“Individual” modeling – what, when and a little bit of how

Richard Brookes

Agenda

- What
  - Some definitions
  - Aggregate vs individual
  - Simple vs complex
- When
  - Model purposes
  - Illustrative example
- How
  - Structural issues
  - Model evaluation

What

Definitions (imprecise)

- An aggregate model gives cash flow or liability predictions for a "large" group of claims only, recognising that there are many different types of claims in the group
- An individual model gives a "genuine" prediction for each claim, using the actual characteristics of the claim

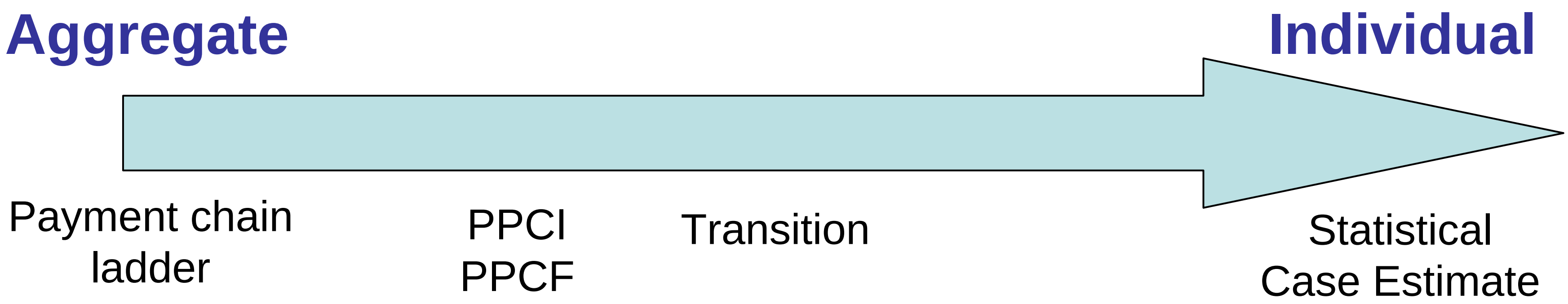
Spectrum

- In fact, models are not often one or the other but somewhere in between
  - Payment chain ladder models are "aggregate"
  - PPCF and PPCI models are thought of as aggregate but can be expressed "per claim" (if we remove IBNR)
    - The claim characteristics are development at the valuation date (or accident year)
    - They make the assumption that all claims with common characteristics are the same

Spectrum

- In fact, models are not often one or the other but somewhere in between
  - Transition models can be expressed “per claim”, again if we remove IBNR
    - Claim characteristics are development year, accident year, state and maybe others
    - Assume that all claims with common characteristics are the same
      - This includes path independence!

Spectrum



- Should we replace “aggregate” with “simple” and “individual” with “complex”?

When

How do we choose?

