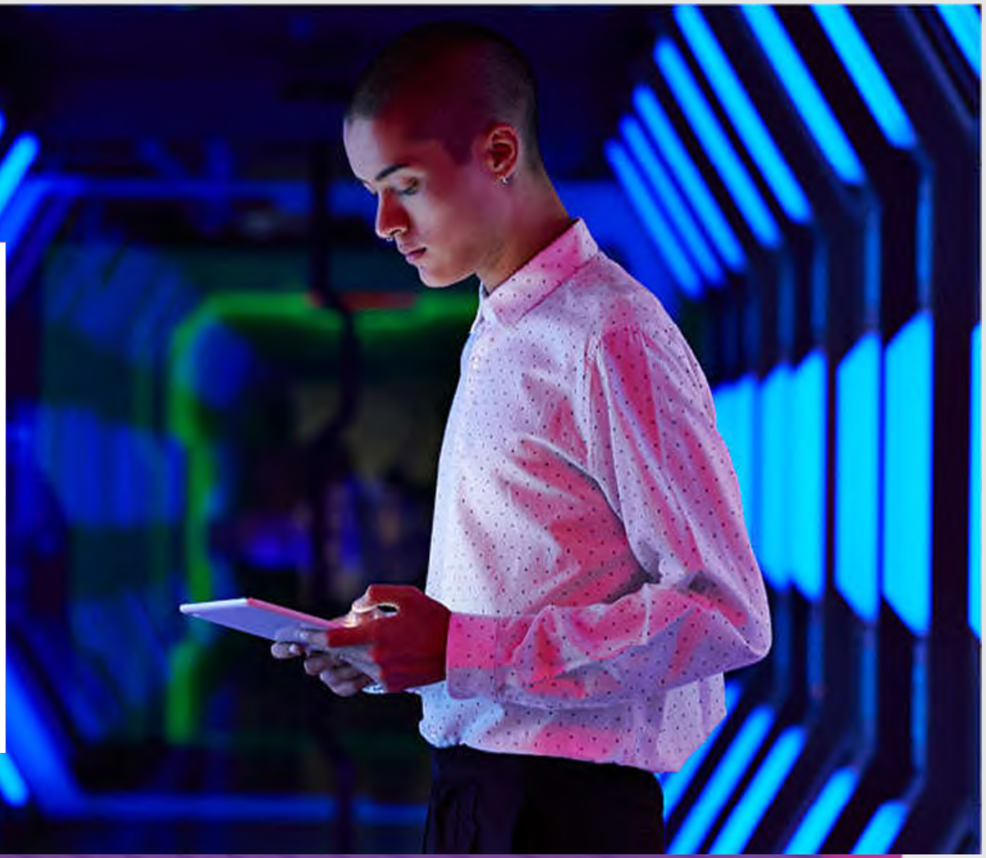
CAS Ratemaking, Product and Modeling Seminar

Advanced Machine Learning in Real Life

Serhat Guven and Michael Chen

March 2023



Agenda: Advanced Machine Learning in Real Life

1. Automated Modeling
2. Rating Plan Development
3. Customer Behavior
4. Simulation and Search

About the Survey

WTW asked property & casualty (P&C) insurers in the U.S. and Canada for their insights on the future of advanced analytics relative to where we are today.

71
U.S. and Canadian insurers

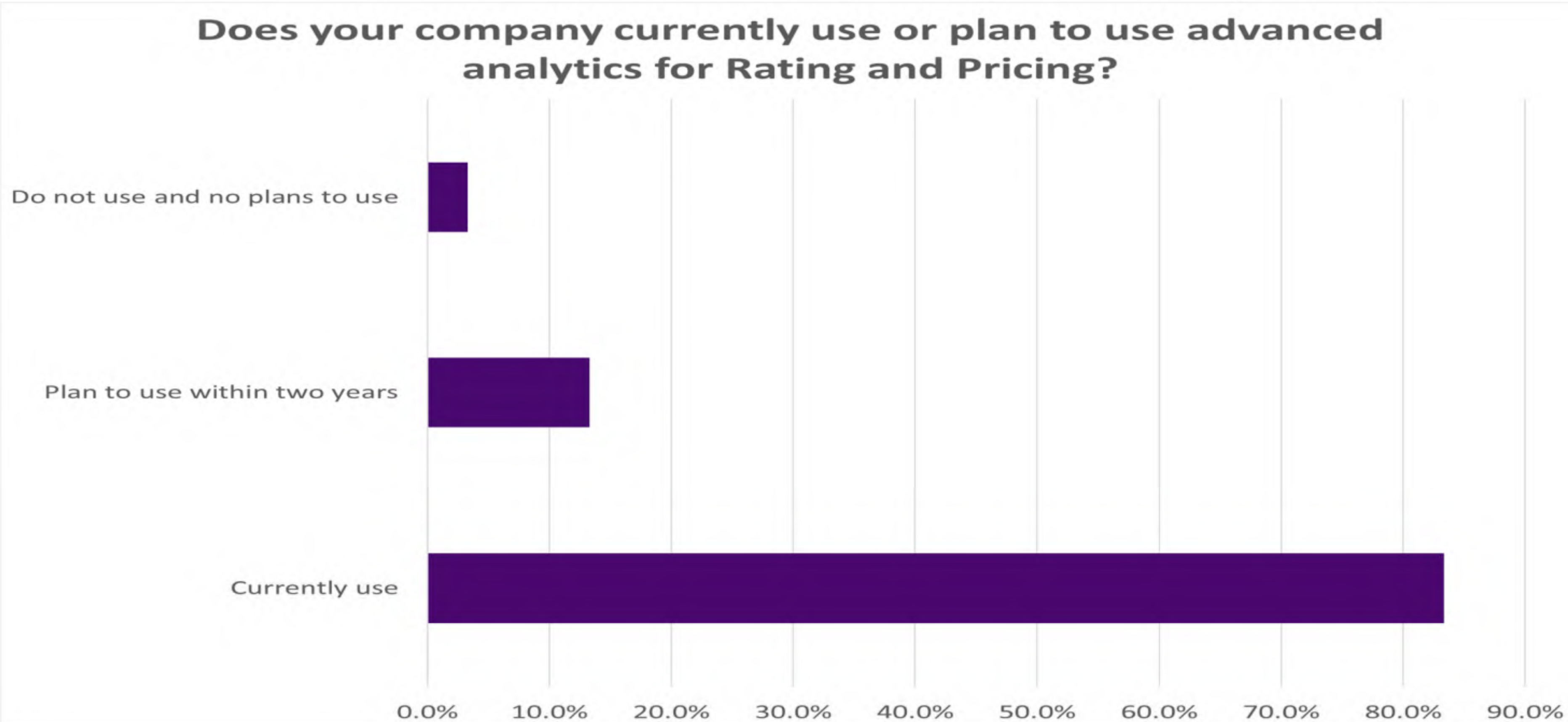
- **62** multiline insurers
- **3** commercial line insurers
- **6** personal line insurers
- **5** **top 10** P&C insurers (**Canada**)
- **6** **top 10** P&C insurers (**U.S.**)

Survey Question

1. Does your company currently use or plan to use advanced analytics for Rating and Pricing?

A. Currently Use
B. Plan to use within two years
C. Do not use and no plans to use

Advanced Analytics Survey Results

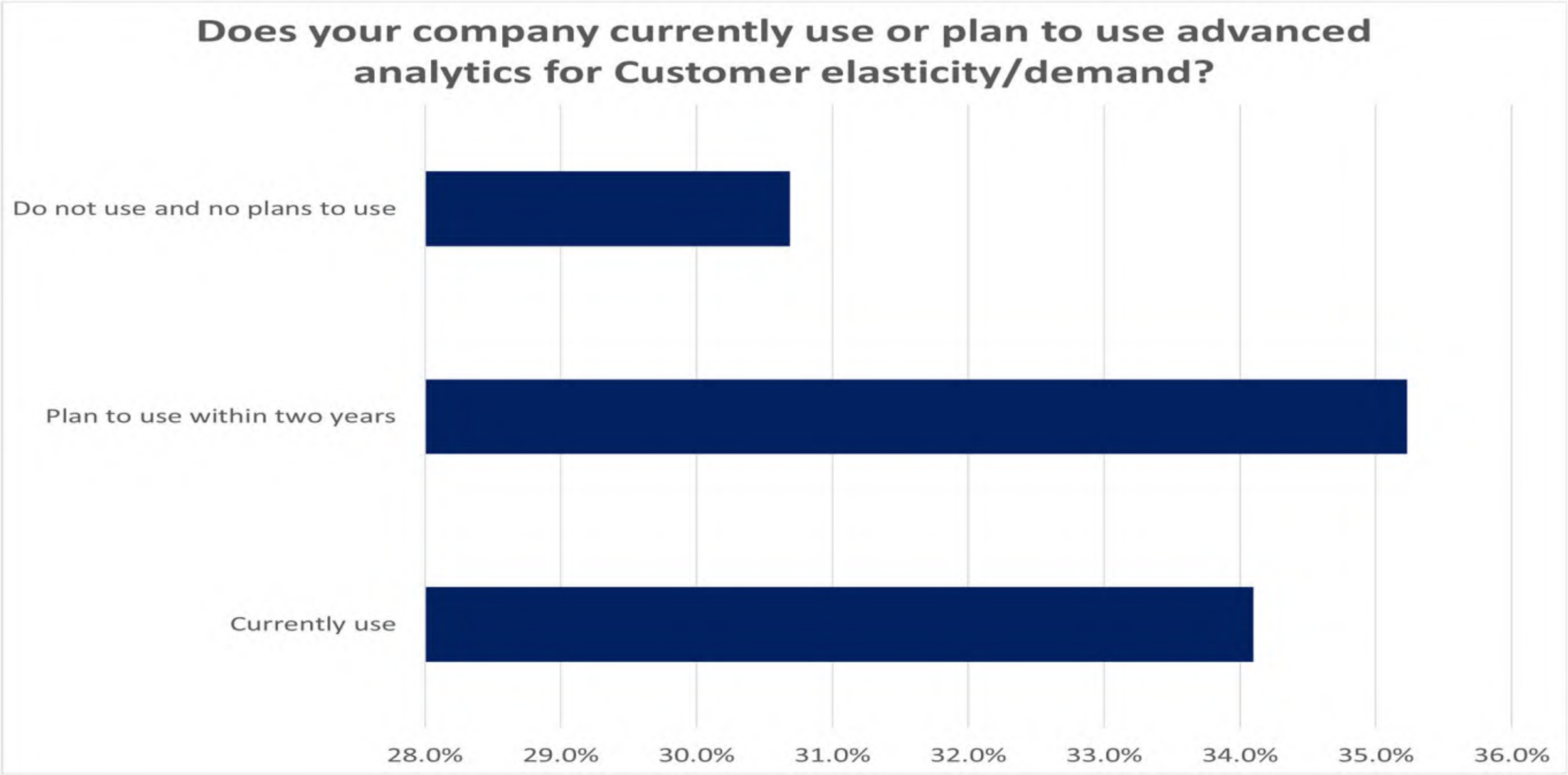


Survey Question

2. Does your company currently use or plan to use advanced analytics for Customer elasticity/demand?

A. Currently Use
B. Plan to use within two years
C. Do not use and no plans to use

Advanced Analytics Survey Results

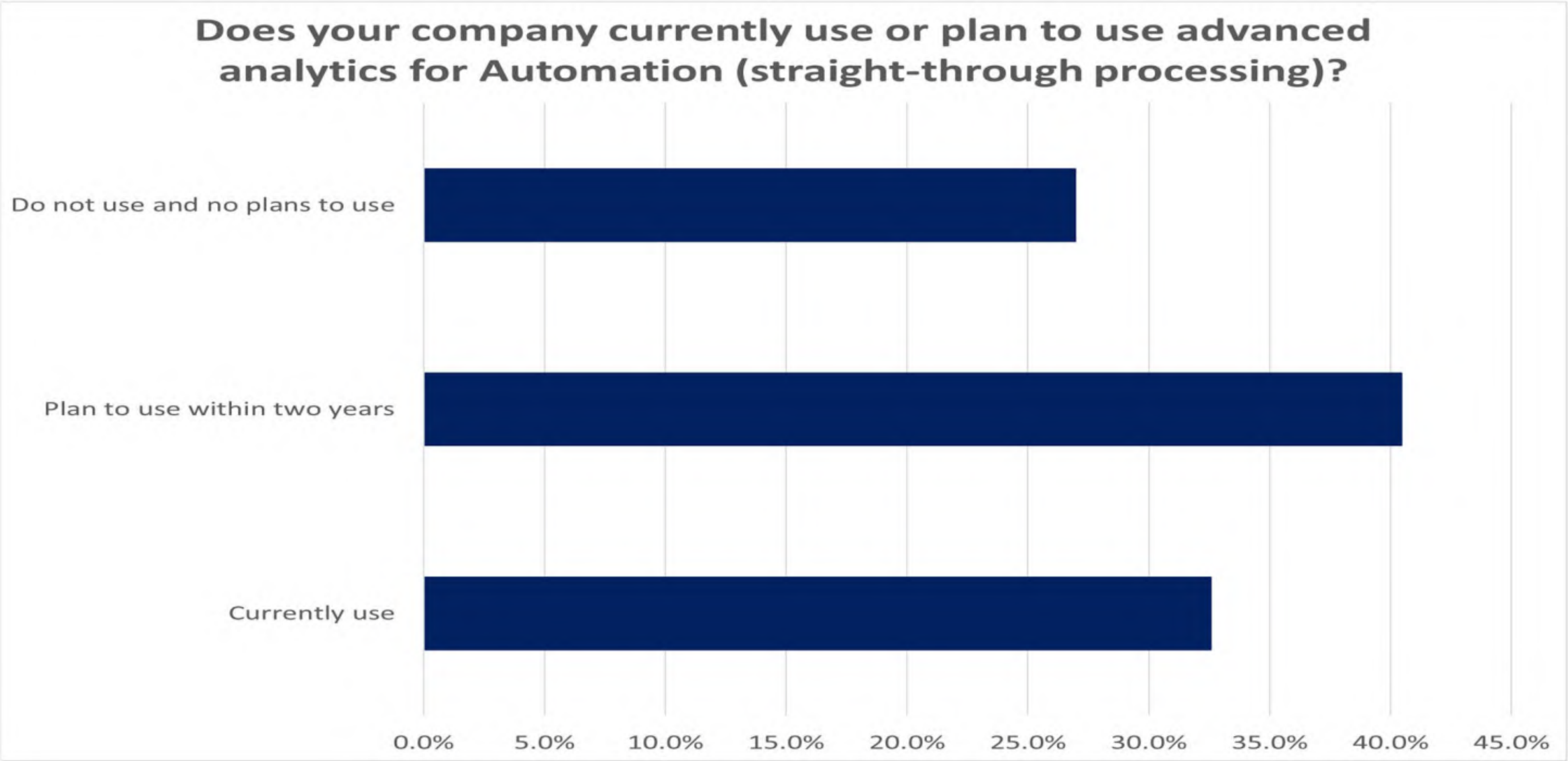


Survey Question

3. Does your company currently use or plan to use advanced analytics for Straight-through processing?

A. Currently Use
B. Plan to use within two years
C. Do not use and no plans to use

Advanced Analytics Survey Results

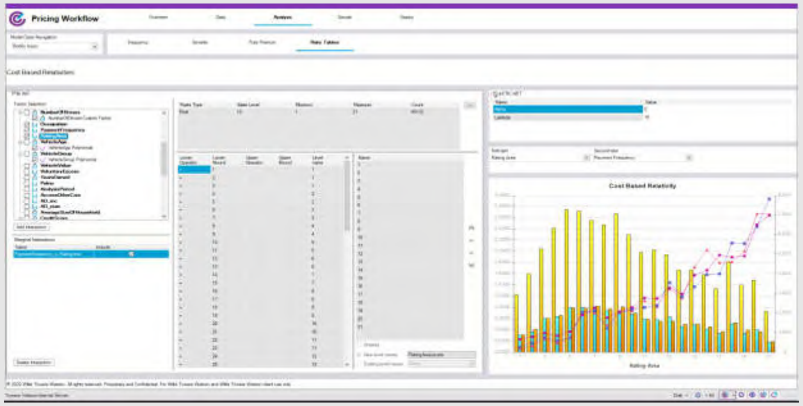


Automated Modeling

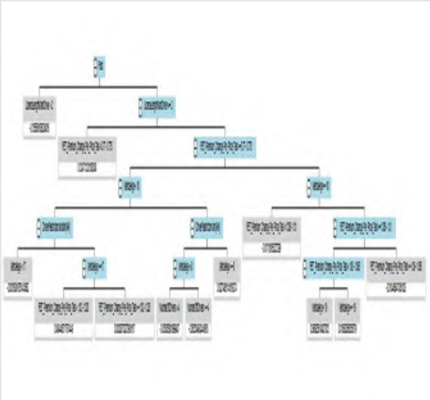
How to Model Customer Demand

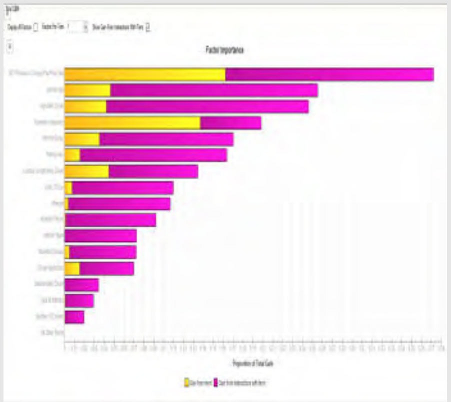
General types of modeling solutions

- Parametric
  - Resulting probability KPIs can be expressed as closed form table based solutions
  - Easier to interpret – although care needs to be taken when articulating relativities
  - Can be more time consuming to create (elastic net strategies do mitigate – but still struggle with lower level interactions
  - Less predictive power



