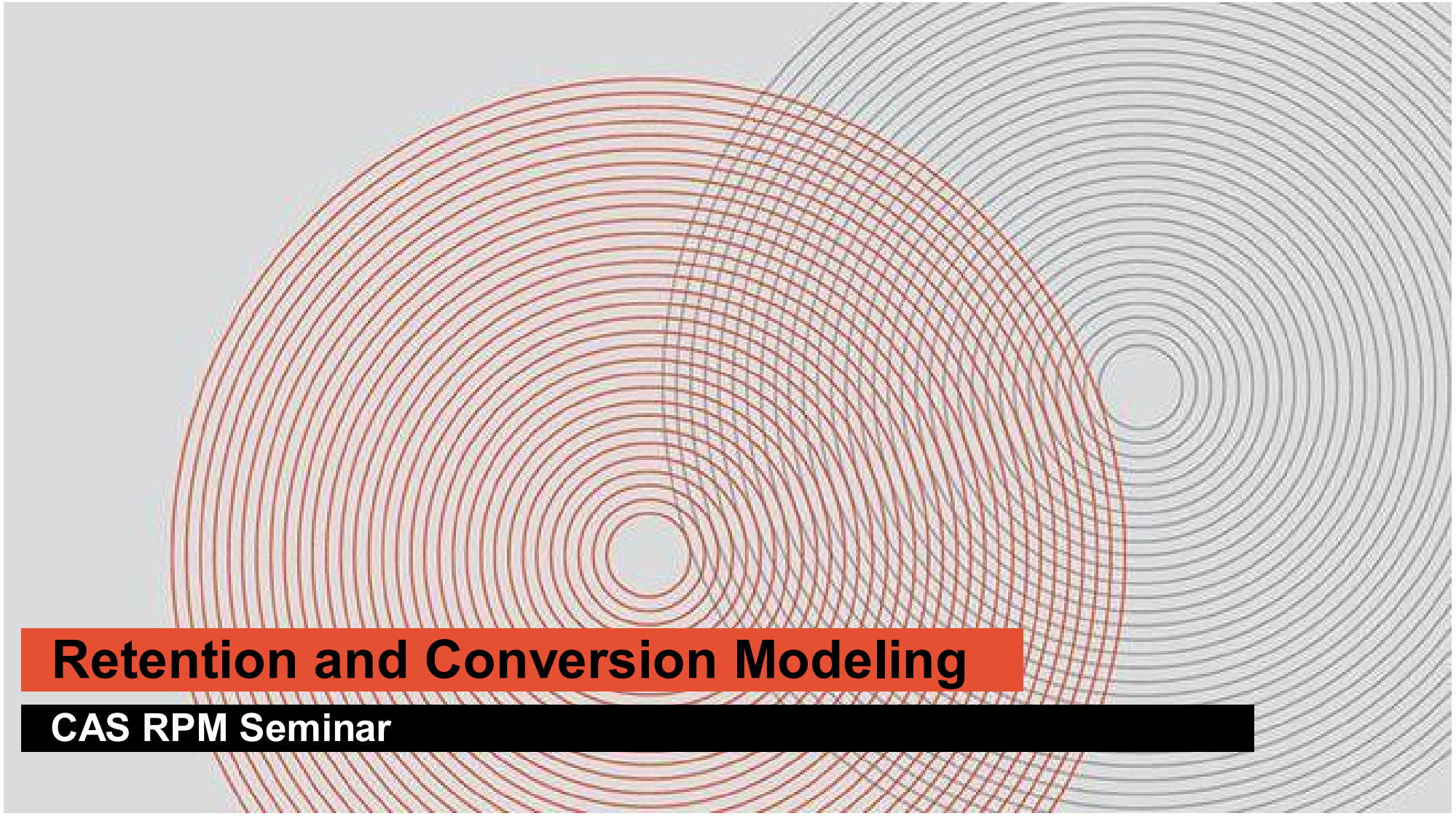


Retention and Conversion Modeling

CAS RPM Seminar

James Tanser

17 March 2010

TOWERS WATSON

Retention and Conversion Modeling

- Introduction
- Drivers
- Measures
- Models

Retention and Conversion Recap

- Want to understand probability that a policy will renew or that a quotation will become a policy
- More widely, understand customer behaviour:
  - Cross selling and Up selling
  - Fraud
- Many uses of these models:
  - What-If? analysis
  - Target marketing spend more accurately
  - Target discounts / price changes
  - Price optimisation
- Good understanding is fundamental to managing your business

Retention and Conversion Modeling

Introduction

Drivers

Measures

Models

What drives customer behaviour?

- How much do they shop around?
- How bothered are they about price differences of differing amounts?
- How much do they value the relationship and brand?
- What is their experience of dealing with the insurer?
- How affected are they by competitors' marketing?
- What else is going on in their lives?

Retention and Conversion Modeling

Introduction

Drivers

Measures

Models

Measurable factors

- Who are your customers?
  - age of policyholder
  - claims history
  - product features
  - other rating factors
  - endorsement activity
  - lifestyle factors
- How do you connect?
  - tenure
  - distribution channel
  - payment method
  - affinity membership
  - other products held
- What have you done to them?
  - proposed rate change
  - last year's rate change
  - cumulative rate changes
  - communications
  - claims service
- What have others done to them?
  - competitors' premium
  - competitors' marketing
  - product differentiation

Controllable factors

- Price
  - Absolute cost of products to customer
  - Component cost of parts
  - Product design
- Claim handling
  - Brand value impact
    - "To tell you the truth, we assume that you do. At Hiscox, we always start by assuming that your claim is valid"
  - Individual customer relationship
- Relationship management
  - Cold calls and direct mail

Measuring

- Price change at renewal
  - Raw price change (New/Old)
  - Adjusted price change (New / Last year's price for this year's risk)
- Market premium
  - Rate filings (can be hard work!)
  - Caller questioning
  - Third party provision (broker quotation systems, mystery shopping)
  - Use index or look at New / Market
- Brand surveys
  - Customer satisfaction by segment
  - Log of marketing activity

Mystery shopping
Competitor Market Analysis (CMA)

- Get a small basket of quotations
  - Few hundred to few thousand
- Design basket so that you have cover key policy characteristics
  - More detail in important segments
  - For large samples, mix designed quotes with random sample of characteristics
- Analyse results to approximate rating structures / produce market price index by segment
- Add to database to help predict customer behaviour

Retention and Conversion Modeling

Introduction

Drivers

Measures

Models

Technical details

- How do you include price change and competitiveness in model?
- How do you set up the model?
- How do I know my model is good?

Linear versus non-linear

- Varying views:
  - Simplistic
  - Complex linear
  - Non-linear
- Consider relative competitiveness for retention as an example
  - Our price / market price

Simplistic

- Treat as variate, assume linear

$$\mu = g^{-1}(\Sigma X_j \beta_j + c(p/m))$$

- Assumes “same” price sensitivity everywhere
  - Central probabilities more sensitive, extremes less sensitive
- No-one does this, but helpful to understand issues

Complex linear

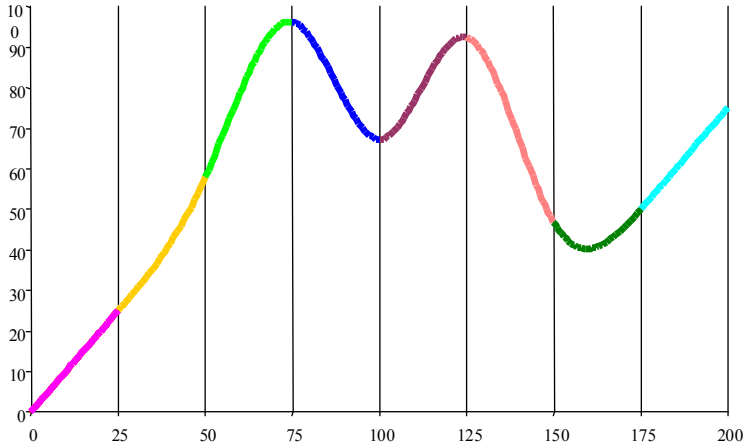
- Use a continuous function of competitiveness

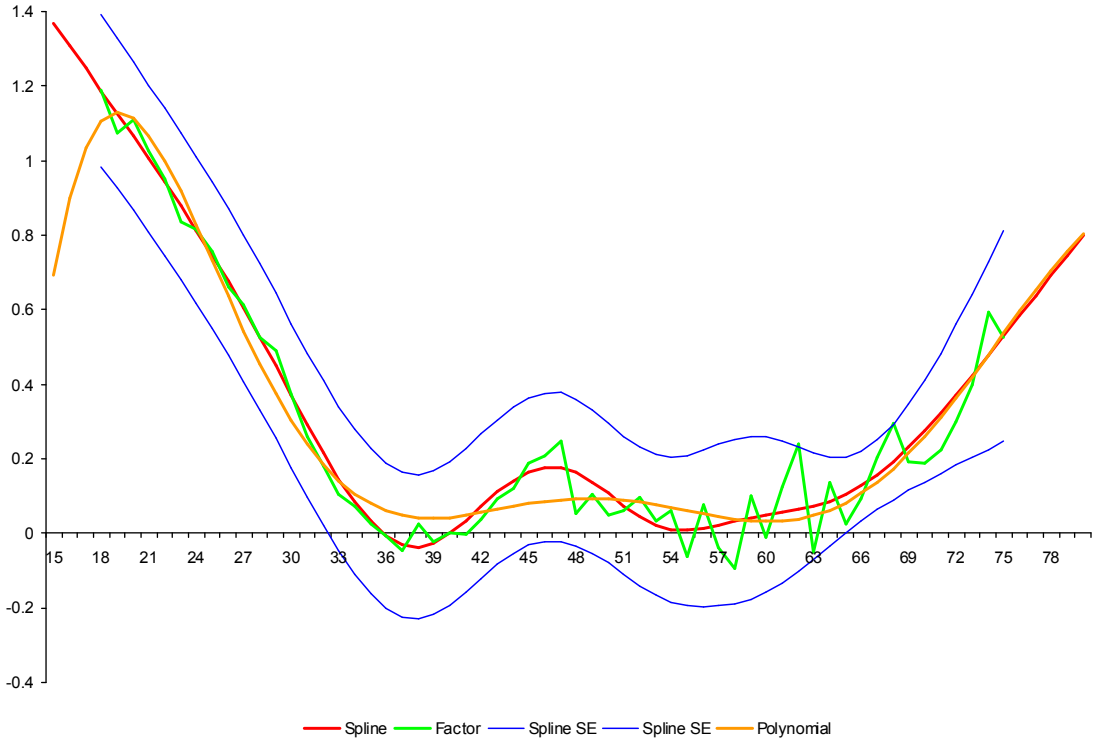
$$\mu = g^{-1}(\Sigma X_j \beta_j + c_k.f_k(p/m))$$

- Function is polynomial or spline
- Can interact with other variables to achieve range of shapes
- Simple to apply with existing tools

