Enhancing **profitability** in a **tariff** market - **scoring** for insurance using **GLMs**

Will Southwell

Will Southwell discusses how profitable personal lines business can be targeted using statistical scoring methods.

Conventional wisdom dictates that in order to maximise profits, companies operating in a tariff market such as Russia's CTPL motor insurance market should write as much business as possible, regardless of portfolio mix. Most insurance companies in Russia have now realised that this is not the case and are looking for more advanced techniques to weight their portfolios with risks that are expected to be profitable. Not only does this create higher profits in the short term, but looking to the future, now is the time to start adjusting the mix of business so that an optimal portfolio is in place when tariffs are eventually removed.

There are a number of methods that have been used in tariff markets around the world to select profitable business. The basic technique is to use a one-way analysis to look at one factor in isolation such as region in order to highlight profitable areas of the country in which to sell a product. In this article we will be looking at a much more sophisticated method that uses generalised linear models (GLMs) to create a statistical score, analysing all significant factors in the same model to create a structure for analysing the profitability of individual risks. We will start by looking at what a score is and how it can be used to target profitable business. We will then go on to explain how to create a statistical score using GLM software (such as Watson Wyatt Pretium®) and give some examples of situations where Watson Wyatt has used this method successfully in the past.

GLMs: Advantages over one-way analysis

A generalized linear model (or GLM) statistically measures the effect that variables have on an observed item. In insurance, GLMs are used to determine the effect rating variables have on claims experience and the effect that rating factors and other factors (for example competitiveness) have on the probability of a policy renewing or new business quotation being accepted.

GLMs have the key advantages of:

- making appropriate allowance for correlations between rating factors enabling the true influence of any individual factor to be isolated.
- providing practical statistical diagnostics which help assess whether a particular feature observed in the experience data is systematic or only random noise
- being robust and able to cope with large quantities of insurance data
- yielding results in a transparent and easily understood way.

GLMs incorporate assumptions about the nature of the random process underlying the claims experience. Unlike iterative one-way methods, GLMs do not assume that all observations come from one fixed distribution. In practice, claims frequency and severity have different distributions, and being able to specify a profitability distribution that matches the observed behaviour increases the accuracy of the analysis.

What is a statistical score?

A statistical score is a value allocated to a particular policy or customer in order to rank it by a particular attribute, for example, profitability, probability of default, probability of lapse, etc.

The banking industry has been using scoring techniques for many years, an obvious example being the credit scores used to measure the likelihood that a person will default on a loan. Credit-based scores are derived by assembling variables from individuals' raw credit reports (such as late payments, bankruptcies, number of inquiries and amounts past due) and deriving from them a composite score that is more predictive of a customer default than any of the variables would be in isolation.

Scores are also being used by many insurance companies and intermediaries as a means to share vital information between the actuarial and statistical departments and the rest of the organisation. For example, scores can be used to predict the profitability of an insurance policy given a certain rating structure, the probability that a policy will lapse or the probability of a quote being converted into an active policy.

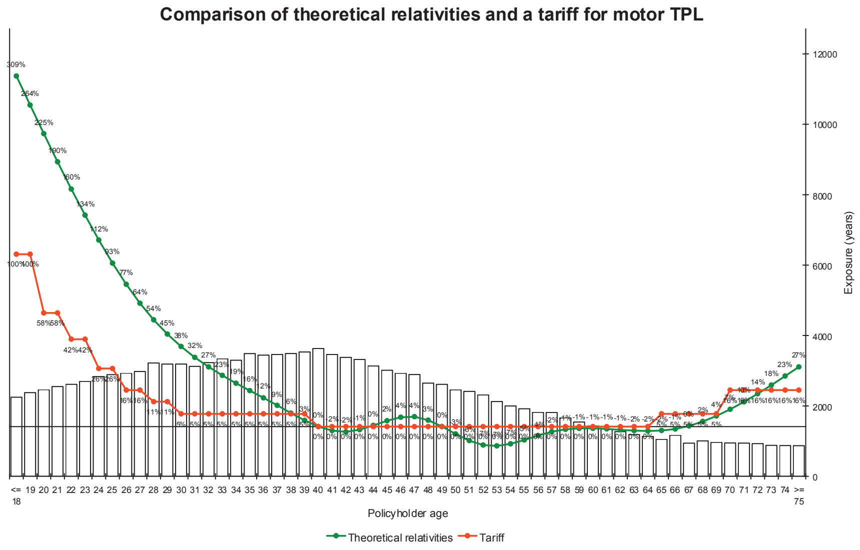
How to use scores to profit in a tariff market

The most direct way to manage the profitability of a personal lines product is through effective premium rating. In the case of the Russian CTPL market the setting of rates is restricted by the enforced tariff. In these circumstances, a score based on the expected loss ratio can be used by insurers to help target their marketing at those people who are likely to be more profitable. The score also provides a common language in order to communicate a strategy throughout the company that everybody is capable of understanding.

