28 November – 1 December 2004

Xth Accident Compensation Seminar 2004

Statistical Case Estimation – An Overview of the NSW WorkCover Model

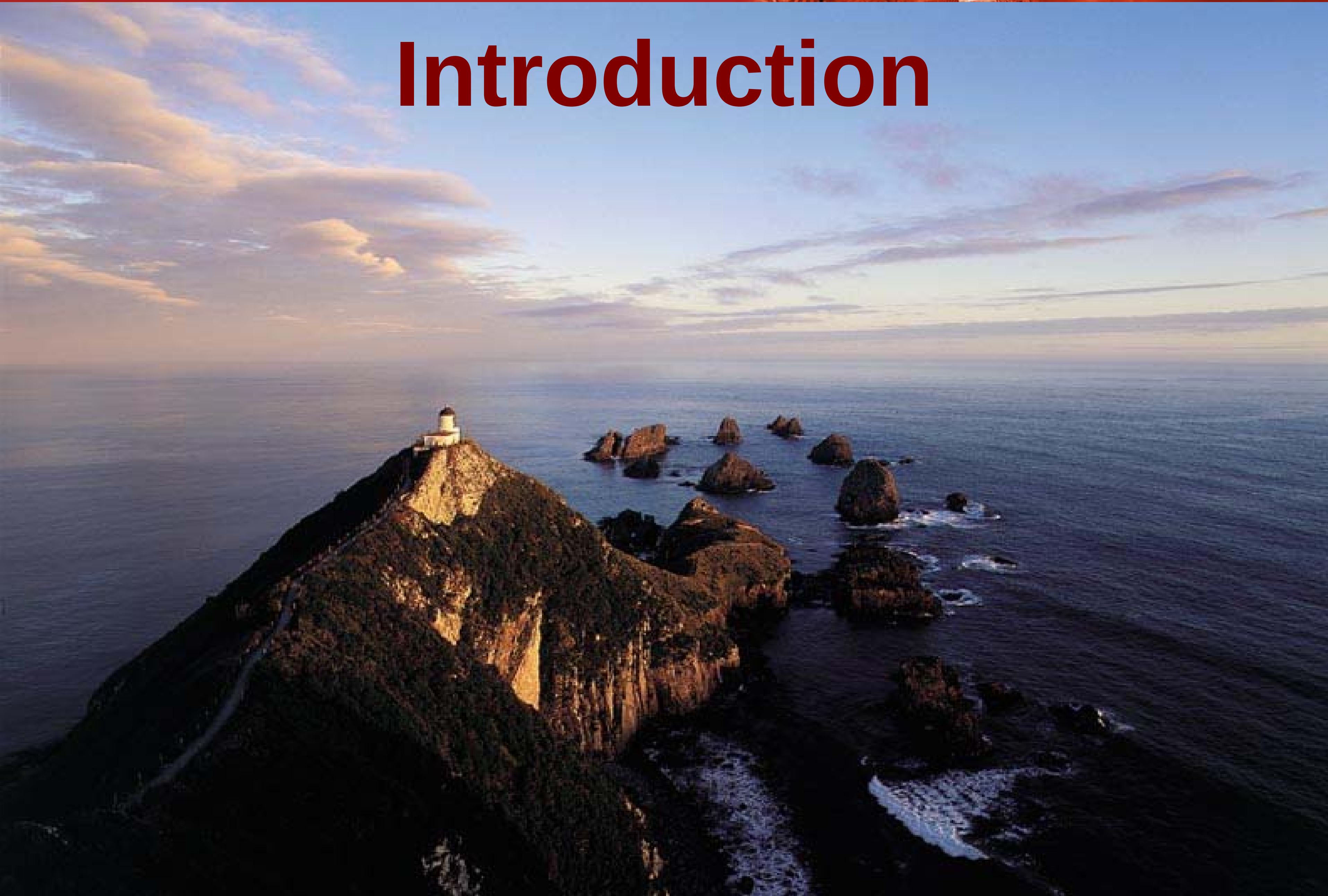
Richard Brookes and Mitchell Prevett

Institute of Actuaries of Australia

Contents

- Introduction
- Dataset and Targets
- Techniques in Modelling
- SCE Model Structure
- Model Performance
- Potential Applications and Refinements

Introduction



Introduction - What is an SCE?

Individual estimates of future claim related costs arising from existing open claims. Estimates are derived via a statistical model using the risk characteristics of the claimant:

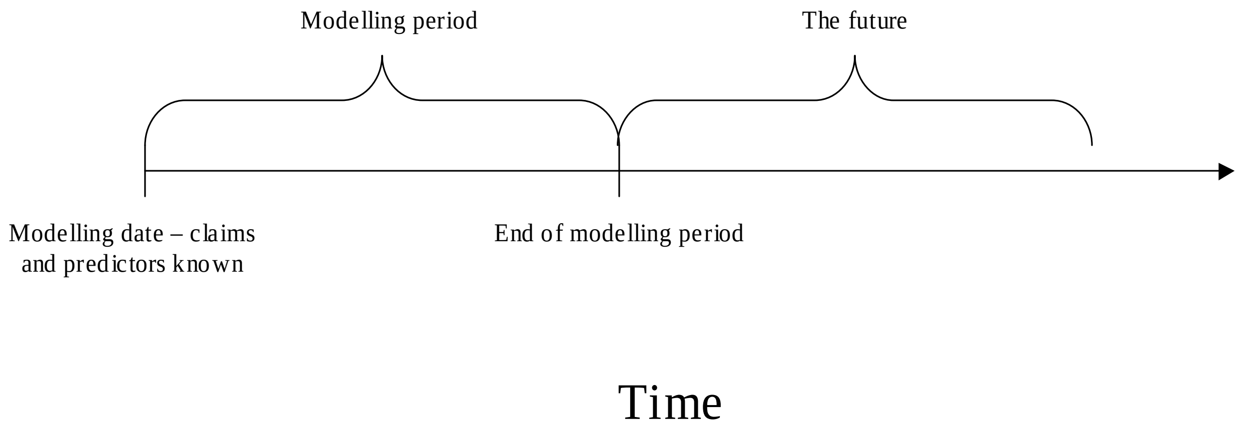
- Claimant characteristics
  - Age, gender, occupation, marital and dependant status, wage rate etc
- Employer characteristics
  - Industry, wages, location, etc
- Claim status
  - Claim is open/closed/reopened/disputed, work status, etc
- Claim characteristics
  - Injury nature, location, etc
- Claim history
  - Payments and rates of payment, time lost, etc

Introduction – What's new?