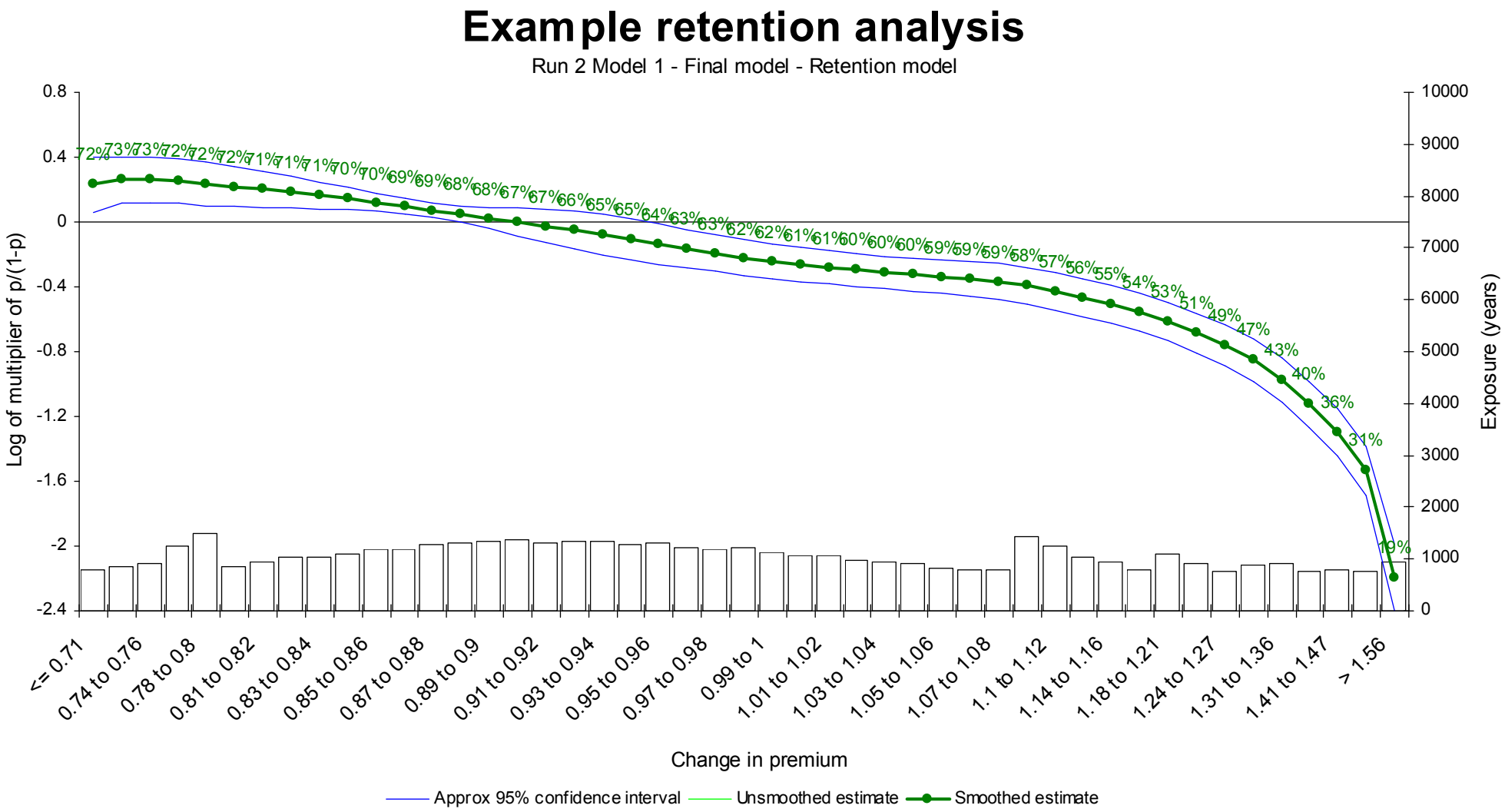
Spline definition

- Regression splines important when modelling elasticity in optimisation context

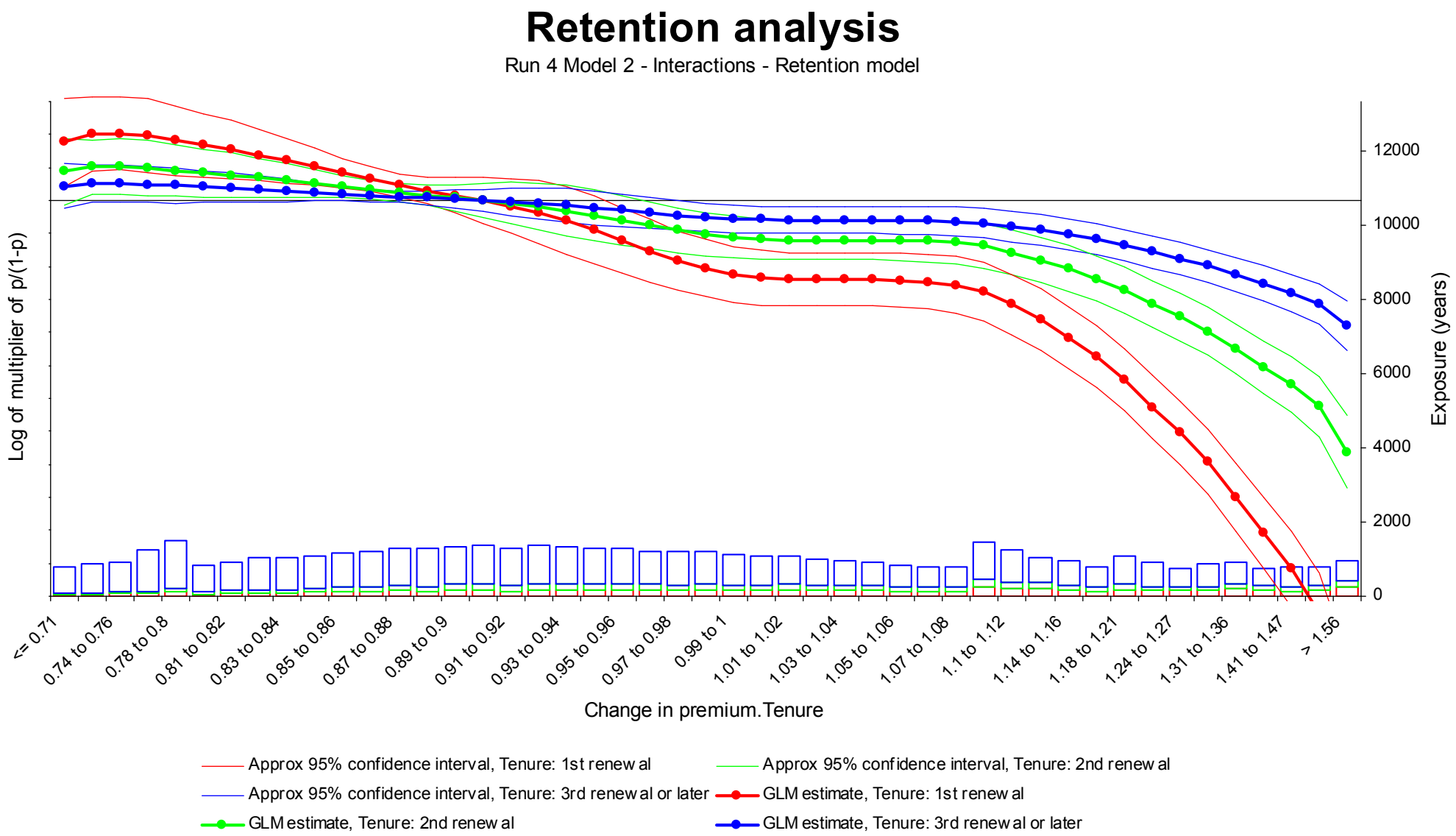




Example retention: elasticity curve



Example retention: elasticity curve



Non-linear

- Linear in competitiveness, gradient varies by segment

$$\mu = g^{-1}(\Sigma X_j \beta_j + (p/m).\exp(\Sigma Z_j \gamma_j))$$

- Similar issues to simplistic, but locally OK
- Hard to fit due to co-linearity of parameters

What is best approach?

- Splines and interactions with splines give wide range of shapes
- Significant evidence that price sensitivity is not “constant”
  - Asymmetric utility curve
- GLM complex linear practical and effective

Negative elasticity

- Interaction curves may not be monotonic
- Negative elasticity can be local or global
- May be:
  - Noise: Reduce number of knots or smooth
  - Genuine effect: Look for consistency and understand segment
    - Eg Large premium fall may increase lapse rate
- Don’t throw the baby out with the bath water

Missing values

- Missing values in a variate cause entire record to be ignored
  - Replace missing values with zeros
- Care is needed to differentiate "real" zeros and "missing" zeros
  - Create a missing flag and include in all models involving variate
  - Remember spline basis functions transform zero to some other (non-zero) value (extrapolation)

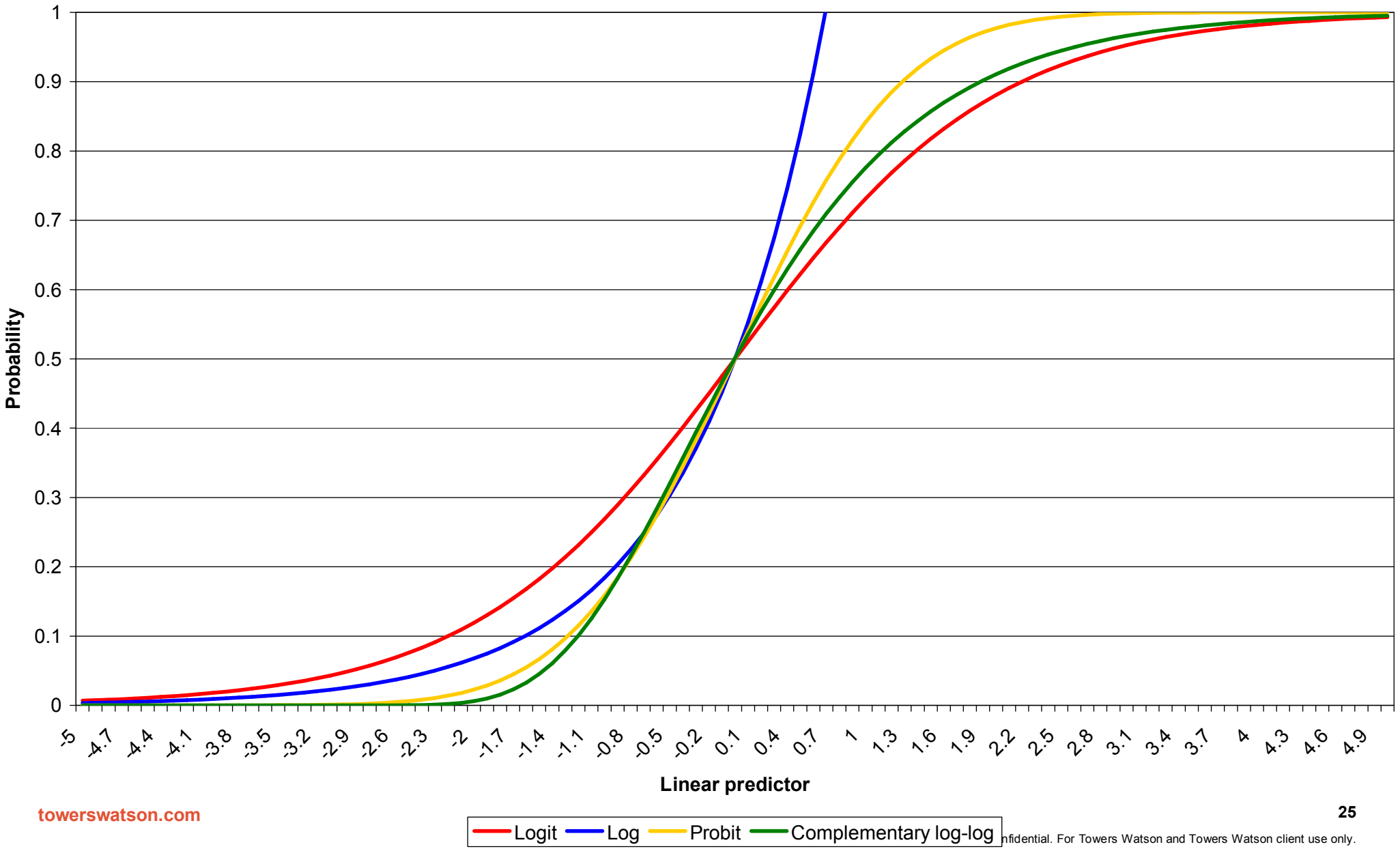
Technical details

- How do you include price change and competitiveness in model?
- How do you set up the model?
- How do I know my model is good?

