SOCIETY OF ACTUARIES®

SOA Predictive Analytics Seminar – Malaysia

27 Aug. 2018 | Kuala Lumpur, Malaysia

Session 2

Predictive Analytics in Policyholder Behavior

Eileen Burns, FSA, MAAA
David Wang, FSA, FIA, MAAA

Predictive Analytics in Policyholder Behavior

Eileen Burns, FSA, MAAA
Principal & Consulting Actuary
Milliman Inc.

David Wang, FIA, FSA, MAAA
Principal & Consulting Actuary
Milliman Inc.

27th August 2018

SOCIETY OF ACTUARIES®

Eileen Burns, FSA, MAAA
Principal & Consulting Actuary
Seattle
Eileen.Burns@milliman.com

Education and Qualifications

University of Washington,
Quantitative Ecology and Resource Management (2008 - 2011)
Masters

Lawrence University (1998 - 2002)
BA, Mathematics

Current responsibilities

- Principal on Milliman's data analytics team
- Product manager for Recon, a Milliman predictive analytics and data product targeted at enhancing experience analysis
- Vice-chair of SOA Predictive Analytics and Futurism section

David Wang, FIA, FSA, MAAA
Principal & Consulting Actuary
Seattle
David.Wang@milliman.com

Education and Qualifications

University of California at Berkeley,
HAAS School of Business (2005 - 2006)
MFE, Financial Engineering

Nanyang Technological University (1994 - 1998)
B. Business

Current responsibilities

- Co-leads Milliman's team specializing in applying data analytics to assist the life and annuity industry in the United States.
- Co-leads Milliman life consulting practice in Seattle

Agenda

- Current state in life and annuity
- Examples of where predictive analytics helps
- Implication on assumption setting process
- Interesting applications

Current State in Life and Annuity

SOCIETY OF ACTUARIES®

What is Predictive Analytics and Predictive Modeling

Predictive analytics uses many techniques from data mining, statistics, modeling, machine learning, and artificial intelligence to analyze current data to make predictions about future. **Predictive modeling** is a process used in **predictive** analytics to create a statistical **model** of future behavior.(*Google Search*)

Policyholder Behavior Modeling: Progression of States

Traditional State

- Traditional one-way actuarial techniques to estimate behavior by age/duration and limited number of other characteristics using experience where it exists
- Primarily macro-oriented… little use of detailed information on policyholder characteristics
- Judgment and guesswork where experience does not exist

Predictive Modeling State

- Next-generation experience studies using policyholder longitudinal data.
- Use much wider set of explanatory variables readily available to company
  - Internal data (Product features, distribution channel, policyholder and contract characteristics)
  - Macro data (Economic data, financial market conditions)
- More sophisticated analysis techniques to find non-linear, multivariate effects, complex interactions

Big Data State

- Employ external consumer/financial/health and big/unstructured data sources in a full Predictive Analytics framework.
- Develop individual policyholder profiles

Applications of Predictive Analytics in Life and Annuity

Behavior modeling	Beyond
• Lapse (mostly annuity) • Withdrawal (variable annuity) • Post level term shock lapse • Premium persistency • Fund transfer	• Predictive underwriting • Target market and lead generation • Producer recruitment and retention • Cross-sell

