2018 Casualty Loss Reserve Seminar

Improving Actuarial Reserve Analysis Through Predictive Analytics

By Manolis Bardis and Paul Bailey

September 6, 2018

Willis Towers Watson

CAS Antitrust Notice

The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.

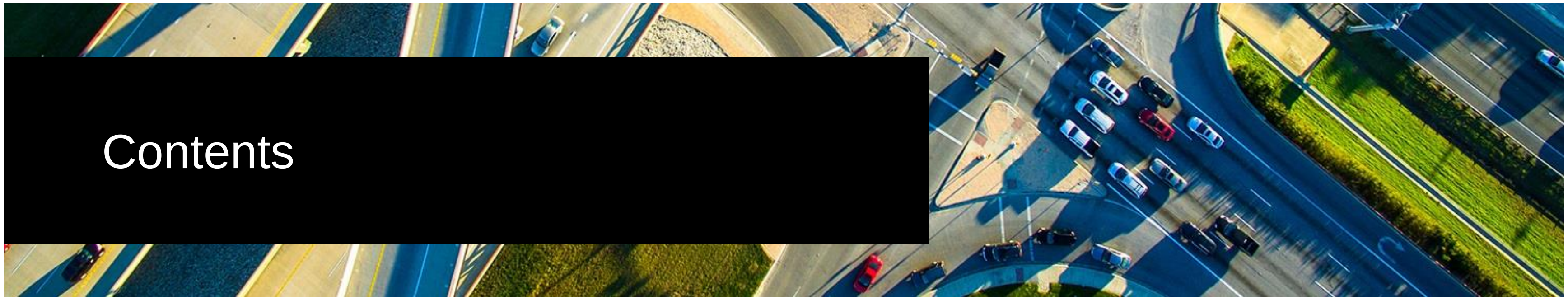
Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com



Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Applications of predictive models in the insurance sector

Marketing and Distribution → Customer Management → Underwriting and Risk Management → Pricing → Reserving and Claims Management → Asset Management

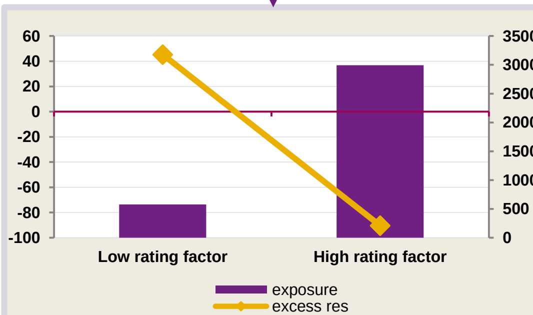
Product development (UBI)

Retention segmentation and program design

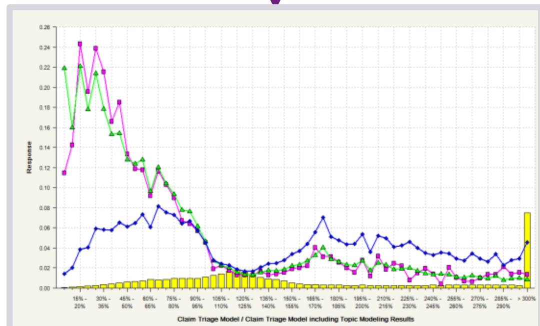
Agent/Broker performance evaluation

Fraud detection

P&C to Life cross-selling



Web-scraping and feature design in commercial lines



Reserving and large loss modeling

How advanced analytics will transform claim management

Current and planned applications

	Now	Two years
Evaluation of claims for fraud potential	26%	82%
Claim triage (identify complex claims to triage workflow)	26%	80%
Evaluation of claims for litigation potential	15%	74%
Evaluation of claims for subrogation potential	13%	62%

Applications of predictive models in reserving

- Validate traditional reserve estimates and assumptions
- Understand the influence of individual claims on reserves
- Assist adjusters to set individual case reserve estimates
- Micro-level stochastic loss reserving
- Predict large losses
  - Underwriting
  - Scenario test effect of different XOL reinsurance treaties
  - Economic capital models
- Synergy with claims triage and other claims analytics efforts

An evolution of predictive models in reserving

Developing comfort, seeking greater insights

Predictive models were fit to the same data contained in a reserve triangle

- Using incremental losses instead of cumulative
- Only two predictors: Report year and lag
- Given appropriate statistical assumptions, the forecasts replicated chain ladder using volume-weighted average of the link ratios

Calendar Year was introduced as a predictor to enable calendar year changes in inflation to be reflected in the reserve estimation process

- Still need to extrapolate beyond the latest diagonal

Reserve accuracy of models fit to aggregated data still hinges on consistency in the data

Natural next steps

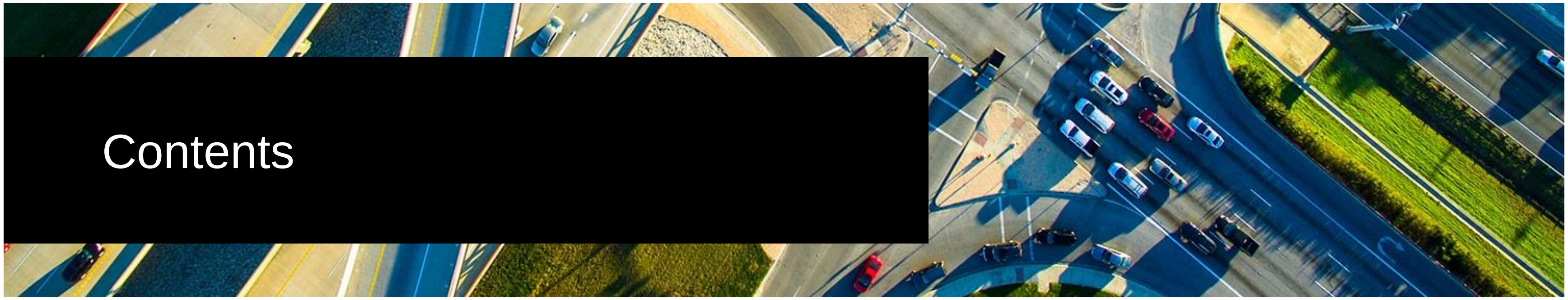
- Predictive modeling could be used to group claims with similar development behavior, such that separate triangles could be analyzed
- Predictive models could be fit to individual claims data

Why consider predictive models in reserving & claims management

Key points

- **Accuracy of traditional reserving methods hinges on consistency**
  - Claim closure rate
  - Case reserve adequacy
  - Inflation
  - Reinsurance
- **Traditional methods do not provide insights into the drivers of claim cost**
  - How much does age affect the cost of WC claims?
  - What is the impact of opioid usage on the cost of claims?
  - How much did reform measures impact claim costs?

Predictive models can address these challenges



Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

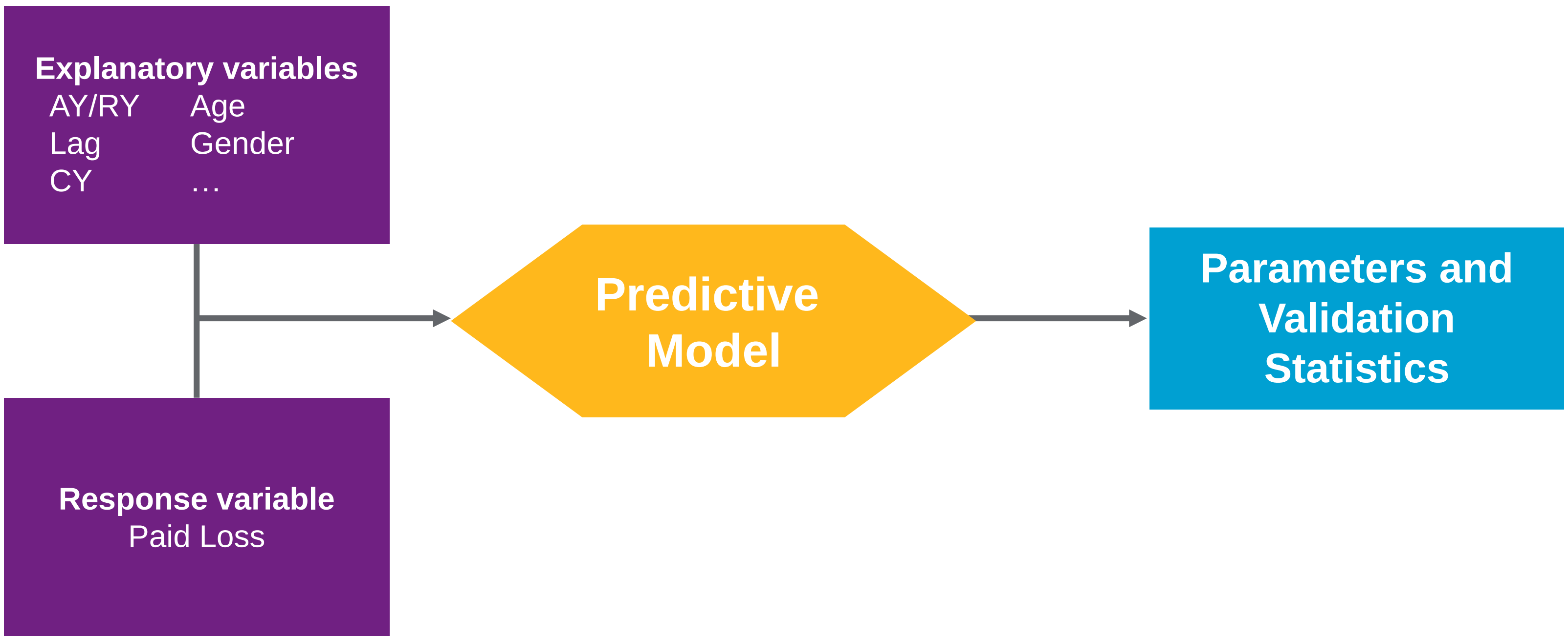
Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Predictive Models

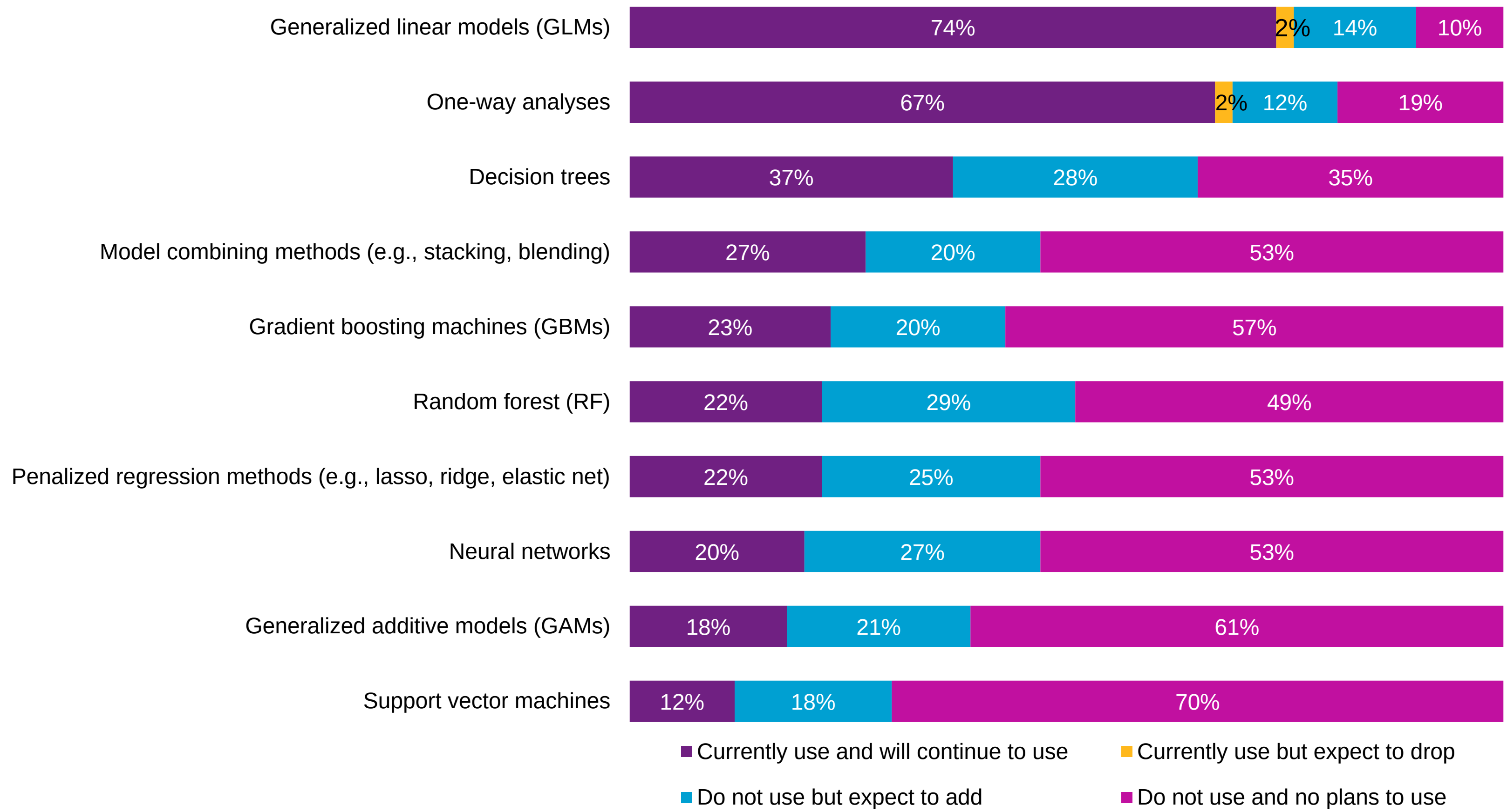
Definition

Statistical model to predict a response variable using a series of explanatory variables