GLM choices

- Could try:
  - Log link
  - Logit link
  - Probit link
  - Complementary log-log link

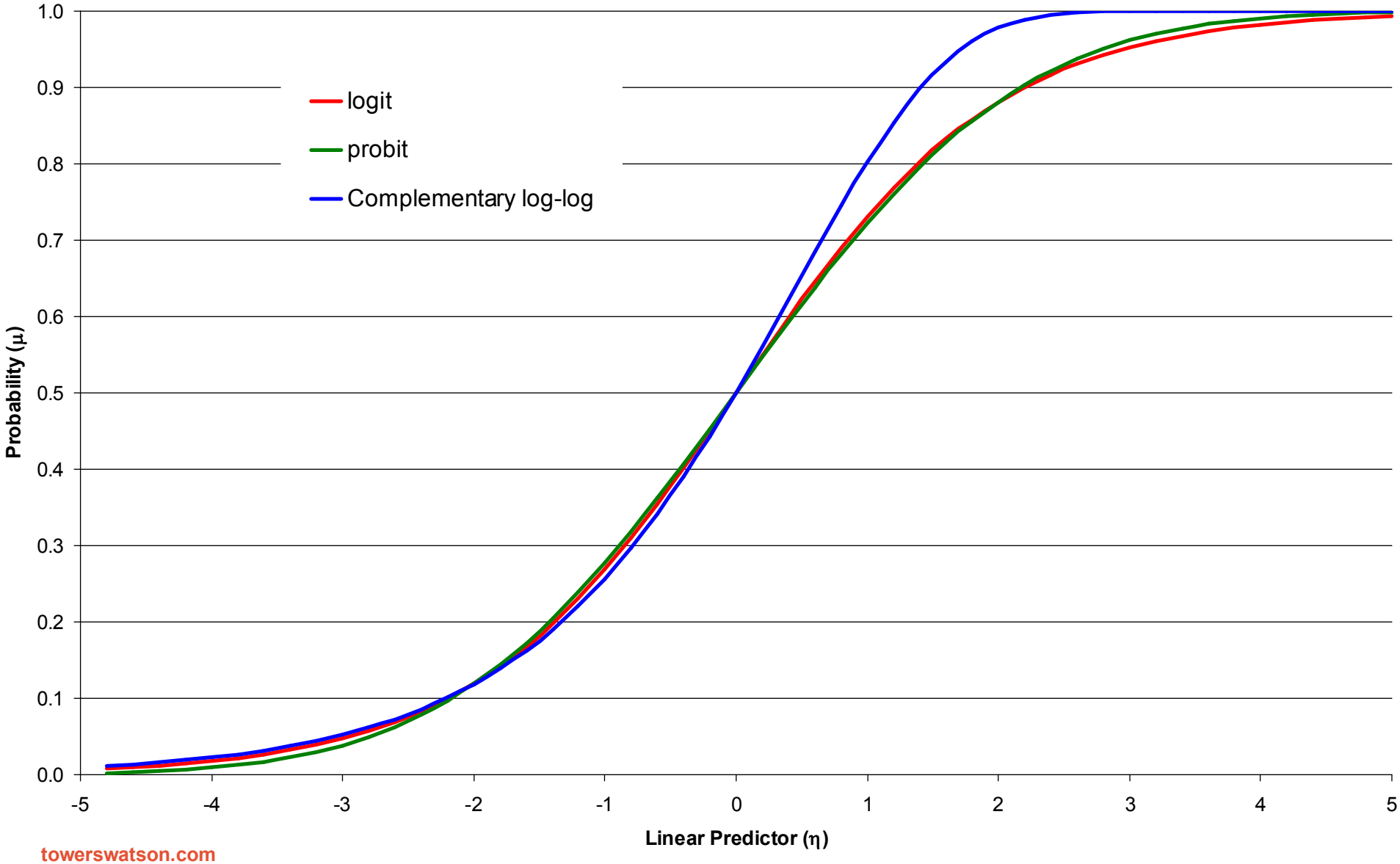
Link functions
Rescaled so 0 => 0.5



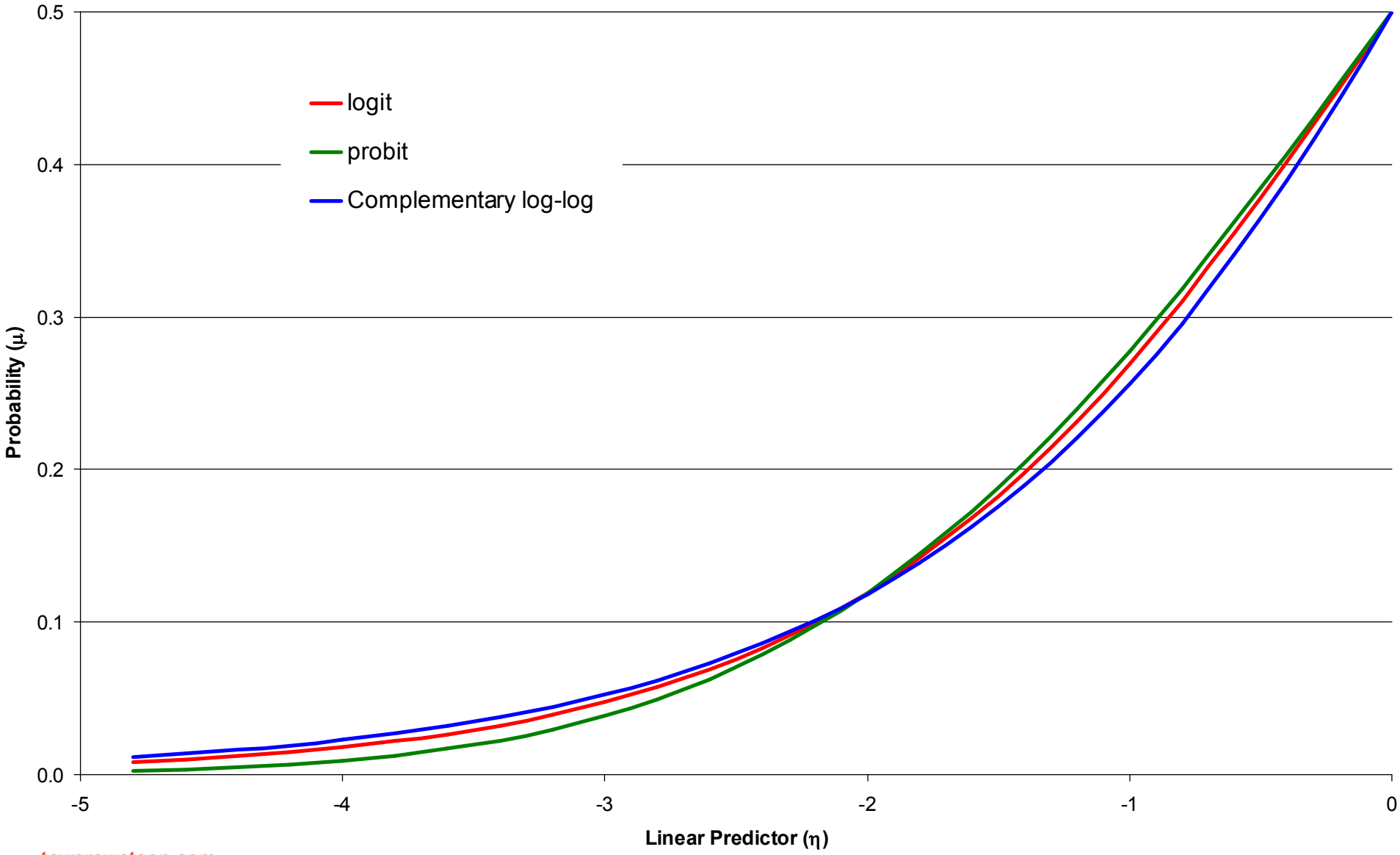
Range of values

- If all probabilities are similar and small (<10%) then any link function can work
  - Logit is asymptotically log
- If range of probabilities is large then choice of link can have big effect
  - Log becomes unsuitable once p>0.1
  - Complementary log-log is not symmetric, so take care what you model

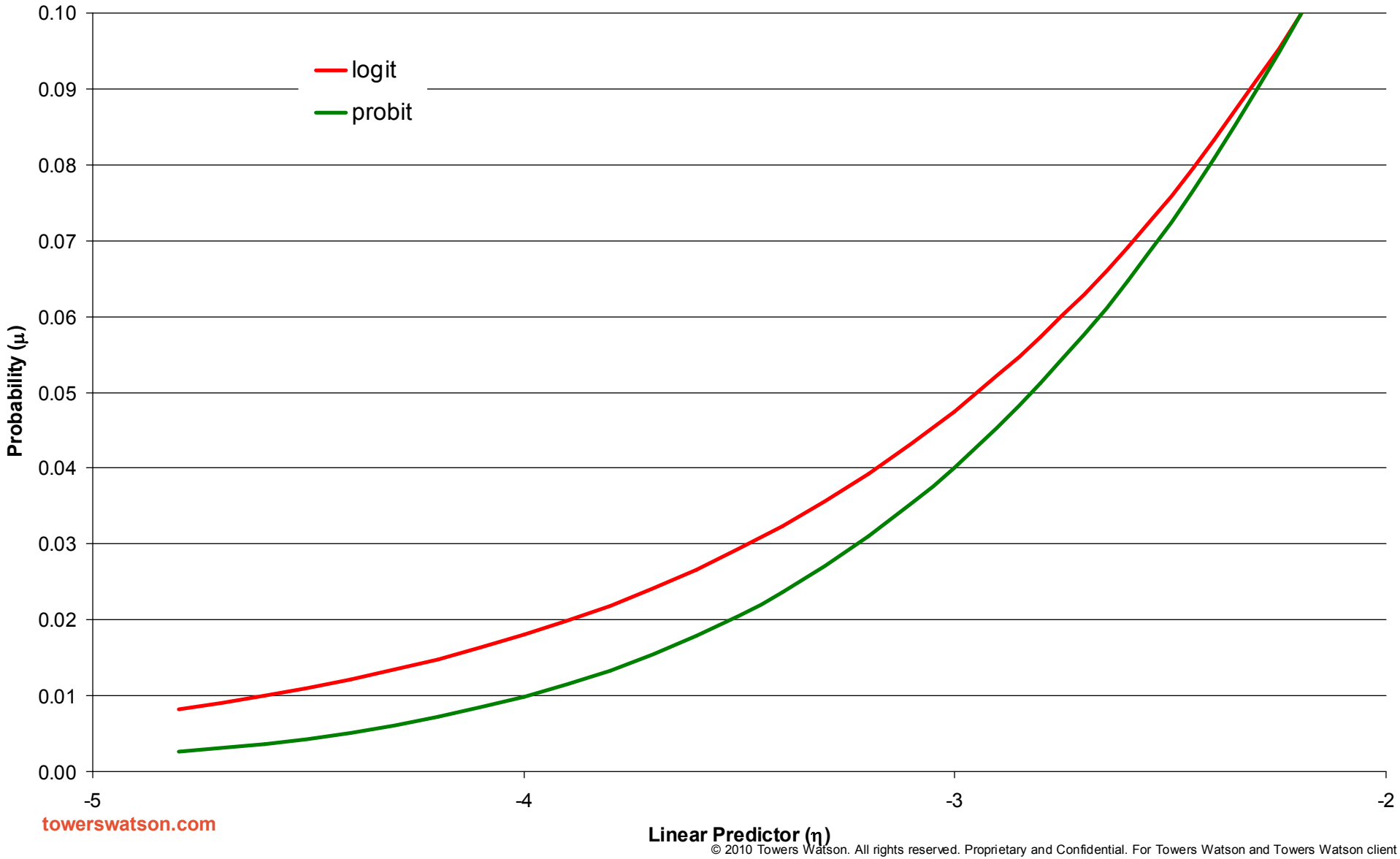
Link functions
Rescaled to be the same at 0 and -2



Link functions
Rescaled to be the same at 0 and -2



Link functions
Rescaled to be the same at p=0.1



Probit and logit

- Probit and logit are different shapes close to 0 or 1
- Both are symmetrical, and similar for central probabilities
- Probit gets close to zero quicker than logit
  - May produce better fit where probabilities ranges from small to very small
  - Try both probit and logit to see which gives better prediction

Technical details

- How do you include price change and competitiveness in model?
- How do you set up the model?
- How do I know my model is good?