These loss ratio scores can be used in many ways depending on how products are distributed to customers and whether you are an insurance company or insurance intermediary. The list is only restricted by the imagination of the people involved. Below are three examples.

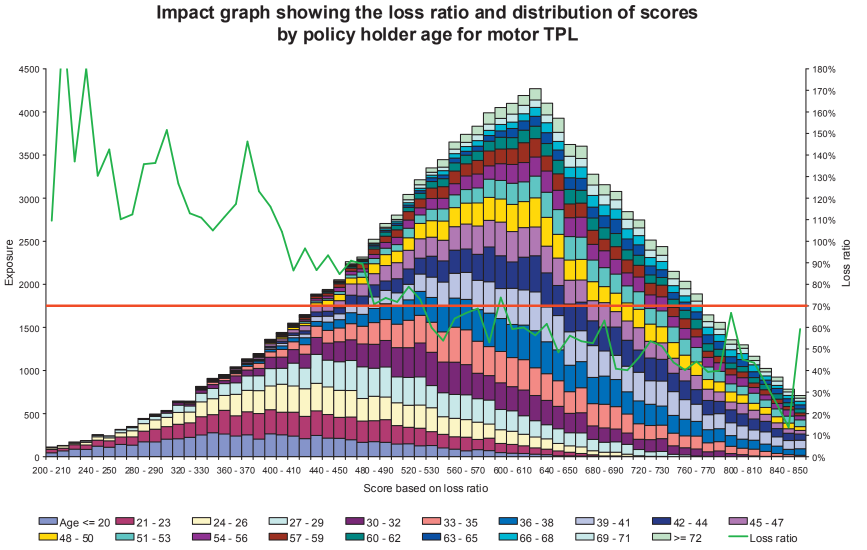
- If insurance is distributed through agents then the score can be used as part of an incentive scheme for direct sales agents, where commission or bonus is linked to the average or total customer score that an agent achieves. When an agent gives a premium quote to a potential customer they not only calculate the appropriate premium to charge but also calculate a customer score based on the information given. This score is produced using a simple table in the same way as premiums. This information can also include policyholder characteristics that are not permitted in the actual premium tariff. Agents will quickly develop their own strategies in order to maximise their average/total score.

Figure 1 | Example



This graph shows the theoretical relativities of risk for drivers of different ages produced using GLMs and smoothed using splines compared with a hypothetical tariff. Relativities are the proportional increase/decrease of a particular factor level when compared with the base level (in this case 40 year olds). The bars show the exposure at different age bands. In this case the tariff is not adequately covering the additional risk for younger drivers.

Figure 2 | Example


Impact graph showing the loss ratio and distribution of scores by policy holder age for motor TPL

This graph is based on the same dataset and tariff as example 1. It shows how a score can be used to distinguish very effectively between profitable and unprofitable business. The shape of the stacks shows the distribution of scores across the range 0 – 1000. Each of the colour bands represents a different policy holder age group and shows the distribution of the scores for this age group. The green line shows the actual loss ratio experienced for business according to these scores. If the company has a target of a 70% loss ratio (red line) then this graph shows that they should be targeting scores greater than 520-530. The graph also shows that the majority of the younger drivers are towards the left of the graph which demonstrates the inadequacy of the tariff over this range as shown in **Figure 1**.

Figure 3 | Example

Example score structure

INTERCEPT
534

Policyholder age

Factor level	Estimate
<= 20	-271
20-21	-231
22-24	-199
25-27	-151
28-30	-95
32-33	-55
34-36	-22
37-39	0
40-42	-6
43-45	-18
46-48	0
49-51	26
52-54	19
55-57	3
58-60	31
61-63	55
64-66	60
67-69	54
>=70	22

Population density

Factor level	Estimate
<= 22	165
23-27	174
28-31	182
32-35	184
36-45	183
46-57	184
58-62	186
63-67	162
68-73	177
74-88	142
89-100	190
101-119	143
120-130	147
131-148	138
149-163	120
164-184	98
185-209	48
210-225	69
226-254	74
255-284	28
>=285	0

Vehicle group

Factor level	Estimate
1	164
2	156
3	112
4	69
5	117
6	53
7	56
8	-6
9	8
10	0
11	-59
12	-72
13	-84
14	-110
15	-88
16	-138
17	-150
18	-143
19	-189
20	-231

Policy duration (years)

Factor level	Estimate
0	0
1	15
2	13
3	16
4	22
5	26
6	26
7	29
8	33
9	34
10	39
11	38
12	42
13	46
14	48
>= 15	52

These tables show an example scoring structure based on the model in the previous 2 examples. The intercept is the starting score for a customer with additions and subtractions to this made as appropriate based on the customers characteristics. For example, a 25 year old person who lives in an area with population density of 100, drives a car of insurance group 5 and has been with the company for 5 years will have a score of 716. Again you can see the low score for younger drivers.

