



Institute  
and Faculty  
of Actuaries

# IFoA GIRO Conference 2024

18–20 November, ICC, Birmingham



Institute  
and Faculty  
of Actuaries

# Transforming Technical Pricing with Limited Data: Secrets to Successful Predictive Modelling

Jake Morris and Jamie Wilson

**IFoA GIRO Conference 2024**

# Introductions



Jamie Wilson

Head of Pricing and  
Innovation

 hyperexponential



Jake Morris

Head of Pricing  
Analytics

**Allianz** 

# Predictive Modelling Experience



Jake Morris

Head of Pricing  
Analytics



2018 – 2020	Argo: International Head of Predictive Analytics	London Market
2021 – 2023	Allianz: Head of Large Corp Predictive Modelling	Global Large Commercial and Specialty
2023 – 2024	Allianz: Head of Pricing Analytics	Global Commercial

# Predictive Modelling Experience



Jamie Wilson

Head of Pricing and  
Innovation

↑ hyperexponential

2016 – 2018

Allianz: Head of Pricing Analytics US

MidMarket US

2019 – 2023

Allianz: Head of Predictive Analytics

Global Large  
Commercial and  
Specialty

2023 – 2024

hx: Head of Pricing and Innovation

Global  
Commercial

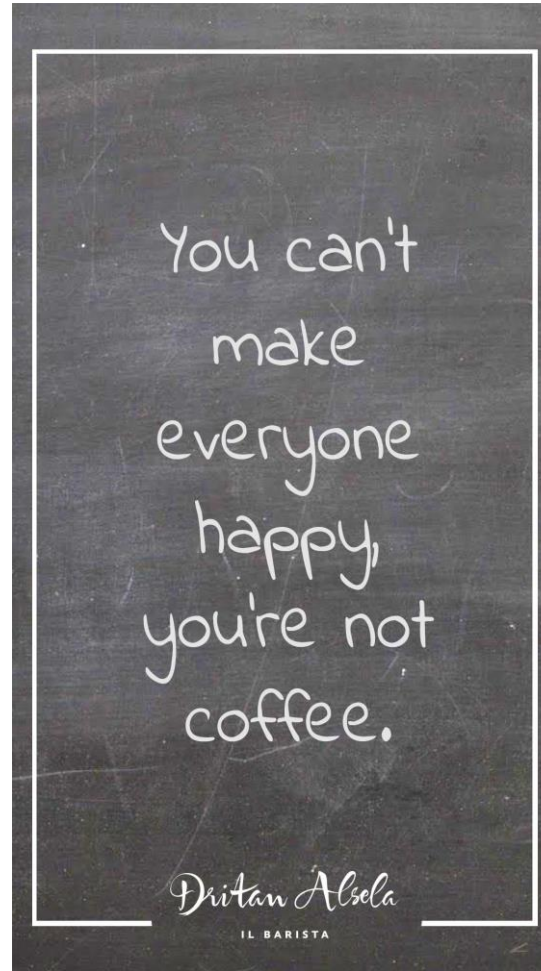
# Over to you

QR Code

What's your experience with predictive modelling to date?

1. *I've never calibrated a pricing model in my life*
2. *I've used 'one analysis' to calibrate models before*
3. *I've used Actuarial modelling tools like Akur8 and Emblem to perform multivariate regression to calibrate a model before*
4. *I'm comfortable building out basic predictive models in R or Python*
5. *Penalised regression? Bayesian methods? Neural nets? I love them all*