- Concept of a 3 year target on a portfolio of open claims
  - Many existing approaches model a quarterly payment or even a daily incapacity rate
- Use of data mining techniques to find structure
  - Other than some high level decisions of which payment types to model and whether to use a one or two stage model, most initial structure is determined by the data mining algorithms
- Use of dynamic predictors such as payment levels and rates
- Robust in-period testing
- Automatic and ongoing assessment of out-of-period model accuracy
- A robust and repeatable process for building SCEs

Dataset and Targets

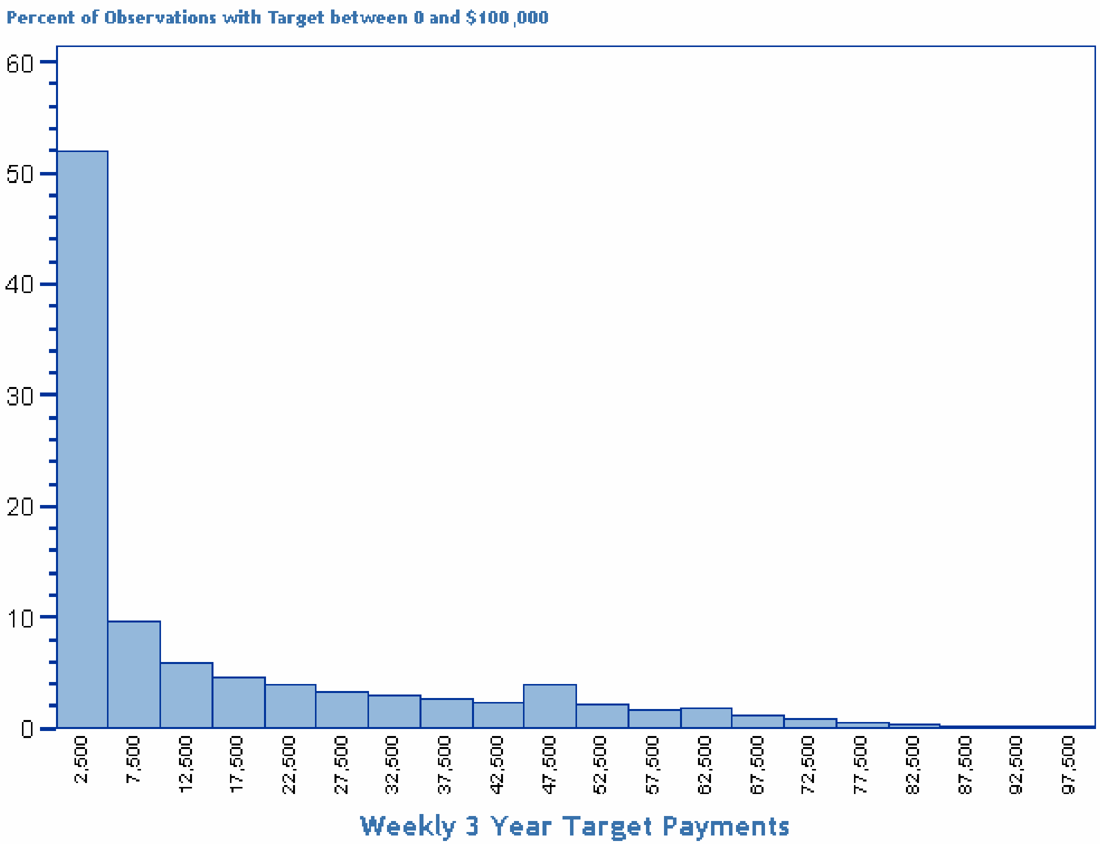
Dataset and Targets



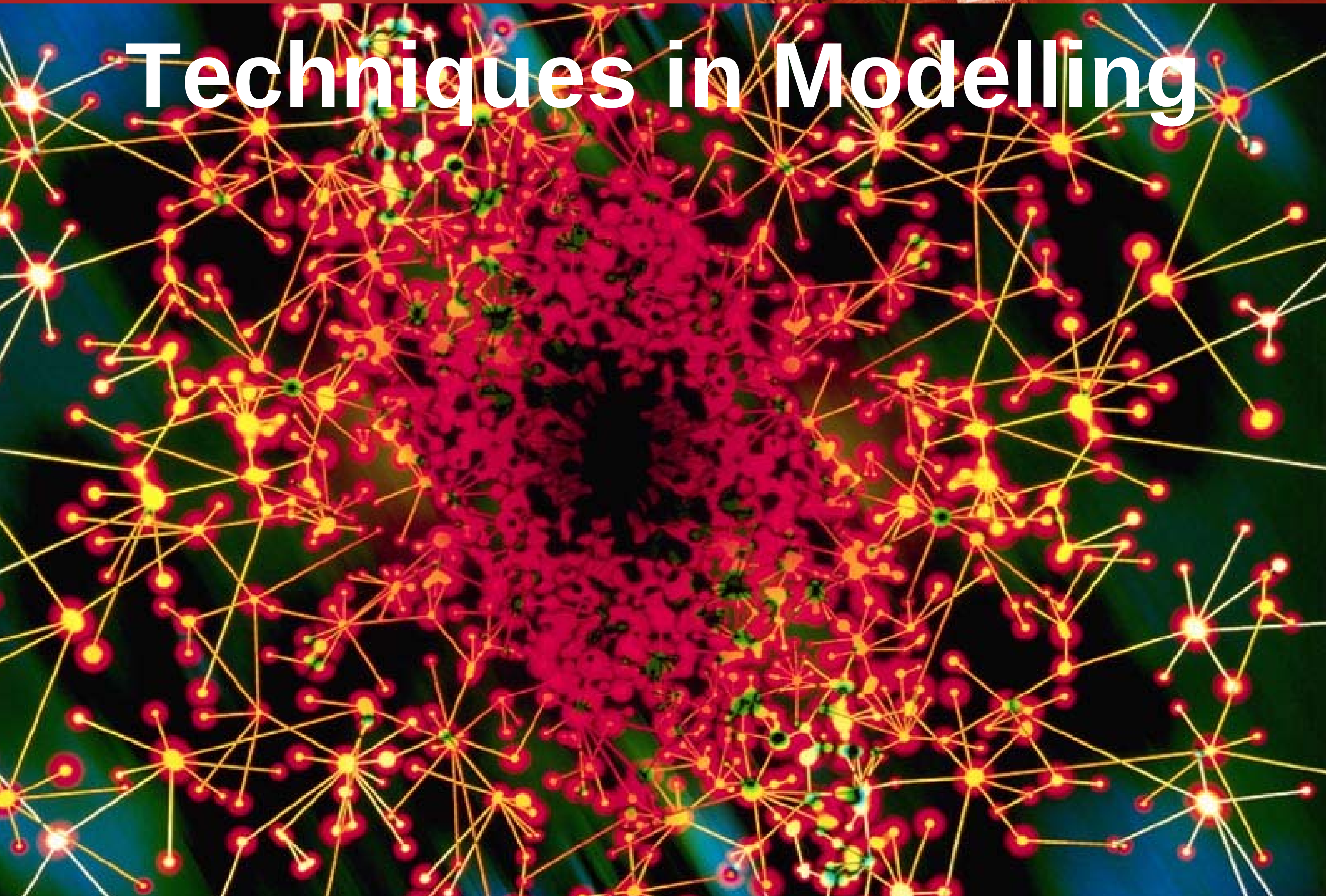
- Open claims at the start of the modelling period, split 70% learning and 30% test
  - 114,000 in the learning sample
- Approximately 200 potential predictors at the start of the modelling period
- 13 payment types to be modelled
- Modelling period of 3 years