- Insurance intermediaries such as banks and brokers that have profit share arrangements with the insurance company can use a score based upon the expected profit that they will receive for each customer. This score can then be used to individually target potential insurance customers from their (generally much larger) banking or insurance customer base. This is a particularly powerful tool for banks since not only can they use the insurance data received from the insurance company but can also use banking and credit type information which can be very predictive of insurance risk.

- On a more basic level, scores can be used to develop an appropriate marketing and advertising campaign. For example if the analysis shows that young females that drive Toyotas, live in St Petersburg and own a dog have a high profitability score then advertising could be placed in young women's magazines and on advertising boards around St Petersburg, insurance could be sold through Toyota garages and there could be special offers on cans of dog food!

Producing the score

One method of deriving a scoring algorithm takes advantage of the ‘linear’ part of GLMs. The output of a GLM is a series of additive parameters which is then transformed via a ‘link function’ to give the expected value for an observation. When calculating a score the link function can be omitted for simplicity, leaving an additive structure which orders the risk. A straightforward calculation can then transform the additive structure into a scoring algorithm which produces scores between 0 and 1000.

To derive a profitability score, the starting point would be a standard analysis of claims experience using GLMs. This would involve fitting a series of GLMs to historic claims data, considering frequency and severity separately for each claim type. These models would include relevant rating factors, as well as any additional information available at the time the score is to be calculated. Such additional information could include geodemographic data, or intelligence on the customers' habits and preferences gathered by monitoring their transactions (for example, using information from banking records if the insurance is sold through a bank).

The expected cost of claims can then be calculated for each record in the data based upon the GLM claims models. For each policy this can be divided by the premium that would be charged under the tariff to yield an expected loss ratio, which can then itself be modelled to derive the profitability score.

The model of expected loss ratio should include only those factors that will be considered in the calculation of the score. For direct marketing campaigns this will usually mean that the traditional insurance rating factors used in the premium will have to be excluded at this point (since they are not known at the time of the campaign).

Case studies

Large US motor insurer – Watson Wyatt conducted an analysis of a US insurer's motor TPL data to produce a statistical score system for the purpose of creating a broker incentive scheme. In the state in which this insurer was operating, the market was highly regulated with the authorities restricting the factors and relativities that could be used in a rating structure. Watson Wyatt produced the theoretical risk based rating structure using their GLM modelling software Pretium which was then used to create a score using the restricted premiums. The insurance company then based its broker commission rates on the customer scores.

Large UK retail bank – The large retail bank in question was acting as an insurance intermediary. The bank received a commission and a share of the insurer's profits but was unable to determine the rates at which the insurance was sold. In order to maximise the profit share it received it wanted to specifically target the customers with the largest profit margins. Watson Wyatt produced a scoring system based on the insurance data and the bank's additional credit scoring data. The bank then embedded this score in their system so that all customers with a high score would be offered the insurance product.

Conclusion

Scores are simple to produce, easy to explain and offer an immediate guide to the prospective market. In a tariff market with increasing loss ratios such as Russia's CTPL market more and more insurers are looking at ways to enhance their profitability. Insurers that are willing to accept whatever business comes their way will continue to suffer anti-selection and will see profit margins decline. A large competitive advantage can be obtained by producing a sophisticated profitable business targeting scheme using a technique such as statistical scoring.

Pretium®

Will Southwell
+44 (0) 1737 241144
william.southwell@watsonwyatt.com

Get ahead of the competition…

…general insurance financial modelling for the 21st century

Watson Wyatt Pretium® – The complete premium rating system including price optimisation

Watson Wyatt Simulum® – Dynamic financial analysis software for non-life insurance

Watson Wyatt IBNRS® – Reserving software for non-life insurers

Pretium	James Tanser	pretium@watsonwyatt.com
Simulum	David Slater	simulum@watsonwyatt.com
IBNRS	Bart Kling	ibnrs@watsonwyatt.com

Telephone +44 (0) 1737 241144 or visit www.watsonwyatt.com/insurancesoftware

Watson Wyatt Worldwide

Watson Wyatt Limited, 21 Tothill Street, Westminster, London SW1H 9LL.
Authorised and regulated by the Financial Services Authority.

© Watson Wyatt Limited 2007-EU-0445