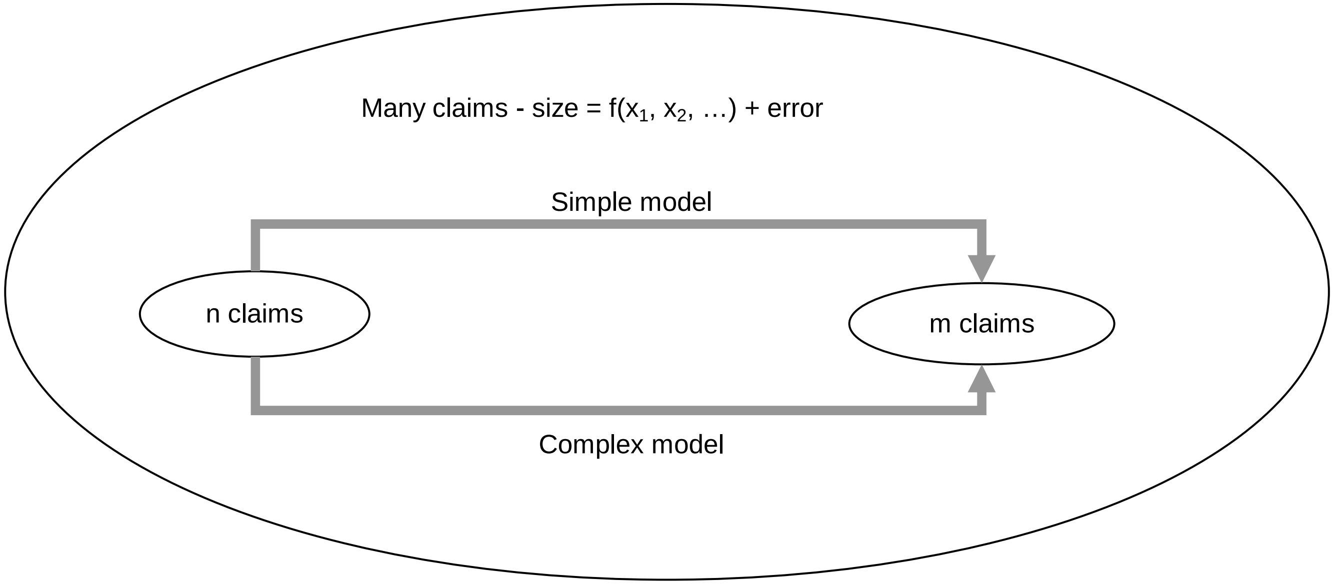
- Purpose is important
  - Reserving; minimise error in aggregate prediction
  - Driver analysis; establish causality
  - Pricing; maximise predicted range of claim size relativity
  - Operational monitoring; benchmark operationally distinct groups;
- It is asking a lot of one model to do all these!

Complex models and reserving

- Illustrative example
  - Assume a large population of claims where the claim size is a function of several time independent claim characteristics plus an error term
  - Take a sample of n claims, drawn independently and fit two models
    - A ”simple” model which is just the mean of realised claim sizes
    - A ”complex” model which has the same structure as the population but with the parameters fitted from the sample

Complex models and reserving

- Now take another sample of m claims and use each model to give an expected liability for the sample



Complex models and reserving

- When n and m are large then both models will likely give a good result
  - The complex model will fit better but both have low prediction errors
- Let's reduce m
  - The probability of the characteristic profile being different than that of the population becomes more significant
  - The complex model gives an increasingly better prediction since it takes account of this
  - Note that if we know in advance what the drivers are then we know when this situation is occurring

Complex models and reserving

- Let's make m large again and reduce n
  - The performance of both models will degrade as parameter error increases
  - As n gets small, the profile of the sample again may be different than the population and again the complex model will perform better
  - How does the number of parameters affect this? Is there a penalty for having many parameters?

Illustrative example – tentative conclusions

- For large stable portfolios, complex models are better but by how much? How much does “random” error matter compared to systemic risks?
- Complex models are most useful for small subgroups of claims or when they incorporate the dependence of a claim characteristic in respect of which the portfolio make-up is changing over time

Complex models and reserving

- Complex models are more likely to be correctly specified (?)
- BUT
    - In practice it can be very difficult to fit complex structures, including interactions, without introducing bias
    - Other risks from model complexity
        - Operational changes/variations in predictors
        - Modeler error
        - Coding/translation error

Purposes revisited

Purpose	Aggregate or Individual?
Reserving	Evaluate risks and implications for capital then decide between simple and complex
Driver analysis	Potentially complex with external rationale
Pricing	Sufficiently complex to take account the differences in claim characteristics between pricing segments Maybe very complex if redefining pricing segments
Operational monitoring	Sufficiently complex to take account the differences in claim characteristics between operational groups