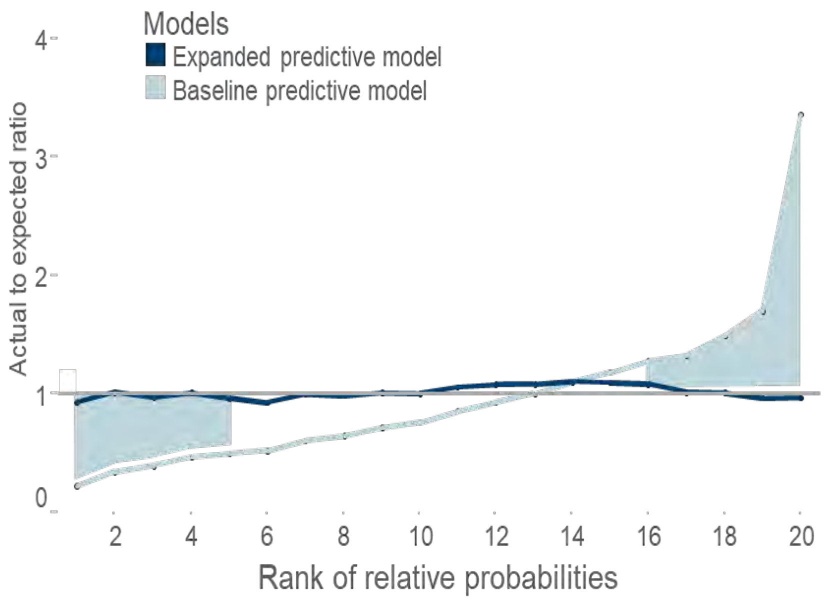
Actuarial ⟷ Data Analytics

Examples of Where Predictive Analytics Helps

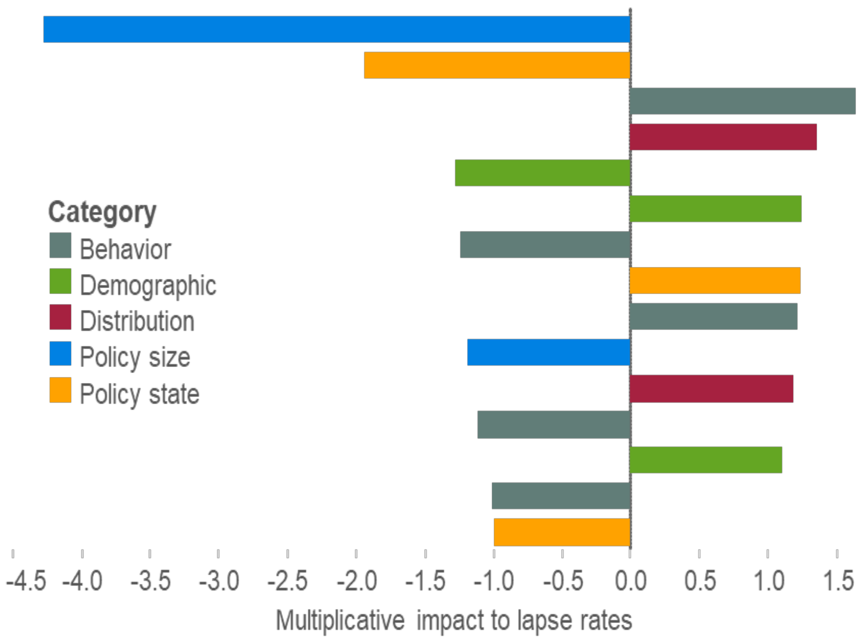
SOCIETY OF ACTUARIES®

Improve Predictions

Overall Improvement in Predictions

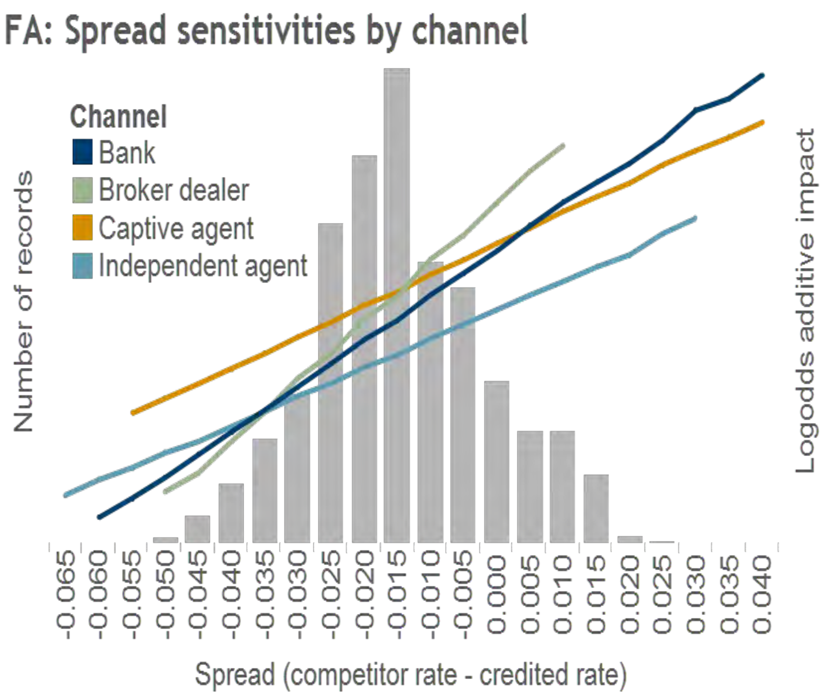


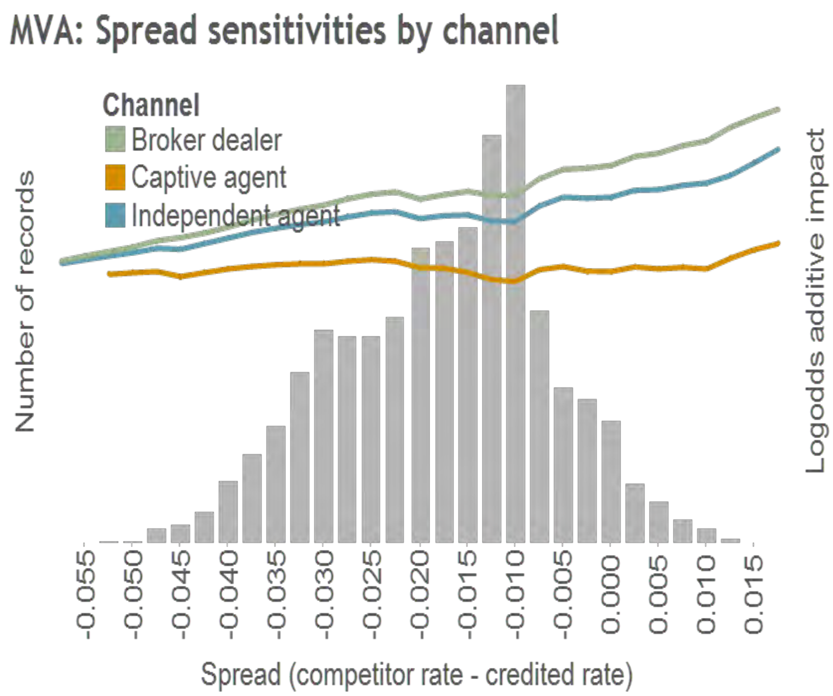
Relative impact from predictors



Test Hypothesis and Answer Question