- Non parametric
  - Resulting probability KPIs expressed as complex series of recursive decision trees
  - Requires a robust hyperparameter tuning process
  - Model interpretation can be very difficult
    - Standard main effect factor importance output
    - Partial dependency plots
    - WTW's layered GBM

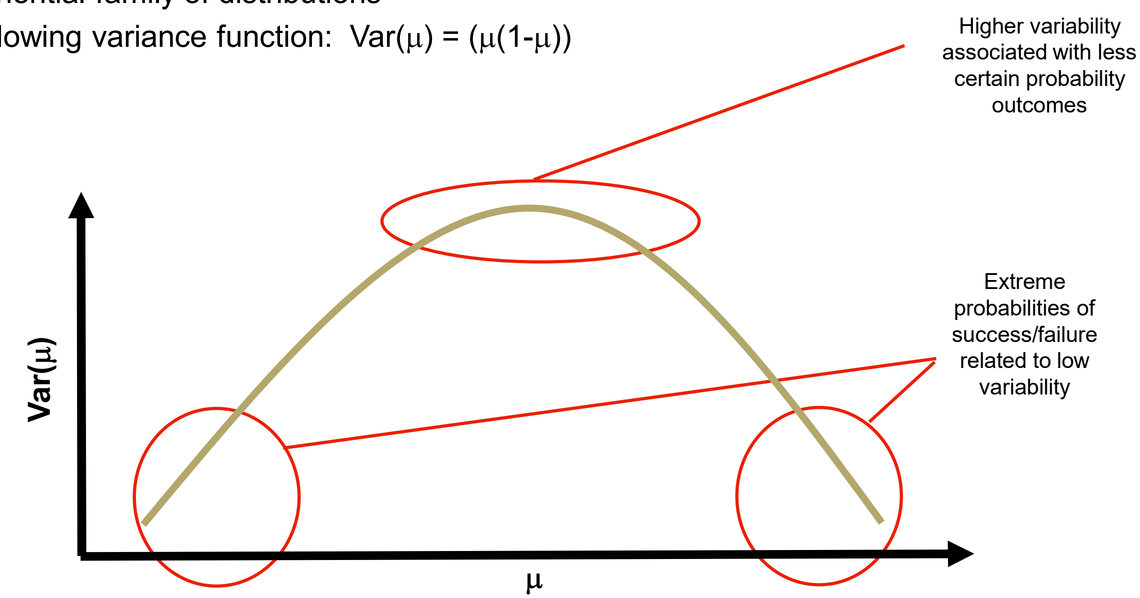




Parametric Modeling of Demand

Using a GLM/GAM/Elastic Net

- Binomial Distribution
  - Basic functional form in decision modeling
  - Belongs to the exponential family of distributions
  - Described by the following variance function: $Var(\mu) = (\mu(1-\mu))$

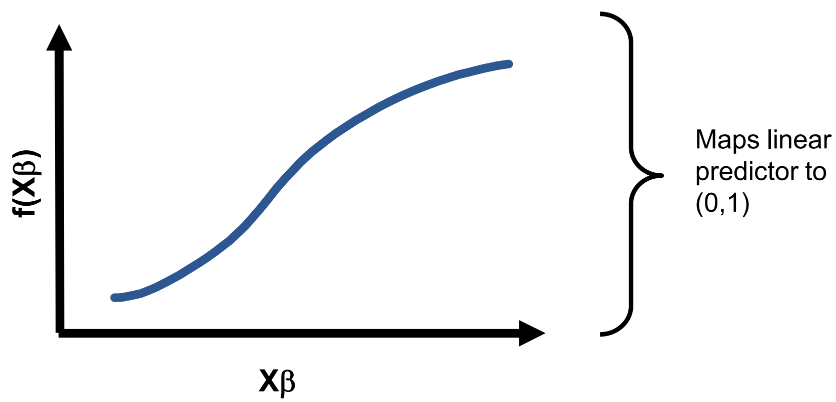


Parametric Modeling of Demand

Using a GLM/GAM/Elastic Nets

- Link function – converts the combination of parameters into an expected demand
  - Logit – canonical link function of the binomial distribution – which can then be expressed as an odds ratio relativity
  - Probit – uses the inverse normal

- General properties of the link function – serves as an S-shape curve with asymptotes at 0 and 1

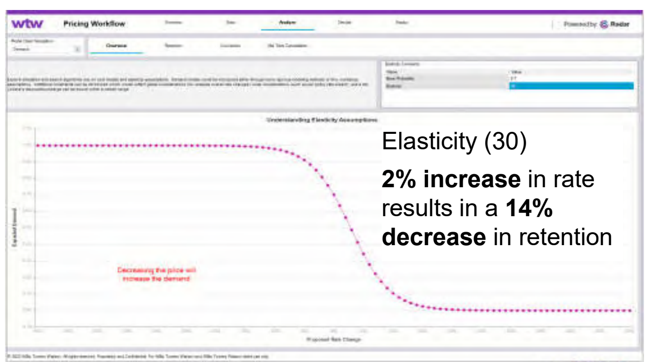


Parametric Modeling of Demand

Using a GLM/GAM/Elastic Nets

- A key interpretative element is to understand how elasticity affects the shape of the demand curve
- Typical skewness insurance retention rates (75-95%) results in rate increases have more impact on retention than decreases

Elasticity Sample Projections



Parametric Modeling of Demand

Using a GLM/GAM/Elastic Nets

- Advantages
  - A known commodity
  - Relatively easy to interpret and explain
  - Reponse variable does not have to be normally distributed

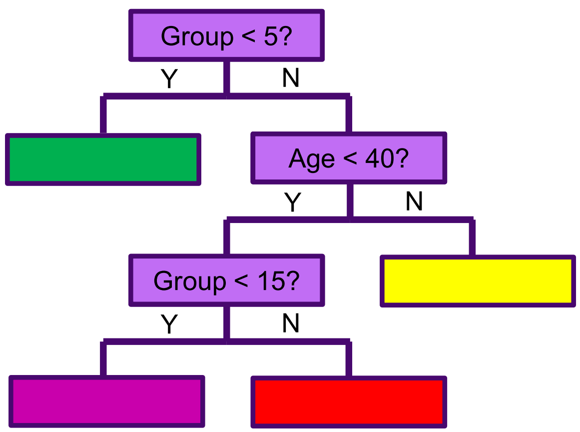
- Disadvantages
  - Does not handle non-linear relationships well
  - Modeling interactions often requires manual adjustments

Non Parametric Modeling of Demand

How do they work?

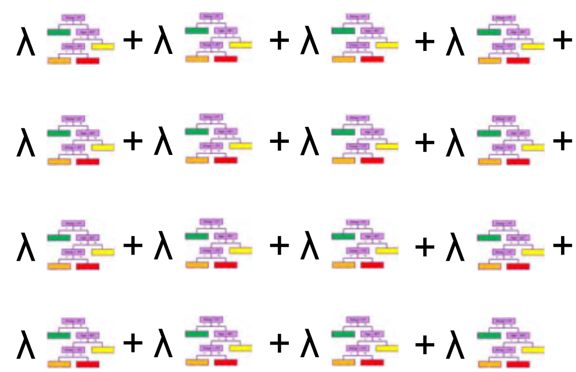
A tree

$f_i(x)$



A GBM

$$f(x) = \lambda \sum_{n=1}^{N} f_n(x)$$



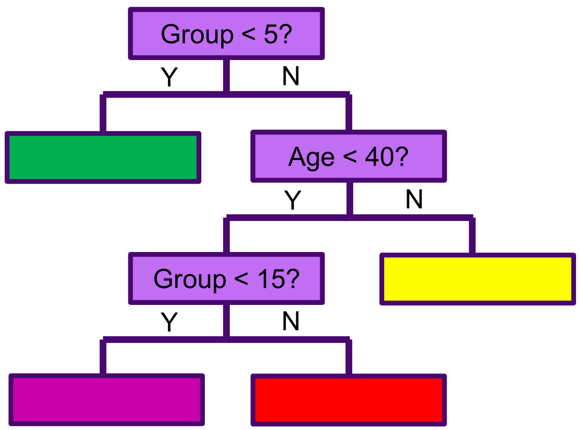
Non Parametric Modeling of Demand

Hyperparameters: the main assumptions of a non-parametric model

- Bag fraction
  - Row sampling: trees are fitted to a subset of the data (the bag fraction) on a randomized basis
  - Column sampling: additional noise-reduction can be achieved by using a random subset of the available factors at each iteration
- N Number of trees (iterations) allowed
- Learning Rate – amount that each iteration contributes to the overall model
- Interaction depth
  - Number of splits allowed on each tree (or the number of terminal nodes – 1)
- Minimum Child Weight
  - Weight required for the tree to continue splitting
- Minimum Split Loss
  - Change in the individual tree fitting objective function

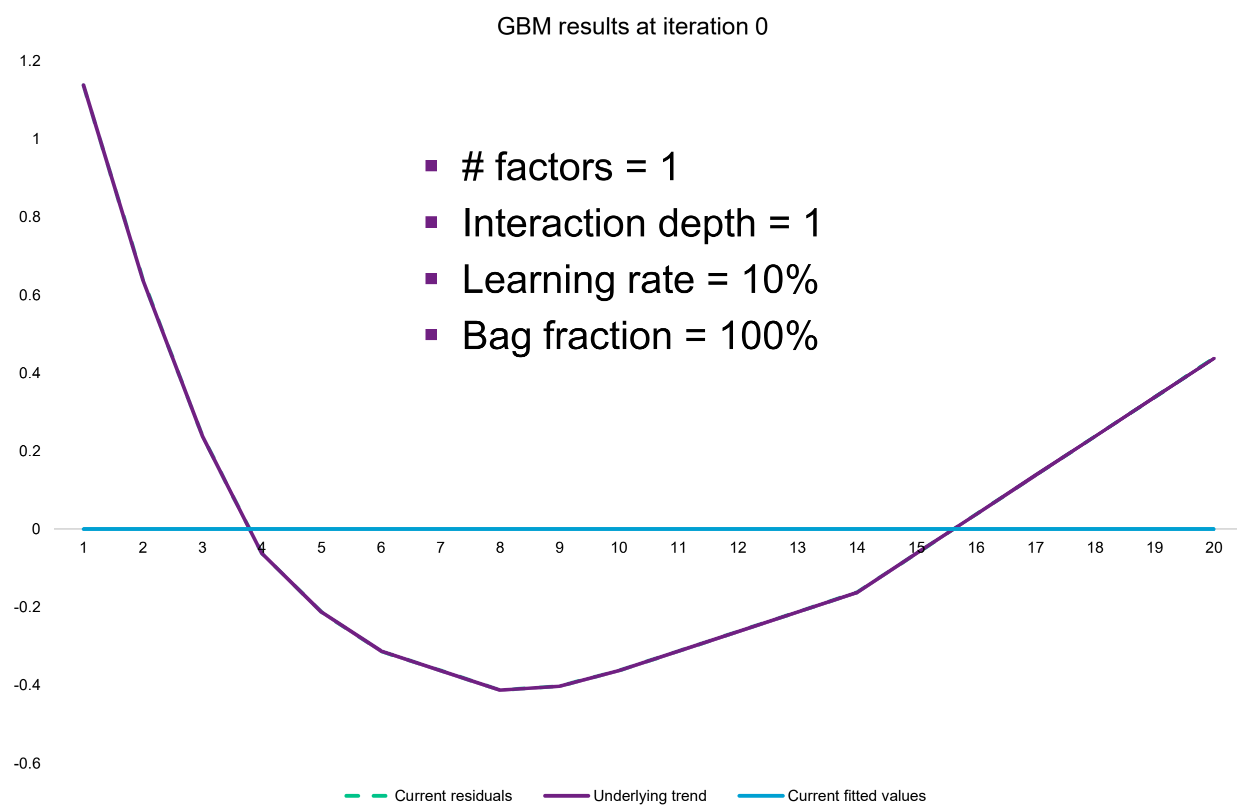
A tree

$f_i(x)$



Non Parametric Modeling of Demand

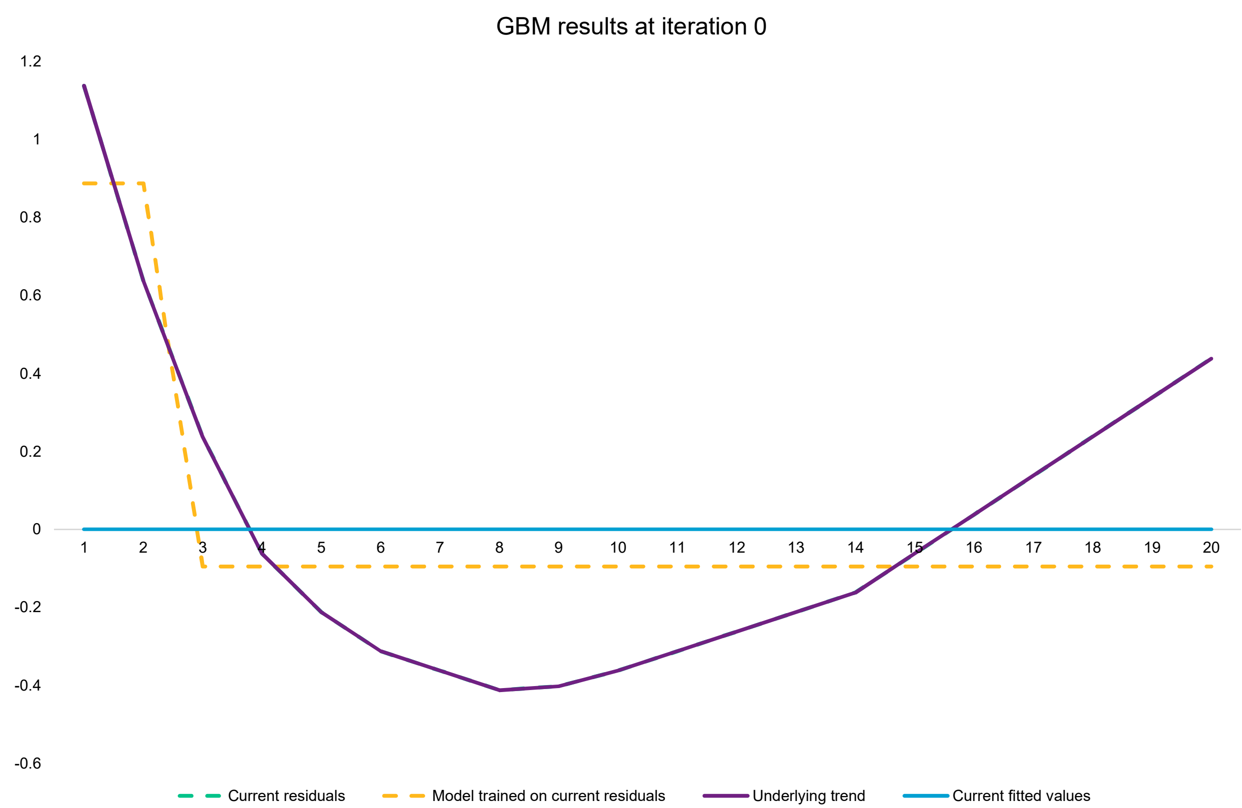
How do they work – a simple example



- # factors = 1
- Interaction depth = 1
- Learning rate = 10%
- Bag fraction = 100%