Structural issues

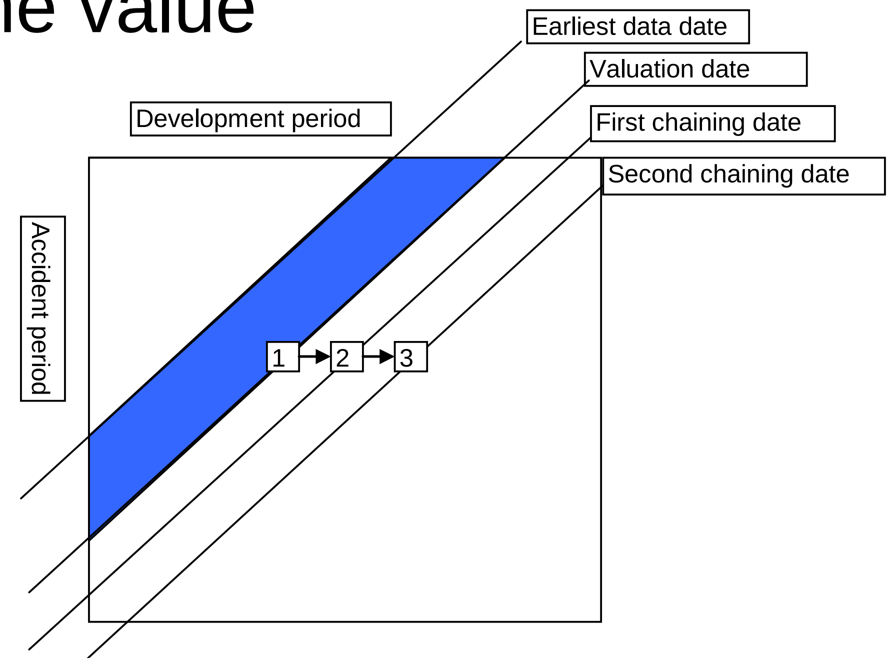
Dynamic and static predictors

Complex models – static and dynamic predictors

- Static predictors
  - Accident quarter, age
- Dynamic predictors
  - Injury severity, medical payments to date, case estimates, date of finalisation
- One of the most important decisions in any complex model is how to handle dynamic predictors

Dynamic predictors – chaining

- “Normal” actuarial process uses “chaining” to get a “lifetime value”