- Is there a difference in sensitivity to crediting spread among distribution channels?
- Does the MVA effectively eliminate sensitivity to crediting spread?





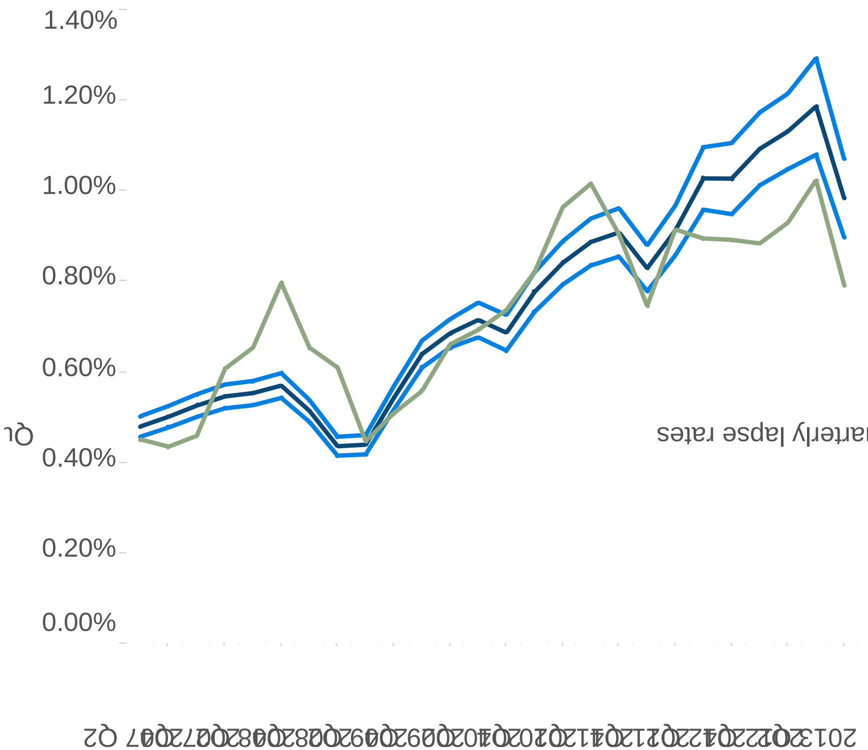
Identify drivers

- Previous behavior – e.g. **withdrawal behavior**
- People – demographics and distribution channel
- Product design – MVA, surrender charge structure, guaranteed minimum
- Macroeconomics – market rates, **unemployment**