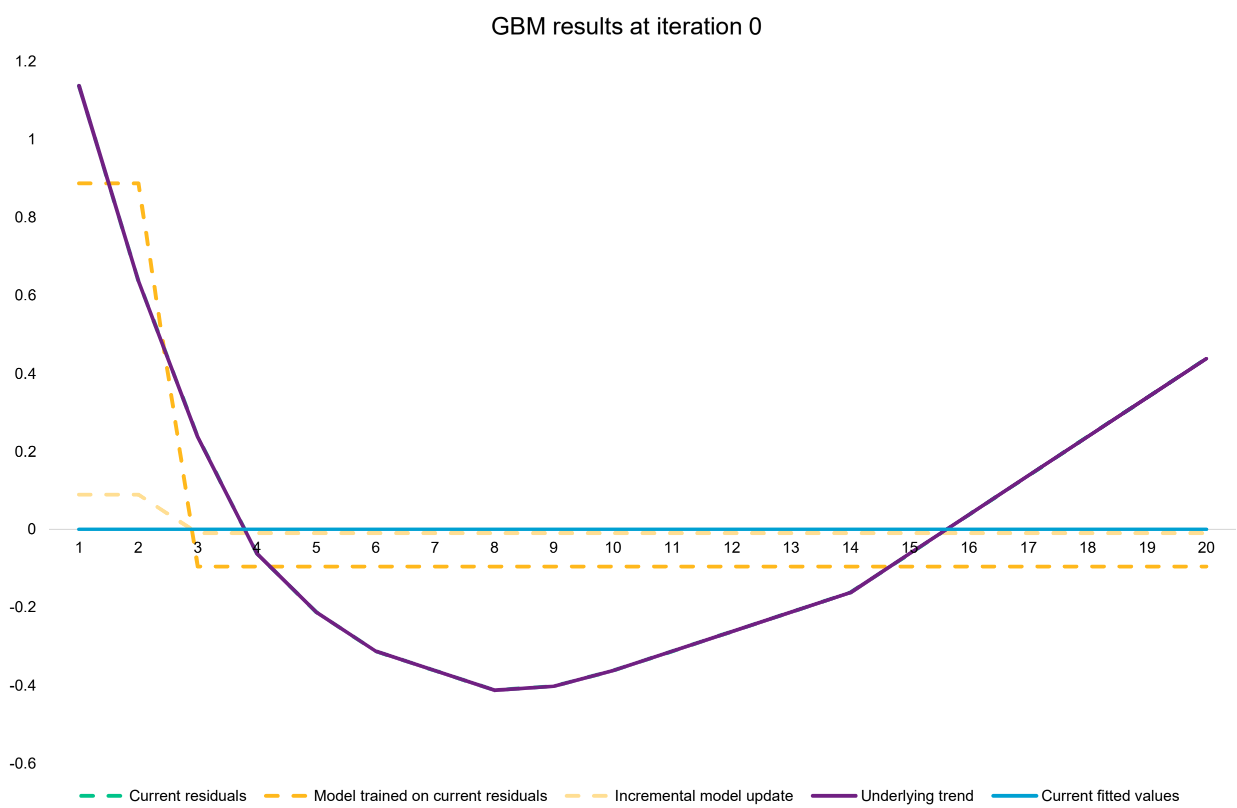
Non Parametric Modeling of Demand

How do they work – a simple example



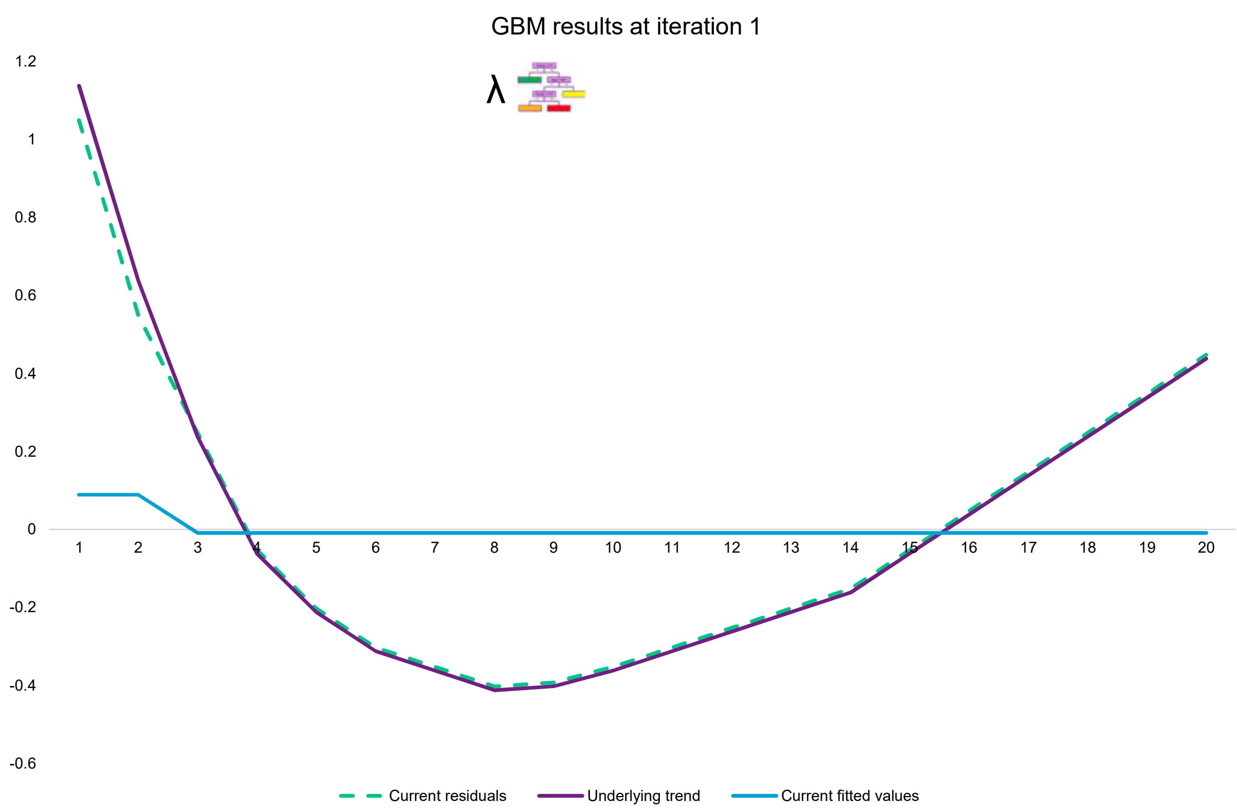
Non Parametric Modeling of Demand

How do they work – a simple example



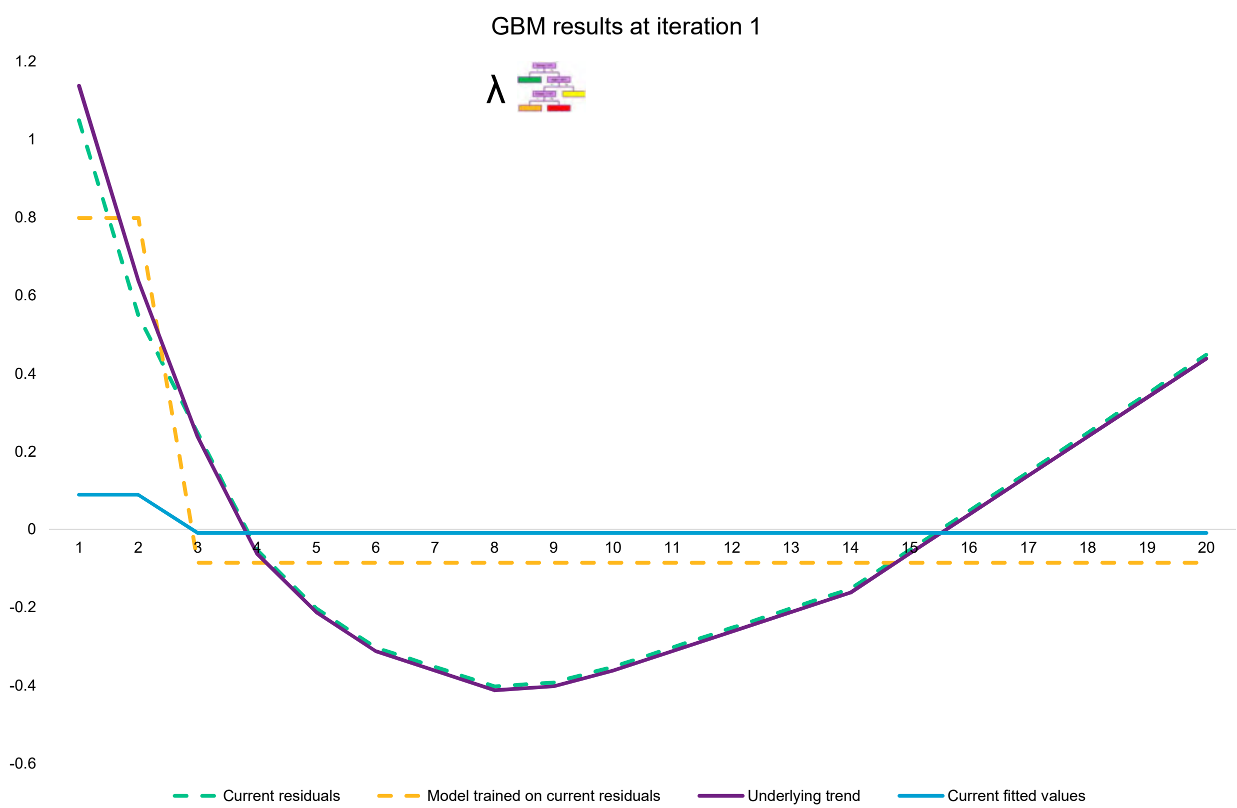
Non Parametric Modeling of Demand

How do they work – a simple example



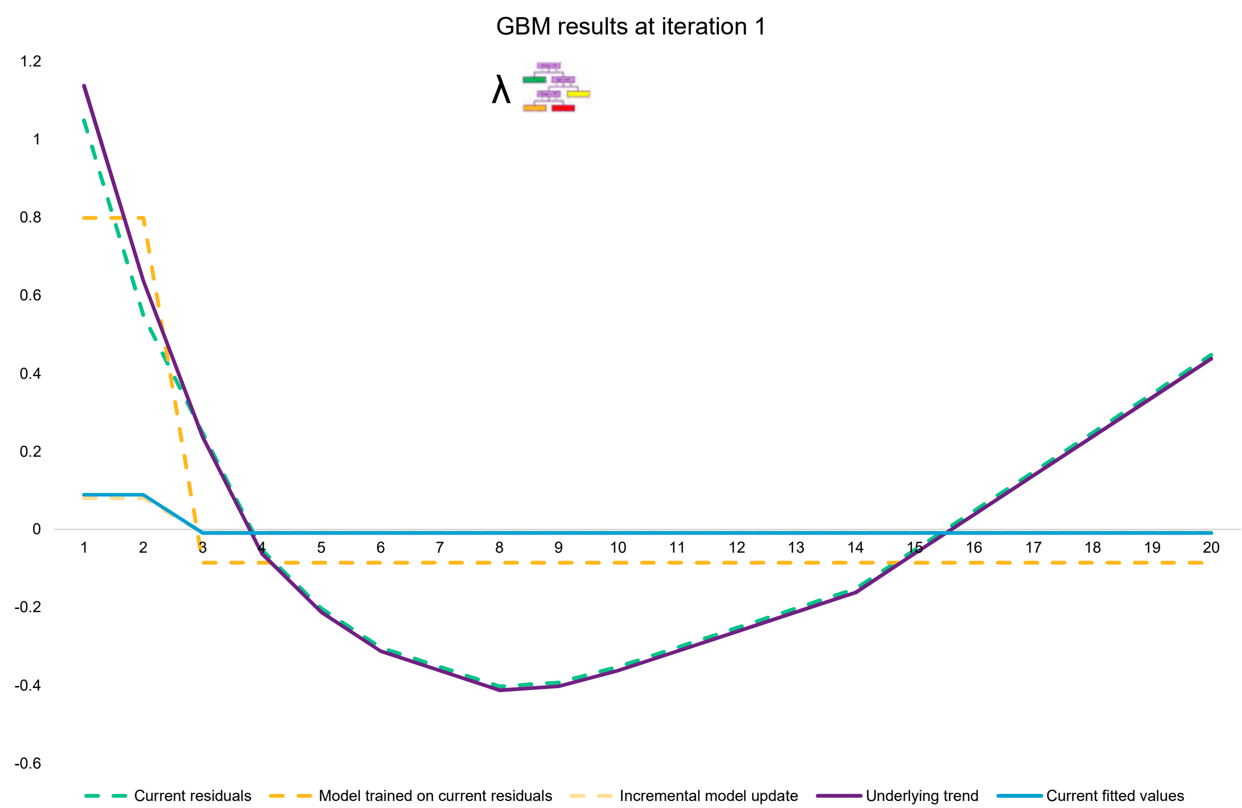
Non Parametric Modeling of Demand

How do they work – a simple example



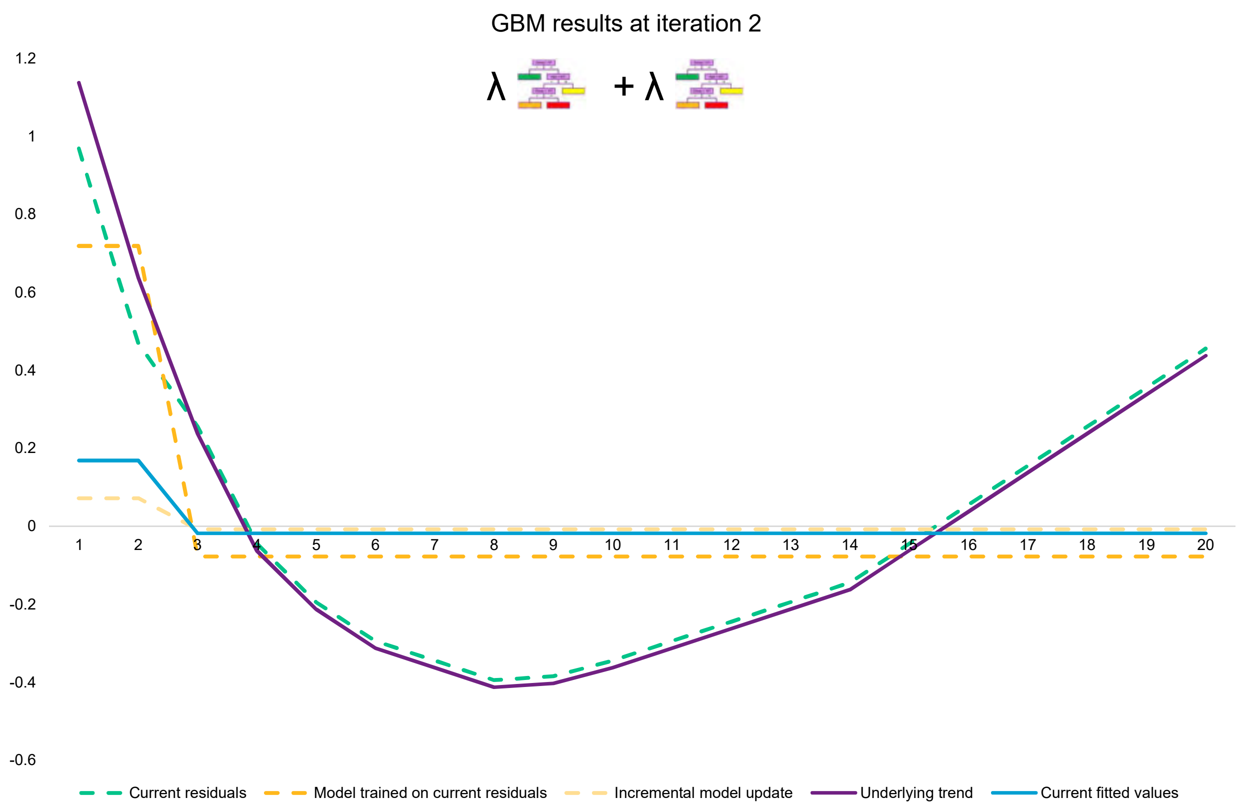
Non Parametric Modeling of Demand

How do they work – a simple example



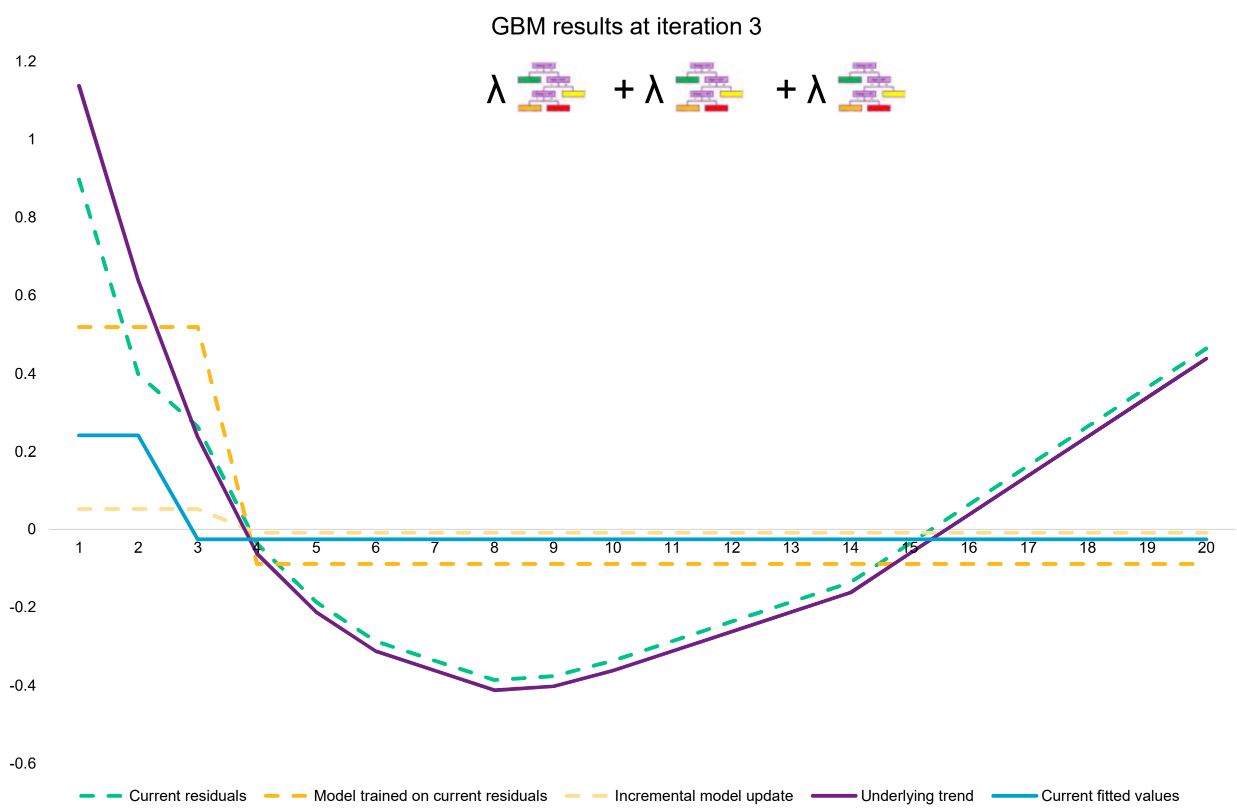
Non Parametric Modeling of Demand

How do they work – a simple example