# Technical Disclaimer



# Agenda

- 
- 01     **Setting the Scene**

---

  - 02     **Why GLMs Fail(ed)**

---

  - 03     **Making GLMs Great Again**

---

  - 04     **It's Not That Hard...**

---

  - 05     **Getting Started**

---

  - 06     **Change Management**

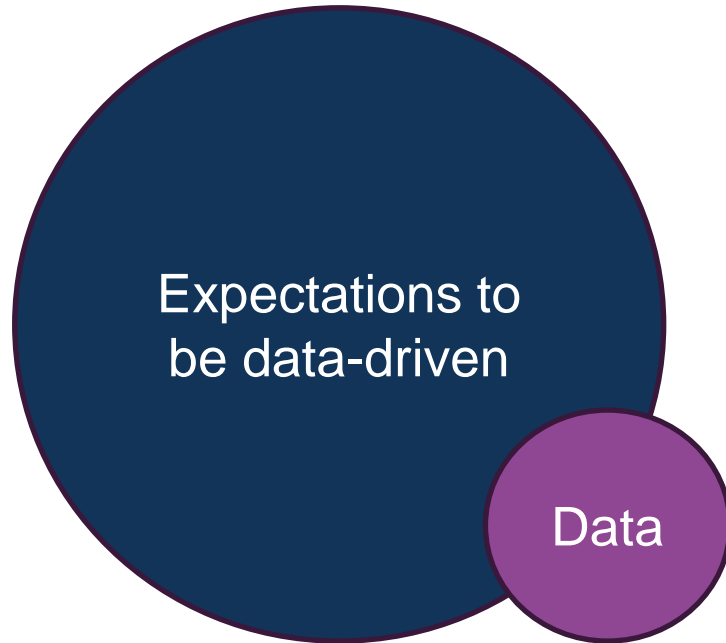
---



# Setting the Scene

What's the current state-of-play for predictive modelling in Large Commercial and Specialty Markets?

# Setting the Scene



- We typically don't have large quantities of data (exposure, loss) in Large Commercial & Specialty lines
- 'Advanced' data-driven techniques (GLMs) don't seem to work
- ...yet we anecdotally hear of some Specialty insurers doing this type of analysis – how?

**A big gap between historic expectation and reality...**

# How did we get here?



2000s



2010s



Today

# How did we get here?

GLM usage takes  
off in Personal  
Lines



2000s



2010s



Today

# How did we get here?

GLM usage takes  
off in Personal  
Lines



2000s

London  
Market Tries  
Adoption



2010s



Today

# How did we get here?

GLM usage takes  
off in Personal  
Lines



2000s

London  
Market Tries  
Adoption



2010s

London  
Market left  
Behind?



Today



# Why GLMs Fail(ed)

What went so wrong with initial applications of GLMs in the London Market?