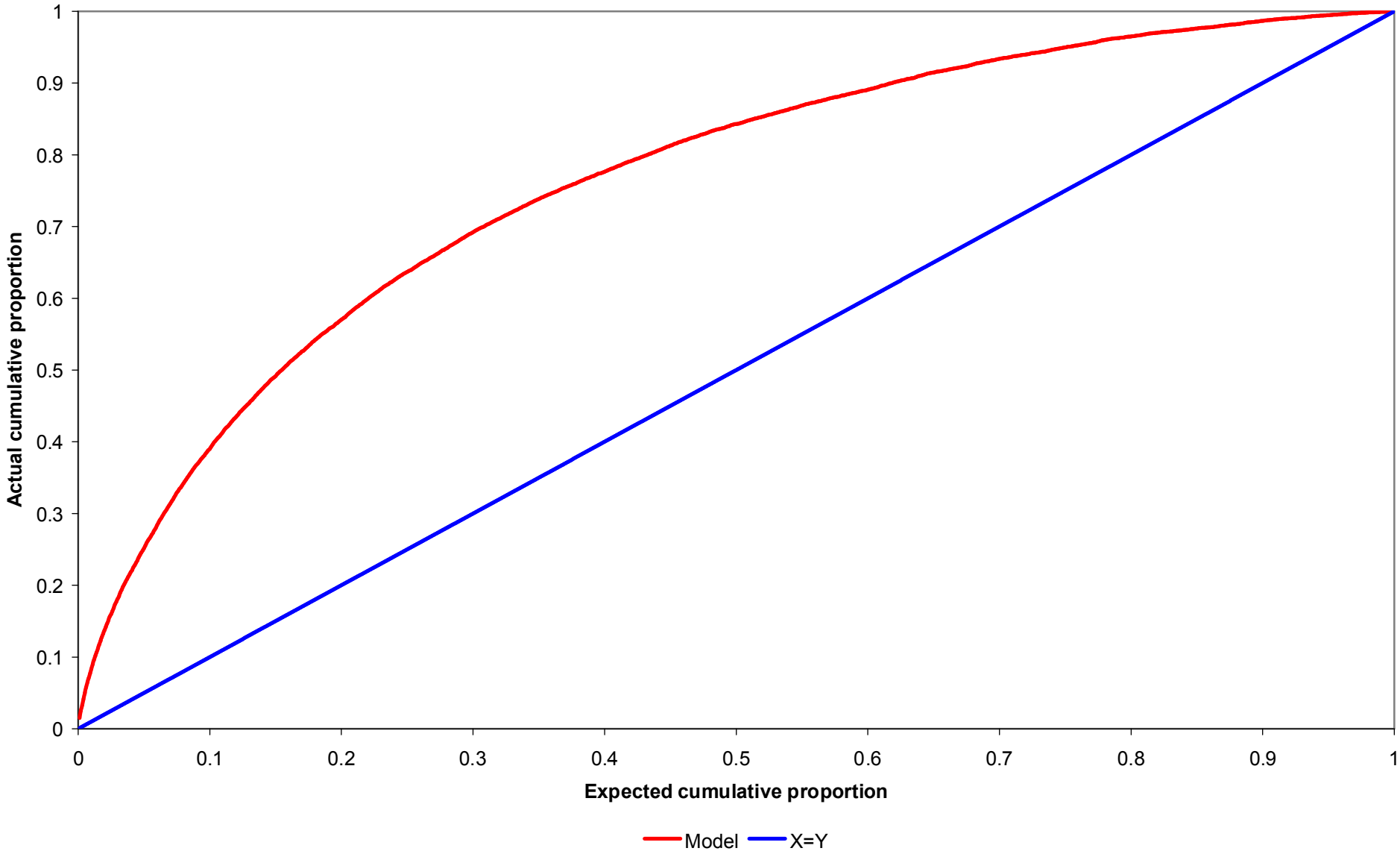
Model validation

- Actual versus expected
  - How well does it fit (predict) the data?
- Lift curves
  - Which rival model separates good and bad risk most effectively
- Loss ratio impact
  - Is the new model better than the old?
- Gains curve
  - Is the model good at separating good and bad risk?
  - Traditionally used for probability models

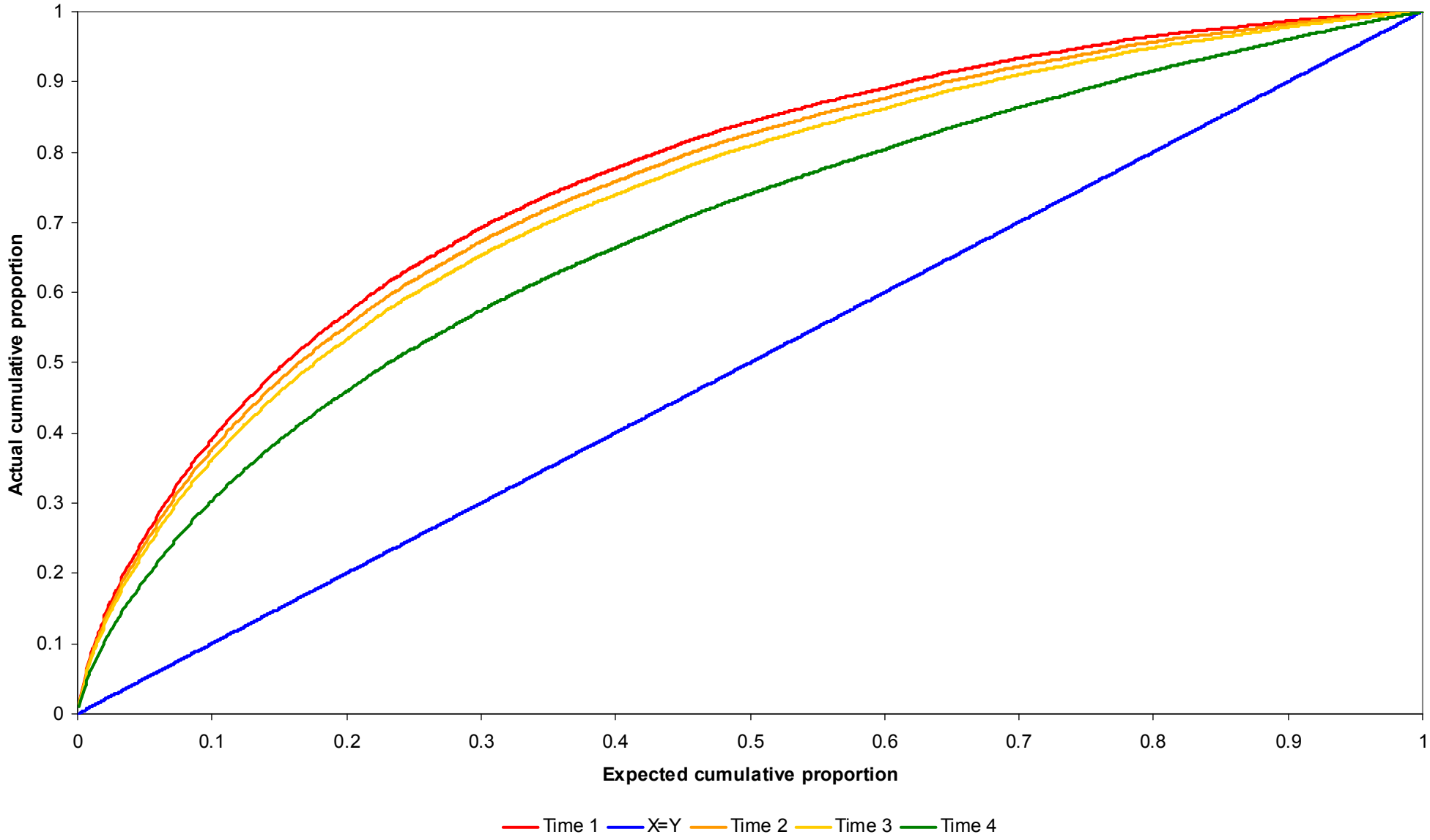
Model validation

- Split data into modelling and validation subsets
- Repeat validation on new data as it comes in to monitor health of model