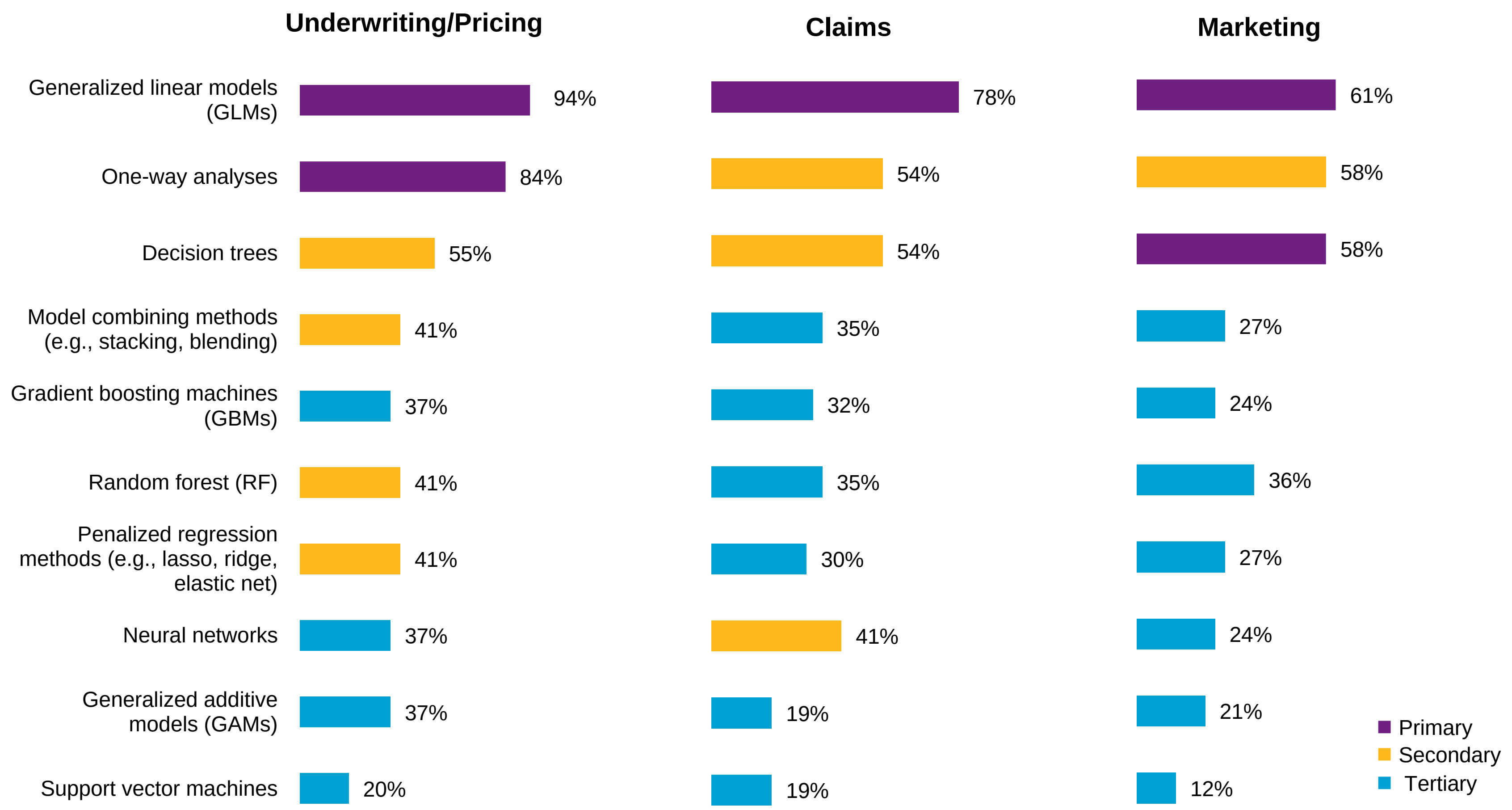
Same techniques apply regardless of what is being modeled

Which advanced analytical modeling techniques do you currently use, and which do you plan to use in the next two years? (Q.16)



Base: U.S. respondents using advanced analytics (n = 51)

For which business applications do you use or plan to use these modeling techniques? (Q.17)



Base: U.S. respondents using advanced analytics for underwriting/pricing (n = 49), claims (n = 37) and/or marketing (n = 33)

Types of predictive models

Statistical regression methods (e.g., GLM)

Response Variable = Systematic Component + Random Component

y = h (Linear Combination of Explanatory Variables) + Error

$g=h^{-1}$ is called the LINK function and is chosen to measure the signal most accurately

Combination of explanatory variables is the model structure

Error should reflect underlying process and comes from the exponential family

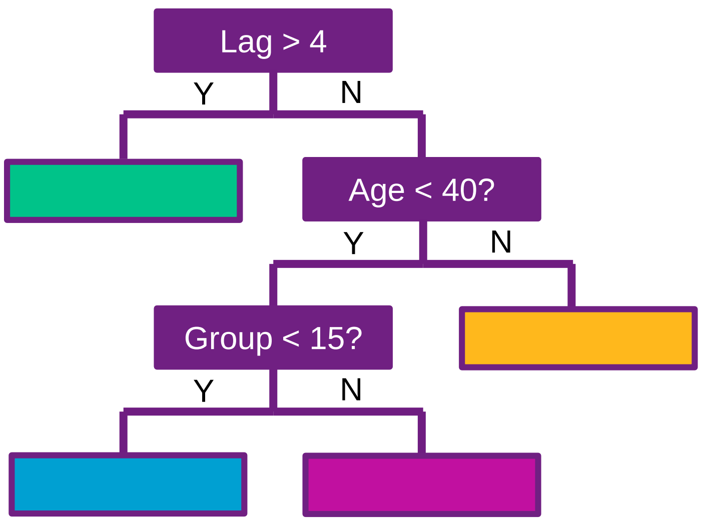
Output is set of parameters and a series of diagnostics

Types of predictive models

Machine learning approaches (e.g., random forest)

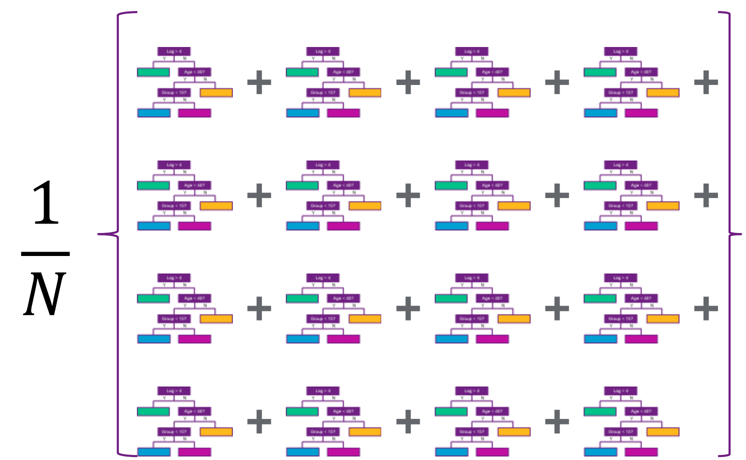
A tree

$$f_i(x)$$



A random forest

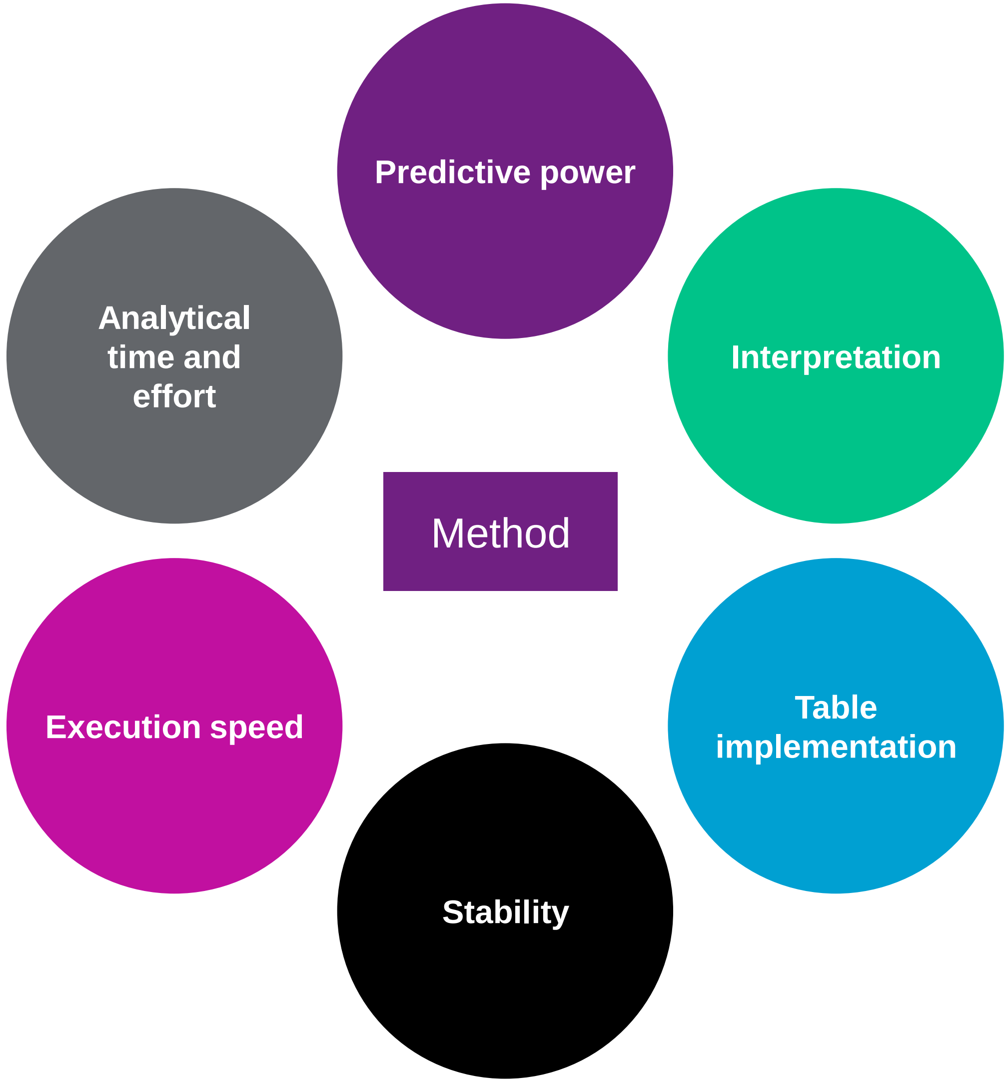
$$f(x) = \frac{1}{N}\sum\nolimits_{n=1}^{N} f_n(x)$$



Output of a random forest is the average of a bunch of independent trees

Considerations

Evaluating predictive models for an application



One model may be most useful aiding another

Machine learning methods can be used in their own right (to forecast development) or can improve certain aspects of the analysis

- Multivariate adaptive regression splines to identify where separate models should be built (e.g., by lag or segment)
- Penalized regression (e.g., elastic net) to select factors to include in analysis
- Topic modeling to create new structured data fields
- GBMs or neural networks to validate regression results

Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

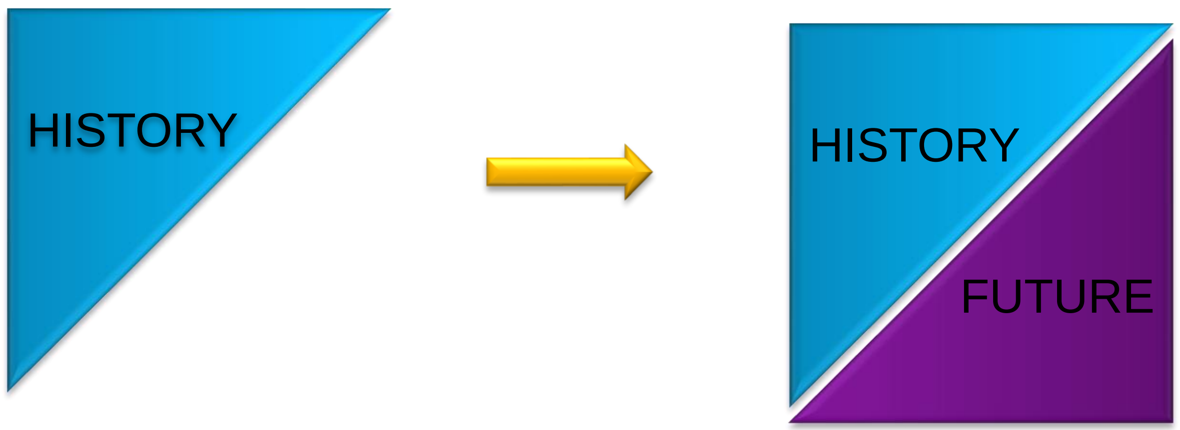
T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Actuarial Reserving in a Nutshell

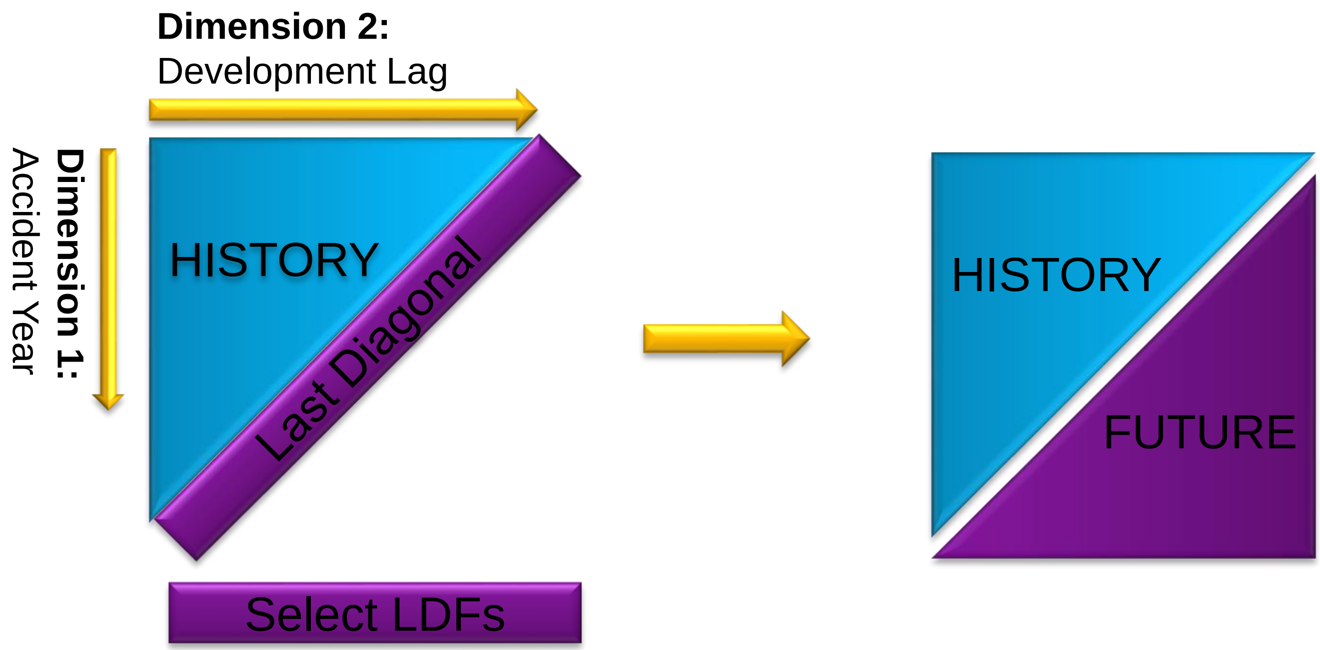
Traditional actuarial reserving methods has been conceptually described as a process of squaring up a triangle:



The **GLM Reserve** method is no different. Estimate future results based on information from historical.

Why GLM?

- Traditional Chain Ladder method focuses on the development Lag dimension to derive estimates:



- Each future estimate can be derived based on the selected development factors.

Why GLM?

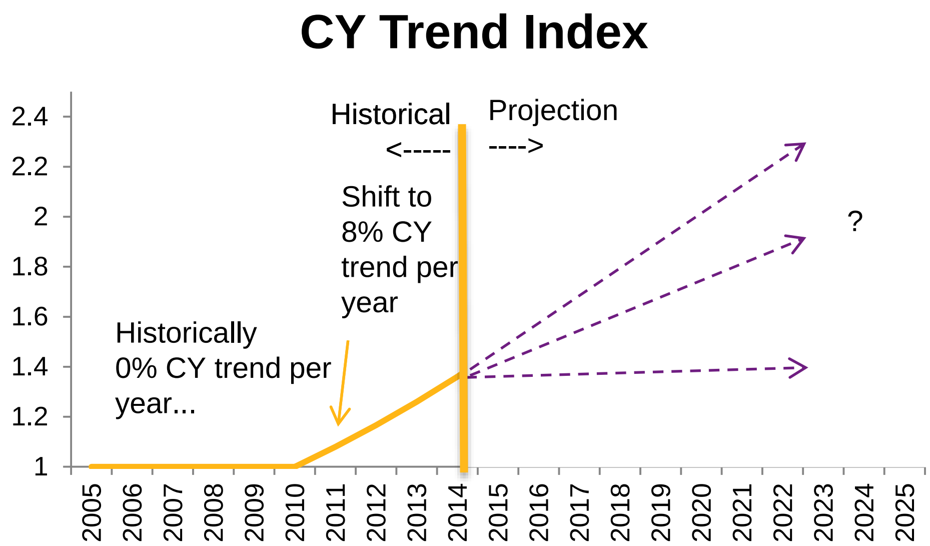
- However, one major limitation with chain ladder is that it does not adjust for accident or calendar year effects

- Examples include:
  - New claims handling process
  - Changing settlement pattern
  - Legislative/Regulatory changes

- GLM Reserving allows us to introduce two additional dimensions
  - **Dimension 1: Accident Year**
  - Dimension 2: Development Lag
  - **Dimension 3: Calendar Year**

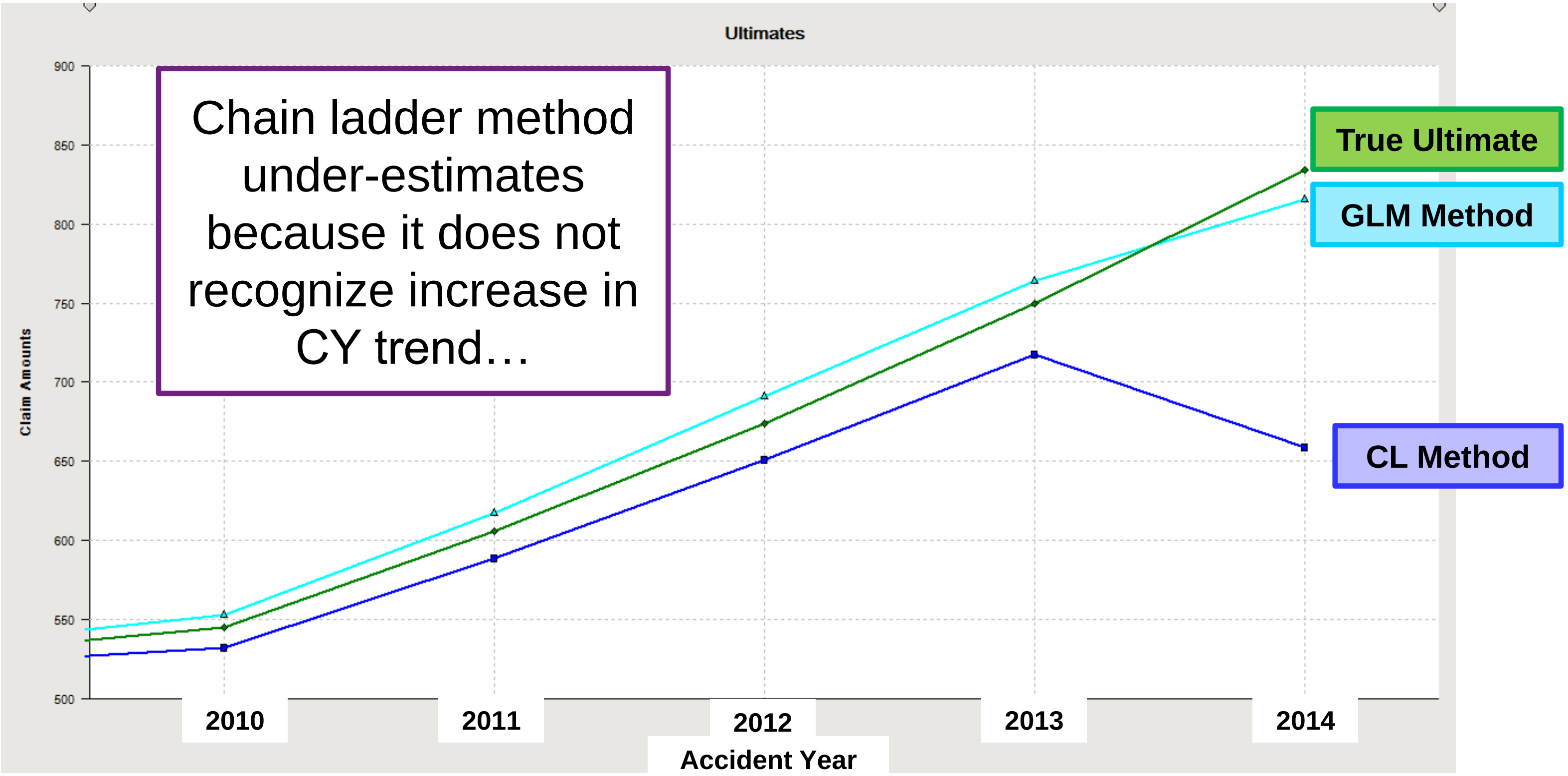
Case Study Example

- Lets quickly go through an illustrative example to demonstrate the impact of calendar year effects using a chain ladder method vs GLM reserving method
- Case Study introduces a calendar year trend in the most recent periods

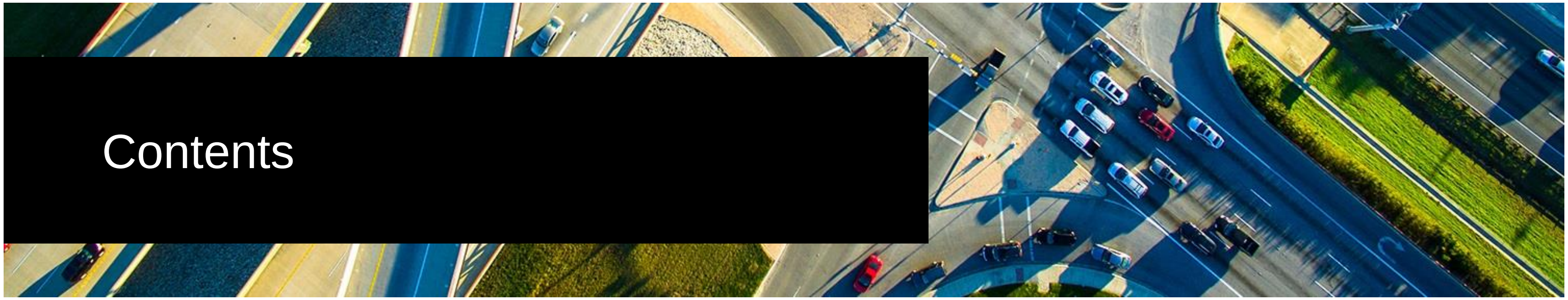


Case Study Example

- Impact can be significant. In this example, the difference from unpaid is only 4% for GLM Method versus -22% difference for Chain Ladder



