# Ensembles – Could we actually file one?

**Christopher Cooksey, FCAS, MAAA, CSPA**

**Head Actuary, Data & Analytics**

**5 November 2020**

GUIDEWIRE

# Agenda

What is an Ensemble?

Ensemble as Feature Creation

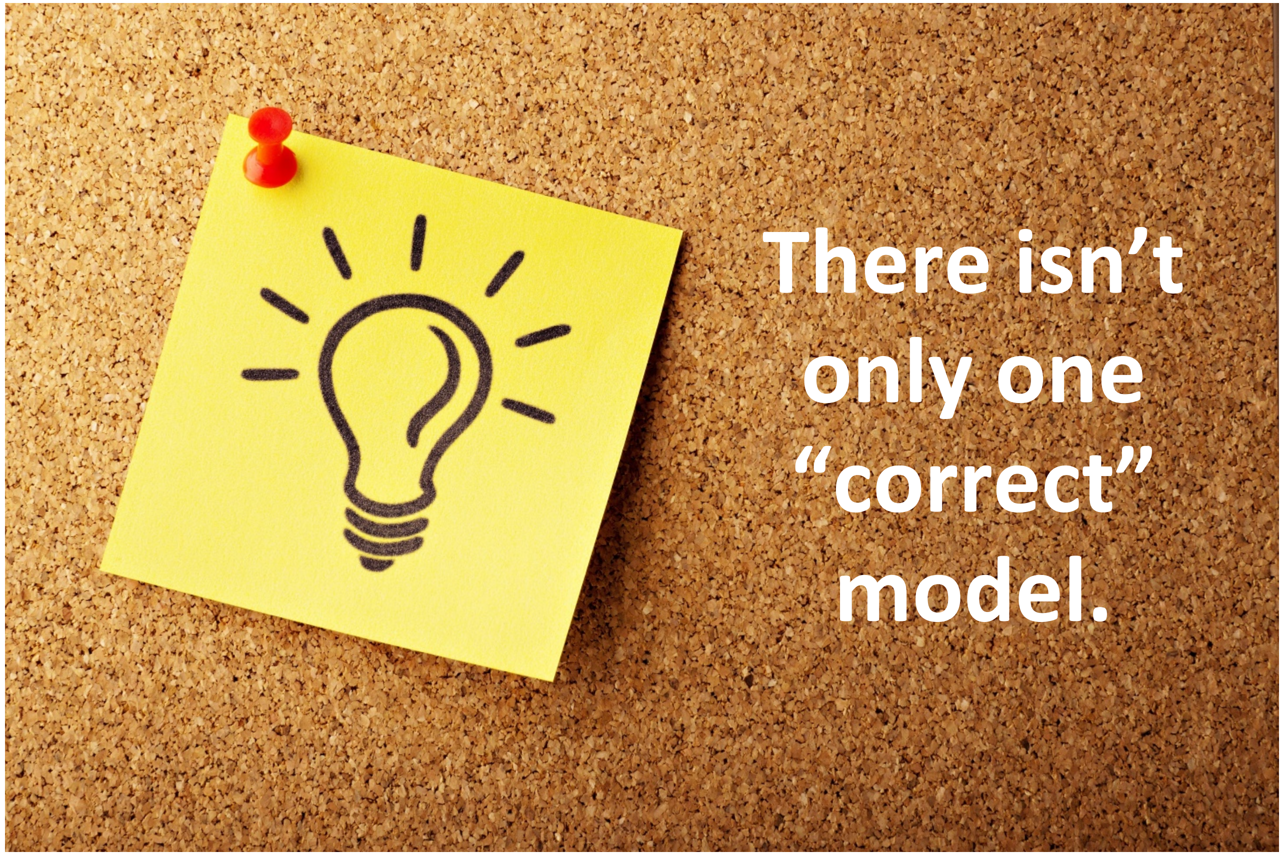
Ensemble as Rate Enhancement

Will Regulators accept Ensembles?

# What is an Ensemble?

# What is an Ensemble?

**An approach to building predictive models.**




**What if two reasonably accurate models tell you something different from each other?**

**Is one right and one wrong?**

# Simple Actuarial Ensemble

**Selecting a statewide loss trend…**

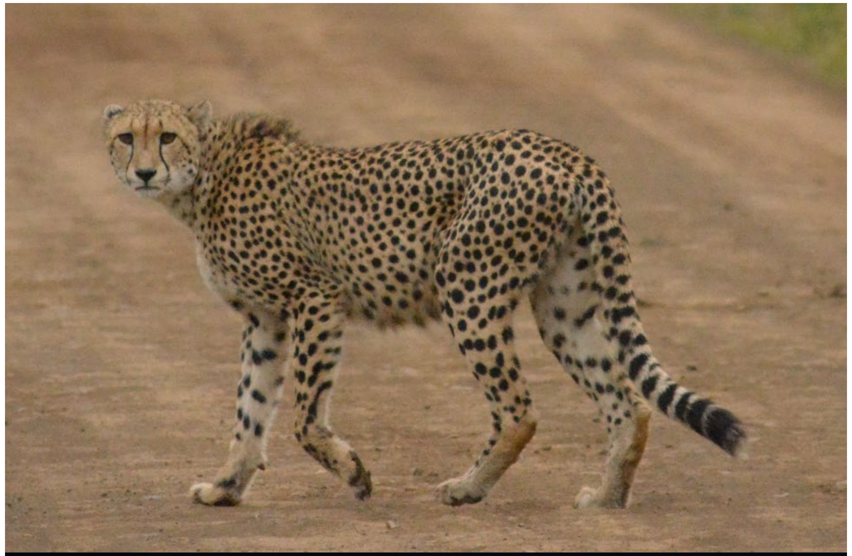
- State Trend – low bias, high variance
- National Trend – high bias, low variance