Dataset and Targets – The Weekly Target

Number in learning dataset	114,127
Mean	5,947
Standard Deviation	15,106
Skewness	3.21
Kurtosis	11.38
Quantiles	
100% Max	200,044
99%	69,489
95%	45,289
90%	22,739
75% Q3	1,469
50% Median	0
25% Q1	0
10%	0
5%	0
1%	0
0% Min	-31,776
Percentage Negative	0.30%
Percentage Equal to Zero	61.07%
Percentage Positive	38.63%



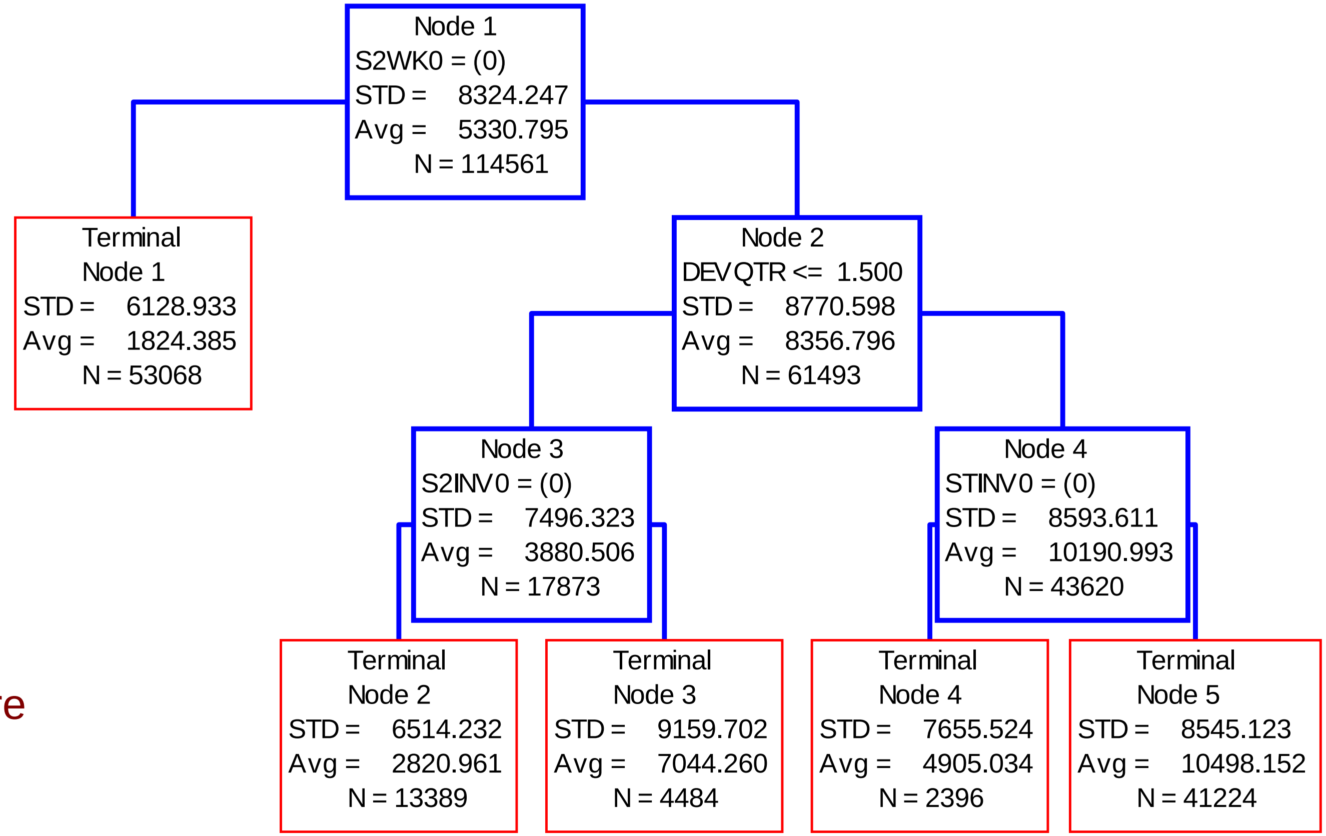
Techniques in Modelling



Techniques in Modelling - CART

Binary recursive partitioning.

- Potential splits are generated with brute force.
- Quality of splits is assessed and the best is chosen.
- Process is repeated until a large tree is grown.
- Poor performing splits on independent data are pruned.



