

GLM Pricing and Machine Learning Rivals or Allies

5th General Insurance Conference

Singapore

30-31 May 2013

Contents

1. Introduction
2. Pricing Best Practice
3. GLM
4. Machine Learning
5. Case Studies (not included in this version)

Q & A

1 Background

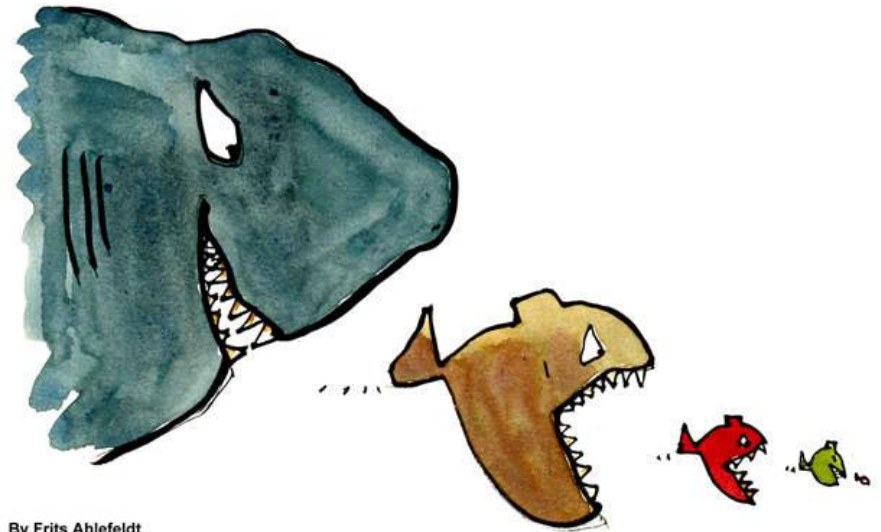


Some Insurers.....

Are happy doing what they have always done.....



Others are happy being
better and bigger than
some.....



By Frits Ahlefeldt

But Today Predictive Analytics Can Give You A Clear Vision of Where You Are and Light The Path Ahead.



Our Path in the Future is Clear

- Increased competition in the marketplaces
- Greater demands for:
 - Better technology
 - Convenience (do it all on a Smartphone)
 - Process Integration
- Data expansion is exponential, but tools are hard and require experts