# Why Don't GLMs Work Out Of The Box?

Changing Environment

Inappropriate Distributions

Dependencies

Multicollinearity

Sparse / Imbalanced Data



# Why Don't GLMs Work Out Of The Box?

Changing Environment

Inappropriate Distributions

Dependencies

Multicollinearity

Sparse / Imbalanced Data

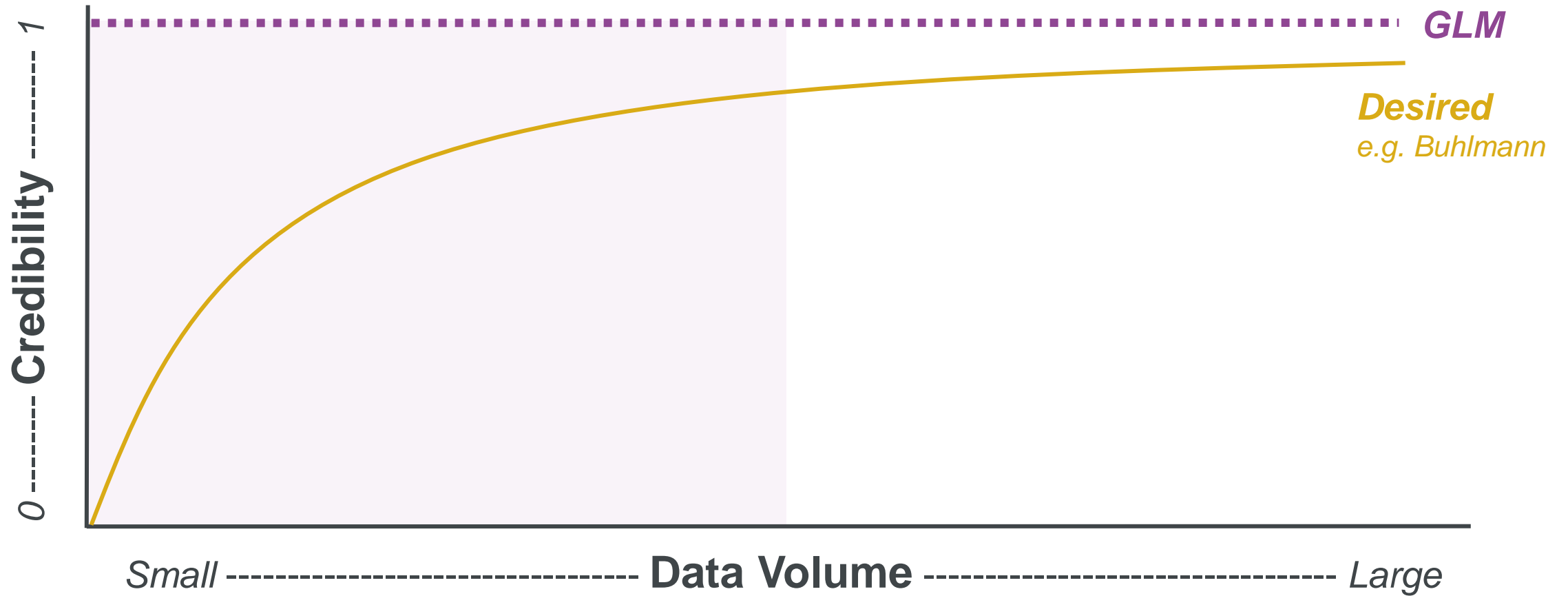
---

All are important, but most prevalent challenge is **data volumes\***

---

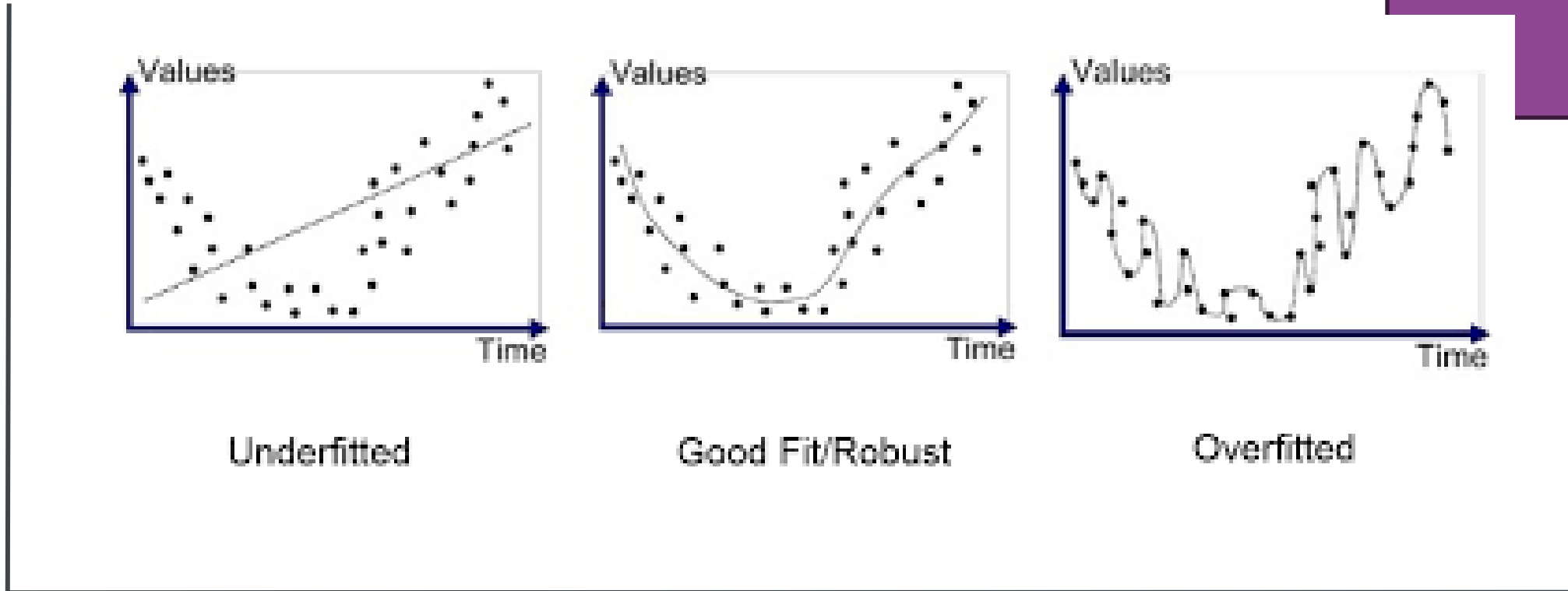
\*and data quality (largely out of scope for today)

# Why Don't GLMs Work Out Of The Box?



# Why Don't GLMs Work Out Of The Box?

Overfitting example



Underfitted

Good Fit/Robust

Overfitted

# Making GLMs Great Again

How can the standard GLM method be modified to make it applicable in data-limited modelling areas?