Techniques in Modelling – MARS

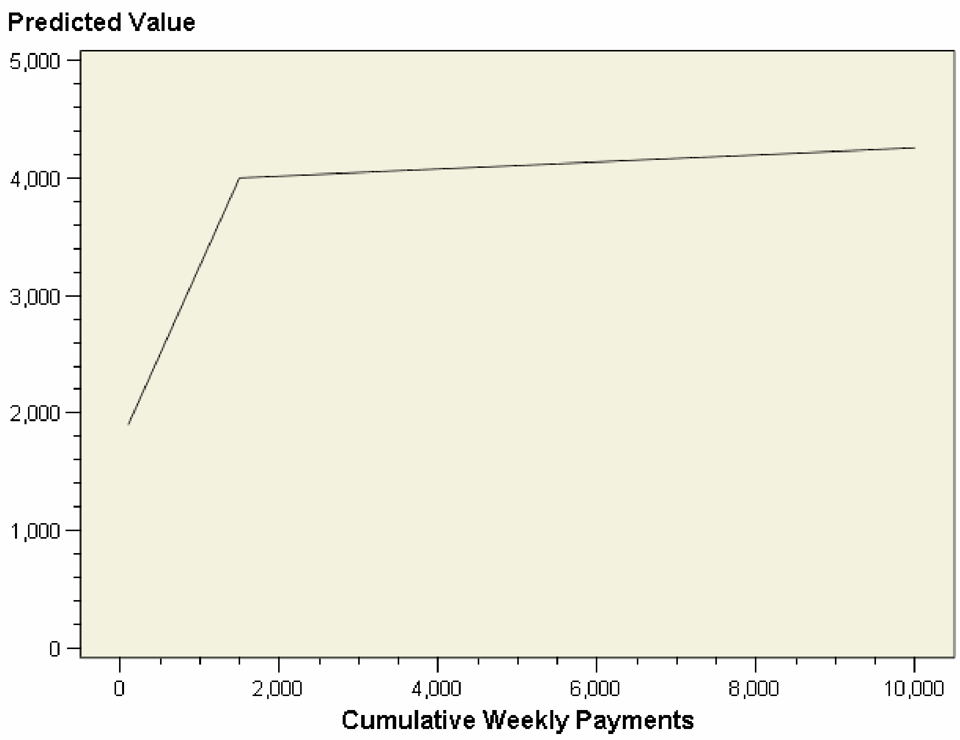
Multivariate Adaptive Regression Splines

- Builds regression models by fitting linear splines to continuous variables (via a combination of basis functions).
- Forward selection is used to select important variables and backward elimination to remove poor fitting functions.
- Example linear spline curve:

BF_1 = max(**0**, WKLYC - **1500**)

BF_2 = max(**0**, **1500** - WKLYC)

Predicted = **4000** + **0.030** * BF_1 - **1.5** * BF_2



Techniques in Modelling – Drawbacks of CART and MARS

CART

- Preference for high level categorical variables (which perform poorly).
- Lower splits in any tree are dependent on higher splits. It is likely that a single different split at the top will result in a drastically different tree.
- Quality of different splits is assessed via least squares for regression trees. As such, trees tend to focus on the higher cost observations.

MARS

- In contrast to CART, MARS is not resistant to outliers.
- Relatively poor treatment of missing values.
- Differentiation of low values can be poor.
- Over-fitting of the learning data is common.

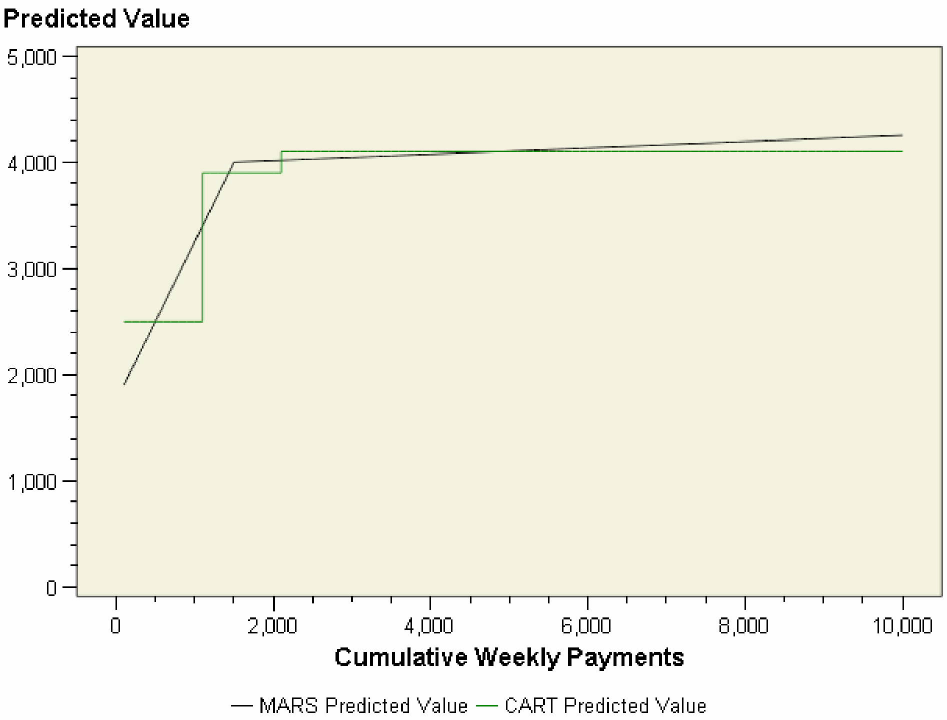
Techniques in Modelling – Hybrid CART/MARS and GLMs

CART/MARS Hybrid

- CART terminal node number is included as a categorical predictor in the MARS model.
- Relative strengths of both techniques can be utilised.
  - ✓ CART treatment of missing values, incorporation of high level interactions, and grouping of categorical variables.
  - ✓ MARS functional form fitting to continuous variables.

CART/MARS/GLM Hybrid

- The appropriate error distribution and link function are determined.
- Basis functions from MARS are refined and reduced (using Type 3 tests) where appropriate.