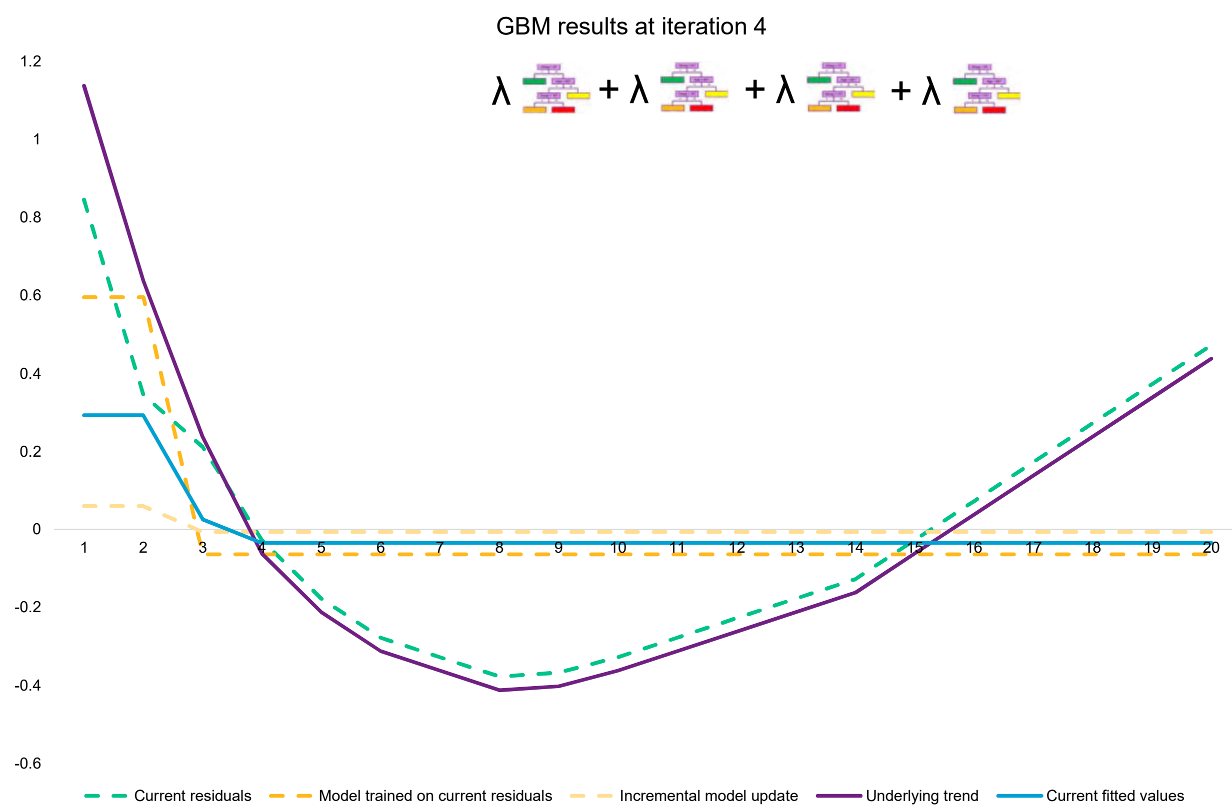
Non Parametric Modeling of Demand

How do they work – a simple example



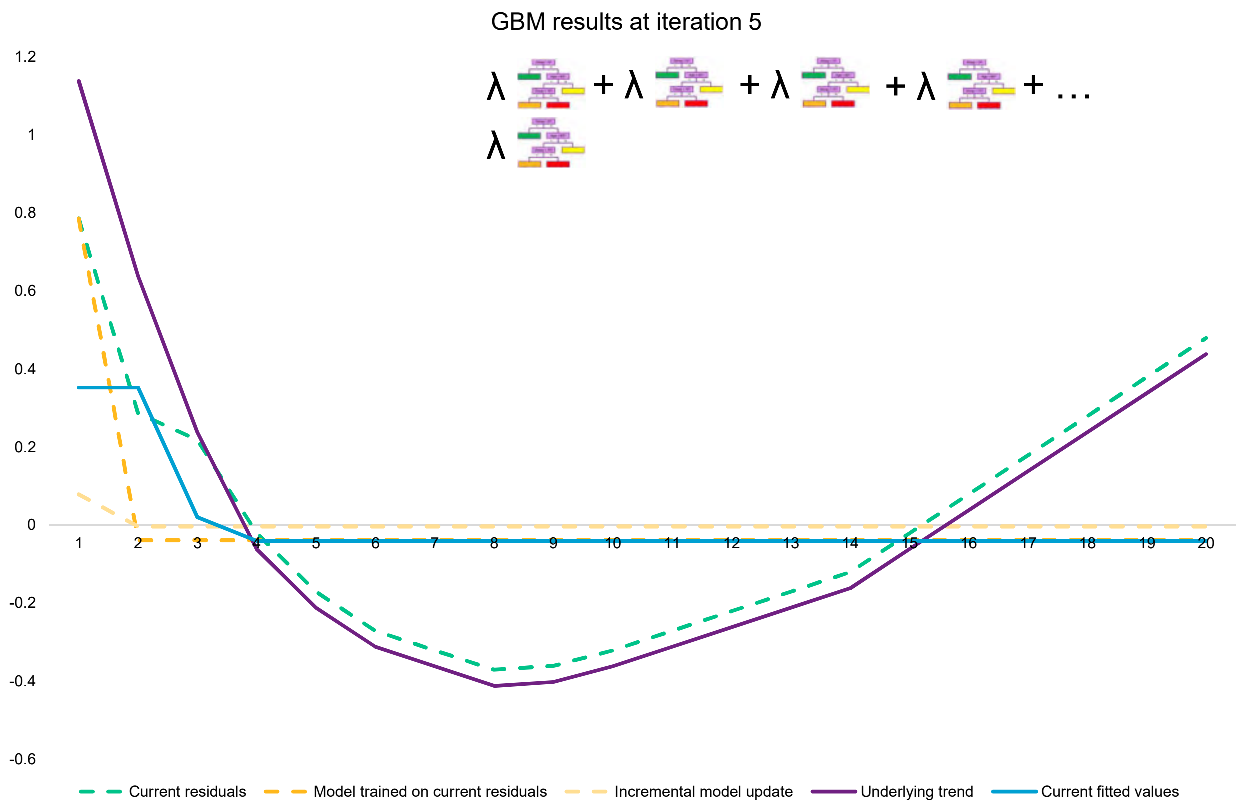
Non Parametric Modeling of Demand

How do they work – a simple example



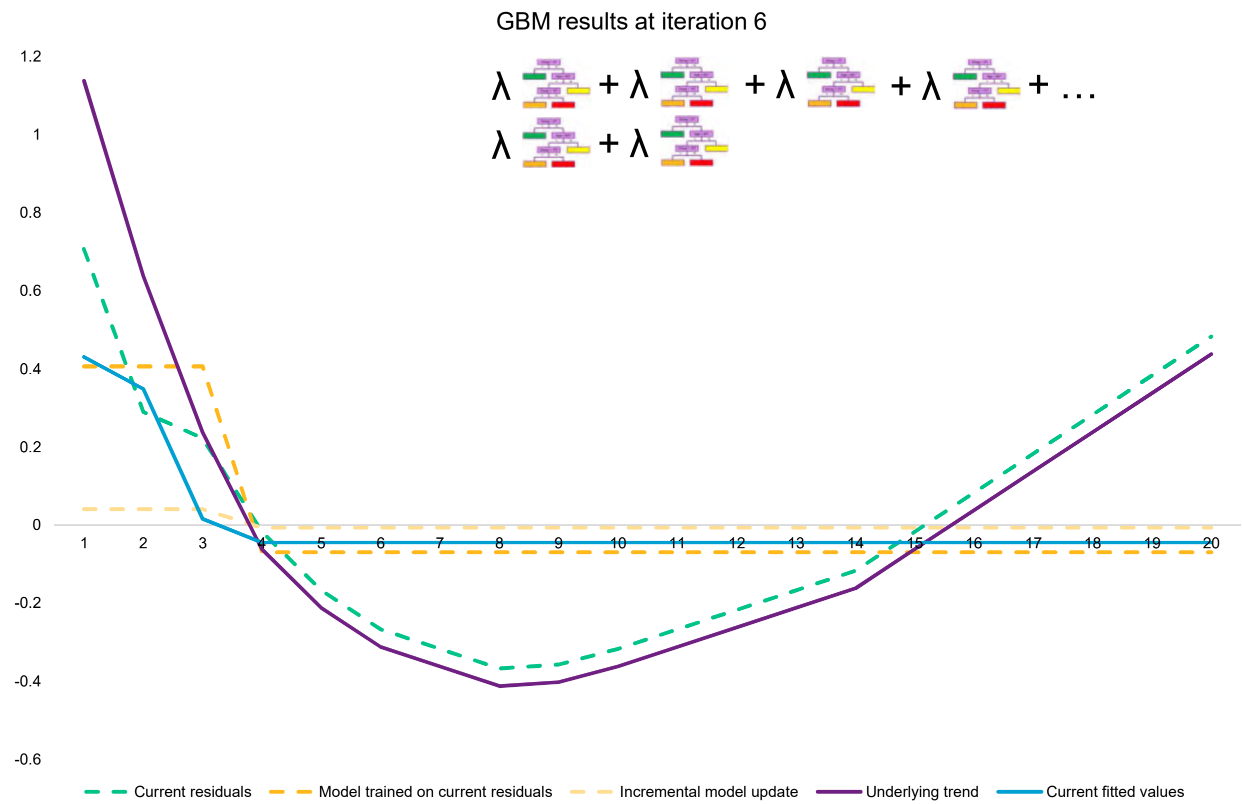
Non Parametric Modeling of Demand

How do they work – a simple example



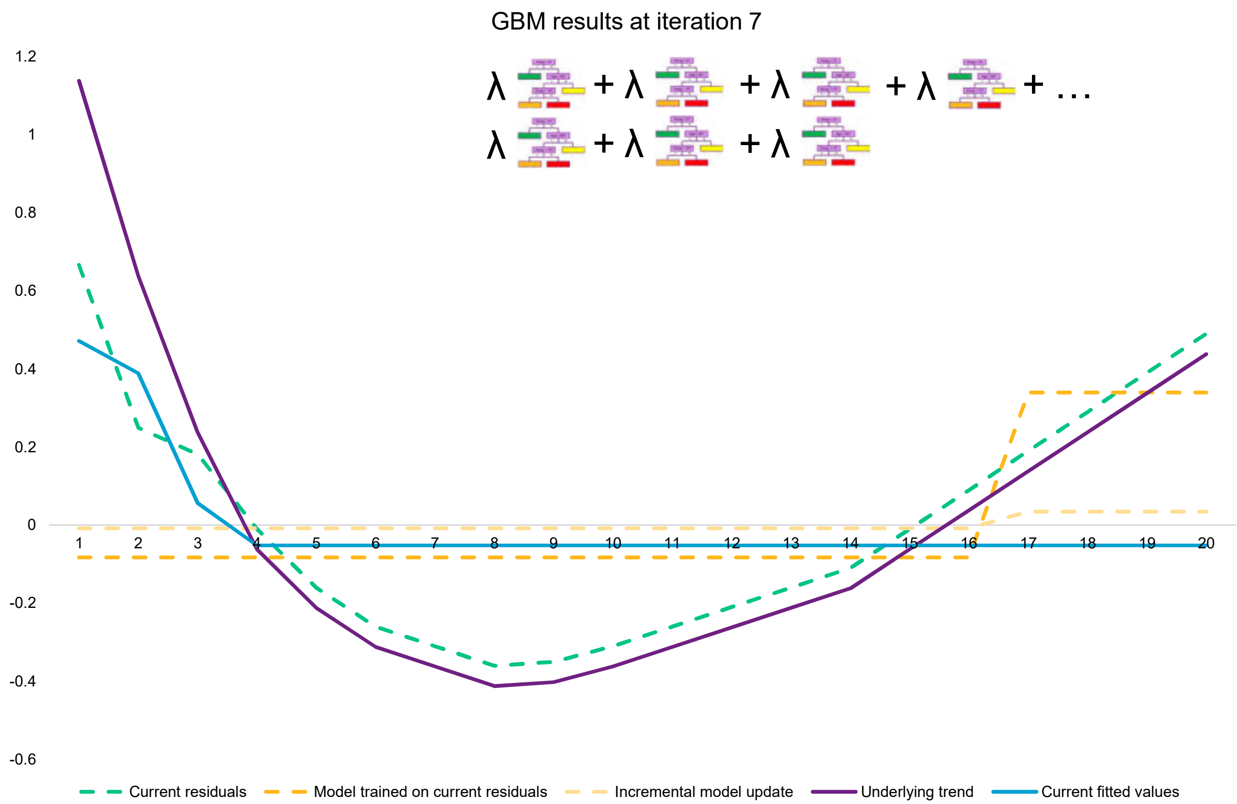
Non Parametric Modeling of Demand

How do they work – a simple example



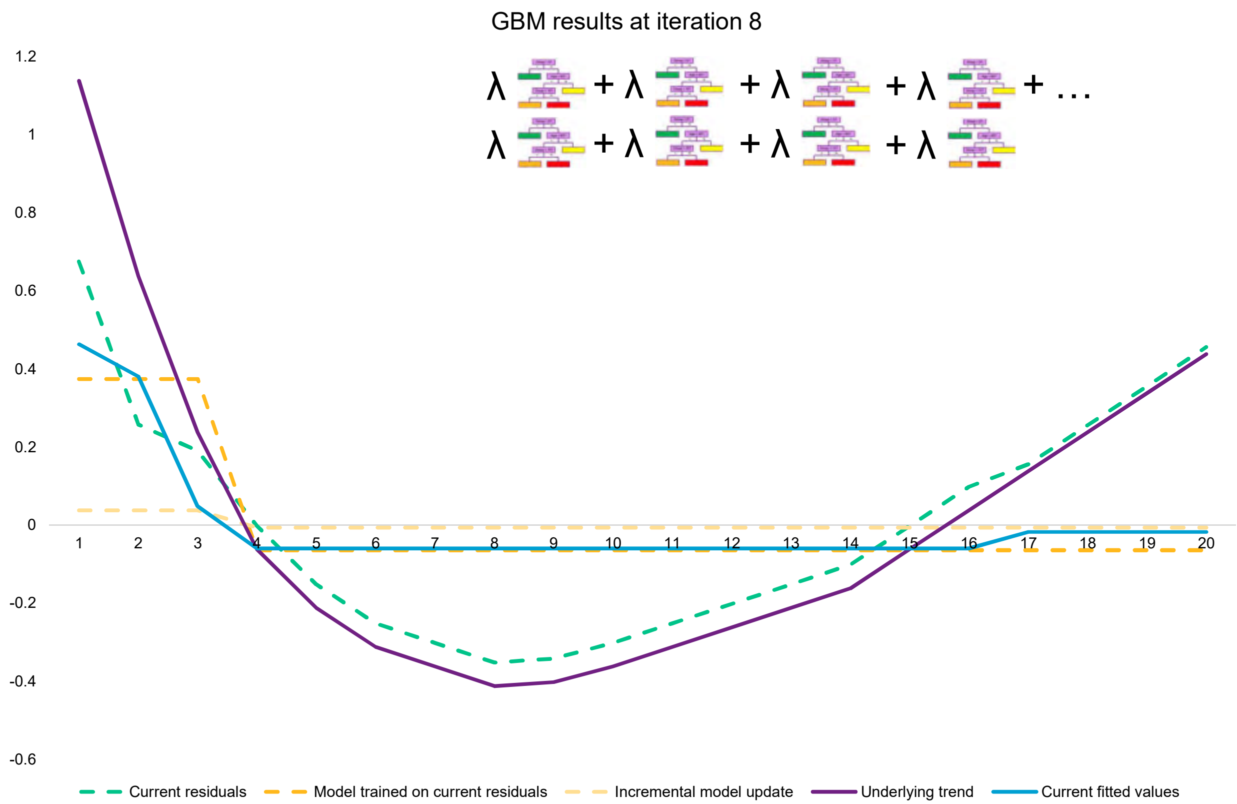
Non Parametric Modeling of Demand

How do they work – a simple example



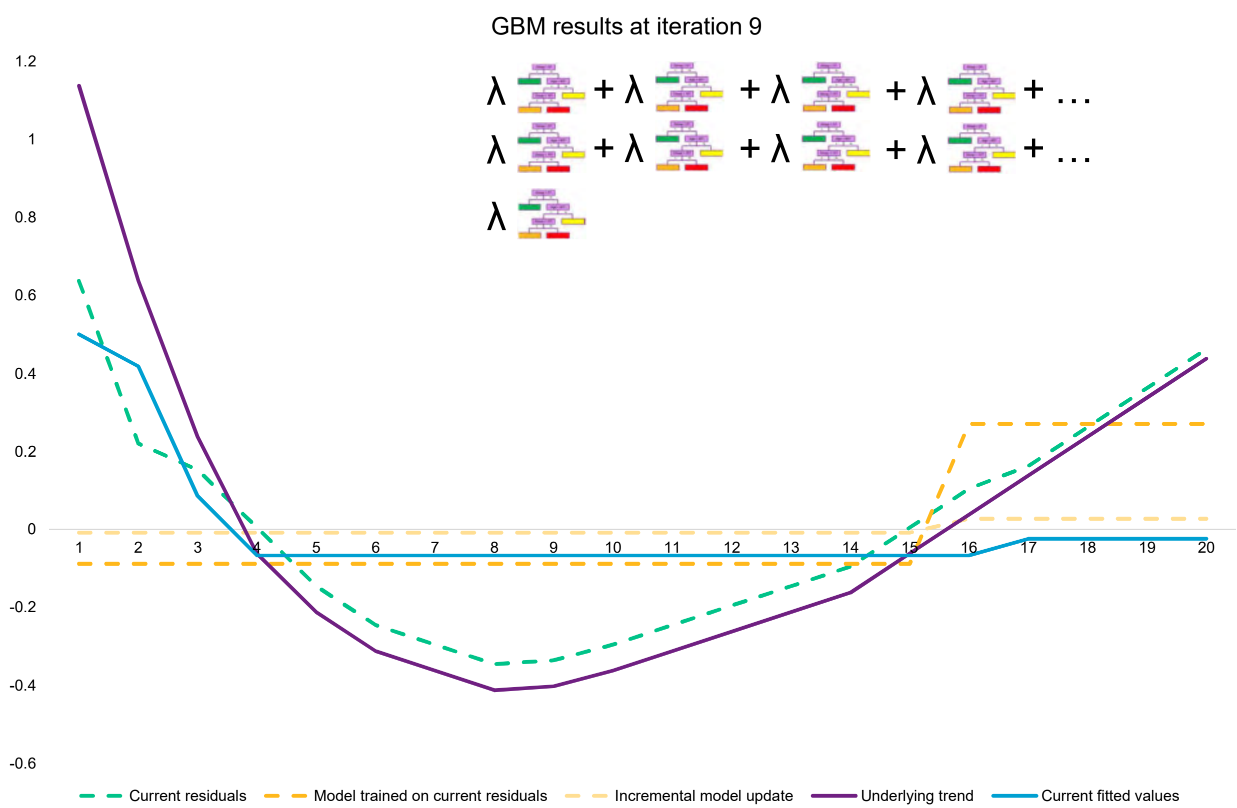
Non Parametric Modeling of Demand

How do they work – a simple example