# Making GLMs Great Again



# Making GLMs Great Again

- **Group** variables (+ simplify continuous variables)
- Existing **Factors as Predictors**
- **Offset** rating components
- **Credibility weight** parameters post-hoc



Traditional  
GLM  
Workarounds

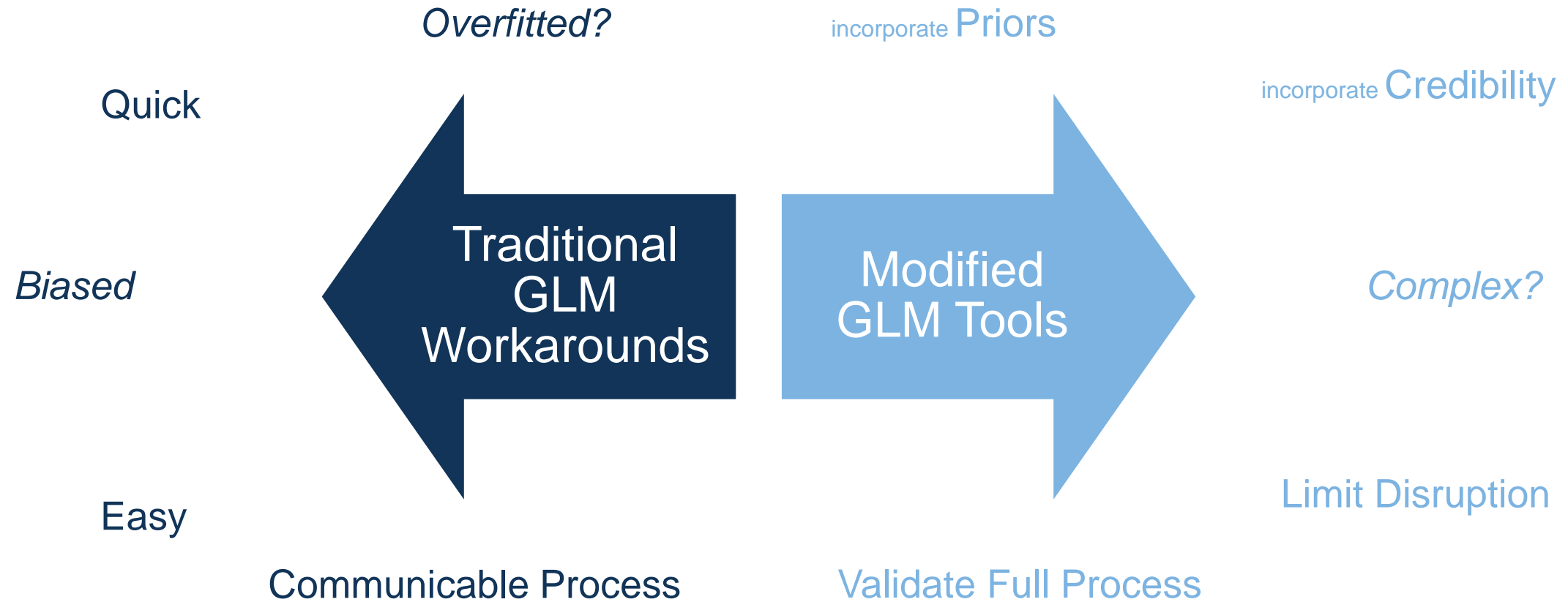
# Making GLMs Great Again



Modified  
GLM Tools

- **Penalised & Bayesian** GLMs
- Incorporate parameter **Priors**
- Consider parameter **Hierarchies**
- **Tune** credibility, e.g. using cross-validation

