329	66469	33951	0
OBS:	193013	1584331	1151882	778980	475203	143352	128612	70845	25077	0
EXP:	425750	1380806	1380806	975161	428440	314654	160233	81720	41741	0
OBS:	376473	1541950	1719509	1032570	289305	382508	270087	108354	23133	0
EXP:	447601	1451671	1451671	636733	650595	330803	168456	85914	43883	0
OBS:	568891	1579158	1277822	734670	680369	217221	147800	57099	64829	0
EXP:	445614	1445229	898519	915576	647708	329335	167709	85533	43689	43689
OBS:	428753	970640	955898	1095771	510072	491853	242995	299845	12280	12280
EXP:	535664	1080091	1559962	1100596	778597	395887	201599	102817	52517	155334
OBS:	458252	989072	1417606	953222	881133	278778	197156	27673	14762	32822
:XP:	275565	1290006	1290006	910134	643858	327378	166712	85024	43429	295165
)BS:	355229	948807	1292900	748003	547288	274367	43841	22884	12207	53323
:XP:	305191	989197	989197	697906	493721	251038	127837	65198	33302	477376
)BS:	282419	688332	1158793	903450	629983	65999	33618	17548	9361	79132
:XP:	323039	1047045	1047045	738719	522593	265719	135313	69011	35249	1027886
)BS:	267600	1044790	1216437	527644	140549	69858	35584	18574	9908	167258
:XP:	370037	1199377	1199377	846194	598624	304378	155000	79051	40378	2023625
)BS:	560307	940002	1185899	221766	160997	80022	40761	21276	11350	300456
:XP:	418225	1355566	1355566	956389	676580	344016	175185	89345	45636	3642717
)BS:	360171	1011773	360708	250646	181963	90443	46069	24047	12827	502218
:XP:	442022	1432697	1432697	1010807	715077	363590	185152	94429	48233	5282681
)BS:	445545	381231	381231	264907	192317	95589	48690	25415	13557	674135
PAYMENT YRS:		4721306	3518808	2235705	1316405	653075	314876	140065	48233	12948473
RD ERRORS:		623018	504462	345451	223516	111688	57849	29752	13557	1030808

APPENDIX D9

TABLE OF OBSERVED AND EXPECTED BY YEAR

ACC. YEAR	EXPECTED	OBSERVED (PAYMENTS IN $1'S)	DIFFERENCE	%ERROR	PMNT YEAR	EXPECTED	OBSERVED	DIFFERENCE (PAYMENTS IN $1's)	%ERROR
69	4355433	4551295	195862	4	69	346295	193013	-153282	-79
70	5189312	5743889	554577	9	70	1548863	1960804	411941	21
71	5267328	5327859	60531	1	71	2951519	3262723	311204	9
72	4849689	4695982	-153707	-3	72	4071263	4506400	435137	9
73	5652397	5175219	-477178	-9	73	4969396	4214487	-754909	-17
74	4736946	4166594	-570352	-13	74	3496670	3467526	-29144	0
75	3475212	3662977	187765	5	75	5166314	4936092	-230222	-4
76	3155847	3056471	-99376	-3	76	4908050	4270279	-637771	-14
77	2768792	2686208	-82584	-3	77	4708472	5166110	457638	8
78	1773792	1371944	-401848	-29	78	4697662	4569353	-128309	-2
79	442022	445545	3523	0	79	4802265	4337196	-465069	-10

NDI)

FORECASTING OUTPUT

ASSUMED FUTURE INFLATION = 0.0000
STANDARD ERROR = 0.0000

EXPECTED PAYMENTS/OBSERVED PAYMENTS (PAYMENTS IN $1 S) FORECAST MEAN PAYMENTS/STANDARD ERROR

EAR											
69	EXP:	346295	1123112	1123112	793172	561671	177321	130329	66469	33951	
	OBS:	193013	1584331	1151882	778980	475203	143352	128612	70845	25077	
70	EXP:	425750	1380806	1380806	975161	428440	314654	160233	81720	41741	
	OBS:	376473	1541950	1719509	1032570	289305	382508	270087	108354	23133	
71	EXP:	447601	1451671	1451671	636733	650595	330803	168456	85914	43883	
	OBS:	568891	1579158	1277822	734670	680369	217221	147800	57099	64829	
72	EXP:	445614	1445229	898519	915576	647708	329335	167709	85533	43689	4368
	OBS:	428753	970640	955898	1095771	510072	491853	242996	299845	12280	1228
73	EXP:	535664	1080091	1559962	1100596	778597	395887	201599	102817	52517	15533
	OBS:	458252	989072	1417606	953222	881133	278778	197156	27673	14762	3282
74	EXP:	275565	1290006	1290006	910134	643858	327378	166712	85024	43429	29516
	OBS:	355229	948807	1292900	748003	547288	274367	43841	22884	12207	5332
75	EXP:	305191	989197	989197	697906	493721	251038	127837	65198	33302	47737
	OBS:	282419	688332	1158793	903450	629983	65999	33618	17548	9361	7913
76	EXP:	323039	1047045	1047045	738719	522593	265719	135313	69011	35249	102788
	OBS:	267600	1044790	1216437	527644	140549	69858	35584	18574	9908	16725
77	EXP:	370037	1199377	1199377	846194	598624	304378	155000	79051	40378	202362
	OBS:	560307	940002	1185899	221766	160997	80022	40761	21276	11350	30045
78	EXP:	418225	1355566	1355566	956389	676580	344016	175185	89345	45636	364271
	OBS:	360171	1011773	360708	250646	181963	90443	46069	24047	12827	50221
79	EXP:	442022	1432697	1432697	1010807	715077	363590	185152	94429	48233	528268
	OBS:	445545	381231	381231	264907	192317	95589	48690	25415	13557	67413
)T.FOR PAYMENT YRS:			4721306	3518808	2235705	1316405	653075	314876	140065	48233	1294847
TANDARD ERRORS:			623018	504462	345451	223516	111688	57849	29752	13557	103080

APPENDIX D10

VALIDATION

REGRESSION TABLE

PARAMETER ESTIMATES

DEV. YEAR	GAMMA	S.E.	T-RATIO	DIFFERENCE IN GAMMA	S.E.	T-RATIO
1	1.2468	0.1076	11.58			
2	0.0000	0.0000	0.00	-1.2468	0.1076	-11.58
3	-0.4024	0.0639	-6.29	-0.4024	0.0639	-6.29
4	-0.4024	0.0639	-6.29	0.0000	0.0000	0.00
5	-0.5544	0.0753	-7.37	-0.1520	0.1213	-1.25
6	-0.5544	0.0753	-7.37	0.0000	0.0000	0.00
7	-0.5544	0.0753	-7.37	0.0000	0.0000	0.00
8	-0.5544	0.0753	-7.37	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

ACCI YEAR	ALPHA	S.E.	T-RATIO	DIFFERENCE IN ALPHA	S.E.	T-RATIO
1969	6.4278	0.0922	69.72			
1970	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1971	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1972	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1973	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1974	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1975	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1976	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1977	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1978	6.4278	0.0922	69.72	0.0000	0.0000	0.00
1979	6.4278	0.0922	69.72	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

PARAMETER ESTIMATES

PMNT YEAR	IOTA	S.E.	T-RATIO	DIFFERENCE IN IOTA	S.E.	T-RATIO
1970	0.0000	0.0000	0.00			
1971	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1972	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1973	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1974	-0.4798	0.1208	-3.97	-0.4798	0.1208	-3.97
1975	0.3087	0.1203	2.57	0.7886	0.2196	3.59
1976	0.0000	0.0000	0.00	-0.3087	0.1203	-2.57
1977	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1978	0.0000	0.0000	0.00	0.0000	0.0000	0.00
1979	0.0000	0.0000	0.00	0.0000	0.0000	0.00

ALL PARAMETERS ARE SIGNIFICANT

APPENDIX D11

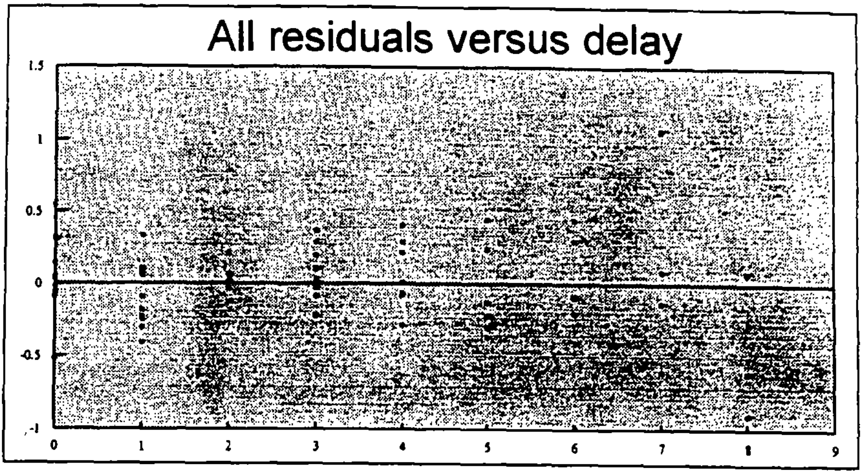


Figure D11.1

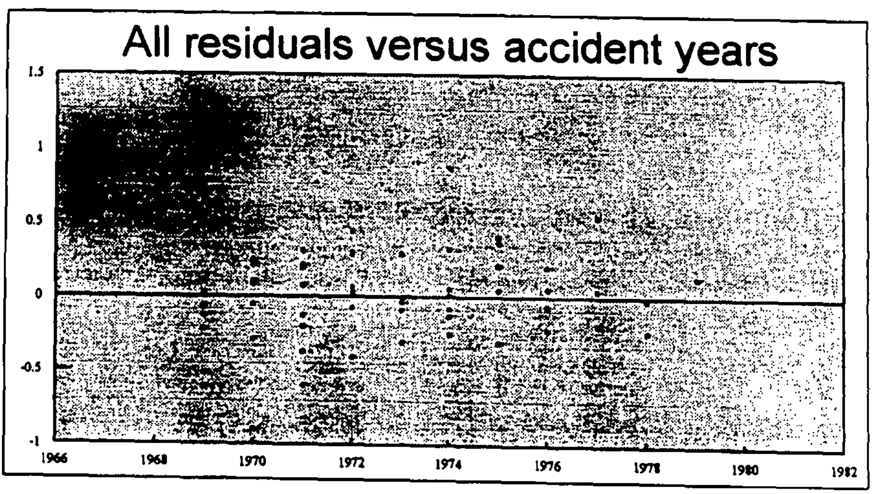


Figure D11.2

APPENDIX D11

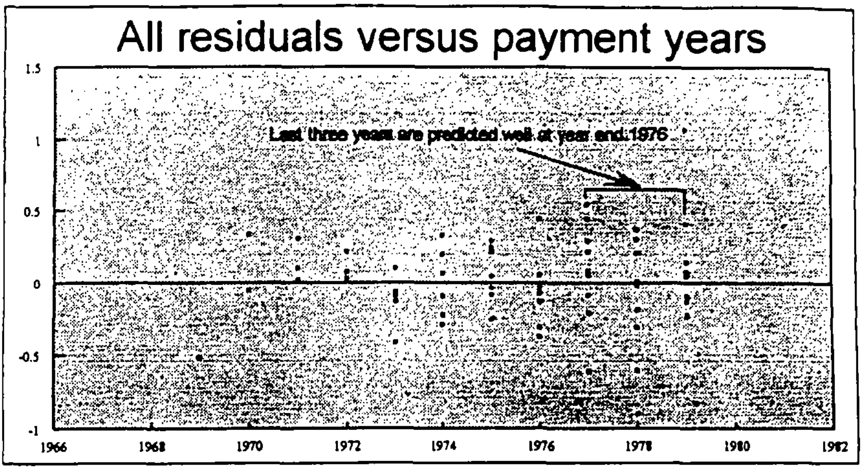


Figure D11.3

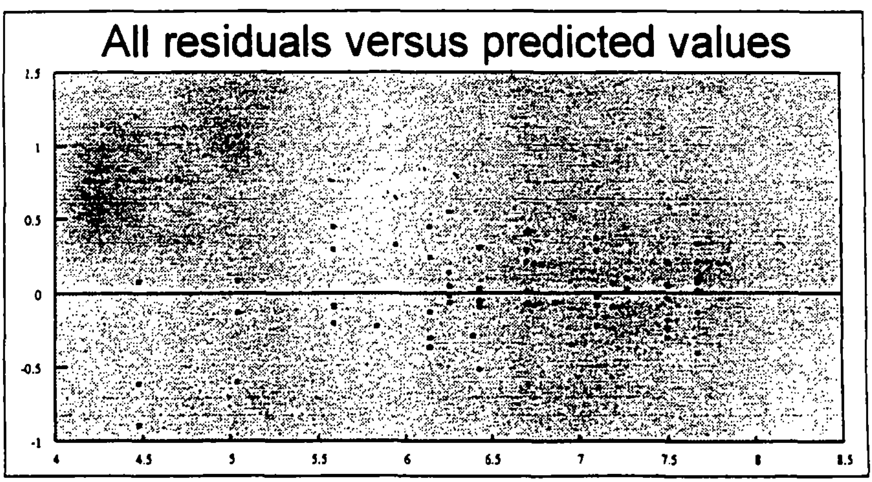


Figure D11.4

APPENDIX D12

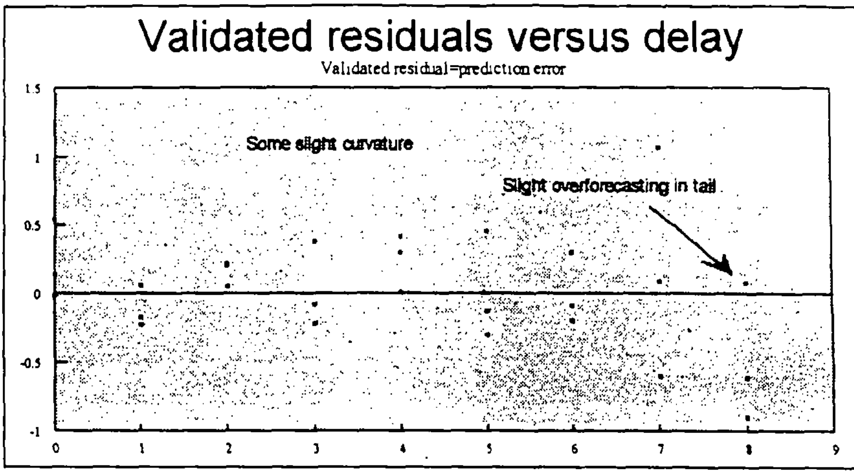


Figure D12.1

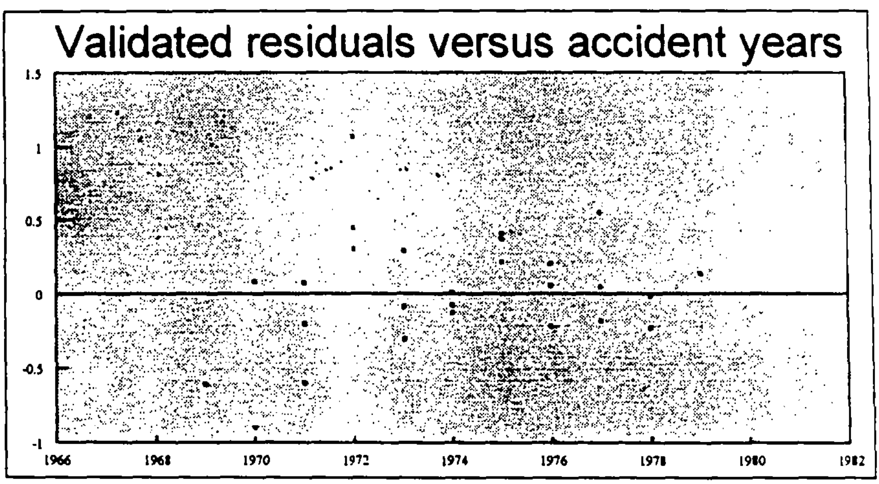


Figure D12.2

APPENDIX D12

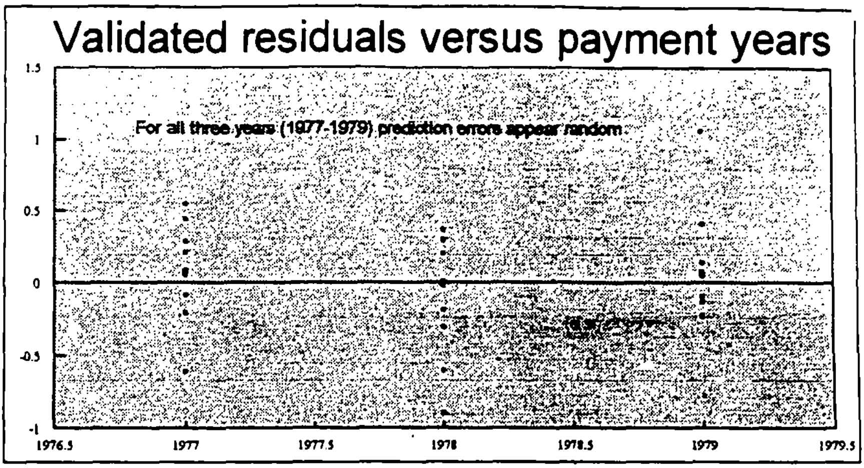


Figure D12.3

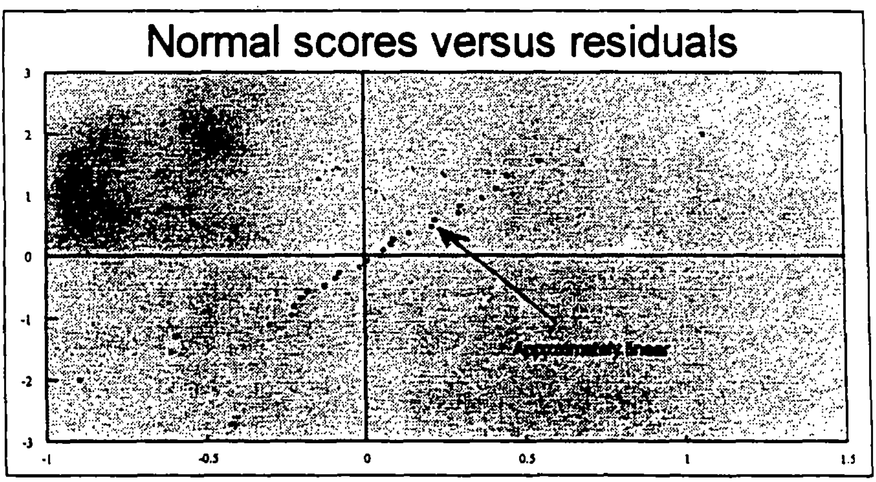


Figure D12.4

FORECASTING OUTPUT

.IDATION MODEL

ASSUMED FUTURE INFLATION = 0.0000

STANDARD ERROR = 0.0000

EXPECTED PAYMENTS/OBSERVED PAYMENTS +------+ **FORECAST MEAN PAYMENTS/STANDARD ERROR:**

(PAYMENTS IN $1 S)

ıR											
9	EXP:	333078	1157384	1157384	774069	519825	184730	144344	83624	48721	0
	OBS:	193013	1584331	1151882	778980	475203	143352	128612	70845	25077	0
)	EXP:	409501	1422941	1422941	951676	395746	308064	177464	102811	59900	0
	OBS:	376473	1541950	1719509	1032570	289305	382508	270087	108354	23133	0
	EXP:	430518	1495969	1495969	620480	565416	323874	186571	108087	62974	0
	OBS:	568891	1579158	1277822	734670	680369	217221	147800	57099	64829	0
	EXP:	428607	1489330	925009	839313	562907	322437	185743	107607	62695	62695
	OBS:	428753	970640	955898	1095771	510072	491853	242995	299845	20985	20985
	EXP:	515220	1111935	1510497	1008922	676660	387595	223278	129353	75364	204717
	OBS:	582529	989072	1417606	953222	881133	278778	197156	36802	25226	52836
	EXP:	264794	1249101	1249101	834325	559561	320521	184639	106968	62322	353929
	OBS:	355229	948807	1292900	748003	547288	274367	46231	30433	20860	72717
	EXP:	275840	957831	957831	639774	429081	245780	141584	82025	47790	517179
	OBS:	282419	688332	1158793	903450	629983	58502	35451	23336	15996	85810
	EXP:	291971	1013844	1013844	677187	454173	260153	149864	86821	50584	1001596
	OBS:	267600	1044790	1216437	527644	113657	61923	37524	24701	16931	147670
	EXP:	334450	1161346	1161346	775710	520250	298003	171668	99453	57944	1923027
	OBS:	560307	940002	1185899	184214	130192	70932	42983	28295	19395	262653
	EXP:	378003	1312583	1312583	876727	588000	336810	194023	112404	65489	3486036
	OBS:	360171	1011773	318349	208204	147147	80169	48581	31980	21920	452821
	EXP:	399511	1387267	1387267	926612	621456	355974	205063	118800	69216	5071655
	OBS:	445545	336463	336463	220050	155519	84731	51345	33799	23168	625691
FOR PAYMENT YRS:			4552200	3368314	2106825	1264545	700034	375411	184289	69216	12620833
ꞁDARD ERRORS:			578766	453677	299060	193348	117101	77112	47562	23168	1072089

A___IDI›___

TABLE OF OBSERVED AND EXPECTED BY YEAR

ACC. YEAR	EXPECTED	OBSERVED	DIFFERENCE	%ER	PMNT YEAR	EXPECTED	OBSERVED	DIFFERENCE	%ER
	(PAYMENTS IN $1'S)					(PAYMENTS IN $1's)			
69	4403160	4551295	148135	3	69	333078	193013	-140065	-72
70	7251043	5743889	492846	8	70	1566886	1960804	393918	20
71	5289859	5327859	38000	0	71	3010843	3262723	251880	7
72	4753347	4695982	-57365	-1	72	4121587	4506400	384813	8
73	5434107	5175219	-258888	-5	73	4972020	4214487	-757533	-17
74	4477402	4166594	-310808	-7	74	3502693	3467526	-35167	-1
75	3260356	3662977	402621	10	75	4892575	4936092	43517	0
76	2996847	3056471	59624	1	76	4655693	4270279	-385414	-9
77	2657142	2686208	29066	1	77	4477648	5166110	688462	13
78	1690586	1371944	-318642	-23	78	4493854	4569353	75499	1
79	399511	445545	46034	10	79	4586481	4337196	-249285	-5

IBNR Reserve Under a Loglinear Location-Scale Regression Model

by Louis Doray

IBNR RESERVE UNDER A LOGLINEAR LOCATION-SCALE REGRESSION MODEL [1]

Abstract

In this paper, we develop models for known claims, when the data are grouped into the usual triangle and the goal is to predict IBNR claims. We assume that the payment for a certain accident and development year is composed of a deterministic part and a multiplicative random error. We use a loglinear location-scale regression model for the amount of claims. The parameters are estimated by maximum likelihood methods, so that their asymptotic properties are well known. The regression model presents many advantages over the chain ladder method: it has fewer parameters, and does not underestimate the reserve. Moreover, it will be possible with a simulation to establish a reserve with a certain level of confidence (for example 80%).

The logarithm of the error is assumed to follow certain known distributions (normal, extreme value, generalized loggamma, logistic and log inverse gaussian). We derive certain theoretical properties of these distributions and prove that the MLE's of the regression and scale parameters exist and are unique, when the error has a log-concave density.

In conclusion, we advocate the use of regression models over the chain ladder method, since they take into account both the error involved in the estimation of the parameters and the statistical error inherent in the prediction of future claims, the fit of the model can be tested statistically and confidence intervals for the reserve can be derived.

Keywords: Chain-ladder method; Weibull-extreme value regression; maximum likelihood; prediction; generalized loggamma; logistic; inverse gaussian; consistency.

[1] The author gratefully acknowledges the financial support of the CAFIR research fund and of the Natural Sciences and Engineering Research Council of Canada.

1 Introduction

1.1 IBNR claims

All insurance companies registered to do business in Canada are required by the regulatory authorities to set up reserves for claims which have been incurred but have not yet been reported as of their financial statement date, usually December 31. In determining the liabilities of the insurance company, the valuation actuary must also estimate the liabilities generated by claims incurred but not enough reserved (IBNER), (also called reported but not settled (RBNS)).

The distinction between these two parts of the loss reserve, the IBNR part and the IBNER part, is not always made in practice, especially when the data are aggregated. In this paper, by IBNR reserve, we will refer to both types of claims.

The primary purpose of those reserves is to ensure the protection of the policyholders: when the insurance company is notified of these claims, it will have the reserves, backed by sufficient assets, to pay those claims.

The delay in reporting the claim may depend on the type of claim (for example, asbestosis may take more than 10 years to manifest itself in a worker). The long delay observed in the settlement of certain claims is sometimes due to the fact that some of them are resisted by the insurance company, putting into motion a long judiciary process. In other cases, there will be a long delay before the ultimate cost of a claim can be determined exactly (in workers' compensation for example, the insurance company will have to wait for an annuity to terminate).

The 1987 Loss Development Study, undertaken by the Reinsurance Association of America, compares the development of losses for various lines of business. Automobile liability was the line where the claims got developed the fastest, while Workers' Compensation was slower to develop. General liability, excluding asbestos claims, had a development pattern similar to Workers' Compensation, but a little bit slower initially. Medical malpractice experienced the slowest development among those lines of business.

Due to this long reporting and settlement lag, it will be extremely important for the valuation actuary to develop adequate statistical models to project known losses to ultimate losses.

1.2 The chain ladder method and its deficiencies

By grouping the claims by accident year (year in which the accident giving rise to the claim occurred) and development year (number of years elapsed since this accident year), the data can be presented in a trapezoidal array.

In this paper, to illustrate the various models proposed, we will use the data in table 1 (taken from CIA Proceedings, Volume 20 no 1, p.183), which represents the liability claims in thousands of dollars incurred by a Canadian insurance company over the ten-year period 1978-1987. We will do the analysis with the incremental claims (in table 2), obtained by differencing successive cumulative amounts.

The problem of estimating IBNR claims consists in predicting, for each accident year,

Table 1: Claims Incurred

Accident year	Development year					
	1	2	3	4	5	6
1978	8489	9785	10709	11289	11535	11661
1979	12970	14766	16201	17060	17714	17979
1980	17522	20305	21774	22797	23220	23872
1981	21754	24338	25501	26284	27171	27526
1982	19208	21549	22769	23388	24229	24932
1983	19604	22073	23296	24543	25155	
1984	21922	24233	25374	26882		
1985	25038	28401	30545			
1986	32532	37006				
1987	39862					

the ultimate amount of claims incurred. The amount paid by the insurance company for those claims is then subtracted, leaving the reserve the insurer should hold for future payments. To calculate the reserve, all methods or models usually assume that the pattern of cumulative or incremental claims incurred or paid is stable across the development years, for each accident year. Since for the last accident year, only one amount will be available, the reserve will be highly sensitive to this amount. Moreover, because of growth experienced by the company, it will be bigger than any other amount in the data set, hence the importance of verifying that the development pattern of the claims has not changed over the years.

One of the earliest methods, and now most commonly used in the actuarial profession, is the chain ladder method. Assuming that for each accident year, the development pattern remains stable, development factors are calculated by dividing cumulative paid or incurred

Table 2: Incremental claims incurred

Accident year	Development year 1	2	3	4	5	6
1978	8489	1296	924	580	246	126
1979	12970	1796	1435	859	654	265
1980	17522	2783	1469	1023	423	652
1981	21754	2584	1163	783	887	355
1982	19208	2341	1220	619	841	703
1983	19604	2469	1223	1247	612	
1984	21922	2311	1141	1508		
1985	25038	3363	2144			
1986	32532	4474				
1987	39862					

claims after j periods since incurral by the cumulative amount after $j - 1$ periods. These factors can be weighted by the amount each year. The year-to-year development factors are then applied to the most recent amount for each accident year, i.e. the amounts on the right-most diagonal.

Using the weighted approach with the cumulative claims of table 1, we obtain the development factors of table 3. Projecting the claims incurred to ultimate amounts with those development factors, we obtain a reserve estimate of 23,919.

Table 3: Loss Development Factors

Year	Development factors
1-2	1.13079
2-3	1.06479
3-4	1.04545
4-5	1.02922
5-6	1.02023

Many variations have been presented for the basic chain ladder method just introduced; a linear trend or an exponential growth may be assumed to be present among the development factors. Instead of taking their weighted average, they would be extrapolated into the future. The chain ladder method can also be adjusted for inflation.

However, the chain ladder method suffers from the following deficiencies:

1- it implicitly assumes too many parameters (one for each column).

2- it does not give any idea of the variability of the reserve estimate, or a confidence interval for the reserve.

3- as will be shown in section 2, it is negatively biased, which could lead to serious underreserving, a threat to the insurer's solvency.

We will therefore develop a stochastic model, which involves only 5 parameters. With this model, we will be able to calculate an amount such that there is an 80% probability that the reserve will be sufficient to cover the liabilities generated by the current block of business.