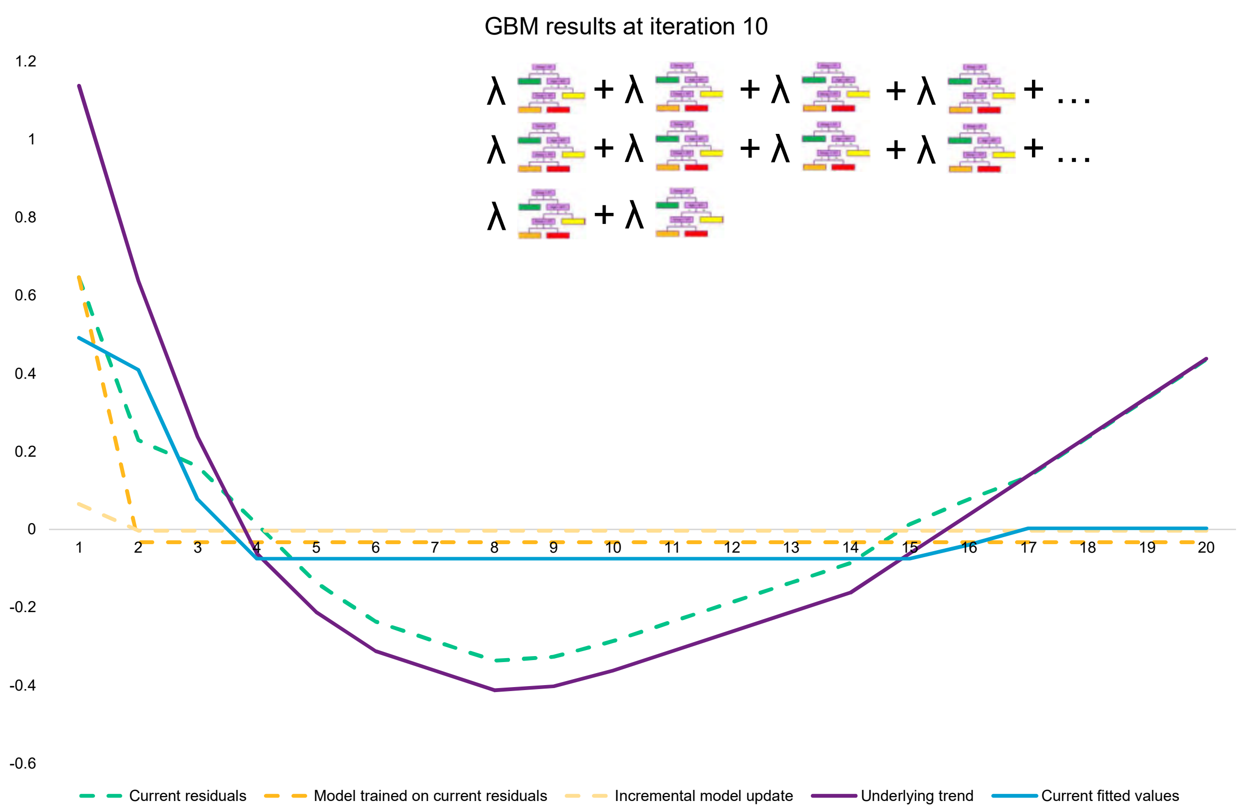
Non Parametric Modeling of Demand

How do they work – a simple example



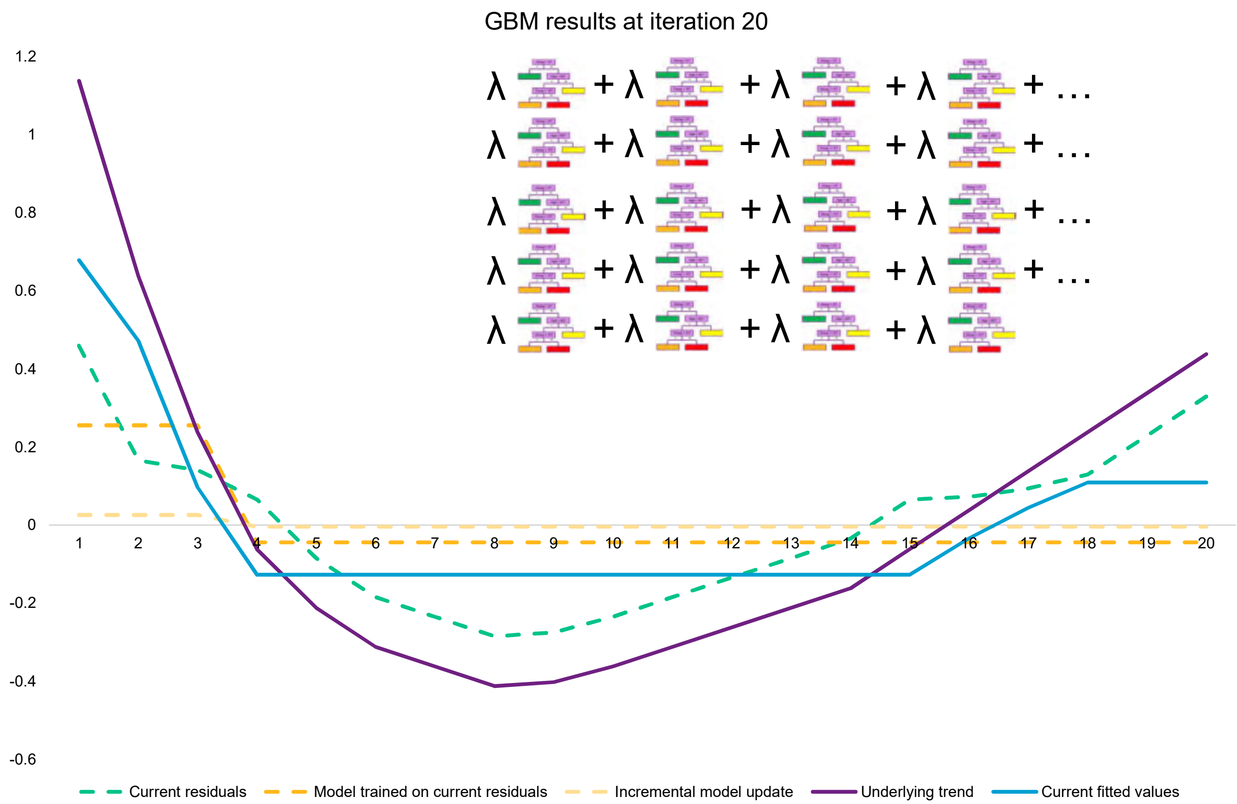
Non Parametric Modeling of Demand

How do they work – a simple example



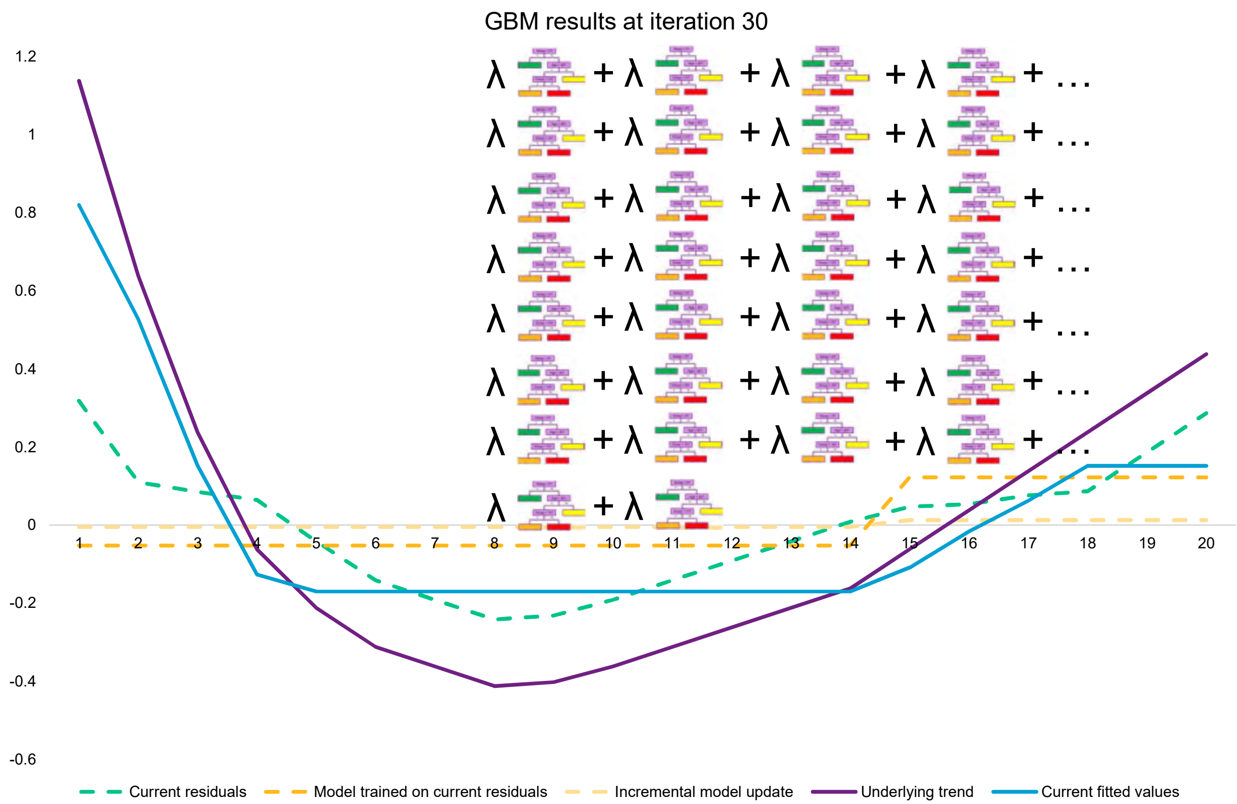
Non Parametric Modeling of Demand

How do they work – a simple example



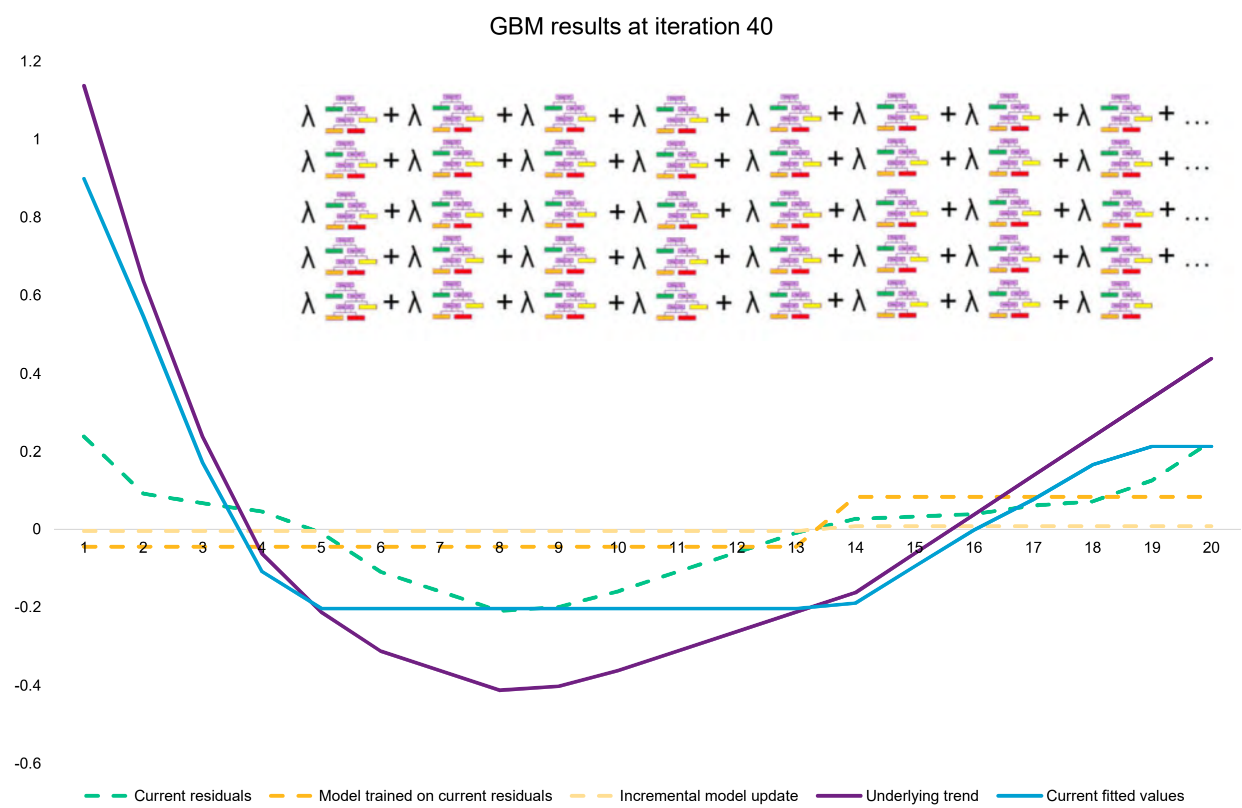
Non Parametric Modeling of Demand

How do they work – a simple example



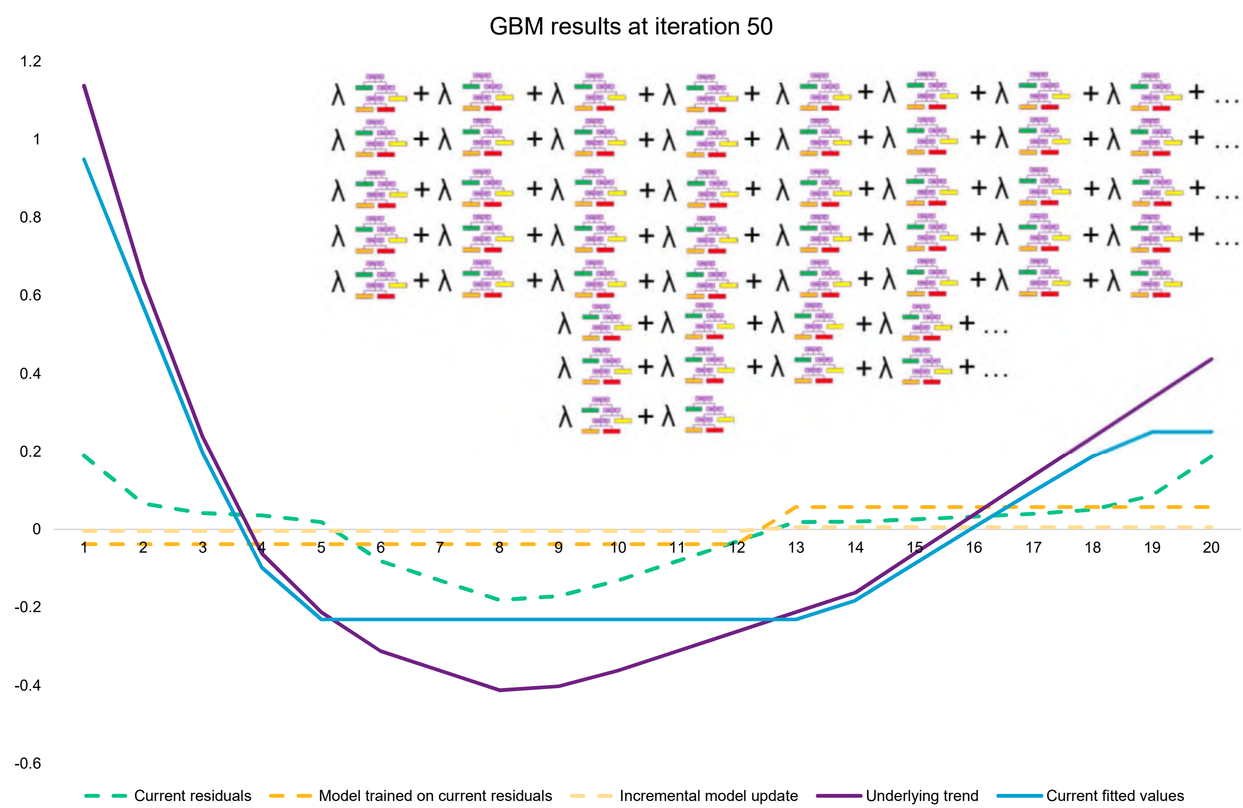
Non Parametric Modeling of Demand

How do they work – a simple example



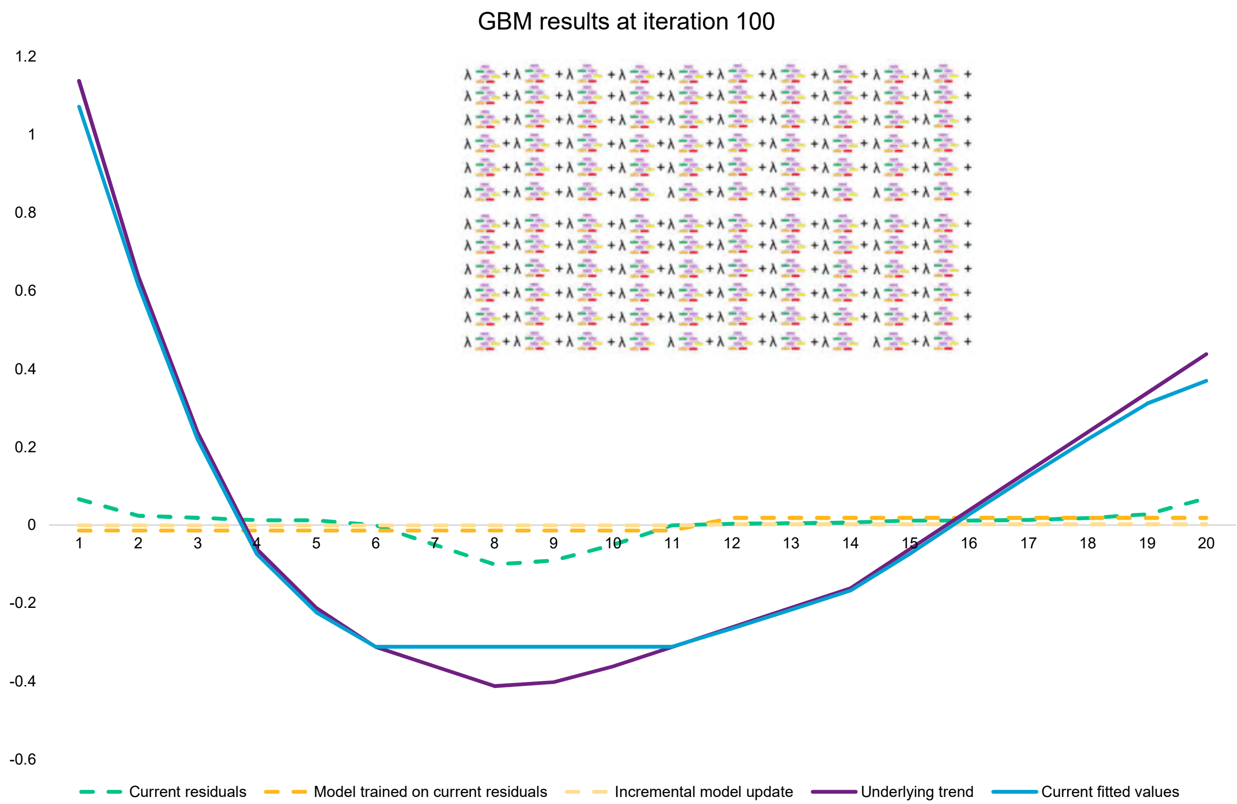
Non Parametric Modeling of Demand