# Making GLMs Great Again





# Making GLMs Great Again: Technical 2 Minutes

Estimated  
Parameter

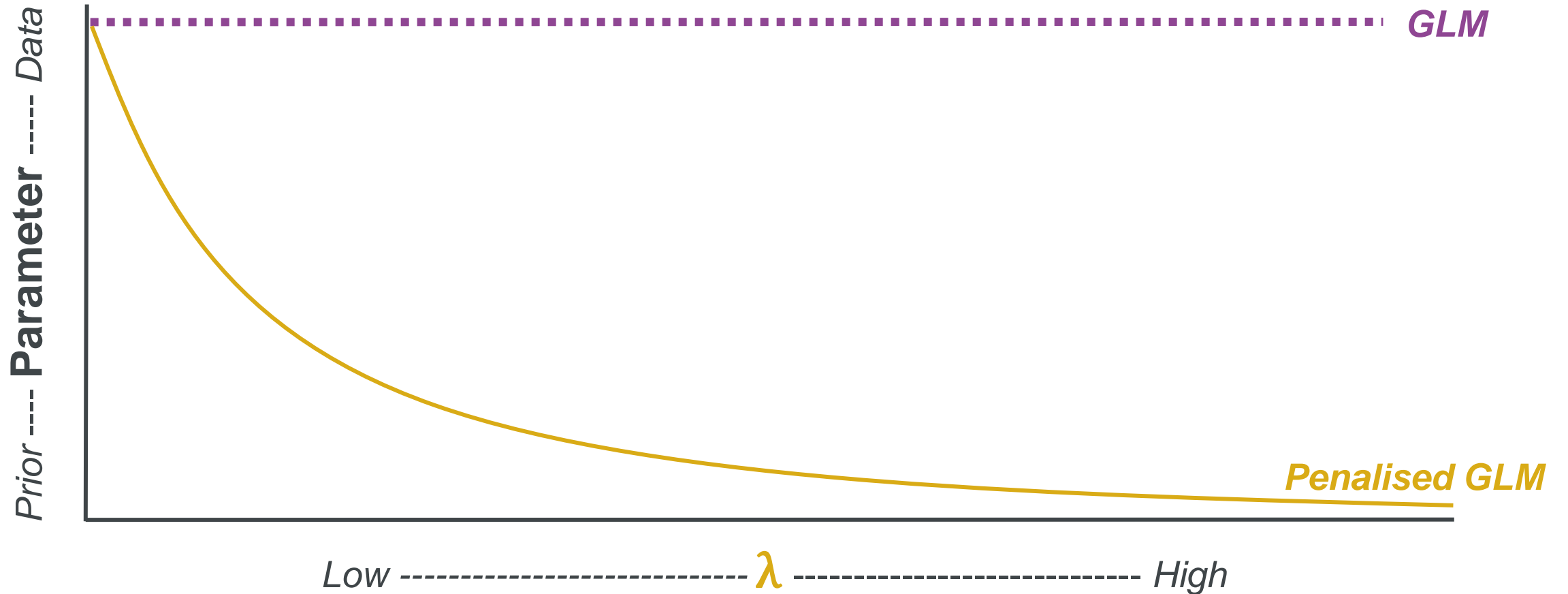
Regularization parameter =  
strength of penalty

$$\hat{\beta} = \mathit{ArgMin}_{\beta} \left[ \underbrace{-\log \mathcal{L}(\beta)}_{\text{Log-Likelihood Function based on data}} + \underbrace{\lambda R(\beta - \beta_{\text{prior}})}_{\text{Penalisation/Regularization function, e.g. Ridge/Lasso/Elastic Net}} \right]$$

