A PRACTICAL STATISTICAL METHOD OF ESTIMATING CLAIMS LIABILITY AND CLAIMS CASH FLOW

T. G. CLARKE, and N. HARLAND
U.K.

INTRODUCTION

1. A few years ago Scurfield (JSS 18) indicated a method used in one U.K. non-Life Office for estimating claims liability for the Motor class of business. Since that paper was written a considerable amount of development has taken place and it is now used in the office as an effective continuous method, using computer techniques, for estimating claims liability. It has also been used for projecting expected cash flow for claims arising in the past.

2. Whilst it is accepted that the method has limitations, it has been found that an automatic method is required especially in the production of financial models in the non-Life field, and its limitations are outweighed by the fact that there are, in our opinion, no easier methods of producing a satisfactory working model.

3. In the following paper it is the intention to set out briefly the details of the method, then concentrating on the known limitations and the methods so far devised to combat these limitations. In the appendices we have reproduced some computer printout and figures which will provide practical information on the method.

METHOD

4. The principle behind the method is that the "Run-off" of claims payments for any "year of claim", or similar cohort of claims, follows a particular pattern which experience has shown to be reasonably stable. We can thus study the pattern of claims payments at each stage of "run-off", say each month and thus ascertain the average and range of values that the historic pattern shows (fig. 1).
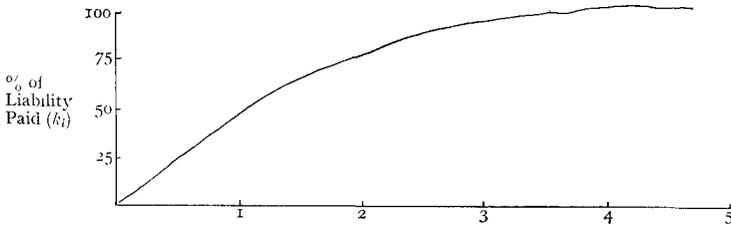
Fig. 1. Period of delay since beginning of claims Period (Years)

5. As indicated in Scurfield's paper, the method as used by one Company covered the whole motor account in total and was carried out at particular points of the year—in fact June, September and December. Whilst the method produced acceptable answers in times of stable business, when the constitution of the account changed, then one could find that the results became less acceptable, thus in recent work we have concentrated on improving the technique by recognising and adjusting for the known limitations as stated in the next section.

MAIN LIMITATIONS

6. *Mix of business*

The "run-off" pattern for specific sections of the business can vary considerably from one another. In fig (2). we show the difference between Comprehensive and non-Comprehensive Private Cars in the U.K. It is therefore necessary to attempt to separate the business into homogeneous groups whilst remembering the sub groups need to be of reasonable size.
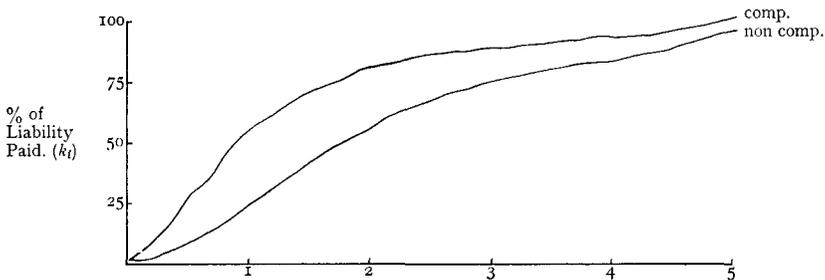

Fig. 2. Period of delay since beginning of claim period (years)

7. *Inflation*

The percentage "Run-off", at any point of time, relies on the fact that the rate of inflation is also stable and thus when the rate of inflation is increasing there is a tendency to undervalue the liability, and when the rate of inflation is reducing, there is a tendency to overvalue the liability. Unfortunately, the movement of the rate of inflation is cyclic and thus at one point of the payment "run-off" we can be undervaluing and using the percentage at another point of the same run-off we can be overvaluing.

8. For an example of this effect we have designed the following crude model. The historic run-off pattern is taken as 40% paid at the end of Year 1, 80% at the end of Year 2 and 100% at the end of year 3. We have used the following inflation factors.

