Data Mining Techniques for Optimizing Québec's Automobile Risk-Sharing Pool

TECHNOLOGICAL WHITE PAPER

Charles Dugas, Ph.D., A.S.A.
Director, insurance solutions

ApSTAT Technologies Inc.
4200 Boul. St-Laurent, Suite 408
Montréal (Québec)
Canada H2W 2R2
1-866-9APSTAT
www.apstat.com

1 Introduction

In the last twenty years, statistical learning algorithms have gained widespread interest among academics of various research communities and practitioners working in diverse industries. They have been successfully implemented to perform predictive tasks related to observed stochastic processes for which multiple explanatory variables can be identified. Recent predictive modeling technologies boast an excellent capacity to capture high-level dependencies, i.e., dependencies that involve many variables simultaneously.

The purpose of this document is to illustrate the profitability that can be gained from the use of predictive modeling for selecting the risks to be ceded to Québec's PRR*. Ratemaking, underwriting and other actuarial tasks must usually reflect strategic concerns as well as analytical ones. The PRR stands out as a purely analytical actuarial problem: the success of the operation is solely based on the use of highly discriminative technologies to power decision tools, allowing insurers to identify undercharged risks. For this reason, it is much easier to build a sound and illustrative business case for the introduction of a new technology by showing how much profit can be gained through its use for PRR purposes. The rest of the document goes as such: we describe technicalities of the PRR in section 2, hypotheses and methodology of our experiments follow in section 3, numerical results appear in section 4 and we conclude in section 5.

**Plan de Répartition des Risques (Québec's Risk-Sharing Pool).*

2 PRR Technicalities

The PRR was created by Québec's GAA† in order to allow insurers to share a portion of their risks, up to 10% of their book of business. Each insurer chooses the risks to cede to the PRR. For those risks, the PRR will reimburse all claims incurred by the insurer. In return, the insurer must pay 75% of the gross premiums that were charged to the ceded risks.

†*Groupement des Assureurs Automobiles (Québec's automobile insurers association).*

The PRR can be viewed as a multi-purpose tool: it was initially conceived to compensate insurers that had been forced to underwrite (and undercharge) risks that had been denied coverage by multiple insurers. One can also view the PRR as a reinsurance plan that allows insurers to cede risks that, although they might have been charged a proper premium, could incur very large claims (fat-tailed loss distributions, e.g., snowbirds travelling to the U.S. where civil responsibility must be covered) and thus add too much volatility to the insurer's book of business. Finally, the PRR can be viewed as a tool to cede those risks that have been undercharged for some reason, be it strategic, legislative, or else. This latter view is the subject of the present document.

In the industry, insurers are thus using what are known as *scoring* systems to identify those undercharged risks, i.e., insureds for which the true mathematical expectation of the claim level is higher than 75% of the gross premium that was charged. Ceding those risks to the PRR is actuarially profitable. Currently, it seems that very few insurers are able to reliably identify parts of their books of business that can be profitably ceded to the PRR. They are thus looking for stronger analytical tools that will allow them to make full and profitable use of Québec's PRR. This document shows how Apstat can, through the use of highly discriminative predictive technologies, help insurers reach that goal.

3 Hypotheses and Methodology

In this section, we describe the hypotheses and the methodology that allowed us to obtain an approximation of the financial benefits that could be obtained from the implementation of predictive models for the selection of the risks to be ceded to the PRR. Numerical results are based on data from a mid-sized automobile insurer. Note that, in Québec, bodily injuries related to car accidents are state-covered by the *Société de l'assurance automobile du Québec* (SAAQ).

In order to accept the transfer of a risk from a member insurer's portfolio to its own, the GAA charges a fee that corresponds to 75% of the gross premium paid by the insured. From the insurer's perspective, the gain of ceding can be measured ex-post as the amount of claims incurred on a risk. The net, actual profit is therefore given as

$$\text{Actual profit} \quad = \quad \text{Claims} - 75\% \times \text{Gross premium}. \tag{1}$$

Note that the GAA considers the gross premium as the one filed by the insurer in its rate plan. Discretionary reductions will be disregarded. In such cases, the gross premium used in equation (1) will actually differ (be higher than) the true charged premium, leading to a lower profit.