- Improved estimates



Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

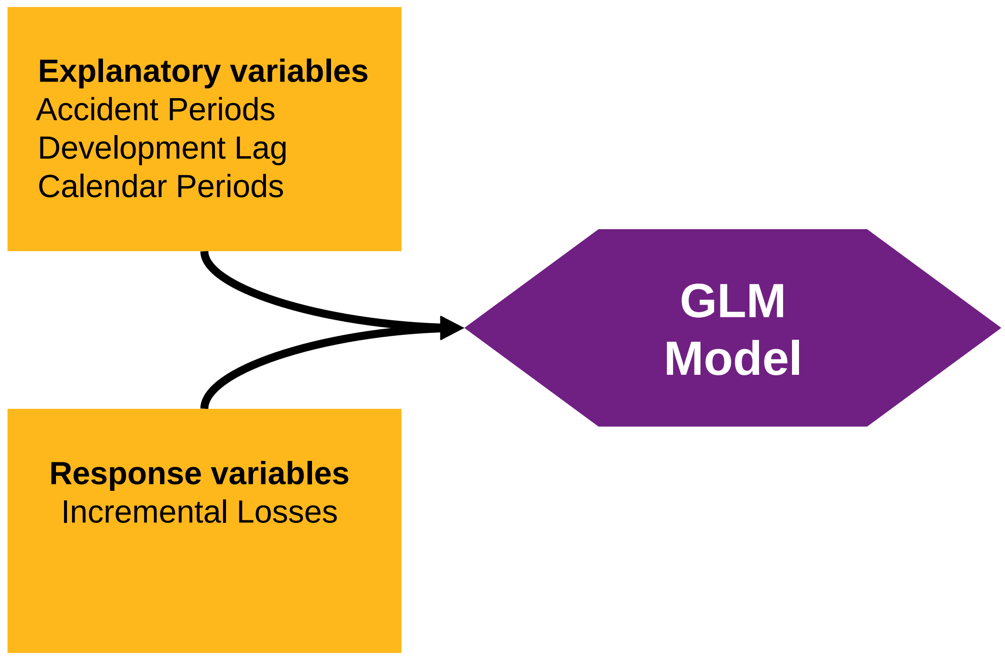
T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Predictive Models

Multivariate statistical model to predict a response variable using a series of explanatory variables

Explanatory variables
- Variables that help explain what we are seeing

Response variables
- Variables that respond to explanatory variables and we are trying to predict



We will use the explanatory variables to try and explain the behavior of incremental losses

Generalized Linear Models (GLMs)

GLMs are a flexible and sophisticated predictive modeling technique. There are two components:

Response Variable **(Incremental Loss)** **=** **Systematic Component** **(Explainable part of model)** **+** **Random Component**

y = h(Linear Combination of Parameters) + Error

1

Link Function

$g=h^{-1}$ is called the LINK function and is chosen to measure the signal most accurately

2

Combination of parameters Used to explain the response variable

3

Error Structure

Error function should reflect underlying process variance

Practical User Considerations

Selecting a Link Function & Error Structure

Options for Link Function

Log

- Assumes parameters interact multiplicatively
- Similar assumption to link ratio approach

Identity

- Use when relationship between parameters are additive in nature rather than multiplicative
- Using additive reserving methods

Practical User Considerations

Selecting a Link Function & Error Structure

Options for Error Structure

Normal or Gamma

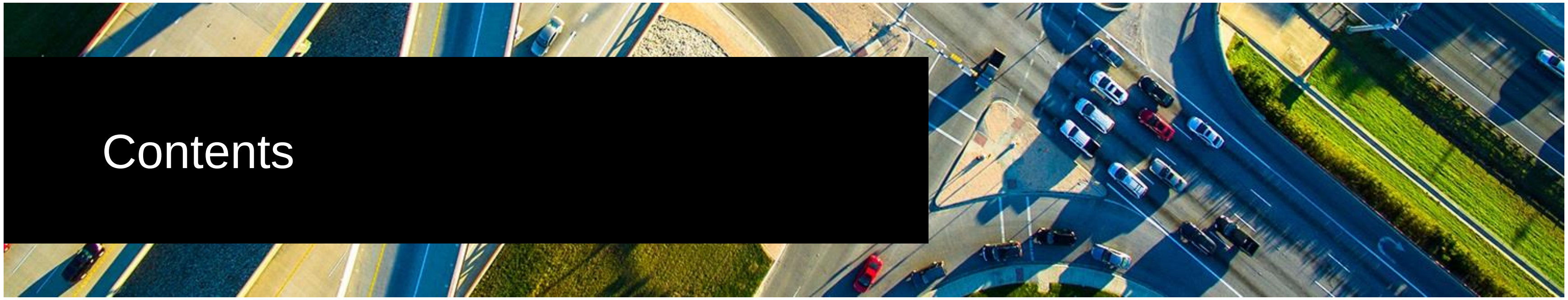
- Normal distribution assumes that all observations have the same fixed variance
- Gamma distribution assumes that the variance increases with the square power of the expected value of each observation

Poisson Scale Free

- A.k.a. "Over-dispersed Poisson" Distribution
- Mean = λ
- Variance = λ x Scale factor
- Allows variance to be lesser/greater than the mean

Poisson – Scale = 1

- Strict definition of Poisson distribution is applied, mean must equal the variance
- It assumes that the variance increases with the expected value of each observation



Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Section Introduction

In this section, we will cover the following:

- Start with 2-dimensional approach
- Show all years volume weighted average vs GLM
- Show how any cell in the historical triangle is linear combination of beta parameters

A simple example

In order to "demystify" the GLM reserve model, we will walk through a basic example and show how future estimates are calculated:

- Start with building a 2 dimensional GLM reserve model:

  Dimension 1 = Accident Year

  Dimension 2 = Development Lag
- Show that results are comparable to Chain Ladder Method using all years volume weighted average

A simple example

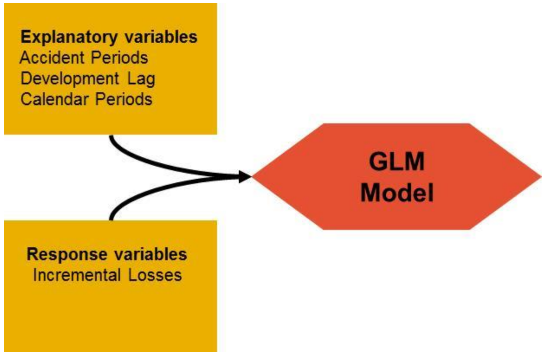
Incremental Paid Loss Triangle

Accident Year	12m	24m	36m	48m	60m	72m	84m	96m	108m	
2005	92	265	47	24	14	7	5	5	6	3
2006	95	273	49	25	12	8	6	6	7	
2007	98	281	50	22	14	9	7	7		
2008	100	290	46	24	15	10	8			
2009	103	288	51	27	17	11				
2010	72	321	57	30	19					
2011	80	357	64	33						
2012	89	397	71							
2013	98	441								
2014	110									

GLM reserve method is based on predicting the response variable, **incremental** losses.

$Y_{AY,DL}$ = Incremental loss
AY = Accident Year AY,
DL = Development Lag, DL

Example:
$Y_{2011,12m} = 80$



A simple example

Incremental Paid Loss Triangle

Accident Year	12m	24m	36m	48m	60m	72m	84m	96m	108m	120m
2005	92	265	47	24	14	7	5	5	6	3
2006	95	273	49	25	12	8	6	6	7	
2007	98	281	50	22	14	9	7	7		
2008	100	290	46	24	15	10	8			
2009	103	288	51	27	17	11				
2010	72	321	57	30	19					
2011	80	357	64	33						
2012	89	397	71							
2013	98	441								
2014	110									

- Any cell in the historical triangle is linear combination of “beta” parameters
- Incremental losses are related to explanatory variables multiplicatively
- Resulting model gives exactly the same forecast as the chain ladder model

$$Y_{AY,DL} = EXP\ (\beta_0 + \beta_{AY} + \beta_{DL}) + \varepsilon$$

Log link function

Linear combination of explanatory variables predicts incremental losses, based on AY and DL

A simple example

		β_{12}	β_{24}	β_{36}	β_{48}	β_{60}	β_{72}	β_{84}	β_{96}	β_{108}	β_{120}
	Accident Year	**12m**	**24m**	**36m**	**48m**	**60m**	**72m**	**84m**	**96m**	**108m**	**120m**
β_{05}	2005	92	265	47	24	14	7	5	5	6	3
β_{06}	2006	95	273	49	25	12	8	6	6	7	
…	2007	98	281	50	22	14	9	7	7		
	2008	100	290	46	24	15	10	8			
β_{09}	2009	103	288	51	27	17	11				
β_{10}	2010	72	321	57	30	19					
β_{11}	2011	80	357	64	33						
…	2012	89	397	71							
β_{13}	2013	98	441								
β_{14}	2014	110									

Begin with a Base Parameter, β_0

We will choose Accident Year 2005, Development Lag 12 months as the base parameter

Why use a Base Parameter?

Needed to allow for model convergence

Setting a base parameter reduces the number of variables by 1

A simple example

	Accident Year	12m	24m	36m	48m	60m	72m	84m	96m	108m	120m
	2005	92	265	47	24	14	7	5	5	6	3
β_{06}	2006	95	273	49	25	12	8	6	6	7	
…	2007	98	281	50	22	14	9	7	7		
	2008	100	290	46	24	15	10	8			
β_{09}	2009	103	288	51	27	17	11				
β_{10}	2010	72	321	57	30	19					
β_{11}	2011	80	357	64	33						
…	2012	89	397	71							
β_{13}	2013	98	441								
β_{14}	2014	110									

Explanatory Variables

Dimension 1 = Accident Year

β_{11} = Multiplicative parameter that describes accident year 2011

$$Y_{11,DL} = EXP(\beta_0 + \beta_{11} + \beta_{DL}) + \varepsilon$$

A simple example

		β_{24}	β_{36}	β_{48}	β_{60}	β_{72}	β_{84}	β_{96}	β_{108}	β_{120}
	12m	**24m**	**36m**	**48m**	**60m**	**72m**	**84m**	**96m**	**108m**	**120m**
2005	92	265	47	24	14	7	5	5	6	3
2006	95	273	49	25	12	8	6	6	7	
2007	98	281	50	22	14	9	7	7		
2008	100	290	46	24	15	10	8			
2009	103	288	51	27	17	11				
2010	72	321	57	30	19					
2011	80	357	64	33						
2012	89	397	71							
2013	98	441								
2014	110									

Explanatory Variables

Dimension 2 = Development Lag

β_{48m} = Multiplicative parameter that describes development lag 48 months

$Y_{AY,48m} = EXP(\beta_0 + \beta_{AY} + \beta_{48m}) + \varepsilon$

A simple example

			β_{24}	β_{36}	β_{48}	β_{60}	β_{72}	β_{84}	β_{96}	β_{108}	β_{120}
	Accident Year	**12m**	**24m**	**36m**	**48m**	**60m**	**72m**	**84m**	**96m**	**108m**	**120m**
β_0	2005	92	265	47	24	14	7	5	5	6	3
β_{06}	2006	95	273	49	25	12	8	6	6	7	
…	2007	98	281	50	22	14	9	7	7		
	2008	100	290	46	24	15	10	8			
β_{09}	2009	103	288	51	27	17	11				
β_{10}	2010	72	321	57	30	19					
β_{11}	2011	80	357	64	**33**						
…	2012	89	397	71							
β_{13}	2013	98	441								
β_{14}	2014	110									

- Use a multivariate regression analysis to solve for the best-fit parameters
- Resulting parameters can be used in formula to predict any cell

Predicted incremental loss for AY 2011 and DL 48 months:

Y11,48m = EXP(β0 + β11 + β48m) + ε

Y11,48m = EXP(4.358 + 0.225 – 1.177) = 30.1 (vs actual = 33)

A simple example

			β_{24}	β_{36}	β_{48}	β_{60}	β_{72}	β_{84}	β_{96}	β_{108}	β_{120}
	Accident Year	**12m**	**24m**	**36m**	**48m**	**60m**	**72m**	**84m**	**96m**	**108m**	**120m**
	2005	β_0 92	265	47	24	14	7	5	5	6	3
β_{06}	2006	95	273	49	25	12	8	6	6	7	
…	2007	98	281	50	22	14	9	7	7		
	2008	100	290	46	24	15	10	8			
β_{09}	2009	103	288	51	27	17	11				
β_{10}	2010	72	321	57	30	19					
β_{11}	2011	80	357	64	33						
…	2012	89	397	71	??						
β_{13}	2013	98	441								
β_{14}	2014	110									

A simple example

Accident Year Parameter	Value
β_{2005}	n/a
β_{2006}	0.029
β_{2007}	0.056
β_{2008}	0.082
β_{2009}	0.105
β_{2010}	0.124
β_{2011}	0.225
β_{2012}	0.325
β_{2013}	0.424
β_{2014}	0.338