Gains curve



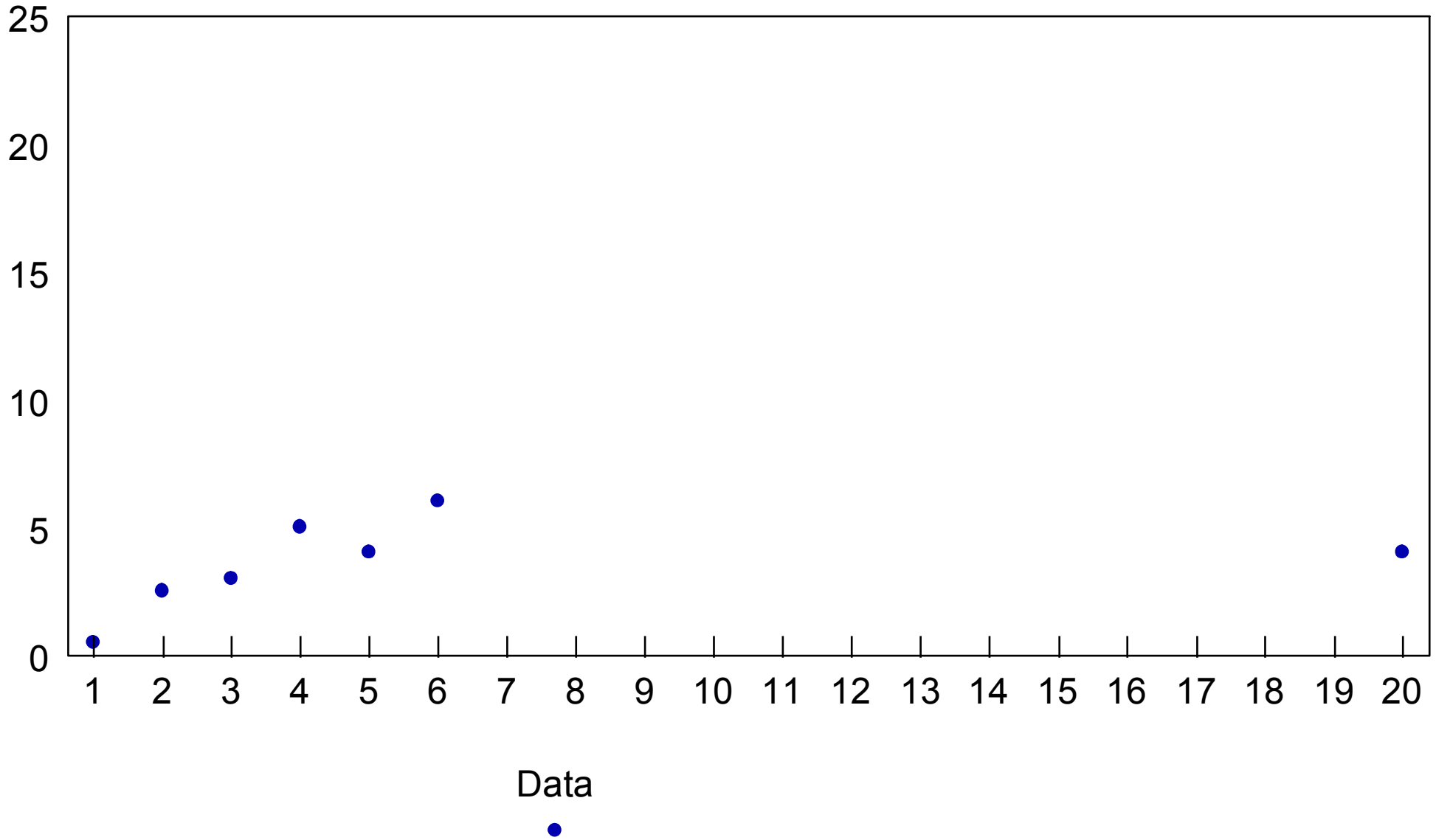
Gains curve: Deteriorating experience



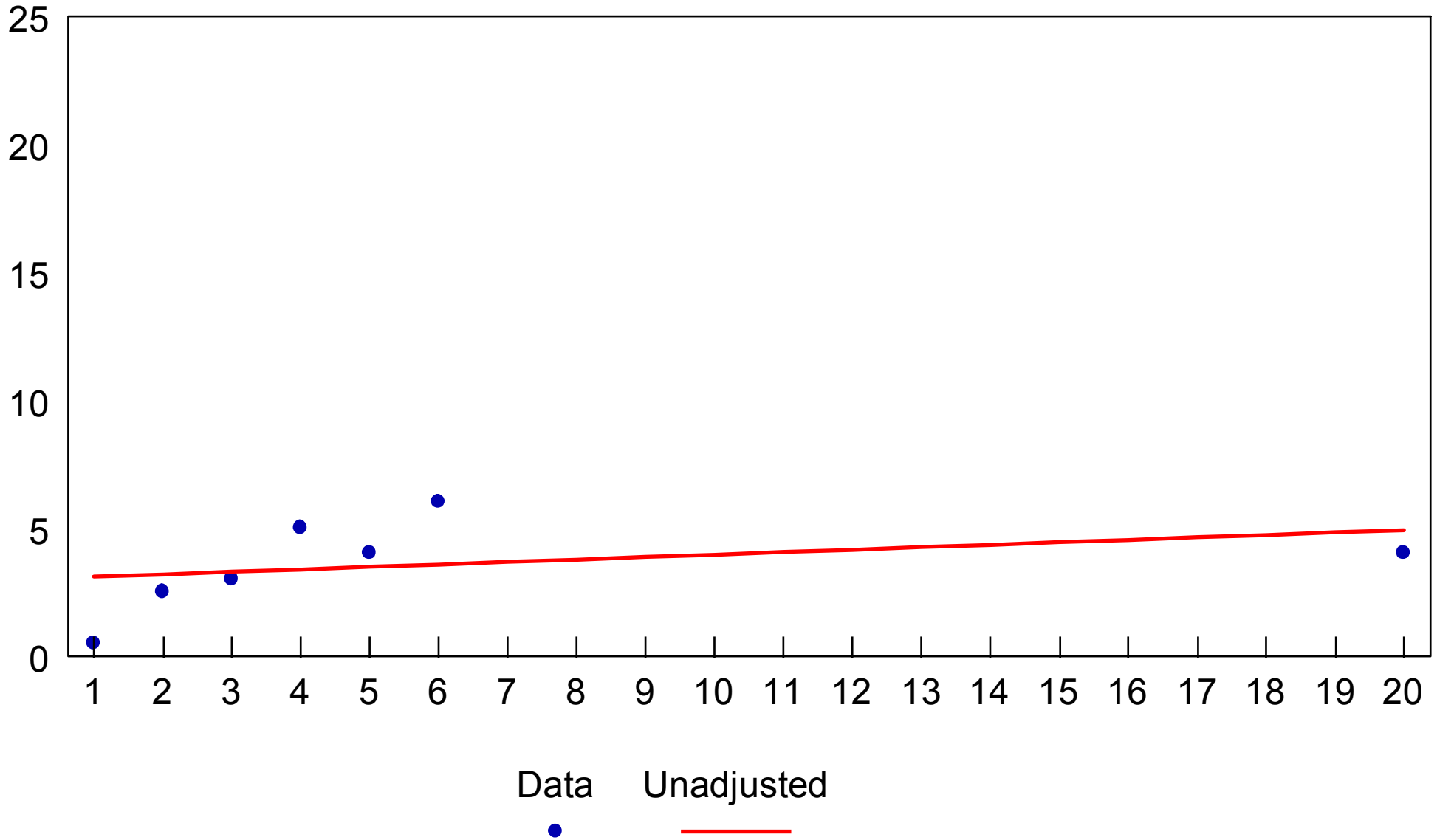
Edge effects

- One or two records with extreme variate values can have a disproportionate effect on the model
  - Look at leverage or Cook's distance
  - Understand your data
  - Consider limiting range of variate
  - Be careful when extrapolating

