SOCIETY OF ACTUARIES®

Article from

Reinsurance News

July 2018

Issue 91

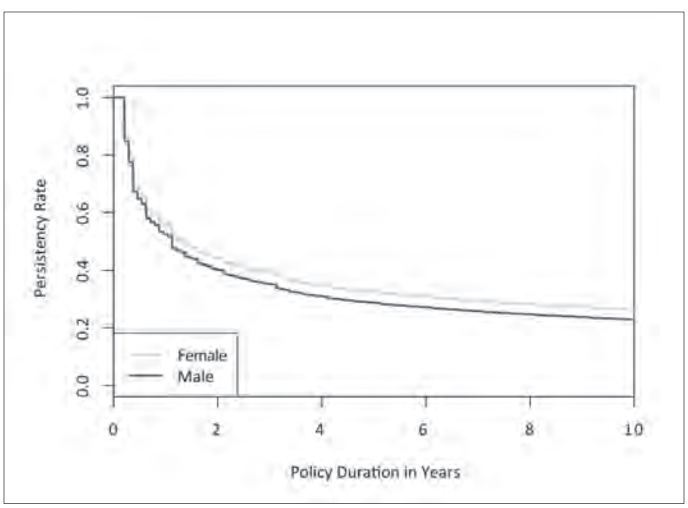
Survival, Persistency And SaleS[1]

By Kai Kaufhold

Reinsurers are becoming more and more involved in the entire value chain of insurance. And rightly so: With the techniques we have developed to understand risk and profitability, actuaries can make a valuable contribution to improving performance at the front end of the insurance business. In this article we explore a case study, in which we applied survival models to analyze the persistency of a life company in Asia and develop a strategy for actively managing the company's sales force.

Besides predictive modeling in life insurance using survival model techniques, there is another theme pervading this three-part series of articles: communicating actuarial concepts and results. The point I was making in my light-hearted introduction to survival models in the first article of this series was that the underlying concepts are very intuitive. Survival models are such a natural fit to the problems of life insurance that they can be a powerful tool for conveying results to non-actuaries—despite the fact that some interesting math is involved in using them.

Figure 1
Kaplan-Meier Curves for Persistency by Gender

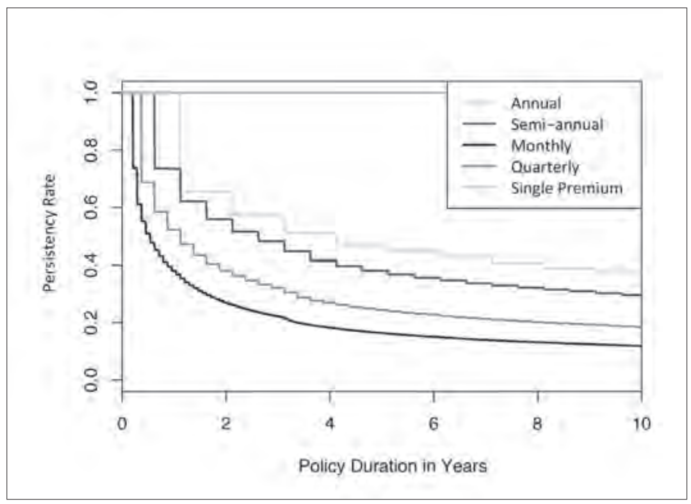


WHY ARE OUR LAPSES SO HIGH?

My team was asked by a life company in Asia to analyze their policy lapses and find out whether we could make any recommendations for improving persistency. The first order of work—after data cleaning, validation and preparation—as with any data analytics project, was to visualize the experience in a way that we could discuss it with management. Enter the tried and trusted Kaplan-Meier curves, a non-parametric method of displaying survival curves which is ubiquitous in data science, clinical research, life sciences, you name it. (See Figure 1.)

One look at the chart in Figure 1 showed the company's CEO what he needed to know: within the first year 40 percent of his business had dropped off the books, and after the second policy year, he had only around 40 percent of policies left. The next chart started to explain why. (See Figure 2.)

Figure 2
Kaplan-Meier Curves for Persistency by Premium Mode



While policies with an annual premium model stuck around for a year and then dropped by 35 percent, monthly premium policies just lasted the premium holiday and then their persistency started dropping like stones. Once you get the hang of using Kaplan-Meier curves, it is just a matter of stepping through all the likely risk factors and identifying which have the most impact. The next thing looked at was premium amount bands, given the financial impact of premiums. You can easily spot the concentration risk when the top 40 percent of policies make for 80 percent of the premium volume (see Figure 3).