$$\boldsymbol{Selected\ Trend = x(State\ Trend) + (1 - x)(National\ Trend)}$$

**Credibility weighting is a simple ensemble.**

# What makes a good Ensemble algorithm?

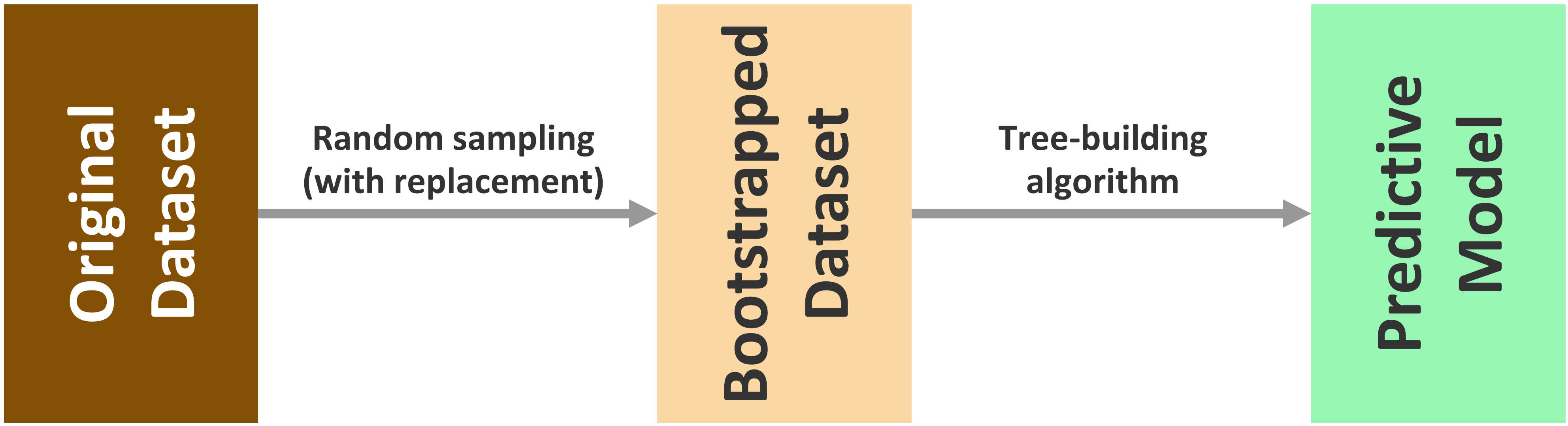
Component models ("learners") must be quick to build and quick to process in production environments.



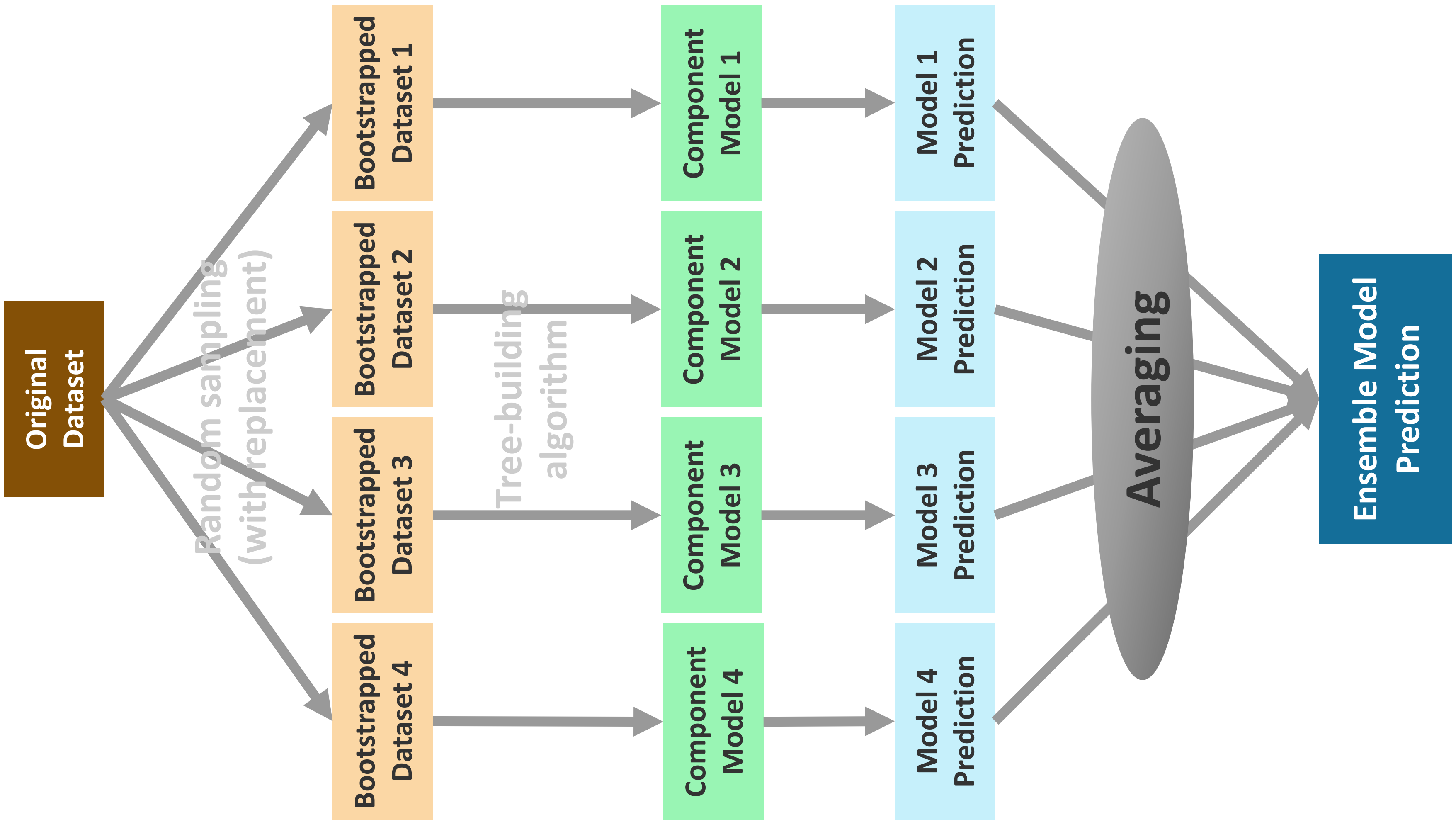
Component models must make *different* errors from each other.