Year 1	10%
Year 2	5%
Year 3	5%
Year 4	10%

The results are as follows:

Claim Yr.	Payments to end Yr. 1.	Payments to end Yr. 2.	Total Liability	Expected Liability Using historic pattern.	
				at end Yr. 1.	at end Yr. 2.
1	44.0	86.0	107.0	110.0	107.5
2	46.2	90.3	113.5	115.5	112.9

9. Although it is dangerous to prove a case by a single example, it can be seen that the wide fluctuations in the rate of inflation do not have too great an effect on the liability estimates. Certainly no more than could be accounted for by statistical fluctuations. Suffice it to say that we have found under recent conditions that the effect of variations in the rate of inflation on the valuation method has been less than was originally anticipated.

10. *Distribution of claims*

As with other factors the method relies on the fact that the distribution of claims, by number, for the period being estimated is similar to that of the period or periods on which the factors are

based. In the case of a portfolio, which is expanding or contracting, it will be found that the number of claims arising will probably also expand or contract, and the longer the period being valued the more important it is to allow for the distribution. This factor is extremely relevant in the first 18 months of a run-off. This point is demonstrated in Para. 27. The method detailed later in the paper concentrates on overcoming this limitation.

11. *Mix of Risk factors within a class of business*

Whilst the run-off pattern for a particular class of business may be similar for the individual risk groups within the class, we still may not project a reasonably accurate liability if during the period being valued there is a change in the mix of the individual risk factors. In this respect we are concerned with the different levels of cost of claims from the different risk groups and thus the effect on the run-off pattern during the period of change will be similar to the changing levels of rate of inflation. This limitation is also allowed for in the following detailed method.

12. *Seasonal effects*

During our work in this field we have notice that there are seasonal changes in the cost of claims from one period to another. If we are using the method over a period of time we should allow for this and the method detailed does this.

13. *Settlement/payment rate*

So far in the development of the method we have not found it possible to automatically allow for the rate of settlement. One difficulty has been to satisfactorily isolate what we should measure. We could measure the number of claims settled but this could be influenced by:

a) the number of claims settled at no cost as a proportion of all claims settled

b) What type of claims have been settled, the more expensive ones or the smaller ones

c) The number of claims which are settled but because of administrative errors are not shown to be settled in computer files.

14. This latter item has been the cause of many of our difficulties because invariably it has been found necessary to have a "clearing-up" operation from time to time and this does not necessarily occur at the same time each year thus the figures change from year to year.

15. A further difficulty has been the fact that we can experience a higher proportion of interim payments thus the amount paid increases but a similar increase in the number settled is not indicated.

16. We have carried out some research into measuring settlement rate by the use of measuring the number of payments per hundred claims. This should be a closer indication of the amount of the liability paid but it also suffers from the drawback that there may from time to time be changes in the number of payments per claim. It is in this direction where we believe that we will be able to find an automatic method of adjusting for settlement rate.

17. Development of Run-off Factors

In developing the factors it is necessary to take into consideration the effects of expansion and contraction of the account as well as inflation and settlement rate. Inflation can be accounted for by using an inflation factor or factors but as explained in the preceding section we have not been able to develop an automatic method to allow for settlement rate. Further factors which we have to allow for in the following formula are seasonal factors and mix of business within a class. We shall now develop the formulae allowing for all the factors except that for settlement rate. The formulae are developed to value the liability for a calendar year, but can easily be adjusted for any other period.

18. Let n_i = number of claims occurring in month "i" of a calendar year.

Let c_i = average cost per claim in month "i".

Let s_r = proportion of liability paid after "r" months measured from the beginning of the month of claim.

Let k_i = proportion of liability paid after "i" months for all claims arising in a calendar year, where "i" is measured from the beginning of that year.