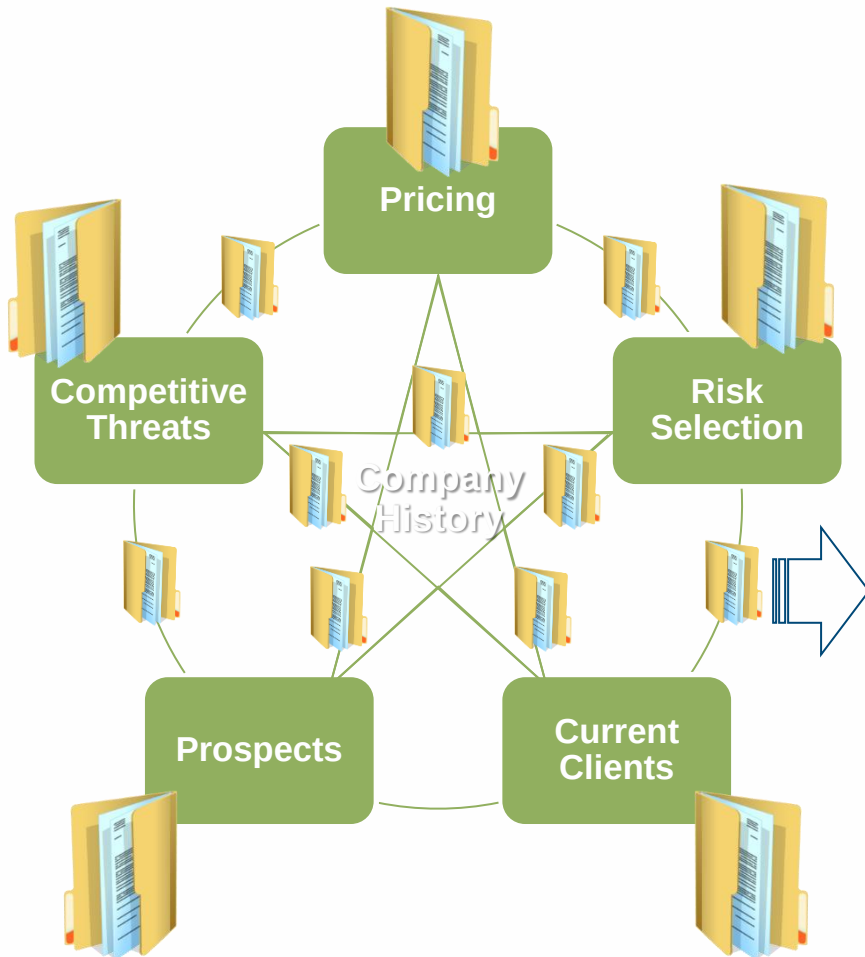
Understanding Predictive Analytics

- **“Predictive analytics** is business intelligence technology that produces a predictive score for each customer or other organizational element. Assigning these predictive scores is the job of a predictive model which has, in turn, been trained over your data.”
- **“Enterprise data** is a priceless strategic asset because it represents the aggregate experience of an organization, the very history of its interactions with customers.”