They must be...decoupled.

# Example:  Bagging (bootstrap aggregation)



# Example: Bagging (bootstrap aggregation)

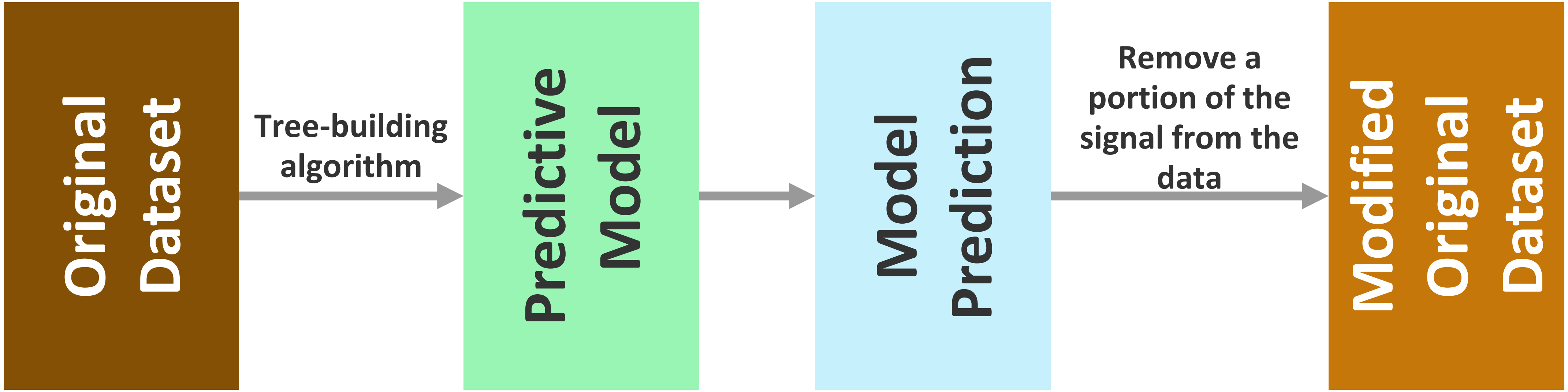



Random Forest uses bagging.

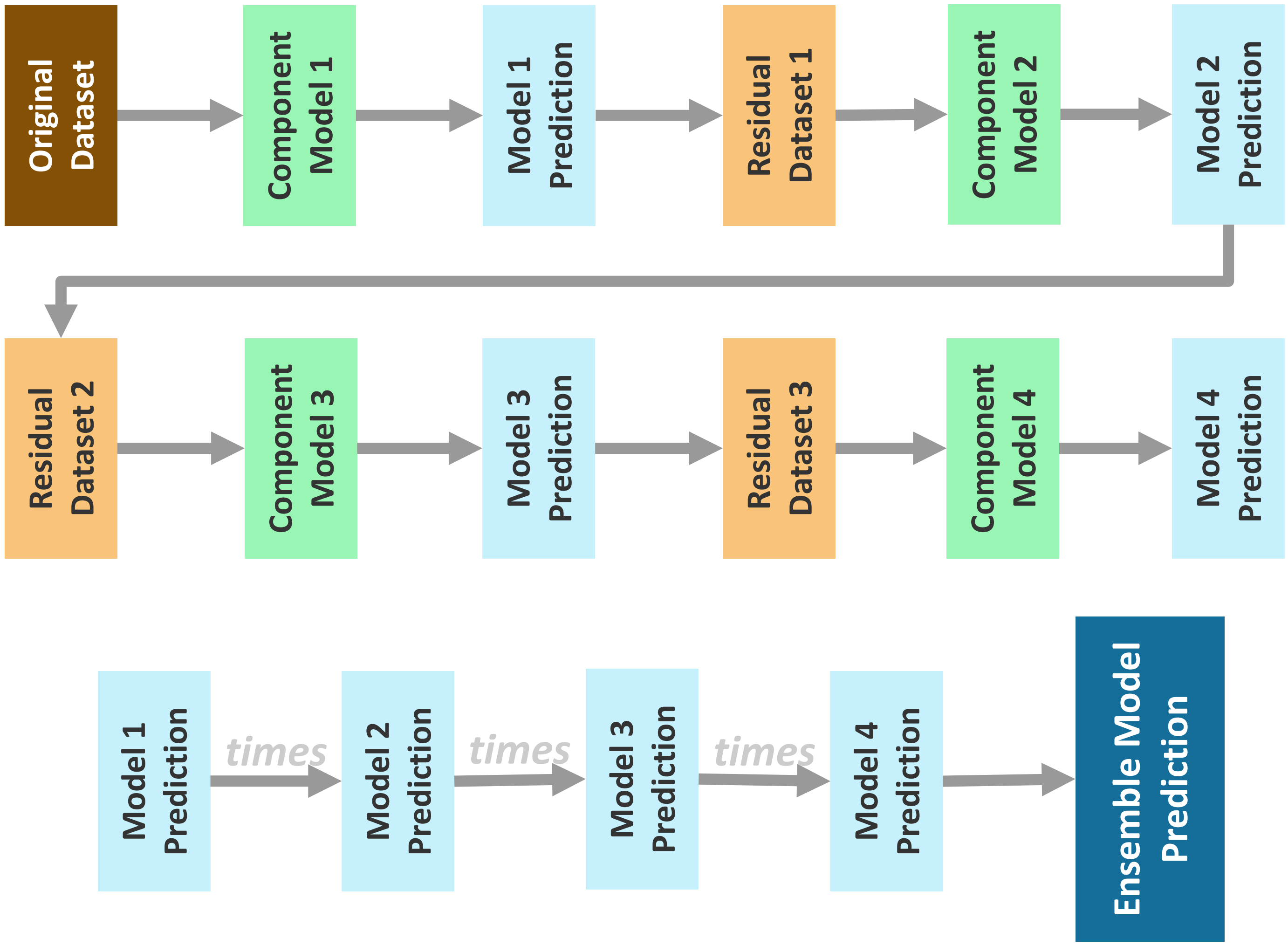
It uses CART to build the trees, modified to include more randomness.