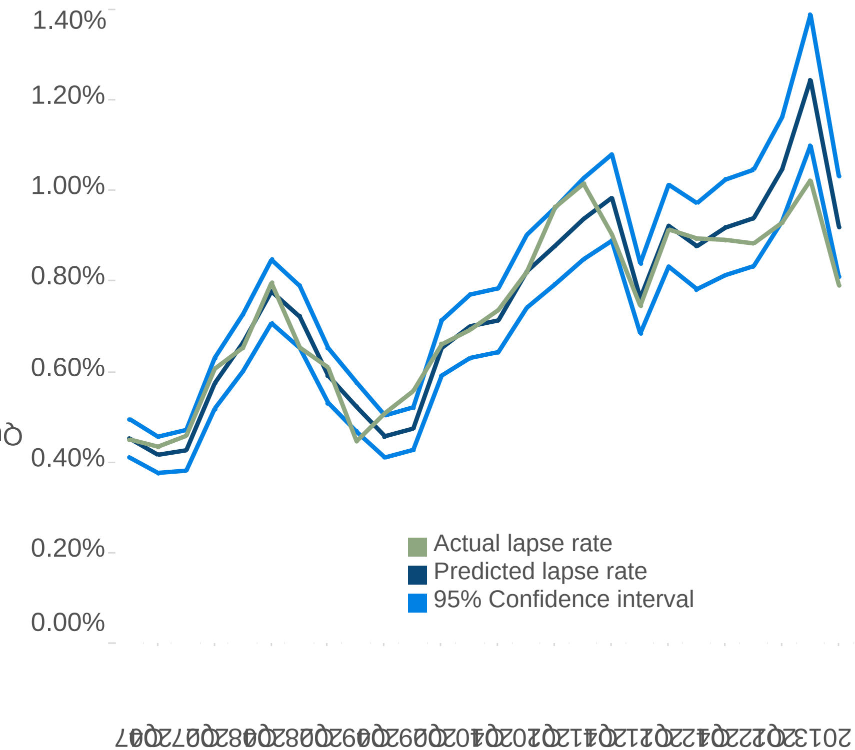
Confidence Intervals

Model predictions and confidence bands versus actual experience

Baseline model



Full model

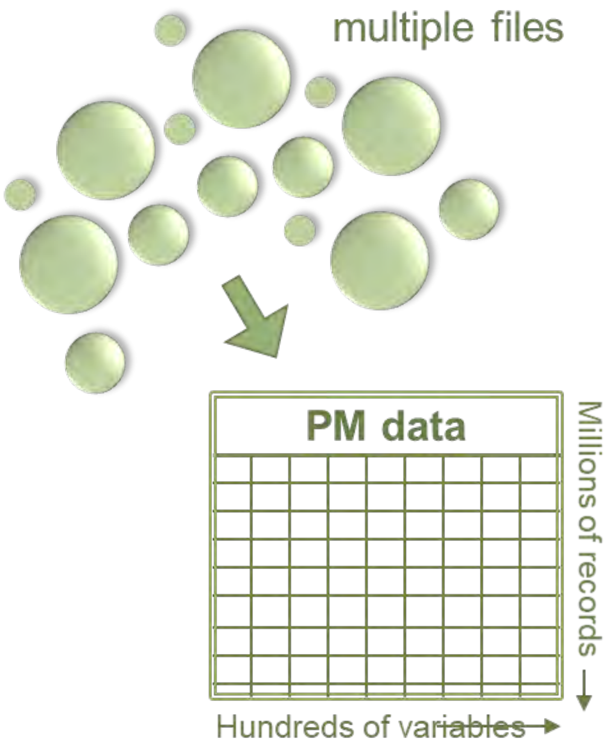


Implication on Assumption Setting Process

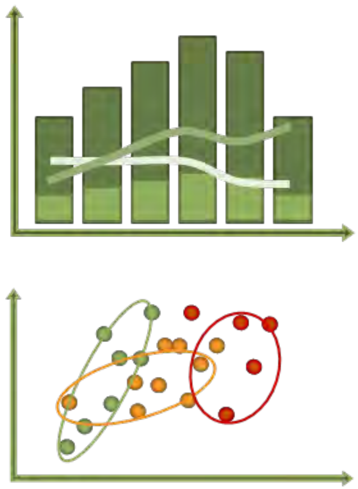
SOCIETY OF ACTUARIES®

Typical predictive modeling process

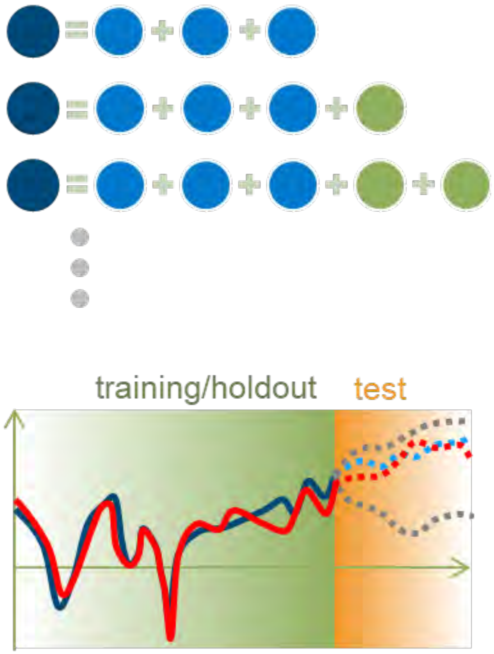
Data Preparation



Data Analysis



Model Building



Era of Big Data has come, but Life Insurers Need to Catch Up!

Challenges the life insurance industry faces

- Little systematic collection and storage of data
- Legacy system inadequate for new data analytics
- Limited data to differentiate customer
- Silos still exist

Data visualization is more than just better pictures

More data, more information, more dimensions, calls for better visualization

Makes traditional date reporting inefficient

Provides guidance and tips on how predictive models should be built

Bring predictive model in assumption setting process

Implementation

Can we model all the predictive drivers in the actuarial cash flow projection?

If not, how do we make compromise and recognize the loss of accuracy.

Communication

How do actuaries convince themselves and management that PM is needed?

How do actuaries communicate model results to senior management?

Assumption Setting

Validation

How is the goodness of fit over different dimensions?

How are we comfortable with confidence intervals?

Domain knowledge is essential to make sense of results.

Control & Governance

Predictive modeling requires new controls & governance. How do we develop appropriate standards?