How do they work – a simple example



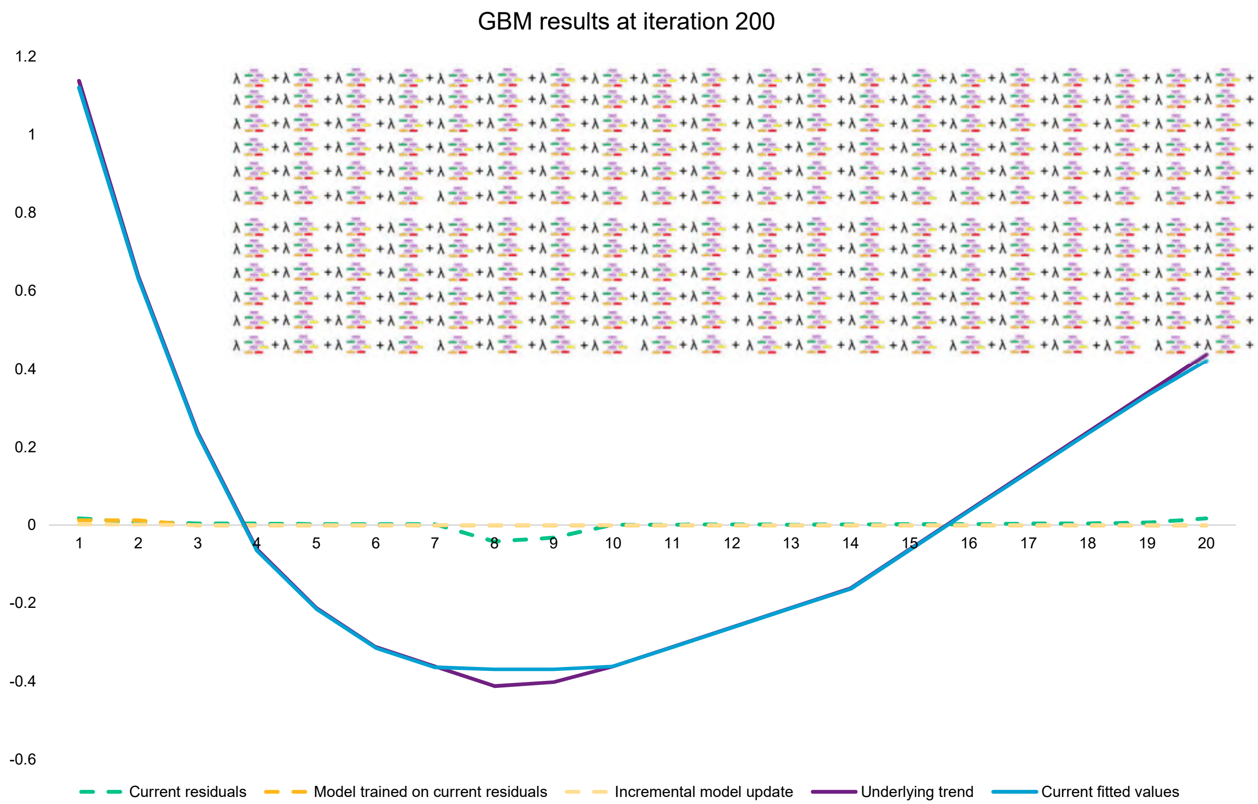
Non Parametric Modeling of Demand

How do they work – a simple example



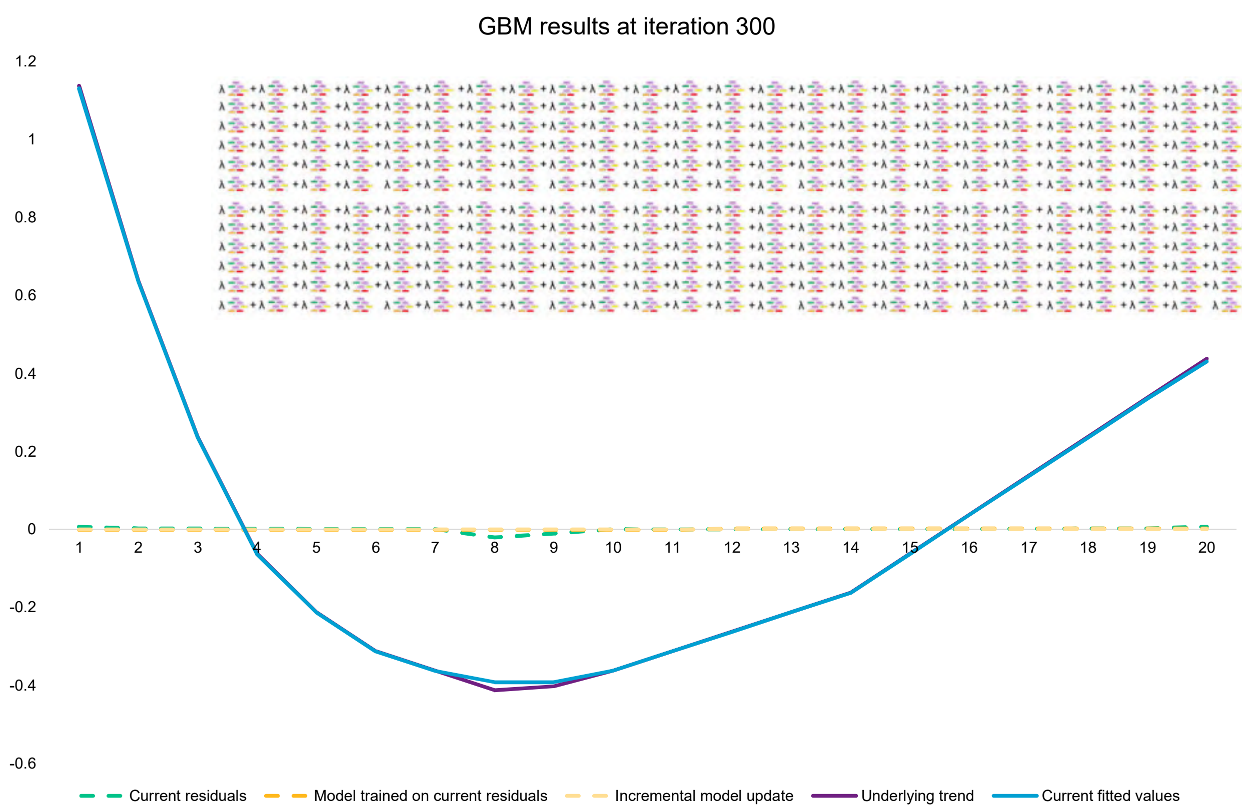
Non Parametric Modeling of Demand

How do they work – a simple example



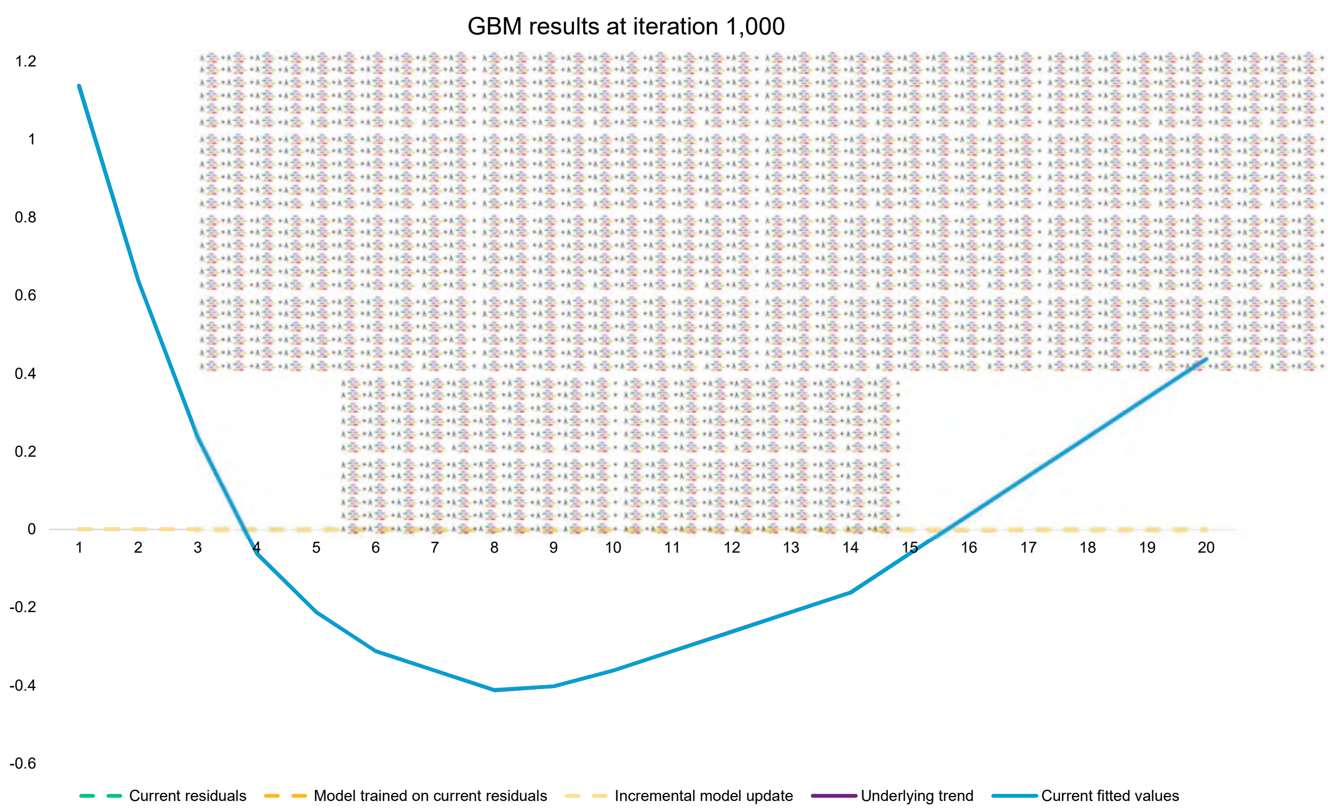
Non Parametric Modeling of Demand

How do they work – a simple example



Non Parametric Modeling of Demand

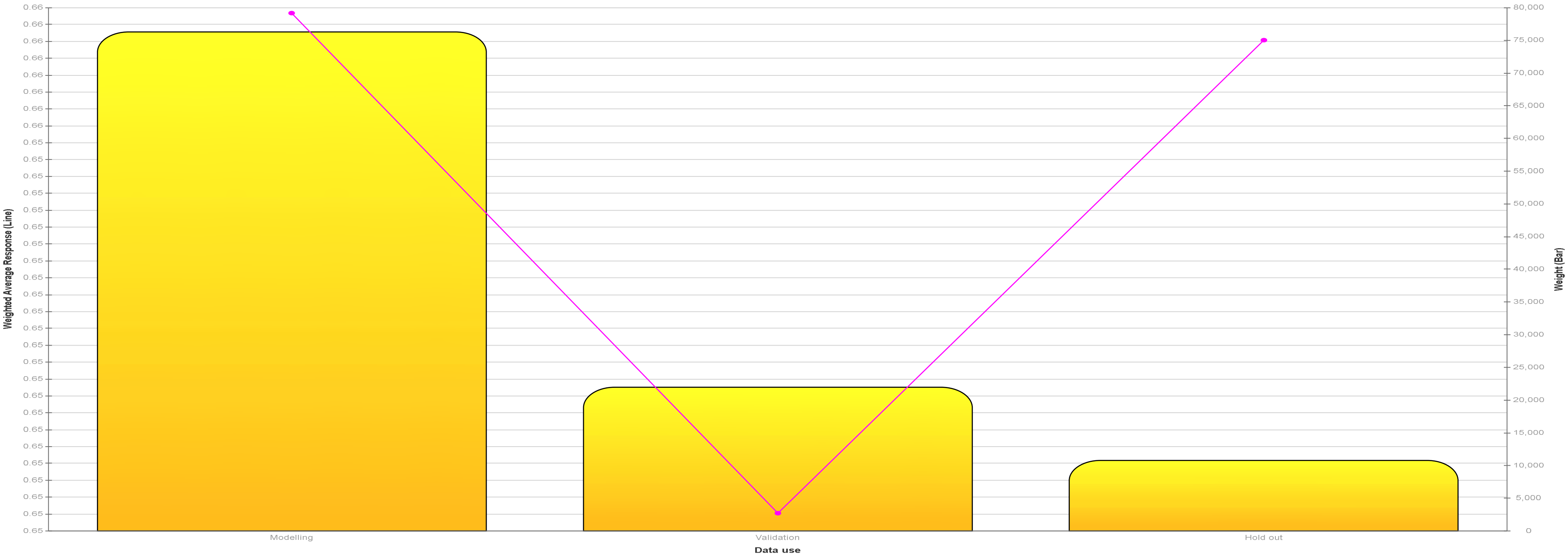
How do they work – a simple example



Non Parametric Modeling of Demand

Sampling is key

- Sampling – splits data into modelling/validation/holdout
  - Sampling needs to have memory if you are wanting to track model changes over time