1.3 The general model

In this paper, we will consider loglinear location-scale regression models of the form

$$Z_i = \ln Y_i = X_i\beta + \sigma\epsilon_i, \quad Y_i > 0$$

where Y_i is the ith element of vector Y (the data), of dimension n,

X is the regression matrix, whose first column contains 1's,

and whose ith row is the vector denoted by X_i

and (i,j) element is denoted X_{ij},

β is the vector of unknown parameters to be estimated,

of dimension p,

$X_i\beta$ is the location parameter for Z_i,

σ is the scale parameter,

and ϵ_i is a random error with known density $f(\epsilon)$.

The loglinear location-scale model has been used extensively in reliability theory and in survival analysis (see for example, Kalbfleisch and Prentice (1980), Lawless (1982), Cohen and Whitten (1988), Bain and Engelhardt (1991)). It is easily shown that the random variable Z_i will have density

$$\frac{1}{\sigma}f(\frac{z_i - X_i\beta}{\sigma}), \quad -\infty < z_i < \infty.$$

As in Zenwirth (1990), for the location parameter, we will use $\alpha + \beta \ln j + \gamma j + \iota(i + j - 2)$, where i is the accident year and j, the development year. Taylor (1986) cautions not to use cumulative claims amounts, but incremental claims in the analysis; otherwise, the estimates obtained would be biased, because of the non-independance of the cumulative amounts.

We will assume that $Y_i > 0$. To model negative values of Y_i, Cohen and Whitten (1988) use modified moment estimators and Cohen (1988), local maximum likelihood methods.

1.4 Outline of the paper

Section 2 considers the lognormal linear regression model and presents the results of a simulation study showing that the chain ladder estimate of the reserve is negatively biased. Other choices possible for the distribution of the random error are the extreme value distribution, leading to the Weibull-extreme value regression model (section 3), the generalized loggamma (section 4), the logistic (section 5), and the log inverse gaussian distribution (section 6). We derive certain theoretical properties of these distributions, such as their moment generating function and moments. We show how the actuary can establish a reserve with a certain level of confidence (for example 80%), with a simulation.

In section 7, we show that the MLE's of the regression and scale parameters exist and are unique when the error ϵ in the loglinear location-scale regression model has a log-concave density. Under misspecification of the error distribution in a linear location-scale model, the MLE's of the regression parameters are shown to be consistent (section 8), while we present a sufficient condition for the consistency of the MLE of the scale parameter, when the postulated model has lognormal errors. Finally, we present some remarks.

2 Lognormal linear regression model

When it is assumed that ϵ_i are independent and identically distributed $N(0,1)$ random variables, we obtain the lognormal linear regression model. Doray (1992) has studied

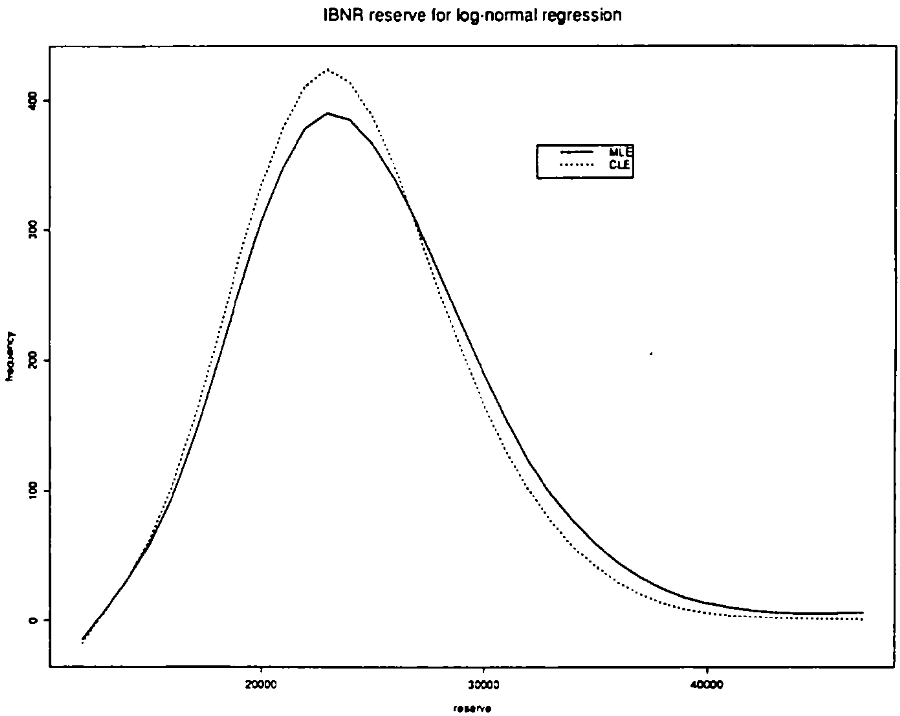
Table 4: Frequency distribution of the IBNR reserve under the normal error assumption

Amount	MLE	CLE	Amount	MLE	CLE
< 13000	0	0	30000-31000	165	152
13000-14000	4	2	31000-32000	150	126
14000-15000	12	11	32000-33000	103	80
15000-16000	33	30	33000-34000	96	68
16000-17000	62	72	34000-35000	76	47
17000-18000	126	131	35000-36000	50	40
18000-19000	191	199	36000-37000	36	26
19000-20000	253	301	37000-38000	28	16
20000-21000	323	376	38000-39000	20	5
21000-22000	372	391	39000-40000	14	2
22000-23000	449	441	40000-41000	13	10
23000-24000	449	498	41000-42000	8	2
24000-25000	393	443	42000-43000	7	3
25000-26000	366	436	43000-44000	7	0
26000-27000	342	375	44000-45000	2	2
27000-28000	334	274	45000-46000	2	1
28000-29000	285	231	46000-47000	6	0
29000-30000	214	207	≥ 47000	9	2

extensively this model, taking into account the estimation error on the parameters and the statistical prediction error in the model. He has derived various estimators for the IBNR reserve, among them the maximum likelihood estimator and the uniformly minimum variance unbiased estimator (UMVUE), as well as an expression for the variance of the latter estimator. The variance of the IBNR reserve is also calculated. The joint distribution of the amounts in each cell of the lower triangle is shown to follow a multivariate lognormal (MLN) distribution.

To compare the traditional chain ladder estimator of the reserve with the MLE, a simulation was performed, assuming the model $\ln Y_{ij} = \alpha_i + \beta_j + \epsilon_{ij}$ is the true model.

Figure 1


IBNR reserve for log-normal regression
MLE
CLE
frequency
0
100
200
300
400
20000
30000
40000
reserve

Five thousand sets of realizations of Y_{ij} in the trapezium were randomly generated, where each Y_{ij} is independent $LN(\hat{\alpha}_i + \hat{\beta}_j, \tilde{\sigma}^2)$, where the values of $\hat{\beta}$ and $\tilde{\sigma}^2$ are the MLE's of the parameters. For each set, we calculated the chain ladder estimate (CLE) and the MLE of the predicted value of IBNR claims using the multivariate lognormal distribution (see appendix 10.1 for the algorithm used for the simulation). The results of the simulation are summarized in table 4 and figure 1. We see from those results that the reserve has a distribution skewed to the right, which comes from the lognormal assumption. The reason why the chain ladder estimate, generally used by actuaries to determine insurance company reserves, underestimates the expected liability, is that it does not capture this long-tail behaviour, as is apparent from table 4.

The MLE of the reserve gives 25,262, while the CLE gives 23,919. The reserve for IBNR claims the insurance company will hold could be set at, for example, the 80-th percentile of the predicted distribution of IBNR claims, that is at 29,019 in our example. The actuary could then state, that in his or her opinion, there is an 80% probability that the reserve will be sufficient to meet the liabilities of the current block of business.

Asymptotically (i.e. as the upper trapezium of data gets larger), the various variables to be predicted will become independent, and from that perspective, we can consider an asymptotic confidence interval for the reserve, using the central limit theorem. The lower bound for the 80% asymptotic confidence interval of the reserve is 29,514, which can be compared with the amount of 29,019 obtained in the simulation.

A provision for adverse deviation could also be defined as equal to the 80-th percentile of the predicted distribution of IBNR claims minus the UMVUE of the reserve (24,403). This gives 4616 as the PAD for the claims of section 1.2.

3 Weibull-extreme value regression model

In this section, we examine the Weibull-extreme value regression model. Let us assume that ϵ follows a standard type I extreme value (or Gumbel) distribution with

probability density function (pdf)	$f(\epsilon) = \exp(\epsilon - e^{\epsilon})$, $\quad -\infty < \epsilon < \infty$,
cumulative distribution function (cdf)	$F(\epsilon) = 1 - \exp(-e^{\epsilon})$,
moment generating function (mgf)	$M_{\epsilon}(t) = \Gamma(1 + t)$, $\quad t > -1$,
mean	$E(\epsilon) = -\gamma = -0.5772156649015329\ldots$,
	where γ is Euler's constant
and variance	$Var(\epsilon) = \pi^2/6$.

The extreme value density is skewed to the left. The probability that a standard normal random variable take a value greater than 1.96 is 0.025, while the corresponding probability for the standard extreme value is only 0.0008256. Lawless (1982, p. 17-19) and Johnson and Kotz (1970) discuss the properties of the extreme value distribution.

Under this assumption for the density of ϵ, Y_i has the pdf

$$\frac{1}{\sigma e^{X_i\beta}}\left(\frac{y_i}{e^{X_i\beta}}\right)^{\frac{1}{\sigma}-1}\exp\left[-\left(\frac{y_i}{e^{X_i\beta}}\right)^{\frac{1}{\sigma}}\right], \quad y_i > 0,$$

which will be recognized as that of a Weibull random variable (Hogg and Klugman (1984)). Under this parametrization, the shape parameter is equal to $1/\sigma$ and the scale parameter to $e^{X_i\beta}$. The hazard rate will be increasing if $\sigma < 1$, decreasing if $\sigma > 1$ and constant if $\sigma = 1$, in which case the Weibull distribution reduces to the exponential distribution. The mean and variance of Y_i are:

$$\begin{aligned} E(Y_i) &= e^{X_i\beta}\Gamma(1+\sigma) \\ Var(Y_i) &= e^{2X_i\beta}[\Gamma(1+2\sigma) - \Gamma(1+\sigma)^2]. \end{aligned}$$

A proof of those results is contained in Lawless (1982).

The likelihood function based on the data $z_i = \ln y_i$, is

$$L(\beta, \sigma) = \prod_{i=1}^{n} \frac{1}{\sigma} \exp\left[\frac{z_i - X_i\beta}{\sigma} - \exp\left(\frac{z_i - X_i\beta}{\sigma}\right)\right],$$

and the log likelihood is

$$l(\beta, \sigma) = \sum_{i=1}^{n} \left[-\ln\sigma + \frac{z_i - X_i\beta}{\sigma} - \exp\left(\frac{z_i - X_i\beta}{\sigma}\right)\right].$$

Let us define $w_i = (z_i - X_i\beta)/\sigma$.

The first and second partial derivatives of l with respect to β_j and σ are

$$\begin{aligned} \frac{\partial l}{\partial \beta_j} &= -\frac{1}{\sigma}\sum_{i=1}^{n} X_{ij} + \frac{1}{\sigma}\sum_{i=1}^{n} X_{ij}e^{w_i}, \quad j = 1, \ldots, p. \\ \frac{\partial l}{\partial \sigma} &= -\frac{n}{\sigma} - \frac{1}{\sigma}\sum_{i=1}^{n} w_i + \frac{1}{\sigma}\sum_{i=1}^{n} w_i e^{w_i}. \\ \frac{\partial^2 l}{\partial \beta_j \partial \beta_k} &= -\frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ij}X_{ik}e^{w_i}, \quad j,k = 1, \ldots, p. \\ \frac{\partial^2 l}{\partial \sigma^2} &= \frac{n}{\sigma^2} + \frac{2}{\sigma^2}\sum_{i=1}^{n} w_i - \frac{2}{\sigma^2}\sum_{i=1}^{n} w_i e^{w_i} - \frac{1}{\sigma^2}\sum_{i=1}^{n} w_i^2 e^{w_i}. \\ \frac{\partial^2 l}{\partial \beta_j \partial \sigma} &= \frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ij} - \frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ij}e^{w_i} - \frac{1}{\sigma^2}\sum_{i=1}^{n} X_{ij}w_i e^{w_i}, \quad j = 1, \ldots, p. \end{aligned}$$

In appendix 10.2, we have listed some asymptotic properties of MLE's. The terms in the observed information matrix can be simplified by using the fact that the MLE's for

β_j and σ satisfy the equations $\frac{\partial l}{\partial \beta_j} = \frac{\partial l}{\partial \sigma} = 0$. The observed information matrix I_0 then becomes

$$\frac{1}{\hat\sigma^2}\begin{pmatrix} n & \sum \ln j & \sum j & \sum i+j-2 & n+\sum \hat w_i \\ \sum \ln j & \sum (\ln j)^2 e^{\hat w_i} & \sum j(\ln j)e^{\hat w_i} & \sum (i+j-2)(\ln j)e^{\hat w_i} & \sum (\ln j)\hat w_i e^{\hat w_i} \\ \sum j & \sum j(\ln j)e^{\hat w_i} & \sum j^2 e^{\hat w_i} & \sum j(i+j-2)e^{\hat w_i} & \sum j\hat w_i e^{\hat w_i} \\ \sum i+j-2 & \sum (i+j-2)(\ln j)e^{\hat w_i} & \sum j(i+j-2)e^{\hat w_i} & \sum (i+j-2)^2 e^{\hat w_i} & \sum (i+j-2)\hat w_i e^{\hat w_i} \\ n+\sum \hat w_i & \sum (\ln j)\hat w_i e^{\hat w_i} & \sum j\hat w_j e^{\hat w_i} & \sum (i+j-2)\hat w_i e^{\hat w_i} & n+\sum \hat w_i^2 e^{\hat w_i} \end{pmatrix}$$

where $\hat w_i = (z_i - X_i\hat\beta)/\hat\sigma$.

The asymptotic variance-covariance matrix of the parameters is equal to the inverse of I_0, and could be found using a symbolic computational language like MAPLE, or evaluated numerically. The expected information matrix can also easily be obtained (ref. Lawless (1982), p. 301-302).

Maximizing the log likelihood with the data of section 1.2 by using the Newton-Raphson algorithm or the SAS (1985) LIFEREG procedure, we find the MLE's, estimated standard errors and correlation matrix appearing in table 5. In section 7, we show that for certain location-scale models, the MLE's exist and are unique; this is true in particular for the Weibull-extreme value regression model.

All parameters are highly significant (at the 0.0001 level). It should also be noticed that the scale parameter estimator $\hat\sigma$ is not independent of the location parameter estimator, as is the case in normal regression. This complicates somewhat the estimation of the IBNR reserve.

Table 5: Weibull-extreme value regression

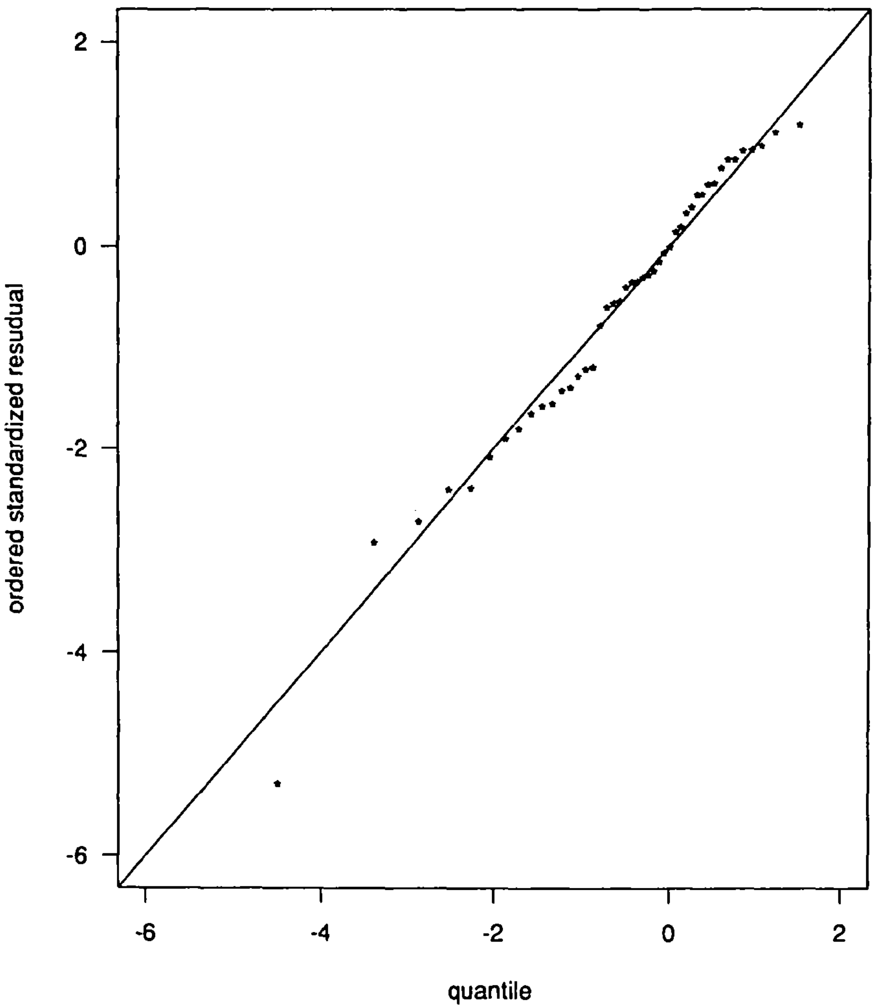
parameter	MLE	std. error	correlation matrix				
α	9.02897	0.11505	1	0.429	-0.515	-0.461	-0.017
β	-3.26637	0.25407	0.429	1	-0.972	0.214	0.0004
γ	0.40378	0.10372	-0.515	-0.972	1	-0.280	-0.006
ι	0.10811	0.01641	-0.461	0.214	-0.280	1	0.011
σ	0.02459	0.00642	-0.017	0.0004	-0.006	0.011	1

A Q-Q plot of the residuals appears in figure 2. It shows no evident departure from the extreme value distribution. It should be noted that the above standard errors and correlation matrix of the parameters are based on the joint asymptotic multivariate normal distribution of the MLE's. This approximation will be appropriate only when the number of cells in the trapezium of data is large enough (in our example, we have 45 cells).

How large is large enough? Bain and Engelhardt (1991) considered this problem for the Weibull distribution, but without covariates in the location parameter. They provide a table giving the bias of the MLE of the shape parameter of the Weibull distribution for different sample sizes. With a sample size of 40, the MLE overestimates the shape parameter by only 3.5%. If the sample size is only 10, care should be taken, since the bias is then around 15%. Those factors were obtained by a simulation study. We will not correct for the bias in our analysis, but we should remember that this might be a good idea for small sample sizes.

To test for $\sigma = 1$ (test of exponentiality of Y_i), we can use the asymptotic normality of the MLE's; unless the sample size is large, Lawless (1982) cautions that the normal approximation might not be very good. A likelihood ratio test can also be performed

Figure 2: Extreme value Q-Q plot of residuals



using the test statistic

$$\Lambda = -2 \log \frac{L(\tilde{\beta}, 1)}{L(\hat{\beta}, \hat{\sigma})},$$

where $\tilde{\beta}$ is the MLE of β under $H_0 : \sigma = 1$; the likelihood ratio statistic Λ has an asymptotic $\chi^2_{(1)}$ distribution. Performing a simple normal test leads us to reject the hypothesis $H_0 : \sigma = 1$. A Weibull distribution is therefore more appropriate for the data than an exponential distribution.

We now turn our attention to the problem of predicting the IBNR reserve. In a log-linear location-scale model, the total error in the log predicted amount Z_{kl} is composed of two parts: an estimation error on the parameters and a statistical prediction error. We saw earlier that in the Weibull-extreme value regression model, the estimators of the parameters have an asymptotic multivariate normal distribution, while the process error has an independent extreme value distribution.

Let Y_{kl} denote the random variable for the amount to be predicted in accident year k and development year l, and let us define $Z_{kl} = \ln Y_{kl}$. The random variable Z_{kl} being equal to $Z_{kl} = \hat{\alpha} + \hat{\beta} \ln k + \hat{\gamma} k + \hat{\iota}(k + l - 2) + \hat{\sigma}\epsilon$, we can appreciate the difficulty involved in trying to get its exact distribution. For this, we would need to find the distribution of the product of a normal and an extreme value random variable ($\hat{\sigma}$ and ϵ) and convolute this with a non-independent normal random variable. To get the distribution of Y_{kl}, the distribution of Z_{kl} is then exponentiated. It is highly doubtful that such a distribution would have a simple density. Instead of trying to accomplish this task, we will perform a simulation study to evaluate IBNR reserves. This will make it possible to find a confidence interval for the reserve.

Table 6: Frequency distribution of the IBNR reserve under the extreme value error assumption

Amount	Frequency
< 15000	0
15000-16000	1
16000-17000	12
17000-18000	54
18000-19000	144
19000-20000	357
20000-21000	664
21000-22000	904
22000-23000	982
23000-24000	791
24000-25000	605
25000-26000	285
26000-27000	142
27000-28000	46
28000-29000	8
29000-30000	4
30000-31000	1
> 31000	0

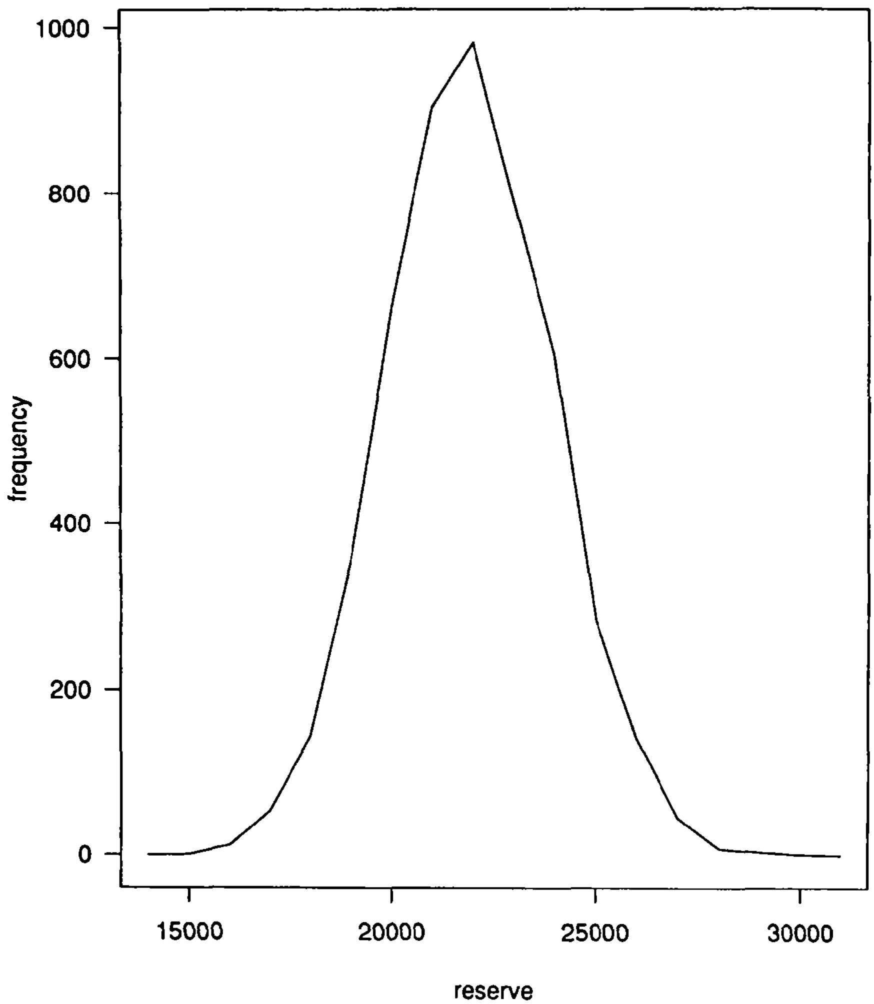
In appendix 10.1, we show how to generate a multivariate normal distribution, using the Choleski decomposition method. To be able to simulate the random variable Y_{kl}, we just need to show how to generate a standard extreme value random variable ϵ, with cdf

$$P[\epsilon \leq \epsilon_0] = 1 - \exp(-e^{\epsilon_0}), \quad -\infty < \epsilon_0 < \infty.$$

This cdf is easily inverted, yielding

$$\epsilon = \ln(-\ln(1 - U)), \quad 0 < U < 1,$$

Figure 3: IBNR reserve for Weibull-extreme value regression



where U is a uniform random variable on the interval $[0,1]$. Note that $1-U$ is also uniform on $[0,1]$, simplifying the algorithm.

Table 6 and figure 3 contain the results of a simulation of 5000 values for the IBNR reserve. The mean of the IBNR claims is 22,402 and the standard deviation of this estimate is 2011. The 80-th percentile for the simulated distribution of the IBNR reserve is 23,980.

Comparison of the extreme value and the normal distributions shows that the former has a heavier left tail and a lighter right tail than the latter. The estimation error on the regression parameters is of the same order in both models, while the stochastic error is smaller in the extreme value case.

4 Generalized loggamma regression model

The regression model used in this section will be the following

$$Z_i = \ln Y_i = X_i\beta + \sigma\epsilon_i,$$

where ϵ_i has a loggamma distribution with pdf

$$f(\epsilon;q) = \frac{|q|}{\Gamma(q^{-2})} q^{-2q^{-2}} \exp[q^{-2}(q\epsilon - e^{q\epsilon})], \quad -\infty < \epsilon < \infty,$$

and the shape parameter q can take any non-zero value (ref. Lawless (1982), p. 322-328). Under this parametrization, as q tends to 0, we obtain the normal distribution with pdf

$$f(\epsilon) = \frac{1}{\sqrt{2\pi}} \exp(-\epsilon^2/2), \quad -\infty < \epsilon < \infty.$$

The following special cases for the random variable Y_i can be obtained for certain values of the parameters q and σ: Weibull ($q = 1$), exponential ($q = \sigma = 1$), lognormal

($q = 0$) and reciprocal Weibull ($q = -1$). The density is negatively skewed for $q > 0$, with absolute skewness and kurtosis increasing as q increases; it is positively skewed for $q > 0$. A likelihood ratio test can be performed to test for the appropriateness of a certain member of the family.

Prentice (1974) and Farewell and Prentice (1977) have studied the properties of this generalized distribution. If we define the parameter $k = q^{-2}$, then it has moment generating function $\Gamma(k + t)$, $t > -k$, mean $\psi(k)$ and variance $\psi'(k)$, where $\psi(\cdot)$ and $\psi'(\cdot)$ are respectively the digamma and trigamma functions, the first and second derivatives of the gamma function. The series expansion for these two functions are:

$$\begin{aligned}\psi(n) &= -\gamma + \sum_{k=1}^{n-1}\frac{1}{k}, \qquad \text{for an integer } n \geq 2\\ \psi'(z) &= \sum_{k=0}^{\infty}(z + k)^{-2}, \qquad z \neq 0, -1, -2, \ldots\end{aligned}$$

The log likelihood function gives

$$l(\beta, \sigma, q) = \sum_{i=1}^{n} \ln f(w_i; q) - \ln \sigma,$$

where $w_i = (z_i - X_i\beta)/\sigma$ and

$$\ln f(w_i;q) = \ln \mid q \mid -2q^{-2} \ln q - \ln \Gamma(q^{-2}) + q^{-2}(qw_i - e^{qw_i}).$$

The first and second partial derivatives of l with respect to β and σ gives

$$\begin{aligned}\frac{\partial l}{\partial \beta_j} &= \sum_{i=1}^{n}\frac{X_{ij}}{q\sigma}[\exp(qw_i) - 1], \quad j = 1, \ldots, p.\\ \frac{\partial l}{\partial \sigma} &= \sum_{i=1}^{n}\{\frac{w_i}{q\sigma}[\exp(qw_i) - 1] - \frac{1}{\sigma}\}\end{aligned}$$

$$\frac{\partial^2 l}{\partial\beta_j\partial\beta_k} = \sum_{i=1}^{n} X_{ij}X_{ik}(\frac{-1}{\sigma^2})\exp(qw_i)$$

$$\frac{\partial^2 l}{\partial\sigma^2} = \sum_{i=1}^{n}\frac{1}{\sigma^2}[1 - w_i^2\exp(qw_i)] - \frac{2w_i}{q\sigma^2}[\exp(qw_i) - 1]$$

$$\frac{\partial^2 l}{\partial\beta_j\partial\sigma} = \sum_{i=1}^{n} X_{ij}(\frac{-1}{\sigma^2})[w_i\exp(qw_i) + \frac{1}{q}(\exp(qw_i) - 1].$$

Again, using the fact that the MLE's satisfy $\frac{\partial l}{\partial\sigma} = \frac{\partial l}{\partial\beta_j} = 0$, we can simplify the last two partial derivatives and obtain

$$\frac{\partial^2 l}{\partial\sigma^2}\bigg|_{(\hat\beta,\hat\sigma)} = -\frac{1}{\hat\sigma^2}[n + \sum\hat{w}_i^2\exp(q\hat{w}_i)]$$

and

$$\frac{\partial^2 l}{\partial\beta_j\partial\sigma}\bigg|_{(\hat\beta,\hat\sigma)} = -\frac{1}{\hat\sigma^2}\sum X_{ij}\hat{w}_i\exp(q\hat{w}_i).$$

To find the MLE's of the parameters, we can use the approach suggested by Farewell and Prentice (1977). The parameter q is fixed at a value q_0 and the profile log likelihood is maximized using the Newton-Raphson algorithm over the regression parameters β and the scale parameter σ. This gives the estimates $(\hat\beta(q_0), \hat\sigma(q_0))$. This procedure of maximizing the profile log likelihood is repeated for many values of q_0, until an overall maximum of the log likelihood over q_0 is attained. This value gives the MLE $\hat{q}$.