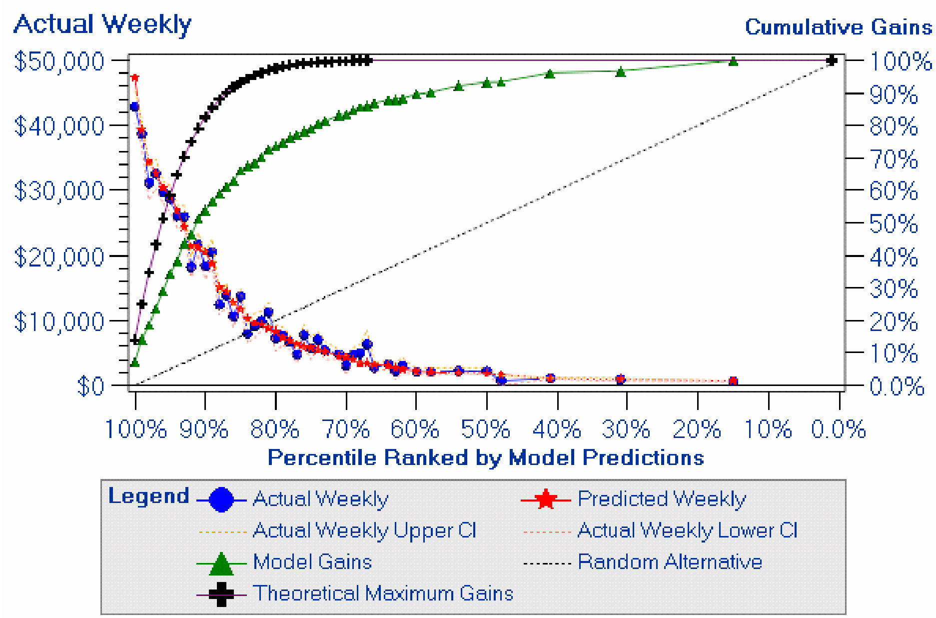
SCE Model Structure

SCE Model Structure – 3 Year Weekly Payments CART Model

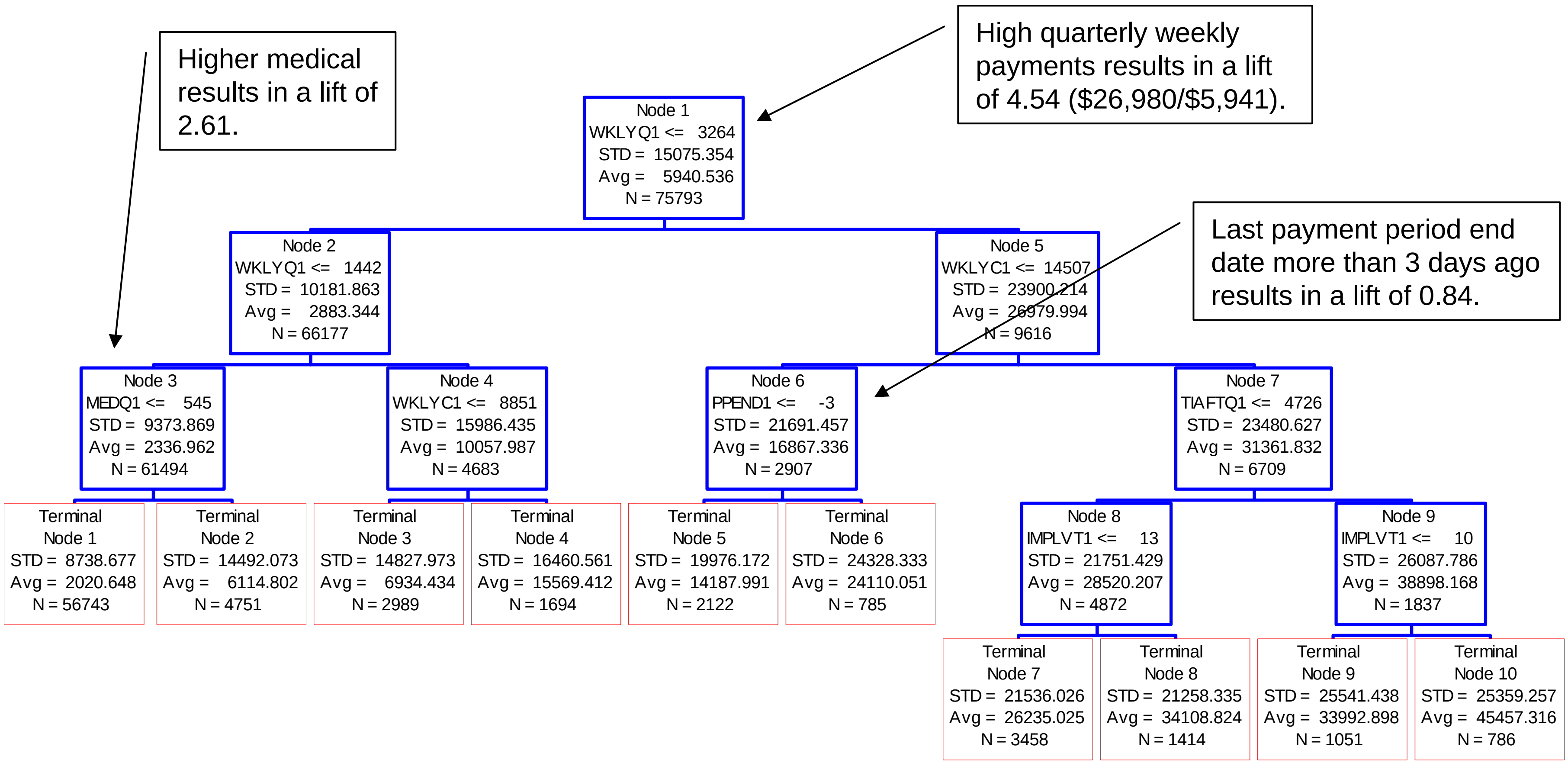
Variable	Importance
Weekly Payments Last Qtr	100
Weekly Payments Cumulative	10.34
Total Incapacity (after 26 wks) Payments Last Qtr	3.09
Impairment Level	2.48
Injury Severity Scale (Weekly)	1.63
Medical Payments Last Qtr	1.63
Days Since Initial Payment Date	1.47
Days Since Last Payment Period End Date	1.26
Weekly Case Estimate Binary	1.1
Interpreter Required Flag	0.74
Injury Location	0.65
Insurer	0.58
Investigation Case Estimate Binary	0.47
Physiotherapy Payments Cumulative	0.44
Policy Premium Experience Modifier	0.41
Rehabilitation Treatment Last Qtr	0.39
Other Payments Cumulative	0.37
Resumed Work Date Binary	0.3
Total Incapacity (first 26 wks) Payments Cumulative	0.27
Investigation Payments Last Qtr	0.27

CART Model

- 68 terminal nodes.
- Lowest cost node predicts $712.
- Highest cost node predicts $52,261.
- Cumulative weekly and last quarter weekly payments are the most important predictors.
- Impairment level based on paid section 66 benefits is an important predictor.
- No consistent bias in actual and expected.
- Top decile captures 51% of the total cost.