“Averaging” here is general – could be weighted, mode, mean…

# Example: Boosting



# Example: Boosting




**Residual modeling is the defining characteristic of boosting methods.**

**Often called gradient boosting because it is a form of gradient descent.**

**An example is GBM (Gradient Boosting Machine). XGBoost is a common implementation package (R, Python, others).**

# Ensembles as Feature Creation

# Can Ensembles be used in Pricing?

# Ensembles as a Rating Plan

**An ensemble model is not likely to be the support for your entire rating plan.**

- **(Base rate * relativities) long precedes GLMs.**
- **Complexity is likely a fatal issue when considering ensembles of trees as the entirety of the rating plan.**

**Could ensemble models be *part* of your rating plan?**

# Feature Creation

**Using a defined set of predictors to combine and create a new predictor.**

- **Geographic information ->  Territory Definitions**
- **Vehicle information ->  Rate Symbols**
- **Credit Information ->  Credit Score**

**Isolates the effects in one place.  Also allows indirectly predictive fields to play a role in rating.**

# Creating Territory Definitions

Trees can be effective alternatives to the traditional smoothing and clustering approach.

Need to have geographical descriptions to use as predictors – for example, zip-code information.

- Demographic info – population density, etc.
- Weather info – number of snow days, etc.

# Creating Territory Definitions

## Smoothing and Clustering – current approach

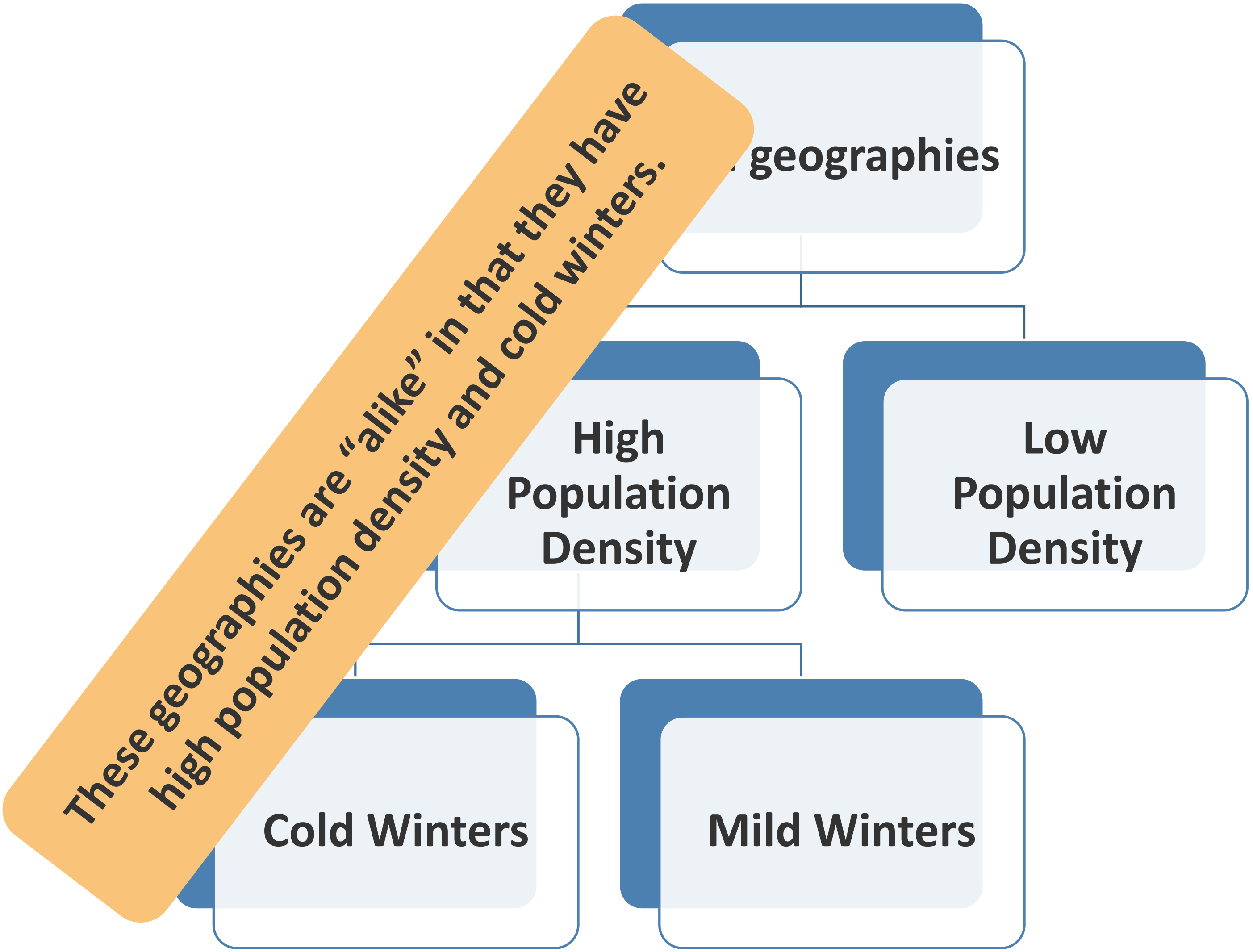
- Experience in credible geographies is attributed to others to the extent that they are "alike".
- "Likeness" is defined by geographical distance.
- User-provided smoothing parameter.
- Cluster smoothed information into groups.