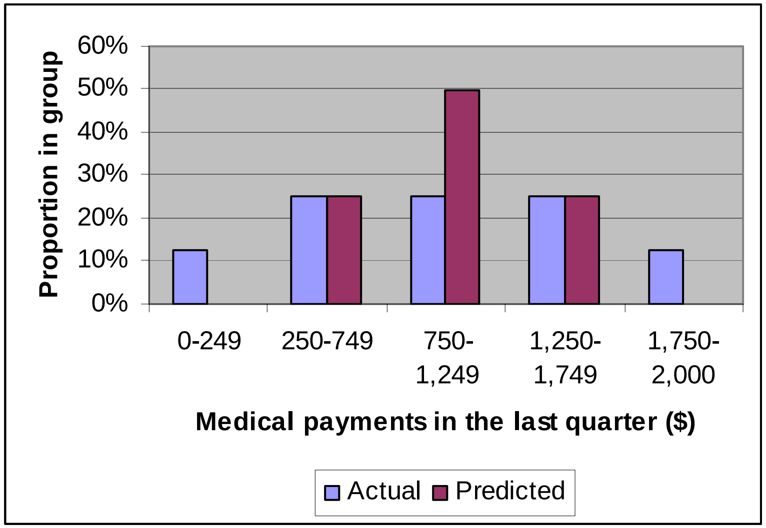


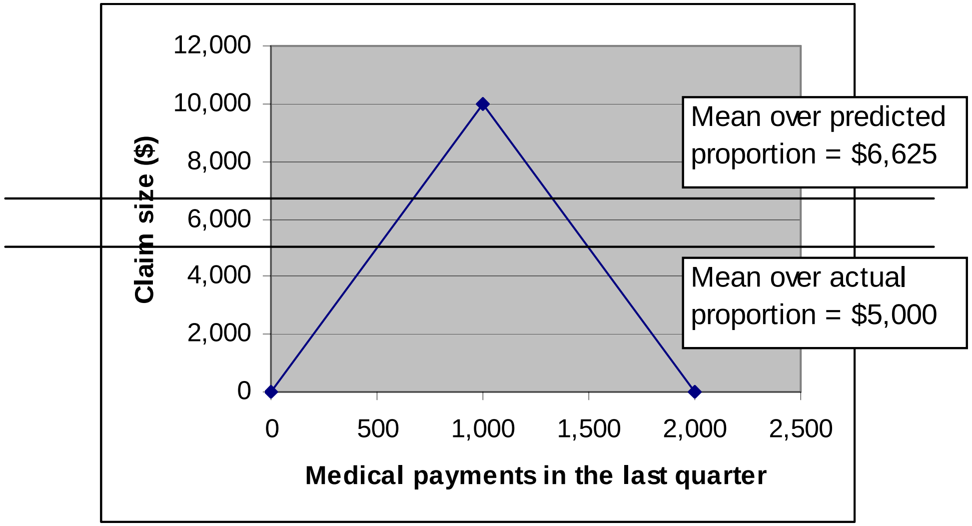
Dynamic predictors – chaining

- Need to predict the value of the dynamic predictor at each chaining point
    - Not so easy if there are several correlated dynamic predictors
    - Considerable danger in using the predicted mean for chaining
        - » Small biases can become very large
        - » Non-linearity of model causes distributional problems

Dynamic predictors – chaining

– Distributional distortion





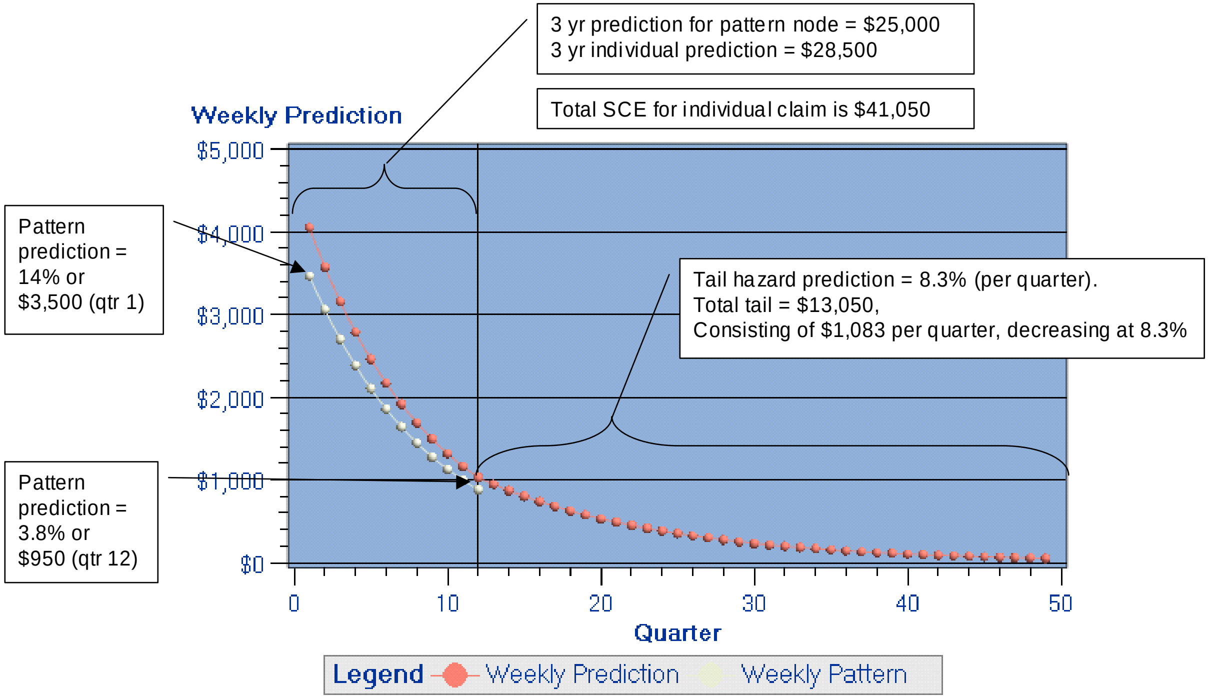
Dynamic predictors – chaining

- Problem is that the distribution of forecast predictor is different from the distribution of actual predictors used to fit the model
- To overcome can
    - Parameterise the distribution of actual predictors and simulate
    - Band the predictor and predict proportions in each band. This often ends up as a transition model
        - Potentially many states and parameters
        - Assumption of path independence