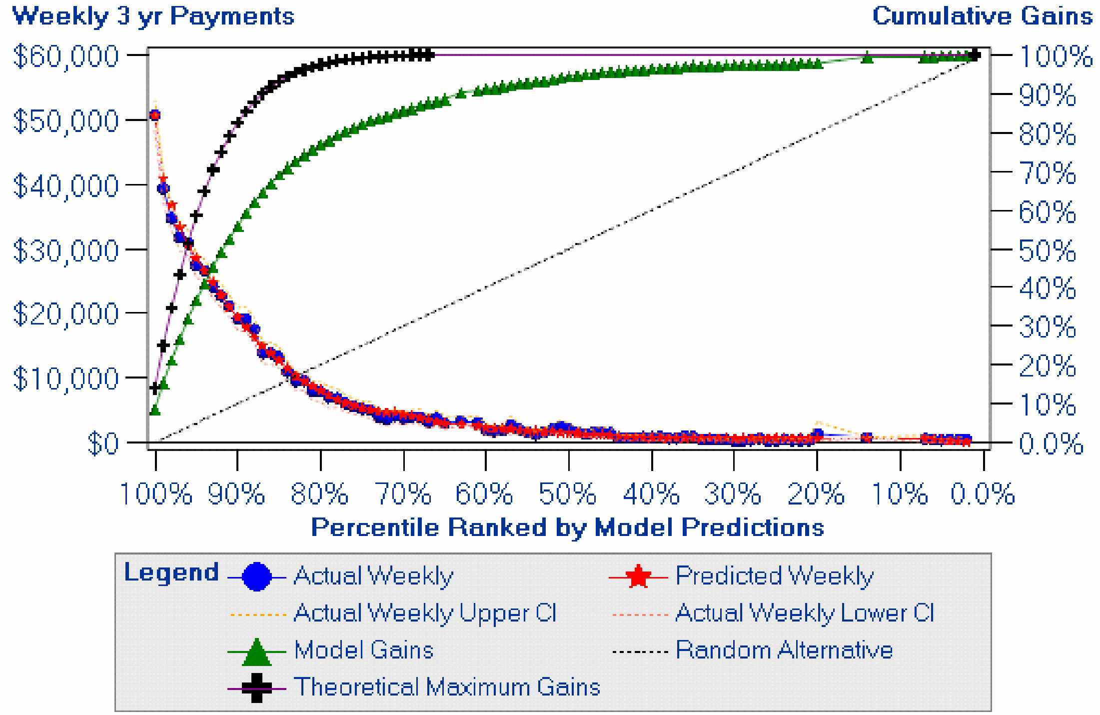
SCE Model Structure – 3 Year Weekly Payments CART Model



SCE Model Structure – 3 Year Weekly Payments MARS Model

MARS Model

- 37 final basis functions.
- Top predictions up to $50,000 (from $44,000 with CART).
- No consistent bias in actual and expected. A smoother match between actual and expected.
- Top decile captures 56% of the total cost.