- **“With predictive analytics**, the enterprise learns from its cumulative experience (data), and takes action to apply what's been learned.”

Although most may not comprehend all the high-level mathematics of predictive analytics, it is not “black magic” but rather a quantification and expansion of your intuition about your organization’s experience.

Source: “Seven Reasons You Need Predictive Analytics Today”, Eric Siegel, PhD, Prediction Impact, Inc.

Key Areas of Business Interaction



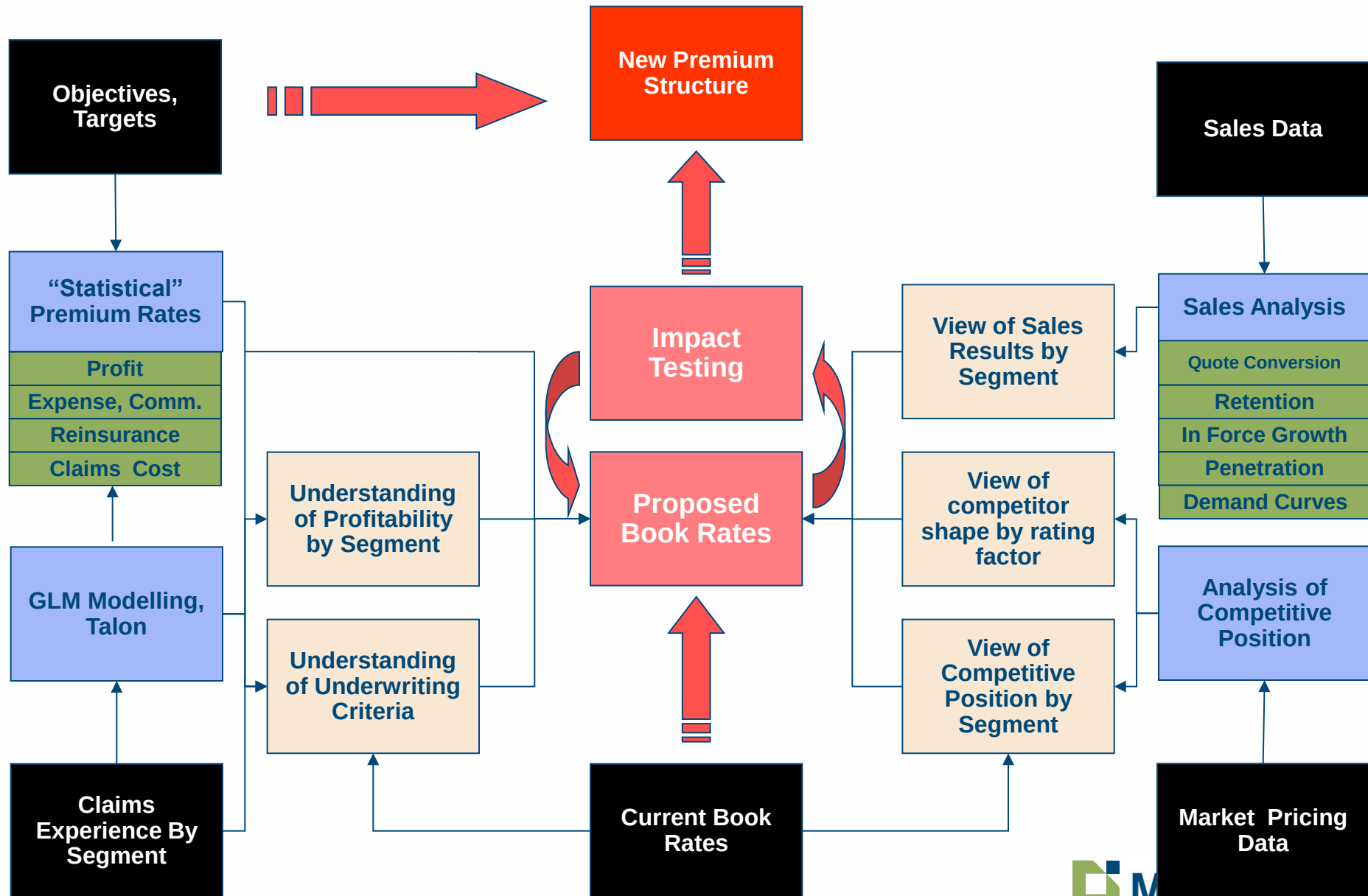
Knowledge gained results in competitive advantage through:

- Greater client satisfaction and retention
- Better risk selection
- Granular, targeted pricing
- More effective marketing

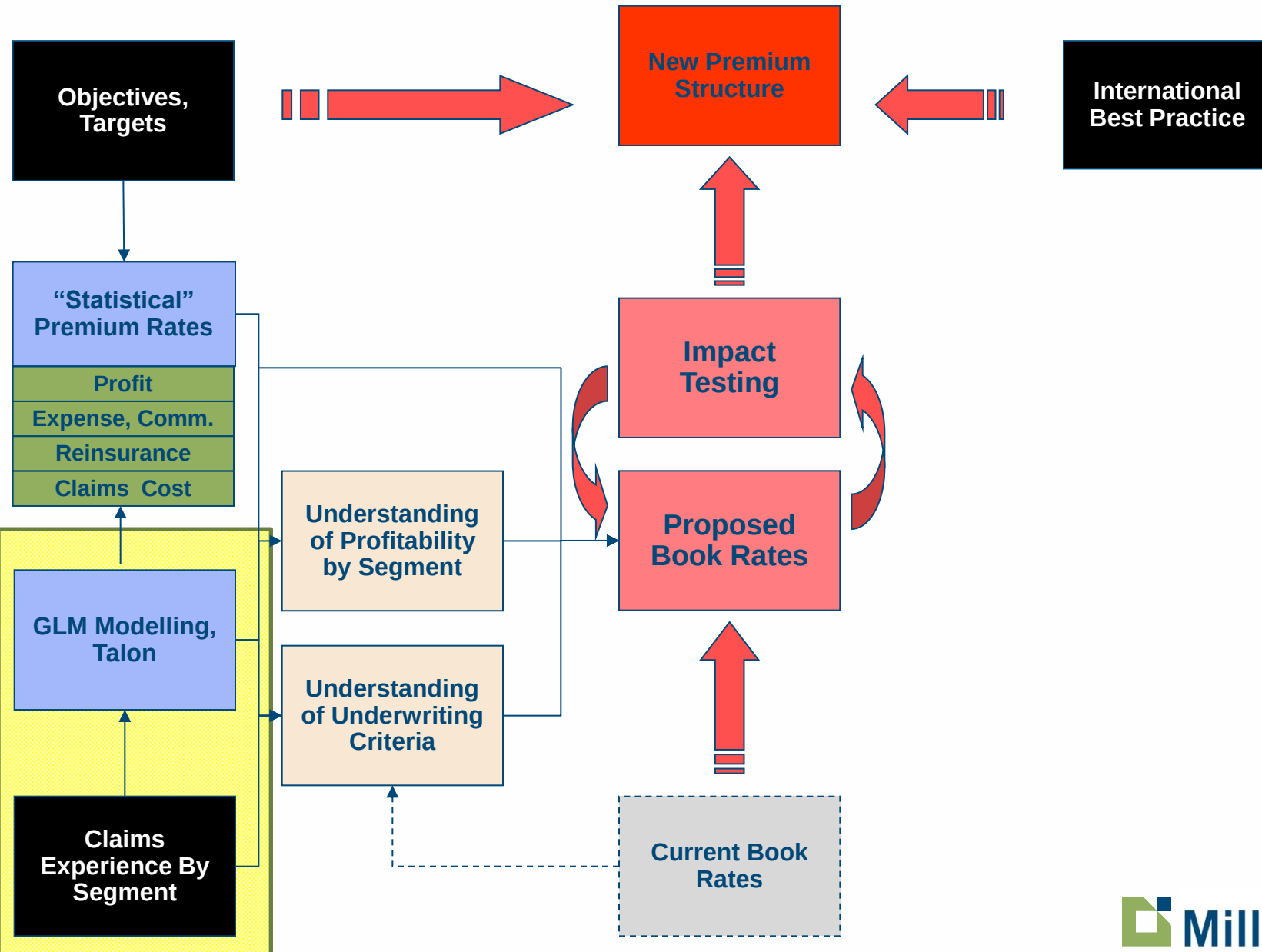
2 Pricing Best Practice



A best practice pricing review process



Claims Cost Modelling



Vehicle Groups

- Perhaps the one of the most challenging aspects of modelling motor lines
- Strong linkage between the underwriting and pricing teams
 - In some cases there may be sufficient data to determine the vehicle grouping based on claims experience
 - In many cases, there is not claims data to credibly assign vehicle models to a group.
 - More commonly the vehicle grouping will blend actuarial and underwriting input. There are various ways of doing this.