Who is qualified to review and sign off?

What type of documentation should be retained?

Some Interesting Applications

SOCIETY OF ACTUARIES®

Evaluation of behavioral tail risk

Types of lapse tail risk

Drift

- Risk that best estimate lapse rates vary under different market conditions
- Captured by a dynamic lapse component

Diffusion

- Risk that estimates of the entire lapse function are off
- Captured by simulation of lapse behaviour using predictive model

Extreme Event

- Risk that some unprecedented events may impact lapse in an extreme way
- Resort to some manner of judgement call

Lapse behavior simulation

$$logodds \sim a + \ b_1 * Variable1 + \ b_2 * Variable2 + \ b_3 * Variable3 + \ \varepsilon$$

Model Assumptions:

- Each coefficient $\boldsymbol{b_x}$ is normally distributed
- The error term $\boldsymbol{\varepsilon}$ is normally distributed with a mean of 0
- Correlation between each predictive $\boldsymbol{VariableX}$ can be given by a correlation matrix
- The standard deviation of $\boldsymbol{\varepsilon}$ denoted by $\boldsymbol{\Omega}$ can be determined from the correlation matrix using numerical methods such as Cholesky decomposition

Lapse behavior simulation – Determine best estimate

- Consider the following model, where the only predictive variable considered is In-The-Moneyness

$$logodds \sim a + \; b_1 * \boldsymbol{ITM} + \; \varepsilon$$

- After fitting your experience to the model, the following best estimate calibration is attained:

$$logodds = \mathbf{0.5} + (-\mathbf{2}) * \boldsymbol{ITM} + \mathbf{0}$$

Resulting best estimate lapse rate (*p*):