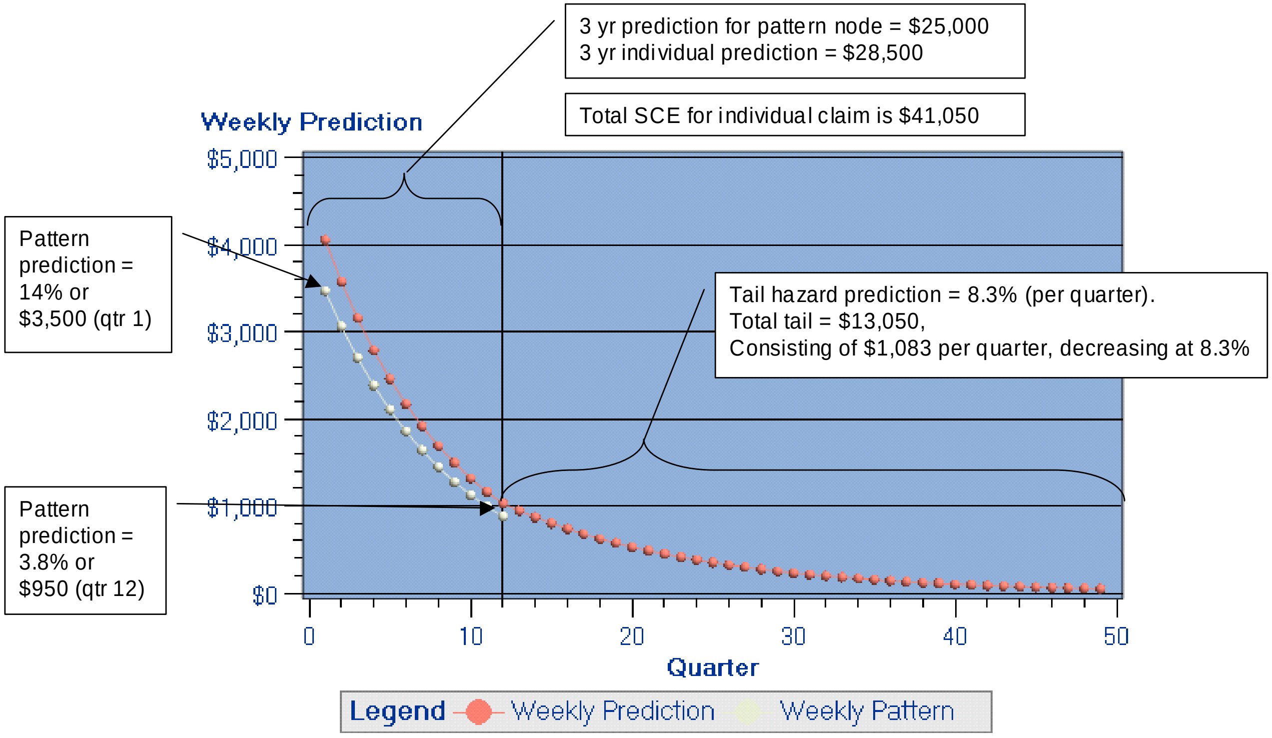
SCE Model Structure

The Modelling Aim

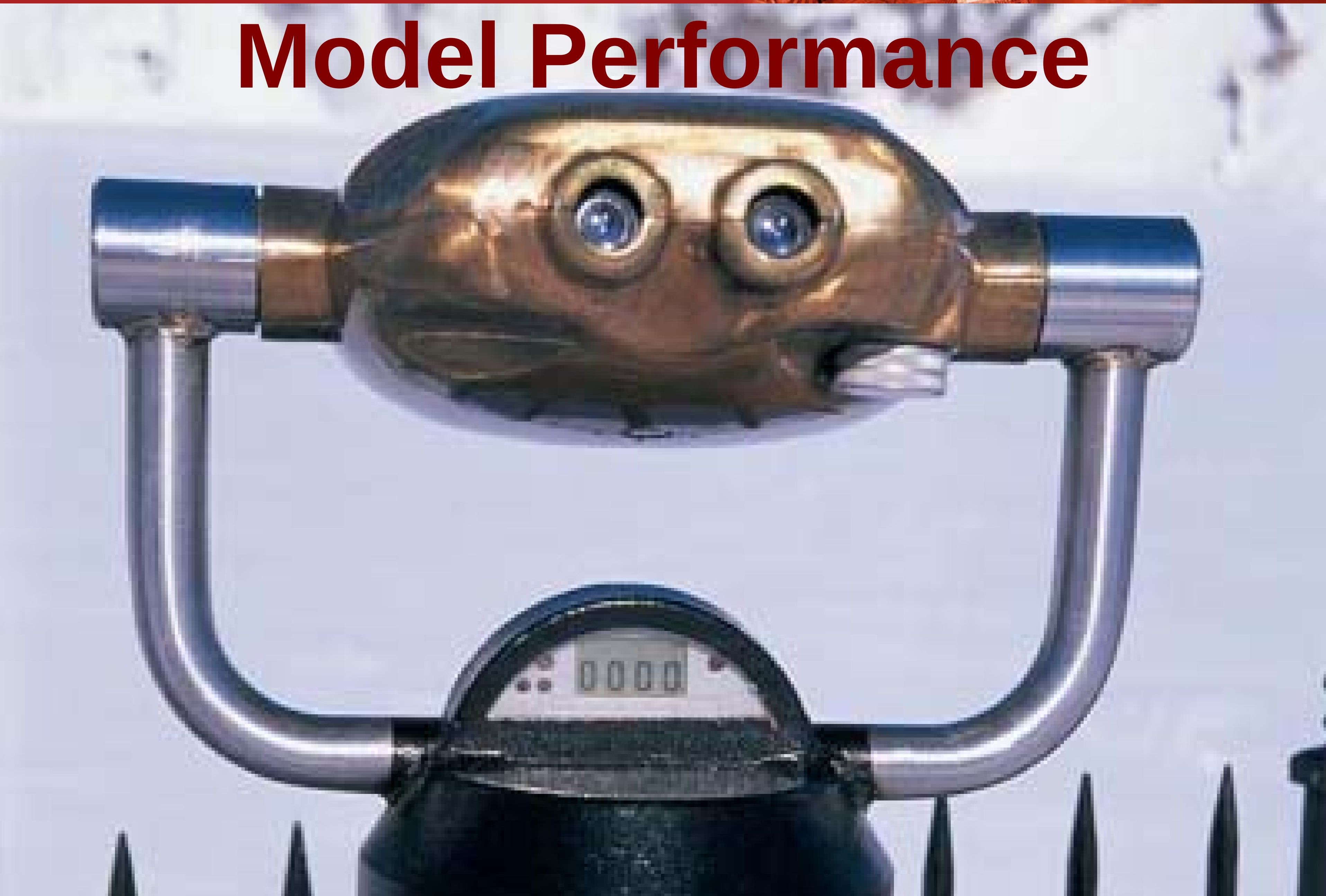
- To produce a series of expected cash flows (the salmon curve) for each claim, by payment type.
- We chose to derive this curve via 3 separate models.

Model Components

1. 3 year payments model.
2. Payment patterns for the first 12 quarters.
3. Tail hazard rate model.



Model Performance

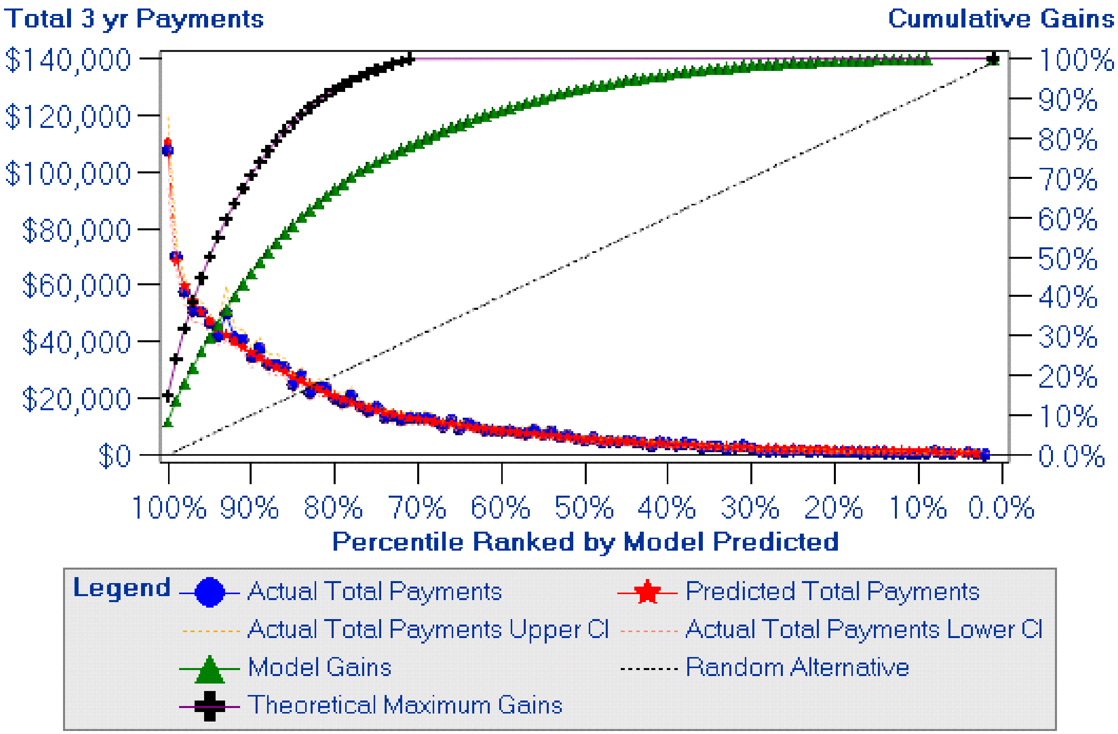


Model Performance – 3 Year Predictions

Predictions at 01 Jan 1999 on the following 3 Year Payments

- Coefficients of Variation of 2.3 and 2.1 for total and weekly.
- CV of nearly 5 for medical (greater target variability).
- Predictions in the top percentile up to $110,000.
- A good match of actual and expected.

