NOTES AND COMMUNICATIONS A Comparative Analysis of Most European And Japanese...
Gary G Venter
Journal of Risk and Insurance (1986-1998); Sep 1991; 58, 3; ABI/INFORM Global
pg. 542

NOTES AND COMMUNICATIONS

A Comparative Analysis of Most European And Japanese Bonus-malus Systems: Extension

Gary G. Venter

ABSTRACT

Upper and lower constraints for the sensitivity of experience rating are discussed by Lemaire(1988). A single widely advocated criterion - predictive accuracy - is shown here to provide both upper and lower constraints. This is illustrated by a simulation of the impact of four bonus-malus systems on the rating of a driver with long term consistent parameters. All four are found to be under-responsive to such a driver's individual experience.

Lemaire's (1988) invited article in this Journal elucidated the functioning of several countries' bonus-malus rating systems for automobile insurance. This comment reviews some aspects of those systems from the perspective of the American actuarial tradition.

A bonus-malus system (BMS) is a particular form of experience rating. Experience rating has been used in most lines of insurance in the US, beginning with workers' compensation in the years before World War I. It is generally used as a supplement to classification rating, in recognition that there are differences among risks, identifiable by individual risk experience, that are not accounted for by classification rates.

The sensitivity of a rating plan has both upper and lower constraints. Lemaire introduces the concept of efficiency, which measures the incremental change in premium induced by an incremental change in an insured's true loss frequency. The ratio of these changes should be 100 percent for perfect efficiency. The sensitivity of the plans reviewed never reaches this level, due to the constraints of market resistance to high surcharges and a social reluctance towards having the plan financially encourage the non-reporting of accidents.

Similar constraints have operated in the Western Hemisphere. For instance, Snader (1980) points out that the sensitivity of a 1940 US workers' compensation experience rating plan was determined in part by specifying that the smallest rated risk could be debited no more than 25 percent for a single accident.

Gary G. Venter is President of the Workers' Compensation Reinsurance Bureau, New York, New York.

The main criterion cited for optimal plan sensitivity, however, has been how well the plan estimates individual insured results. For instance, Meyers (1985) says "The purpose of experience rating is to estimate the expected loss ratio," Venter (1987) says "the degree to which an insured should be charged for past loss experience is the degree to which that experience is predictive of future loss experience," or Freifelder (1985) says "The accurate prediction of an insured's true loss potential is the goal of the ratemaking process." This note will explore BMS plan sensitivity from the viewpoint of predictive accuracy.

Predictive accuracy provides both an upper and lower constraint on the sensitivity of a plan to risk experience. If the plan is not sensitive enough, high risk insureds will be undercharged and good risks overcharged. If the plan is too sensitive, the opposite will occur, as random fluctuations in experience will be accorded too much weight. This happens currently with experience rating of large commercial risks, which are sometimes given full credibility.

Lemaire's concept of efficiency is related to predictive accuracy, in that a 100 percent efficient plan in actuarial balance should be highly accurate. Nonetheless, as will be seen below, plans with comparable measures of efficiency do not necessarily predict equally well.

A direct measure of predictive accuracy could be developed by using a risk's relative charge from the BMS as an estimate of its relative frequency. By simulation, the estimated frequency could be compared to the true frequency, and the average of the squared errors, for example, could provide a measure of the predictive accuracy of the plan. An analytic result using the expected squared error approach to evaluating experience rating can be found in Freifelder (1985) where, rather than using BMS type rules, experience rating is based on Bayes estimation in the gamma-Poisson model, introduced by Greenwood and Yule (1920) and by Keffer (1929).

Applying this approach to bonus-malus systems, a simulation was performed for four countries' BMS's at five assumed claim frequencies (λ). Each risk kept its assumed frequency over time. For each risk, 20 years of simulated experience were put through the BMS rules to determine a rating level, and hence an implied estimated frequency level. At each claim frequency 10,000 such risks were simulated, and for each risk the estimated frequency was compared to the assumed frequency that generated the experience. The resulting mean squared errors for each assumed frequency are shown in Table 1.