ITM	*p*
225%	1.8%
175%	4.7%
125%	11.9%
75%	26.9%
25%	50.0%

Lapse behavior simulation – Simulating the risk of model misestimation

- Alternatively, we can simulate lapse rates by allowing the coefficients to vary according to their standard deviation, assuming a multivariate normal distribution

$$logodds = \mathbf{0.5} + N_1(0,\Omega) + (-\mathbf{2} + N_2(0,\Omega)) * ITM$$

Resulting simulated lapse rate (p):

Best Estimate		
ε(i)	ITM	p
0	225%	1.8%
0	175%	4.7%
0	125%	11.9%
0	75%	26.9%
0	25%	50.0%

ε = {-0.2, -0.1}		
ε	ITM	P
-0.2, -0.1	225%	1.2%
-0.2, -0.1	175%	3.3%
-0.2, -0.1	125%	8.9%
-0.2, -0.1	75%	21.8%
-0.2, -0.1	25%	44.4%

ε = {0.2, 0.1}		
ε	ITM	p
0.2, 0.1	225%	2.8%
0.2, 0.1	175%	6.8%
0.2, 0.1	125%	15.8%
0.2, 0.1	75%	32.6%
0.2, 0.1	25%	55.6%

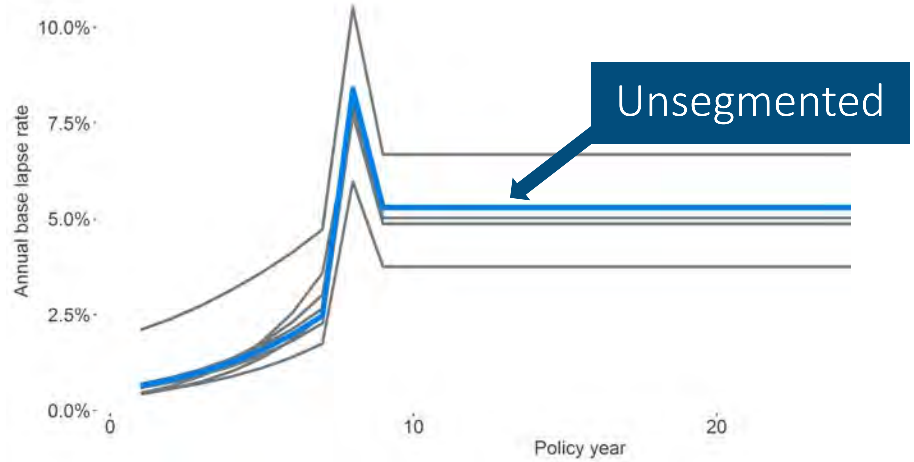
Customer segmentation

Identify segments of policyholders

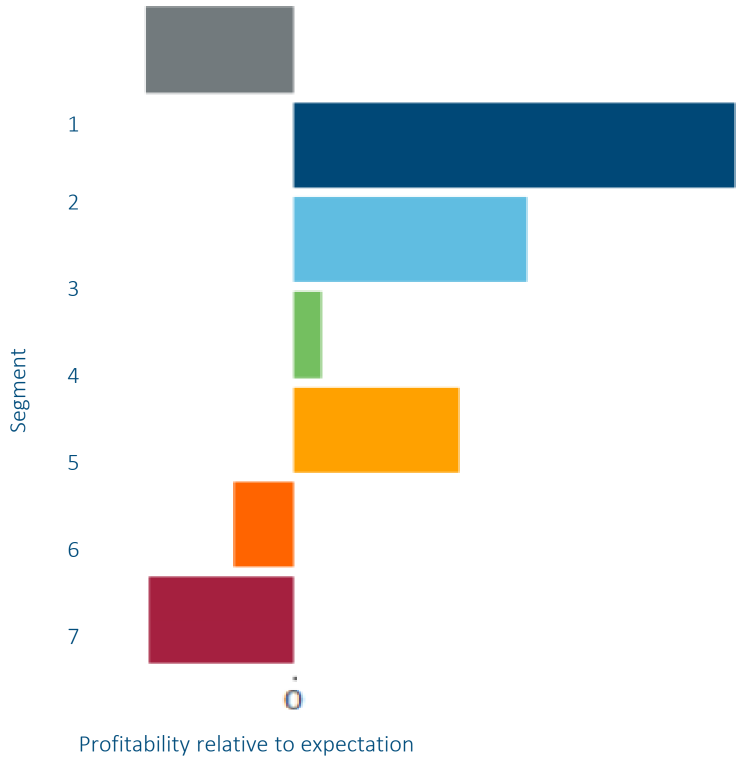
Data-driven segments identify policyholders likely to behave in similar ways