The SAS package fits generalized loggamma regression models. Using the SAS LIFEREG procedure for complete data, we find the results appearing in Table 7.

The default convergence criterion used by SAS is that a maximum is assumed to have occurred if the relative change in the parameters is less than 0.001. However, as can be seen from table 8, the likelihood keeps increasing beyond this value of $\hat{q}$. The convergence criterion we used is that the score statistic with respect to each parameter should be of

Table 7: Generalized loggamma regression (SAS program)

parameter	MLE	std. error	correlation matrix					
α	9.32243	0.02789	1	0.469	-0.521	-0.160	-0.497	0.497
β	-3.12566	0.07028	0.469	1	-0.991	0.645	-0.150	0.150
γ	0.35670	0.02969	-0.521	-0.991	1	-0.626	0.124	-0.123
ι	0.10058	0.00357	-0.160	0.645	-0.626	1	-0.087	0.086
σ	0.04035	0.03187	-0.497	-0.150	0.124	-0.087	1	-0.981
q	9.99342	7.63421	0.497	0.150	-0.123	0.086	-0.981	1

the order of 10^{-6}. Past the value of $q_0 = 31.623$ (corresponding to $k = q_0^{-2} = 0.001$), some elements of the information matrix become so large that it cannot be inverted and the standard Newton-Raphson algorithm fails.

Table 8: Generalized loggamma regression for various values of q_0

q_0	$\hat{\alpha}(q_0)$	$\hat{\beta}(q_0)$	$\hat{\gamma}(q_0)$	$\hat{\iota}(q_0)$	$\hat{\sigma}(q_0)$	$l(q_0)$
0	8.97986	-3.14641	0.30881	0.12298	0.31380	-11.70862
1	9.02897	-3.26637	0.40378	0.10811	0.24588	-8.66845
2	9.15105	-3.19165	0.38375	0.10369	0.17552	-7.82173
3	9.24020	-3.13178	0.35787	0.10264	0.12742	-7.23110
4	9.27974	-3.12132	0.35336	0.10188	0.09803	-6.64823
6	9.30818	-3.12572	0.35608	0.10088	0.06590	-5.68347
8	9.31835	-3.12611	0.35666	0.10061	0.04950	-5.03186
10	9.32308	-3.12419	0.35609	0.10063	0.03964	-4.62194
20	9.33019	-3.11565	0.35296	0.10088	0.01986	-3.87515
30	9.33340	-3.11023	0.35061	0.10092	0.01324	-3.68571

A few remarks should be made here.

1- the likelihood is so flat that it makes the standard error of $\hat{q}$ (7.63421), calculated assuming asymptotic normality, totally unreliable. Bain and Engelhardt (1991, p. 393) report that the asymptotic normal distribution for $\hat{k}$ will not be very accurate unless the sample size is greater than 200 or 400. Farewell and Prentice (1977) note that the skewness in the $\hat{q}$ distribution is related to an asymptotic variance that *increases rapidly as* $|q|$ *increases. To get a confidence interval for* $\hat{q}$, *a likelihood ratio* test would be preferable. This interval for $\hat{q}$ would include all the values q_0 satisfying

$$-2[\ln l(\hat{q}, \hat{\beta}, \hat{\sigma}) - \ln l(q_0, \hat{\beta}(q_0), \hat{\sigma}(q_0)] \leq 3.841.$$

2- the correlation between $\hat{\sigma}$ and $\hat{q}$ almost equal to -1 should be noted. From table 8, we can see that as q_0 increases, $\hat{\sigma}(q_0)$ decreases. Cox and Hinkley (1968) have shown that in the general regression model $Z = \alpha + X\beta + \sigma\epsilon(q)$, $(\hat{\alpha}, \hat{\sigma}, \hat{q})$ are asymptotically independent of $\hat{\beta}$, if the columns of X add to zero.

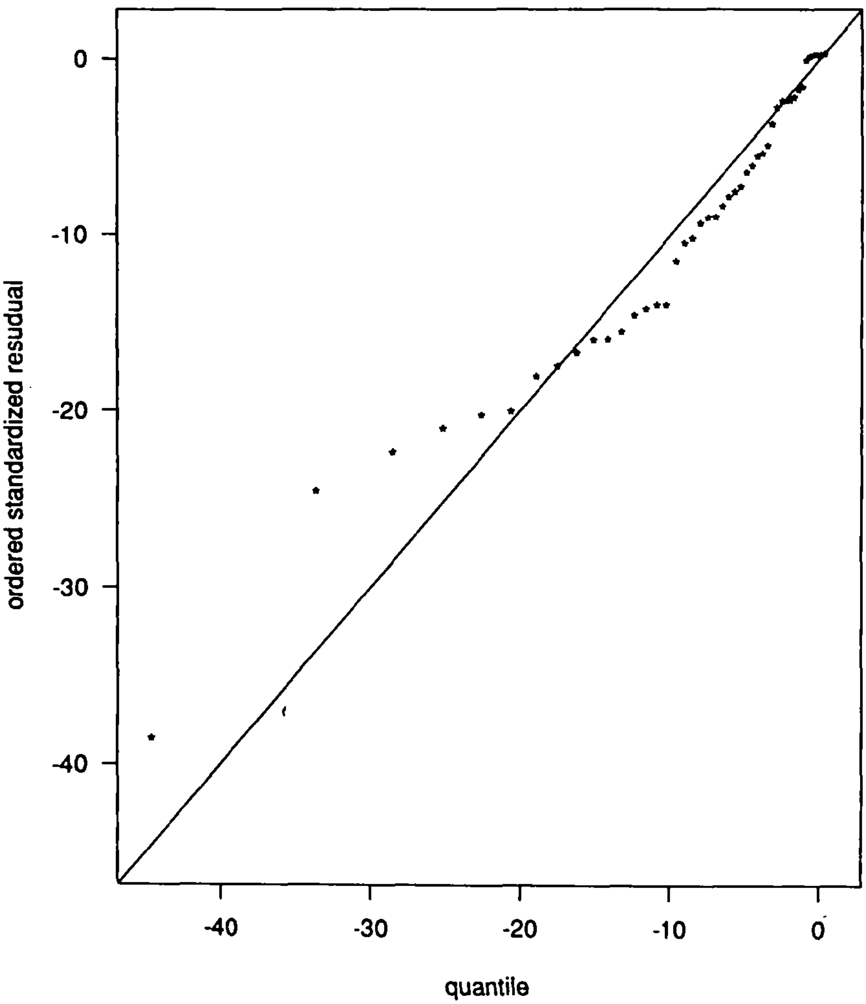
3- The regression parameters $(\alpha, \beta, \gamma, \iota)$ for any fixed value of q_0 are very close to those obtained in the normal and extreme value regression, and so is their standard error and their correlation matrix.

It should be remembered however that, although the MLE $\hat{q}$ cannot be found accurately, we know that it exists and is unique, because of the log-concavity of the loggamma distribution (see section 7).

If the exact value of $\hat{q}$, *was available, this would make the estimation of* E(*IBNR* claims) much more complicated than in the normal or extreme value cases, because of the non-independence of $\hat{q}$ with $\hat{\beta}$ and $\hat{\sigma}$. In this model, Y_{kl} is equal to

$$Y_{kl} = e^{\hat{\alpha} + \hat{\beta} \ln k + \hat{\gamma} k + \hat{\iota}(k+l-2) + \hat{\sigma}\epsilon(\hat{q})},$$

Figure 4: Loggamma (q=10) Q-Q plot of residuals



and we can see that the estimation error on the parameters is not independent of the process error $\epsilon(\hat{q})$, since $\hat{\beta}$, $\hat{\sigma}$ are estimated using the same set of past data which is used in estimating $\hat{q}$.

To assess the adequacy of the loggamma regression model, we fitted that model with a fixed q value, $q = 10$. Figure 4 presents the corresponding Q-Q plot. Since the left tail of the distribution is too short, we will not simulate the IBNR reserve; however, Devroye (1986) presents many algorithms to generate gamma random variables.

5 Logistic regression model

The logistic linear model is

$$Z_i = \ln Y_i = X_i\beta + \sigma\epsilon_i,$$

where ϵ has a standard logistic distribution with (see Lawless (1982), p. 46)

pdf	$f(\epsilon) = \frac{e^{\epsilon}}{(1+e^{\epsilon})^2}, \quad -\infty < \epsilon < \infty,$
cdf	$F(\epsilon) = 1 - (1 + e^{\epsilon})^{-1},$
mgf	$\Gamma(1+t)\Gamma(1-t), \mid t \mid < 1,$
mean	$E(\epsilon) = 0,$
variance	$Var(\epsilon) = \pi^2/3.$

The density of the logistic distribution somewhat looks like the standard normal density. The symmetry of the pdf around $\epsilon = 0$ implies that there is probability 1/2 that the amount Y_i be understated or overstated. The probability that a standard logistic random variable exceeds 1.96 is 0.12347. The logistic distribution has thick tails, which behave like that of the exponential distribution. The loglogistic is a special case of the Burr distribution, with the parameter α equal to 1 (ref. Panjer and Willmot (1992), p. 120).

The random variable Z_i has density

$$f_{Z_i}(z_i) = \frac{1}{\sigma}\frac{\exp[\frac{z_i - X_i\beta}{\sigma}]}{[1 + \exp(\frac{z_i - X_i\beta}{\sigma})]^2}, \quad -\infty < z_i < \infty,$$

and Y_i has the loglogistic density

$$\frac{1}{\sigma e^{X_i\beta}}\left(\frac{y_i}{e^{X_i\beta}}\right)^{\frac{1}{\sigma}-1}\left[1 + \left(\frac{y_i}{e^{X_i\beta}}\right)^{\frac{1}{\sigma}}\right]^{-2}, \quad y_i > 0, \qquad (5.1)$$

where again $e^{X_i\beta}$ is the scale parameter and $1/\sigma$ the shape parameter. In proposition 5.1, we derive the moments of order k of a loglogistic random variable with density 5.1 and show that its moment generating function does not exist.

Proposition 5.1: If Y has density

$$f_Y(y) = \frac{\delta^{1/\sigma}}{\sigma}\frac{y^{1/\sigma - 1}}{[1 + \delta^{1/\sigma}y^{1/\sigma}]^2}, \quad y > 0,$$

then

$$E(Y^k) = \delta^{\frac{1}{\sigma}-(k+1)}[1 - \sigma(k+1)]\pi \ \mathrm{cosec}[\pi\sigma(k+1)],$$

for all k such that $\frac{2}{\sigma} - 1 < k < \frac{4}{\sigma} - 1$, and the moment generating function of Y does not exist.

Proof: $E(Y^k) = \int_0^\infty y^k \frac{\delta^{1/\sigma}}{\sigma}\frac{y^{1/\sigma-1}}{[1+\delta^{1/\sigma}y^{1/\sigma}]^2}dy$.

By letting $y^{1/\sigma} = v$, we obtain

$$E(Y^k) = \delta^{1/\sigma}\int_0^\infty \frac{v^{\sigma(k+1)-1}}{[1 + \delta^{1/\sigma}v]^2}dv.$$

Using the formula

$$\int_0^\infty \frac{x^{\mu-1}}{(1+\beta x)^2}dx = \frac{1-\mu}{\beta^\mu}\pi \text{ cosec } \mu\pi,$$

the result is easily obtained. The integral will have a finite value iff

$$-1 < (k+1)\sigma - 3 < 1$$

or

$$\frac{2}{\sigma} - 1 < k < \frac{4}{\sigma} - 1.$$

The moments of all positive orders do not exist; therefore, the moment generating function of Y does not exist. □

The likelihood function is

$$L(\beta,\sigma) = \prod_{i=1}^{n} \frac{1}{\sigma}\frac{\exp(w_i)}{[1+\exp(w_i)]^2}, \quad -\infty < w_i < \infty,$$

where $w_i = \frac{z_i - X_i\beta}{\sigma}$, from which we get the log likelihood

$$l(\beta,\sigma) = \sum_{i=1}^{n}[w_i - 2\ln(1+e^{w_i}) - \ln\sigma].$$

For first and second order partial derivatives with respect to the parameters, see Kalbfleisch and Prentice (1980; p. 54-57). The SAS procedure LIFEREG was used to fit a logistic regression model to the data of section 1.3. The MLE's of the parameters, their estimated standard error and the estimated correlation matrix appear in table 3.5.

A Q-Q plot of the residuals in figure 5 shows that the logistic distribution does not provide a very good fit for the right tail. We will therefore not attempt to predict the IBNR reserve, but just indicate how it could easily be done by simulation, if it was appropriate to do so.

Figure 5: Logistic Q-Q plot of residuals

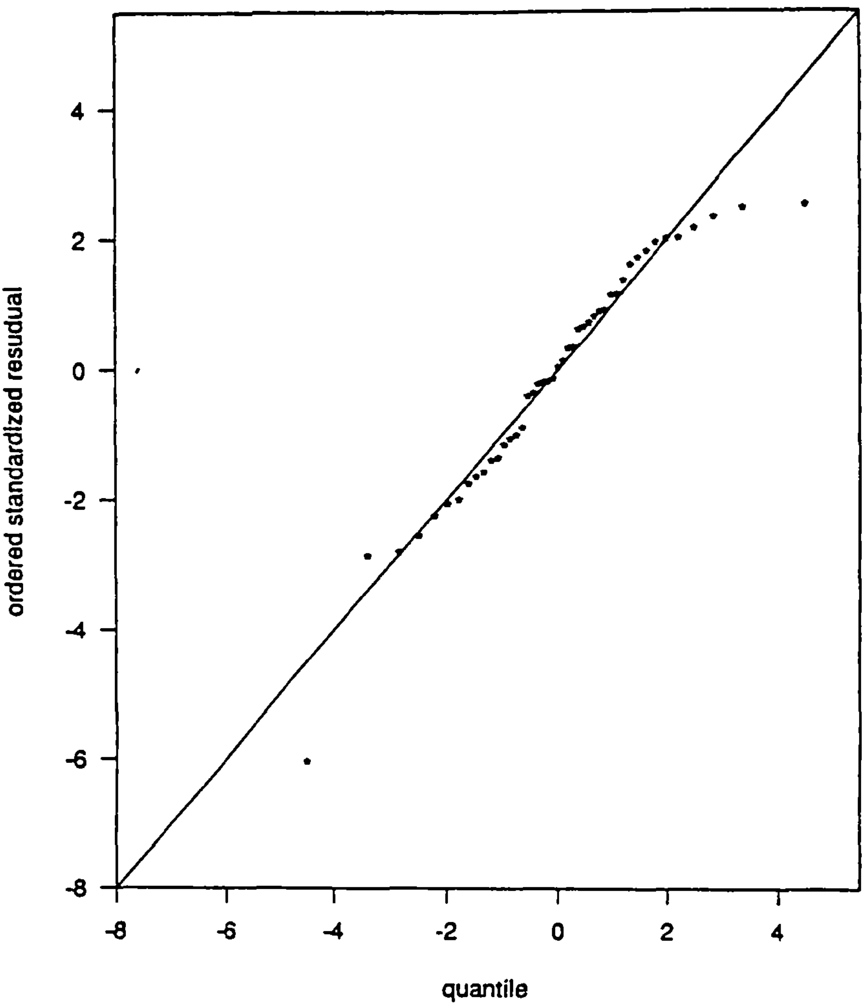


Table 9: Logistic regression

parameter	MLE	std. error	correlation matrix				
α	8.94023	0.13799	1	0.437	-0.516	-0.540	0.039
β	-3.31681	0.30143	0.437	1	-0.964	0.078	0.072
γ	0.38904	0.12058	-0.516	-0.964	1	-0.169	-0.083
ι	0.11789	0.02004	-0.540	0.078	-0.169	1	0.025
σ	0.17957	0.02203	0.039	0.072	-0.083	0.025	1

The loglogistic model for Y_{kl} is $Y_{kl} = e^{\hat{\alpha}+\hat{\beta}\ln k+\hat{\gamma}k+\hat{\iota}(k+l-2)+\hat{\sigma}\epsilon}$. The joint asymptotic distribution for $(\hat{\beta}, \hat{\sigma})$ is multivariate normal with parameter estimates given in table 9 and can be easily simulated (see Appendix 10.1). Inverting the cdf of the logistic random variable ϵ yields

$$\epsilon = \ln(\frac{1-U}{U}), \quad \text{where } U \text{ is uniform } [0,1].$$

The value of is then exponentiated to give Y_{kl}.

6 Log Inverse Gaussian regression model

The inverse gaussian regression model for Y_i is $Y_i = e^{X_i\beta+\epsilon_i}$, where the multiplicative error e^{ϵ} is assumed to have a standard inverse gaussian (IG), or Wald distribution, with density

$$f_V(v) = (2\pi\lambda v^3)^{-1/2}\exp\{-\frac{(v-1)^2}{2\lambda v}\}, \quad v > 0, \quad \lambda > 0.$$

This long-tail positively skewed distribution with exponential tails has a shape similar to that of the lognormal distribution (ref. Cohen and Whitten (1988), p. 77) and is located between the gamma and lognormal in Pearson's system of distributions, which shows possible regions of variation of the skewness and kurtosis (Jorgensen (1982), p. 19). To learn more about the inverse gaussian distribution, see Chhikara and Folks (1989) and Jorgensen (1982). Here are some of its important properties. The mean equals 1 and the variance λ. It is unimodal and a member of the exponential family. If V is $IG(1, \lambda)$, and $a > 0$ is a constant, aV is $IG(a, a\lambda)$. The sum of n independent $IG(1, \lambda)$ is $IG(n, \lambda)$.

Taking the log of Y_i, we obtain the loglinear model

$$Z_i = \ln Y_i = X_i\beta + \epsilon_i,$$

where ϵ has a log inverse gaussian (LIG) distribution. The pdf of ϵ is now derived.

Let $\epsilon = \ln V$, where V is $IG(1, \lambda)$. Then $V = e^\epsilon$ and $dV/d\epsilon = e^\epsilon$. It follows that

$$\begin{aligned} f(\epsilon) &= e^\epsilon(2\pi\lambda e^{3\epsilon})^{-1/2}\exp[-\frac{(e^\epsilon - 1)^2}{2\lambda e^\epsilon}] \\ &= (2\pi\lambda e^\epsilon)^{-1/2}\exp[-\frac{(e^\epsilon - 2 + e^{-\epsilon})}{2\lambda}] \\ &= (2\pi\lambda)^{-1/2}e^{-\epsilon/2}e^{1/\lambda}\exp[-\frac{1}{\lambda}\cosh\epsilon], \end{aligned} \tag{6.1}$$

where $\cosh\epsilon = (e^\epsilon + e^{-\epsilon})/2$.

In the next two propositions, we derive the moment generating function and the mean of the LIG distribution.

Proposition 6.1: The mgf of the LIG distribution with pdf (6.1) is

$$M_\epsilon(t) = (2\pi\lambda)^{-1/2}e^{1/\lambda}2K_{1/2-t}(1/\lambda).$$

Proof: Let the constant $C = (2\pi\lambda)^{-1/2}e^{1/\lambda}$. Then

$$\begin{aligned} M_\epsilon(t) &= E(e^{t\epsilon}) = \int_{-\infty}^{\infty} e^{t\epsilon}f(\epsilon)d\epsilon \\ &= C\int_{-\infty}^{\infty} e^{\epsilon(t-1/2)}\exp[-\frac{1}{\lambda}\cosh\epsilon]d. \end{aligned}$$

Using the formula

$$\int_{-\infty}^{\infty} \exp[-\alpha x - \frac{1}{\lambda}\cosh x]dx = 2K_\alpha(1/\lambda),$$

on page 309 of Gradshteyn and Ryzhik (1980), we get

$$M_\epsilon(t) = (2\pi\lambda)^{-1/2}e^{1/\lambda}2K_{1/2-t}(1/\lambda),$$

for $t\epsilon[-\infty, 1/2]$, where $K_\alpha(\cdot)$ denotes the Bessel function of the third kind of order α. □

Proposition 6.2

$$E(\epsilon) = e^{2/\lambda}\{-\gamma - \ln(2/\lambda) - \sum_{n=1}^{\infty}\frac{(-1)^n(2/\lambda)^n}{n \cdot n!}\}$$

Proof: We know that $E(\epsilon) = M_\epsilon'(t)\mid_{t=0}$.

The reader will appreciate the difficulty involved in taking the derivative of $M_\epsilon(t)$ with respect to t, since we need to differentiate with respect to the order of the Bessel function. From Abramowitz and Stegun (1972), p. 445, we get

$$\frac{\partial}{\partial\alpha}K_\alpha(x)\mid_{\alpha=1/2} = -\sqrt{\frac{\pi}{2x}}E_i(-2x)e^x,$$

where $-E_i(-x) = E_1(x) = \int_x^\infty \frac{e^{-t}}{t}dt$. So

$$\begin{aligned} E(\epsilon) &= (2\pi\lambda)^{-1/2}e^{1/\lambda}\cdot 2\sqrt{\pi\lambda/2}E_1(2/\lambda)e^{1/\lambda} \\ &= e^{2/\lambda}E_1(2/\lambda), \end{aligned}$$

where the series expansion for $E_1(x)$ is

$$E_1(x) = -\gamma - \ln x - \sum_{n=1}^{\infty}\frac{(-1)^n x^n}{n \cdot n!}$$

□

Let us now consider the estimation of the parameters λ and β. Y_i has an inverse gaussian distribution with parameters $(e^{X_i\beta}, \lambda e^{X_i\beta})$. The likelihood function is

$$L(\beta,\lambda) = \prod_{i=1}^{n} e^{X_i\beta}(2\pi\lambda e^{X_i\beta} y_i^3)^{-1/2} \cdot \exp\left\{-\frac{(y_i - e^{X_i\beta})^2}{2\lambda e^{X_i\beta} y_i}\right\},$$

and the log likelihood is

$$l(\beta,\lambda) = \sum_{i=1}^{n} X_i\beta - \frac{1}{2}\ln\lambda - X_i\beta/2 - \frac{3}{2}\ln y_i - \frac{(y_i - e^{X_i\beta})^2}{2\lambda e^{X_i\beta} y_i}.$$

The partial derivatives are

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n} \frac{-1}{2\lambda} + \text{[illegible]} \frac{\text{[illegible]}^{X_i\beta})^2}{\text{[illegible]}^{X_i\beta} y_i},$$

so that $\hat{\lambda} = \sum \frac{(y_i - e^{X_i\beta})^2}{n e^{X_i\beta} y_i}$.

$$\begin{aligned}
\frac{\partial l}{\partial \beta_j} &= \sum_{i=1}^{n} \frac{X_{ij}}{2\lambda}[\lambda + y_i e^{-X_i\beta} - e^{X_i\beta}] \\
\frac{\partial^2 l}{\partial \lambda^2} &= \sum_{i=1}^{n} \frac{1}{2\lambda^2} - \frac{(y_i - e^{X_i\beta})^2}{\lambda^3 e^{X_i\beta} y_i} \\
\frac{\partial^2 l}{\partial \lambda \partial \beta_j} &= \sum_{i=1}^{n} \frac{-X_{ij}}{2\lambda^2}[y_i e^{-X_i\beta} - e^{X_i\beta}/y_i] \\
\frac{\partial^2 l}{\partial \beta_j \partial \beta_k} &= \sum_{i=1}^{n} \frac{X_{ij}X_{ik}}{2\lambda}[-y_i e^{-X_i\beta} - e^{X_i\beta}/y_i]
\end{aligned}$$

To find the MLE's of β and λ, one could use the Newton-Raphson algorithm. The log-concavity of the LIG distribution will guarantee the existence of unique MLE's (see section 7).

The quantiles of this distribution could be obtained from the IG distribution, since

$$P[\epsilon \le \epsilon_0] = P[e^{\epsilon} \le e^{\epsilon_0}] = P[Y \le e^{\epsilon_0}],$$

where $Y \sim IG$. Therefore the q quantile of the LIG distribution is equal to the log of the q quantile of the IG distribution. Those can be calculated or obtained from a table, e.g. Koziol (1989). If an inverse gaussian regression model was found to be appropriate, to simulate $Y_{kl} = e^{\hat{\alpha}+\hat{\beta}\ln k+\hat{\gamma}k+\hat{\iota}(k+l-2)+\epsilon}$, we would need to simulate e^{ϵ}, which is $IG(1,\lambda)$. Michael, Schucany and Haas (1976) developed an algorithm to simulate such a distribution.

7 Existence and uniqueness of MLE's

In this section, we show that all the distributions used in this chapter for the error ϵ are log-concave. A consequence of this fact is that the MLE's will exist and be unique, although they need not be finite (ref. Burridge (1981)). When convergence is achieved in the Newton-Raphson algorithm, this implies that we found a global maximum, not just a local maximum.

Let us consider the loglinear location-scale model

$$Z_i = \ln Y_i = X_i\beta + \sigma\epsilon_i.$$

If we reparametrize to $\phi = 1/\sigma$, the log-likelihood of the data becomes

$$l(\sigma,\beta) = n\ln\phi + \sum_{i=1}^{n}\ln f(w_i)$$

where $w_i = (z_i - X_i\beta)\phi$ and $f(\cdot)$ is the density function of the error ϵ_i. Since w_i is a linear function of each of the parameters β and ϕ and is therefore concave, and the function ln is concave, l will be concave provided $\ln f(\cdot)$ is concave (ref. Burride (1981)). We have therefore shown the remarkable property that, in a loglinear location-scale regression model, the existence of the MLE's does not depend on the data but only on the log-concavity of the density of the error ϵ. We now show this is indeed the case for the five distributions used so far.

1- If $\epsilon \sim N(0,1)$, $f(\epsilon) = \frac{1}{\sqrt{2\pi}}\exp(-\epsilon^2/2)$, and $\ln f(\epsilon) = K - \epsilon^2/2$; so $\frac{\partial^2}{\partial\epsilon^2}\ln f(\epsilon) = -1 < 0\ \forall\epsilon$.

2- If $\epsilon \sim$ extreme value, $f(\epsilon) = \exp(\epsilon - e^\epsilon)$, and $\ln f(\epsilon) = \epsilon - e^\epsilon$; so $\frac{\partial^2}{\partial\epsilon^2}\ln f(\epsilon) = -e^\epsilon < 0\ \forall\epsilon$.

3- If $\epsilon \sim$ generalized loggamma,

$$f(\epsilon;q) = \frac{|\,q\,|}{\Gamma(q^{-2})}q^{-2q^{-2}}\exp[q^{-2}(\epsilon q - e^{q\epsilon})],$$

and $\ln f(\epsilon;q) = K + q^{-2}(\epsilon q - e^{q\epsilon})$; then $\frac{\partial^2}{\partial\epsilon^2}\ln f(\epsilon;q) = -e^{q\epsilon} < 0,\ \ \forall\epsilon$.

4- If $\epsilon \sim$ logistic, $f(\epsilon) = \frac{e^\epsilon}{(1+e^\epsilon)^2}$; then $\ln f(\epsilon) = \epsilon - 2\ln(1+e^\epsilon)$ and $\frac{\partial^2}{\partial\epsilon^2}\ln f(\epsilon) = \frac{-2e^{-\epsilon}}{(1+e^{-\epsilon})^2} < 0\ \forall\epsilon$.

5- If $\epsilon \sim LIG$, $f(\epsilon) = (2\pi\beta e^\epsilon)^{-\frac{1}{2}}\exp[\frac{-(e^{\epsilon/2}-e^{-\epsilon/2})^2}{2\beta}]$; so $\ln f(\epsilon) = K - \frac{\epsilon}{2} - \frac{(e^{\epsilon/2}-e^{-\epsilon/2})^2}{2\beta}$,

$$\frac{\partial \ln f(\epsilon)}{2\epsilon} = -\frac{1}{2} - \frac{e^\epsilon - e^{-\epsilon}}{2\beta}$$

and $\frac{\partial^2}{\partial\epsilon^2}\ln f(\epsilon) = -(\frac{e^\epsilon+e^{-\epsilon}}{2\beta}) < 0\ \forall\epsilon$.

An example of a distribution for ϵ which does not have the property of log-concavity for all ϵ is the Student's t distribution with n degrees of freedom, and density

$$f(\epsilon) = \frac{(1+\epsilon^2/2n)^{-(n+1)/2}}{\sqrt{n}\beta(1/2,n/2)}.$$

Then $\ln f(\epsilon) = K - \frac{1}{2}(n+1)\ln(1+\epsilon^2/n)$,

$$\frac{\partial}{\partial\epsilon}\ln f(\epsilon) = -(n+1)\epsilon/(\epsilon^2+n),$$

and $\frac{\partial}{\partial\epsilon^2}\ln f(\epsilon) = -(n+1)\frac{n-\epsilon^2}{(n+\epsilon^2)^2}$, which is positive for $\epsilon > \sqrt{n}$ or $\epsilon < -\sqrt{n}$.