# Creating Territory Definitions

## Tree-based Ensemble – alternative approach

- Experience in credible geographies is attributed to others to the extent that they are “alike”.
- “Likeness” is defined by having similar characteristics – being in the same branch of the tree.

# Creating Territory Definitions




**This tree structure is defined by the geographies where you have business.**

**However, it would apply to all geographies, even those where you don’t have business.**

# Creating Territory Definitions

## Tree-based Ensemble – alternative approach

- Experience in credible geographies is attributed to others to the extent that they are “alike”.
- “Likeness” is defined by having similar characteristics – being in the same branch of the tree.
- No user-provided smoothing parameter.
- Cluster ensemble predictions into groups.

# Creating Territory Definitions

## 2011 RPM talk – *Territorial Ratemaking* (Eliade Micu)

Conclusions from that talk – Smoothing/Clustering v Trees

- Similar results overall; similar territories produced.
- Trees provided slightly better lift and fit.
- Trees require less information from the user (distance measure, smoothing parameter, number of clusters).
- Tree-based models can be applied to other states, and the approach can generalize to other problems.

# Creating Rate Symbols

**With "likeness" defined by characteristics, not physical distance, can take the same approach to vehicle information.**



**Need general information about vehicles as predictors – Engine size, vehicle weight, body style, etc.**

# Creating Rate Symbols

## Tree-based Ensemble – approach to Rate Symbols

- Experience in credible vehicle types is attributed to others to the extent that they are "alike".
- "Likeness" is defined by having similar characteristics – being in the same branch of the tree.
- No user-provided smoothing parameter.
- Cluster ensemble prediction (smoothed information) into groups.

# Feature Creation

Useful when you want to compartmentalize effects

- Auto – vehicle effects, geographic effects,…driver effects?
- Commercial – area of operations; company characteristics
- Any line – credit characteristics

# Credit as Feature Creation

**Why create a feature out of credit? Why not let number of late payments, etc. play as individual predictors?**

**Credit gets more scrutiny than other features.**

- **Relating predictors to outcomes in general**
- **Reason codes for the specific**

**Note that ensembles can provide this information, despite their complexity.**

# Ensembles as Rate Enhancement

# Are current rating plans perfect?

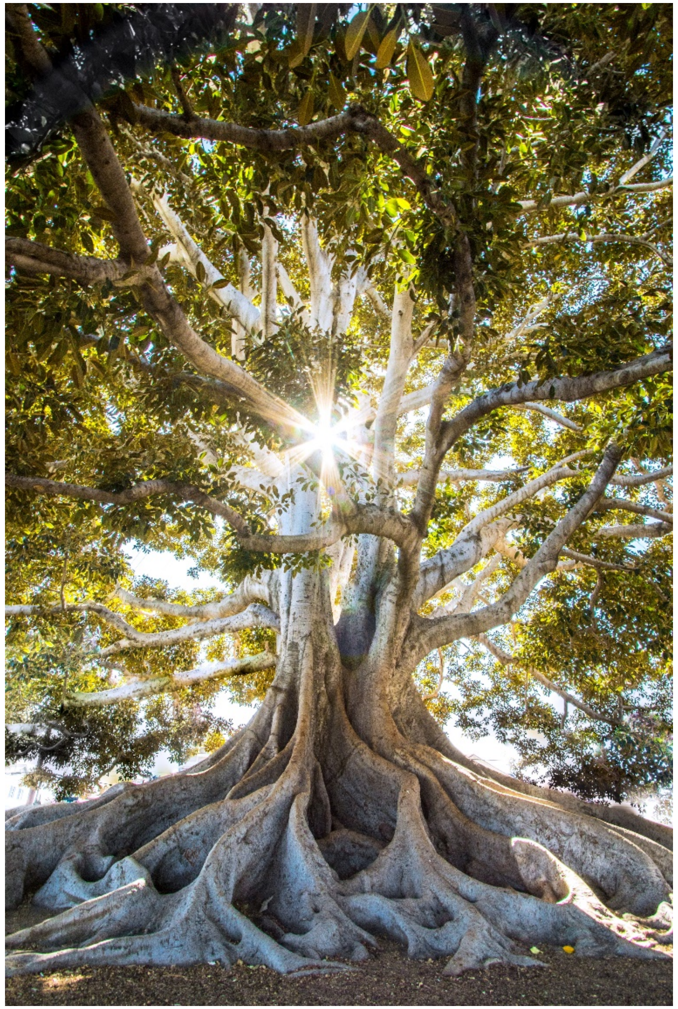

**If perfect, residuals from rating plans should be random, with no discernable signals. To test, try to model the residuals.**