3 GLM

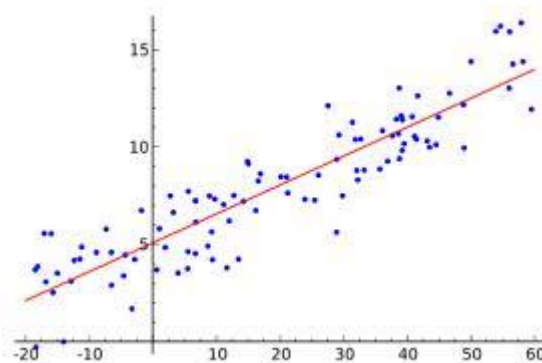


GLM: revealing the “true” risk shape

- Linear Model

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \varepsilon$$

ε = random error

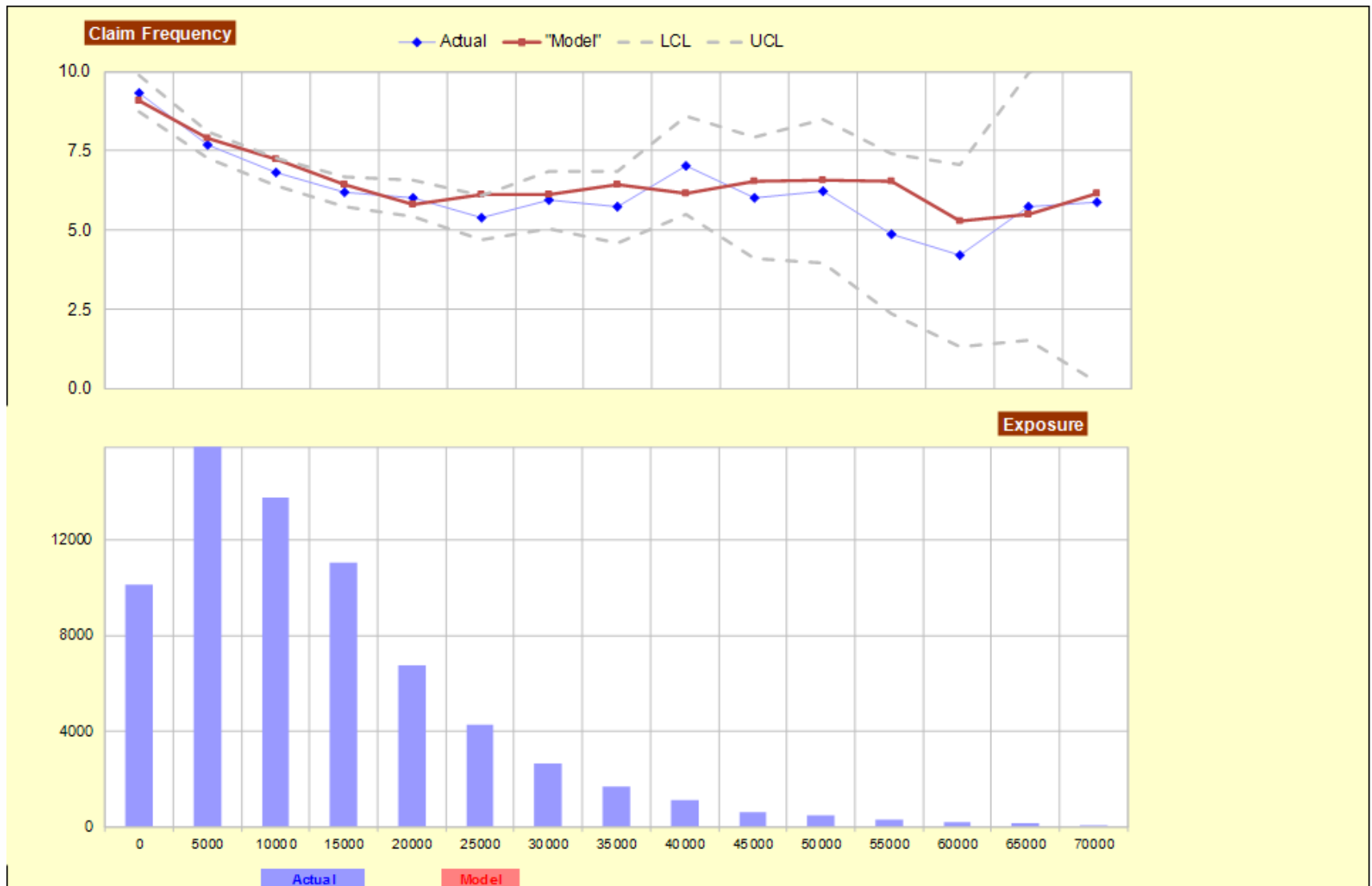


- Deals with all factors simultaneously
- Quantifies the “pure” contribution for each factor
- “Generalised” models enable more flexibility in model structure

GLM: what is a good model?

- Consistent over time and withstand random sampling tests
- Strikes a balance between fitting well and over-fitting the data
- Various measures and tests can be done using a combination of :
 - AIC/BIC
 - Residual plots
 - Cramer's V - test the correlation of two categorical factors
 - Deviance
 - Chi-square
 - Confidence interval of fitted values for each factor
 - Gini

GLM: revealing the risk shape



4 Machine Learning



Machine Learning

What is Machine Learning

- *“a branch of artificial intelligence, is a scientific discipline concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data, such as from sensor data or databases.” Wikipedia*

Addresses the fundamental questions directly

- Where are we making money?
- Where are we losing money?
- Can we be confident?