Non Parametric Modeling of Demand

Calibrating the Hyperparameters

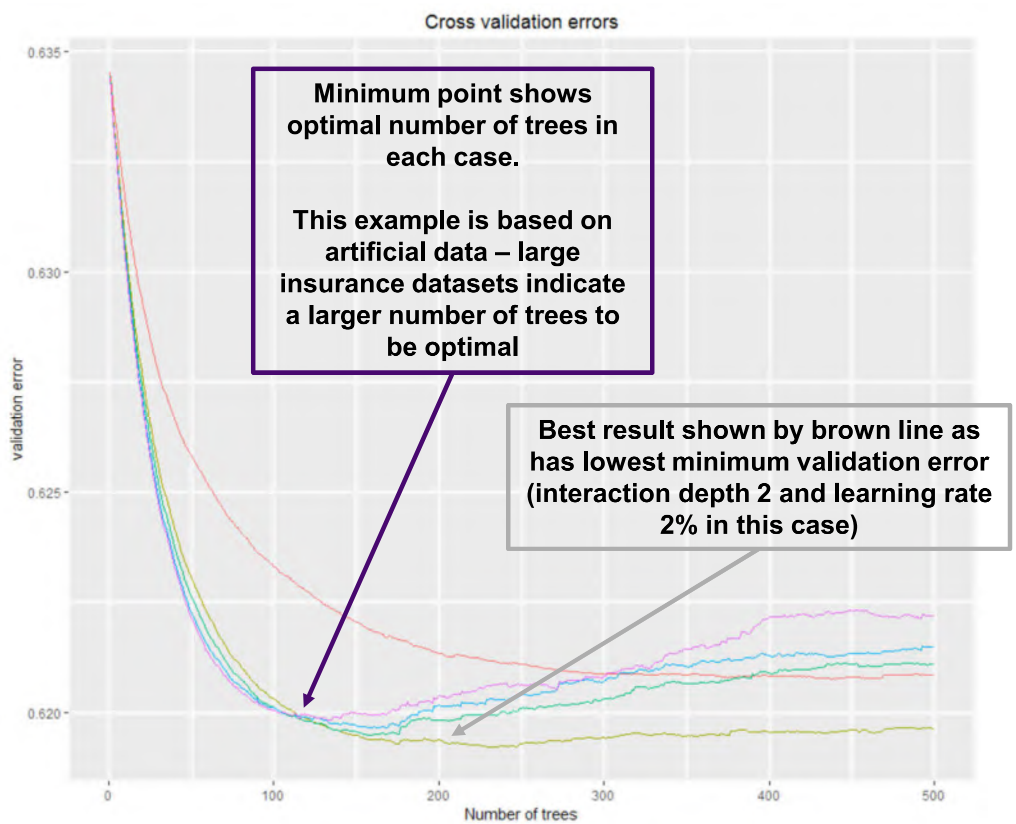
- Cross Validation
  - Further splits to the modelling data that is used to search for hyperparameters
  - Example – 3 fold validation – split modelling data into 3, fit on the purple, test on the pink, take the average

1	2	3
Fit	Fit	Test
Fit	Test	Fit
Test	Fit	Fit

- Resulting negative log-likelihood plots can be used to identify optimal choice

Non Parametric Modeling of Demand

Calibrating the Hyperparameters

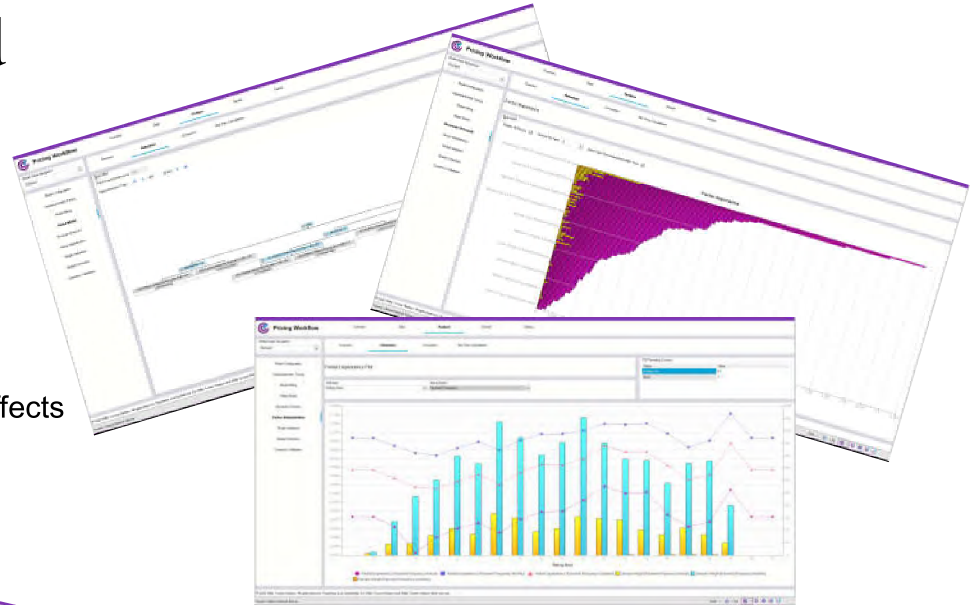


- Grid searches are used with more hyperparameter dimensions

	Min Eval Metric	Rank	Learning Rate	Max Depth	Min Child Weight	Min Split Loss	Number Of Trees
0	0.60035	9	0.120	4	3.700	5.630	400
1	0.60003	2	0.110	6	8.670	5.330	465
2	0.60019	6	0.140	6	2.570	5.600	446
3	0.60074	17	0.150	4	1.460	5.340	364
4	0.60022	7	0.130	5	1.790	5.440	423
5	0.60085	23	0.190	4	7.270	5.890	342
6	0.60049	12	0.100	4	9.410	4.090	400
7	0.60080	20	0.160	4	8.390	5.820	355
8	0.60045	11	0.200	5	7.880	5.470	392
9	0.60066	14	0.180	5	1.440	4.600	373
10	0.60135	27	0.160	3	1.270	5.880	300
11	0.60157	29	0.130	3	7.380	5.130	300
12	0.60089	24	0.190	4	5.830	5.970	351
13	0.60023	8	0.120	4	4.440	4.290	400
14	0.59980	1	0.100	6	2.730	4.820	457
15	0.60041	10	0.140	4	9.410	5.260	400
16	0.60076	18	0.180	4	9.150	4.110	352
17	0.60112	25	0.200	6	6.160	5.840	352
18	0.60059	13	0.160	6	9.400	4.980	415
19	0.60072	16	0.190	5	2.850	5.780	354
20	0.60136	28	0.150	3	3.950	5.190	300
21	0.60010	4	0.110	5	4.620	4.850	431
22	0.60083	22	0.190	4	6.380	4.860	354
23	0.60176	30	0.110	3	7.010	4.960	300
24	0.60122	26	0.190	3	8.040	4.910	300
25	0.60013	5	0.160	6	5.370	4.810	437
26	0.60082	21	0.180	5	9.460	4.680	374
27	0.60079	19	0.200	5	2.950	5.470	352
28	0.60066	15	0.160	5	2.700	4.770	384
29	0.60010	3	0.110	6	8.200	4.800	450

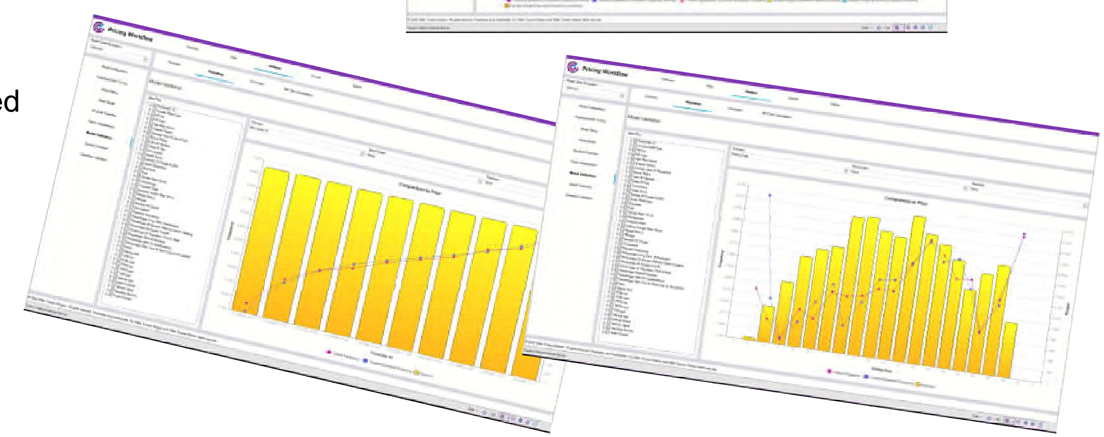
Non Parametric Modeling of Demand

Model Interpretation



- Structure
  - Resulting tree can be difficult to interpret
  - Using a layered GBM identifies contributions from each main effect and interaction term
  - Partial dependency plots are methods to express results across main effects and interactions

- Validation
  - Model should be assessed on the validation data set
  - Aggregate average observed vs. aggregate average expected
  - Lift charts for single model comparisons
  - Double lift charts for multiple model comparisons



Non Parametric Modeling of Demand

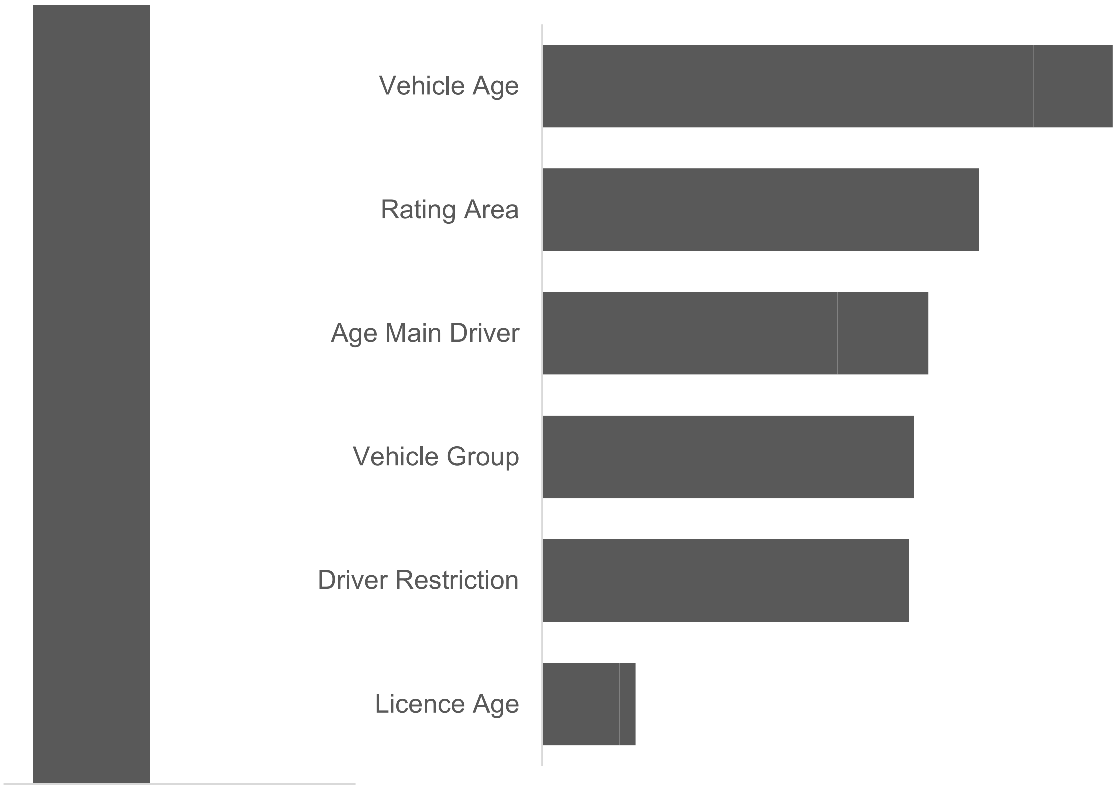
Using decision trees/GBMs/ML

- Advantages
  - Can handle non-linear relationships
  - Naturally models interactions without manual adjustments
  - Tends to produce better model fits

- Disadvantages
  - Not as common in the insurance industry
  - Can overfit the data without proper tuning
  - Large trees can be difficult to intepret

Non Parametric Modeling of Demand

GBM: Lack of Transparency

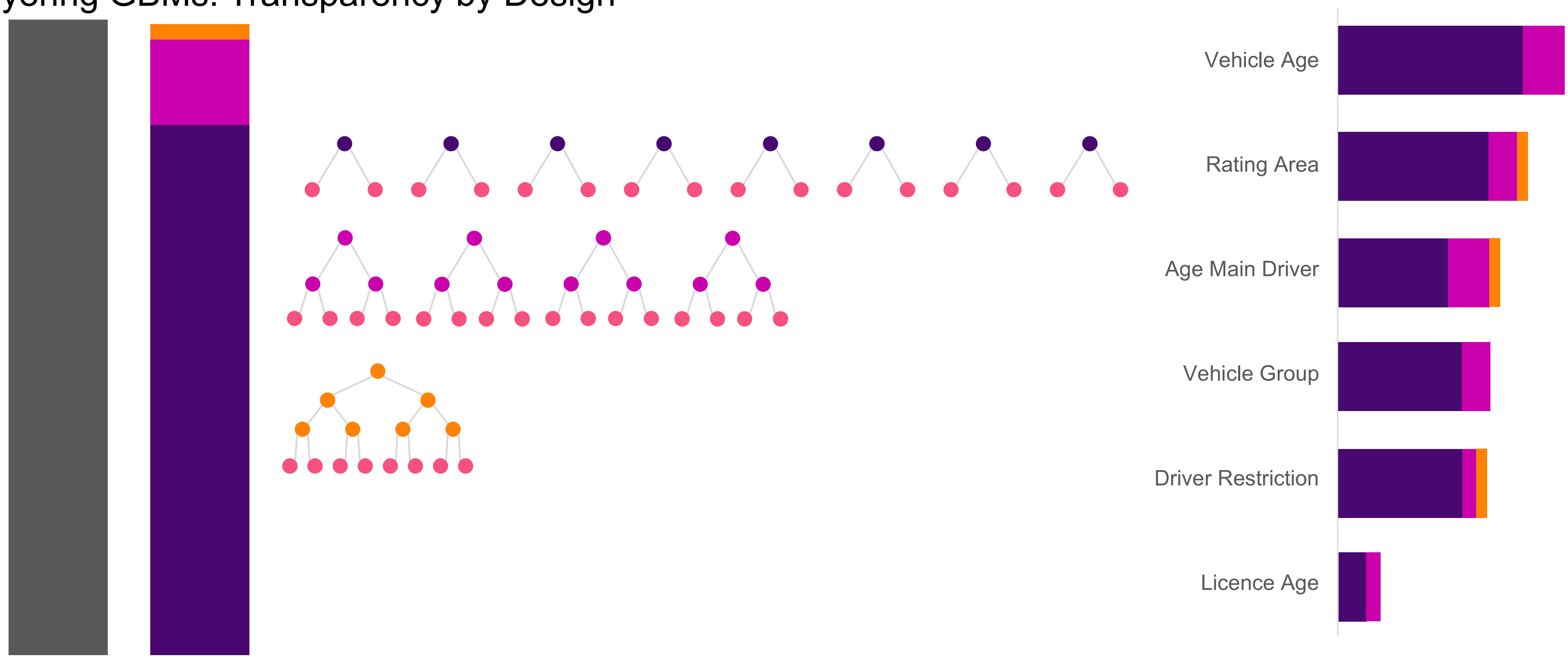


Limitations of Interpretation

- Large iterations results in too complex of a decision trees
- Factor importance: identifies single factor effects but does not distinguish the interaction terms
- Current tools are approximations
  - Partial Dependency Plots
  - SHAP values
  - H-statistics
  - ICE charts