Dynamic predictors – not chaining

- – PPCF models can be thought of as transition models
- Choose long chaining period and minimise chaining iterations (maybe zero!)
  - – E.g. build complex model to predict claim payments over three years and then just extrapolate
  - – Will likely get the relativity between claims correct
  - – The extrapolation procedure is suspect but of lesser importance if the initial modeling period is long enough

Dynamic predictors – not chaining



Dynamic predictors - prediction

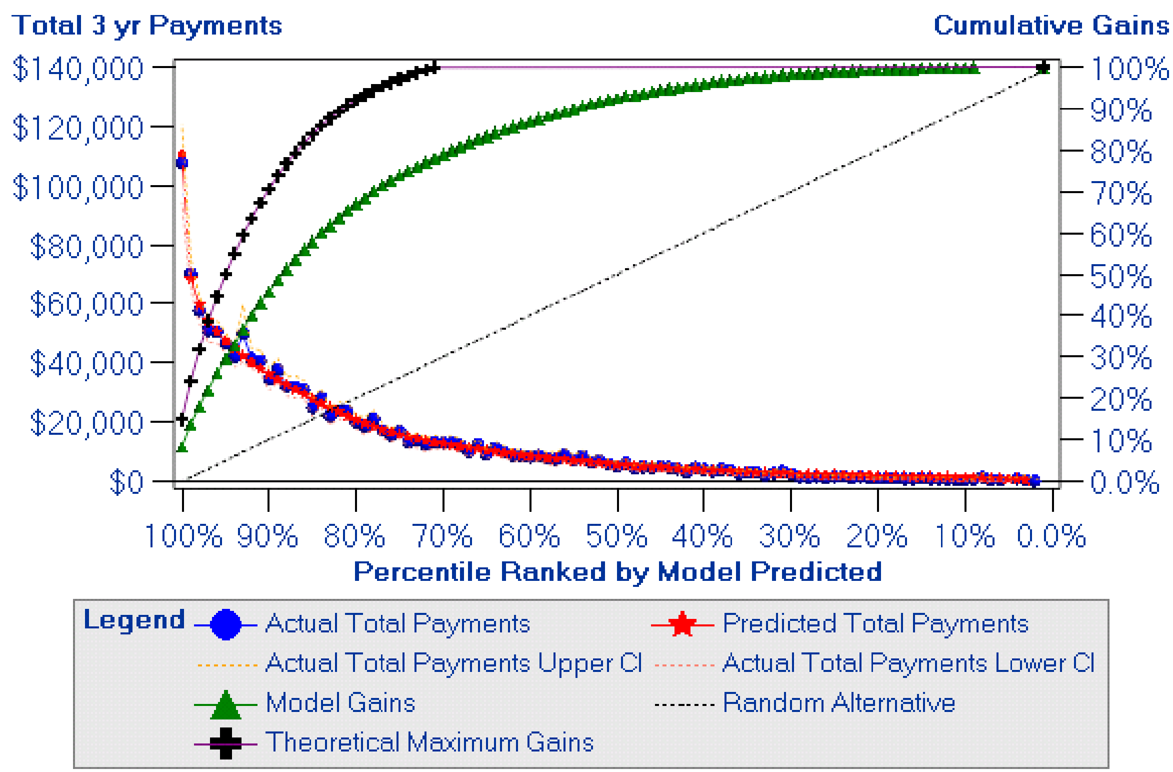
- If we do use dynamic predictors it is vital to understand the uncertainty associated with their prediction
    - Finalisation rates for PPCF
    - Compare with finalisation rates for PPCF in operational time
  - – Consider possibility of specification error as well as random error, e.g. operational delays