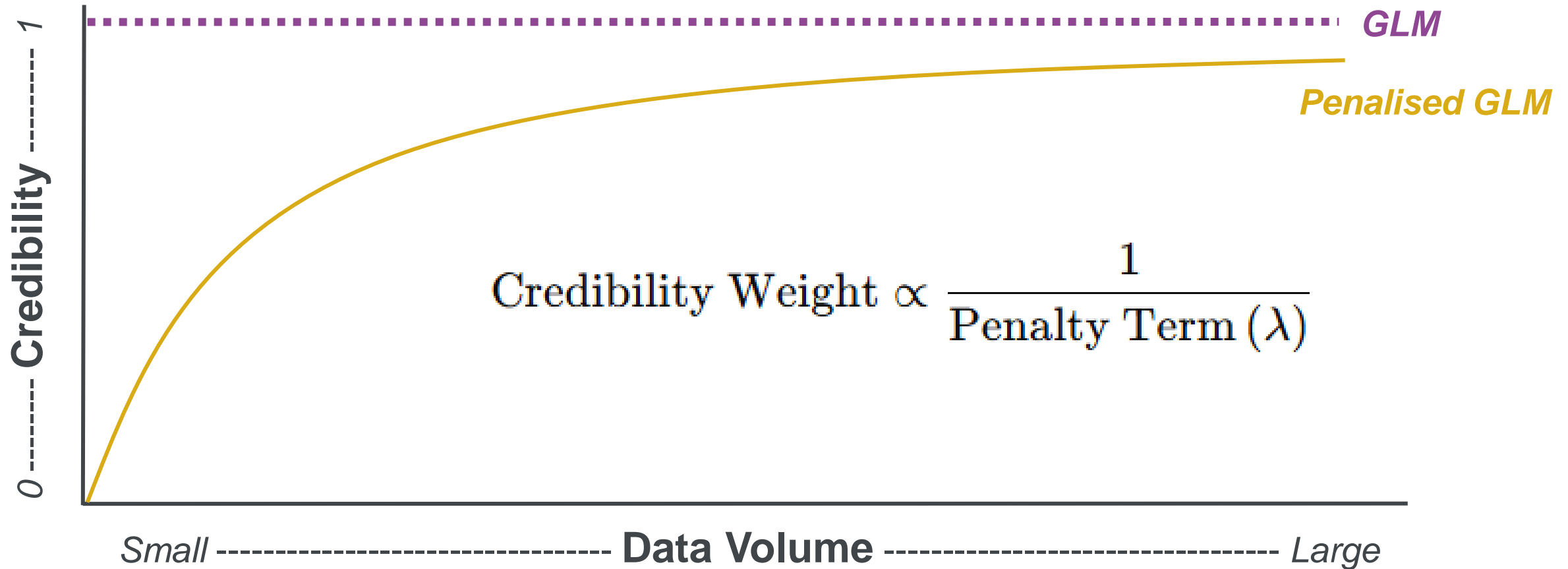
Log-Likelihood Function  
based on data

Penalisation/Regularization function,  
e.g. Ridge/Lasso/Elastic Net

# Making GLMs Great Again: Technical 2 Minutes



# Making GLMs Great Again: Technical 2 Minutes



# Making GLMs Great Again: Technical 2 Minutes

Relativities Plot with Low / Med / High regularisation

**It's not that  
hard...**

You'll need to work at it, but there's never been a better time to get into modelling...

# It's not that hard – Ask ChatGPT to get started!



Flowchart:

- Prep data
- Visualise data
- Code model
- Validate model

# It's not that hard – Ask ChatGPT to get started!

- Prep data Prompts: xxx, yyy, x
- Visualise data Prompts: xxx, yyy, x
- Code model Prompts: xxx, yyy, x
- Validate model Prompts: xxx, yyy, x

# It's not that hard – Ask ChatGPT to get started!

Example Prompt output image





# Debunking the Difficulty: General Recommendations

1. Get **more data** if you can, e.g. from external sources
2. Refresh insurance **GLM fundamentals**: Frequency vs. Severity vs. Pure Premium, distributions, link functions, statistical significance, transforms, visualisation, control variables, capping, weights, offsets, validation...
3. Be aware of the **risks of overfitting**
4. Understand your **complement of credibility**: existing rating plan, premiums, UW judgement, ...
5. **Talk** to the business

# Getting Started

You're excited to get going, so how to get started with this all?

# Skillsets and Resources

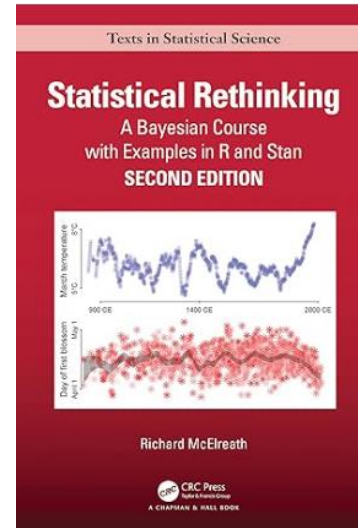
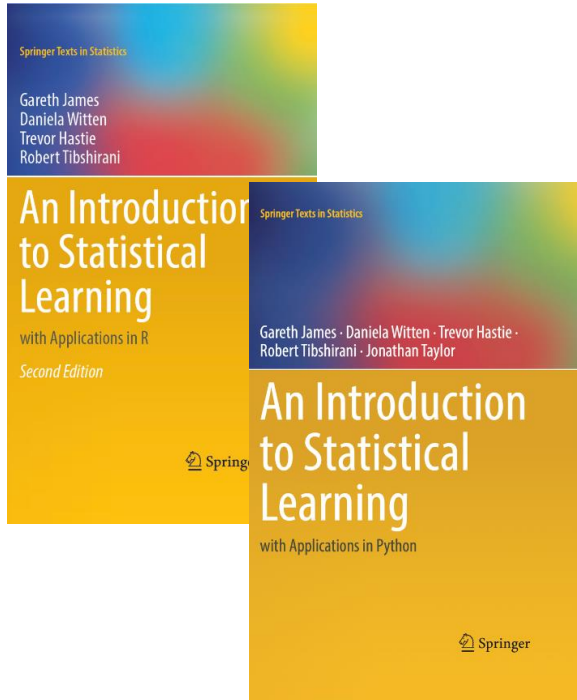
## Placeholder slide:

- Best to get started now
- Find a community (Bayesian mixer, kaggle, etc.)
  - See what you can do where you are

## Skills:

- Data wrangler
- Statistical understanding
  - Coding or Tooling

# Resources



New monograph



\*Links to all resources can be found in the appendix



# Change Management

And it all comes crumbling down...

# Making an impact

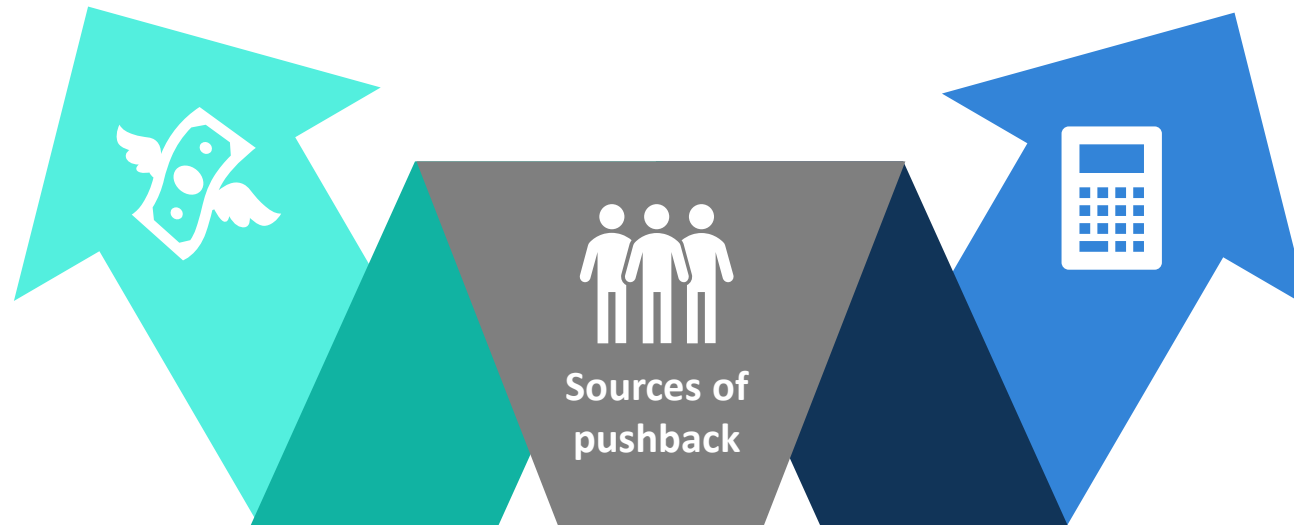
*“If a model makes a prediction in the woods, and no pays attention to it, is it a good model?”*

## *Underwriters*

- *Models don't work in my field*
- *How do I use the model?*

## *Actuaries*

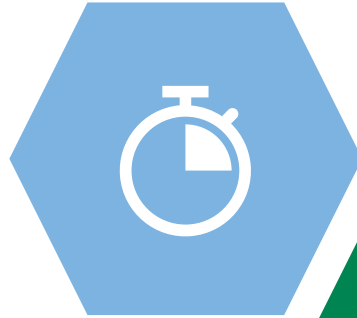
- *Models don't work in my field*
- *The Pricing Tool belongs to me!*