Thus we can say that

$$k_1 = \frac{n_1\, c_1\, s_1}{\sum\limits_{i=1}^{12} n_i c_i}$$

$$k_2 = \frac{n_1\, c_1\, s_2 + n_2\, c_2\, s_1}{\sum\limits_{i=1}^{12} n_i c_i}$$

and generally

$$k_r = \frac{\sum\limits_{i=1}^{r} n_i\, c_i\, s_{r+1-i}}{\sum\limits_{i=1}^{12} n_i c_i} \qquad r \leqslant 12$$

$$k_r = \frac{\sum\limits_{i=1}^{12} n_i\, c_i\, s_{r+1-i}}{\sum\limits_{i=1}^{12} n_i c_i} \qquad r > 12 \qquad [A]$$

19. Just to emphasise the principle involved we could explain k_2 as follows $n_1\, c_1\, s_2$ is the amount paid by the end of February for these claims arising in January.
$n_2\, c_2\, s_1$ is the amount paid by the end of February for these claims arising in February.

Adding the above two together and dividing by $\overset{12}{\underset{i=1}{\Sigma}} n_i c_i$ the estimated total liability for the whole year we obtain the percentage paid at the end of February. Although the variables c_i appear in the formula it is not necessary to determine them absolutely since it is only the relationship between successive c_i which is important. This relationship is now derived.

20. Now Let:

$$j_i = \text{inflation factor} \qquad (j_1 = 0)$$
$$p_i = \text{seasonal factor} \qquad (p_1 = 1)$$
$$Z_i = \text{Mix of risk factor} \qquad (z_1 = 1)$$

The factor j_i which is the inflationary increase in the cost per claim in month "i", over month "$i-1$" can be variable over time as it almost certainly is. The seasonal factor is included because

for many classes it is apparent that there is some change in the level of cost depending on the time of the year. The mix of risk factor allows for possible substantial movement within a class of business between different risk levels. This factor is very relevant in a fast expanding or contracting account.

From the above we have:

$$c_2 = c_1(1 + j_2)\, p_2\, z_2$$
$$c_3 = c_1(1 + j_2)\,(1 + j_3)\, p_3\, z_3 \text{ etc.} \qquad [B]$$

thus

$$k_1 = \frac{n_1\, p_1\, z_1\, c_1\, s_1}{c_1[n_1\, p_1\, z_1 + n_2\, p_2\, z_2(1 + j_2) \, \ldots\ldots\, n_{12}\, p_{12}\, z_{12}((1 + j_2)\,(1 + j_{12}))]} \qquad [C]$$

$$k_2 = \frac{c_1[n_1\, p_1\, z_1\, s_2 + n_2\, p_2\, z_2(1 + j_2)\, s_1]}{c_1[n_1 p_1 z_1 + n_2 p_2 z_2(1 + j_2) \, \ldots\ldots\, n_{12}\, p_{12}\, z_{12}((1 + j_2) \ldots (1 + j_{12}))]}$$

etc.

Simplifying:

$$k_1 = \frac{n_1\, p_1\, z_1\, s_1}{\displaystyle\sum_{i=1}^{12} n_i\, p_i\, z_i\,\Big[\prod_{t=1}^{t=i}(1 + j_t)\Big]}$$

$$k_2 = \frac{n_1\, p_1\, z_1\, s_2 + n_2\, p_2\, z_2(1 + j_2)\, s_1}{\displaystyle\sum_{i=1}^{12} n_i\, p_i\, z_i\,\Big[\prod_{t=1}^{t=i}(1 + j_t)\Big]} \qquad [D]$$

etc.

from past historic data we will know the values of k_i and n_i and will have good estimates of p_i, z_i and j_i thus we can now estimate s_r as follows:

$$s_1 = \frac{k_1 \displaystyle\sum_{i=1}^{12} n_i\, p_i\, z_i\,\Big[\prod_{t=1}^{t=i}(1 + j_t)\Big]}{n_1\, p_1\, z_1}$$

$$s_2 = \frac{k_2 \displaystyle\sum_{i=1}^{12} n_i\, p_i\, z_i\,\Big[\prod_{t=1}^{t=i}(1 + j_t)\Big] - n_2\, p_2\, z_2\,(1 + j_2)\, s_1}{n_1\, p_1\, z_1} \qquad [E]$$

etc.