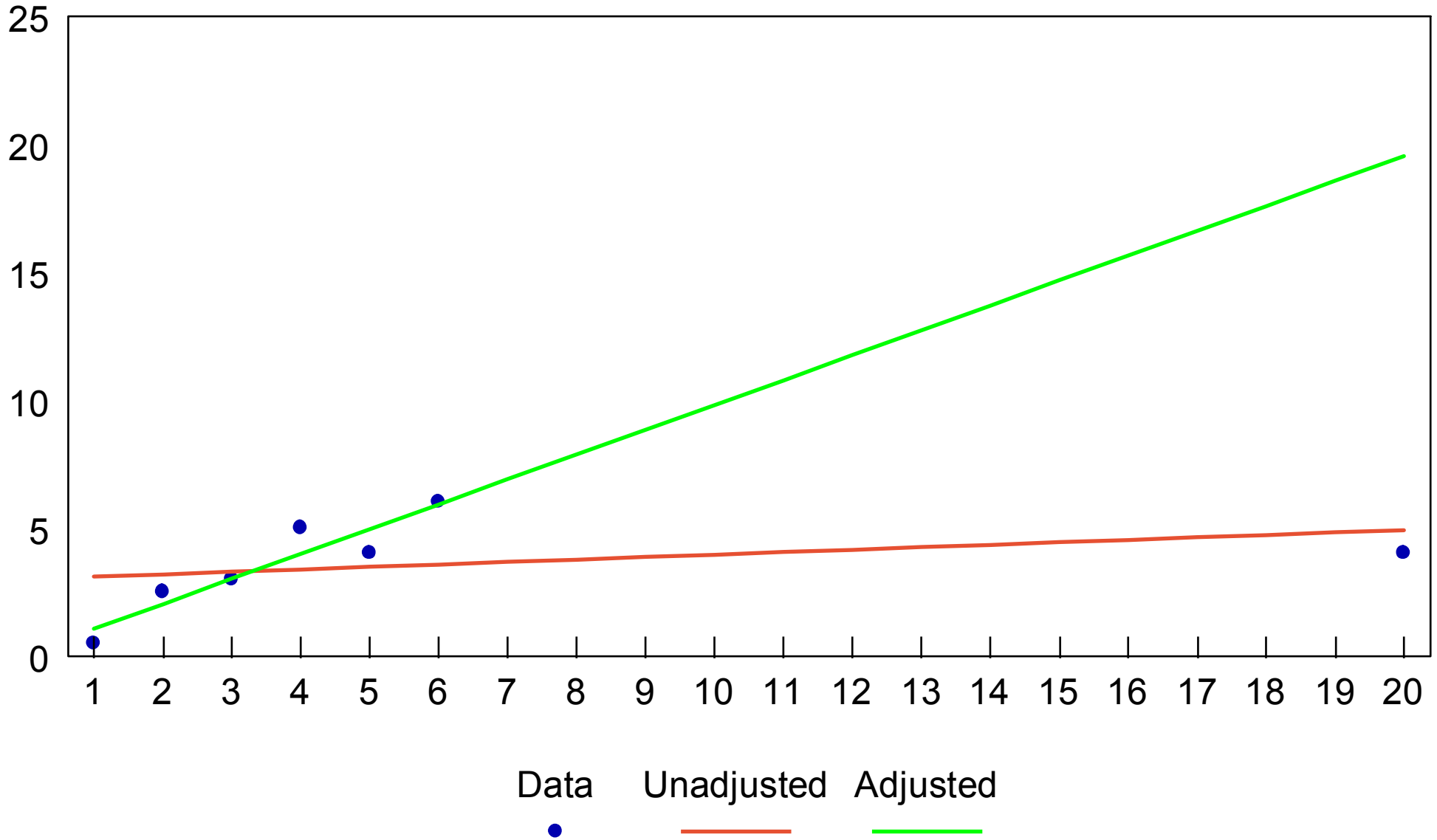
Leverage



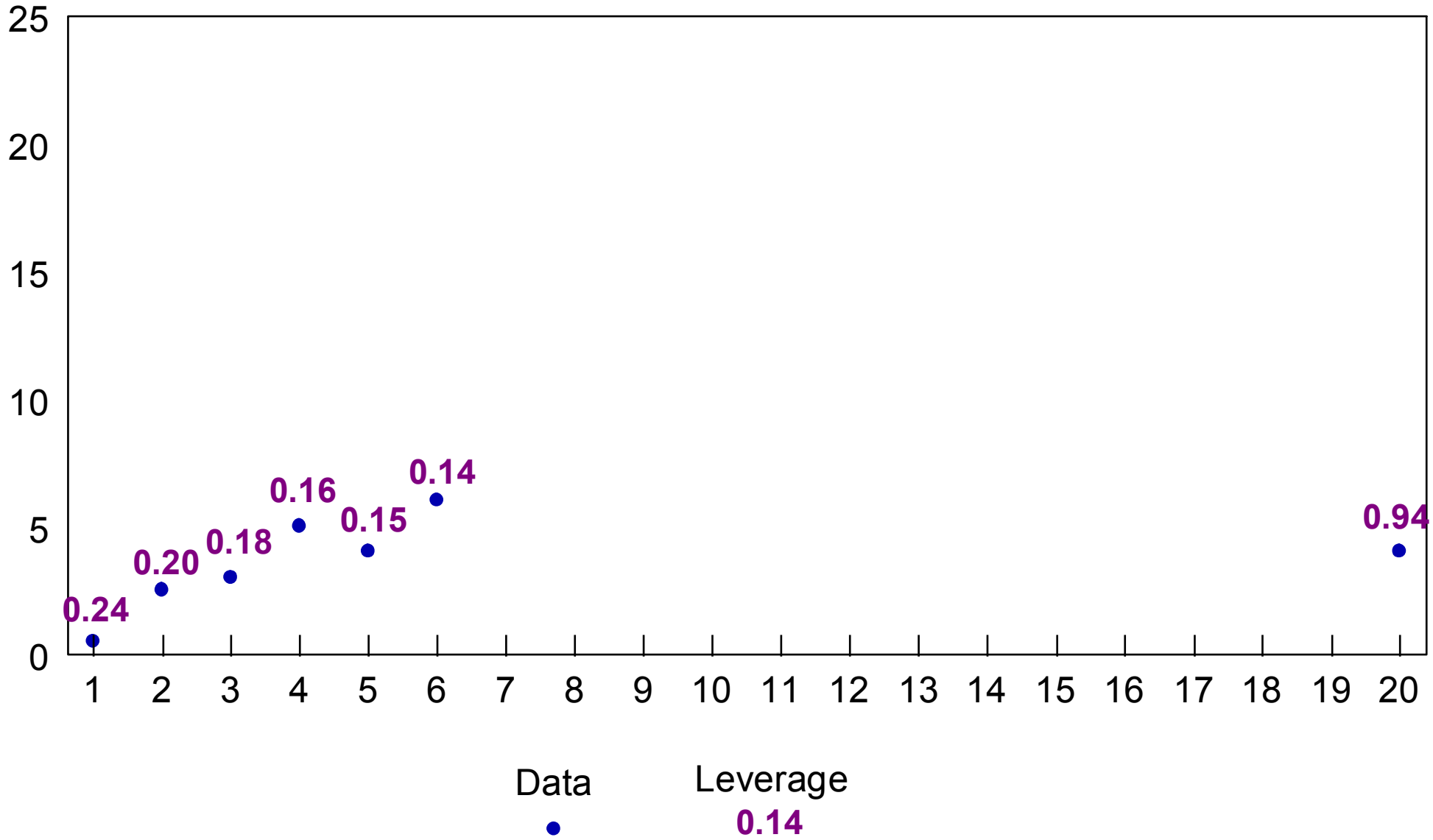
Leverage



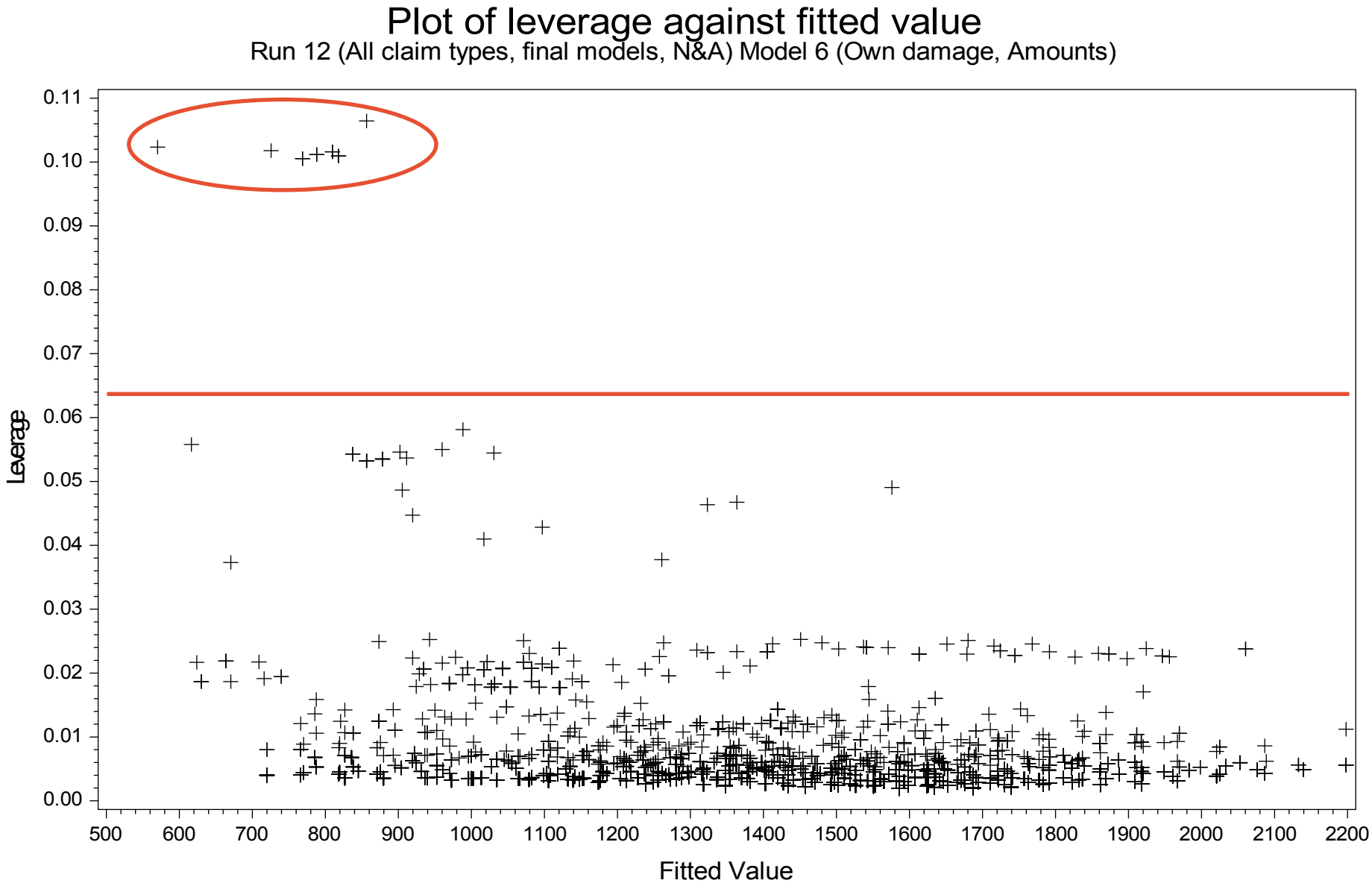
Leverage



Leverage



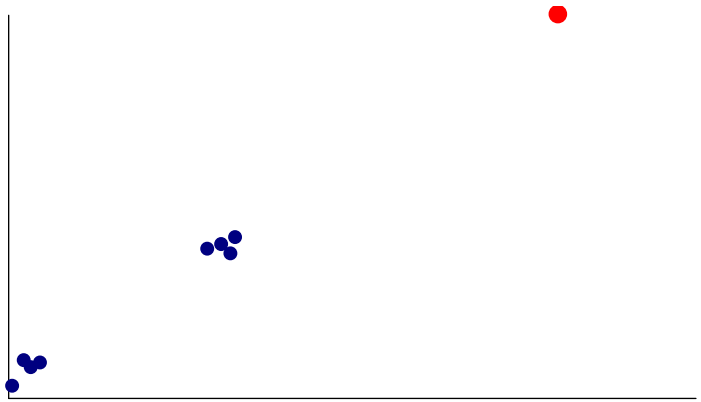
Leverage



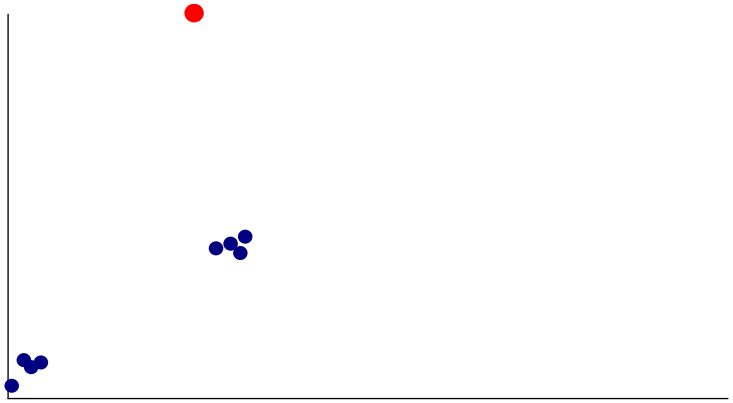
Pretium 08/01/2004 12:32

Leverage cf Cook's Distance

- Large leverage but small Cook's distance



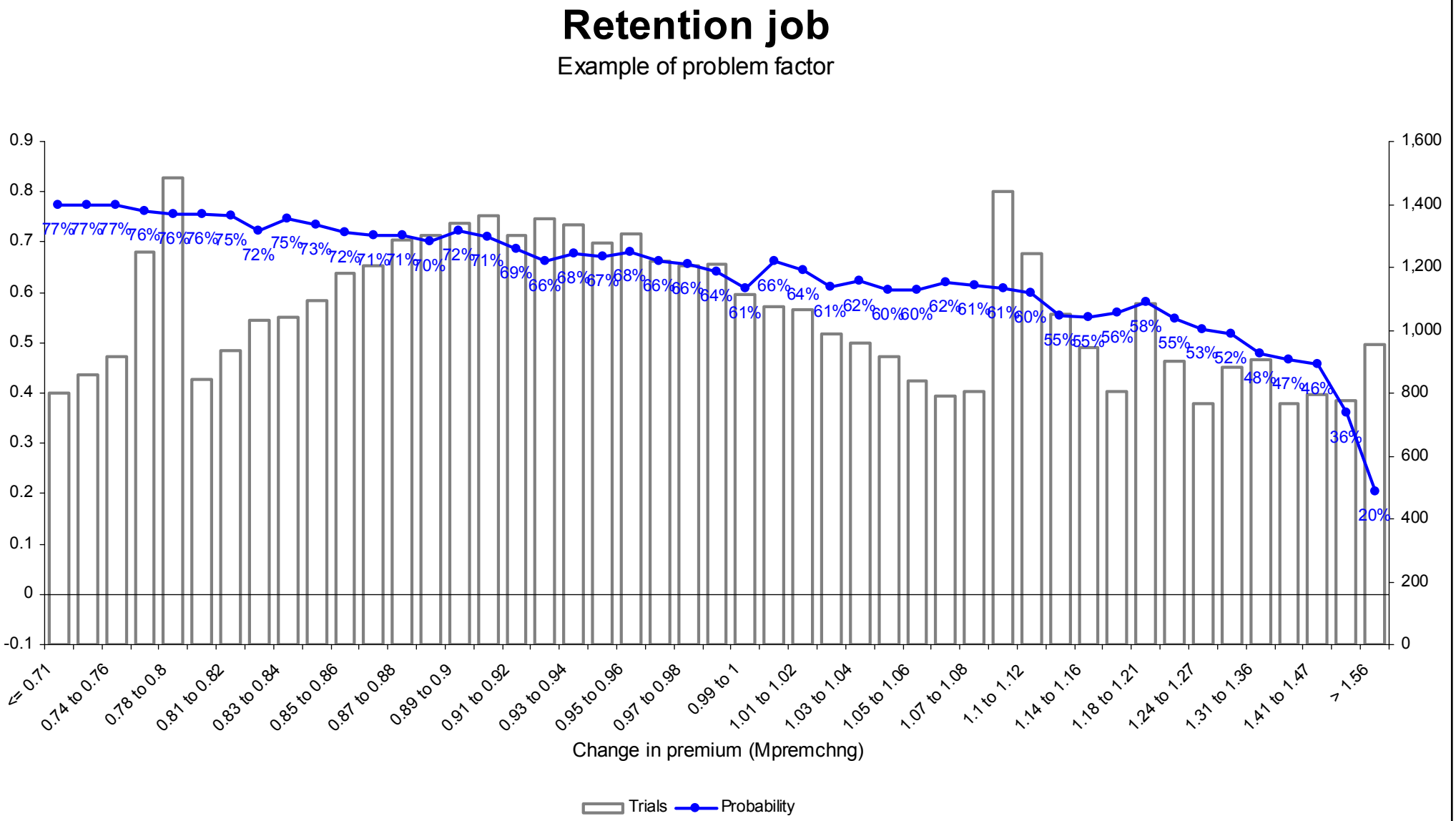
- Large Cook's Distance but small leverage



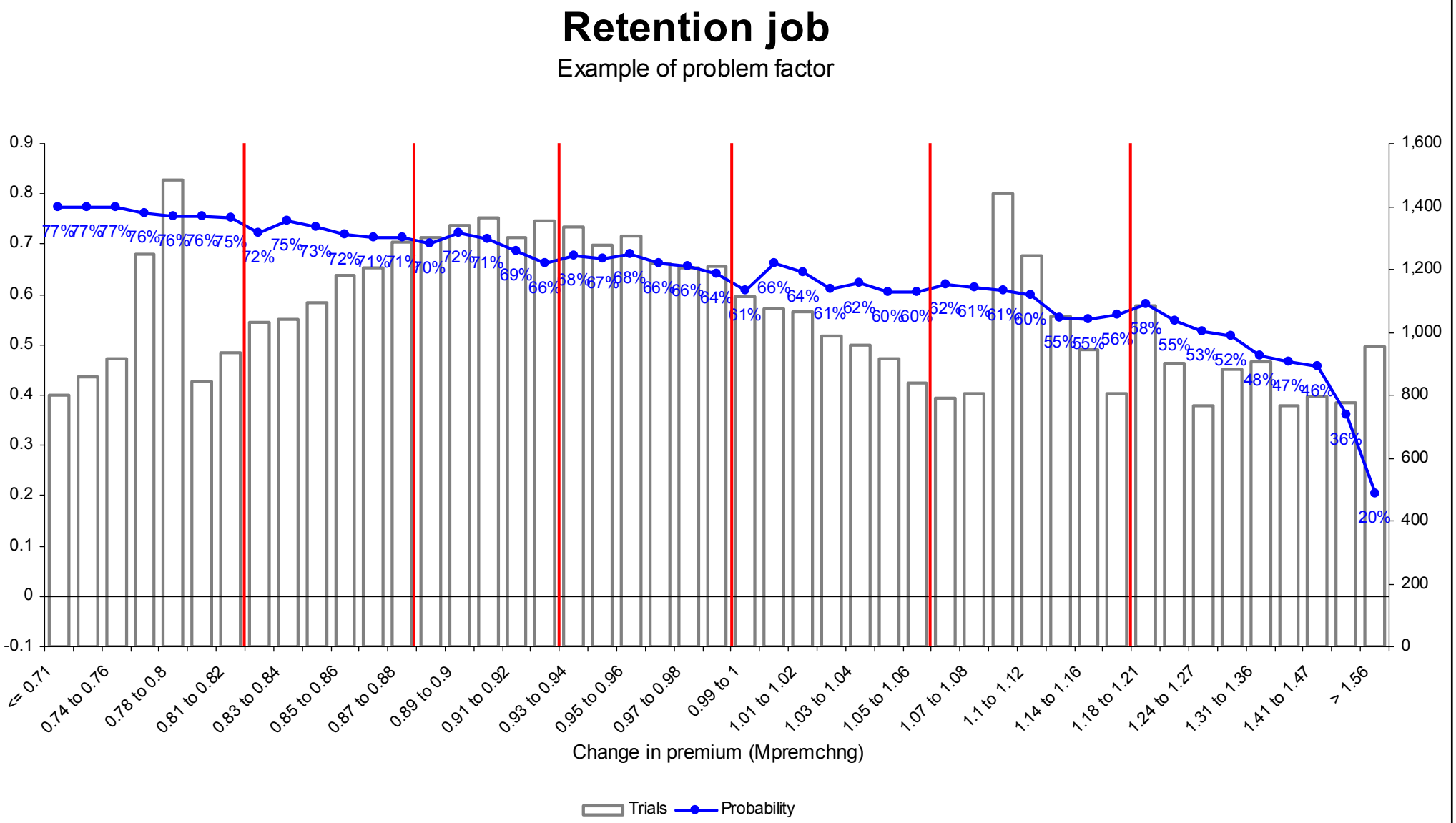
Cautionary example

- Artificial data, loosely based on actual naive analysis
- Retention analysis containing three records with incorrect premium change, all of which renewed
- Problems:
  - Overfitting to edges
  - Knot placement