Model evaluation

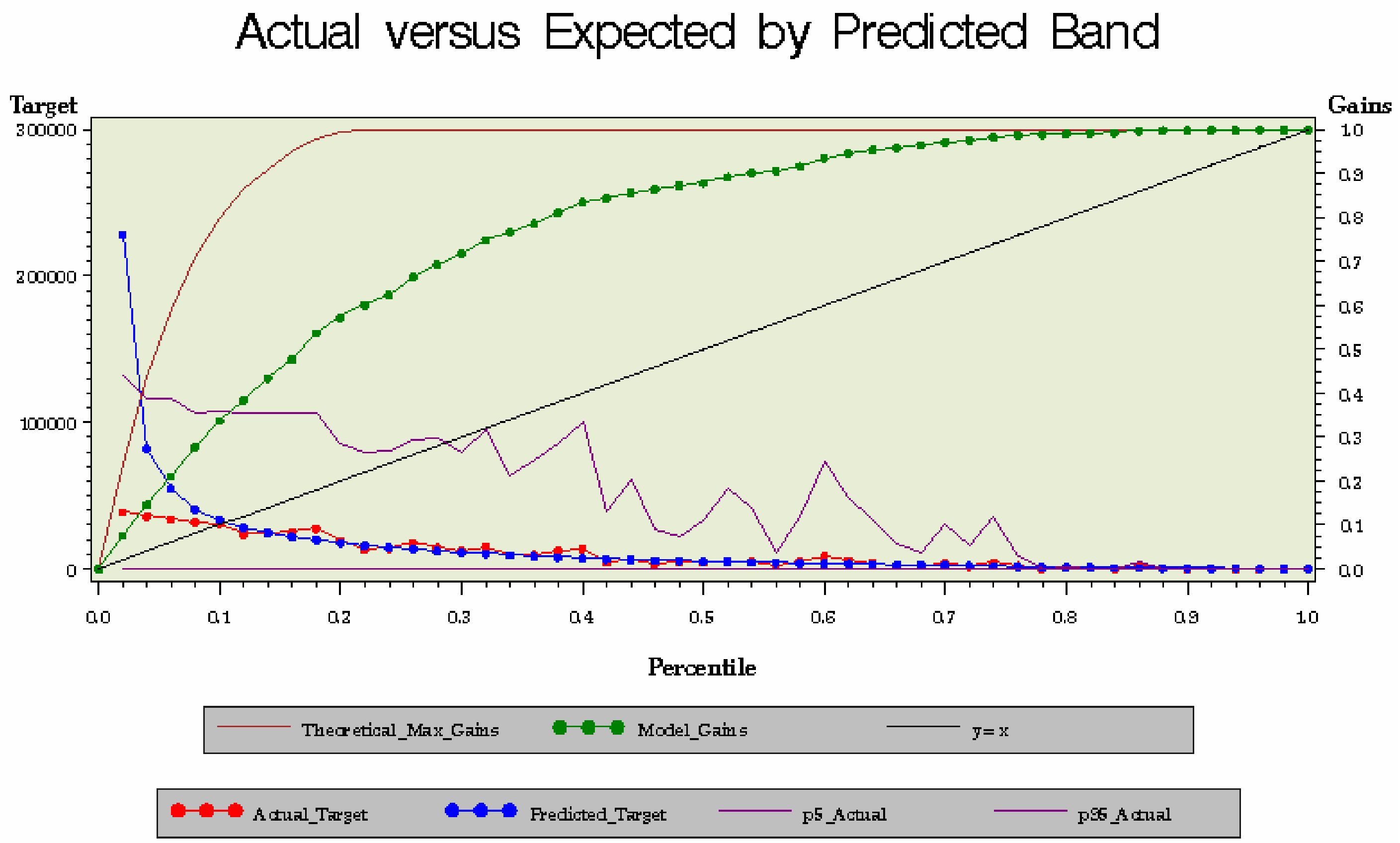
How to evaluate individual models

- Use statistical tests if fitting statistical model
  - Typically only useful for “component” models
  - Not always applicable if fitting data-mining type models
- Don’t use R^2
- Use learn/test/validation framework
  - On component models
  - On incurred cost for combined model