and generally.

$$s_r = \frac{k_r \sum_{i=1}^{12} n_1 p_i z_i [\prod_{t=1}^{i-i} (1+j_t)] - \sum_{i=2}^{r} n_i p_i z_i [\prod_{t=2}^{i-i} (1+j_t)]s_{r+1-i}}{n_1 p_1 z_1} \qquad r \leqslant 12$$

or $\qquad\qquad$ [F]

$$s_r = \frac{k_r \sum_{i=1}^{12} n_i p_i z_i [\prod_{t=1}^{i-i} (1+j_t)] - \sum_{i=2}^{12} n_i p_i z_i [\prod_{t=2}^{i-i} (1+j_t)]s_{r+1-i}}{n_1 p_1 z_1} \qquad r > 12$$

21. Using the above basic formulae we can assess the values of s_r for a number of years of historic information. From this information we can derive a standard table of s_r. It is preferable that the historic information used is in respect of years where the possible drift in claims still outstanding will not affect the values significantly.

22. The s_r table is then used in formula (D) together with either known or projected data regarding n_i, p_i, z_i, j_i to project the liability in respect of more recent years of claim.

23. It is appreciated that we could derive the s_r table from payment information collected for each month of claim, however the amount of historic data stored would be excessive. We believe that as this process can be computerized and the results derived from holding annual information are acceptable, the method is to be preferred.

PRACTICAL RESULTS

24. In our opinion in a live situation there will be a certain amount of computer output which will be required. This output can be divided into two main categories namely:

25. a) *Monthly output*

This provides an up to date projection of the liability using the latest month's claim payment information together with the projections as at the end of previous months of the run-off. (example in appendix 1).

26. b) *Periodic output*

Clearly the total liability in respect of historic years will be adjusted as time progresses. In order to see the effect of this on the

s_r factors we provide once or twice a year a tabulation showing the historic range of these factors. (example in appendix 2).

27. An indication of the effect the method developed in this paper has on the projected liability as compared with the crude method referred to by Scurfield is given below. This example has been produced from actual experience and covers a period of fast expansion after a period of stable business. We have related the projected liability by both methods to the actual experienced liability.
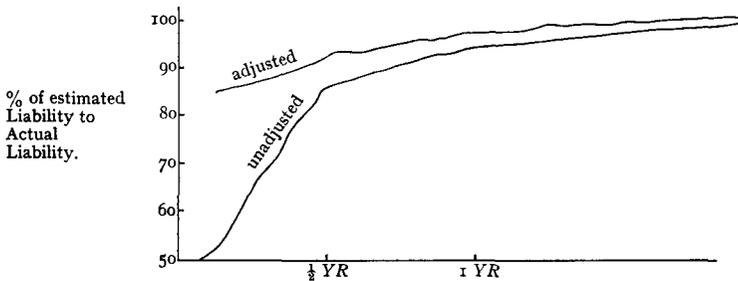

Fig. 3

28. In a practical situation one would not always place too much weight on the results obtained in the first three to six months development of the current year being projected because they will be sensitive to the assumptions and the statistical variation of the historic information, which will be large due to the small numbers involved.

29. During the first year of run-off when we are projecting the claims liability for claims arising during the whole year the method relies heavily on the accuracy of the projected number of claims arising each month. This projection of the number of claims takes into consideration increases in the level of business (known and projected) as well as projected claims frequency, and therefore the error in the projected liability during the period may be as much a function of the error in the projected number of claims as it is in the error due to random fluctuations within the method itself.

30. The other factors which have to be allowed for namely inflation, seasonal factors, mix of risk factors, all have to be projected. As mentioned earlier inflation has been found not to be of such importance as it was originally thought but in our opinion an attempt should be made to project the changes in the rate of inflation and used accordingly in the method. Both the seasonal factor and mix of risk factor can have varying degrees of effect on the outcome of the method depending on the class of business involved and one should attempt to be as realistic as possible in their projection whilst keeping a sense of proportion especially if the underlying historic data is subject to large random fluctuations.