Hopefully the higher premium bands had lower lapse rates ... No such luck! The top four premium deciles had higher than average lapses, as Kaplan and Meier show us in Figure 4. What

Figure 3
Policies by Premium (Deciles)

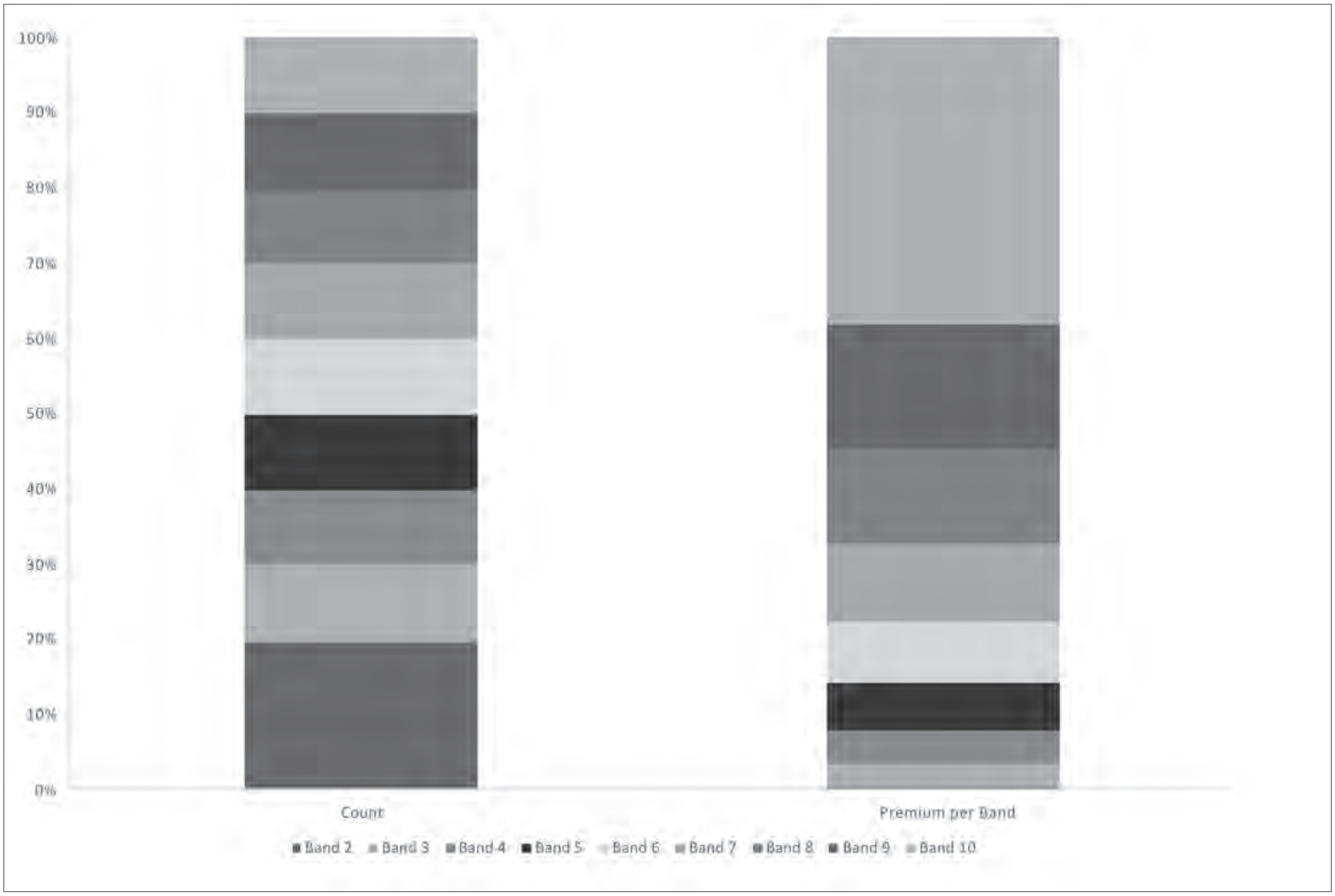


Figure 4
Kaplan-Meier Curves for Persistency by Premium Band

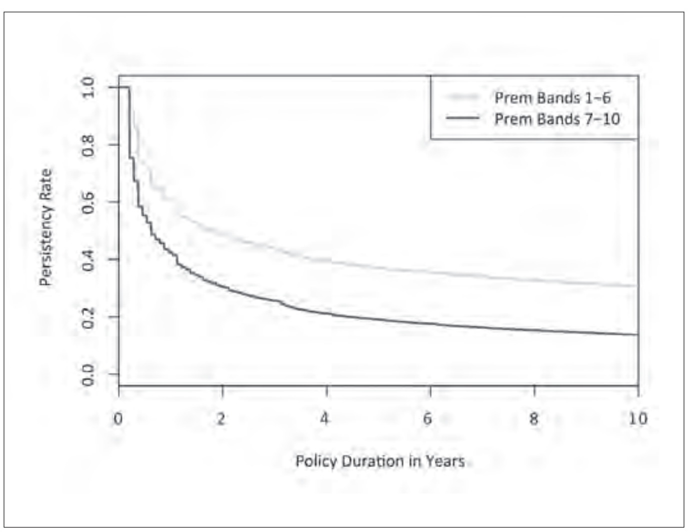
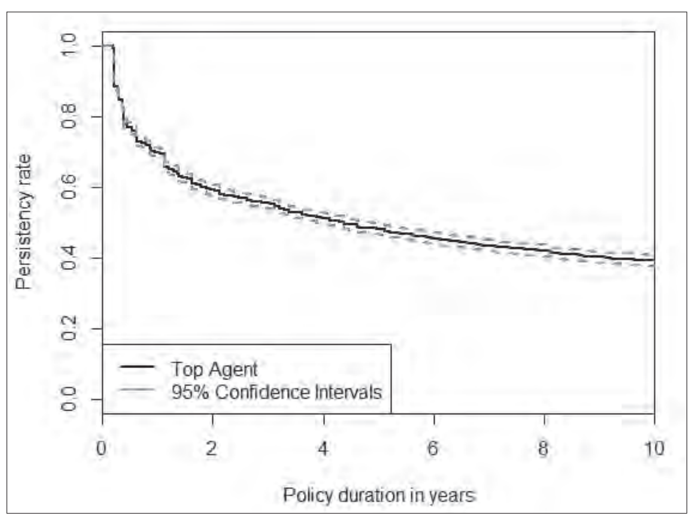
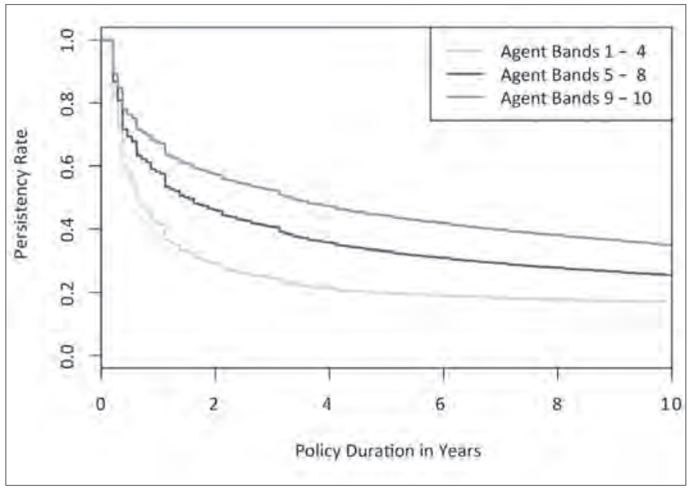


Figure 5
Kaplan-Meier Curves for Persistency for the Agent with Highest Premium Volume



to do next? Well, usually persistency has something to do with the insurance agents who sell the policies. So we took a look at the persistency patterns for individual agents. And finally we were on to something. In Figure 5 (pg. 41), we encouter the company's top sales agent with the highest overall premium production. This agent had a much higher persistency for the policies he had sold. After two years, nearly 60 percent of policies were still left, which is more than twice the average persistency. We could tell that this was a reliable result, because you can easily[2] generate confidence bands for Kaplan-Meier curves, and ours were really tight.