**Turns out, it is easy to create credible, stable models from the residuals of GLM models using ensemble techniques.**

# Why does residual modeling work?

**Probably more to do with the tree-based approach.**

**Trees are complementary to GLMs and focus well on precisely the signal that GLMs have trouble representing.**



# Rate Adjustment Factor

$$GLM\ Pred.\ Pure\ Prem = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n$$

$$Residual\ Loss\ Ratio = \frac{Actual\ Loss}{Pred.\ Pure\ Prem}$$

**Any systemic variation in the residual LR can be implemented as an adjustment to the predicted pure premium from the GLM.**

**In other words, just one more relativity.**

## Rate Adjustment Factors

**In some sense, we've combined the idea of feature creation and residual modeling to create a single adjustment factor that can be used as part of a rating plan.**

**To simplify things, express your residual loss ratio model with a 3-digit score and associated predictions.**

# Rate Adjustment Factors

| Ensemble Group | Earned Exposures |
|---|---|
| 0-99 | 10% |
| 100-249 | 15% |
| 250-499 | 25% |
| 500-749 | 25% |
| 749-899 | 15% |
| 900-999 | 10% |
| Total | 100% |

# Rate Adjustment Factors

| Ensemble Group | Earned Exposures | Predicted Pure Premium |
|---|---|---|
| 0-99 | 10% | 11% |
| 100-249 | 15% | 16% |
| 250-499 | 25% | 25% |
| 500-749 | 25% | 25% |
| 749-899 | 15% | 14% |
| 900-999 | 10% | 9% |
| Total | 100% | 100% |

# Rate Adjustment Factors

| Ensemble Group | Earned Exposures | Predicted Pure Premium | Actual Incurred Loss |
|---|---|---|---|
| 0-99 | 10% | 11% | 14% |
| 100-249 | 15% | 16% | 18% |
| 250-499 | 25% | 25% | 26% |
| 500-749 | 25% | 25% | 24% |
| 749-899 | 15% | 14% | 12% |
| 900-999 | 10% | 9% | 7% |
| Total | 100% | 100% | 100% |

# Rate Adjustment Factors

| Ensemble Group | Earned Exposures | Predicted Pure Premium | Actual Incurred Loss | Residual Loss Ratio | Residual Loss Ratio Relativity |
|---|---|---|---|---|---|
| 0-99 | 10% | 11% | 14% | 127.3% | 1.273 |
| 100-249 | 15% | 16% | 18% | 112.5% | 1.125 |
| 250-499 | 25% | 25% | 26% | 104.0% | 1.040 |
| 500-749 | 25% | 25% | 24% | 96.0% | 0.960 |
| 749-899 | 15% | 14% | 12% | 78.6% | 0.786 |
| 900-999 | 10% | 9% | 7% | 77.8% | 0.778 |
| Total | 100% | 100% | 100% | 100% | 1.000 |

# Rate Adjustment Factors

**The factor is filed as another rating variable. Actuarial justification comes from the model output.**

| Rate Adjustment Factor Score | Indicated Relativity | Selected Relativity |
|---|---|---|
| 0-99 | 1.273 | 1.20 |
| 100-249 | 1.125 | 1.10 |
| 250-499 | 1.040 | 1.04 |
| 500-749 | 0.960 | 0.96 |
| 749-899 | 0.786 | 0.80 |
| 900-999 | 0.778 | 0.78 |
| Total | 1.000 | 1.000 |

**This approach does not change the GLM indicated relativities.**

**It is an adjustment which improves the accuracy of the GLM predictions.**

# Rate Adjustment Factors – Alternate Approach

The previous approach is entirely internally consistent. The model is built on the GLM residuals and so is applied straight-forwardly to the GLM prediction.

Another option is to re-fit the GLM with the 3-digit score as an additional predictor and let it solve for new relativities.

With this approach we are specifically viewing the 3-digit score from the residual model as a created feature. This feature can be viewed as field that captures the nonlinear relationships.

# Will Regulators Accept Ensembles?

# The Short Answer

**Yes.  We've seen it many times.**

**Companies are using it as the basis of their territory groupings.**

**Companies are also using it as an adjustment factor to their GLM-based rates.**

# However…

**You still need to concern yourself with having acceptable predictors and, possibly, multiple models across states.**

**It is possible that a regulator will not be comfortable with the model.**

# Transparency is Key

**The nice thing about trees, and therefore ensembles of trees, is that while they can be complex, they are completely *transparent*.**

**Make sure you can print out every split of every tree.**

**Some states need to have "the calculation" on file.**

# Summary

- Ensembles work by combining information from multiple models.
- The most common ensemble techniques are based on decision trees.
- Ensembles of trees are powerful techniques for creating features – territory groups, rate symbols and, potentially, many others.
- Ensembles of trees are powerful techniques for finding signal in the residual of GLM models.  The results can be used as adjustment factors in rating plans.
- You will get questions from regulators, but the questions generally have satisfactory answers.