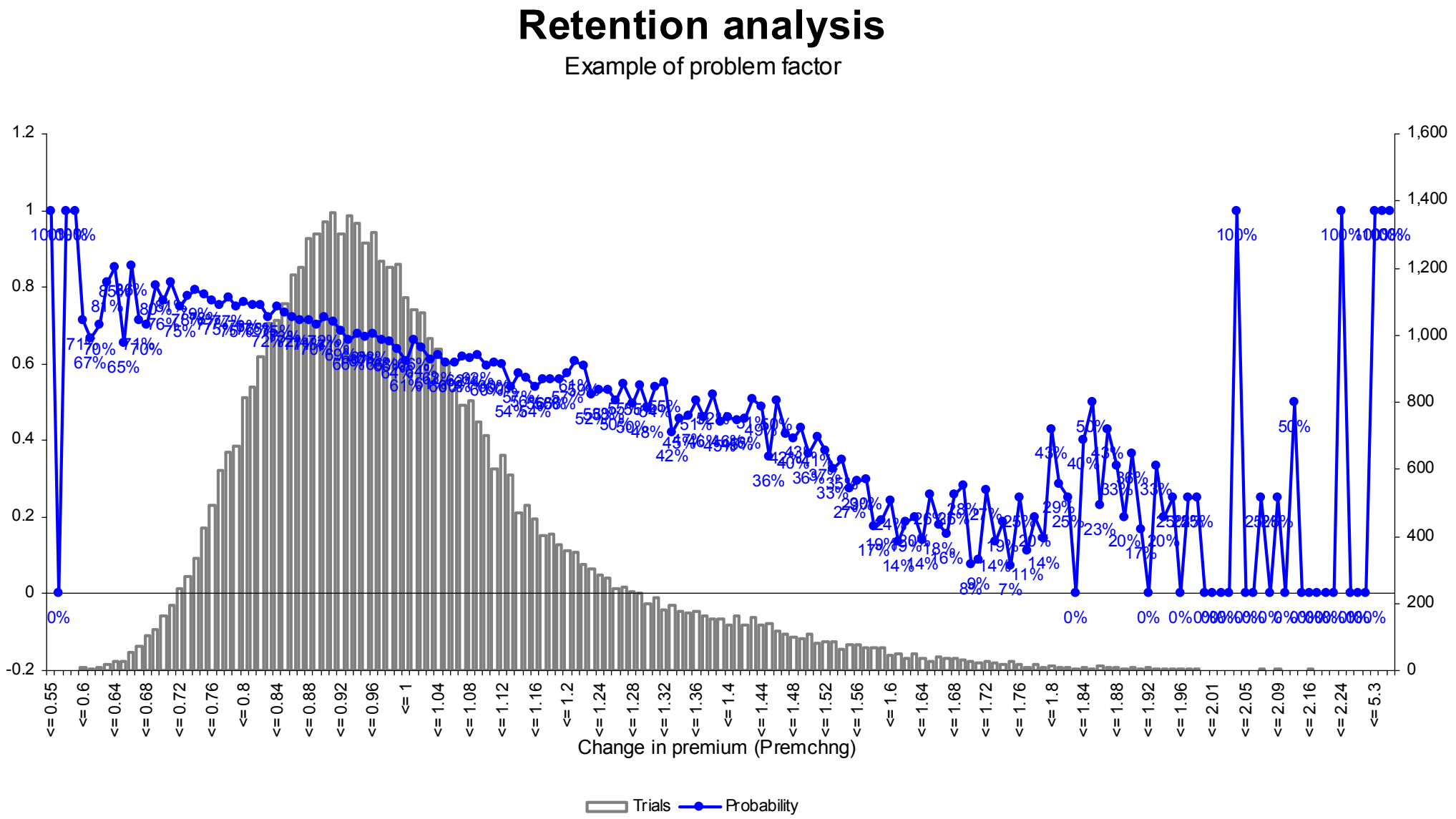
Simple grouped oneway



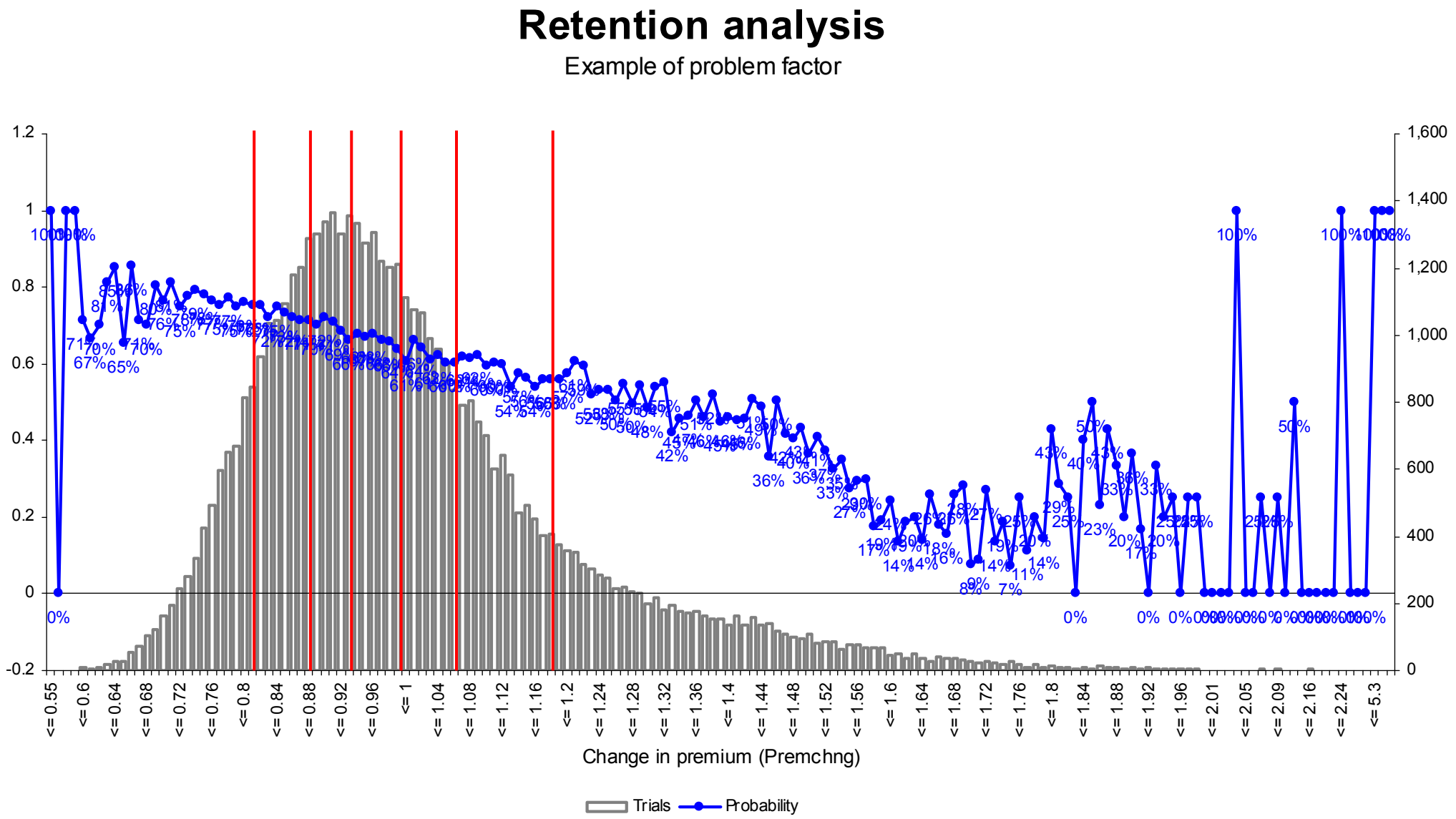
Simple grouped oneway



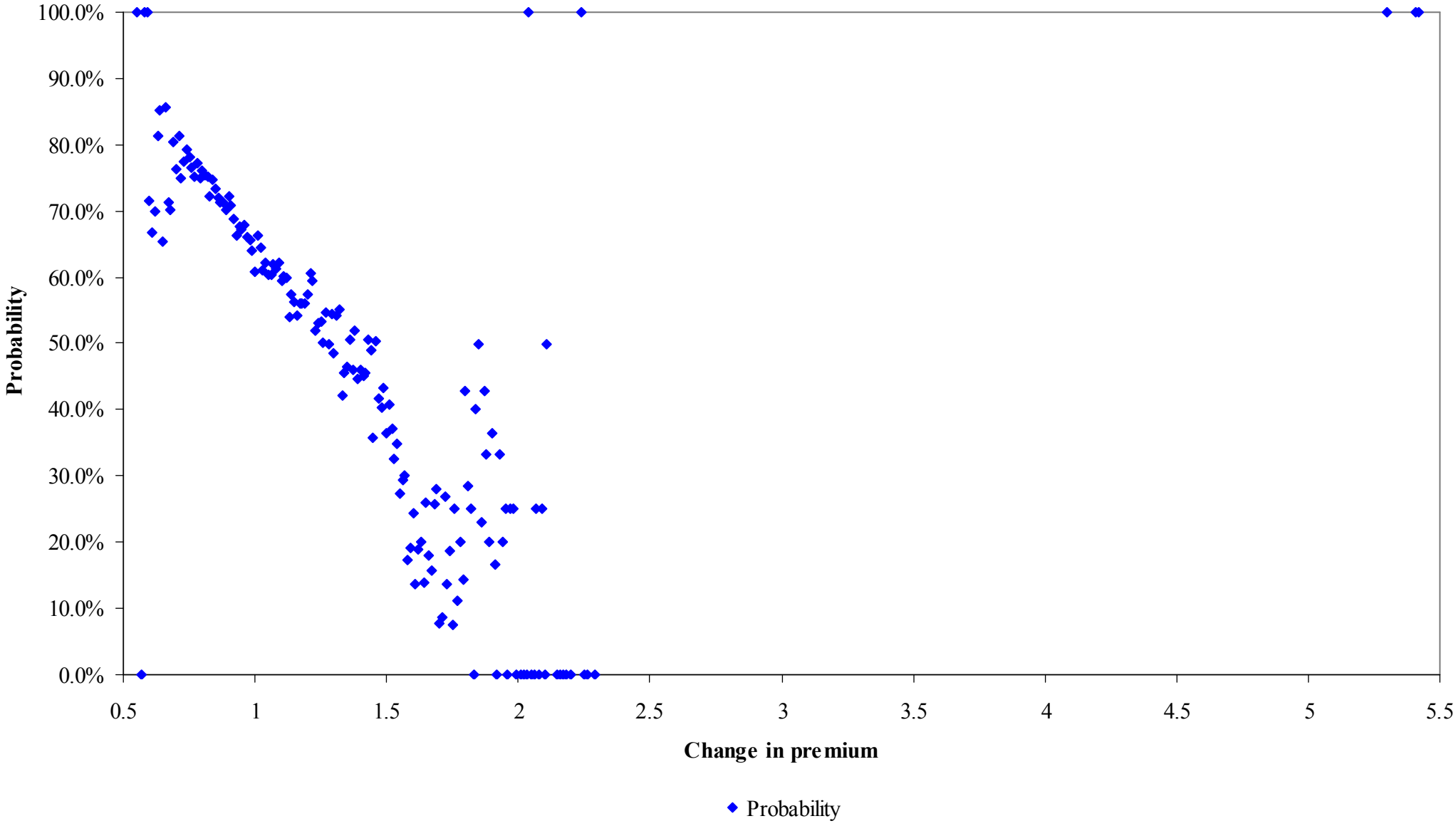
Detailed oneway



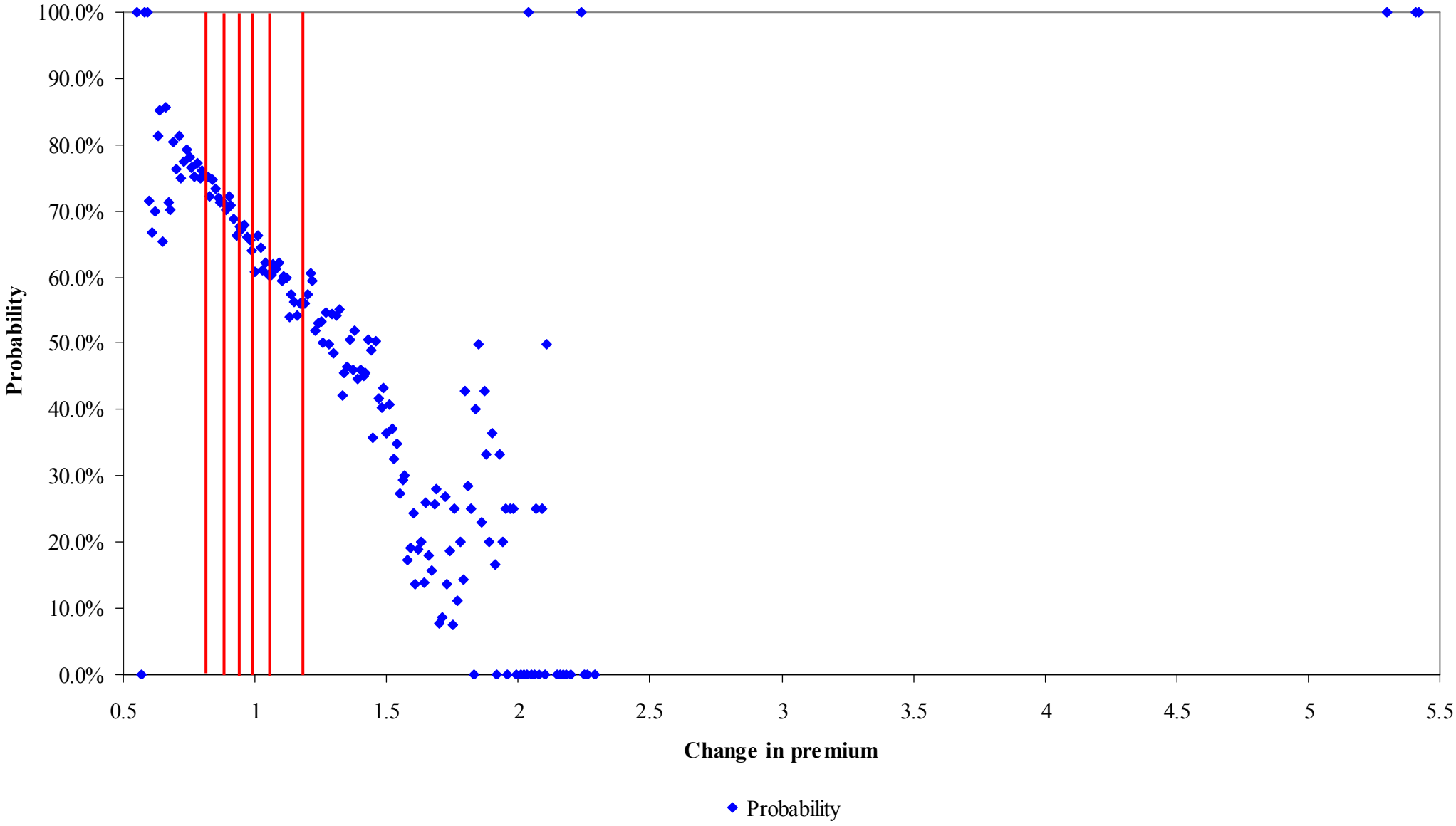
Detailed oneway



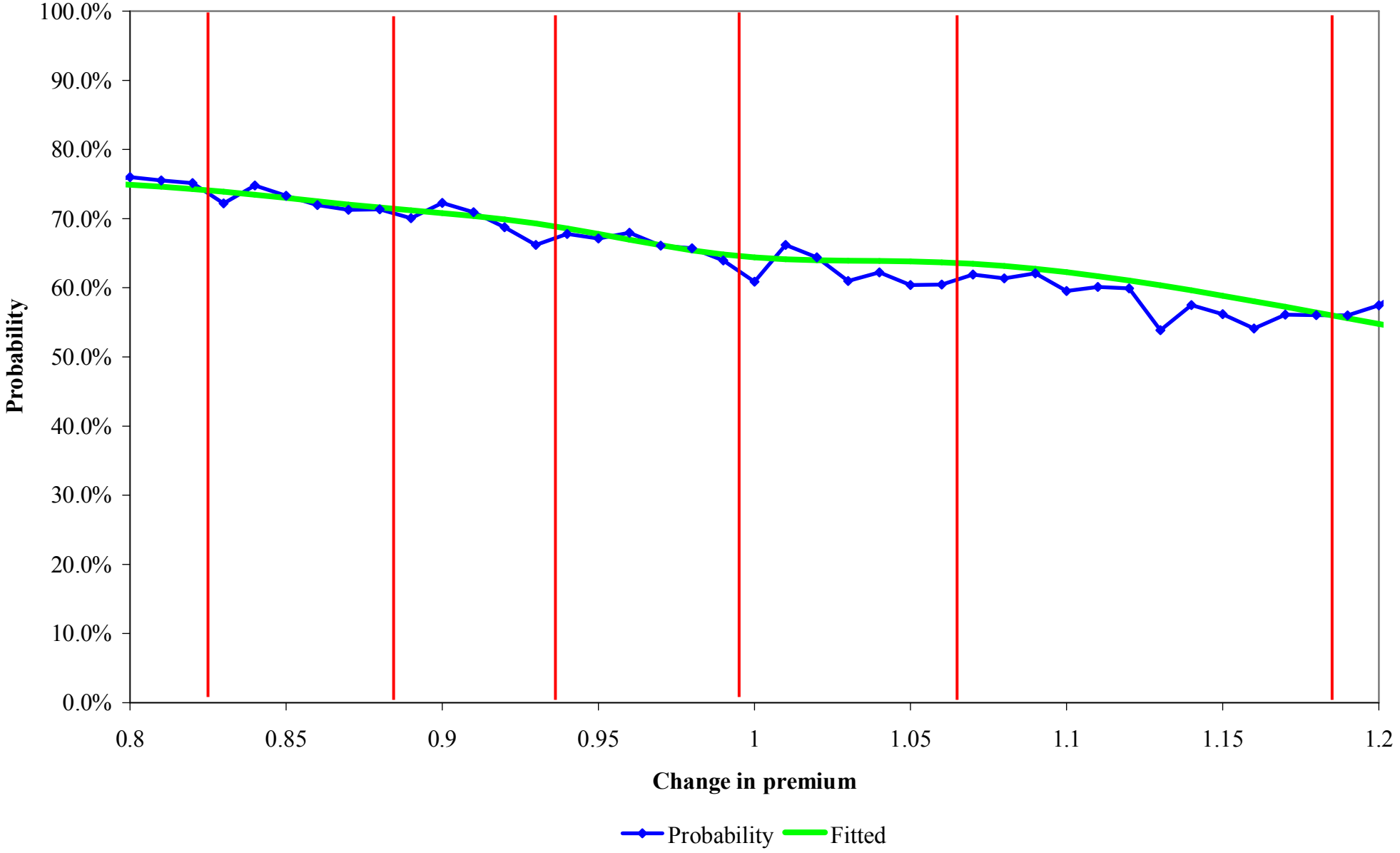
X-Y plot



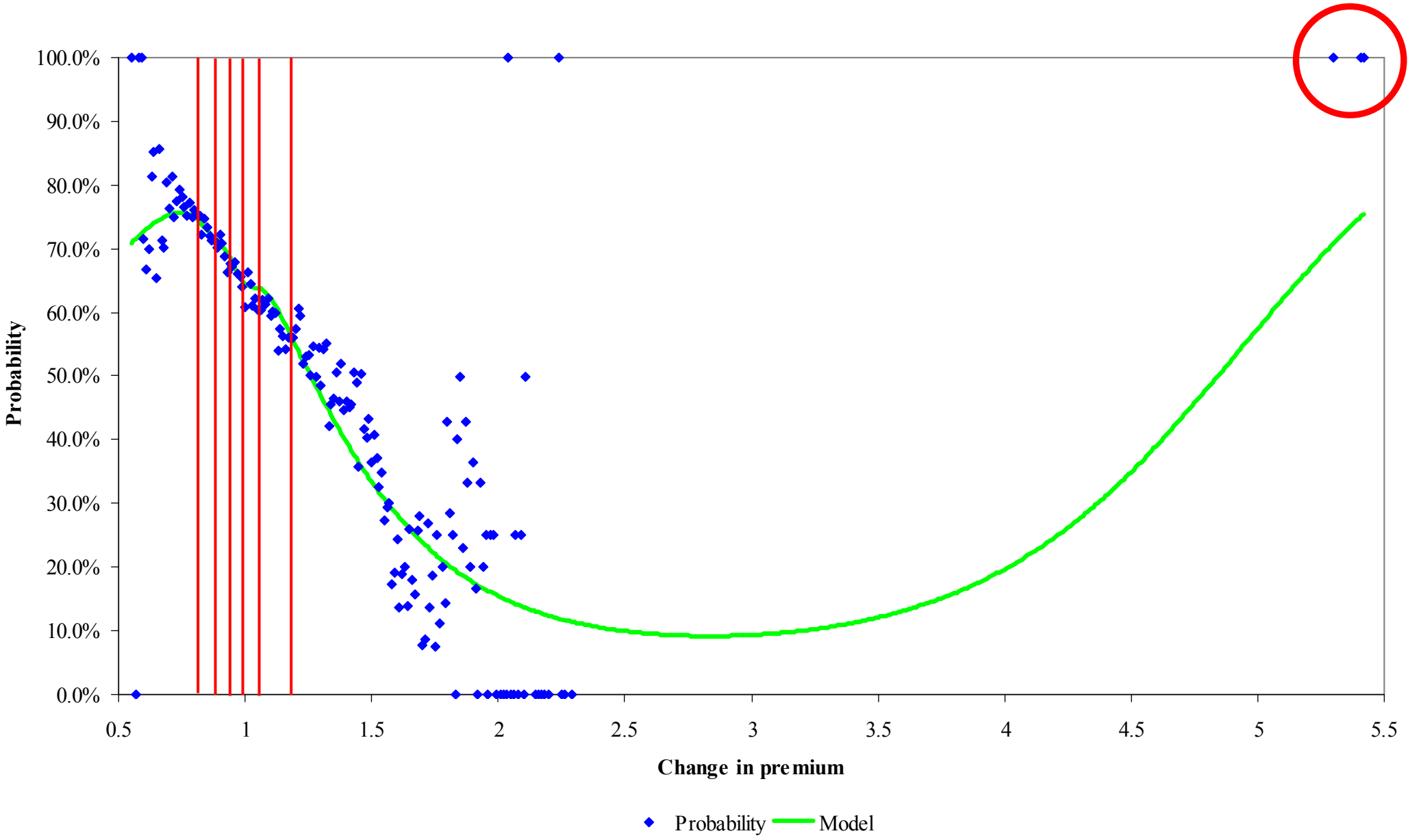
X-Y plot



Model results



Model results



Solutions

- Cap and collar price changes in data
- Understand your data and results
  - Thought needed
- Look at leverage
- Take care in extrapolation

Questions?

- James Tanser
  - Senior Consultant, Towers Watson
  - james.tanser@towerswatson.com
  - +44 1737 274249