Identifies risk segments that are credible and produce consistent results from year to year

- Iterative, artificial intelligence process
- User defines the degree of credibility within segments

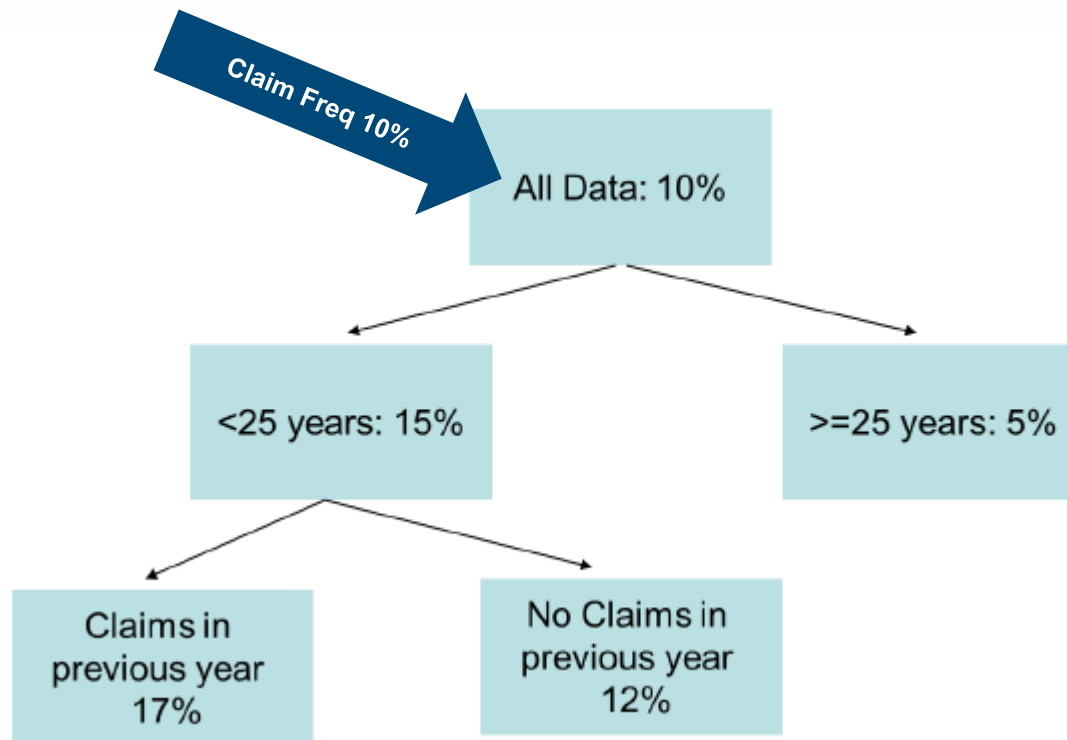
Results are only as good as the algorithm

- A good algorithm will maximise the number of segments identified

What is Machine Learning

Segmentation Using Machine Learning

- A decision tree is a recursive partitioning algorithm
- Build a segmentation of the data by first splitting the data into two sub-regions in a way which maximizes the difference in the target variable.
- This splitting approach is re-applied to the sub-regions over several iterations until we have built up a whole “tree”.
- Stopping rule to ensure we have a sensible number of segments