- Number and defining characteristics of segments will be specific to the particular dataset
- Likely defining values for segments include credit score, income, home value, home mortgage loan-to-value, etc.

Segment specific behavior modeling reveals how people use insurance differently



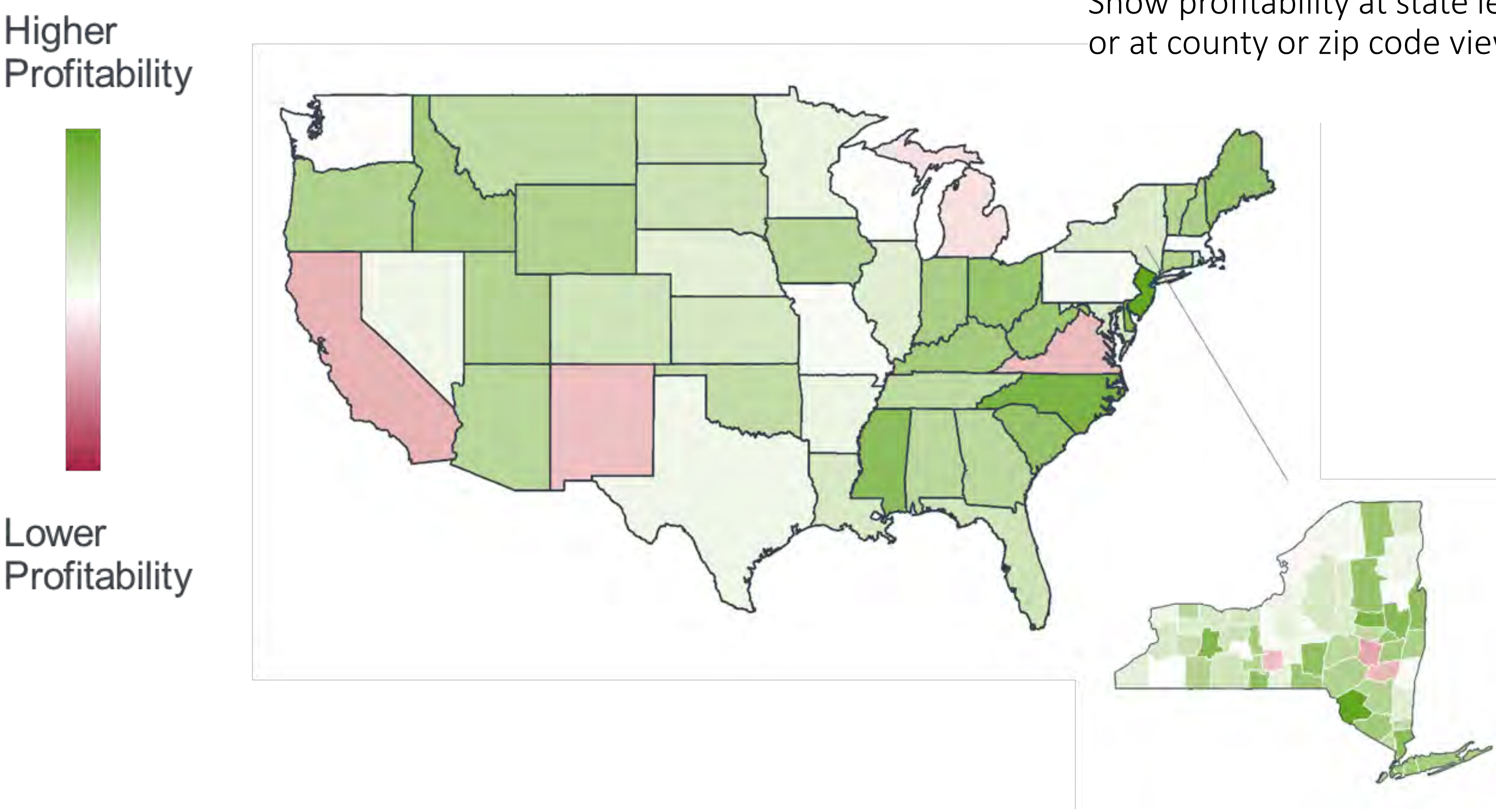
Differentiation between policyholder behavior and corresponding profitability