8 Consistency of the parameters under error misspecificati(

Gould and Lawless (1988) investigated the consistency of the maximum likelihood estimators of the regression parameters under misspecification of the error distribution in a linear location-scale model.

The postulated model is

$$Z = \alpha + X\beta + \sigma\epsilon, \quad -\infty < \epsilon < \infty, \tag{8.1}$$

where σ is a scale parameter and ϵ has a specified distribution with density $f(\epsilon)$. They assume that the true unknown model is given by

$$Z = \mu_0 + X\mu + \tau w, \quad -\infty < w < \infty, \tag{8.2}$$

where w has density $g(w)$. The location-scale structure of the postulated model has the correct form; only the error distribution is misspecified.

If the following three assumptions are satisfied,

1- the covariates are centered;

2- all the expectations below exist and

3- $n^{-1}(X'X)$ is bounded as $n \to \infty$,

White (1982) proves that the MLE's of (α, β, σ) converge in probability to a unique limit $(\alpha^*, \beta^*, \sigma^*)$. Gould and Lawless (1988) then show that $\hat{\beta} = \mu^*$ and $\hat{\beta}$ is therefore a consistent estimator of μ. In addition, for $\hat{\alpha}$ and $\hat{\sigma}$ to be consistent estimators of μ_0 and τ, they must satisfy the two equations

$$E_T(\frac{\partial}{\partial W} \log W) = 0$$

and

$$E_T(W \cdot \frac{\partial}{\partial W} \log(W) + 1) = 0 \tag{8.3}$$

where $W = (\tau w + \mu_0 - \alpha^*)/\sigma^*$ and E_T indicates that the expectation is taken with respect to the true error distribution $g(w)$.

Gould and Lawless (1988) also analyze the asymptotic efficiency of the MLE based on the correct model. We will derive conditions that $g(w)$ must satisfy in order for $\hat{\alpha}$ and $\hat{\sigma}$ to be consistent estimators of μ_0 and τ, when the error ϵ in the postulated model (8.1) has a normal $N(0,1)$ distribution.

Lemma 8.1: Under the assumption of standard normal errors in model (8.1), a sufficient condition for $\hat{\alpha}$ and $\hat{\sigma}$ to be consistent estimators of μ_0 and τ is that $E(w) = 0$ and $Var(w) = 1$.

Proof: If $f(\epsilon) = \frac{1}{\sqrt{2\pi}} e^{-\epsilon^2/2}$, then $\frac{\partial}{\partial \epsilon} \log f(\epsilon) = -\epsilon$, and the equations (8.3) become $E_T(W) = 0$ and $E_T(W^2) = 1$.

Since $W = (\tau w + \mu_0 - \alpha^*)/\sigma^*$, the condition $E_T(W) = 0$ implies that $\mu_0 = \alpha^*$ i.e. $\hat{\alpha}$ is a consistent estimator of μ_0. If $E_T(W) = 0$, then $E_T(W^2) = Var_T(W) = (\tau/\sigma^*)^2 Var(w) =$

1. The condition $Var(w) = 1$ will imply that $\tau = \sigma^*$, i.e. that $\hat{\sigma}$ is a consistent estimator of τ. □

The consistency of $\hat{\alpha}$ and $\hat{\sigma}$ therefore depends only on the first two moments of the distribution of w, when the postulated model is lognormal linear.

We must point out here that one of the assumptions for the above development to be valid is that $n^{-1}(X'X)$ be bounded as $n \to \infty$. This condition is not verified in the model

$$Y_{ij} = \alpha + \beta \ln j + \gamma j + \iota(i + j - 2) + \epsilon_{ij}.$$

The covariate i would need to be removed from the model, for example by normalizing the amounts Y_{ij}, in order for $n^{-1}(X'X)$ to be bounded as $n \to \infty$.

9 Conclusion

In this paper, we have presented an anthology of models differing between them only in the distribution assumed for the error ϵ. To discriminate between the normal, extreme value, logistic and loggamma distribution for ϵ, we can assume that ϵ belongs to the generalized $\log F$ distribution (Prentice (1974)), with pdf

$$f(\epsilon) = (m_1/m_2)^{m_1} e^{wm_1}[1 + m_1 e^{w}/m_2]^{-(m_1/m_2)}.$$

After finding the MLE's $(\hat{m}_1, \hat{m}_2)$, we can perform a likelihood ratio test for

$(m_1, m_2) = (1, 1)$: logistic distribution

$(m_1, m_2) = (1, \infty)$: extreme value distribution

$m_2 = \infty$: generalized loggamma distribution

$(m_1, m_2) \rightarrow (\infty, \infty)$: normal distribution,

to select one particular member of the family. Gould (1986) did an extensive study of the location-scale model with the error ϵ following the $\log F$ distribution. Her conclusions are that if one tries to estimate two shape parameters as in the $\log F$ family, the precision of the estimates may be so low as to make them virtually uninformative. However, as we have also observed, the MLE $\hat{\beta}$ of the regression parameters is quite robust with respect to misspecification of the distribution of ϵ.

Numerous other researchers have in the past also encountered difficulty when trying to estimate the shape parameter of the generalized loggamma distribution. Lawless (1982, p. 237), observed that, even with sample sizes of 200 or 300, it is not uncommon for the Newton-Raphson algorithm not to converge to the MLE's. Because in usual insurance situations, the trapezium of data contains a small number of cells (in our case, 45 observations with 5 parameters to estimate), the actuary might encounter problems with this distribution. According to Prentice (1974), two distributions in the loggamma family with very different values of the shape parameter k, will look very similar, creating estimation problems. The extreme value distribution ($q = 1$) is difficult to discriminate from the normal distribution ($q = 0$), when the sample size is small.

In view of these facts, we therefore recommend that a simple distribution be assumed for ϵ, like the extreme value or the normal. After comparing the log likelihood, fit can be assessed by a Q-Q plot. If a symmetric distribution is needed, the normal distribution should be assumed for ϵ, since it is the only symmetric member of the generalized loggamma family. Fitting the normal model is useful for finding initial parameter estimates for the extreme value model. The estimated IBNR reserve can then be easily calculated under both assumptions.

The assumption of a normal distribution for ϵ presents one advantage over that of the extreme value distribution. When reserves are to be discounted for interest, we can still find the distribution of the present value of the future payments. If the force of interest δ is constant over a year, it follows from a property of the lognormal distribution that the joint distribution of the discounted value of the future payments is also multivariate lognormal. Stochastic interest rates could also be built into the model and the reserve estimated by simulation.

In conclusion, regression models present many advantages over the chain ladder method: they have fewer parameters and do not underestimate the reserve; the properties of the estimators of the parameters have been well studied; they take into account both the error involved in the estimation of the parameters and the statistical error inherent in the prediction of future claims; the fit of the model can be tested statistically by a $Q-Q$ plot; and confidence intervals for the reserve can be calculated with a simulation. We therefore strongly advocate the use of regression models.

10 Appendices

10.1 Algorithm to generate a multinormal random variable

To simulate the distribution of the IBNR reserve, we need to generate a $MLN(\mu, \sum)$ random variable. The following algorithm was used.

1. Generate $Z \sim MN(0, I)$, using the Box-Muller transformation

$$Z_1 = (-2 \ln U_1) \cos(2\pi U_2)$$

$$Z_2 = (-2 \ln U_2) \cos(2\pi U_2),$$

   where U_1 and U_2 are i.i.d., uniform on $(0, 1)$.

2. Transform Z to Y, a $MN(\mu, \sum)$ distribution:

$$Y = \mu + CZ,$$

   where $\sum = CC'$ and C is calculated from the Choleski factorization algorithm (ref. Kellison (1975)):

$$\begin{aligned} c_{11} &= \sqrt{\sigma_{11}} \\ c_{ij} &= \frac{1}{c_{jj}}(\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{kj}) \\ c_{ii} &= \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} c_{ik}^2} \end{aligned}$$

3. Exponentiate each component of Y

$$e^Y = (e^{Y_{kl}}) \sim MLN(\mu, \sum).$$

10.2 Asymptotic properties of MLE's

If $X_1, \ldots, X_n$ is a random sample of size n from the density $f(x; \underline{\theta})$, where $\underline{\theta} = (\theta_1, \ldots, \theta_{p+1})$ contains the regression parameter vector β and the scale parameter σ, then under certain regularity conditions, the following results hold.

1- The MLE $\hat{\underline{\theta}} = (\hat{\theta}_1, \ldots, \hat{\theta}_k)$ exists.

2- It is a consistent estimator of θ.

3- $\hat{\theta}_1, \ldots, \hat{\theta}_{p+1}$ are asymptotically efficient,

$$\text{i.e.} \quad \lim_{n \to \infty} \frac{Var(\hat{\theta}_j)}{CRLB(\hat{\theta}_j)} = 1,$$

where $CRLB(\hat{\theta}_j)$ is the Cramér-Rao lower bound, obtained as $1/nE[\frac{\partial \log f}{\partial \theta_j}]^2$.

4- $\sqrt{n}(\hat{\theta} - \theta)$ has an asymptotically multivariate normal $MN(\underline{0}, I_0^{-1})$ distribution where I_0 is the observed information matrix, with element

$$I_{ij} = -\frac{\partial^2}{\partial \theta_i \partial \theta_j} \log L(\underline{\theta}; x_1, \ldots, x_n) \mid_{\underline{\theta} = \hat{\underline{\theta}}} .$$

11 References

Abramowitz, M. and Stegun, I.A. (1972). Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, U.S. Department of Commerce, Washington.

Bain, L.J. and Engelhardt, M. (1991). Statistical Analysis of Reliability and Life-Testing Models, Theory and Methods, Marcel Dekker Inc., New York.

Burridge, J. (1981). A note on maximum likelihood estimation for regression models using grouped data. JRSS B, 43, 41-45.

Chhikara, R.S. and Folks, J.L. (1989). The Inverse Gaussian Distribution: Theory, Methodology and Applications. Marcel Dekker Inc., New York.

Cohen, C. (1988). Three-parameter estimation. In Lognormal Distributions, Theory and Applications, edited by Crow, E.L. and Shimizu, K., Marcel Dekker, Inc., New York.

Cohen, A.C. and Whitten, B.J. (1988). Parameter estimation in reliability and life span models. Marcel Dekker, Inc., New York.

Cox, D.R. and Hinkley, D.V. (1968). A note on the efficiency of least-squares estimates. JRSS B, 30, 284-289.

Devroye, L. (1986). Non-Uniform Random Variate Generation. Springer-Verlag, New York.

Doray, L.G. (1992). UMVUE of the IBNR in a Lognormal Linear Regression Model, submitted to Insurance: Mathematics and Economics.

Farewell, V.T. and Prentice, R.L. (1977). A study of distribution shape in life testing. Technometrics, 19, 69-75.

Gould, A. (1986). Some issues in the regression analysis of survival data. Ph.D. thesis, University of Waterloo.

Gould, A. and Lawless, J.F. (1988). Consistency and efficiency of regression coefficient estimates in location-scale models. Biometrika, 75, 535-540.

Gradshteyn, I.S. and Ryzhik, I.M. (1980). Tables of Integrals, Series and Products. Academic Press, New York.

Hogg, R.V. and Klugman, S.A. (1984). Loss Distributions. John Wiley and Sons, Inc., New York.

Johnson, N.L. and Kotz, S. (1970). Distributions in Statistics: Continuous Univariate Distributions, Vol. 1 and 2, Houghton Mifflin Company, Boston.

Jörgensen, B. (1982). Statistical Properties of the Generalized Inverse Gaussian Distribution. Springer-Verlag, New York.

Kalbfleisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data. Wiley, New York.

Kellison, S.G. (1975). Fundamentals of Numerical Analysis. Richard D. Irwin, Inc., Illinois

Koziol, J.A. (1989). Handbook of Percentage Points of the Inverse Gaussian Distribution. CRC Press, Inc., Boca Raton, Florida.

Lawless, J.F. (1982). Statistical Models and Methods for Lifetime Data. Wiley, New York.

Michael, J.R., Schucany, W.R. and Haas, R.W. (1976). Generating random variates using transformations with multiple roots. The American Statistician, 30, 88-90.

Panjer, H.H. and Willmot G.E. (1992). Insurance risk models. Society of Actuaries, Schaumburg.

Prentice, R.L. (1974). A loggamma model and its maximum likelihood estimation. Biometrika, 61, 539-544.

Reinsurance Association of America (1987). Loss Development Study. In Proceedings of the XXIth Astin Colloquium, New York, 1989.

SAS (1985) User's Guide: Statistics, Version 5 Edition, SAS Insitute Inc. Cory, NC.

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

White, H. (1982). Maximum likelihood estimation of misspecified models. Econometrica, 50, 1-25.

Zenwirth, B. (1990). Interactive Claims Reserving Forecasting System, Version 5.3, Manual Volume 1, Benhar Nominees Pty. Ltd., Australia.

A Generalized Framework for the Stochastic Loss Reserving

by Changseob Joe Kim

A GENERALIZED FRAMEWORK FOR THE STOCHASTIC LOSS RESERVING

The traditional actuarial methods like loss (paid and incurred) development methods, Bornheutter-Ferguson method, or Berquist-Sherman method have been served well as long as point estimates are concerned. Since they are not stochastic approaches, they do not provide confidence intervals which are getting more attention connected to the risk-based capital requirements, explicit discounting the future liabilities, etc. So far, most of the stochastic reserving models which are either in the developing stage or are being used by some companies or organizations, have been explanatory models. The Hoerl curve fitting is their basic formulation. These types of models are fundamentally deficient, because they fit the Hoerl curve to the loss history data. Hoerl curve fitting may be fine, as long as it fits a simple, one dimensional, small series of data to obtain a fitted curve without any statistical implications. If the Hoerl curve fitting method is used with some statistical perspectives in mind, it may produce inconsistent estimtates which may not make any sense. In this article, the author suggests a generalized framework which starts by understanding the unique data characteristics of the insurance data. By expanding a Box-Jenkins type time-series model, we developed a generalized framework for modelizing a stochastic process on the loss history data. It turned out that some lines require more complex specifications than the others. We may presume that some lines are more sensitive to the insurance business cycle than the others. Our contributions will be to provide a generalized framework to derive confidence intervals in which the business cycle was taken into account as well as to provide future estimates for the planning process. This paper is the first step to that direction.

I. INTRODUCTION

Insurance data arranged to evaluate future liabilities takes a unique form which is different from ordinary non-insurance data. The ordinary non-insurance data usually takes a one-dimensional time-series form. For example, monthly unemployment figures for the period January 1948 – October 1977 was used to forecast November 1977 and onward monthly unemployment rate. On the while, the insurance data has to be arranged either by accident year, policy year or report year and development year in order to figure out the future liabilities of each of those years separately. Because of this, the typical insurance data takes an upper triangular form.

The traditional actuarial methods like loss (paid and incurred) development methods, Bornheutter-Ferguson method, or Berquist-Sherman method have been served well as long as point estimates are concerned. Since they are not stochastic approaches, they do not provide confidence intervals which are getting more attention connected to the risk-based capital requirements, explicit discounting the future liabilities, etc.

There have been hundreds of methods which were contended to provide confidence intervals. The fundamental problems of these methods are they are lacking in theoretical backgrounds because these methods are intended to apply to the one-dimensional data array. Minor adjustments are added to solve the problems. However, they have never been successful.

In this article, the author suggests a generalized framework which starts by understanding the unique data characteristics of the insurance data. In the next chapter, we provide the critics regarding the problems of those suggested stochastic methods. In chapter III, we articulate the characteristics of the insurance data. We also state how these characteristcs have been incorporated in the traditional actuarial methods. In chapter IV, the theoretical framework will be provided. We will show some applicaitons in chapter V and conclude in chapter VI.

II. Critics on suggested stochastic models

Makridakis and Wheelwright (1985) suggested:

> If the user wants to increase forecasting accuracy, a time series method should be used. If the objective is to understanding better the factors that influence forecasting (prediction) accuracy, then an explanatory model should be selected.

So far, most of the stochastic reserving models which are either in the developing stage or are being used by some companies or organizations, have been explanatory models. The Hoerl curve fitting is their basic formulation. First of all, the explanatory variables in their models are either the number of development years and its functional variations, the number of accident years, the number of calendar years or a combination of these. Because of these formulations, their explanatory variables do not explain the dependent variable quite well. For example, "increase one unit of log transformed development years will decrease .3 unit of total loss paid" does not provide any valuable information.

Secondly, normally it is assumed that the time series data consists of four parts of components. They are trend, seasonality, cycle and ramdom components. If we use time and its functional variation as only explanatory variables, we are ignoring the seasonal and cyclical components of data. If the annual data is used, we may ignore the seasonality, but not the cyclical component. Since some insurance business is sensitive to the business cycle, we may expect that the cyclical movement is a critical component of the data.

Thirdly, since one of the explanatory variables is a functional variation of the other, these two explanatory variables are highly correlated. This problem is called multicollinearity. If one of these two variables is deleted, there will be an autocorrelation problem because the remaining explanatory variables will not fully explain the dependent variable. The consequences of these problems include: unstable estimates, spurious predictions, inconsistent estimation of standard errors and confidence intervals.

Some argue that as long as the autocorrelations between the two explanatory variables are lower than that bewteen the dependent and explanatory variables, we do not have to worry about this problem. This may be true if the two explanatory variables are independently created. This is why explanatory variables are sometimes called independent variables. They are supposed to be independent. However, as long as correlations between these explanatory variables are not high compared to correlations between dependent variable and explanatory variables, the problem may not be that serious. The issue here is whether we should use models which contain multicollinearity problems due to the model formulation (one of the explanatory variables is a functional variation of another).

The other problem of these types of explanatory models is what type of indicator we should use for the accident year trends. Some authors normalized all incremental payments based on some readily available index of inflation. We cannot simply divide incremental payments by some indices, because these indices are estimated with their own variances. Consequently, it requires to assume that these indices are deterministic. However, this assumption is hardly persuasive at all. Because of this problem, some authors divide the payments by some types of exposures. The problem of this approach is we need to find an alternative if there isn't any exposure data available, which is often the case. Still others introduce level parameters which are assigned same values to each accident years. Since the level parameters themselves have to be estimated, this automatically violate the assumption that explanatory variables are supposedly nonrandom variables which are the cases of the other two variables. Others create another explanatory variable using the sum of the accident year and the development year. They chose this as another explanatory variable because they could not use the number of accident years as their explanatory variable due to the perfect linearity with the number of development years. This choice is as bad as choosing the number of development years as an explanatory variable.

Still another problem of this type of model is that they do not provide any method that deals with interrelationships between series of incremental payments and incremental claims reported. Other things being equal, we expect more incremental payments if there are more claims reported. Therefore, if claims reported data is available, we should utilize these data assuming that this is also a stochastic process. So far no method has been suggested to deal with this situation. Some authors apply traditional loss development approach in obtaining ultimate claims reported. They treat them as a determinstic variable to divide incremental payments by these estimated ultimate claims reported.

What if we need to analize quarterly data instead of annual data? Quite possibly that quarterly data may contain seasonal patterns. No methods have been suggested to deal with this seasonality problem.

These types of models are fundamentally deficient, because they fit the Hoerl curve to the loss history data. Hoerl curve fitting may be fine, as long as it fits a simple, one dimensional, small series of data to obtain a fitted curve without any statistical implications. If the Hoerl curve fitting method is used with some statistical perspectives in mind, it may produce inconsistent estimtates which may not make any sense.

III. Insurance Data as a Two-Dimensional Time-series

1. Data itself.

Insurance loss or claim history data can be considered as a two dimensional time series data. Loss or claim development, in which additional losses or claims are paid/reported in chronological order upon accidents occurred or claims reported is one dimension. A chronogical order of claims grouped by date of occurence is another dimension. As a result, a typical insurance loss or claim history takes an upper triangle form. A prediction of future loss payments or claims reported corresponds to filling out the bottom lower triangle area assuming that the first accident or reported year losses or claims are fully developed.

There are at least two factors which cause loss history data as time-series through the accident years. The first factor is inflation. Ever increasing price levels (at least prior to the current recession) is called economic inflation. Increased tendency to file more claims helped by trial lawyers or increasing amount of jury awards is called social inflation. Some authors have tried to catch these inflations by either normalizing the incremental payments or by inserting a level parameter. The indices used were either general price indices or at most industry-specific indicator. Because of ever increasing tendencies of the loss payment and these general indices, you may obtain significant t-values for the estimated coefficient of these indices. These t-values are disguising. Even if you insert any series which is increasing, you may still obtain significant t-values. Instead of inserting or dividng by an extraneous series, we should use the data's own indices! We should look at every trend and/or cyclical pattern of incremental payment of each development year. Interestingly, there is an approach which utilizes these trends to estimate ultimate losses. The problem is it is not a stochastic approach. We cannot obtain confidence intervals based on this approach. We will present this approach later.

As more consumers or insureds are getting more information on their insurance policy provisions, and as more trial lawyers are eagerly recruiting their clients, we can expect more claims to be reported over the accident year horizon. As overall population grows, there will be more policies written. Other things being equal, consequently there will be more claims reported. These utilization increase and additional new polcies will be the main driving force for the consistent upward trend through the accident year horizon.

For the development horizon, since there is a fixed number of policies written during the policy effective period, there is a fixed number of occurrence of accidents for each accident year. There may be some incurred but not reported claims which are reported later. There may be some cumulative injury claims which take many years to be closed. Still every claim will be closed eventually. In a mathematical term, total cumulative loss payments or total reported claims will be converged to certain levels. Because of this characteristic, all incremental payments and all incremental reported claims will be automatically satisfied with the stability condition of the time-series analysis. This stability is a necessary condition in applying Box-Jenkins types time-series framework.

The traditional actuarial method called the "loss or claim development method ", utilizes the development period dimension in a simple manner. The accident period dimension in this method is partially utilized by taking current cumulative payments as "given". Recently proposed regressional approaches are lacking in these two dimensional features. As in the traditional actuarial loss development (LD) method, these new methods reflect the loss development dimension by using "age" of loss development. However, the other dimension is either completely ignored or grouped together by assigning dummy variables or filled with a so-called level parameter. There is an inherent autocorrelation problem which may not be significant in some lines due to negligence of the time related features in the loss history data, especially for long tail lines in which regulators or company's executives are most interested.

In the traditional development approach, by multiplying the selected factors for each development year, some sort of time-series conception was used in a simple fashion. For instance, assuming that there are no additional payments after ten years of development, the ultimate factor for the 1982 accident year will be obtained by taking a ratio of the 10th year development to the 9th year of development. Notice that only the accident year 1981 and prior provides the information required to obtain a factor for the 9th to 10th development. The ultimate factor for 1983 is derived through multiplying the selected factor from the 8th to 9th year of development by the selected factor from the 9th to 10th year of development. Again the selected factor for 8th to 9th year of development is based on the factors which are available in 1982 and prior accident years. Although it is a simple fashion, without a consideration of cyclical patterns, the development method reflects time series characteristic through development years. In the accident year direction, the LD method simply takes most current actual payments as selected estimates. If these values are outliers, the LD method will generate biased estimates. Otherwise, the LD method will produce reasonable estimates. For the older accident years, the actual values are fairly close to the estimates which are supposed to be compared to its maturity because the payments have already been made quite a few times (approximately more than 3 or 4 years for short tail lines). The problem is most recent immature accident years. Bornheutter-Ferguson (B-F, 1978) and Berquist-Sherman (B-S, 1979) suggested a couple of methods to get over these problems.

2. Time-series Reflected in B-F Method.

In the adjusted development method suggested by Bornheutter and Ferguson, a two-year average of total payment at a particular development adjusted by the increase or decrease in the second year's exposure relative to the two-year average exposure was replaced for total payment. The ultimate factors derived in the development method is then applied to these adjusted losses. This method will correct some irregularities of the data. However, the adjustments contain too short memory (one year backward). The probability of two data points being outliers is only half of the probability of one data point being an outlier. Consequently, this does not provide appropriate remedies to correct the problem in the development method. This may be the reason why this method is seldomly used in the ordinary actuarial analysis.

In the well-known B-F approach, the expected losses are first derived. Unpaid factors are then calculated from the ultimate development factors. The ultimate losses are estimated as the sum of total payment and indicated reserve, where indicated reserve is expected loss times the unpaid factor. Two methods are suggested to calculate the expected loss. The undiscounted loss provisions in the rates multiplied by the units of exposure is one, trending, or otherwise extrapolating, $\frac{\text{ultimate loss}}{\text{ultimate claim count (or premium)}}$ relationships of the prior accident years is the other. The author prefers the latter methods based on two reasons. First, it is very difficlut to obtain the undiscounted loss provision. One of the major reasons is the differences in line-breakdown between pricing and reserving. Second, by trending the past history, we can glean the time-series nature of the loss history data. You may notice that in LD method, only the time-series nature across the development years was recognized. By applying trending or extrapolating method to $\frac{\text{ultimate loss}}{\text{ultimate claim count}}$ across the accident years, we are able to utilize the time-series nature in another dimension at least partially (cosidering only trend factors).

This indicated (B-F) method is one of the most popular methods in the actuarial analysis because this method can be used to correct the estimated ultimate loss for the recent accident years produced by the development method.

Although these two methods are a little more advanced than loss development methods in terms of utilizing the time-series nature across the accident years, the method is not sophisticated and also performed partially (only trend factors are considered). Instead of trending a whole loss history across the accident years, only the indicated severity for each accident year was used. Since the indicated severity is also estimated, it may be contaminated with estimating errors. Berquist and Sherman suggested a few methods which utilize a whole loss history in a simple fashion.

3. Time-series Reflected in B-S Method.

Berquist and Sherman suggested six methods (Method I through VI) except for Method II which is exactly paid loss development method applying weighted average to loss development factors in order to obtain ultimate development factors, all methods assume that there are some trends to be utilized across the accident years. Method I applies a straight linear regression to the loss development factors for each development years as long as there are at least three factors. For columns with two factors, a straight average is taken for all future development factors. For columns which only one factor, that factor is used.

In Method III, the total payments per ultimate claim count ($CS_{i,j}$) by accident year (i) and by development year (j) are calculated. By applying a exponential fit to $CS_{i,j}$ for each j, a growth rate B_j for each development year j is estimated. Then by multiplying e^{B_j} by $DS_{i,j}$ where $DS_{i,j}$ is the incremental payment for the accident year i and development year j, we obtain a incremental payment on current cost level $IS_{i,j}$. After applying appropriate weights to these $IS_{i,j}$, the estimated incremental payments evaluated as of current date $WS_{i,m-i+1}$, where $i = m, m-1, \ldots, l$, the oldest accident year and m the latest accident year are calculated. By applying growth rate e^{B_j} to $WS_{i,m-i+1}$, future incremental payment per claim is produced. After adding them up across the development years to obtain ultimate loss per claim, ultimate loss is derived by multiplying the ultimate claim count.

In Method IV, overall growth rate is calculated by weighting various column growth rates calculated in Method III, in proportion to the square of number of rows of that column. The adjusted column growth rate is then calculated by applying the formula $B'_j = \frac{W_j R_j + (W_1 - W_j) R}{W_1}$ where W_j is the weight for the particular column, W_1 is that for the initial colmun (development year 1) and R_j is column growth rate. The same procedure with the Method III is then applied to produce the ultimate loss.

In Method V, the paid loss development factors minus unity are used instead of total payment per claim in Method IV to derive growth factor for the development factors. After applying the same steps as in Method IV to derive future factors (minus one), adding one to each of the results and applying resulting factors to total payments, the ultimate losses are derived. In Method VI, the incremental payments per claim are used to estimate growth rate. The exact same steps as Method IV are then used.

Notice that in the various Berquist-Sherman methods except for Method II, more emphases are levied on the trends across the accident years. In Method I and Method III, the trend factors (growth rates) are estimated by development years. Each trend factor for a particular development year is independent of those of the other development years. On the while, in the Method IV, V, and VI, the overall trend factor was calculated by the weighted average of all the trends for each development years. The adjusted trend for individual development year was then calculated as a weighted average of its own trend and the overall trend. Since these methods are focused on the time-series nature of the loss history across the accident years ignoring possible cyclical patterns, by combining the ultimate loss based on these method and the ultimate loss based on the loss development method, we can produce relatively reasonable selected ultimate loss.

As we have seen in this chapter, even if the word of time-series has never been spelled out, one way or the other, every method tried to utilize the time-series concept. The trouble was that the concept was utilized partially. Except for Berquist-Sherman methods, more weights were given to the claim development process. Even in one direction, only the trend component of the time-series was reflected. A cyclical movement and seasonal pattern were completely ignored. In our approach, the two dimensions are explicitly taken into account. Today's loss payment is not only a function of losses paid in the past loss development periods, but also a function of losses paid in the past accident periods. The implication of various statistics in the time series method are also considered in a two dimensional perspective. Empirical results based on various lines of industry total are shown.

IV. A Framework of Two Dimensional Time Series Model

1. The Univariate Model.

1) Assumptions

In this univariate model, we assume that only the payment series is available. There is no reliable case reserve, exposure or reported claim information available. More often than not, actuaries, especially consulting actuaries, have to provide ultimate loss payment based on exclusively loss payment series.