# Building Credibility with Underwriters

## Involve Underwriting early

*Get their input throughout and incorporate it into the model*



## Show them the potential value

*Share exhibits on historic anti selection or profitable growth options*



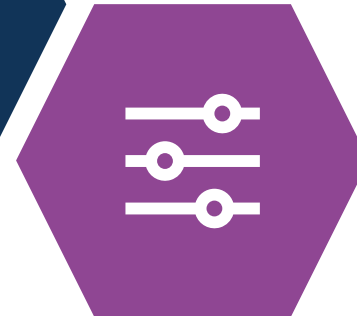
## Show them it works

*Loss ratio bases – avoid statistical terminology*



## Discuss adoption options

*They don't/shouldn't just charge technical price!*

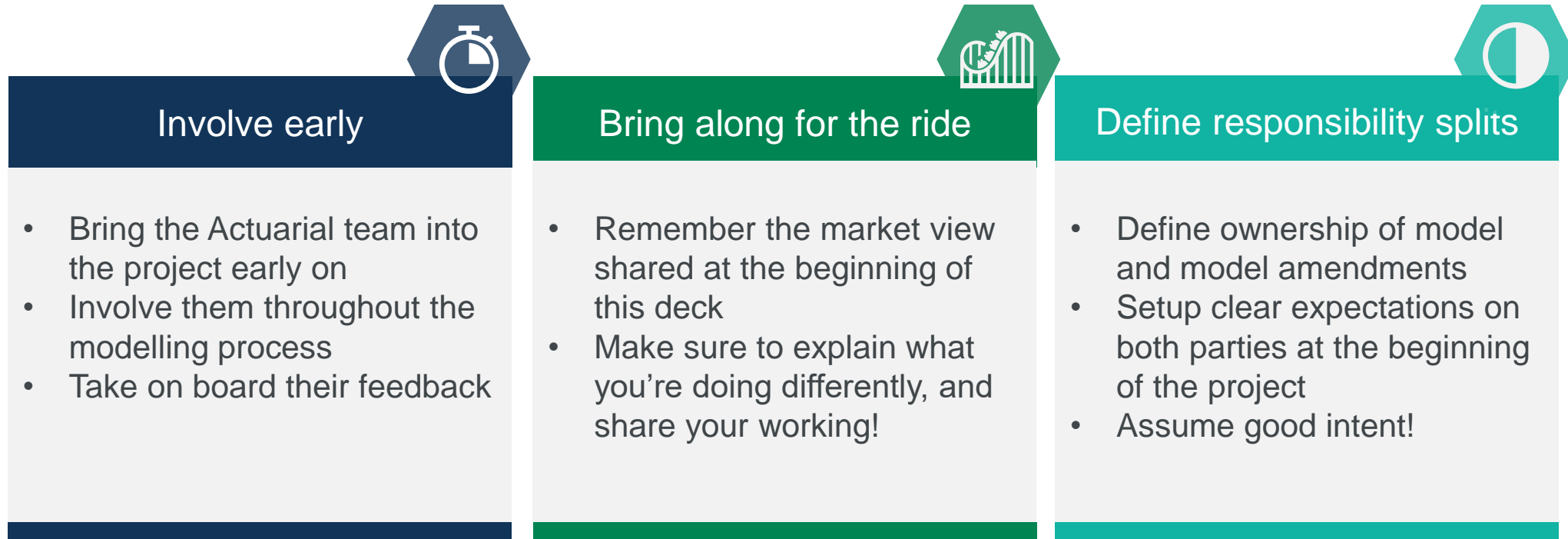


**And remember – your model isn't right!**

# Building Credibility with Actuaries

*(Often harder!)*

**Potential Challenge:** Modelling responsibility split from Actuarial owner of Pricing Tool



**Secondments can be great!**



# Questions

# Comments

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenter.

# Making GLMs Great Again: Technical 2 Minutes

What about Bayesian methods?

- These also effectively penalise, but instead of point estimates they penalise parameter distributions far from a given prior distribution
  - Can result in similar behaviours to penalised regression – in fact, under certain assumptions, bayesian regression models can produce the same results as penalised regression models (taking a 'MAP' point estimate)
  - We use these in comm/specialty to quan uncertainty, but also to implement custom likelihoods, e.g. to account for layer truncation and censoring for severity models
- 
- Buhlmann credibility can still give too much credibility for smaller sample sizes
  - Out of scope for this talk, but there are extensions (e.g. per-parameter penalties)
- 
- Also --- may prefer nested cross validation to cross validation

# Making GLMs Great Again: Why GLMs At All?

- Other predictive modelling methods are available, e.g. Machine Learning, which can incorporate regularisation/penalisation. However:
  - Complex link between model hyperparameters and amount of 'traditional' credibility
  - Underwriters/markets hold predictive information about complex risks external to available data
  - Simpler algorithms (generally) preferred for communication with brokers / interpretability / clear adjustment points
  - GLMs can (often) capture much of the predictive signal found through ML
- Machine Learning helpful for quantifying predictivity upper bounds, feature selection (where we have a volume of potential factors) and identifying patterns for GLMs
- There will be exceptions!