- Plots show profitability differences driven purely by behavioral difference due to belonging to different segments.
- Help identify groups of people whose needs are not served properly by current product offerings and identify need for new products

A new perspective of product profitability

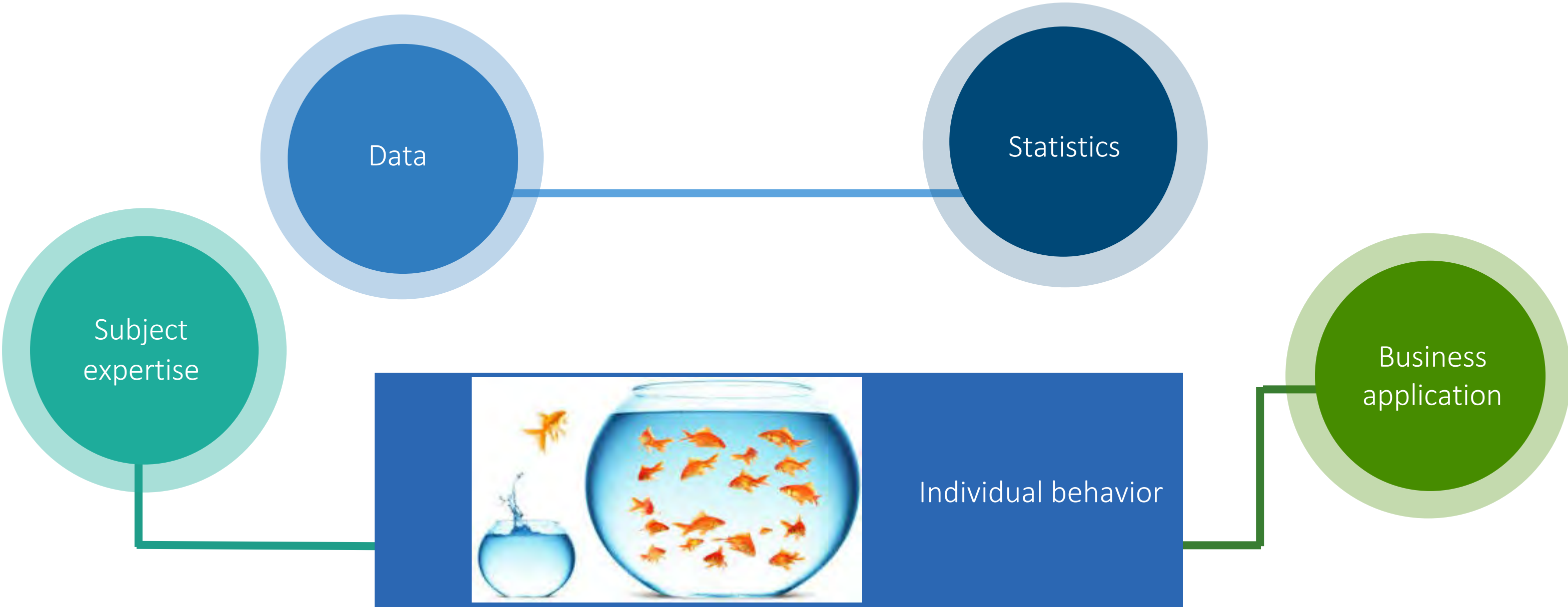
Show profitability at state level, or at county or zip code view



Final thoughts

Goal and elements of predictive analytics in policyholder behavior

To predict (individual) policyholder behavior by applying rigorous statistical techniques to large amounts of data under the guided framework designed by subject experts



SOCIETY OF ACTUARIES®