We also assume that the available data is not separable to the individual claim level. In other words, we treat the incremental payment for a particular accident period and development period itself as a random variable. This is a realistic assumption because most loss history data takes an upper triangular form in which the incremental payment is a minimum unit of counting.

We assume that the tail of the loss payment development is known. This assumption may not be realistic. However, it is at least practical. Whenever we fit any distributional curve to the loss payment developments, the estimated curve converges to the ultimate level a lot more slowly than we ever expect in actual loss developments. Unless we assume a certain cut-off point, the estimated length of the development will be extremely long.

We assume that any payment in a certain point is affected only orthognally. For example, total or incremental payment in [accident year 83 – third development year] is a function of [accident year 83 – second deveopment year] and [accident year 82 – third development year]. This is a reasonable assumption to simplify the algorithms and also consistent with the average norm. We can expect the incremental payment at [accident year 83 – third development year] will be high if the incremental payment at [accident year 83 – first and second development years] due to either volume increase or frequency/severity increase. Also we can expect the incremental payment at [accident year 83 – third development year] will be high if the incremental payments at [third development year – accident year 81 or 82] are high. The former tendency may be related to the inflation, exposure, and frequency/severity change. The latter may be related to the company's individual line characteristics – like a liability line develops more slowly than a property line.

Finally, we assume that the selected model is the true model. In others words, specification error is ignored. This error exits only in a hypothetical sense. Since in reality the true model is never known, you can never measure the direct error. This assumption is consistent with most econometric or time-series literatures. By assigning higher probability confidence intervals than what is necessary, we can eliminate the specifiaction error problem. For example, if the confidence intervals with 90% probability is required, then by raising the probability to the 95% level, we may take into consideration the specification error problem.

2) Model

Parzen suggested a very powerful time-series forecasting model. It extends the Box-Jenkins methodology and provides a more practical alternative to the time-series forecasting model. Also the theoretical supports of "$ARAMA$" models are solid and their potential contribution to good forecasting is excellent.

Contrary to the Box-Jenkins methodology, Parzen's approach is not as concerned with parsimony. Parzen's model is willing to sacrifice the parsimony that would result from introducing the moving average terms, and simply includes more autoregressive terms. The MA terms are available but used only for special cases when a scheme cannot be used to produce random residuals.

We utilize Parzen's view of Box-Jenkins time-series methodology. The main reason is the tractability without giving away any theoretical merits. In our application, the stability may not be an important issue. In the development period horizon, because any open claim will be closed eventually, the convergence of the time-series is guranteed. In the accident period, due to the regulation constraint of premium-surplus ratio, there exists a limit of maximum expansion. Consequently, as long as there are enough data points, we expect the stability condition will be met in the average insurance data.

Across the accident year we restrictly use AR terms. However, across the development year, we first take differencing on the total payments and then take log transformation if it is possible. After transforming long memory time series across the development years, the AR terms are used to produce white noise errors.

It is a matter of semantic, whether you need a differencing operation or not across the development years. If you start with incremental payment data, there is no need of differencing. However, if you start with the total payment data, you do need differencing due to the conspicuous cumulative nature of the payment data.

In a general form we can express the model as:

$$F(IP_{i,j})) = \sum_{l,k} \phi_{l,k} F(IP_{i-l,j-k}) + e_{i,j} \quad l = 0, 1, 2, ..., i-1$$
$$\text{and} \quad k = 0, 1, 2, ..., j-1 \quad \text{excluding} \quad l = 0 \;\;\&\;\; k = 0 \qquad (4\text{-}1)$$

where F(.) notates any functional form (most of the case log operator if it is possible, otherwise identity operator), IP denotes incremental payment for the accident year i - development year j. Since we assumed any non-orthogonal lag variables can be ignored, equation 4.1 can take much simpler form as:

$$F(IP_{i,j}) = \sum_{l,k} \phi_{l,k} F(IP_{i-l,j-k}) + e_{i,j} \quad l = 1, 2, ..., i-1 \;\;\&\;\; k = 0$$
$$\text{or} \quad k = 1, 2, ..., j-1 \;\;\&\;\; l = 0 \quad \text{excluding} \quad l = 0 \;\;\&\;\; k = 0 \qquad (4\text{-}2)$$

Note that since no nonlinearity is invloved, we can use Ordinary Least Square Method to estimate $\phi_{l,k}$. This is a whole advantage expressing the model with AR terms only. The most simple case will be:

$$IP_{i,j} = \phi_{1,0} IP_{i-1,j} + \phi_{0,1} IP_{i,j-1} + e_{i,j} \tag{4-3}$$

where the incremental payment for the accident i - development j is explained the incremental payment of the one year previous accident year and the incremental payment of the one year previous development year.

For a better understanding, an example will be followed. Say you allow two lags in each direction as explanatory variables. Then there are eight possible explanatory variables. They are [No lag in accident year(AY) – 1 lag in development year(DY)], [No lag in AY – 2 lag in DY], [1 lag in AY – 1 lag in DY], [1 lag in AY - 2 lag],[2 lag in AY – 1 lag in DY],[2 lag in AY – 2 lag in DY], [1 lag in AY – no lag in DY], [2 lag in AY – no lag in DY]. Out of these eight combinations, the set of DY lag only is orthogonal to the set of AY lag only (four cases).

First of all, it does make sense modelizing the fact that the current incremental payments is explained by previous incremental payment series by accident and development year-wise because the current payment can be explained or can be a function of prior payments. Second, it does not have any multicollinearity problem because there is no functional relationship between the explanatory variables (note that accident year series are orthognal to the development series). Third, because it does not involve any nonlinearity, it is fairly easy to estimate parameters. Even we can use Lotus 1-2-3 to estimate these parameters. Fourth, most importantly, it provides a reasonable fit and also is also stable.

3) Interval Forecasts

Since the major contribution of the stochastic method in loss reserving is providing the confidence intervals, the variance of the forecast errors should be well defined. In order to derive the variance of the forecast errors, we first express $AR(l,k)$ process in the error-shock form by successive substitution for $\sum \phi_{l,k} IP_{i-l,j-k}$. By doing this, we can write the model in terms of current and past errors only as:

$$IP_{i,j} = e_{i,j} + \xi_{0,1} e_{i,j-1} + \xi_{1,0} e_{i-1,j} + \xi_{1,1} e_{i-1,j-1} + \cdots \tag{4-4}$$

The values of the parameters $(\xi_{0,1}, \xi_{1,0}, \xi_{1,1}, \ldots)$ depend upon the particular $AR(l,k)$ model and are called *error learning coefficients*.

The selected forecast $IP_{i,j}(g,h)$ can also be expressed using the equation 4-4 in terms of current and past errors:

$$IP_{i,j}(g,h) = \xi_{g,h} e_{i,j} + \xi_{g+1,h} e_{i-1,j} + \xi_{g,h+1} e_{i,j-1} + \cdots \tag{4-5}$$

As a result, the (g,h) step ahead forecast error can be expressed as:

$$e_{i,j}(g,h) = IP_{i+g,j+h} - IP_{i,j}(g,h) \tag{4-6}$$

Again the equation 4-6 can be written as:

$$e_{i,j}(g,h) = e_{i+g,j+h} + \xi_{1,0} e_{i+g-1,j+h} + \xi_{0,1} e_{i+g,j+h-1} + \xi_{1,1} e_{i+g-1,j+h-1} + \cdots \tag{4-7}$$

Because the errors are independent, it follows from the equation 4-7 that $e_{i,j}(g,h)$ is an $MA(g-1,h-1)$ process. From the equation 4-7, the forecast errors $e_{i,j}(g,h)$ have mean 0 and variance equal to

$$V[e_{i,j}(g,h)] = E[e_{i,j}^2(g,h)] = \sigma_e^2 \sum_{p,q=0}^{g,h} \xi_{p,q}^2 \quad \text{excluding} \quad (p,q) = (g,h) \tag{4-8}$$

Based on the model, not only can the future development year forecast be performed, but also the accident year forecast. However, since our main objective is to obtain confidence intervals for the future liabilities, we can focus on the development year horizon only.

4) Some Examples

For example, the one year ahead forecast to the development period horizon of the $AR(1,1)$ model can be expressed using equation 4-3 as:

$$IP_{i,j+1} = \phi_{1,0} IP_{i-1,j+1} + \phi_{0,1} IP_{i,j} + e_{i,j+1} \tag{4-9}$$

Then the equation 4-9 can be expressed as:

$$\begin{aligned} IP_{i,j+1} =& \phi_{1,0}(\phi_{1,0} IP_{i-2,j+1} + \phi_{0,1} IP_{i-1,j} + e_{i-1,j+1}) \\ & \phi_{0,1}(\phi_{1,0} IP_{i-1,j} + \phi_{0,1} IP_{i,j-1} + e_{i,j}) + e_{i,j+1} \end{aligned} \tag{4-10}$$

Since the only errors terms $e_{i-1,j+1}$, $e_{i,j}$ and $e_{i,j+1}$ are unkown and their variances are σ_e^2, the variace of $IP_{i,j+1}$ can be expressed as:

$$V(IP_{i,j+1}) = (\phi_{1,0}^2 + \phi_{0,1}^2 + 1)\sigma_e^2 \tag{4-11}$$

The two year ahead forecast to the development period will be:

$$IP_{i,j+2} = \phi_{1,0} IP_{i-1,j+2} + \phi_{0,1} IP_{i,j+1} + e_{i,j+2} \tag{4-12}$$

Again, the equation 4-12 can be expressed as:

$$\begin{aligned} IP_{i,j+2} =& \phi_{1,0}(\phi_{1,0} IP_{i-2,j+2} + \phi_{0,1} IP_{i-1,j+1} + e_{i-1,j+2}) \\ =& \phi_{0,1}(\phi_{1,0} IP_{i-1,j+1} \phi_{0,1} IP_{i,j} + e_{i,j+1}) + e_{i,j+2} \end{aligned} \tag{4-13}$$

By applying the equation 4-10, we can obtain a two year ahead forecast variance to the development period as:

$$V(IP_{i,j+2}) = ((\phi_{1,0}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + (\phi_{0,1}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + 1)\sigma_e^2 \tag{4-14}$$

Similarly we can obtain an n year ahead forecast variance to the development period by applying a inductive procedure as:

$$V(IP_{i,j+n}) = ((\phi_{1,0}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + (\phi_{0,1}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + 1)\sigma_e^2 \tag{4-15}$$

We can also apply the same inductive process to the $AR(2,1)$ or $AR(3,1)$ model. For the $AR(2,1)$ model, one year head, two year ahead and n year ahead forecast variances are given as:

$$V(IP_{i,j+1}) = (\phi_{1,0}^2 + \phi_{0,1}^2 + 1)\sigma_e^2 \tag{4-16}$$

$$V(IP_{i,j+2}) = ((\phi_{1,0}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + (\phi_{0,1}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + \phi_{0,2}^2 + 1)\sigma_e^2 \tag{4-17}$$

$$\begin{aligned} V(IP_{i,j+n}) = &((\phi_{1,0}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + (\phi_{0,1}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + \\ &(\phi_{0,2}^2)(\frac{V(IP_{i,j+n-2})}{\sigma_e^2}) + 1)\sigma_e^2 \end{aligned} \tag{4-18}$$

For the $AR(3,1)$ model, one year head, two year ahead, three year ahead and n year ahead forecast variances are given as:

$$V(IP_{i,j+1}) = (\phi_{1,0}^2 + \phi_{0,1}^2 + 1)\sigma_e^2 \tag{4-19}$$

$$V(IP_{i,j+2}) = ((\phi_{1,0}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + (\phi_{0,1}^2)(\phi_{1,0}^2 + \phi_{0,1}^2 + 1) + \phi_{0,2}^2 + 1)\sigma_e^2 \tag{4-20}$$

$$\begin{aligned} V(IP_{i,j+3}) = &((\phi_{1,0}^2)(\frac{V(IP_{i,j+2}}{\sigma_e^2})) + (\phi_{0,1}^2)(\frac{V(IP_{i,j+2}}{\sigma_e^2}) + \\ &\phi_{0,2}^2(\frac{V(IP_{i,j+1}}{\sigma_e^2}) + \phi_{0,3}^2 + 1)\sigma_e^2 \end{aligned} \tag{4-21}$$

$$\begin{aligned} V(IP_{i,j+n}) = &((\phi_{1,0}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + (\phi_{0,1}^2)(\frac{V(IP_{i,j+n-1})}{\sigma_e^2}) + \\ &(\phi_{0,2}^2)(\frac{V(IP_{i,j+n-2})}{\sigma_e^2})(\phi_{0,3}^2)(\frac{V(IP_{i,j+n-3})}{\sigma_e^2}) + 1)\sigma_e^2 \end{aligned} \tag{4-22}$$

If we expect any seasonality either across the development horizon or across the accident horizon or both, by inserting $\phi_{0,m}$ or $\phi_{m,0}$ or both lags, we can take care of seasonality, where m is the seasonality interval.

2. The Multivariate Model.

By applying either vector autoregressive model or transfer function model, we can expand the univariate model to the multivariate mode. Either closed counts development or reported counts development will be a good candidate for the right-hand side variable because we can presume that the claim counts will have a impact on the loss development; not vice versa. It is theoretically possible to derive the formula for the variances. However, we decided to postpone further articulation of the model due to the time constraint.

V. Model Selection Process with Empirical Data

1. Statistics to be used.

In order to find a right (or reasonable) model, we need certain criteria to identify whether the estimated errors are not correlated. Since we are going to use the $AR(l,k)$ model, we need to estimate partial autocorrelations (PCAF) of the residuals. We also use Q-statistic to verify overall randomness of errors. Since these statistics are intended to serve for the one-dimensional data, we have to apply these statistics to each accident year and development year separately. Because of this, we may have to be a little lenient when we reject the null hypothesis.

1). Partial Autocorrelation.

In practice, we never know the population values of autocorrelations and partial autocorrelation of the underlying stochastic process. Consequently, in identifying a tentative model, we must use the estimated autocorrelation and estimated partial autocorrelation to see if they are similar to those of typical models for which the parameters are known. Notice that since we do not have any MA terms in our model, there is no need to calculate estimated autocorrelations. However, partial autocorrelations are calculated from a solution of the Yule-Walker equation system, expressing the partial autocorrelation as a function of the autocorrelation. We need to calculate estimated autocorrelation.

In any time series textbook, an estimate of autocorrelation $r(h)$ is defined as:

$$r_h = \frac{c_h}{c_0} \qquad (5\text{-}1)$$

where c_h defined as $c_h = 1/n \times \sum z_t z_{t+h} \quad h \geq 0$, and c_h is the estimate of the autocovariance. For our model we can redefine this estimated autocorrelation for the development year dimension of the accident year n as:

$$r_{n,k} = \frac{c_{n,k}}{c_{n,0}} \tag{5-2}$$

in which $c_{n,k} = 1/m \sum_{j=1}^{m} z_{n,j} z_{n,j+k} \quad k \geq 0$ where m is the number of development years. For the accident year dimension of the development year m, the estimated correlation can be defined as:

$$r_{l,m} = \frac{c_{l,m}}{c_0} \tag{5-3}$$

where $c_{l,m} = 1/n \sum_{i=1}^{n} z_{i,m} z_{i+l,m} l \geq 0$. And n is the number of accident years.

The Yule-Walker equation is expressed as:

$$\begin{pmatrix} \rho_1 = & \phi_1 & + & \phi_2\rho_1 & + & \cdots & + & \phi_p\rho_{p-1} \\ \rho_2 = & \phi_1\rho_1 & + & \phi_2 & + & \cdots & + & \phi_p\rho_{p-2} \\ \vdots & \vdots & + & \vdots & + & \ddots & + & \cdots \\ \rho_p = & \phi_1\rho_{p-1} & + & \phi_2\rho_{p-2} & + & \cdots & + & \phi_p \end{pmatrix} \tag{5-4}$$

The equation 5-4 can be written as:

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{pmatrix} \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{pmatrix} \tag{5-5}$$

Hence, as soon as we calculate these autocorrelation, we can derive the estimated partial autocorrelations by applying Box and Jenkins's recursive method, which are due to Durbin(1960):

$$\hat{\phi}_{p+1,j} = \hat{\phi}_{p,j} - \hat{\phi}_{p+1,p+1}\hat{\phi}_{p,p-j+1} \qquad j = 1, 2, \ldots, p \tag{5-6}$$

$$\hat{\phi}_{p+1,p+1} = \frac{r_{p+1} - \sum_{j=1}^{p} \hat{\phi}_{p,j} r_{p+1-j}}{1 - \sum_{j=1}^{p} \hat{\phi}_{p,j} r_j} \tag{5-7}$$

In order to identify the exact form of the model, we need to find out when population partial autocorrelations can be considered to be zero. We therefore need to evaluate the standard error of the estimated partial autocollreations. Quenouille (1949) showed that the variance of the estimate of the partial autocorrelations is approximately equal to

$$V(\phi_{hh}) \approx 1/n, \qquad h > 0 \tag{5-8}$$

where n equals the number of observations after suitable differencing and transformation, and ϕ represents the partial autocorrelations that are assumed to be zero. Equation 5.8 provides a way, after identifying the tentative model, by calcuating ϕ_{hh} on the estimated residuals, to evaluate if all other estimated partial sutocorrelations are different from zero. We can also define the variance of the estimate of the partial autocorrelation for the development year dimension as:

$$V(\phi_{n,kk}) \approx 1/m, \qquad k > 0 \tag{5-9}$$

and for the accident year dimension as:

$$V(\phi_{ll,m}) \approx 1/n, \qquad l > 0 \tag{5-10}$$

2). Q-test.

Box and Pierce (1970) showed that for a purely random process, that is, a model with all $\rho_k = 0$, the statistic called Q-statistc:

$$Q(K) = n(n+2)\sum_{k=1}^{K} \frac{1}{n-k}\hat{r}_k^2 \approx \chi^2(K) \tag{5-11}$$

where $\hat{r}_k$ is defined as

$$\hat{r}_k = \frac{\sum_{t=k+1}^{n} \hat{e}_t \hat{e}_{t+k}}{\sum_1^n \hat{e}_t^2} \tag{5-12}$$

with $\hat{e}$ is a fitted residual. It should be noted that the Q-test is not a very powerful test for detecting specific departures from white noise. However, it is useful to check how a series of autocorrelations (first order, second order and third order autocorrelations etc.) is white noise or not in an overall sense. Furthermore, the Q-test is also sensitive to the values of

K, the number of autocorrelations used to calculate Q-test. For economic data, $K = 12$ and $K = 24$ have proven to be useful. Since insurance data have fewer data points, $K = 4$ may be sufficient. Since the Q-statistic was also designed to apply to the one dimensional data points, we performed the Q-test on each accident year and each developmemt year.

2. Creation of Auxiliary Observations.

We first calculate age-to-age factors for each dvelopment years. We then select age-to-age factor for each development years based on the last 5 years average method. We assume that payments of the Homeowner/Farmowners (HOMFAM), Private Passenger Automobile Liability/Medical (PRVAUT), Commercial Auto/Truck Liability/Medical (COMAUT), Commercial Multiple Peril (COMMUL), Workers' Compensation (WOKCOM), Medical Malpractice (MEDMAL), Special Liability (SPELIA), Other Liability (OTHLIA) and Product Liability (PROLIA) are paid off at 10th, 11th, 13th, 13th, 14th, 16th, 11th, 15th and 16th years of development, respectively. With this tail-factor assumption we create future incremental payments based on the LD method. In other words, we fill out the lower part of triangles.

There are two purposes in creating these auxiliary observations. The first purpose is creating initial values of lag variables based on the backward forecasting. Since we started with small amount of data points, we cannot afford to lose any data elements by the intializing process. By running Oridnary Least Squares with logarithms of incremental payments as dependent variables and development years for each accident year as explanatory variables, we were able to create development year initial lag values. For the accident year initial lag values we ran OLS on accident years for each development years. The second purpose was to obtain tentative models. We did not attempt to use upper triangle angle only because the model utilize the whole data at once, this will put too much emphasis on the earlier years which contain more data points. This is a major disadvantage of any stochastic model which fits the entire data at once without filling up the lower triangle portion. Even though the development method does not provide confidence intervals, it does provide at least an approximate estimate. It is also consistent with the NAIC model act for the liability discount which explicitly specifies the future payout patterns.

3. Model Selection.

We started with $AR(1,1)$ model for all nine lines we used for this analysis. Estimated coefficients are listed in Table 1. Estimated Q-test on the residuals by accident year and by development years are listed in Table 2. Due to small data points, we only estimated up to four years. Estimated partial autocorrelations on the residuals by accident year and development year are shown in Table 3. The thresholds with 95% confidence level for Q-tests are 7.81 with K=3, 9.49 with K=4, 11.1 with K=5 and 12.6 with K=6, 14.1 with K=7. Most of the cases, Q-tests do not reject the Null Hypothesis that the errors are not white noise. Applying the $\frac{1}{n^{1/2}}$ formula, the thresholds with 95% confidence level for PCAF are 0.653 with n=9, 0.693 with n=8, 0.741 with n=7 and 0.800 with n=6. Except for few cases, there aren't any such cases that reject the whiteness of the errors.

Identifying a model as $AR(1,1)$ is equivalent to saying that the loss history can be explained as a combination of constant trends through accident period and development period. Since the coefficients of all lines are less than 1, we can say that data satisfies the stability condition. This is a desirable condition, otherwise, the estimated variances will be blown up. You may also notice that in every case, the coefficents for the accident year are a lot higher than those of development years. This indicates that the trends through the accident periods are much more important than those through the development years.

You may want to stop here because all the PACF are satisfactory and because the parsimony dictates the fewer the coefficients are, the better the model is. However, since the model with more coefficients will provide more stable forecastings, we tried up to $AR(3,2)$. Except for COMMUL, since the coefficients for development years are already small, we didn't bother to try more development lag coefficients except COMMUL. When we tried $AR(3,2)$ for COMMUL, the second development lag term became very close to the zero. Hence we selected the $AR(3,1)$ for COMMUL. The second lag term indicates that there are more than just straight trend. We may interpret this as a simple cycle. If we require a third lag term, this will indicate that the data contains a complicate cycle.

When we tried $AR(2,1)$ for HOMFAM, suprisingly the second lag term for the accident year became bigger than the first term. Consequently, we tried $AR(3,1)$. Even though the coefficient for the third lag term is still high, we decided to stop here due to the limitation of the data points. We also didn't want those artificially generated initial values to dominate the whole actual data.

For PRVAUT, we tried up to $AR(3,1)$. Since the third lag term of accident years wasn't big enough, we decided to go with $AR(2,1)$. The same was true for PROLIA. For COMAUT as soon as we tried $AR(2,1)$ the second lag became relatively small. Hence, we selected $AR(1,1)$ for COMAUT. The same was true for MEDMAL, SPELIA. For WOKCOM, as soon as we added one more lag term, the first lag term became bigger than 1.0 (which became unstable). Consequently, we chose $AR(1,1)$ for WOKCOM. Finally, for OTHLIA, we chose $AR(3,1)$ as a selected model as HOMFAM. Interestingly, the coefficent of the third lag term was highest. We showed estimated coefficients of the $AR(2,1)$ models, their Q-statistics and PCAFs on the residuals in Table 4, 5 and 6, respectively. Estimated coefficents of the $AR(3,1)$ models, their Q-statistics and PCAFs on the residuals are shown in Table 7, 8 and 9, respectively.

As you may noticed, the process of personal lines like HOMFAM and PRVAUT ar either more complicated or as complicated as comercial lines. Secondly, the longer tail lines like MEDMAL do not necessarily possess a more complicated process.

4. Point Estimates and Confidence Intervals.

After we selected each model based on the rectangular form of data, we eliminated auxiliary observations in the lower triangular area. We filled the lower triangle with forcast values. By adding up row-wise we obtained ultimate loss based on the selected model. Based on the variance formula mentioned on the prior chapter, we estimated each variance for the forecast value.

In Table 10, in the first column, the upper limit of the estimated ultimate loss with 95% probability (one-tail test) are shown. This indicates that if we repeatedly estimate the ultimate loss with different samples, but with same formula, and in each case we construct confidence intervals, then 95% of all the cases of the interval given will inclcude the true parameter. Thus, the probability statement is not about population parameter but estimated parameter.

The distance of the interval is determined by the size of the estimated variance for the error, the complexity of the model and the size of the tail. In the third column the relative distance of the confidence interval in terms of the ultimate loss are provided. In the fifth and seventh column, the upper limit of the estimated future expected liability and its relative distance of the confidence interval are shown, respectively.

If we look at the relative size of the confidence interval in terms of ultimate loss, personal lines' (HOMFAM and PRVAUT) sizes are a lot smaller than commercial lines'. Among the commercial lines, WOKCOM's relative size of the confidence interval is the smallest even though its tail is longer than either COMAUT, COMMUL or SPELIA. The WOKCOM's relative size of the confidence interval may be the smallest because its stability of the exposure growth as well as as its stable payment pattern. SPELIA's relative size of the confidence interval is bigger than either COMAUT or COMMUL or WOKCOM, even though its tail is the shortest among the commercial lines. As we expected, MEDMAL's relative size is biggest among all lines, despite of its simplicity of the model. HOMFAM and SPELIA's relative size of the confidence interval in terms of the future liability are extremely high compared to their size in terms of ultimate loss due to their large estimated variance of the error terms. Other lines' relative size are consistent with their counterparts.

Except for the cases of COMMUL and SPELIA whose estimated constant coefficients' signs are negative, all point estimates based on the models are slightly smaller than those based on the loss development methods. This does not necesarily indicate that model-created estimates are understated. One of the evidences are shown column (9) through column (13). We reserved column (9) of actual paid loss as of 12/91 for the comparison purpose. In column (10), we provided the estimated paid loss as of 12/91 based on the models and in column (11) the projected paid loss as of 12/91 are shown based on the development method. The performances of five lines out of nine lines were better with the models rather than the loss development methods. To the contrary of the ultimate loss comparison cases, where seven out of nine cases, the model estimates were bigger than the actuals. While five out nine cases, the estimates of loss development methods were bigger than the actuals.

One of the main advantages of our model is that it provide future estimates for the future accident years with confidence intervals. Neither ordinary regressional models nor loss development methods provide these estimates, which are valuable for planning purposes. The last rows of column (10) are future accident year estimates and their confidence intervals. Compared to the actual values in column (9), the estimates seem to be reasonable.

By looking at columns (1) through (4), you may notice that every case, the ultimate losses based on the development method has fallen inside of the confidence intervals. This is a small evidence showing that our estimated confidence intervals are reasonable. However, figures on lower rows of the columns (9) and (10) indicate that one out of nine cases, the actual payment located outside the confidence interval with a probability of 97.5%, and two out of nine cases the actual payments laying outside the confidence interval with the probability of 95%. These appear to show that our confidence intervals for the accident year may be too narrow because the actual probabilities indicate that 77.8% and 88.9% instead of the theoretical values of 95% and 97.5%, respectively. This is not the case because the confidence interval with 95% probability means that there is a 95% chance that the interval includes the **true parameter (true mean)** not the actual value. Consequently, the 77.8% and 88.9% regarding the actual values are reasonable considering that the population possesses its own distribution. This is the main reason why the theoretical probability with the normality assumption was larger than the empirical one in Gardner (1988).

In Table 11, the actual cumulative payment triangles, age-to-age factors and ultimate losses based on the loss development methods are shown.

IV. Conclusion

By expanding a Box-Jenkins type time-series model, we developed a generalized framework for modelizing a stochastic process on the loss history data. It turned out that some lines require more complex specifications than the others. We may presume that some lines are more sensitive to the insurance business cycle than the others. Our contributions will be to provide a generalized framework to derive confidence intervals in which the business cycle was taken into account as well as to provide future estimates for the planning process. This paper is the first step to that direction.

We would like to incorporate claim count estimates into our framework by utilizing vector autoregressive model in the near future. We may also incorporate outstanding reserve which is also a valuable information.

BIBLIOGRAPHY

Berquist, J. R. and Sherman, R. E., "Loss Reserve Adequacy Testing: A Comprehensive, Systematic Approach," PCAS LXIV, 1977, p. 123.

Bornheutter, R. L. and Ferguson, R. E., "The Actuary and IBNR," PCAS LXI, 1972 p. 181.

Box, G. E. P. and Jenkins, G. M., *Time Series Analysis: Forecasting and Control,* 1976.

Box, G. E. P. and Pierce, D. A., "Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models" *Journal of the American Statistical Association, 65 332,* p. 1509.

Durbin, J., "The Fitting of Time Series Models," *Review of the International Institute of Statistic, 28, 3,* p. 233.

Gardner, Jr., E. S., "A Simple Method of Computing Prediction Intervals fro Time Series Forecasts," *Management Science, 34, 4,* p. 541.

Maddala, G. S., *Econometrics,* 1977.

Makridakis, S. and Wheelwright, S. C., *Forecasting Methods for Management (Fifth Edition),* 1989.

Quenouille, M. H., "Approximate Tests of Correlation in Time Series," *Journal of the Royal Statistical Society, Series B, 11, 1,* p. 68.

Vandaele, W., *Applied Time Series and Box-Jenkins Models,* 1983.