Figure 6
Kaplan-Meier Curves for Persistency by Agent Premium Band



If you now cluster the agents into groups with similar persistency patterns, you find out that you can categorize them by the amount of premium volume which they contributed to the overall book of business. We found three distinct groups of agents, grouped by total amount of business sold. As you can see in Figure 6 there are clear differences between the three groups of agents. The difference in performance of these three groups was so large that this was likely the key to unlocking the persistency riddle.

Figure 7
Crude Hazard Rates for Male Lapses with Log-linear Trendline

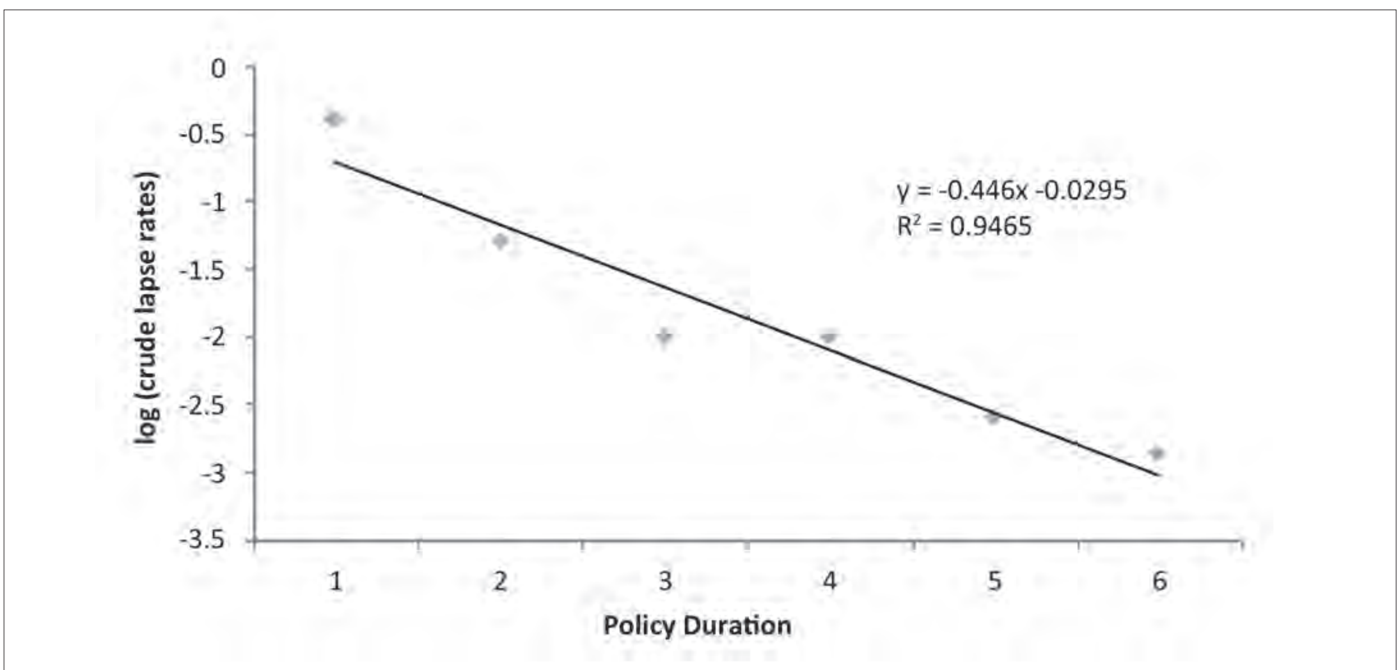
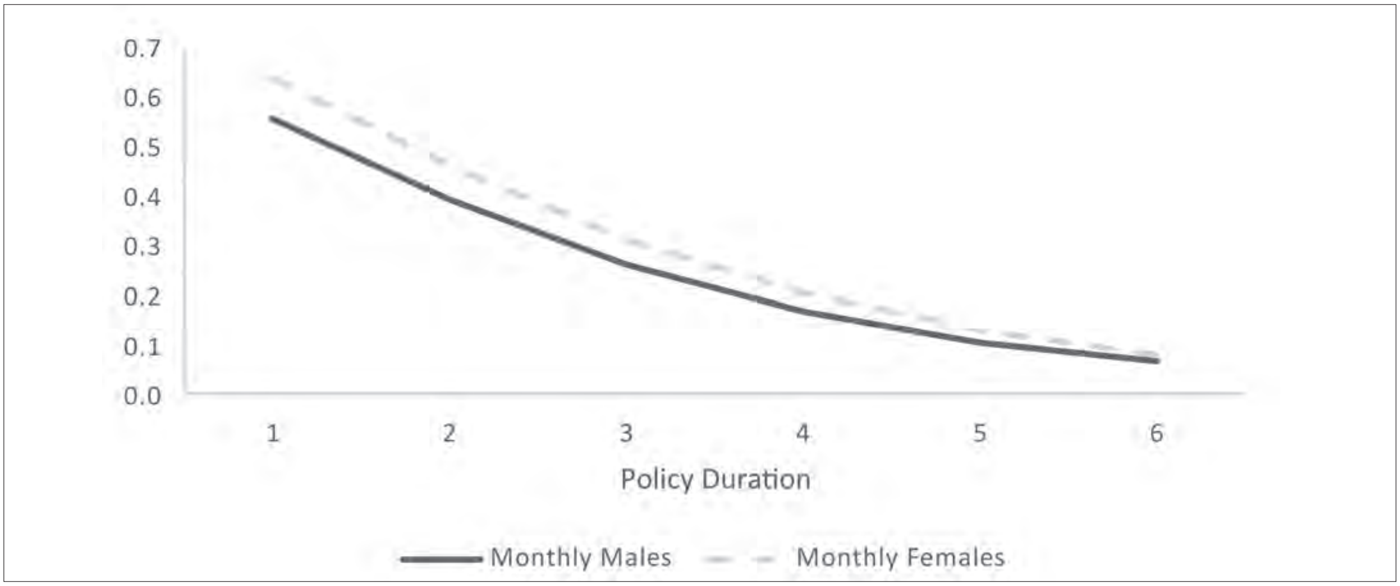