Development Lag Parameter	Value
β_{12m}	n/a
β_{24m}	1.260
β_{36m}	(0.485)
β_{48m}	(1.177)
β_{60m}	(1.704)
β_{72m}	(2.244)
β_{84m}	(2.533)
β_{96m}	(2.612)
β_{108m}	(2.470)
β_{120m}	(3.143)

Base Parameter	Value
β_0	4.358

Example 1:

$Y_{12,36m}$ = EXP($\beta_0 + \beta_{12} + \beta_{36m}$)

= EXP (4.358 + 0.325 - 0.485)

= 67 (vs actual 71)

Example 2:

$Y_{12,48m}$ = EXP($\beta_0 + \beta_{12} + \beta_{48m}$)

= EXP (4.358 + 0.325 – 1.177)

= 33

A simple example

			β_{24}	β_{36}	β_{48}	β_{60}	β_{72}	β_{84}	β_{96}	β_{108}	β_{120}
	Accident Year	**12m**	**24m**	**36m**	**48m**	**60m**	**72m**	**84m**	**96m**	**108m**	**120m**
	2		265	47	24	14	7	5	5	6	3
β_{06}	2		273	49	25	12	8	6	6	7	3
…	2007	98	281	50	22	14	9	7	7	7	4
	2008	100	290	46	24	15	10	8	6	7	4
β_{09}	2009	103	288	51	27	17	11	7	6	7	4
β_{10}	2010	72	321	57	30	19	9	7	6	7	
β_{11}	2011	80	357	64	33	18	10	8	7	8	
…	2012	89	397	71	33	20	11	9	8	9	
β_{13}	2013	98	441	73	37	22	13	9	9	10	5
β_{14}	2014	110	386	67	34	20	12	9	8	9	5

Historical Values

Predicted Unpaid Values

- Reserve triangle can squared up based on previous slide to develop to ultimate
- Each predicted value is calculated using its beta parameters

A simple example

Accident Year	2-D GLM Unpaid	Chain Ladder Unpaid	Difference
Prior	470	470	0
2008	484	484	0
2009	497	497	0
2010	510	510	0
2011	522	522	0
2012	532	532	0
2013	589	589	0
2014	651	651	0
Total	**5,632**	**5,632**	**0**

- When excluding the calendar year dimension, as we did in this example, the results are same as chain ladder method using all year volume weighted average

Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Section Introduction

- Traditional Reserving Development Methods
  - Key Points
  - Challenges
- Reserving with Predictive Modeling
  - Advantages
- Aggregate Reserving Methods
  - Aggregate Incremental Paid Method
  - Calendar Year Method
- Individual Claim Reserving Methods
  - Incremental Paid Method
  - Claim Closure Rate Method
  - Open Claim Method
  - Frequency/Severity Method

Traditional Development Methods

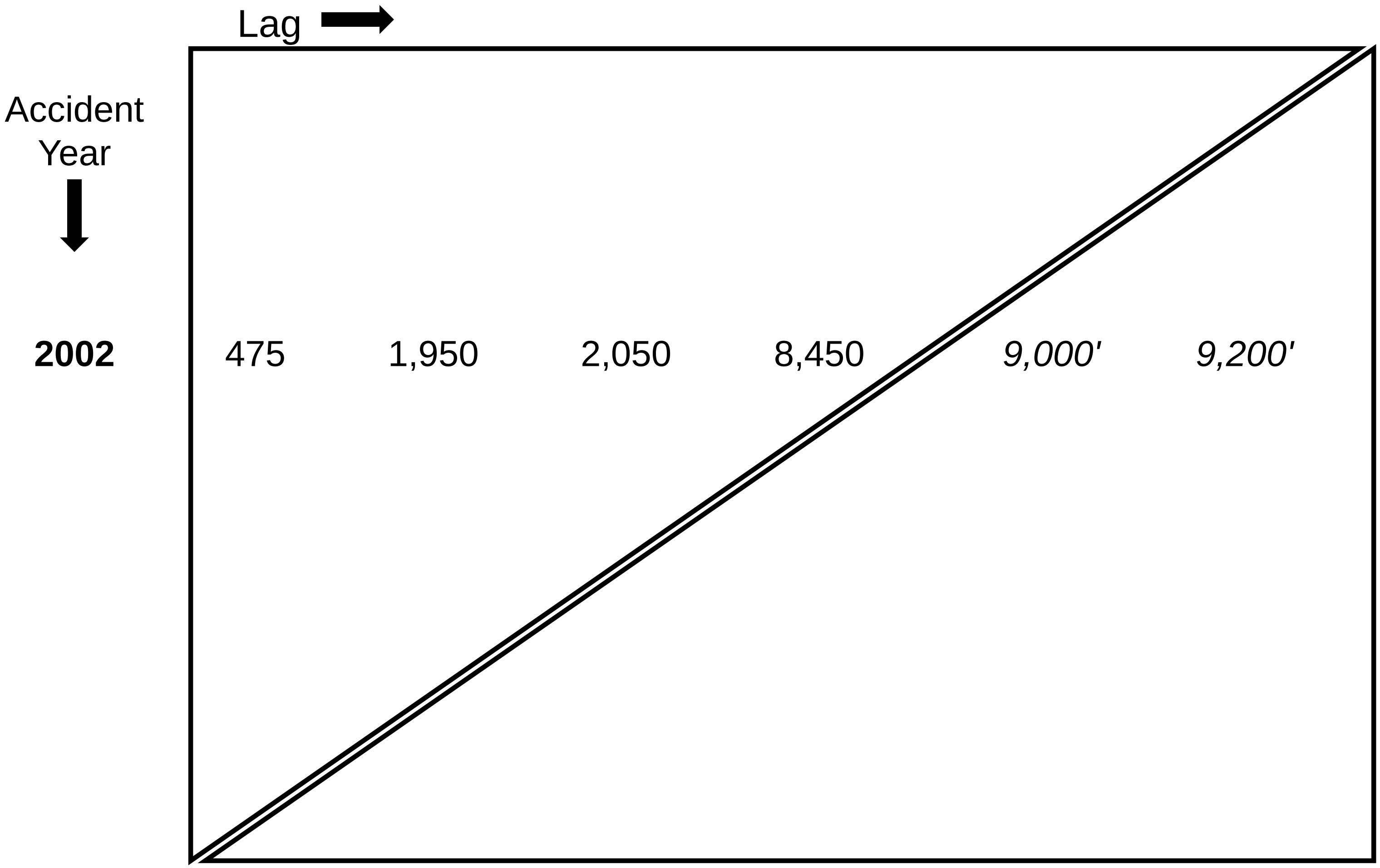
Traditional methods **aggregate** all claims in each cell within the historical triangle on a **cumulative** basis

Accident Year 2002

Claim	12	24	36	48
000001	0	1,000	1,000	5,000
000021	50	50	50	50
000060	0	0	0	250
000124	300	500	500	750
000328	125	400	400	400
000443	0	0	100	2,000

Traditional Loss Development Methods

Goal is to square up the triangle using link ratios



Traditional Development Methods

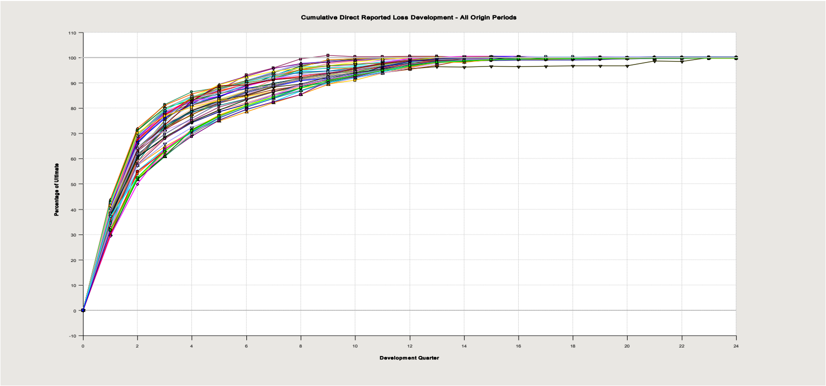
Key Points

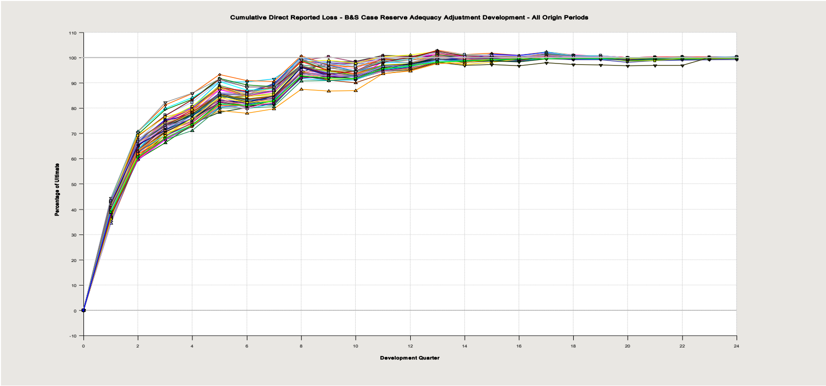
- Aggregated Data
  - Forfeit almost all information unique to each claim
  - Paid, case, reported, open, closed
- Evaluates across only two dimensions: Year and Lag
- Estimates IBNER and pure IBNR together
- Accuracy hinges on consistency
  - Claim closure rate
  - Case reserve adequacy
  - Inflation
  - Reinsurance
- Traditional development methods work quite well when the historical data is consistent, reasonably credible and contains sufficient history

Traditional Development Methods

Challenges

- Challenge is dealing with inconsistency
  - Can consistency/inconsistency be measured?
    - Few cells within triangle make it challenging to measure
    - Small changes are oftentimes masked by random volatility but can impact indications significantly
    - Especially difficult with low frequency/high severity business
  - When measurable, can historical data be adjusted to be consistent?
    - Traditional adjustment approaches tend to produce patterns that are difficult to interpret