TABLE 1. ESTIMATED COEFFICENTS FOR AR(1,1) MODEL

	1ST YEAR AY LAG	1ST YEAR DY LAG	CONST
HOMFAM	0.85250	0.13494	0.11621
PRVAUT	0.99250	0.00708	0.11526
COMAUT	0.98074	0.01818	0.09425
COMMUL	0.73432	0.27660	-0.21894
WOKCOM	0.99844	0.00328	0.09810
MEDMAL	0.85550	0.14628	-0.07682
OTHLIA	0.97503	0.02445	0.11304
SPELIA	0.97018	0.02990	0.10406
PROLIA	0.97063	0.03365	0.06065

TABLE 2. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(1,1) MODEL

ACCIDENT YEAR = 82

	K=3	K=4	K=5	K=6	K=7
HOMFAM	2.38778	2.68698	5.43214	6.40956	7.67962
PRVAUT	6.20165	7.43330	8.03333	9.57421	10.27192
COMAUT	8.02664	9.08966	12.78114	22.70667	27.73444
COMMUL	15.59455	18.74024	21.77020	24.20164	26.17824
WOKCOM	17.29664	24.02996	24.85543	32.13953	34.81509
MEDMAL	3.63634	4.52361	9.18822	13.88175	14.61208
OTHLIA	4.38933	6.13802	6.52584	6.80700	6.81674
SPELIA	2.00036	2.33159	3.48908	3.51597	3.51782
PROLIA	10.63477	11.35506	11.47956	11.52169	11.52889

ACCIDENT YEAR = 83

	K=3	K=4	K=5	K=6
HOMFAM	2.54875	2.76390	2.93312	3.76485
PRVAUT	3.19666	4.15370	4.68083	5.11533
COMAUT	5.94915	7.45970	7.67292	23.55856
COMMUL	9.28121	12.03609	16.41462	17.97051
WOKCOM	7.81576	14.92529	16.12265	17.08352
MEDMAL	20.22335	25.45722	30.65844	39.76625
OTHLIA	7.81660	7.94727	10.83099	10.87109
SPELIA	1.58167	2.12018	3.56477	3.96429
PROLIA	9.95443	16.92331	18.41628	21.73013

	ACCIDENT YEAR = 84			ACCIDENT YEAR = 85	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	1.50912	1.84325	2.69574	12.44707	14.18820
PRVAUT	0.90452	1.73380	3.31919	8.57997	8.92221
COMAUT	11.85483	18.02801	19.35910	23.75158	30.68252
COMMUL	19.31421	19.80757	20.31336	15.62485	17.12087
WOKCOM	15.00407	16.46119	16.83647	5.94221	6.27584
MEDMAL	1.52935	2.59451	13.99429	1.81445	2.17930
OTHLIA	7.44905	8.13170	9.67102	12.64123	17.46448
SPELIA	8.21914	10.63992	23.36301	4.13378	4.18345
PROLIA	19.23100	26.05147	33.40982	9.72884	11.05814

TABLE 2. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(1,1) MODEL

DEVELOPMENT YEAR = 1

	K=3	K=4	K=5	K=6	K=7
HOMFAM	20.86283	27.43541	29.97995	39.16037	44.28323
PRVAUT	16.65263	24.27383	31.32747	36.31636	38.33991
COMAUT	10.11426	14.08209	19.90366	35.09818	39.43475
COMMUL	17.38610	26.24465	29.21483	32.85728	36.79327
WOKCOM	13.65747	21.10487	22.18290	24.12949	24.29261
MEDMAL	9.07254	11.45357	12.16451	12.43951	12.53369
OTHLIA	14.13229	17.98698	23.45365	24.57243	28.50565
SPELIA	8.23842	8.89819	9.60571	10.40635	10.46272
PROLIA	10.28675	11.52355	12.65645	14.36246	14.92514

DEVELOPMENT YEAR = 2

	K=3	K=4	K=5	K=6
HOMFAM	15.80416	17.02433	24.92092	34.06265
PRVAUT	14.36262	16.41183	19.37089	24.11920
COMAUT	9.50703	11.75657	14.57927	22.44170
COMMUL	11.90035	15.55383	16.78860	30.58926
WOKCOM	10.04670	18.99859	22.83892	25.65263
MEDMAL	17.35611	22.35855	24.53940	26.06088
OTHLIA	14.20316	15.72022	16.72064	16.99232
SPELIA	24.34332	30.12124	36.38168	38.53166
PROLIA	9.35144	13.16147	13.46168	13.71009

	DEVELOPMET YEAR = 3			DEVELOPMENT YEAR = 4	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	12.64103	13.35973	13.49182	6.16684	7.02828
PRVAUT	11.42169	13.92889	19.69748	13.35642	15.11712
COMAUT	10.18653	12.17216	17.63906	8.03854	10.13738
COMMUL	14.08152	16.70407	17.94427	10.95556	13.88891
WOKCOM	6.13730	7.06503	7.34507	9.18472	9.82891
MEDMAL	5.66534	12.20602	14.21097	5.38781	7.75356
OTHLIA	14.29288	22.40355	27.73785	10.06279	14.94903
SPELIA	18.25537	21.90669	27.88511	6.28131	6.59398
PROLIA	15.05529	17.17875	18.72870	7.20772	8.26060

TABLE 3. PCAF OF THE ESTIMATED RESIDUALS FOR AR(1,1) MODEL

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 82						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 83					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.37076	-0.03665	-0.07018	0.00852	-0.17023	-0.25534	-0.28725	-0.44010	-2.88914	2.02737	0.63195	0.14616
PRVAUT	-0.34487	-0.00364	-0.00112	-0.00223	0.00111	0.00201	-0.37356	-0.01520	-0.00476	-0.01134	-0.00362	0.00860
COMAUT	-0.10285	0.00115	-0.00226	-0.00255	-0.00431	-0.00255	-0.30156	0.00355	-0.00414	-0.01290	-0.00372	-0.02474
COMMUL	0.09514	-0.01340	-0.00179	-0.00324	-0.00126	-0.00043	-0.43576	0.00571	-0.00843	-0.00719	-0.02689	-0.00648
WOKCOM	0.16951	0.00892	-0.00934	-0.00782	-0.00939	-0.00908	0.12489	0.00460	-0.01047	-0.01656	-0.00316	-0.00483
MEDMAL	-0.14126	-0.10140	-0.11312	-0.12456	0.02052	0.00782	0.15254	-0.27062	-0.19031	-0.04040	-0.09085	-0.08738
OTHLIA	0.44427	-0.00304	-0.00352	-0.00171	0.00037	0.00209	0.10988	0.00077	-0.00673	-0.00701	-0.00103	-0.00195
SPELIA	-0.22599	-0.03193	-0.00833	-0.01076	-0.04414	-0.00522	-0.12508	-0.07599	-0.07512	-0.11286	-0.28879	-0.27434
PROLIA	0.25349	-0.01981	-0.00540	0.00476	0.00134	-0.00018	0.03450	-0.01090	-0.01662	-0.05356	-0.00867	0.00463

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 84						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 85					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.18333	-0.02987	-0.04104	-0.05068	0.08708	-0.17562	-0.17577	-0.12642	0.07477	0.01706	-0.04701	-0.02116
PRVAUT	-0.02935	0.00670	-0.00169	-0.00649	-0.00482	-0.01315	0.02938	-0.00062	-0.00214	0.00043	0.00000	-0.00000
COMAUT	-0.47491	-0.00447	0.00382	-0.00309	-0.00181	0.00121	-0.46176	-0.00270	0.00132	-0.00348	0.00029	0.00101
COMMUL	0.25051	-0.01295	-0.01081	-0.00404	-0.00315	-0.00162	0.06822	0.00182	-0.00738	-0.00228	-0.00159	0.00008
WOKCOM	0.36364	-0.01617	-0.00188	-0.02261	-0.04645	-0.00271	-0.02170	-0.05033	-0.11816	-0.05752	-0.21919	-0.03886
MEDMAL	-0.57419	0.01834	-0.03797	-0.03747	-0.01874	-0.00296	-0.20607	-0.00680	-0.00968	-0.01695	0.01045	0.00267
OTHLIA	0.30091	-0.00597	-0.00298	0.00018	0.00095	-0.00036	-0.44420	-0.00020	-0.00140	-0.00011	-0.00003	-0.00016
SPELIA	-0.18716	-0.01487	0.01288	-0.01212	-0.01795	0.00782	-0.46475	-0.01362	-0.00260	-0.00066	0.00161	0.00082
PROLIA	-0.70515	0.00668	-0.02618	-0.00267	-0.00430	-0.01450	0.02055	-0.00099	-0.02622	0.00490	-0.02459	-0.00013

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 1						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 2					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.18377	-0.03396	0.00009	0.01791	-0.00445	-0.01272	-0.57390	0.01386	-0.07921	-0.00814	-0.01905	0.04257
PRVAUT	-0.45164	-0.00008	-0.00018	0.00010	-0.00002	0.00003	0.25633	-0.00027	-0.00065	-0.00004	-0.00105	-0.00041
COMAUT	-0.21997	-0.02538	-0.00589	0.01806	0.00802	-0.01858	-0.11445	0.01353	-0.01414	-0.00924	-0.01754	0.00189
COMMUL	-0.10355	-0.03000	-0.00079	0.00855	-0.00545	-0.00335	-0.46323	-0.01815	-0.08047	0.01340	0.00891	0.05430
WOKCOM	-0.35143	-0.09390	-0.37586	-0.77899	-1.58942	2.35511	0.04069	-0.02385	-0.01581	-0.00568	-0.00437	0.00722
MEDMAL	-0.10960	-0.03756	-0.01395	0.00318	0.02590	-0.00361	-0.06726	-0.03033	-0.08582	0.03129	0.00811	0.01127
OTHLIA	-0.13521	-0.01166	-0.00083	0.00683	-0.00041	-0.00166	-0.03812	-0.00776	-0.00575	0.00938	-0.00156	-0.00708
SPELIA	-0.14748	-0.36557	0.02584	-0.05500	-0.04185	0.03707	-0.30293	-0.00657	0.00221	-0.00923	0.00570	0.00422
PROLIA	-0.46299	-0.22962	0.09621	-0.01529	-0.06378	0.00849	-0.20689	-0.00140	-0.06373	-0.03099	0.02172	0.00733

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 3						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 4					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.49241	-0.05105	0.03597	-0.02773	-0.00072	0.02678	-0.33917	-0.00705	-0.00676	-0.00166	-0.00732	0.01159
PRVAUT	0.12929	0.00019	-0.00200	-0.00348	-0.00112	-0.00004	0.31263	-0.00149	-0.00539	-0.00035	-0.00200	0.00086
COMAUT	-0.12691	-0.00915	-0.00679	-0.00327	0.00723	-0.02852	0.10055	0.00706	-0.02027	-0.00191	-0.02065	-0.00043
COMMUL	-0.20871	-0.01654	-0.00699	0.00686	0.01078	-0.00357	-0.25202	-0.01538	-0.01829	0.00486	0.01590	-0.00168
WOKCOM	0.23198	-0.02239	-0.02952	-0.01093	0.00816	0.01704	0.24741	-0.02784	-0.01471	-0.02061	0.01033	-0.00124
MEDMAL	0.10842	-0.01029	-0.05544	-0.02723	0.02407	-0.00287	0.03956	-0.04484	-0.02243	0.04137	-0.02612	-0.03618
OTHLIA	0.05596	-0.01590	0.00112	0.00798	-0.00592	-0.02045	0.12130	-0.01779	-0.00050	0.00182	0.00055	-0.01890
SPELIA	-0.30055	-0.01253	-0.00252	-0.01689	0.01012	0.00566	-0.20675	-0.03803	-0.01484	0.01032	0.00542	-0.00049
PROLIA	-0.29443	-0.08523	0.03853	0.02532	-0.05277	0.00191	-0.05020	-0.12462	-0.04818	0.02097	0.01188	-0.00370

TABLE 4. ESTIMATED COEFFICENTS FOR AR(2,1) MODEL

	1ST YEAR AY LAG	2ND YEAR AY LAG	1ST YEAR DY LAG	CONST
HOMFAM	0.30030	0.63392	0.06093	0.13195
PRVAUT	0.55930	0.44051	-0.00025	0.17295
COMAUT	0.96540	0.01553	0.01800	0.09608
COMMUL	0.53940	0.20832	0.26344	-0.19422
WOKCOM	1.05840	-0.08517	0.02632	0.09982
MEDMAL	0.94113	0.05838	0.00222	0.10451
OTHLIA	0.52058	0.46175	0.01822	0.16178
SPELIA	0.73300	0.13460	0.13427	-0.06073
PROLIA	0.76355	0.20860	0.03330	0.07551

TABLE 5. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(2,1) MODEL

	ACCIDENT YEAR = 82				
	K=3	K=4	K=5	K=6	K=7
HOMFAM	1.98996	2.16336	3.36277	3.89974	3.99259
PRVAUT	5.67318	6.52377	6.88369	8.01349	8.49348
COMAUT	8.26154	9.46501	13.10957	23.54066	28.98261
COMMUL	15.91527	19.07583	22.34244	24.80556	27.03102
WOKCOM	17.42113	24.40334	25.35317	32.47001	35.49136
MEDMAL	3.48411	4.34488	8.93030	13.57876	14.25841
OTHLIA	4.28978	5.97255	6.34765	6.62104	6.62712
SPELIA	1.91596	2.24384	3.28078	3.30467	3.31895
PROLIA	10.68277	11.42674	11.54250	11.63724	11.65731

	ACCIDENT YEAR = 83			
	K=3	K=4	K=5	K=6
HOMFAM	2.76251	2.95369	3.21404	4.18011
PRVAUT	3.03098	4.00322	4.60045	4.81416
COMAUT	5.97649	7.44506	7.64307	23.31011
COMMUL	9.54224	12.34727	16.56810	17.93802
WOKCOM	7.98863	15.33981	16.52810	17.33435
MEDMAL	20.13986	25.52529	30.47728	39.19085
OTHLIA	8.11471	8.28323	11.36782	11.42579
SPELIA	1.47905	1.97616	3.32353	3.71314
PROLIA	10.47838	17.39252	19.15444	22.12186

	ACCIDENT YEAR = 84			ACCIDENT YEAR = 85	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	2.05900	6.74590	7.66354	5.50849	5.96703
PRVAUT	2.83376	5.38396	7.92933	7.46222	8.40334
COMAUT	10.06888	16.17177	16.83847	24.26775	31.77923
COMMUL	18.55787	19.17945	19.70395	15.43779	16.71569
WOKCOM	15.31477	16.47673	16.96448	6.42555	6.96612
MEDMAL	1.54353	2.58499	13.71380	1.85397	2.21131
OTHLIA	5.51755	6.19861	7.76043	11.90048	13.22891
SPELIA	8.40713	10.13824	23.01675	3.15066	3.23195
PROLIA	18.54066	25.60754	34.80326	11.18560	13.16175

TABLE 5. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(2,1) MODEL

DEVELOPMENT YEAR = 1

	K=3	K=4	K=5	K=6	K=7
HOMFAM	14.63397	19.64255	20.59734	23.40826	24.88186
PRVAUT	11.98079	15.66238	21.10133	24.99560	27.26196
COMAUT	8.08051	11.72902	16.20662	31.76165	36.63307
COMMUL	20.63175	26.85926	30.63580	32.90663	37.20689
WOKCOM	14.93724	22.40663	23.92360	25.82408	25.96615
MEDMAL	8.96038	11.68971	12.57961	12.90309	13.02912
OTHLIA	17.73336	22.70700	27.48417	29.51769	33.14230
SPELIA	18.02518	18.93740	19.53405	20.26905	20.59254
PROLIA	7.84471	9.18015	9.98272	11.64293	11.87074

DEVELOPMENT YEAR = 2

	K=3	K=4	K=5	K=6
HOMFAM	12.67954	15.72676	18.39232	32.04126
PRVAUT	13.09267	16.42352	20.74574	22.55493
COMAUT	7.63526	9.50123	11.62120	18.42212
COMMUL	10.80210	14.60958	14.63837	27.08606
WOKCOM	10.45595	19.07627	23.99037	26.13308
MEDMAL	16.61186	21.45131	23.27465	24.53169
OTHLIA	16.80799	18.60625	19.69520	20.20313
SPELIA	14.71297	15.86508	18.18733	19.24083
PROLIA	9.03563	11.83704	12.33777	12.70747

	DEVELOPMET YEAR = 3			DEVELOPMENT YEAR = 4	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	12.59476	13.41527	13.83793	6.64678	11.21330
PRVAUT	13.30942	18.24313	20.58350	13.25334	13.34350
COMAUT	11.90572	14.96047	20.73482	7.95809	11.02764
COMMUL	14.98182	18.14605	19.30833	8.16434	10.32852
WOKCOM	6.89509	7.86649	8.11614	10.03878	10.80485
MEDMAL	6.65933	13.80642	16.30530	6.24661	8.74575
OTHLIA	15.36807	24.11587	28.18779	10.07870	14.35889
SPELIA	7.65181	9.61350	11.73637	6.15198	6.24645
PROLIA	15.55548	18.15833	20.22100	6.75053	7.67302

TABLE 6. PCAF OF THE ESTIMATED RESIDUALS FOR AR(2,1) MODEL

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 82						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 83					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.29839	-0.03024	-0.08139	-0.02566	-0.14819	-0.20061	-0.31339	-0.21206	-0.59748	-1.99424	1.94932	0.56558
PRVAUT	-0.35057	-0.00372	-0.00093	-0.00169	0.00082	0.00181	-0.38088	-0.00887	-0.00269	-0.00790	-0.00272	0.00433
COMAUT	-0.10371	0.00114	-0.00225	-0.00254	-0.00430	-0.00254	-0.30283	0.00347	-0.00404	-0.01270	-0.00360	-0.02437
COMMUL	0.09390	-0.01189	-0.00157	-0.00268	-0.00151	-0.00045	-0.44897	0.00397	-0.00569	-0.00561	-0.02227	-0.00574
WOKCOM	0.45075	-0.00315	-0.00362	-0.00178	0.00035	0.00214	0.12863	0.00068	-0.00764	-0.00770	-0.00093	-0.00190
MEDMAL	0.16519	0.00912	-0.00928	-0.00782	-0.00957	-0.00932	0.11239	0.00401	-0.00926	-0.01538	-0.00271	-0.00427
OTHLIA	-0.21331	-0.03158	-0.00947	-0.01174	-0.04263	-0.01222	-0.13467	-0.03399	-0.04098	-0.06096	-0.16086	-0.12331
SPELIA	-0.12282	-0.09918	-0.11554	-0.11675	0.02340	0.00631	0.15227	-0.22705	-0.15534	-0.02243	-0.05877	-0.06188
PROLIA	0.25919	-0.01946	-0.00555	0.00416	0.00140	0.00024	-0.01047	-0.00492	-0.01241	-0.04244	-0.00831	0.00169

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 84						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 85					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.005058	-0.006	-0.01433	-0.03493	0.014516	-0.01964	-0.11976	-0.05905	0.040935	0.016009	-0.0523	-0.0090964
PRVAUT	0.0156667	0.003217	-0.00109	-0.00263	-0.00098	-0.00571	0.24675	-0.00133	-0.00292	-0.00016	0.000178	2.210E-06
COMAUT	-0.46568	-0.00452	0.003665	-0.00293	-0.0018	0.001218	-0.45725	-0.00256	0.001196	-0.00344	0.00032	0.0009861
COMMUL	0.2397241	-0.00472	-0.01994	-0.00238	-0.00305	-0.00132	0.195429	-0.00175	-0.00777	-0.00161	-0.00073	0.000206
WOKCOM	0.2756712	-0.00533	-0.00256	0.000319	0.000797	-0.00034	-0.53336	-0.00045	-0.00097	-2.4E-05	0.000084	-0.0002153
MEDMAL	0.3682351	-0.01603	-0.0045	-0.02342	-0.04879	-0.00312	-0.026	-0.04862	-0.11593	-0.05468	-0.20505	-0.0372327
OTHLIA	0.0061007	-0.01844	0.019313	-0.01696	-0.02119	0.003093	-0.41031	-0.00536	-0.00605	-0.00085	0.000995	0.0001095
SPELIA	-0.473081	0.041366	-0.07167	-0.05588	-0.03586	-0.00849	-0.17857	-0.00338	-0.01325	-0.0101	0.007809	0.0014824
PROLIA	-0.601772	0.009743	-0.0404	0.003948	-0.01889	-0.00817	0.094585	-0.00995	-0.00978	0.000682	-0.0181	0.0017349

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 1						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 2					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	0.10441	-0.05122	-0.00100	0.01980	-0.01004	-0.00983	0.09371	-0.01559	-0.06582	-0.01545	-0.00559	0.03722
PRVAUT	-0.12470	-0.00017	-0.00012	0.00008	-0.00002	0.00002	0.44546	-0.00104	-0.00089	-0.00055	-0.00123	-0.00032
COMAUT	-0.21823	-0.02537	-0.00582	0.01807	0.00780	-0.01841	-0.09615	0.01338	-0.01452	-0.00935	-0.01747	0.00200
COMMUL	-0.01338	-0.03375	-0.00120	0.00760	-0.00828	-0.00234	-0.30425	-0.02329	-0.07694	0.01252	0.01284	0.04627
WOKCOM	-0.15666	-0.01164	-0.00086	0.00692	-0.00051	-0.00156	-0.04033	-0.00734	-0.00633	0.00939	-0.00130	-0.00678
MEDMAL	-0.30151	-0.08300	-0.37307	-0.68592	-0.84343	-4.48305	0.06494	-0.02510	-0.01789	-0.00599	-0.00269	0.00850
OTHLIA	0.00015	-0.39626	0.05815	-0.11715	-0.05436	0.01327	-0.11245	-0.00791	-0.00253	-0.00697	0.00876	0.00234
SPELIA	-0.06512	-0.04156	-0.01475	0.00451	0.02518	-0.00397	-0.01628	-0.04230	-0.08082	0.03488	0.00849	0.01003
PROLIA	-0.38248	-0.18376	0.07492	-0.03374	-0.05593	0.01576	-0.00534	-0.00816	-0.06943	-0.02396	0.01872	0.00306

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 3						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 4					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.18953	-0.04044	0.00878	-0.03115	0.00606	0.02137	0.17352	-0.00906	-0.01119	-0.00946	-0.00037	0.00912
PRVAUT	0.44993	-0.00274	-0.00413	-0.00426	-0.00055	0.00079	0.53043	-0.00635	-0.00715	-0.00195	-0.00060	0.00190
COMAUT	-0.11604	-0.00928	-0.00675	-0.00334	0.00681	-0.02858	0.11468	0.00666	-0.02046	-0.00220	-0.02066	-0.00040
COMMUL	-0.12782	-0.01971	-0.00771	0.00885	0.01224	-0.00631	-0.11636	-0.01955	-0.02103	0.00670	0.01472	-0.00424
WOKCOM	0.02123	-0.01555	0.00073	0.00846	-0.00461	-0.02015	0.08589	-0.01757	-0.00009	0.00203	0.00214	-0.01867
MEDMAL	0.25400	-0.02553	-0.03058	-0.01050	0.00958	0.01747	0.26950	-0.03012	-0.01741	-0.01940	0.01161	-0.00188
OTHLIA	0.04945	-0.01640	-0.01122	-0.01399	0.01147	0.00525	0.04172	-0.05498	-0.01581	0.01260	0.00645	-0.00000
SPELIA	0.17825	-0.01958	-0.06156	-0.02226	0.02695	0.00058	0.08795	-0.05177	-0.01969	0.04080	-0.03065	-0.04013
PROLIA	-0.20953	-0.07372	0.03710	0.01521	-0.05075	-0.00047	0.05577	-0.13833	-0.04611	0.01823	0.00958	-0.00575

TABLE 7. ESTIMATED COEFFICENTS FOR AR(3,1) MODEL

	1ST YEAR AY LAG	2ND YEAR AY LAG	3RD YEAR AY LAG	1ST YEAR DY LAG	CONST
HOMFAM	0.02596	0.47760	0.44232	0.05052	0.17460
PRVAUT	0.52211	0.39606	0.08301	-0.00161	0.18837
COMAUT	0.96374	-0.03759	0.05602	0.01672	0.10371
COMMUL	0.57237	-0.15216	0.35489	0.23524	-0.14156
WOKCOM	1.04169	-0.72885	0.69056	-0.00487	0.23127
MEDMAL	0.94271	0.06672	-0.01021	0.00256	0.10270
OTHLIA	0.32960	0.24380	0.41686	0.01021	0.24194
SPELIA	0.67767	-0.16442	0.39012	0.09871	-0.00733
PROLIA	0.69942	-0.20058	0.47181	0.03626	0.11847

TABLE 8. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(3,1) MODEL

ACCIDENT YEAR = 82

	K=3	K=4	K=5	K=6	K=7
HOMFAM	1.89961	2.04500	2.97544	3.43541	3.46336
PRVAUT	5.56150	6.32262	6.64156	7.64102	8.05960
COMAUT	8.14380	9.29051	12.72664	22.87711	28.23228
COMMUL	16.53650	19.75038	23.49269	25.95561	28.53064
WOKCOM	19.11617	24.39896	25.27048	32.04011	33.54338
MEDMAL	3.49878	4.36072	8.94902	13.59271	14.28146
OTHLIA	4.18413	5.66641	6.06382	6.31699	6.32329
SPELIA	1.80286	2.09206	2.94262	3.02819	3.02980
PROLIA	10.78075	11.56971	11.72503	11.82286	11.84372

ACCIDENT YEAR = 83

	K=3	K=4	K=5	K=6
HOMFAM	3.00390	3.19140	3.47231	4.55810
PRVAUT	3.02413	4.02694	4.63506	4.84792
COMAUT	6.02254	7.46375	7.66840	23.16026
COMMUL	7.74866	10.01261	13.87038	15.45627
WOKCOM	8.21531	15.54758	16.70029	16.93862
MEDMAL	20.14428	25.51812	30.48372	39.23301
OTHLIA	7.27898	7.38832	10.29606	10.34080
SPELIA	1.46445	1.97280	3.36948	3.68359
PROLIA	10.85263	17.76777	19.67225	22.63624

	ACCIDENT YEAR = 84			ACCIDENT YEAR = 85	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	3.47006	11.40593	12.13779	2.39762	3.26036
PRVAUT	2.77704	5.20330	7.77981	7.33106	8.51191
COMAUT	9.32228	15.07253	15.79316	21.97061	30.34663
COMMUL	18.38219	18.77972	19.25388	9.07117	10.22135
WOKCOM	7.16027	7.49055	7.90569	4.38527	5.29524
MEDMAL	1.54678	2.59282	13.74100	1.85422	2.21259
OTHLIA	7.29723	7.87348	9.74389	10.06132	10.63347
SPELIA	10.78656	13.38671	25.73312	6.81150	7.12185
PROLIA	18.87935	26.42670	34.61267	6.16343	6.58304

TABLE 8. ESTIMATED Q-STATISTICS OF THE RESIDUALS FOR AR(3,1) MODEL

DEVELOPMENT YEAR = 1

	K=3	K=4	K=5	K=6	K=7
HOMFAM	17.82687	21.04413	23.07158	24.98393	27.15608
PRVAUT	11.03701	14.39888	19.69780	23.07588	24.94225
COMAUT	6.57248	8.72126	13.03380	28.91539	35.40280
COMMUL	20.03356	25.71469	31.46851	33.42196	39.33703
WOKCOM	8.08914	9.10014	9.93583	11.51103	11.76411
MEDMAL	8.74491	11.48403	12.35361	12.67248	12.79704
OTHLIA	10.38935	14.33840	17.31665	17.82988	18.07093
SPELIA	18.52091	19.44997	20.03453	20.46299	20.93029
PROLIA	15.27199	19.91718	21.19592	23.37152	24.59059

DEVELOPMENT YEAR = 2

	K=3	K=4	K=5	K=6
HOMFAM	10.88493	11.93953	12.21178	20.84214
PRVAUT	13.52875	17.50895	21.48084	23.50739
COMAUT	7.98087	9.05037	11.05467	17.58706
COMMUL	11.90663	15.01756	16.33284	29.18758
WOKCOM	8.21686	14.10539	24.69604	28.53668
MEDMAL	16.56766	21.43637	23.26615	24.55222
OTHLIA	14.50624	15.88895	17.12689	17.76154
SPELIA	10.22151	11.52239	14.49992	16.72538
PROLIA	8.03753	9.61459	10.63513	11.23291

	DEVELOPMET YEAR = 3			DEVELOPMENT YEAR = 4	
	K=3	K=4	K=5	K=3	K=4
HOMFAM	9.64579	10.36975	10.58742	11.04662	18.18216
PRVAUT	13.42957	18.21212	20.00778	12.42276	12.60368
COMAUT	11.80267	15.14643	21.79389	6.32948	8.06215
COMMUL	16.69330	17.66878	20.20201	12.39557	15.46746
WOKCOM	12.18456	19.46557	22.84906	5.10041	10.99673
MEDMAL	6.65049	13.76803	16.19655	6.22484	8.72807
OTHLIA	15.37184	24.15344	26.06879	8.66626	10.23055
SPELIA	14.85186	18.34796	23.94580	4.27263	4.47297
PROLIA	15.87020	18.06519	19.86379	5.08702	5.69904