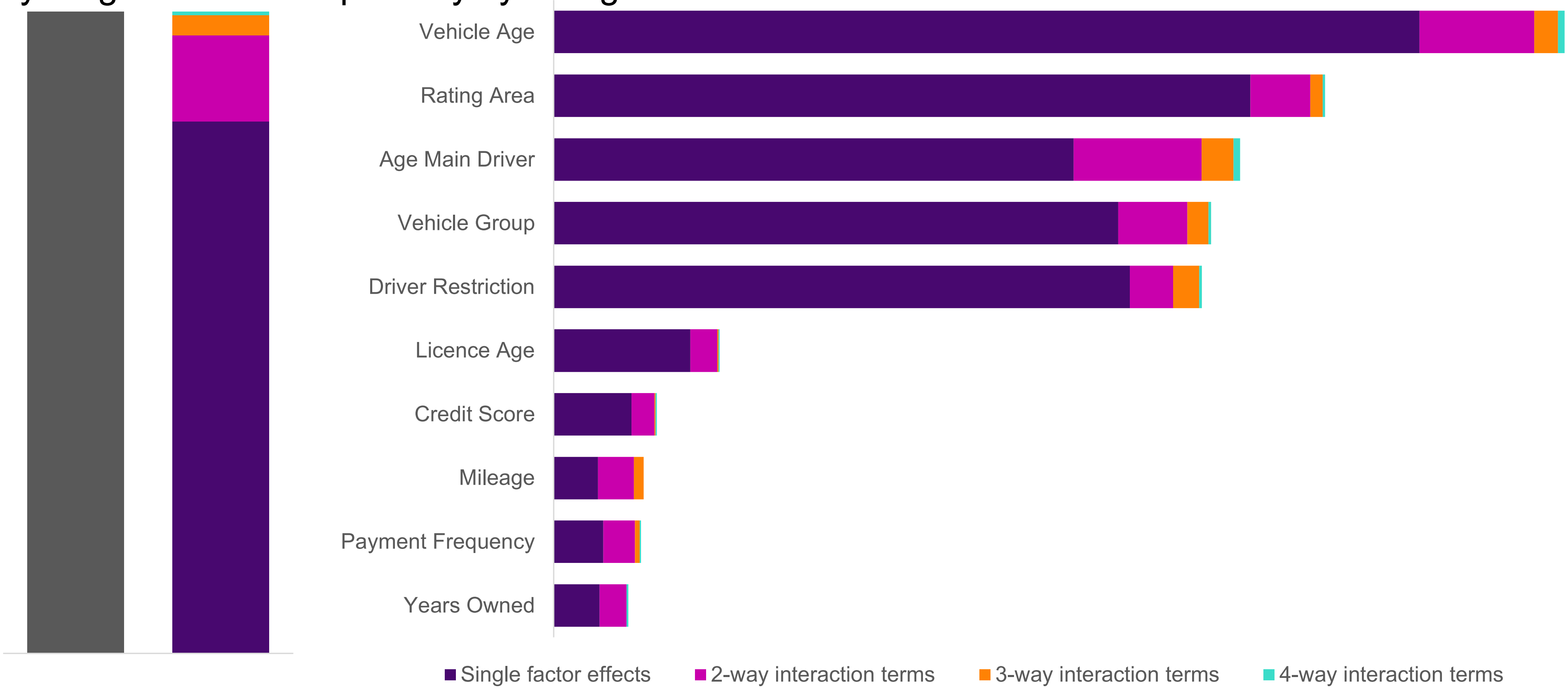
Non Parametric Modeling of Demand

Layering GBMs: Transparency by Design



Non Parametric Modeling of Demand

Layering GBMs: Transparency by Design



Non Parametric Modeling of Demand

Layering GBMs: Transparency by Design

Non Parametric Modeling of Demand

Layering GBMs: Transparency by Design

Vehicle Age
Rating Area
Driver Restriction
Vehicle Group
Age Main Driver
Licence Length Main Driver
Age Main Driver x Vehicle Age
Credit Score
Payment Frequency
Years Owned
Mileage
Rating Area x Vehicle Group
Age Main Driver x Mileage
Age Main Driver x Payment Frequency
Licence Length Main Driver x Vehicle Age
Driver Restriction x Vehicle Group
Age Main Driver x Gender Main Driver
Age Main Driver x Vehicle Group
Age Main Driver x Vehicle Age x Vehicle Group
Driver Restriction x Rating Area
Vehicle Age x Years Owned
Age Main Driver x Driver Restriction x Vehicle Age
Vehicle Age x Vehicle Group
Rating Area x Vehicle Age
Driver Restriction x Vehicle Age
Age Main Driver x Driver Restriction x Vehicle Group
Mileage x Payment Frequency
Licence Length Main Driver x Rating Area
Driver Restriction x Mileage x Vehicle Age
Credit Score x Rating Area
Mileage x Rating Area x Vehicle Age

Age Main Driver x Rating Area x Vehicle Group
Mileage x Vehicle Age
Credit Score x Vehicle Group
Credit Score x Vehicle Age
Gender Main Driver x Vehicle Group
Age Main Driver x Driver Restriction x Rating Area
Gender Main Driver
Age Main Driver x Driver Restriction x Gender Main Driver
Age Main Driver x Payment Frequency x Rating Area
Age Main Driver x Years Owned
Licence Length Main Driver x Years Owned
Credit Score x Mileage
Driver Restriction x Gender Main Driver
Age Main Driver x Driver Restriction
Driver Restriction x Mileage x Vehicle Group
Driver Restriction x Payment Frequency x Vehicle Group
Age Main Driver x Driver Restriction x Vehicle Age x Years Owned
Driver Restriction x Years Owned
Age Main Driver x Rating Area
Age Main Driver x Credit Score x Vehicle Age x Vehicle Group
Payment Frequency x Rating Area
Age Main Driver x Payment Frequency x Rating Area x Vehicle Age
Credit Score x Years Owned
Mileage x Rating Area
Age Main Driver x Credit Score
Age Main Driver x Rating Area x Vehicle Age x Vehicle Group
Rating Area x Years Owned
Payment Frequency x Vehicle Group
Payment Frequency x Vehicle Age
Driver Restriction x Licence Length Main Driver x Vehicle Group
Age Main Driver x Driver Restriction x Licence Length x Vehicle Age
Payment Frequency x Years Owned
Age Main Driver x Credit Score x Vehicle Age
Driver Restriction x Licence Length Main Driver

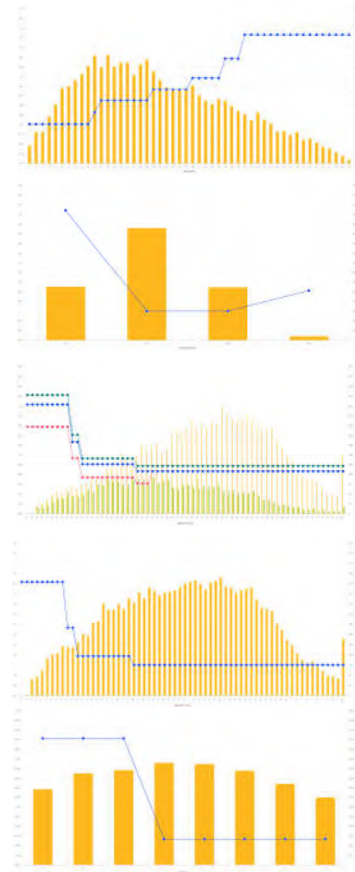
Non Parametric Modeling of Demand

Layering GBMs: Transparency by Design

Vehicle Age
Rating Area
Driver Restriction
Vehicle Group
Age Main Driver
Licence Length Main Driver
Age Main Driver x Vehicle Age
Credit Score
Payment Frequency
Years Owned
Mileage
Rating Area x Vehicle Group
Age Main Driver x Mileage
Age Main Driver x Payment Frequency
Licence Length Main Driver x Vehicle Age
Driver Restriction x Vehicle Group
Age Main Driver x Gender Main Driver
Age Main Driver x Vehicle Group
Driver Restriction x Rating Area
Vehicle Age x Years Owned
Vehicle Age x Vehicle Group
Rating Area x Vehicle Age
Driver Restriction x Vehicle Age
Mileage x Payment Frequency
Licence Length Main Driver x Rating Area
Credit Score x Rating Area

Mileage x Vehicle Age
Credit Score x Vehicle Group
Credit Score x Vehicle Age
Gender Main Driver x Vehicle Group
Gender Main Driver
Age Main Driver x Years Owned
Licence Length Main Driver x Years Owned
Credit Score x Mileage
Driver Restriction x Gender Main Driver
Age Main Driver x Driver Restriction
Driver Restriction x Years Owned
Age Main Driver x Rating Area
Payment Frequency x Rating Area
Credit Score x Years Owned
Mileage x Rating Area
Age Main Driver x Credit Score
Rating Area x Years Owned
Payment Frequency x Vehicle Group
Payment Frequency x Vehicle Age
Payment Frequency x Years Owned
Driver Restriction x Licence Length Main Driver

Elastic Net GLM with automatic factor simplification



Rating Plan Development

Customer Behavior

Application 1: Offbalancing a 5% Rate Increase

Comparing traditional to demand weighted

- Inforce

Inforce Policy	Current Premium	Proposed Premium	Expected Losses	Retention
1	974	1,023	682	0.88
2	958	1,006	680	0.89
3	950	998	684	0.90
4	968	1,016	707	0.91
5	986	1,035	730	0.92
6	955	1,003	716	0.93
7	965	1,013	733	0.94
8	963	1,011	742	0.95
9	973	1,022	759	0.96
10	961	1,009	807	0.97
Total	9,653	10,136	7,240	

- Quote

Quote	Proposed Premium	Expected Losses	Conversion	Retention
1	1,044	835	0.20	0.80
2	1,048	891	0.22	0.82
3	1,063	914	0.24	0.85
4	1,079	950	0.26	0.88
5	1,095	986	0.28	0.92
Total	5,329	4,575		

Application 1: Offbalancing a 5% Rate Increase

Comparing traditional to demand weighted

- Traditional view focuses on inforce dataset
  - Current loss ratio = 7,240/9,653 = 0.750
  - Proposed loss ratio = 7,240/10,136 = 0.714

- Alternative view focuses on inforce and quote datasets AND considers demand
  - Demand-weighted proposed premium on inforce dataset = 1,023(0.88) + … + 1,009(0.97) = 9,376
  - Demand-weighted expected losses on inforce dataset = 6,707
  - Demand-weighted proposed premium on quote dataset
    = 1,044(0.20)(0.80) + … + 1,095(0.28)(0.92) = 1,102
  - Demand-weighted expected loss on quote dataset = 952
  - Current loss ratio = 7,240/9,653 = 0.75
  - Demand-weighted proposed loss ratio = (6,707 + 952)/(9,376 + 1,102)
    = 0.731

- **Key point: Traditional off-balance approach may lead to insufficient rate**

Simulation and Search

Application 2: Multi period Simulations

Short and long term impacts of rate decisions

- Renewal dataset & quote dataset
- Time Horizon – four periods, each lasting six months
- Quote growth rate – 5% each period
- Quotes do not enter simulation until new rates go into effect at the beginning of period 1
- Quote distribution constant over time
- Aging assumptions
  - Property age by 1 every other period
  - Property values increase by local CPI
- Current loss ratio is 75%
- Ignore trend
- Scenarios
  - 5% base rate decrease
  - 15% decrease small business owners off-balanced to an overall 5%

Application 2: Multi period Simulations

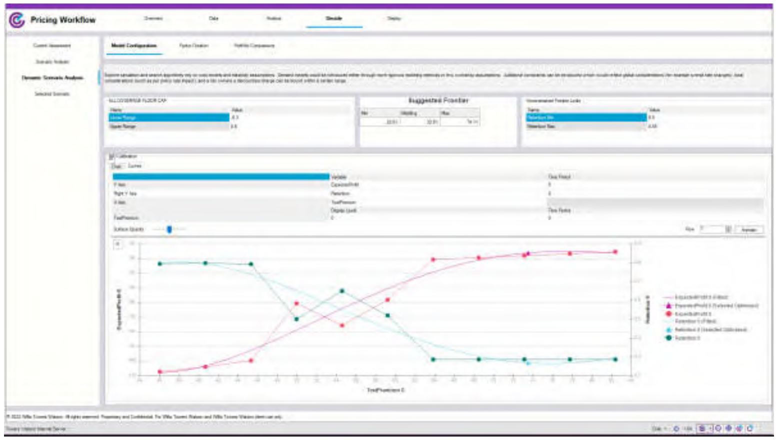
Key point: A multi-period view is often needed to make the best rate decision

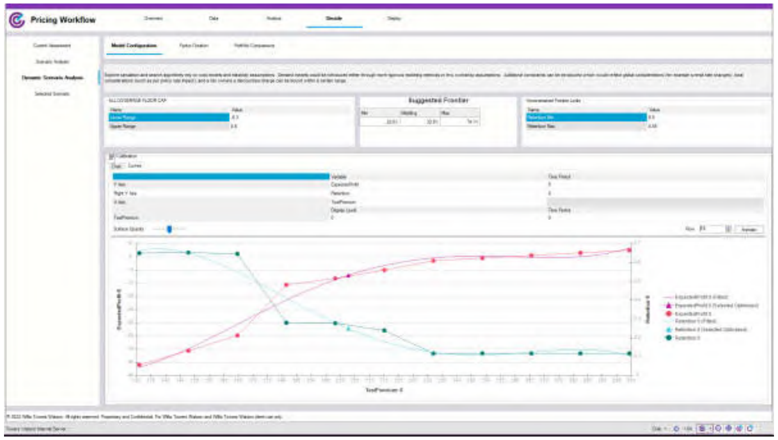
Quotes + Renewals							
	Period	**Policies Offered**	**Policies Written**	**Policies Retained**	**Earned Premium**	**Profit Margin**	**Absolute Profit**
Scenario 1	0	50,000	50,000	44,000	$35,250,000	2.5%	$881,250
	1	64,000	49,493	45,956	$34,486,258	2.3%	$810,152
	2	66,956	51,723	49,064	$36,412,258	2.3%	$822,930
	3	71,114	55,122	52,296	$38,800,399	2.2%	$842,146
	4	75,449	58,657	54,722	$40,949,798	2.2%	$888,759
Scenario 2	0	50,000	50,000	44,000	$35,250,000	2.5%	$881,250
	1	64,000	49,646	46,030	$34,729,064	2.3%	$812,026
	2	67,030	51,958	49,135	$36,692,114	2.4%	$891,271
	3	71,185	55,363	52,352	$39,087,466	2.5%	$985,029
	4	75,505	58,890	54,755	$41,236,423	2.6%	$1,076,159

Application 3: Dynamic Scenario Analysis

Simulating rate scenarios

- Simulate premium change activity for each individual policy/quote in the data set
- Track KPIs (profit/volume/etc) for each simulation
- Observe how the simulation impacts different policies differently
- Compile all simulations into an array

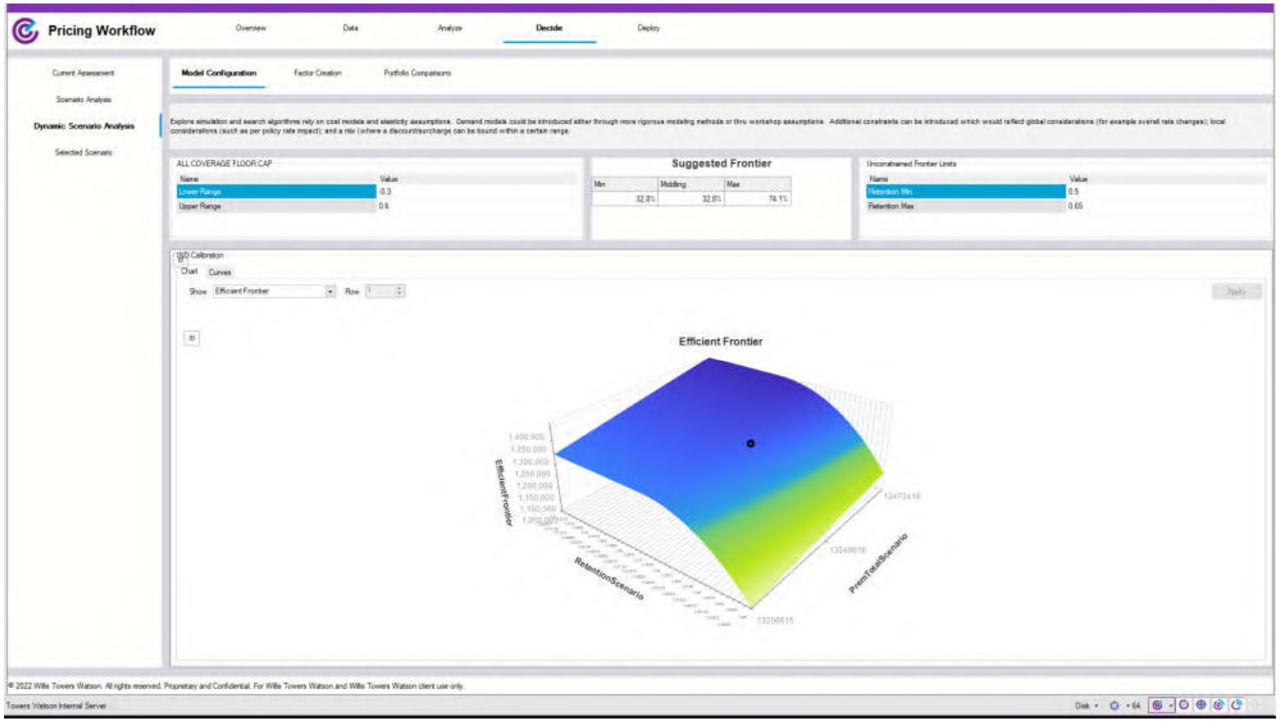




Application 3: Dynamic Scenario Analysis

Simulating rate scenarios

- Incorporate other constraints into the simulation space
  - Cross sell preferences
  - Traditional actuarial balancing rules (e.g. SUM(Current) = SUM(Proposed)
- Use simulations and constraints to assess portfolio options to better determine the right rate for the risk



Thank you

Serhat Guven

Managing Director

Serhat Guven is the global leader of WTW's Insurance Consulting and Technology's product, pricing, claims, and underwriting technology suite for personal lines.

He and his team are responsible for the design and development of the analytical, decisioning and deployment technology platform, Radar. Radar is a technology solution that is uniquely designed to help insurers respond to significant industry trends.

Prior to his current role, Serhat was WTW's Regional Line of Business Leader for the Americas and was responsible for a full range of consulting services and software solutions to insurance companies.

Michael Chen

Director

Michael Chen is a Director working in WTW’s Chicago office. Michael primarily works in predictive analytics and has 20 years of actuarial experience in the property and casualty insurance industry.

He is a Fellow of the Casualty Actuarial Society, a CSPA and a member of the American Academy of Actuaries.

He currently serves as a member of the Committee on Professionalism Education.