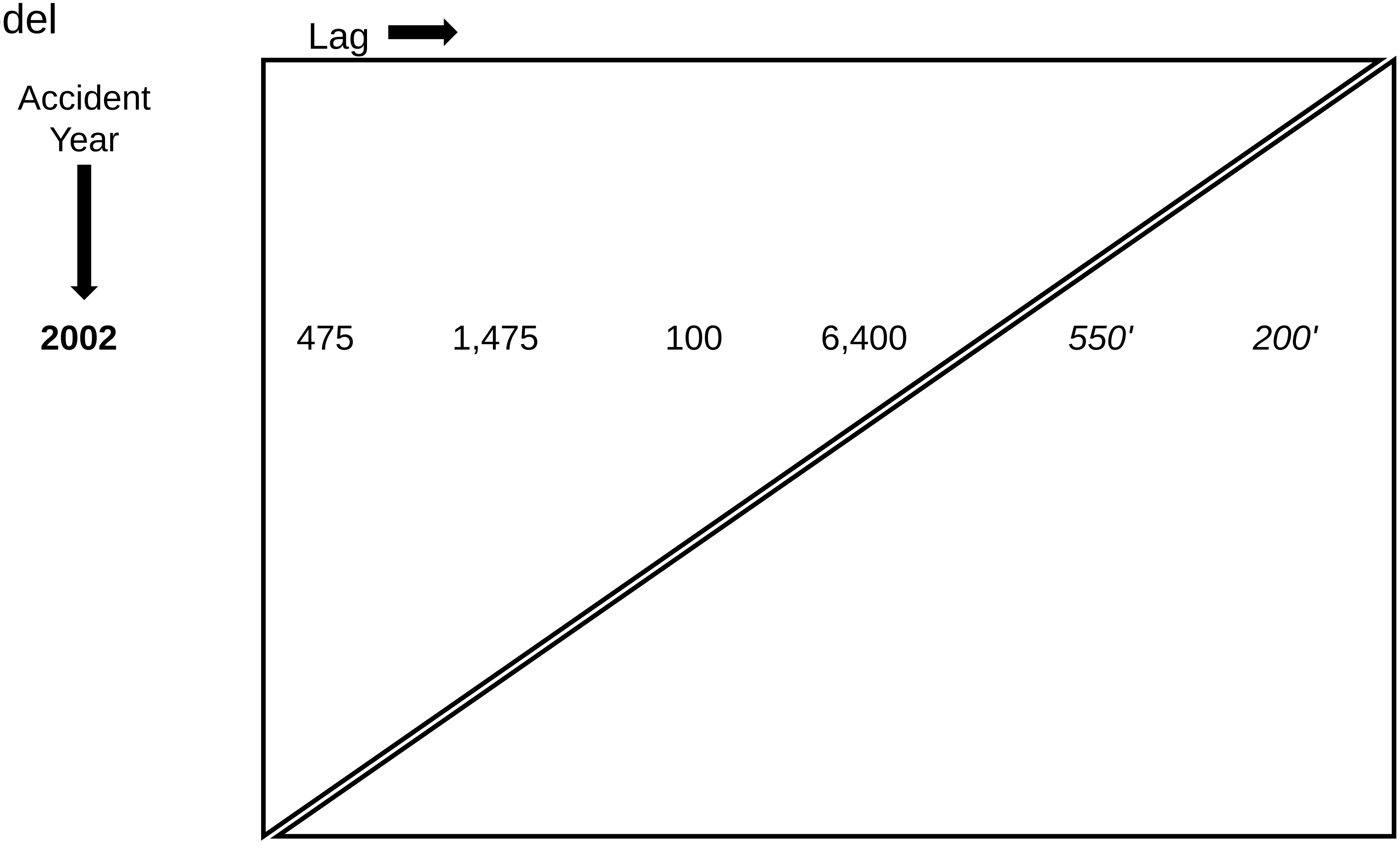




Aggregate Reserving Methods

Aggregate Incremental Paid Method

Goal in GLM is the same: square up the triangle using parameters from the model



Aggregate Incremental Paid Method — GLM Structure

Response Variable = Systematic Component + Random Component

y = h (Linear Combination of Rating Factors) + Error

Incremental Paid Loss = Log Link Function (Accident Year + Lag) + Poisson Error

Aggregate Incremental Paid Method

Key Points

- Aggregated Data
  - Forfeit almost all information unique to each claim
  - Paid, case, reported, open, closed
- Evaluates across only two dimensions: Year and Lag
- Estimates IBNER and pure IBNR together
- Accuracy hinges on consistency
  - Claim closure rate
  - Case reserve adequacy
  - Inflation
  - Reinsurance
- Replicates a traditional paid loss development method using volume weighted average link ratios

Calendar Year Method – GLM Structure

Response Variable = Systematic Component + Random Component

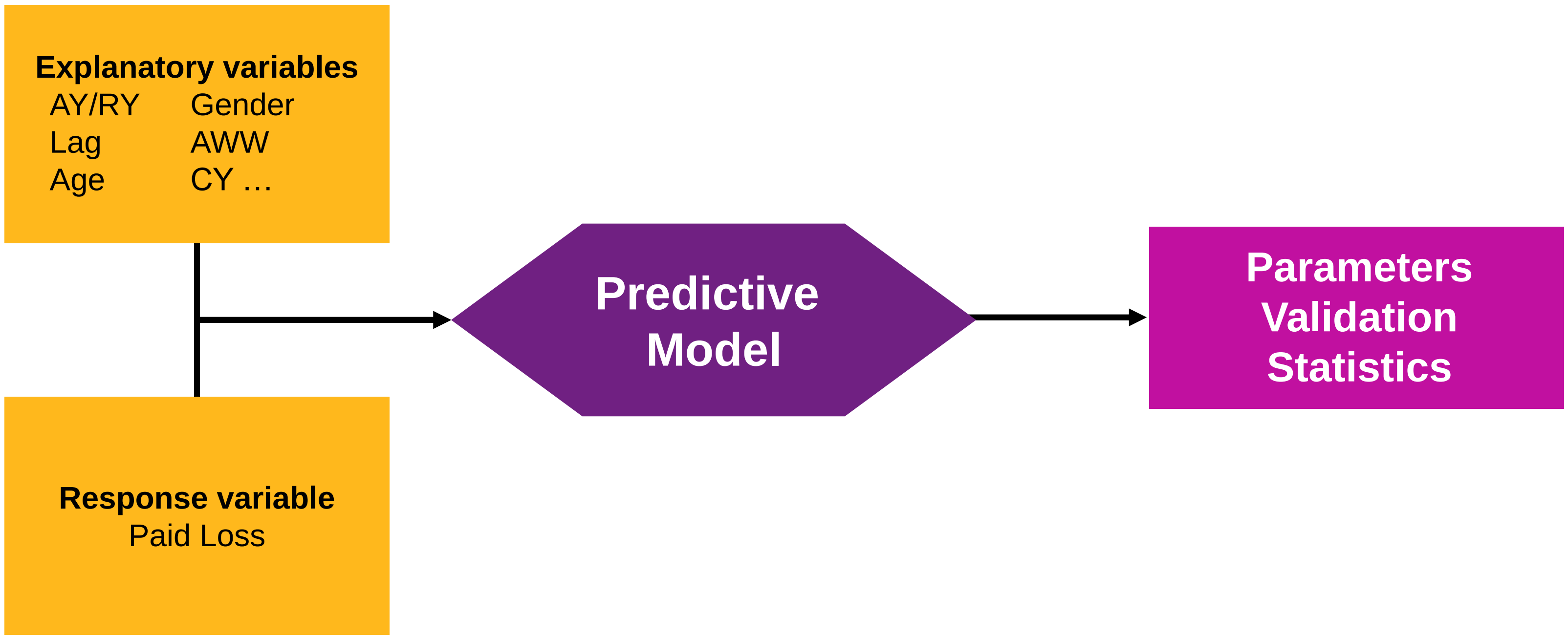
y = h (Linear Combination of Rating Factors) + Error

Incremental Paid Loss = Log Link Function (Accident Year + Lag + Calendar Year) + Poisson Error

Individual Claim Reserving Methods

Individual Claim Reserving Methods

- Now that the data is configured by claim instead of in aggregate, we can introduce additional explanatory variables that are unique to each claim:



Individual Claim Reserving Methods

WC Data Utilized

Insured Characteristics
- Policy start/ end date
- Months with company
- Number of employees
- Limits/ Deductible
- Class code
- Hazard group
- Prior Experience

Claimant Characteristics
- Age
- Number of dependents
- Previous disability
- Marital status
- Years employed
- Job level
- Performance evaluations
- Weekly Wage

Billing
- Payment history
- Payment plan
- Times delinquent

Loss Characteristics
- Cause of injury
- Body part injured
- Day of week
- Month
- State
- Report lag
- Loss date
- Type of loss

Service Providers
- Medical service provider
- Pharmaceuticals
- Attorney involvement

Geography
- Urban/Rural
- Distance to work

Incremental Paid Method

While previous examples used aggregated data, GLM's also work with individual claim data

Incremental 2002 Claims

Claim	12	24	36	48
000001	0	1,000	0	4,000
000021	50	0	0	0
000060	0	0	0	250
000124	300	200	0	250
000328	125	275	0	0
000443	0	0	100	1,900
2002 Total	475	1,475	100	6,400

Incremental Paid Method

Goal: square up the triangle with respect to each individual claim

Lag →

Report Year	Claim				
2002	000001	0	1,000	0	4,000
2002	000021	50	0	0	0
2002	000060	0	0	0	250
2002	000124	300	200	0	250
2002	000328	125	275	0	0
2002	000443	0	0	100	1,900

Results can still replicate a traditional aggregate development method

Incremental Paid Method – GLM Structure

Response Variable = **Systematic Component** + **Random Component**

y = h (Linear Combination of Rating Factors) + Error

Incremental Paid Loss = **Log Link Function** **(Report Year + Lag + Many More)** + **Poisson Error**

Incremental Paid Method

Key Points

- Aggregate incremental paid method blends the estimation of IBNER and pure IBNR into one single estimate
- Individual Incremental Paid method models individual claim data and as a result focuses solely on forecasting IBNER
  - Pure IBNR must be estimated separately
    - Model to predict the frequency of IBNR claims
    - Model to predict the severity of IBNR claims
- Individual claim characteristics used as explanatory variables must be static or known throughout the forecasted periods
  - Med-only/Lost-time
  - Open/Closed

Claim Closure Rate Method

- Models closed claim data and expands on the Calendar Year method by adding a fourth dimension:
  - Year
  - Lag
  - Calendar Year
  - Claim Closure Rate
- Discussed in a paper by Greg Taylor and Grianne McGuire
- Advantages
  - Ideal for high frequency / low severity business where minor changes in claim closure rate affect aggregate methods
  - Estimates total IBNR
- Challenge
  - Method for forecasting future closed claims restricts ability to incorporate unique claim characteristics

Open Claim Method

- Open Claim method builds a series of models that takes advantage of all information known about the claims, including:
  - Calendar year – builds upon previous method
  - Latest paid/incurred to date
  - Individual claim characteristics
- Models reserves for each open claim
- Advantage
  - Claim information is not limited to being static or known
- Challenge
  - Multiple models need to be built
  - Credibility concerns can occur in the tail

Frequency / Severity Method

- Aggregate ultimate severity by year estimated through traditional approaches
- Robust severity model is built using all available claim information and latest known information
  - Development is normalized across data
- Ultimate Severity x Severity Model applied to known and IBNR claims individually to produce ultimate

- Advantages
  - Ideal for low frequency / high severity business where aggregate loss development methods are volatile

Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Taylor & McGuire Method

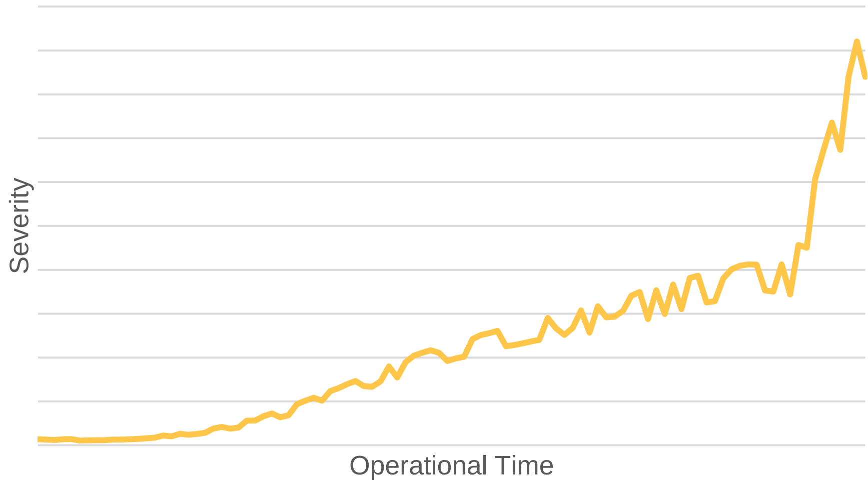
Key Points

- Based on the paper "Loss Reserving with GLMs: A Case Study" written by Greg Taylor and Gráinne McGuire
- A frequency & severity model where:
  - Frequency is estimated from traditional aggregate claim count development methods.
    - Uses "with payment" claim counts
  - Severity is a GLM model using "Operational Time"
- Models individual closed claim data:
  - Accident Period
  - Calendar Period
  - Claim Closure Rate = Operational Time

Taylor & McGuire Method

Operational Time

- Calculating operational time is a multi-step process:
  - Estimate the ultimate closed with payment claim count by accident period
  - Within each accident period, sort closed claims in ascending order of claim closure date
  - Then, the first 1% of ultimate claims to close for a given accident period will have operational time equal to 1; the second 1% will have operational time equal to 2, etc.