# Questions?

**Christopher Cooksey, FCAS, MAAA, CSPA**
**Head Actuary, Data and Analytics**

**Guidewire Software**

**ccooksey@guidewire.com**

***All photos from Unsplash.com***

INDIANA FARM BUREAU INSURANCE®

# *Ensembles – Could we actually file one?*

Brian Poole FCAS
Actuarial Data Manager
Indiana Farm Bureau Insurance

March 24 2020

# *Agenda*

- Benefits
- Inputs
- Internal Acceptance
- Implementation
- Filing

# *Benefits of Adding Ensemble Model*

- Interactions missed in rating formula
- Secondary
  - Add new variables
  - Simplify future changes

# Determining Variables

- Start with everything
  - Data quality
  - Regulatory concerns
  - Choice Variables
  - Availability
- Let the analysis do the rest

# *Internal acceptance*

- Input variables
  - Clearly defined
  - Simple is good
  - Univariate analysis
- Lift charts
- Typical risks
- New information vs acceptance

# *Implementation Concerns*

- New variables ➔ Learning curve
- To see or not to see
  - Actionable?
  - Residual model
  - Companion GLM models
- Refreshing Models

# *Filing the Model*

- Identify technique
- Simple description
- Overview of analysis
- Define variables
- Model Results → Rating

# *From Model to Rating Factor*

- Model → loss ratio buckets
- Rank buckets
- Assign scores
- Create score ranges → Rating groups
- Groups → Rating factors

# Compare Training to Validation

**Summary of Aggregate Experience Training Data**
Policy Years 2013-2017 as of 12/31/2017

| (1) Risk Score Band | (2) Earned Exposures | (3) Earned Premium | (4) Case Incurred Losses & ALAE | (5) Claim Count | (6) Case Incurred Loss & ALAE Ratio | Actual Loss Ratio Relativity | Indicated Loss Ratio Relativity |
|---|---|---|---|---|---|---|---|
| 1-99 | 128,645 | 190,010 | 161,067 | 19,448 | 84.8% | 1.699 | 1.618 |
| 100-199 | 128,645 | 187,525 | 123,724 | 16,553 | 66.0% | 1.323 | 1.319 |
| 200-299 | 128,645 | 195,095 | 111,340 | 15,700 | 57.1% | 1.144 | 1.151 |
| 300-399 | 128,645 | 188,359 | 97,836 | 14,370 | 51.9% | 1.041 | 1.050 |
| 400-499 | 128,645 | 182,834 | 88,566 | 12,961 | 48.4% | 0.971 | 0.958 |
| 500-599 | 128,645 | 169,750 | 72,814 | 11,072 | 42.9% | 0.860 | 0.873 |
| 600-699 | 128,645 | 155,691 | 59,090 | 9,280 | 38.0% | 0.761 | 0.801 |
| 700-799 | 128,645 | 141,331 | 47,431 | 7,401 | 33.6% | 0.673 | 0.706 |
| 800-899 | 128,645 | 122,472 | 33,808 | 5,362 | 27.6% | 0.553 | 0.601 |
| 900-1000 | 128,645 | 101,552 | 19,739 | 2,625 | 19.4% | 0.390 | 0.352 |
| Overall | 1,286,450 | 1,634,619 | 815,415 | 114,772 | 49.9% | 1.000 | 1.000 |

**Summary of Aggregate Experience Validation Data**
Policy Years 2013-2017 as of 12/31/2017

| (1) Risk Score Band | (2) Earned Exposures | (3) Earned Premium | (4) Case Incurred Losses & ALAE | (5) Claim Count | (6) Case Incurred Loss & ALAE Ratio | Actual Loss Ratio Relativity | Indicated Loss Ratio Relativity |
|---|---|---|---|---|---|---|---|
| 1-99 | 55,284 | 81,841 | 61,573 | 8,168 | 75.2% | 1.508 | 1.618 |
| 100-199 | 55,069 | 80,302 | 54,081 | 7,212 | 67.3% | 1.350 | 1.319 |
| 200-299 | 55,189 | 83,594 | 48,285 | 6,715 | 57.8% | 1.158 | 1.151 |
| 300-399 | 55,089 | 81,162 | 44,047 | 6,152 | 54.3% | 1.088 | 1.050 |
| 400-499 | 55,264 | 78,250 | 37,511 | 5,432 | 47.9% | 0.961 | 0.958 |
| 500-599 | 55,164 | 72,675 | 32,356 | 4,812 | 44.5% | 0.892 | 0.873 |
| 600-699 | 54,989 | 66,663 | 27,338 | 4,098 | 41.0% | 0.822 | 0.801 |
| 700-799 | 55,264 | 60,800 | 22,116 | 3,246 | 36.4% | 0.729 | 0.706 |
| 800-899 | 55,114 | 52,538 | 16,165 | 2,312 | 30.8% | 0.617 | 0.601 |
| 900-1000 | 55,214 | 43,333 | 8,180 | 1,107 | 18.9% | 0.378 | 0.352 |
| Overall | 551,638 | 701,158 | 351,650 | 49,255 | 50.2% | 1.005 | 1.000 |