Table 1

Mean Squared Error

λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.0046	.0014	.0027	.0019
.10	.00014	.0015	.00070	.0013
.25	.026	.019	.020	.015
.50	.15	.11	.075	.041
.90	.52	.49	.33	.30
Average	.0069	.0051	.0049	.0040

The average for each country is a weighted average of the five frequencies shown. To compute the weights, it was assumed that the mean risk frequency was .10 in each country, and that the risk Poisson frequencies were gamma distributed with a variance of .007. (This results in an overall negative binomial distribution of claims, with mean .10 and variance .107. This was also the distribution Lemaire used, although this may be somewhat unclear in the paper.)

If the mean squared error for every frequency were known, the average would be their integral, weighted by the gamma density. To approximate this numerically, the five frequencies were chosen to be near the abscissas for five point Laguerre integration, so the weights for this integration were combined with the gamma density function to yield respective weights for the five points of: .3325, .5218, .1483, .0094, and .0001. The gamma percentiles for these 5 points are 11, 61, 94, 99.8, and 99.999, respectively.

The overall estimated frequency should be close to .10. The average of the estimated frequencies for each λ and overall are given in Table 2.

Table 2

Estimated Frequencies

λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.088	.057	.072	.064
.10	.088	.073	.077	.077
.25	.090	.12	.12	.17
.50	.11	.17	.24	.31
.90	.18	.20	.33	.36
Average	.090	.076	.084	.089

A downward bias is evident in all of these plans, mitigated only by a capping on the credit of the lower frequency risks. The Italian plan fails to differentiate risks to any appreciable degree, giving it the lowest bias overall, but the highest average estimation errors. The plans of Luxembourg and the Netherlands have roughly comparable errors, even though they operate quite differently. Lemaire rated the Dutch plan at double the efficiency of that of the Grand Duchy. The Swiss plan works the best of these, according to the squared error criterion. It is not assumed that the under-responsiveness of these plans is inherent in the BMS in general. It is more likely a result of the particular operating characteristics of the individual plans. Lemaire suggests that policyholder resistance to actuarially balanced surcharges may have contributed to this effect.

To explore the extent to which random error within the simulation may have contributed to the results, additional simulation are shown in Table 3. First, only 1000 drivers were simulated at each frequency. Second, ten years of driving rather than 20 was simulated for each driver. The general conclusions about the relative and absolute levels of responsiveness of the plans were the same in each case.

Table 3

Additional Simulations

20 Years per driver, 1000 drivers per

		Estimated Frequencies		
λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.088	.057	.072	.064
.10	.088	.070	.075	.074
.25	.089	.12	.11	.15
.50	.11	.17	.23	.31
Average	.090	.075	.082	.086

		Mean Squared Error		
λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.0046	.0014	.0027	.0019
.10	.00014	.0014	.00071	.0012
.25	.026	.020	.021	.017
.50	.15	.11	.081	.044
.90	.52	.49	.33	.30
Average	.0069	.0052	.0051	.0042

10 Years per driver, 10,000 drivers per

		Estimated Frequencies		
λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.088	.070	.077	.068
.10	.089	.093	.097	.094
.25	.097	.14	.13	.16
.50	.12	.18	.21	.27
.90	.18	.20	.31	.35
Average	.091	.094	.098	.098

		Mean Squared Error		
λ	Italy	Netherlands	Luxembourg	Switzerland
.02	.0046	.0027	.0034	.0025
.10	.00014	.0014	.00059	.0015
.25	.024	.015	.016	.014
.50	.14	.11	.092	.059
.90	.52	.49	.35	.31
Average	.0065	.0048	.0046	.0043

Operating characteristics of the plans that contribute to predictive accuracy include the following:

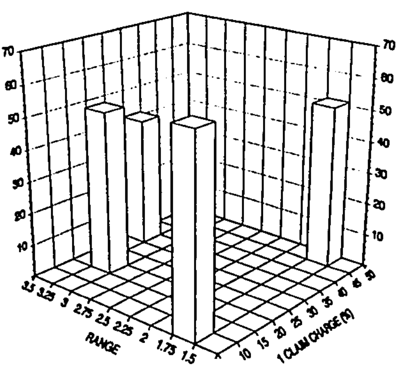
	Italy	Netherlands	Luxembourg	Switzerland
Relative High Charge	2.5	2.2	3.6	3.8
Average Charge	1.0	1.0	1.0	1.0
Relative Low Charge	.88	.55	.71	.63
Range of Relative Charges	1.6	1.6	2.9	3.2
Charge for a Single Claim	6%	45%	14%	29%

Having a wide range of charges and a large upward movement from a single claim both make a plan more responsive. The plans of the Netherlands and Luxembourg appear to have a trade off between range and surcharge for a single claim that makes their effects similar. The Swiss plan seems to gain in accuracy from its wide range and relatively large step size, while the Italian plan works in the opposite direction to equalize charges across risks. Incorporating larger single claim surcharges and/or a wider range between high and low limits would thus appear to have potential for improved BMS predictive accuracy for long term stable risks.

To help visualize these effects, Figure 1 shows the squared error as a function of range and single claim surcharge for these four plans.

Extensions of this method could give a more complete picture of the effects of BMS plans. As noted, the above results depend on the model assumption of constant risk frequency over 20 years. A potentially more realistic model could allow the frequency some random variation, as driving patterns might

Figure 1



Mean Squared Error (ten thousands)

fluctuate with the ups and downs of life, along with a downward drift over time, at least in the initial years, when driving skill is improving. Such a model would make it reasonable to measure the accuracy each year, not just after 20 years. However, this would require additional information about changes in accident frequency over time. In addition, the variance across risks should be reflective of the variance within the classes that are used in conjunction with the BMS, rather than the variance across all risks, but this would again require additional information to perform. The gamma assumption could be varied as well, in that distributions very close to the negative binomial can be formed by mixing Poissons by other structure functions.

Although, given data availability, a more realistic simulation would be possible, and probably should be performed, the problems inherent in modeling the distribution of risks have produced a preference for actual data among many actuaries. A method of using available data to measure predictive accuracy was prescribed by Dorweiler (1934). Applied to a BMS, this method would involve looking at the next year's experience of all risks that are found in each BMS level. The relative loss experience should be compatible with the charge for the level, so that the loss ratios are equal across levels. This test could also be performed with simulated data, if a realistic model were used. Given the apparent under-responsiveness of current BMS's, this test would probably indicate that the risks that are charged the most are nonetheless subsidized by the risks that are charged the least.

REFERENCES

Dorweiler, Paul, 1934, A Survey of Risk Credibility in Experience Rating, Proceedings of the Casualty Actuarial Society, 21: 1.

Freifelder, Leonard R., 1985, Measuring the Impact of Merit Rating on Ratemaking Efficiency, Journal of Risk and Insurance, 52: 607–26.

Greenwood, M. and G.U. Yule, 1920, An Inquiry into the Nature of Frequency Distribution of Multiple Happenings, Journal of the Royal Statistical Society, A, 83: 255–79.

Keffer, Ralph, 1929), An Experience Rating Formula, Transactions of the Society of Actuaries, 30: 130–9.

Lemaire, Jean, 1988, A Comparative Analysis of Most European and Japanese Bonus-malus Systems, Journal of Risk and Insurance, 56: 660–81.

Meyers, Glenn G., 1985, An Analysis of Experience Rating, Proceedings of the Casualty Actuarial Society, 72: 278–317.

Snader, Richard H., 1980, Fundamentals of Individual Risk Rating and Related Topics, New York, Casualty Actuarial Society.

Venter, Gary G., 1987, Experience Rating Equity and Predictive Accuracy, NCCI Digest, April: 27–35.