TABLE 9. PCAF OF THE ESTIMATED RESIDUALS FOR AR(3,1) MODEL

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 82						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 83					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	-0.26716	-0.05608	-0.09823	-0.03268	-0.15005	-0.19583	-0.34105	-0.14883	-0.28184	-0.54963	-1.00750	165.03700
PRVAUT	-0.35392	-0.00374	-0.00087	-0.00155	0.00078	0.00171	-0.37978	-0.00842	-0.00266	-0.00779	-0.00259	0.00420
COMAUT	-0.10556	0.00113	-0.00222	-0.00252	-0.00424	-0.00254	-0.30055	0.00354	-0.00412	-0.01273	-0.00370	-0.02435
COMMUL	0.07346	-0.00944	-0.00113	-0.00155	-0.00185	-0.00011	-0.43714	0.00220	-0.00590	-0.00454	-0.02300	-0.00490
WOKCOM	0.42938	-0.00243	-0.00288	-0.00118	0.00076	0.00140	0.21134	-0.00119	-0.00924	-0.00818	-0.00126	-0.00064
MEDMAL	0.16511	0.00910	-0.00928	-0.00781	-0.00955	-0.00930	0.11287	0.00404	-0.00929	-0.01541	-0.00274	-0.00430
OTHLIA	-0.20786	-0.03110	-0.01044	-0.01123	-0.03803	-0.01716	-0.14054	-0.02174	-0.03029	-0.04656	-0.12731	-0.08358
SPELIA	-0.06205	-0.09696	-0.12202	-0.09324	0.03265	0.00158	0.14931	-0.20989	-0.13631	-0.02612	-0.06189	-0.04440
PROLIA	0.26115	-0.01899	-0.00559	0.00428	0.00163	0.00018	-0.02798	-0.00369	-0.01099	-0.04032	-0.00852	0.00083

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 84						RESIDUAL PARTIAL AUTOCORRELATIONS FOR AY 85					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	0.07795	-0.00494	-0.01208	-0.02316	0.00795	-0.00433	-0.17042	-0.03758	0.02303	0.01861	-0.04912	-0.01055
PRVAUT	0.01303	0.00306	-0.00103	-0.00244	-0.00097	-0.00552	0.26761	-0.00172	-0.00316	-0.00024	0.00028	-0.00001
COMAUT	-0.48487	-0.00436	0.00393	-0.00318	-0.00176	0.00115	-0.42019	-0.00234	0.00080	-0.00406	0.00055	0.00102
COMMUL	0.25694	-0.01070	-0.00853	-0.00239	-0.00244	-0.00148	-0.07603	0.00781	-0.00963	-0.00349	-0.00504	-0.00044
WOKCOM	0.01346	-0.00281	-0.00141	0.00052	0.00050	0.00004	-0.10101	0.00206	-0.00480	-0.00246	-0.00149	-0.00015
MEDMAL	0.36923	-0.01605	-0.00473	-0.02360	-0.04905	-0.00309	-0.02576	-0.04798	-0.11721	-0.05509	-0.20639	-0.03768
OTHLIA	-0.02931	-0.01409	0.01604	-0.01398	-0.01551	0.00166	-0.04590	-0.00408	-0.01476	0.00086	-0.00257	-0.00033
SPELIA	-0.58797	0.01067	-0.02147	-0.02795	-0.01883	-0.00098	0.06009	0.04129	-0.05578	-0.01232	-0.03542	0.00085
PROLIA	-0.71820	0.00228	-0.01896	0.00097	-0.00332	-0.01319	-0.01319	-0.00640	-0.02216	0.00609	-0.00750	-0.00051

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 1						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 2					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	0.23170	-0.05298	-0.00917	0.01944	-0.00427	-0.01357	0.10000	-0.01616	-0.10801	-0.00389	-0.00775	0.03960
PRVAUT	-0.14311	-0.00015	-0.00013	0.00008	-0.00002	0.00002	0.36253	-0.00109	-0.00092	-0.00055	-0.00025	0.00051
COMAUT	-0.23192	-0.02294	-0.00552	0.01620	0.00802	-0.01801	-0.18645	0.00064	-0.00954	-0.00842	-0.01235	0.00811
COMMUL	0.02573	-0.01432	-0.00745	-0.00118	-0.00065	-0.00504	-0.40701	0.01556	-0.08795	0.00244	-0.00734	0.04367
WOKCOM	-0.29811	-0.09055	-0.37332	-0.68566	-0.84675	-4.64384	0.14061	-0.03839	-0.01055	0.01633	0.00467	-0.00571
MEDMAL	-0.12728	-0.02112	-0.01534	-0.00478	0.01630	-0.00126	-0.00312	0.01149	-0.04696	0.00255	-0.01053	0.00084
OTHLIA	-0.50265	-0.00245	-0.00075	0.00121	-0.00105	0.00052	-0.51547	-0.00022	-0.00081	0.00116	-0.00069	0.00018
SPELIA	0.23052	-0.25837	-0.00921	-0.05344	-0.06884	0.01160	-0.11998	-0.00057	-0.00311	-0.00269	0.00245	-0.00151
PROLIA	-0.38036	-0.05829	0.08004	-0.07761	-0.06791	0.01103	-0.02004	-0.00532	-0.06729	-0.04119	0.01474	0.00795

	RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 3						RESIDUAL PARTIAL AUTOCORRELATIONS FOR DY 4					
	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG	1ST LAG	2ND LAG	3RD LAG	4TH LAG	5TH LAG	6TH LAG
HOMFAM	0.10623	-0.02944	-0.00545	-0.00213	0.00105	0.00075	0.49432	-0.00582	-0.00624	-0.00158	0.00199	0.00023
PRVAUT	0.25605	-0.00344	-0.00390	-0.00250	0.00136	0.00258	0.27579	-0.00612	-0.00425	0.00164	0.00180	-0.00055
COMAUT	-0.26974	-0.01250	-0.01730	-0.00502	0.01743	-0.00279	-0.05468	-0.00942	-0.01159	0.00530	-0.00062	-0.00006
COMMUL	-0.02071	-0.00495	-0.00786	-0.00020	0.00301	0.00026	0.19221	-0.01643	-0.00646	0.00501	0.00242	-0.00062
WOKCOM	0.19224	-0.00518	-0.01352	-0.00875	0.00337	0.00103	0.01502	-0.01277	0.00102	-0.01046	-0.00261	-0.00022
MEDMAL	0.17799	-0.01470	-0.02706	-0.00848	0.01132	-0.00330	0.04140	-0.02958	-0.02602	0.01337	-0.00174	0.00002
OTHLIA	-0.40906	-0.00249	0.00018	-0.00629	0.00166	0.00011	-0.19410	-0.00397	-0.00276	-0.01209	0.00510	0.00139
SPELIA	0.20983	-0.01628	-0.01917	-0.01090	0.01290	0.00292	0.01638	-0.03931	-0.01341	0.00531	0.00591	0.00015
PROLIA	-0.42933	-0.01915	0.01781	-0.00325	-0.01107	-0.00225	0.21993	-0.02346	-0.04364	-0.00565	-0.01262	0.00146

Table 10. HOMFAM Comparison of Estimates

Accident Year	(1) 95 % Limit	(2) Point Estimate	(3) (1-2)/(2)	(4) Loss Dev Method	(5) 95 % Limit	(6) Point Estimate	(7) (5-6)/(6)	(8) LDF Method	(9) Actual Paid L @12/91	(10) Model @12/91	(11) Loss Dev Method @12/91	(12) (10)-(9)	(13) (11)-(9)
	Ultimate Loss Comparison				Liability Comparison								
1982	8,227,483	8,222,584	0.06%	8,222,506	10,775	5,876	83.37%	5,798	8,224,257	8,222,584	8,222,506	(1,673)	(1,751)
1983	8,894,303	8,883,211	0.12%	8,884,462	28,618	17,526	63.29%	18,777	8,883,252	8,877,052	8,878,197	(6,200)	(5,055)
1984	9,223,736	9,198,101	0.28%	9,195,274	70,010	44,375	57.77%	41,548	9,183,429	9,178,764	9,175,840	(4,665)	(7,589)
1985	10,440,020	10,376,815	0.61%	10,299,264	179,966	116,761	54.13%	39,210	10,314,312	10,344,095	10,252,727	29,783	(61,585)
1986	9,756,424	9,631,000	1.30%	9,597,963	354,433	229,009	54.77%	195,972	9,497,598	9,515,095	9,477,972	17,497	(19,626)
1987	10,259,092	10,038,562	2.20%	10,008,421	618,497	397,967	55.41%	367,826	9,789,919	9,827,809	9,804,068	37,890	14,149
1988	11,486,361	11,100,605	3.48%	11,098,940	1,076,049	690,293	55.88%	688,628	10,656,496	10,699,876	10,691,036	43,380	34,540
1989	14,651,688	13,968,085	4.89%	14,199,606	1,906,785	1,223,182	55.89%	1,454,703	13,254,760	13,272,218	13,318,598	17,458	63,838
1990	15,710,658	13,473,811	16.60%	13,819,411	6,473,740	4,236,893	52.79%	4,582,493	12,358,709	12,249,744	12,403,657	(108,965)	44,948
Total	98,649,765	94,892,774	3.96%	95,325,847	10,718,873	6,961,882	53.97%	7,394,955	92,162,732	92,187,239	92,224,603	24,507	61,871
1991									10,670,718	9,411,233			
Upper Limit with 97.5 % Two-Tail Test										4,746,733			
Lower Limit with 97.5 % Two-Tail Test										14,075,732			
Upper Limit with 95 % Two-Tail Test										5,944,205			
Lower Limit with 95 % Two-Tail Test										12,878,260			

Table 10. PRVAUT Comparison of Estimates

Accident Year	(1) 95 % Limit	(2) Ultimate Loss Comparison: Point Estimate	(3) (1-2)/(2)	(4) Loss Dev Method	(5) 95 % Limit	(6) Liability Comparison: Point Estimate	(7) (5-6)/(6)	(8) LDF Method	(9) Actual Paid L @12/91	(10) Model @12/91	(11) Loss Dev Method @12/91	(12) (10)-(9)	(13) (11)-(9)
1982	15,782,753	15,777,808	0.03%	15,776,929	40,393	35,448	13.95%	34,569	15,779,034	15,765,978	15,765,395	(13,056)	(13,639)
1983	17,927,403	17,917,921	0.05%	17,921,001	97,846	88,364	10.73%	91,444	17,901,737	17,881,747	17,881,735	(19,990)	(20,002)
1984	20,670,924	20,653,401	0.08%	20,672,629	198,108	180,585	9.70%	199,813	20,622,934	20,564,722	20,567,144	(58,212)	(55,790)
1985	23,488,419	23,449,125	0.17%	23,508,711	475,428	436,134	9.01%	495,720	23,320,319	23,264,891	23,281,485	(55,428)	(38,834)
1986	26,412,360	26,317,875	0.36%	26,419,114	1,178,061	1,083,576	8.72%	1,184,815	25,881,852	25,866,542	25,862,024	(15,310)	(19,828)
1987	29,571,320	29,353,848	0.74%	29,531,112	2,737,075	2,519,603	8.63%	2,696,867	28,250,991	28,264,057	28,206,733	13,066	(44,258)
1988	33,027,267	32,549,404	1.47%	32,925,117	6,020,076	5,542,213	8.62%	5,917,926	29,844,056	30,007,063	29,918,300	163,007	74,244
1989	36,606,510	35,613,939	2.79%	36,497,086	12,477,209	11,484,638	8.64%	12,367,785	29,852,941	30,043,508	29,937,138	190,567	84,197
1990	40,625,515	38,461,978	5.63%	40,181,987	27,283,305	25,119,768	8.61%	26,839,777	26,102,083	26,936,751	26,565,498	834,668	463,415
Total	244,112,471	240,095,299	1.67%	243,433,686	50,507,501	46,490,329	8.64%	49,828,716	217,555,947	218,595,259	217,985,451	1,039,312	429,504
1991									13,340,803	14,876,242			

Upper Limit with 97.5 % Two-Tail Test	16,270,389
Lower Limit with 97.5 % Two-Tail Test	13,482,096
Upper Limit with 95 % Two-Tail Test	15,994,073
Lower Limit with 95 % Two-Tail Test	13,758,411

Table 10. COMAUT Comparison of Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Ultimate Loss Comparison				Liability Comparison				Actual		Loss Dev		
Accident	95 %	Point		Loss Dev	95 %	Point		LDF	Paid L	Model	Method	(10)-(9)	(11)-(9)
Year	Limit	Estimate	(1-2)/(2)	Method	Limit	Estimate	(5-6)/(6)	Method	@12/91	@12/91	@12/91		
1982	4,105,218	4,092,216	0.32%	4,058,434	76,713	63,711	20.41%	29,929	4,042,160	4,062,119	4,044,430	19,959	2,270
1983	4,666,126	4,643,961	0.48%	4,615,709	120,396	98,231	22.57%	69,979	4,577,032	4,580,530	4,581,670	3,498	4,638
1984	5,713,126	5,673,248	0.70%	5,673,773	210,465	170,587	23.38%	171,112	5,583,276	5,575,213	5,587,753	(8,063)	4,477
1985	6,606,130	6,524,735	1.25%	6,557,468	432,697	351,302	23.17%	384,035	6,360,828	6,353,997	6,359,705	(6,831)	(1,123)
1986	7,325,185	7,161,108	2.29%	7,235,420	870,487	706,410	23.23%	780,722	6,839,937	6,809,762	6,811,681	(30,175)	(28,256)
1987	8,188,251	7,850,104	4.31%	7,933,205	1,797,660	1,459,513	23.17%	1,542,614	7,085,223	7,143,635	7,077,190	58,412	(8,033)
1988	8,982,215	8,334,791	7.77%	8,427,419	3,401,045	2,753,621	23.51%	2,846,249	6,815,728	6,878,664	6,788,705	62,936	(27,023)
1989	10,081,724	8,955,483	12.58%	9,280,319	5,787,786	4,661,545	24.16%	4,986,381	6,220,537	6,215,497	6,146,015	(5,040)	(74,522)
1990	10,817,614	9,015,129	19.99%	9,205,528	9,022,573	7,220,088	24.96%	7,410,487	4,195,956	4,386,378	4,259,333	190,422	63,377
Total	66,485,590	62,250,775	6.80%	62,987,274	21,719,823	17,485,008	24.22%	18,221,507	51,720,677	52,005,795	51,656,481	285,118	(64,196)
1991									1,704,288	1,997,109			
Upper Limit with 97.5 % Two-Tail Test										2,383,303			
Lower Limit with 97.5 % Two-Tail Test										1,610,914			
Upper Limit with 95 % Two-Tail Test										2,292,035			
Lower Limit with 95 % Two-Tail Test										1,702,182			

Table 10. COMMUL Comparison of Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
		Ultimate Loss Comparison				Liability Comparison			Actual		Loss Dev		
Accident Year	95 % Limit	Point Estimate	(1-2)/(2)	Loss Dev Method	95 % Limit	Point Estimate	(5-6)/(6)	LDF Method	Paid L @12/91	Model @12/91	Method @12/91	(10)-(9)	(11)-(9)
1982	5,437,398	5,422,912	0.27%	5,417,230	85,087	70,601	20.52%	64,919	5,381,291	5,389,061	5,386,804	7,770	5,513
1983	6,354,958	6,321,652	0.53%	6,316,166	193,894	160,588	20.74%	155,102	6,206,690	6,234,573	6,240,475	27,883	33,785
1984	7,305,313	7,236,856	0.95%	7,225,004	395,729	327,272	20.92%	315,420	7,053,579	7,044,965	7,047,584	(8,614)	(5,995)
1985	7,999,620	7,864,431	1.72%	7,832,537	777,351	642,162	21.05%	610,268	7,492,393	7,479,635	7,490,595	(12,758)	(1,798)
1986	7,681,575	7,434,025	3.33%	7,200,161	1,413,268	1,165,718	21.24%	931,854	6,660,445	6,681,956	6,639,164	21,511	(21,281)
1987	8,505,365	8,078,574	5.28%	7,634,479	2,415,107	1,988,316	21.46%	1,544,221	6,715,892	6,692,439	6,646,415	(23,453)	(69,477)
1988	9,909,739	9,220,743	7.47%	8,619,542	3,864,267	3,175,271	21.70%	2,574,070	6,914,450	6,884,622	6,876,073	(29,828)	(38,377)
1989	12,567,415	11,485,820	9.42%	11,191,586	6,031,693	4,950,098	21.85%	4,655,864	7,763,973	7,800,080	7,849,421	36,107	85,448
1990	14,158,039	12,282,635	15.27%	10,497,573	10,517,534	8,642,130	21.70%	6,857,068	6,133,380	6,400,062	6,130,429	266,682	(2,951)
Total	79,919,422	75,347,647	6.07%	71,934,279	25,693,930	21,122,155	21.64%	17,708,787	60,322,093	60,607,392	60,306,960	285,299	(15,133)
1991									3,906,165	4,080,413			
Upper Limit with 97.5 % Two-Tail Test										4,860,506			
Lower Limit with 97.5 % Two-Tail Test										3,300,321			
Upper Limit with 95 % Two-Tail Test										4,676,277			
Lower Limit with 95 % Two-Tail Test										3,484,550			

Table 10. WOKCOM Comparison of Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Ultimate Loss Comparison				Liability Comparison				Actual		Loss Dev		
Accident	95 %	Point		Loss Dev	95 %	Point		LDF	Paid L	Model	Method	(10)-(9)	(11)-(9)
Year	Limit	Estimate	(1-2)/(2)	Method	Limit	Estimate	(5-6)/(6)	Method	@12/91	@12/91	@12/91		
1982	9,213,514	9,146,789	0.73%	8,942,805	466,195	399,470	16.70%	195,486	8,893,778	8,949,043	8,847,496	55,265	(46,282)
1983	10,598,467	10,486,982	1.06%	10,317,945	732,055	620,570	17.96%	451,533	10,059,841	10,086,206	10,092,399	26,365	32,558
1984	13,069,409	12,893,512	1.36%	12,879,912	1,110,585	934,688	18.82%	921,088	12,296,335	12,271,112	12,316,262	(25,223)	19,927
1985	14,643,669	14,365,071	1.94%	14,450,883	1,728,692	1,450,094	19.21%	1,535,906	13,439,155	13,426,764	13,417,449	(12,391)	(21,706)
1986	16,006,922	15,570,589	2.80%	15,752,839	2,676,469	2,240,136	19.48%	2,422,386	14,105,048	14,114,942	14,078,555	9,894	(26,493)
1987	18,214,288	17,527,814	3.92%	18,033,056	4,191,669	3,505,195	19.58%	4,010,437	15,266,334	15,278,500	15,260,031	12,166	(6,303)
1988	21,159,960	20,044,447	5.57%	21,345,500	6,850,868	5,735,355	19.45%	7,036,408	16,587,748	16,521,593	16,598,396	(66,155)	10,648
1989	23,809,901	21,896,318	8.74%	23,820,266	11,928,420	10,014,837	19.11%	11,938,785	16,069,736	16,124,360	15,968,068	54,624	(101,668)
1990	26,395,660	23,044,713	14.54%	24,455,565	21,095,488	17,744,541	18.88%	19,155,393	12,900,611	12,964,768	12,198,366	64,157	(702,245)
Total	153,111,789	144,976,235	5.61%	149,998,771	50,780,440	42,644,886	19.08%	47,667,422	119,618,586	119,737,287	118,777,022	118,701	(841,564)
1991									5,488,466	6,046,709			
Upper Limit with 97.5 % Two-Tail Test										6,947,791			
Lower Limit with 97.5 % Two-Tail Test										5,145,628			
Upper Limit with 95 % Two-Tail Test										6,737,861			
Lower Limit with 95 % Two-Tail Test										5,355,558			

Table 10. MEDMAL Comparison of Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Ultimate Loss Comparison				Liability Comparison				Actual		Loss Dev		
Accident Year	95 % Limit	Point Estimate	(1-2)/(2)	Loss Dev Method	95 % Limit	Point Estimate	(5-6)/(6)	LDF Method	Paid L @12/91	Model @12/91	Method @12/91	(10)-(9)	(11)-(9)
1982	1,996,508	1,873,257	6.58%	1,755,479	365,564	242,313	50.86%	124,535	1,706,116	1,747,222	1,692,177	41,106	(13,939)
1983	2,350,126	2,147,099	9.46%	2,049,968	578,607	375,580	54.06%	278,449	1,898,418	1,899,876	1,904,541	1,458	6,123
1984	2,730,133	2,408,021	13.38%	2,336,516	894,527	572,415	56.27%	500,910	2,000,148	2,024,658	2,019,146	24,510	18,998
1985	3,062,136	2,573,302	19.00%	2,537,082	1,325,694	836,860	58.41%	800,640	1,923,757	1,988,787	1,993,174	65,030	69,417
1986	3,323,696	2,589,890	28.33%	2,460,406	1,961,478	1,227,672	59.77%	1,098,188	1,621,187	1,734,912	1,683,963	113,725	62,776
1987	3,719,757	2,685,264	38.52%	2,568,473	2,690,760	1,656,267	62.46%	1,539,476	1,347,593	1,430,709	1,422,050	83,116	74,457
1988	4,270,137	2,870,635	48.75%	2,991,231	3,536,882	2,137,380	65.48%	2,257,976	1,091,623	1,177,769	1,198,365	86,146	106,742
1989	4,843,472	3,026,259	60.05%	3,665,678	4,447,701	2,630,488	69.08%	3,269,907	852,508	841,429	898,586	(11,079)	46,078
1990	5,383,128	3,132,196	71.86%	4,744,960	5,295,116	3,044,184	73.94%	4,656,948	444,715	443,456	512,297	(1,259)	67,582
Total	31,679,093	23,305,923	35.93%	25,109,792	21,096,329	12,723,159	65.81%	14,527,028	12,886,065	13,288,818	13,324,299	402,753	438,234
1991										97,729	99,978		
Upper Limit with 97.5 % Two-Tail Test											140,418		
Lower Limit with 97.5 % Two-Tail Test											59,538		
Upper Limit with 95 % Two-Tail Test											130,268		
Lower Limit with 95 % Two-Tail Test											69,687		

Table 10. OTHLIA Comparison of Estimates

	Ultimate Loss Comparison				Liability Comparison								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Accident Year	95 % Limit	Point Estimate	(1-2)/(2)	Loss Dev Method	95 % Limit	Point Estimate	(5-6)/(6)	LDF Method	Actual Paid L @12/91	Model @12/91	Loss Dev Method @12/91	(10)-(9)	(11)-(9)
1982	4,604,844	4,551,164	1.18%	4,546,694	236,338	182,658	29.39%	178,188	4,494,388	4,460,185	4,457,828	(34,203)	(36,560)
1983	5,082,057	4,968,894	2.28%	4,966,155	497,989	384,826	29.41%	382,087	5,077,919	4,778,656	4,771,528	(299,263)	(306,391)
1984	6,300,135	6,125,941	2.84%	6,261,412	757,147	582,953	29.88%	718,424	5,952,007	5,726,084	5,779,671	(225,923)	(172,336)
1985	7,456,156	7,097,853	5.05%	7,416,525	1,565,652	1,207,349	29.68%	1,526,021	6,568,768	6,486,247	6,565,565	(82,521)	(3,203)
1986	7,930,839	7,287,319	8.83%	7,630,093	2,821,185	2,177,665	29.55%	2,520,439	5,983,973	6,083,384	6,060,128	99,411	76,155
1987	8,889,403	7,803,392	13.92%	7,944,806	4,721,624	3,635,613	29.87%	3,777,027	5,317,321	5,473,511	5,320,409	156,190	3,088
1988	10,460,042	8,817,891	18.62%	9,536,151	7,084,123	5,441,972	30.18%	6,160,232	4,917,109	5,124,972	5,002,585	207,863	85,476
1989	11,476,012	9,278,585	23.68%	10,978,423	9,391,578	7,194,151	30.54%	8,893,989	3,770,531	3,868,661	3,886,502	98,130	115,971
1990	12,226,490	9,507,210	28.60%	12,404,269	11,480,261	8,760,981	31.04%	11,658,040	2,170,377	2,248,189	2,355,154	77,812	184,777
Total	74,425,977	65,438,249	13.73%	71,684,526	38,555,896	29,568,168	30.40%	35,814,445	44,252,393	44,249,889	44,199,370	(2,504)	(53,023)
1991									745,429	960,584			
Upper Limit with 97.5 % Two-Tail Test										1,231,186			
Lower Limit with 97.5 % Two-Tail Test										689,981			
Upper Limit with 95 % Two-Tail Test										1,165,511			
Lower Limit with 95 % Two-Tail Test										755,657			

Table 10. SPELIA Comparison of Estimates

Accident Year	(1) 95 % Limit	(2) Ultimate Loss Comparison: Point Estimate	(3) (1-2)/(2)	(4) Loss Dev Method	(5) 95 % Limit	(6) Liability Comparison: Point Estimate	(7) (5-6)/(6)	(8) LDF Method	(9) Actual Paid L @12/91	(10) Model @12/91	(11) Loss Dev Method @12/91	(12) (10)-(9)	(13) (11)-(9)
1982	1,129,276	1,126,763	0.22%	1,124,166	7,876	5,363	46.87%	2,766	1,124,673	1,124,848	1,123,243	175	(1,430)
1983	1,279,849	1,274,976	0.38%	1,272,515	14,624	9,751	49.97%	7,290	1,273,497	1,269,268	1,269,384	(4,229)	(4,113)
1984	1,370,003	1,358,845	0.82%	1,357,735	33,498	22,340	49.95%	21,230	1,355,884	1,347,930	1,349,957	(7,954)	(5,927)
1985	1,382,886	1,357,925	1.84%	1,354,815	74,656	49,695	50.23%	46,585	1,327,123	1,332,671	1,333,631	5,548	6,508
1986	1,381,856	1,332,072	3.74%	1,313,246	147,088	97,304	51.16%	78,478	1,283,582	1,276,620	1,268,090	(6,962)	(15,492)
1987	1,580,592	1,490,415	6.05%	1,469,089	262,730	172,553	52.26%	151,227	1,393,829	1,383,098	1,381,298	(10,731)	(12,531)
1988	1,883,734	1,714,565	9.87%	1,699,189	492,560	323,391	52.31%	308,015	1,535,560	1,522,262	1,524,276	(13,298)	(11,284)
1989	2,200,011	1,865,627	17.92%	1,821,966	978,618	644,234	51.90%	600,573	1,479,785	1,501,293	1,491,695	21,508	11,910
1990	2,521,771	1,830,600	37.76%	1,538,000	2,034,359	1,343,188	51.46%	1,050,588	1,102,659	1,098,583	1,031,030	(4,076)	(71,629)
Total	14,729,978	13,351,788	10.32%	12,950,722	4,046,009	2,667,819	51.66%	2,266,753	11,876,592	11,856,573	11,772,604	(20,019)	(103,988)
1991									576,235	541,668			
Upper Limit with 97.5 % Two-Tail Test										778,117			
Lower Limit with 97.5 % Two-Tail Test										305,219			
Upper Limit with 95 % Two-Tail Test										718,289			
Lower Limit with 95 % Two-Tail Test										365,047			

Table 10. PROLIA Comparison of Estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Ultimate Loss Comparison				Liability Comparison				Actual		Loss Dev		
Accident Year	95 % Limit	Point Estimate	(1-2)/(2)	Loss Dev Method	95 % Limit	Point Estimate	(5-6)/(6)	LDF Method	Paid L @12/91	Model @12/91	Method @12/91	(10)-(9)	(11)-(9)
1982	989,769	966,017	2.46%	967,054	88,316	64,564	36.79%	65,601	943,316	933,372	933,745	(9,944)	(9,571)
1983	1,145,144	1,094,816	4.60%	1,101,458	187,046	136,718	36.81%	143,360	1,033,765	1,025,827	1,026,740	(7,938)	(7,025)
1984	1,298,761	1,207,311	7.57%	1,223,989	335,820	244,370	37.42%	261,048	1,097,869	1,061,559	1,064,681	(36,310)	(33,188)
1985	1,440,065	1,291,714	11.48%	1,328,108	537,273	388,922	38.14%	425,316	1,066,652	1,032,314	1,044,854	(34,338)	(21,798)
1986	1,621,608	1,389,740	16.68%	1,431,263	832,621	600,753	38.60%	642,276	978,806	976,726	972,913	(2,080)	(5,893)
1987	1,770,715	1,418,404	24.84%	1,310,343	1,257,356	905,045	38.93%	796,984	729,495	780,349	722,330	50,854	(7,165)
1988	2,066,074	1,592,504	29.74%	1,755,785	1,660,121	1,186,551	39.91%	1,349,832	661,341	629,879	687,872	(31,462)	26,531
1989	2,341,619	1,739,953	34.58%	2,370,755	2,078,041	1,476,375	40.75%	2,107,177	497,061	484,900	548,140	(12,161)	51,079
1990	2,582,807	1,849,230	39.67%	2,523,669	2,500,189	1,766,612	41.52%	2,441,051	260,440	282,845	280,579	22,405	20,139
Total	15,256,563	12,549,689	21.57%	14,012,422	9,476,784	6,769,910	39.98%	8,232,643	7,268,745	7,207,772	7,281,852	(60,973)	13,107
1991									102,397	88,792			
Upper Limit with 97.5 % Two-Tail Test										132,583			
Lower Limit with 97.5 % Two-Tail Test										45,000			
Upper Limit with 95 % Two-Tail Test										124,386			
Lower Limit with 95 % Two-Tail Test										53,198			

Table 11. Cumulative Loss and OLAE Payment Triangle

ACC YEAR	HOMEOWNERS/FARMOOWNERS 1	2	3	4	5	6	7	8	9	10
1982	5,693,422	7,434,119	7,714,556	7,910,584	8,029,886	8,122,750	8,169,220	8,206,128	8,216,708	8,222,508
1983	5,564,828	7,906,906	8,264,634	8,491,603	8,653,403	8,747,620	8,861,620	8,866,685	8,878,197	8,884,462
1984	6,213,606	8,213,101	8,617,067	8,849,286	9,003,560	9,103,862	9,153,728	9,175,840	9,188,790	9,195,274
1985	7,150,829	9,424,965	9,773,928	10,017,275	10,183,439	10,170,505	10,252,727	10,277,498	10,292,001	10,299,264
1986	6,592,556	8,589,695	8,998,577	9,248,197	9,401,991	9,477,972	9,554,595	9,577,678	9,591,195	9,597,963
1987	6,571,191	8,957,742	9,380,020	9,640,598	9,804,068	9,883,298	9,963,199	9,987,288	10,001,363	10,008,421
1988	7,415,249	9,970,199	10,410,312	10,691,036	10,872,322	10,960,184	11,048,791	11,075,483	11,091,114	11,098,940
1989	9,199,159	12,744,903	13,318,598	13,677,747	13,909,678	14,022,086	14,135,446	14,169,595	14,189,593	14,199,606
1990	9,236,916	12,403,667	12,961,992	13,311,525	13,537,245	13,646,644	13,758,969	13,790,204	13,809,666	13,819,411
1991	10,670,718	12,358,709	13,254,760	10,656,496	9,769,919	9,497,598	10,314,312	9,163,429	8,663,252	8,224,257
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT	
1982	1.3067	1.0377	1.0254	1.0151	1.0116	1.0057	1.0044	1.0014		
1983	1.4161	1.0450	1.0275	1.0191	1.0109	1.0131	1.0004			
1984	1.3218	1.0492	1.0269	1.0174	1.0111	1.0055				
1985	1.3180	1.0370	1.0249	1.0166	0.9987					
1986	1.3030	1.0476	1.0277	1.0166						
1987	1.3632	1.0471	1.0278							
1988	1.3446	1.0441								
1989	1.3654									
LAST 5 AVG	1.3428	1.0450	1.0270	1.0170	1.0081	1.0081	1.0024	1.0014	1.0007	
AGE-TO-ULT	1.4961	1.1141	1.0661	1.0382	1.0208	1.0127	1.0045	1.0021	1.0007	
EST LAST L		12,403,667	13,318,598	10,691,036	9,804,068	9,477,972	10,252,727	9,175,840	8,878,197	8,222,508
EST ULT LOSS		13,819,411	14,199,606	11,098,940	10,008,421	9,597,963	10,299,264	9,195,274	8,884,462	8,222,508

Note: Amount in AY 1985 – DY 6th adjusted to prevent from being a negative increments payment in our model.
It appears to be a typographical error in Best's publication.