Exposure Weighted Correlation = 99%

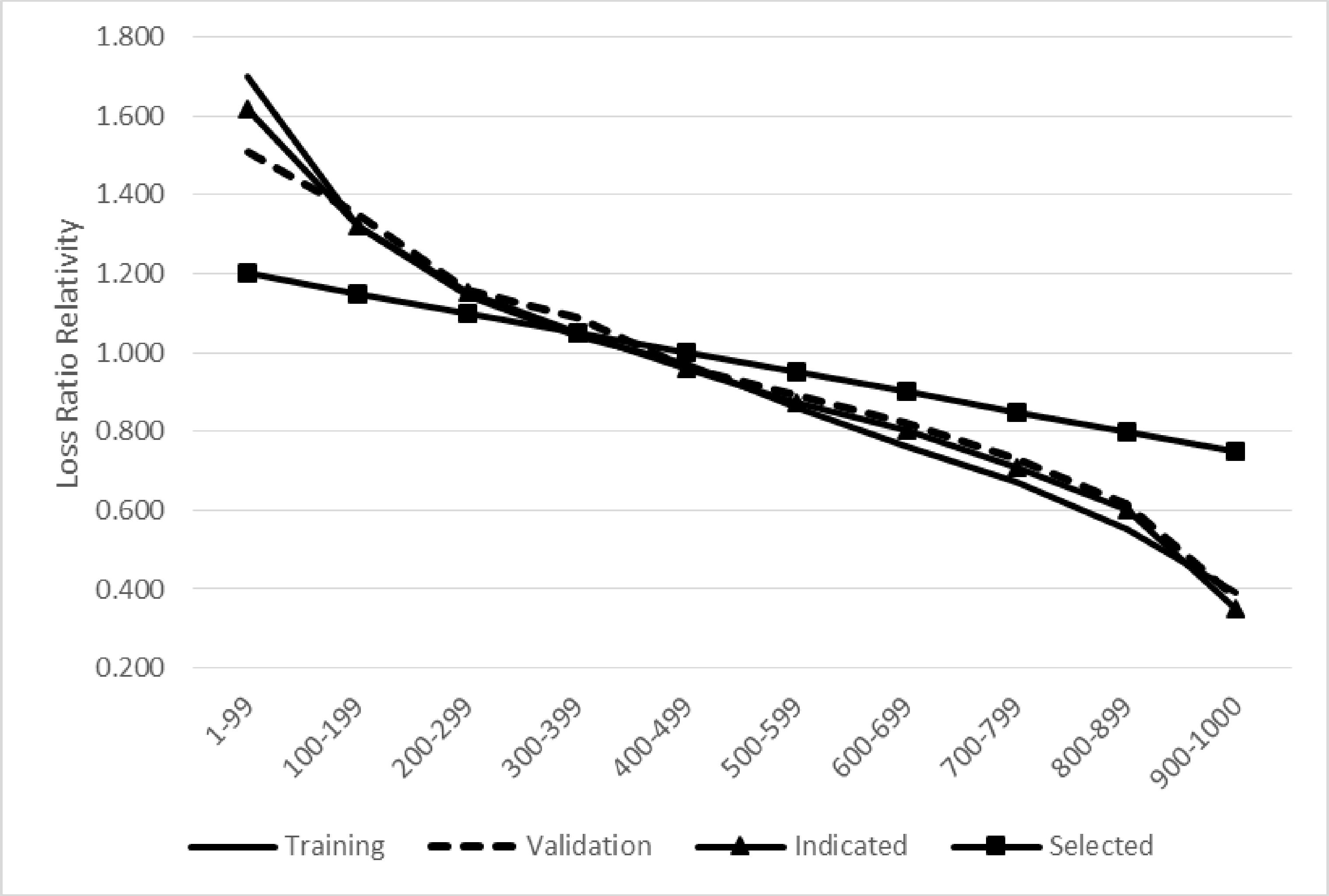
# Compare Indicated to Selected

**Summary of Aggregate Experience**

Policy Years 2013-2017 as of 12/31/2017

| (1) Risk Score Band | (2) Earned Exposures | (3) Earned Premium | (4) Case Incurred Losses & ALAE | (5) Claim Count | (6) Case Incurred Loss & ALAE Ratio | (7) Loss Ratio Relativity (Indicated Factor) * | (8) Indicated Factor Change | (9) Current Factor | (10) Proposed Factor | (11) Proposed Factor Change |
|---|---|---|---|---|---|---|---|---|---|---|
| 1-99 | 183,929 | 271,851 | 222,639 | 27,616 | 81.9% | 1.618 | 61.8% | 1.000 | 1.200 | 20.0% |
| 100-199 | 183,714 | 267,827 | 177,805 | 23,765 | 66.4% | 1.319 | 31.9% | 1.000 | 1.150 | 15.0% |
| 200-299 | 183,834 | 278,689 | 159,624 | 22,415 | 57.3% | 1.151 | 15.1% | 1.000 | 1.100 | 10.0% |
| 300-399 | 183,734 | 269,521 | 141,883 | 20,523 | 52.6% | 1.050 | 5.0% | 1.000 | 1.050 | 5.0% |
| 400-499 | 183,909 | 261,084 | 126,077 | 18,394 | 48.3% | 0.958 | -4.2% | 1.000 | 1.000 | 0.0% |
| 500-599 | 183,809 | 242,425 | 105,170 | 15,884 | 43.4% | 0.873 | -12.7% | 1.000 | 0.950 | -5.0% |
| 600-699 | 183,634 | 222,355 | 86,428 | 13,378 | 38.9% | 0.801 | -19.9% | 1.000 | 0.900 | -10.0% |
| 700-799 | 183,909 | 202,131 | 69,547 | 10,646 | 34.4% | 0.706 | -29.4% | 1.000 | 0.850 | -15.0% |
| 800-899 | 183,759 | 175,010 | 49,972 | 7,675 | 28.6% | 0.601 | -39.9% | 1.000 | 0.800 | -20.0% |
| 900-1000 | 183,859 | 144,885 | 27,920 | 3,732 | 19.3% | 0.352 | -64.8% | 1.000 | 0.750 | -25.0% |
| Overall | 1,838,088 | 2,335,777 | 1,167,065 | 164,027 | 50.0% | 1.000 | 0.0% | 1.000 | 1.000 | 0.0% |

# More Comparisons



## *Filing Questions*

- What is “the calculation”?
- Double dipping?
- Typical Risks?
- GLM Output

# Summary

- Internal and regulatory acceptance share common features

- Keep description of model and process simple and non-technical (but disclose and document both)

- Lift comparison of modeled vs hold out data are key