31. *Conclusion*

In this paper we have developed a method of assessing the liability of different classes of general insurance by using the claim payment information. This method, accepting the limitations noted, provides a method which has been shown in one company to be readily adaptable to computer techniques.

32. Another advantage of the method is the fact that when the process is used in reverse we can estimate the expected future cash flow for claims arising in the past, a requirement which is most important when projections of interest income are to be made. These calculations are based on the net cash flow and an important element in its calculation is the cash flow of claims.

33. One disappointment to ourselves is the fact that we have been unable to find an acceptable method of allowing for the rate of settlement other than a subjective adjustment.

34. *Acknowledgements*

CLASS: — £000's

		Claim year 1970 No. of claims 2056		Month of print:— March 1972 Claim year 1971 No. of claims 4530	
		Liability	CPN	Liability	CPN
Year	PD				
1	1	195.3	95.0	419.4	92.6
	2	203.9	99.2	483.2	106.7
	3	187.8	91.3	624.0	137.7
	4	203.9	99.2	566.9	125.1
	5	211.0	102.6	587.6	129.7
	6	213.8	104.0	594.4	131.2
	7	212.1	103.2	586.5	129.5
	8	213.0	103.6	610.1	134.7
	9	214.4	104.3	597.3	131.9
	10	212.7	103.5	588.7	130.0
	11	210.0	102.1	594.0	131.1
	12	210.2	102.2	578.3	127.7
2	1	209.3	101.8	587.9	129.8
	2	202.9	98.7	597.9	132.0
	3	206.0	100.2	600.2	132.5
	4	210.5	102.4	0.0	0.0
	5	209.0	101.7	0.0	0.0
	6	212.5	103.4	0.0	0.0
	7	213.7	103.9	0.0	0.0
	8	213.0	103.6	0.0	0.0
	9	210.3	102.3	0.0	0.0
	10	209.8	102.0	0.0	0.0
	11	210.0	102.1	0.0	0.0
	12	209.0	101.7	0.0	0.0
3	1	206.6	100.5	0.0	0.0
	2	208.5	101.4	0.0	0.0
	3	208.0	101.2	0.0	0.0
	4	0.0	0.0	0.0	0.0

APPENDIX 2

CLASS: — £ooo's

			"S_r"	Month of print:—Dec 1972	
		1966	1967	1968	1969
Year	PD				
1	1	1.59	2.86	2.01	2.54
	2	7.53	9.57	8.21	9.06
	3	15.97	25.64	19.19	23.22
	4	20.77	31.26	24.27	28.64
	5	27.96	37.83	31.25	35.36
	6	27.99	39.84	31.94	36.88
	7	33.46	45.24	37.38	42.30
	8	31.30	50.58	37.73	45.76
	9	36.48	38.43	37.13	37.94
	10	40.61	50.90	44.04	48.33
	11	43.70	58.09	48.50	54.49
	12	37.50	46.02	40.34	43.89
2	1	45.73	45.26	45.57	45.38
	2	37.80	63.50	46.37	57.08
	3	45.70	59.17	50.17	55.80
	4	47.93	72.08	55.98	66.04
	5	62.12	60.57	61.60	60.96
	6	57.94	65.24	60.37	63.42
	7	60.52	66.47	62.50	64.98
	8	49.42	69.49	56.11	64.47
	9	55.80	67.89	59.83	64.87
	10	79.23	71.19	76.55	73.20
	11	75.83	78.53	76.73	77.86
	12	61.30	63.16	61.72	62.70
3	1	57.68	57.06	57.47	57.22
	2	58.66	71.25	62.86	68.10
	3	56.74	62.23	64.57	63.11
	4	64.31	142.43	90.35	83.84
	5	73.32	67.01	71.22	68.59
	6	88.37	79.19	85.31	81.84
	7	92.72	149.47	111.64	106.91
	8	71.54	78.72	73.93	76.93
	9	58.29	75.63	64.07	71.30
	10	89.84	83.71	87.80	85.24
	11	101.72	82.71	95.38	87.46
	12	82.32	84.98	83.21	84.32