The decision to cede a risk must be made at policy inception or renewal. At that time, the gross premium is determined but the claims amount remains unknown. Thus, the insurer must rely on the conditional expectation of these future claims, in order to compute a *projected* profit.

$$\text{Projected profit} \;=\; E[\text{Claims}] - 75\% \times \text{Gross premium}. \qquad (2)$$

Once the projected profit has been calculated, we rank the risks in decreasing order of the ratio

$$\frac{\text{Projected profit}}{\text{Gross premium}}$$

and then cede the risks with the highest ratio until the 10% limit is attained or when the projected profit falls below zero.

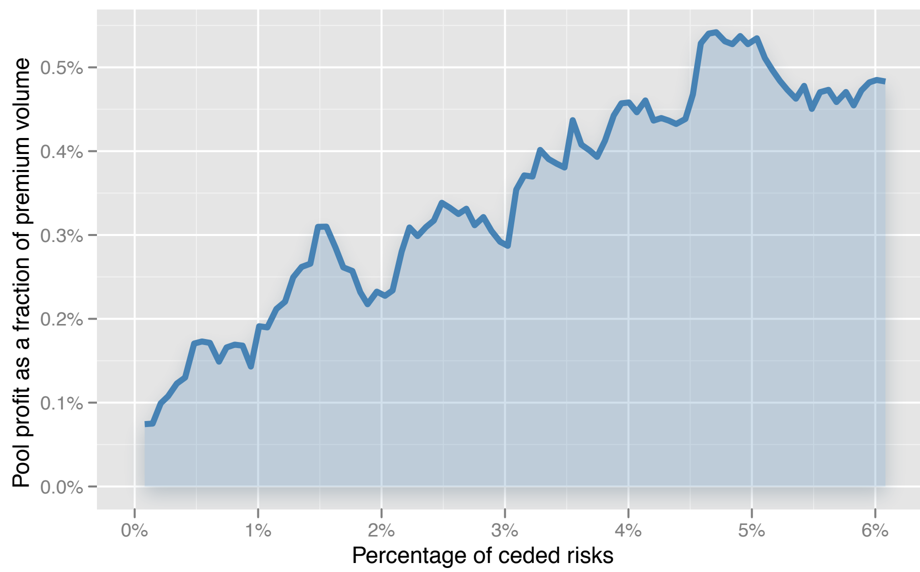
These simple formulas summarize the strategy that we adopted in our PRR experiments. Better precision could be obtained by accounting for interest discounting of cashflow due to timing of transactions between the insurer and the insured or the PRR.

4 Numerical Results

We used a data set of insurance policies representing $165M in earned premiums in the years 2005 through 2007. The overall loss ratio on these policies is 59.4%.

In order to estimate the profitability of our predictive modeling techniques, we implemented a process of *sequential validation* whereby model parameters are estimated using a *training set* of policies. These model parameters are used to determine which policies belonging to a disjoint subsequent *test set* should be ceded. Numerous comparative experiments have lead us to choose a training set length of 12 months and a test set length of 1 month. A lag of 3 months separates both sets. For example, in order to decide which policies to cede in January 2005, we used policies and claims information from October 2003 to September 2004 incl. This process of

► **Figure 1.** PRR *profitability as a function of the percentage of ceded risks.*



parameter estimation (*training*) and policy cessions (*testing*) was repeated for every month from January 2005 to December 2007 incl., i.e., 36 times.

Figure 1 illustrates the PRR profitability as a function of the percentage of the book that has been ceded. If the insurer chooses to cede anywhere between 4 and 6% of its book of business, the profit is always above 0.4% of the volume of business. For a mid-sized insurer with a volume of $100M, this represents $400M of annual savings.

5 Conclusion

According to our experiments, our predictive modeling techniques allow for the reliable and profitable detection of undercharged risks. Our results show that an insurer with a 100M$ book of business in Québec's automobile insurance market can expect gains in the order of $400K, annually.

We believe that recent advances in machine learning technologies such as those based on Neural Networks will soon enter the insurance market just as GLM*s have now become popular. These are powerful modeling technologies essentially because they allow

Generalized Linear Model.

to automate the detection of high-level nonlinear dependencies between explanatory variables. In the long run, the first companies to implement these new technologies will be the ones to benefit the most from them.