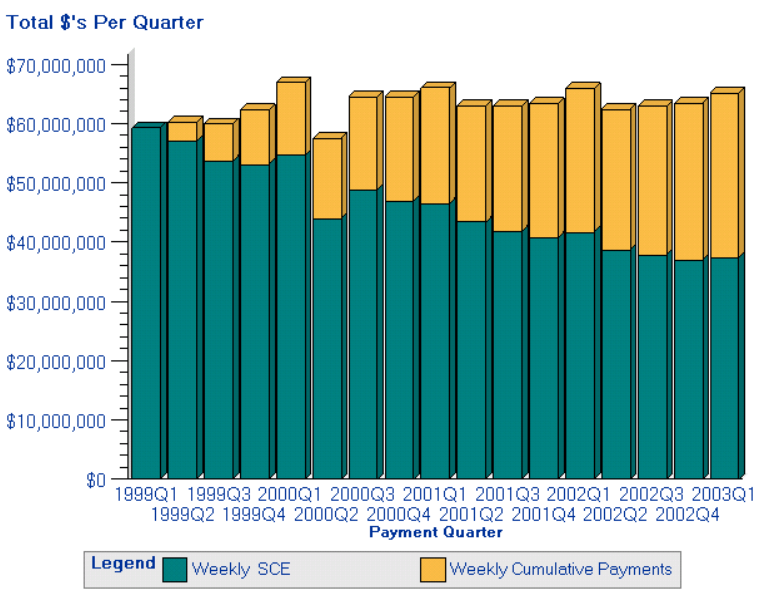
Actual vs expected for component model

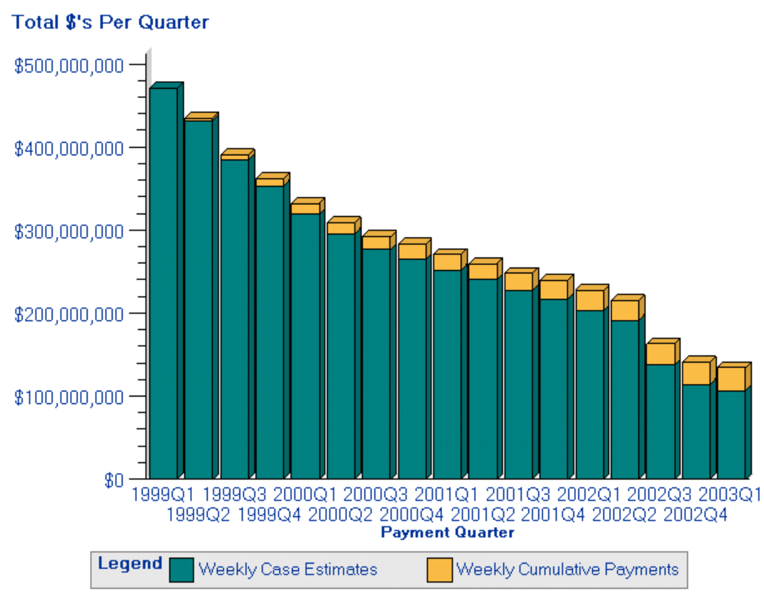


Actual vs expected – interaction bias



Incurred cost development – for claim subgroups





Automated validation

- Individual models can be very complex and evaluation is often “avoided”
- Validation tests comparing actual versus expected over past time periods, using current parameters should be standard, frequent practice
- Consider making part of “model code”

Overall conclusions

- Don’t build/use a complex model just because you can!
- If you do
  - Keep a clear idea of purpose
  - Think hard about the model structure and dynamic predictors
  - Assess the error from using dynamic predictors
  - In addition to statistical tests on component models
    - Learn/test/validation discipline
    - Incurred cost test on overall model