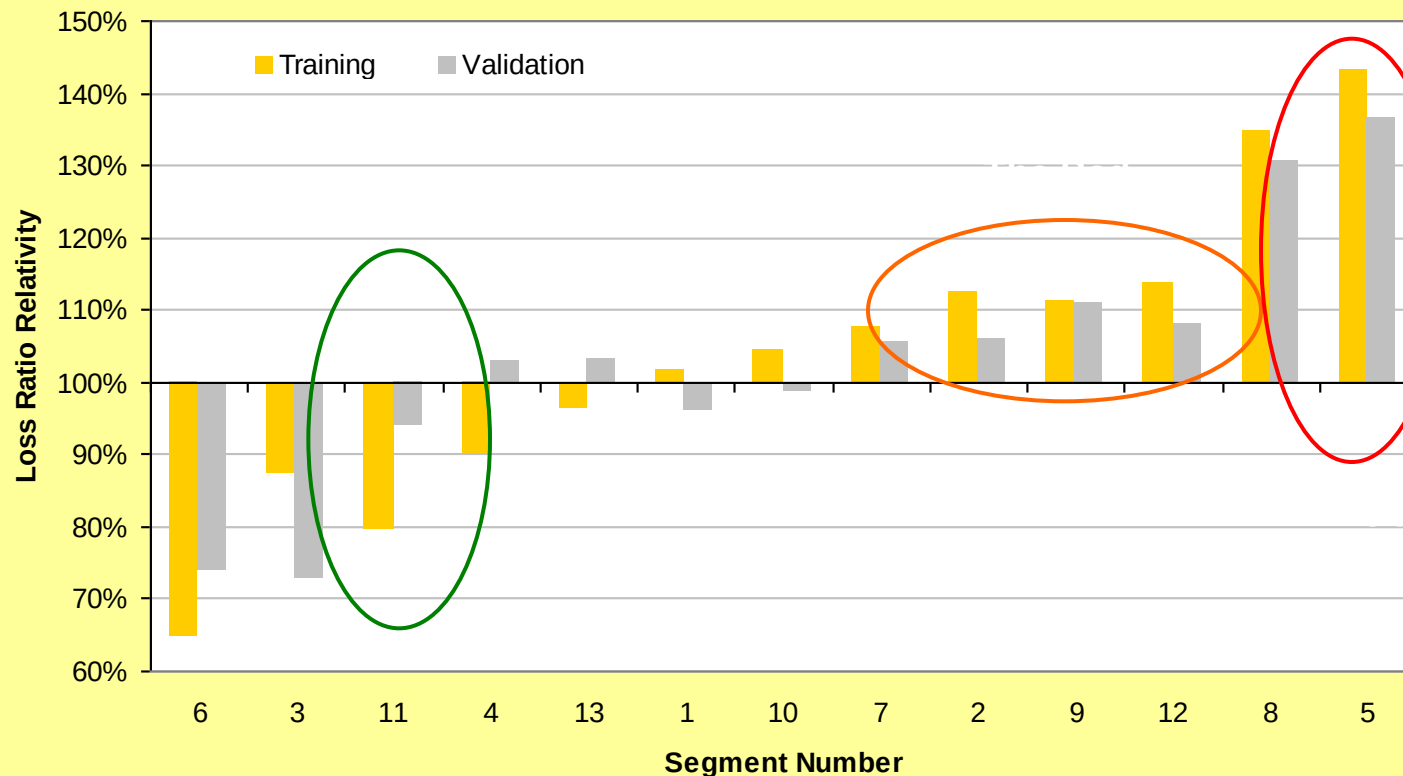
Technology for Insurers – Computing Power



Use Talon to show where fit can be improved

Aim of the Model is to replicate the mechanics of the real world which lead to the observed outcomes

- Training & validation is key
- Best practice involves the use of machine learning algorithms to analyse the residuals from the GLM, and thereby improving the overall model fit



Talon output includes Risk Segment Descriptions

- Segments may be multi-way combinations of variables that can subsequently be built into the model

Segment			% of Exposure	Loss Ratio Relativity	Vehicle Type	Location	Duration	Excess
Good	1st	#6	11%	68%	✓		✓	✓
	2nd	#3	9%	83%	✓		✓	✓
<....7 segments in between....>								
Bad	3rd	#9	6%	111%	✓	✓	✓	✓
	4th	#12	12%	112%	✓	✓		
Ugly	2nd	#8	8%	134%	✓	✓		✓
	1st	#5	6%	141%	✓	✓	✓	

Recap

- ✓ Need for Predictive Modelling is Today
- ✓ GLM is a robust pricing approach

BUT

- ✓ Machine Learning will
 - ✓ Tackle GLM's **shortcomings**
 - ✓ **Identify** critical hidden “gems” and “pitfalls”
 - ✓ **Speed up** the model build process **systematically**

Rivals or Allies?
Why not take Best of Both Worlds

Q & A ?

Pang Chye FIA

Principal & Consulting Actuary, Non-Life

Tel + 852 2598 1516

Mobile + 852 9262 8805

Email pang.chye@milliman.com

Jack Jiang FIAA

Director & Consulting Actuary, Non-Life

Tel + 852 2598 1516

Mobile + 852 9127 3174

Email jack.jiang@milliman.com