Taylor & McGuire Method

GLM Structure

Response Variable = Systematic Component + Random Component

y = h (Linear Combination of Rating Factors) + Error

Cumulative Paid Loss Per Closed Claim = Log Link Function (Accident Period + Operational Time + Calendar Period) + Gamma Error

Taylor & McGuire Method

Advantages

- Ideal for high frequency / low severity business where minor changes in claim closure rate, inflation, and/or claims handling practices affect aggregate methods
- Independent of the loss projection methods
- Estimates IBNER and pure IBNR
- Does not consider case reserves
- Does not consider payments on open claims

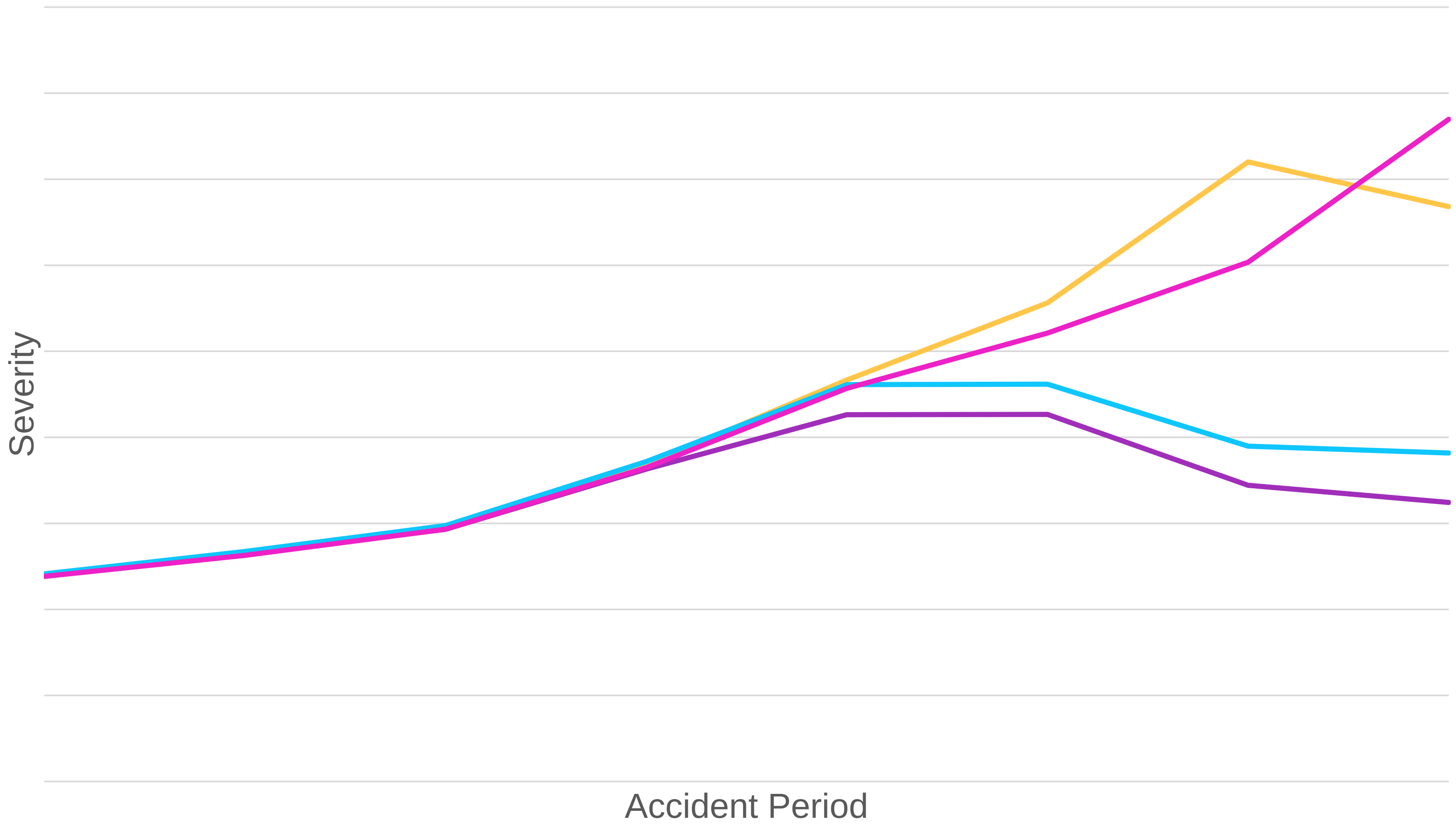
Taylor & McGuire Method

Disadvantages

- Method for forecasting future closed claims restricts ability to incorporate unique claim characteristics
- Results are only valid when aggregated to accident period
- Does not consider case reserves
- Does not consider payments on open claims

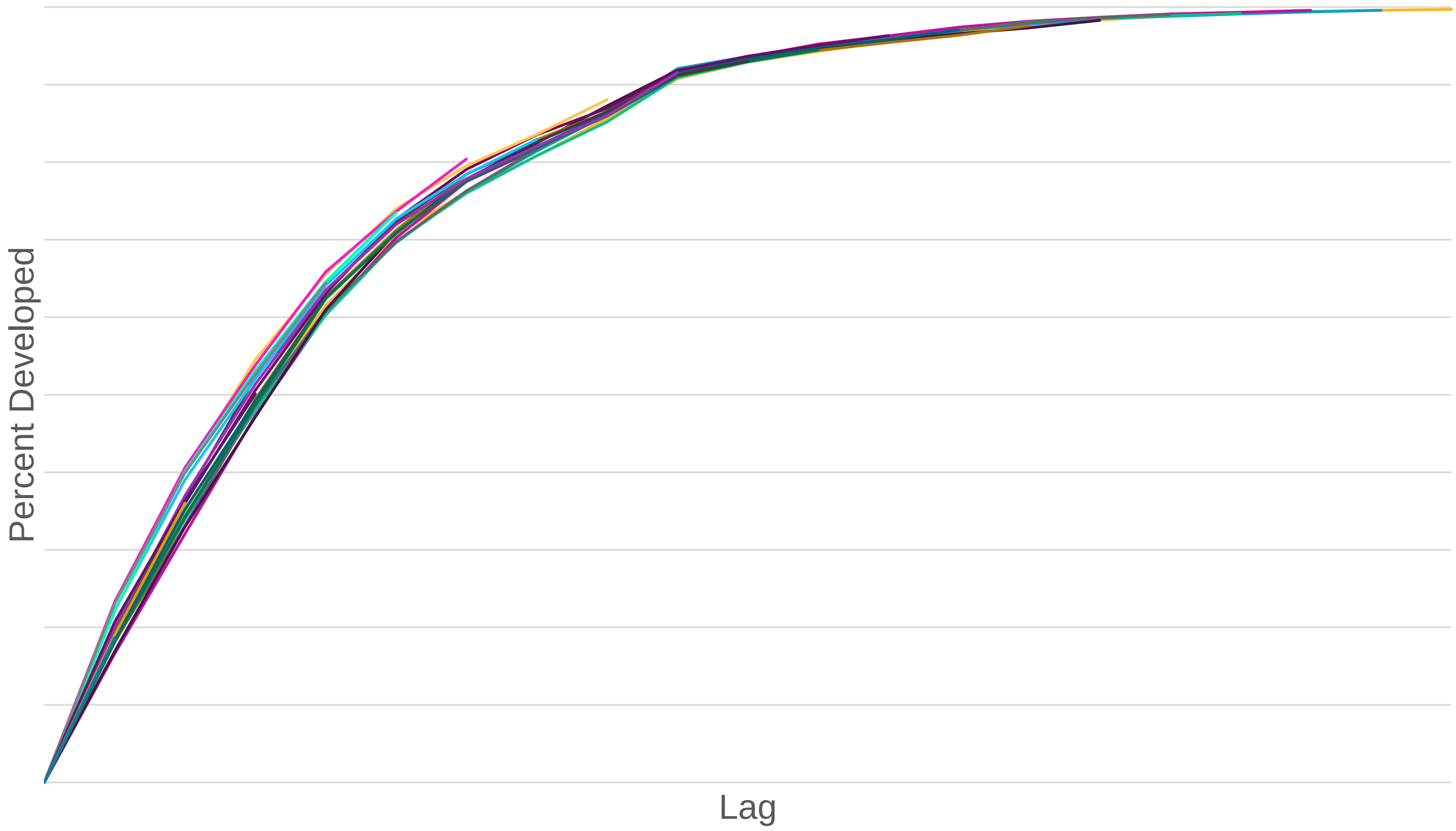
Applications of the Taylor & McGuire Method

Traditional Methods



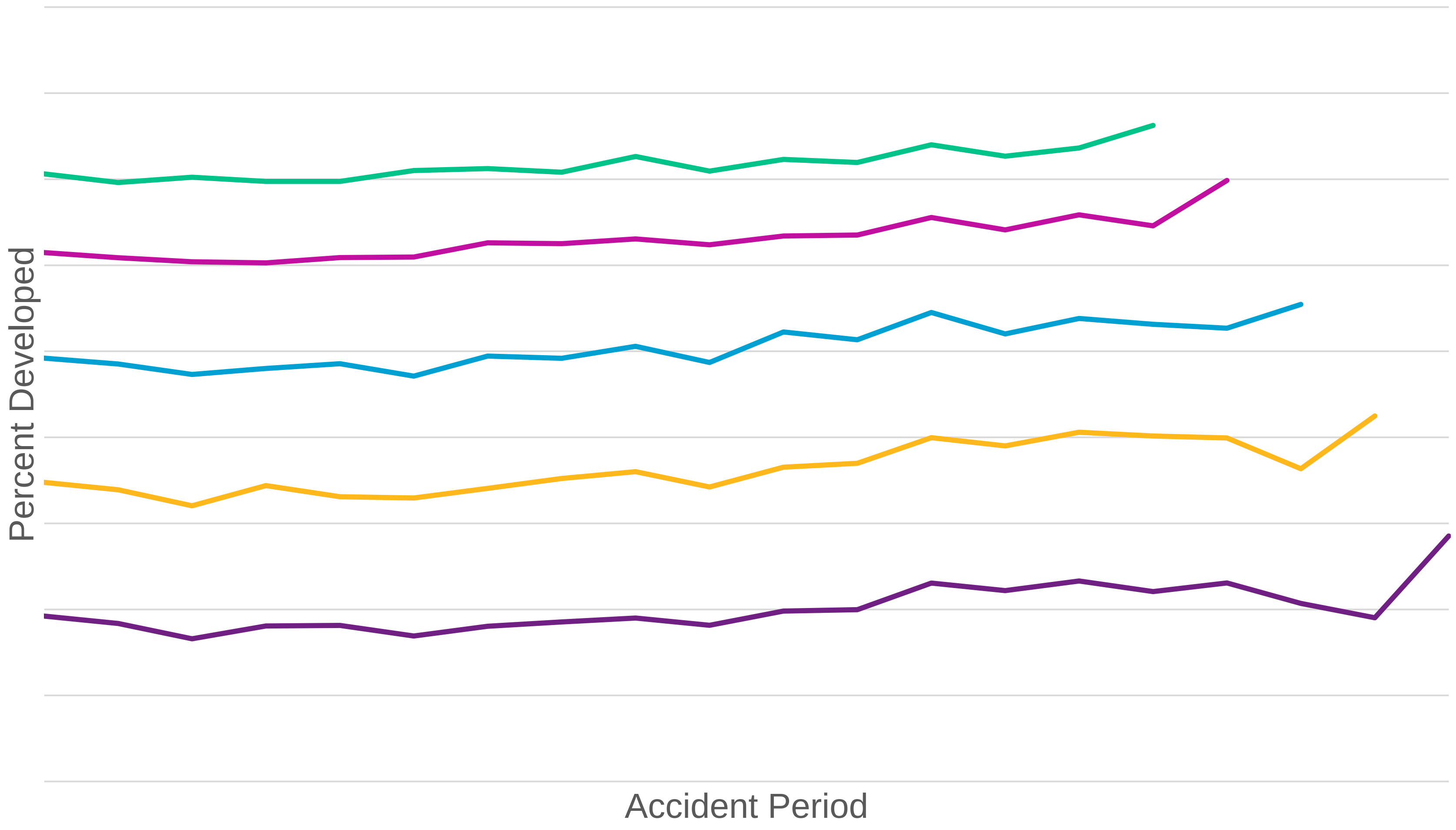
Applications of the Taylor & McGuire Method