Figure 8
Lapse Hazard by Policy Duration Using a Log-linear Model



However, in order to quantify the impact of these different risk factors which we had identified, gender, premium payment mode and agent group, it was necessary to depart from Kaplan-Meier curves, because they give only a univariate view of the risk. We needed to find out what the impact of each risk factor was in the presence of all others. So we decided to model the persistency using a parametric hazard rate function, which is the basis of parametric survival models explained in part 1 of this series in the previous issue of *Reinsurance News*. The choice of hazard rate function was fairly simple. The chart in Figure 7 suggests that on a logarithmic scale, a straight line already accounts for a large portion of the shape, because the R^2 statistic is high at nearly 95 percent.

We can parametrise a straight line using only two parameters, and by letting the parameters vary by risk factor (gender, premium mode, agent amount band) we can measure the impact of each risk factor exactly. As it turns out, it was worthwhile doing the multivariate analysis. In Figure 8 we see that if we control for the premium model (here: monthly premium payment), males end up showing lower lapse rates than females, despite the fact that the Kaplan-Meier curves in Figure 1 suggested just the opposite. We can go even further: Using the simple survival model, we can not only diagnose which different risk groups display different persistency results, but we are also able to predict what the impact would be of changing the incentives for different agents. By training the agents in the intermediate category to behave in a similar way to the top agents, we can predict how many million rupies the company wins from higher profits due to higher persistency for this group. We cannot disclose the figure here, but what we can say is that the prospective gain easily merits the cost of enhanced training of the agency sales force and collecting more data on the agents themselves.

Now that we have a statistical model which describes the persistency rates and allows us to predict the financial outcome of interventions, it will also make sense to run simulations to quantify the volatility of this book of business. But that is the topic of part three of this series that will appear in the next issue of *Reinsurance News*. Stay tuned if you want to find out why Australian disability income business is so volatile. And if you are interested in the math behind Kaplan-Meier curves and parametric survival models, why not take a look at the newly published book on mortality modeling by Angus MacDonald, Stephen Richards and Iain Currie?[3] ■

Kai Kaufhold, Aktuar DAV, is partner, Prediction Consulting and Longevity with NMG. He can be contacted at *kai.kaufhold@nmg-group.com*.

ENDNOTES

1 SPASS is the German word for fun.

2 All Kaplan-Meier curves shown here were generated using the R package *survival* with a single line of code.

3 Angus S. MacDonald, Stephen J. Richards and Iain D. Currie (2018) *Modelling Mortality with Actuarial Applications*, Cambridge University Press, April 2018.