Table 11. Cumulative Loss and OLAE Payment Triangle

PRIVATE PASSENGER AUTO LIABILITY/MEDICAL

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11
1982	5,757,145	10,773,841	13,072,279	14,372,676	15,063,154	15,432,107	15,628,671	15,696,425	15,742,360	15,765,395	15,776,929
1983	6,348,149	12,107,894	14,841,844	16,348,266	17,147,869	17,569,399	17,742,976	17,829,557	17,881,735	17,907,900	17,921,001
1984	7,124,946	13,777,714	16,965,364	18,744,239	19,728,053	20,222,374	20,472,816	20,567,144	20,627,333	20,657,516	20,672,629
1985	7,829,961	15,464,796	19,180,333	21,318,963	22,443,562	23,012,991	23,281,465	23,388,756	23,457,201	23,491,524	23,508,711
1986	8,706,107	17,266,158	21,563,704	23,961,790	25,234,299	25,862,024	26,163,756	26,264,307	26,361,227	26,399,800	26,419,114
1987	9,706,311	19,421,045	24,167,366	26,834,246	28,206,733	28,908,399	29,246,676	29,380,425	29,466,406	29,509,522	29,531,112
1988	10,929,661	21,779,669	27,007,191	29,916,300	31,446,526	32,230,637	32,606,877	32,757,112	32,862,975	32,901,046	32,925,117
1989	12,057,063	24,129,301	29,937,136	33,164,066	34,860,306	35,727,484	36,144,320	36,310,654	36,417,117	36,470,403	36,497,066
1990	13,342,210	26,565,496	32,959,719	36,512,462	38,370,949	39,334,663	39,760,604	39,976,953	40,083,944	40,152,611	40,161,967
1991	13,340,603	26,102,063	26,102,063	29,844,066	28,250,991	25,881,852	23,320,319	20,622,934	17,901,737	15,779,034	
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT		
1982	1.6714	1.2133	1.0996	1.0460	1.0245	1.0127	1.0043	1.0029			
1983	1.9073	1.2258	1.1015	1.0489	1.0246	1.0099	1.0049				
1984	1.9337	1.2314	1.1049	1.0525	1.0251	1.0124					
1985	1.9751	1.2403	1.1115	1.0528	1.0254						
1986	1.9855	1.2475	1.1107	1.0535							
1987	2.0009	1.2444	1.1104								
1988	1.9927	1.2400									
1989	2.0013										
LAST 5 AVG	1.9911	1.2407	1.1078	1.0511	1.0249	1.0117	1.0046	1.0029	1.0015	1.0007	
AGE-TO-ULT	3.0116	1.5128	1.2191	1.1005	1.0470	1.0215	1.0098	1.0061	1.0022	1.0007	
EST LAST L		26,565,496	29,937,136	29,916,300	28,206,733	25,862,024	23,281,465	20,567,144	17,881,735	15,765,395	
EST ULT LOSS		40,161,967	36,497,066	32,925,117	29,531,112	26,419,114	23,508,711	20,672,629	17,921,001	15,776,929	

Table 11. Cumulative Loss and OLAE Payment Triangle

COMMERCAL AUTO/TRUCK LIABILITY/MEDICAL

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11	12	13
1982	980,339	2,036,928	2,782,237	3,309,035	3,633,950	3,829,788	3,932,958	3,996,904	4,028,506	4,044,430	4,052,425	4,056,430	4,058,434
1983	995,644	2,201,489	3,079,672	3,713,367	4,114,104	4,337,825	4,480,017	4,545,730	4,581,670	4,599,782	4,608,874	4,613,429	4,615,709
1984	1,221,564	2,641,449	3,791,230	4,574,653	5,046,690	5,338,235	5,502,661	5,587,753	5,631,932	5,654,196	5,665,372	5,670,971	5,673,773
1985	1,368,162	3,059,678	4,324,870	5,277,826	5,850,344	6,173,433	6,359,705	6,458,050	6,509,110	6,534,841	6,547,758	6,554,229	6,557,466
1986	1,372,336	3,255,267	4,725,715	5,766,301	6,454,696	6,811,681	7,017,211	7,125,723	7,182,082	7,210,454	7,224,708	7,231,846	7,236,420
1987	1,432,429	3,732,418	5,200,463	6,390,591	7,077,190	7,468,601	7,693,952	7,812,929	7,874,701	7,905,631	7,921,456	7,929,267	7,933,205
1988	1,608,157	3,815,196	5,581,170	6,788,705	7,518,077	7,933,671	8,173,261	8,299,650	8,366,270	8,398,340	8,414,940	8,423,256	8,427,419
1989	1,885,297	4,293,936	6,146,015	7,476,756	8,278,947	8,736,821	9,000,439	9,139,620	9,211,661	9,248,297	9,266,577	9,275,736	9,280,319
1990	1,796,041	4,259,333	6,096,483	7,415,810	8,212,226	8,666,410	8,927,903	9,066,962	9,137,641	9,173,784	9,191,697	9,200,961	9,205,528
1991	1,704,266	4,195,956	6,220,537	6,815,726	7,065,223	6,839,937	6,360,626	5,563,276	4,577,032	4,042,160			
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT				
1982	2.0778	1.3659	1.1893	1.0982	1.0539	1.0269	1.0163	1.0079					
1983	2.2111	1.3989	1.2058	1.1079	1.0544	1.0328	1.0147						
1984	2.1623	1.4353	1.2066	1.1032	1.0577	1.0308							
1985	2.2365	1.4134	1.2203	1.1085	1.0552								
1986	2.3721	1.4517	1.2202	1.1194									
1987	2.6057	1.3933	1.2288										
1988	2.3724	1.4629											
1989	2.2776												
LAST 5 AVG	2.3728	1.4313	1.2164	1.1074	1.0553	1.0302	1.0155	1.0079	1.0040	1.0020	1.0010	1.0005	
AGE-TO-ULT	5.1283	2.1613	1.5100	1.2414	1.1210	1.0622	1.0311	1.0154	1.0074	1.0035	1.0015	1.0005	
EST LAST L		4,259,333	6,146,015	6,788,705	7,077,190	6,811,681	6,359,705	5,587,753	4,581,670	4,044,430			
EST ULT LOSS		9,205,528	9,280,319	8,427,419	7,933,205	7,236,420	6,557,466	5,673,773	4,615,709	4,058,434			

Table 11. Cumulative Loss and OLAE Payment Triangle

COMMERCAILL MULTI PERIL

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11	12	13
1982	2,236,250	3,391,657	3,907,911	4,356,420	4,732,145	4,993,411	5,161,217	5,264,202	5,352,311	5,386,604	5,404,162	5,412,669	5,417,230
1983	2,203,710	3,844,439	4,505,309	5,046,232	5,493,609	5,829,400	6,039,657	6,161,064	6,240,475	6,280,692	6,300,931	6,311,062	6,316,166
1984	2,579,181	4,260,446	5,065,162	5,755,156	6,277,324	6,656,273	6,909,564	7,047,564	7,136,421	7,164,426	7,207,576	7,219,166	7,225,004
1985	2,859,289	4,658,915	5,516,966	6,256,130	6,615,616	7,222,269	7,490,595	7,640,199	7,738,674	7,786,547	7,813,644	7,826,233	7,832,537
1986	2,517,583	4,107,624	4,971,444	5,666,695	6,266,307	6,639,164	6,866,626	7,023,351	7,113,676	7,159,722	7,182,793	7,194,365	7,200,161
1987	2,532,566	4,364,152	5,296,936	6,090,256	6,646,415	7,039,642	7,301,183	7,447,003	7,542,989	7,591,601	7,616,063	7,628,334	7,634,479
1988	3,005,661	4,996,132	6,045,472	6,676,073	7,503,990	7,947,966	8,243,242	8,407,677	8,516,246	8,571,132	8,596,750	8,612,604	8,619,642
1989	3,684,676	6,536,722	7,649,421	8,927,671	9,743,157	10,319,599	10,702,998	10,916,761	11,057,469	11,128,730	11,164,590	11,182,576	11,191,566
1990	3,640,505	6,130,429	7,362,663	8,374,236	9,136,964	9,679,660	10,039,264	10,239,790	10,371,773	10,436,615	10,472,251	10,489,123	10,497,573
1991	3,906,165	6,133,360	7,763,973	6,914,450	6,716,692	6,660,445	7,492,393	7,063,579	6,206,690	5,361,291			

	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT			
1982	1.5174	1.1522	1.1146	1.0662	1.0552	1.0376	1.0199	1.0129				
1983	1.7446	1.1719	1.1201	1.0687	1.0611	1.0361	1.0201					
1984	1.6519	1.1865	1.1365	1.0907	1.0607	1.0377						
1985	1.6294	1.1842	1.1343	1.0691	1.0596							
1986	1.6316	1.2103	1.1443	1.1018								
1987	1.7232	1.2137	1.1496									
1988	1.6619	1.2103										
1989	1.7737											
LAST 5 AVG	1.6640	1.2010	1.1374	1.0913	1.0592	1.0372	1.0200	1.0129	1.0064	1.0032	1.0016	1.0008
AGE-TO-ULT	2.5635	1.7124	1.4259	1.2536	1.1487	1.0845	1.0456	1.0252	1.0121	1.0056	1.0024	1.0008
EST LAST L		6,130,429	7,649,421	6,676,073	6,646,415	6,639,164	7,490,595	7,047,564	6,240,475	5,386,604		
EST ULT LOSS		10,497,573	11,191,566	8,619,542	7,634,479	7,200,161	7,832,537	7,225,004	6,316,166	5,417,230		

Table 11. Cumulative Loss and OLAE Payment Triangle

ACC YEAR	WORKERS' COMPENSATION 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1982	2,409,057	4,824,645	6,173,199	7,009,277	7,578,163	7,990,402	8,295,488	8,551,451	8,747,319	8,847,496	8,898,156	8,923,634	8,936,409	8,942,805
1983	2,601,939	5,429,395	7,009,517	8,037,067	8,737,532	9,229,591	9,569,075	9,866,412	10,092,399	10,207,980	10,268,433	10,295,826	10,310,565	10,317,945
1984	3,105,061	6,494,525	8,617,264	9,977,567	10,886,426	11,503,425	11,958,824	12,316,262	12,598,362	12,742,842	12,815,609	12,852,301	12,870,899	12,879,912
1985	3,316,436	7,241,296	9,633,975	11,202,492	12,217,466	12,914,977	13,417,449	13,818,464	14,134,992	14,296,870	14,378,737	14,419,904	14,440,547	14,450,863
1986	3,416,136	7,701,036	10,438,662	12,216,194	13,330,453	14,078,555	14,628,298	15,063,464	15,406,466	15,584,951	15,674,193	15,719,069	15,741,672	15,752,839
1987	3,830,969	8,934,312	12,063,264	14,022,619	15,260,031	16,116,420	16,743,449	17,243,895	17,638,860	17,840,866	17,943,026	17,994,396	18,020,156	18,033,056
1988	4,527,081	10,546,801	14,309,092	16,598,396	18,063,106	19,076,802	19,819,008	20,411,379	20,878,895	21,118,007	21,238,932	21,299,741	21,330,232	21,346,500
1989	4,934,351	11,881,481	15,968,066	18,522,791	20,157,316	21,288,540	22,116,796	22,777,846	23,299,564	23,569,396	23,701,343	23,769,202	23,803,229	23,820,266
1990	5,300,172	12,198,366	16,393,944	19,016,803	20,694,921	21,858,316	22,708,661	23,385,342	23,920,975	24,194,926	24,333,470	24,403,139	24,438,073	24,455,565
1991	5,488,466	12,900,811	16,069,736	16,567,748	15,268,334	14,106,048	13,439,155	12,298,335	10,059,841	8,893,778				
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT					
1982	2.0027	1.2795	1.1364	1.0812	1.0544	1.0382	1.0309	1.0229						
1983	2.0867	1.2910	1.1466	1.0872	1.0563	1.0369	1.0269							
1984	2.0916	1.3269	1.1679	1.0911	1.0567	1.0396								
1985	2.1836	1.3304	1.1628	1.0906	1.0571									
1986	2.2543	1.3555	1.1703	1.0912										
1987	2.3321	1.3502	1.1624											
1988	2.3297	1.3567												
1989	2.4079													
LAST 6 AVG	2.3015	1.3439	1.1600	1.0882	1.0561	1.0369	1.0299	1.0229	1.0115	1.0057	1.0029	1.0014	1.0007	
AGE-TO-ULT	4.6141	2.0046	1.4917	1.2860	1.1817	1.1189	1.0770	1.0456	1.0223	1.0108	1.0050	1.0021	1.0007	
EST LAST L		12,198,366	15,968,066	16,598,396	15,260,031	14,078,555	13,417,449	12,316,262	10,092,399	8,847,496				
EST ULT LOSS		24,455,565	23,820,266	21,346,500	18,033,056	15,752,839	14,450,883	12,879,912	10,317,945	8,942,805				

Table 11. Cumulative Loss and OLAE Payment Triangle

ACC YEAR	MEDICAL MALPRACTICE 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1982	50,001	172,083	383,096	675,119	951,902	1,197,254	1,384,910	1,517,031	1,630,944	1,692,177	1,723,943	1,740,125	1,748,291	1,752,394	1,754,450	1,756,479
1983	66,658	218,433	487,207	800,339	1,120,576	1,396,990	1,603,799	1,771,519	1,904,541	1,976,047	2,013,142	2,032,038	2,041,574	2,046,365	2,048,766	2,049,968
1984	104,213	298,416	609,084	973,318	1,337,378	1,611,920	1,836,606	2,019,146	2,170,782	2,252,263	2,294,543	2,316,080	2,326,950	2,332,410	2,335,146	2,336,516
1985	42,799	253,767	602,290	1,024,935	1,405,956	1,736,442	1,993,174	2,192,466	2,357,099	2,445,596	2,491,506	2,514,891	2,526,694	2,532,623	2,535,594	2,537,082
1986	52,427	261,420	625,567	1,006,077	1,362,218	1,683,963	1,932,936	2,126,207	2,266,662	2,371,685	2,418,207	2,436,886	2,450,332	2,456,061	2,458,963	2,460,406
1987	37,440	267,368	634,861	1,028,997	1,422,050	1,757,927	2,017,836	2,219,596	2,366,264	2,475,856	2,522,333	2,546,008	2,557,957	2,563,959	2,566,967	2,568,473
1988	55,867	337,606	733,256	1,196,365	1,656,113	2,047,273	2,349,961	2,564,930	2,779,031	2,863,369	2,937,496	2,966,068	2,978,964	2,985,974	2,989,477	2,991,231
1989	79,156	396,771	898,566	1,468,566	2,029,524	2,506,862	2,879,816	3,167,767	3,405,633	3,533,496	3,599,628	3,633,617	3,650,670	3,659,236	3,663,529	3,665,678
1990	88,012	512,297	1,163,155	1,900,954	2,627,075	3,247,569	3,727,719	4,100,448	4,408,349	4,573,859	4,659,721	4,703,456	4,725,532	4,736,620	4,742,176	4,744,960
1991	97,729	444,715	652,506	1,091,623	1,347,593	1,621,167	1,923,757	2,000,146	1,898,418	1,706,116						
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT							
1982	3.4412	2.2265	1.7623	1.4100	1.2577	1.1567	1.0954	1 0751								
1983	3.2770	2.2305	1.6427	1.4001	1.2467	1.1460	1.1046									
1984	2.8635	2.0410	1.5960	1.3740	1.2053	1.1366										
1985	5.9293	2.3734	1.7017	1.3718	1.2351											
1986	4.9864	2.3930	1.6063	1.3640												
1987	7.1417	2.3743	1.6206													
1988	6 0466	2.1706														
1989	4.9999															
LAST 5 AVG	6.8206	2.2705	1.6343	1.3620	1.2362	1.1478	1.1000	1.0751	1.0375	1.0166	1.0094	1.0047	1 0023	1 0012	1 0006	
AGE-TO-ULT	53.9126	9.2621	4.0794	2.4961	1.8062	1.4611	1.2729	1.1572	1.0764	1.0374	1.0183	1.0088	1 0041	1 0018	1.0006	
EST LAST L		512,297	898,566	1,196,365	1,422,050	1,683,963	1,993,174	2,019,146	1,904,541	1,692,177						
EST ULT LOSS		4,744,960	3,665,678	2,991,231	2,568,473	2,460,406	2,537,082	2,336,516	2,049,968	1,755,479						

Table 11. Cumulative Loss and OLAE Payment Triangle

SPECIAL LIABILITY

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11
1982	416,835	802,732	966,246	1,032,263	1,073,777	1,084,781	1,105,001	1,117,726	1,121,400	1,123,243	1,124,166
1983	463,774	907,517	1,082,762	1,141,067	1,197,957	1,228,120	1,254,419	1,266,225	1,269,384	1,271,470	1,272,515
1984	497,367	942,863	1,103,446	1,208,034	1,262,751	1,312,622	1,336,506	1,349,957	1,354,394	1,356,620	1,357,735
1985	471,503	920,190	1,091,495	1,205,696	1,266,407	1,308,230	1,333,631	1,347,054	1,351,482	1,353,703	1,354,815
1986	447,374	852,741	1,063,333	1,172,349	1,234,766	1,268,090	1,292,711	1,306,723	1,310,015	1,312,166	1,313,246
1987	407,167	934,308	1,193,111	1,317,862	1,361,298	1,418,574	1,446,117	1,460,673	1,466,474	1,467,862	1,469,069
1988	519,294	1,134,815	1,391,174	1,524,276	1,597,846	1,640,762	1,672,620	1,689,455	1,696,008	1,697,794	1,699,169
1989	545,437	1,221,393	1,491,896	1,634,415	1,713,066	1,759,316	1,793,477	1,811,529	1,817,463	1,820,470	1,821,966
1990	467,412	1,031,030	1,259,204	1,379,860	1,446,091	1,485,116	1,513,951	1,529,169	1,534,216	1,536,737	1,538,000
1991	676,236	1,102,659	1,479,786	1,536,560	1,393,629	1,283,562	1,327,123	1,356,864	1,273,497	1,124,673	

	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT	
1982	1.9258	1.2062	1.0661	1.0402	1.0102	1.0186	1.0115	1.0033		
1985	1.9516	1.1862	1.1046	1.0504	1.0330					
1986	1.9061	1.2470	1.1025	1.0532						
1987	2.2947	1.2770	1.1046							
1988	2.1849	1.2261								
1989	2.2393									
LAST 5 AVG	2.1153	1.2213	1.0957	1.0461	1.0270	1.0194	1.0101	1.0033	1.0016	1.0008
AGE-TO-ULT	3.1554	1.4917	1.2214	1.1148	1.0636	1.0356	1.0159	1.0058	1.0025	1.0008
EST LAST L		1,031,030	1,491,896	1,524,276	1,361,298	1,268,090	1,333,631	1,349,957	1,269,384	1,123,243
EST ULT LOSS		1,538,000	1,821,966	1,699,169	1,469,069	1,313,246	1,354,815	1,357,735	1,272,515	1,124,166

Table 11. Cumulative Loss and OLAE Payment Triangle

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
			OTHER LIABILITY												
1982	383,327	1,027,139	1,656,502	2,376,564	3,030,570	3,559,969	3,936,249	4,196,660	4,366,506	4,457,626	4,503,402	4,528,423	4,537,991	4,543,791	4,546,694
1983	377,761	1,059,724	1,840,764	2,662,610	3,392,365	3,994,025	4,495,365	4,564,066	4,771,526	4,869,090	4,918,669	4,944,013	4,956,649	4,962,963	4,966,155
1984	430,659	1,224,747	2,243,516	3,216,960	4,136,576	4,984,622	5,542,966	5,779,671	6,016,023	6,139,031	6,201,793	6,233,496	6,249,427	6,257,413	6,261,412
1985	621,770	1,452,577	2,543,634	3,925,197	4,960,053	5,890,504	6,566,565	6,645,911	7,125,866	7,271,566	7,345,906	7,363,456	7,402,329	7,411,769	7,418,525
1986	385,193	1,348,095	2,663,462	3,978,014	5,109,654	6,060,126	6,754,629	7,043,048	7,331,064	7,480,961	7,557,442	7,596,074	7,615,486	7,625,220	7,630,093
1987	347,135	1,445,695	2,792,912	4,167,779	5,320,409	6,310,067	7,033,233	7,333,546	7,633,444	7,769,524	7,869,159	7,909,364	7,929,569	7,939,733	7,944,806
1988	700,769	1,842,748	3,376,919	5,002,565	6,396,067	7,573,997	8,441,969	8,602,456	9,162,423	9,349,766	9,445,352	9,493,634	9,517,896	9,530,062	9,536,151
1989	771,309	2,084,434	3,886,502	5,759,169	7,361,936	8,719,506	9,718,777	10,130,764	10,546,171	10,763,846	10,873,891	10,929,476	10,957,410	10,971,413	10,976,423
1990	746,229	2,355,154	4,391,269	6,507,176	8,306,763	9,861,972	10,961,024	11,449,906	11,916,137	12,161,625	12,286,161	12,348,964	12,360,526	12,396,346	12,404,269
1991	745,429	2,170,377	3,770,531	4,917,109	5,317,321	5,963,973	6,566,766	5,952,007	5,077,919	4,494,388					
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT						
1982	2.6795	1.6127	1.4369	1.2741	1.1747	1.1063	1.0667	1.0409							
1983	2.8053	1.7370	1.4466	1.2741	1.1774	1.1265	1.0197								
1984	2.8439	1.8318	1.4339	1.2865	1.2044	1.1120									
1985	2.7839	1.7513	1.5430	1.2636	1.1876										
1986	3.4998	1.9757	1.4936	1.2845											
1987	4.1646	1.9319	1.4923												
1988	2.6296	1.8320													
1989	2.7025														
LAST 5AVG	3.1561	1.8645	1.4818	1.2765	1.1860	1.1146	1.0427	1.0409	1.0204	1.0102	1.0051	1.0026	1.0013	1.0006	
AGETULT	16.6226	5.2669	2.8246	1.9062	1.4933	1.2591	1.1296	1.0834	1.0408	1.0199	1.0098	1.0046	1.0019	1.0006	
EST LAST		2,355,154	3,886,502	5,002,565	5,320,409	6,060,126	6,566,565	5,779,671	4,771,526	4,457,626					
EST ULT		12,404,269	10,976,423	9,536,151	7,944,806	7,630,093	7,418,525	6,261,412	4,966,155	4,546,694					

Table 11. Cumulative Loss and OLAE Payment Triangle

PRODUCTS LIABILITY

ACC YEAR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1982	32,484	107,953	227,048	376,814	511,223	636,491	755,163	841,187	901,453	933,745	950,469	958,981	963,275	965,432	966,513	967,054
1983	33,238	121,563	262,823	431,183	621,698	754,467	873,065	958,098	1,028,740	1,063,520	1,082,569	1,092,263	1,097,154	1,099,611	1,100,842	1,101,458
1984	34,625	138,548	266,293	475,859	674,936	853,343	962,941	1,084,681	1,140,959	1,181,830	1,202,996	1,213,771	1,219,208	1,221,936	1,223,304	1,223,989
1985	45,661	148,932	309,119	542,320	746,394	902,792	1,044,854	1,155,248	1,238,015	1,282,363	1,306,331	1,317,021	1,322,919	1,325,860	1,327,365	1,328,108
1986	38,912	134,337	317,168	547,096	786,967	972,913	1,128,009	1,244,977	1,334,172	1,381,965	1,408,718	1,419,315	1,425,671	1,428,863	1,430,462	1,431,263
1987	41,498	157,286	327,941	513,359	722,330	890,717	1,030,878	1,139,798	1,221,456	1,265,211	1,287,872	1,299,406	1,305,224	1,308,146	1,309,610	1,310,343
1988	78,342	207,702	405,963	667,872	967,860	1,193,509	1,361,318	1,527,261	1,636,680	1,695,309	1,725,674	1,741,128	1,748,925	1,752,840	1,754,802	1,755,786
1989	66,867	263,578	548,140	926,801	1,306,864	1,611,540	1,866,130	2,062,190	2,209,934	2,289,098	2,330,096	2,350,966	2,361,493	2,366,780	2,369,429	2,370,755
1990	62,618	280,579	583,496	986,709	1,391,178	1,715,485	1,965,430	2,195,202	2,352,475	2,436,745	2,480,369	2,502,603	2,513,809	2,519,437	2,522,257	2,523,669
1991	102,397	260,440	497,061	661,341	729,495	976,806	1,088,652	1,097,869	1,033,765	943,316						
	1 TO 2	2 TO 3	3 TO 4	4 TO 5	5 TO 6	6 TO 7	7 TO 8	8 TO 9	9 TO ULT							
1982	3.3233	2.1032	1.6598	1.3567	1.2450	1.1864	1.1139	1.0716								
1983	3.6574	2.1620	1.6408	1.4418	1.2138	1.1572	1.0974									
1984	4.0014	1.9220	1.7870	1.4184	1.2643	1.1284										
1985	3.2600	2.0756	1.7544	1.3763	1.2095											
1986	3.4523	2.3610	1.7249	1.4421												
1987	3.7904	2.0850	1.5654													
1988	2.6512	1.9546														
1989	3.6262															
LAST 5 AVG	3.3961	2.0798	1.6945	1.4071	1.2331	1.1574	1.1057	1.0716	1.0358	1.0179	1.0090	1.0045	1.0022	1.0011	1.0008	
AGE-TO-ULT	30.5462	8.9946	4.3261	2.5528	1.8141	1.4711	1.2711	1.1498	1.0728	1.0357	1.0174	1.0084	1.0039	1.0017	1.0008	
EST LAST L		280,579	548,140	667,872	722,330	972,913	1,044,854	1,084,681	1,028,740	933,745						
EST ULT LOSS		2,523,669	2,370,755	1,755,785	1,310,343	1,431,263	1,328,108	1,223,989	1,101,458	967,054						