Closure Rate Diagnostics



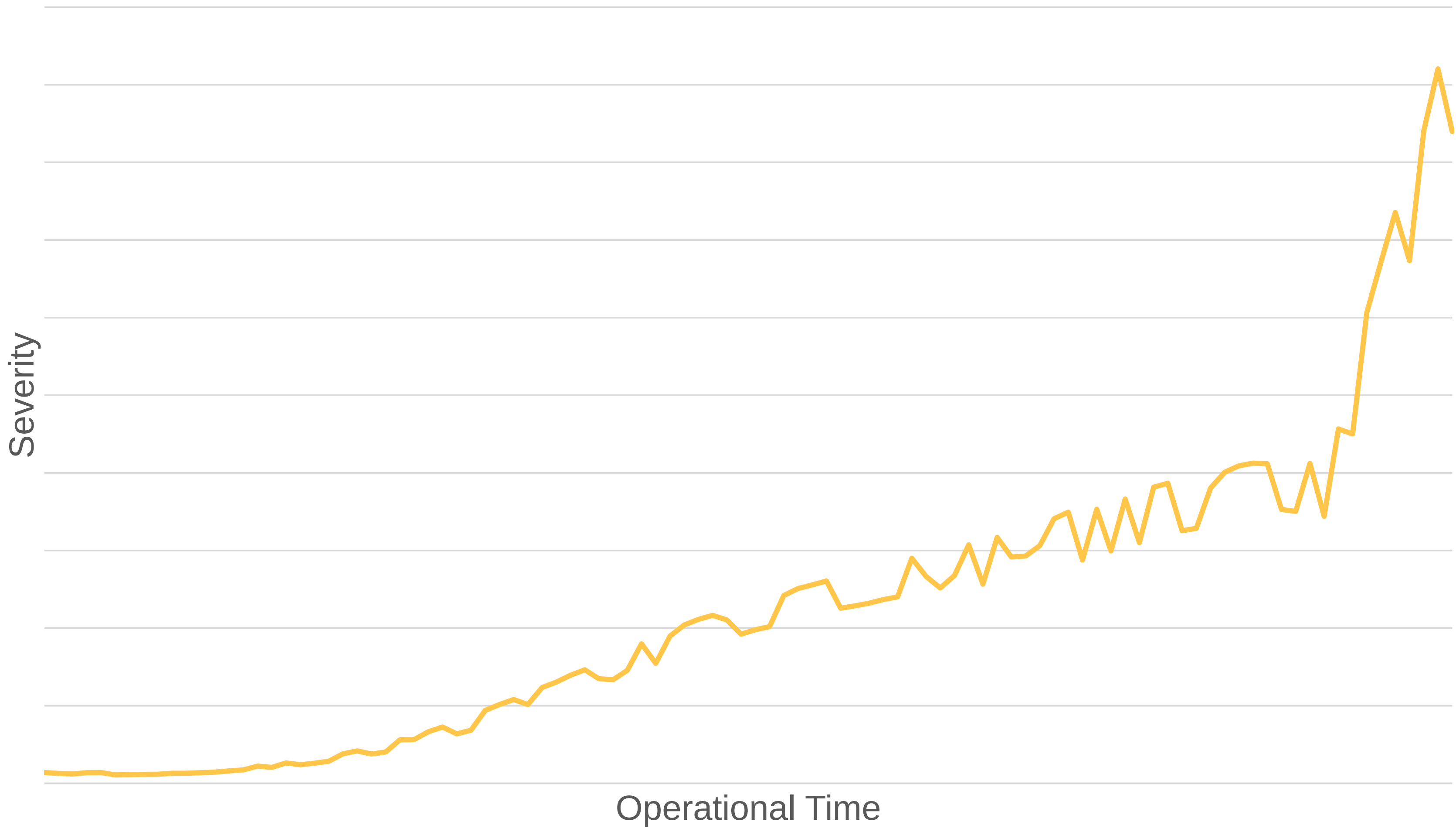
Applications of the Taylor & McGuire Method

Closure Rate Diagnostics



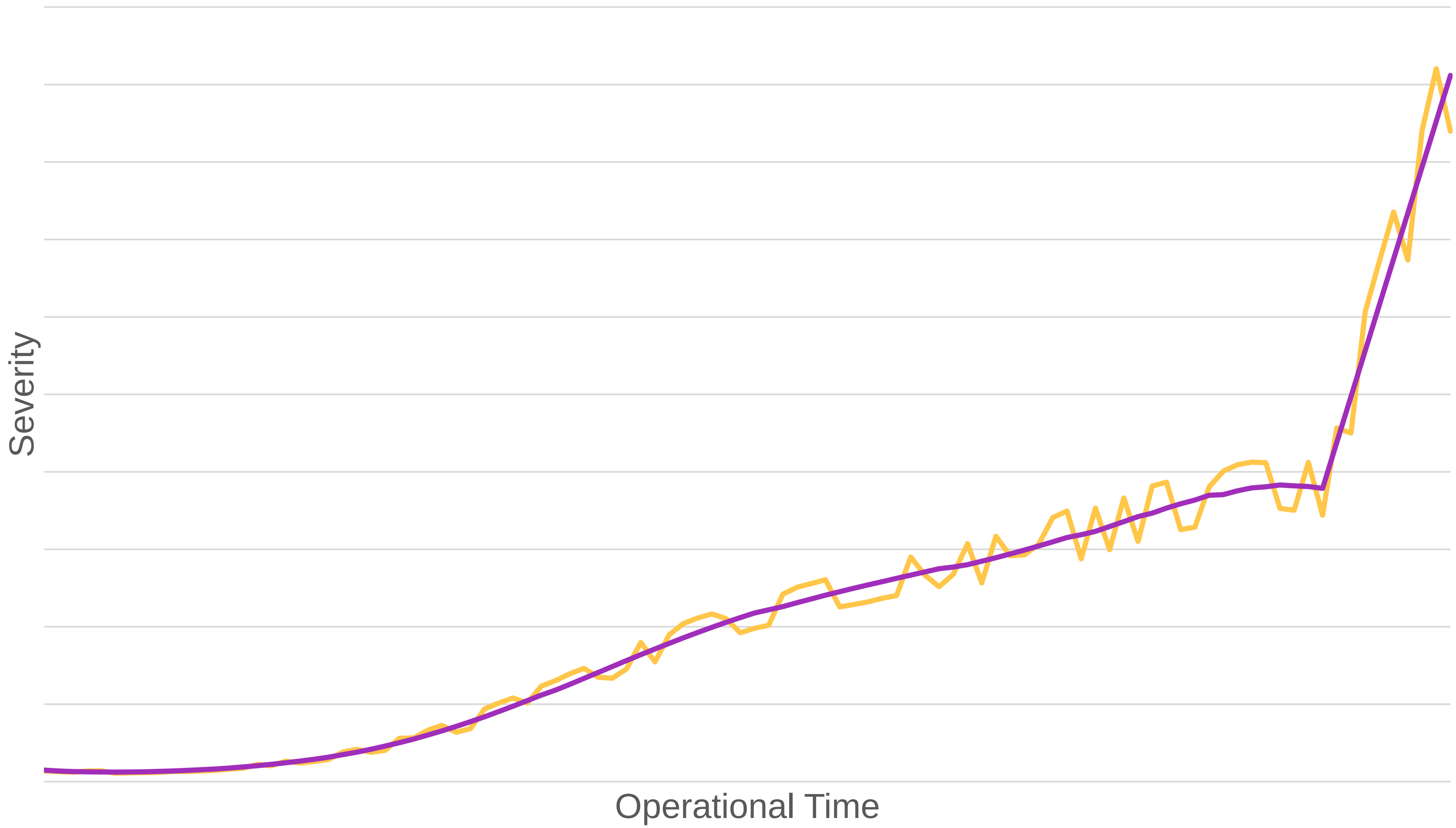
Applications of the Taylor & McGuire Method

Taylor & McGuire



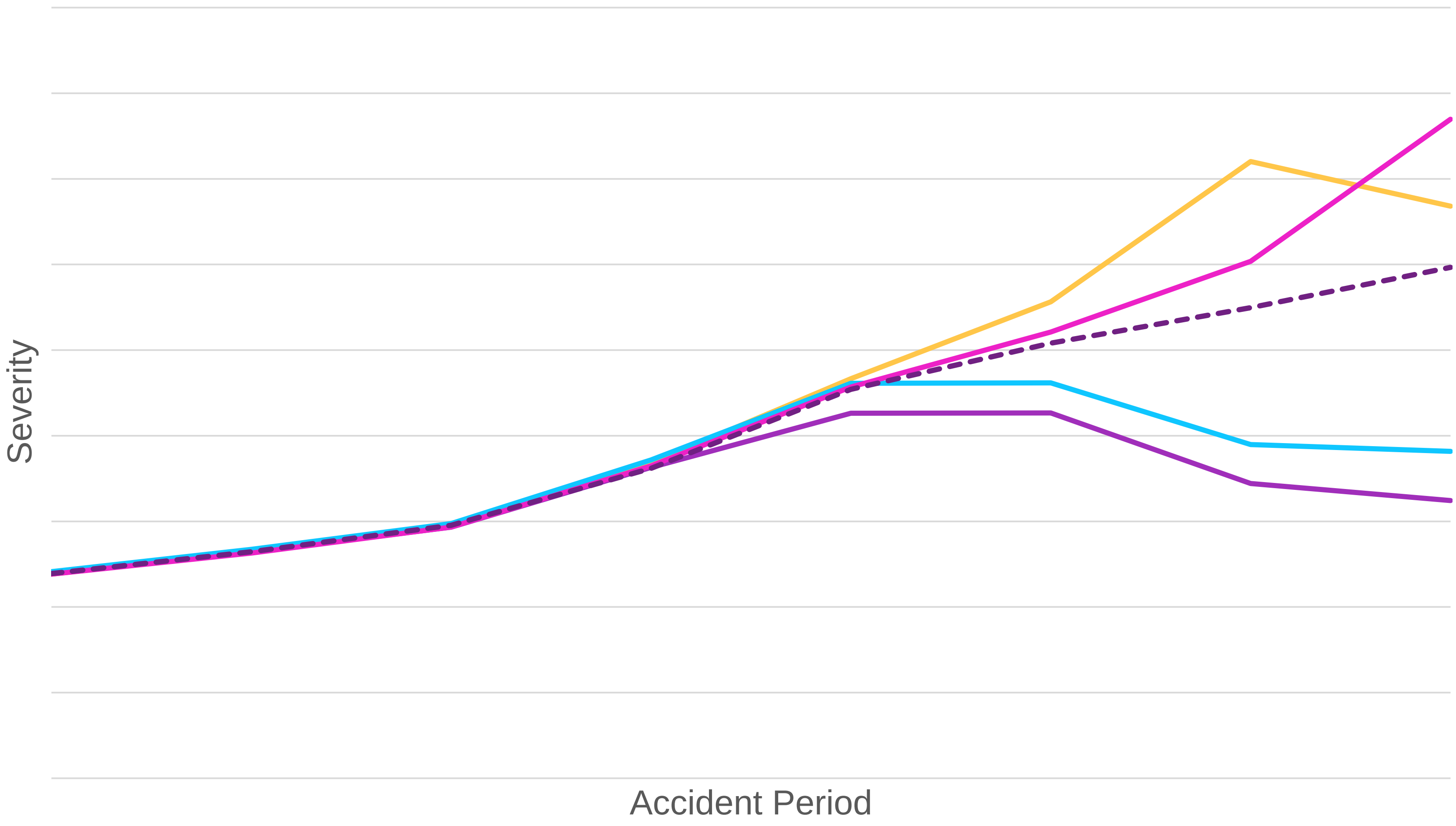
Applications of the Taylor & McGuire Method

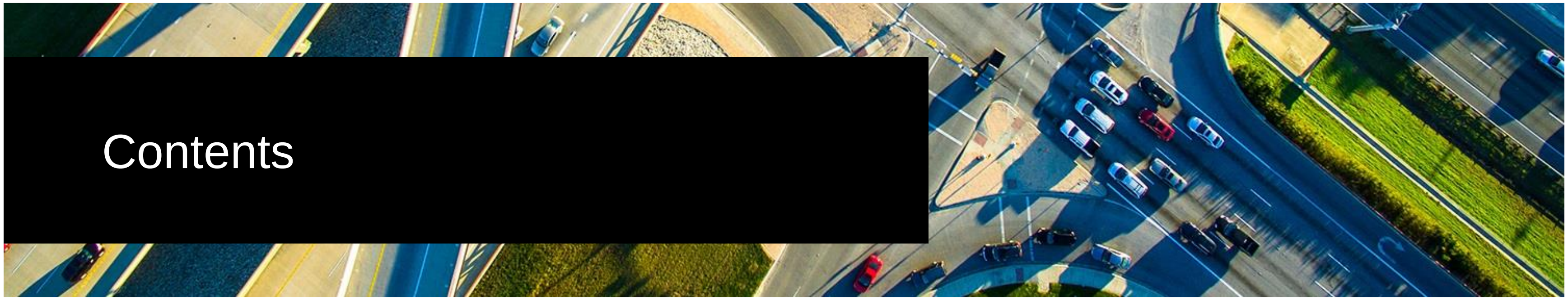
Taylor & McGuire



Applications of the Taylor & McGuire Method

Traditional Methods with Taylor & McGuire





Contents

Manolis Bardis, FCAS, MAAA, PhD, MBA
Director
Insurance Consulting and Technology

T +1 617 638 3807
E Manolis.Bardis@willistowerswatson.com

Paul Bailey
Manager
Insurance Consulting and Technology

T +1 858 678 2160
E Paul.Bailey@willistowerswatson.com

Conclusions

- There is appetite to use predictive models in reserving to address inconsistencies in (aggregated) data and to provide additional insights into cost drivers
- Structuring data for modeling individual claims requires careful planning including – cause of loss coding, claim-level predictors at points in time and opportunities for additional data enrichment
- Model forms include statistical and machine learning, and often one model improves (rather than replaces) another
- Applications include reserving analyses validation, case reserve estimation, large loss prediction in UW, reinsurance, economic capital models and claims triage
- Domain experts must weigh predictive power with critical deployment considerations

Willis Towers Watson