Variants of the Chain Ladder

Murphy considered three calculations of chain ladder factors, namely regression, ratio of sums, and average of ratios. The three methods can have quite different impacts from extreme observations. Table 17 shows the first two columns of incremental losses from the Facultative General Liability Excess triangle of Mack (1993).

	0	1
0	5,012	3,257
1	106	4,179
2	3,410	5,582
3	5,655	5,900
4	1,092	8,473
5	1,513	4,932
6	557	3,463
7	1,351	5,596
8	3,133	2,262
sum	21,829	43,644

Table 17

Years 1 and 6 start low, so have high development factors. The respective factors from the three methods are 1.217, 1.999 and 7.206. To see their sensitivity to $q_{1,0}$, doubling it to 212 changes the factors to 1.222, 1.990 and 5.106. The implied fitted lines for the original factors along with the data points are graphed in Figure 9. The average residual is zero for the ratio of the sums, but it has a higher sum of squared residuals than the regression.

Again denoting the incrementals by y_i and the previous cumulatives by x_i, the derivatives of the respective factors f with respect to x_j are $[y_j - 2fx_j]/\Sigma x^2$, $-f/\Sigma x$, and $-y_j/nx_j^2$. For positive incremental losses, these are negative except for the regression estimate if y is at least double the factor times x. For $q_{0,1}$ the derivatives come out 4.66E-05, –9.16E-05, and –4.13E-02 respectively, showing greater sensitivity to the outlier as the power of x in the residual variance increases.

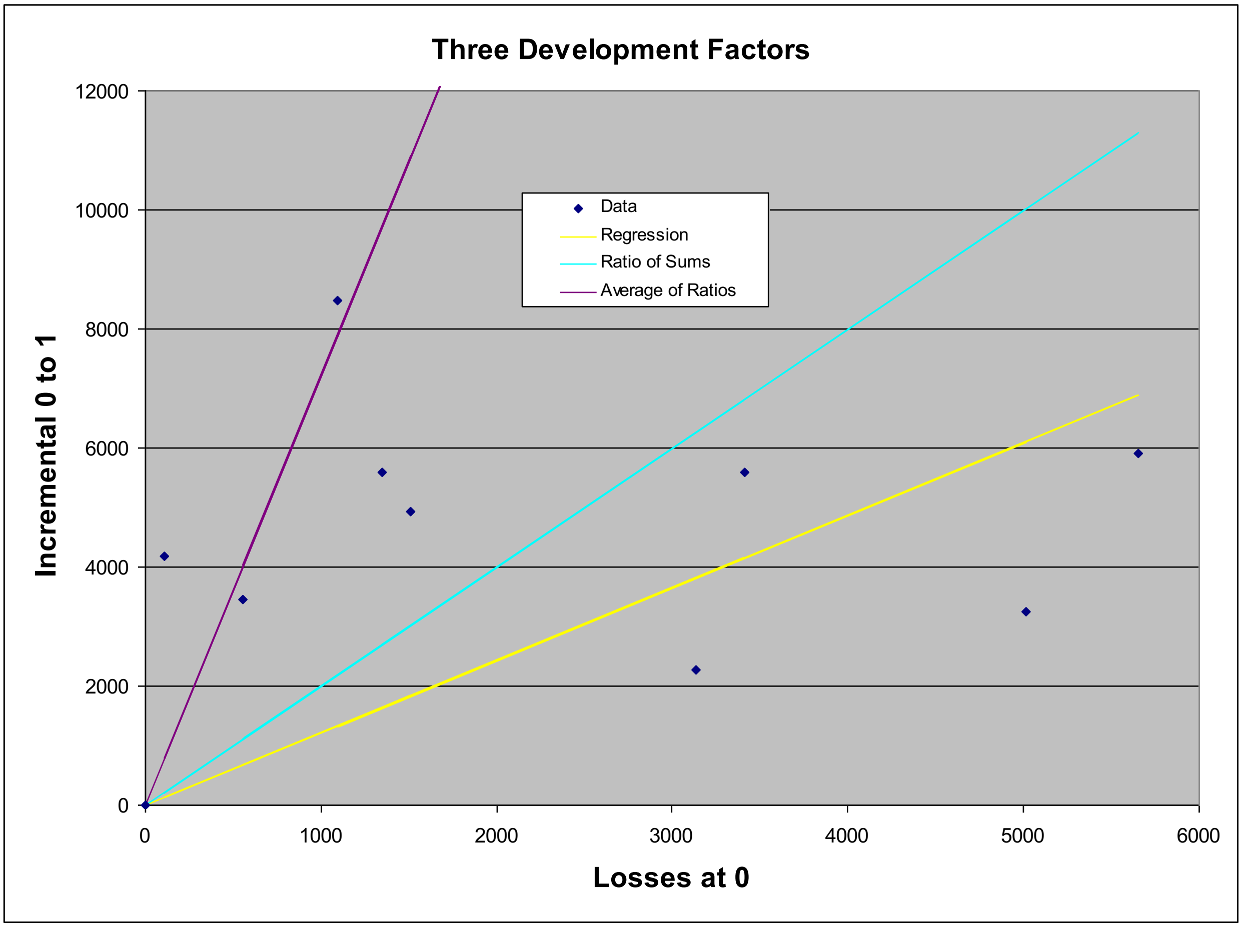

Figure 9

The ratio of sums and average of ratios estimates can be put into regression format by dividing both columns by the previous cumulative, for average of ratios, or its square root, for ratio of sums. For average of ratios, the individual ratios are modeled by a constant, and for ratio of sums the regression estimate $\Sigma xy/\Sigma x^2$ becomes $\Sigma(x/x^{1/2})(y/x^{1/2})/\Sigma x = \Sigma y/\Sigma x$.

This adjustment can be done for multiple regression as well. There is only one previous cumulative in each row of the matrix of independent and dependent variables, so the entire row, including the dummy variables and the 1 for the constant term if included, can be divided by the previous cumulative or its square root. Thus calendar-year effects can be modeled with any variant of the chain ladder. This adjustment is not likely to remove heteroscedasticity from the regressions, however, as the smallest incremental losses are still going to be factors times the largest previous cumulative.

Other variants of the chain ladder are possible. For any power s, dividing x and y by x^s gives the regression estimated factor of $\Sigma x^{1-2s}y/\Sigma x^{2-2s}$. This factor is applied to future estimates of x^{1-s} to estimate y/x^s, so would then be multiplied by x^s, but this is the same estimate as multiplying the factor by x. Even a lognormal chain ladder could be used, with $\log y = \log f + \log x + \varepsilon$, where ε is normal. Care is needed in this case, however, when exponentiating, to include the factor of $\exp(\frac{1}{2}\sigma^2)$ to get the expected value of the development factor. With a diagonal factor this model becomes $\log y = \log f + \log x + \log h + \varepsilon$.

Further variants of the chain ladder using generalized linear models are also possible. Generalized linear models replace the normal distribution assumption of the residuals with other distributions. The PCS could be used, for example, which would have variance proportional to mean for the entire multiple regression. This could in itself eliminate the problem of heteroscedasticity.