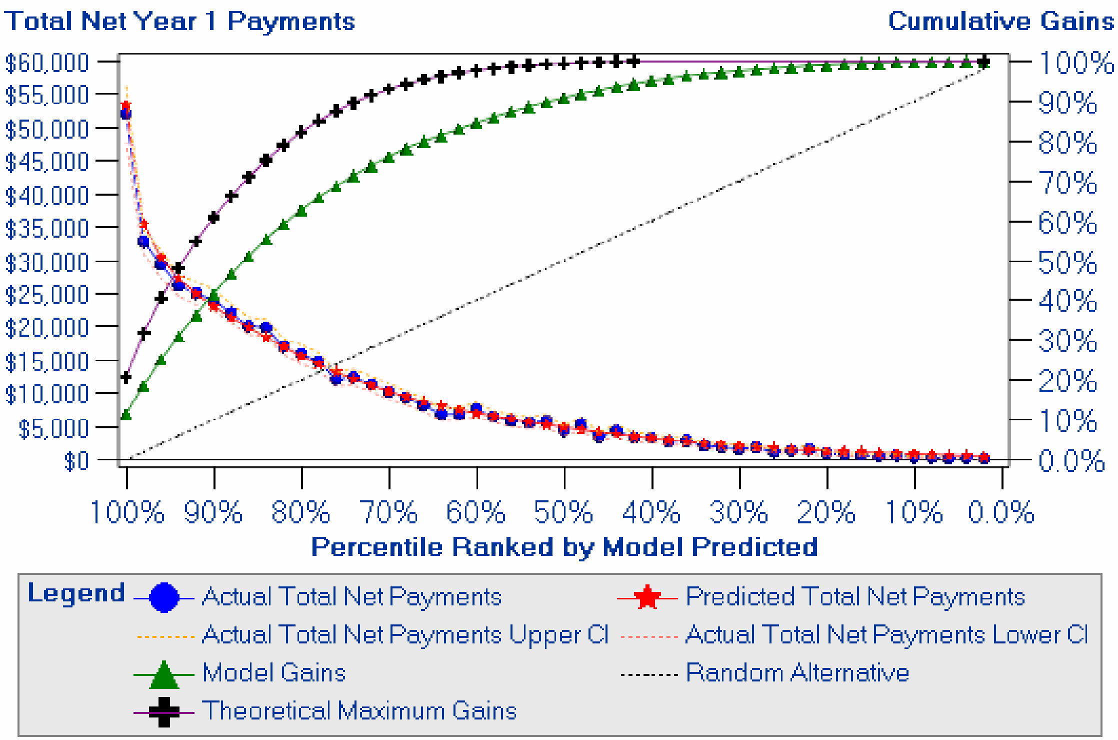
Payment Type	Payment Period	Mean Predicted	Root Average Square Error	Coefficient of Variation	R-Square
Total Net Payments	3 Years	12,913	30,071	2.33	42.7%
Weekly	3 Years	5,871	12,317	2.10	53.1%
Medical	3 Years	2,217	10,957	4.94	50.0%



Model Performance – Pattern Predictions

Predictions at 01 Jan 1999 on the following 1 Year Payments

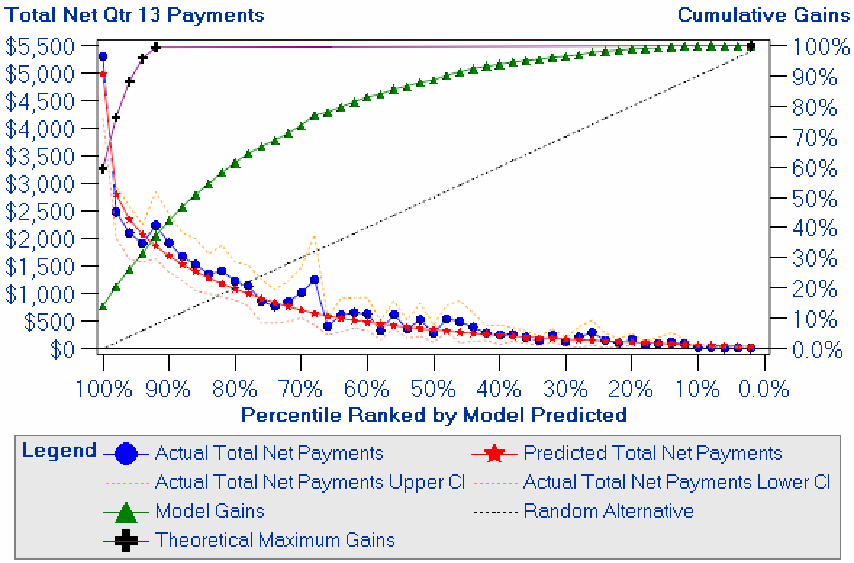
- Predictions up to almost $55,000.
- A good match of actual and expected.

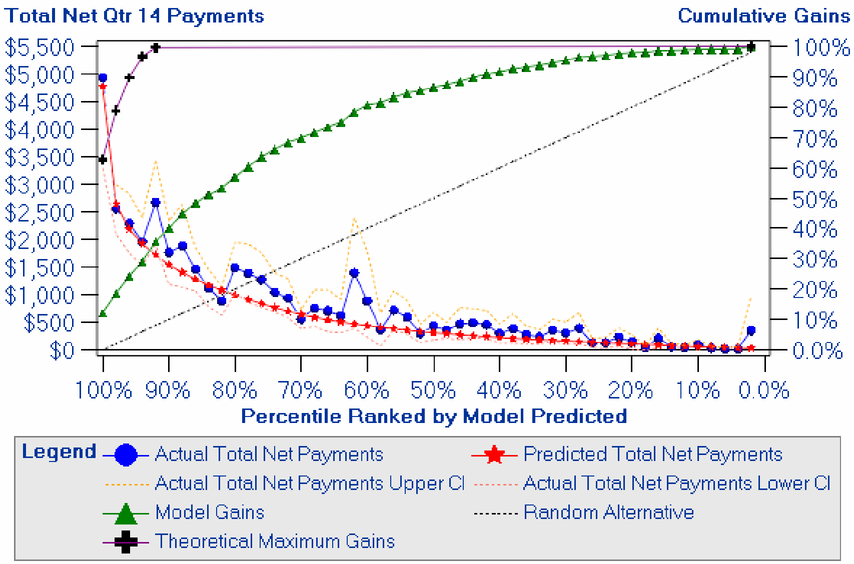


Model Performance – Tail Predictions

Predictions at 01 Jan 1999 on Payments in Projection Quarters 13 and 14.

- Greater variation in total payments exists 3 to 3.5 years after that projection date.
- Predictions are still differentiated between $5,000 and close to $0.
- Actual and expected match reasonably well across the range. Some evidence of superimposed inflation may be present.



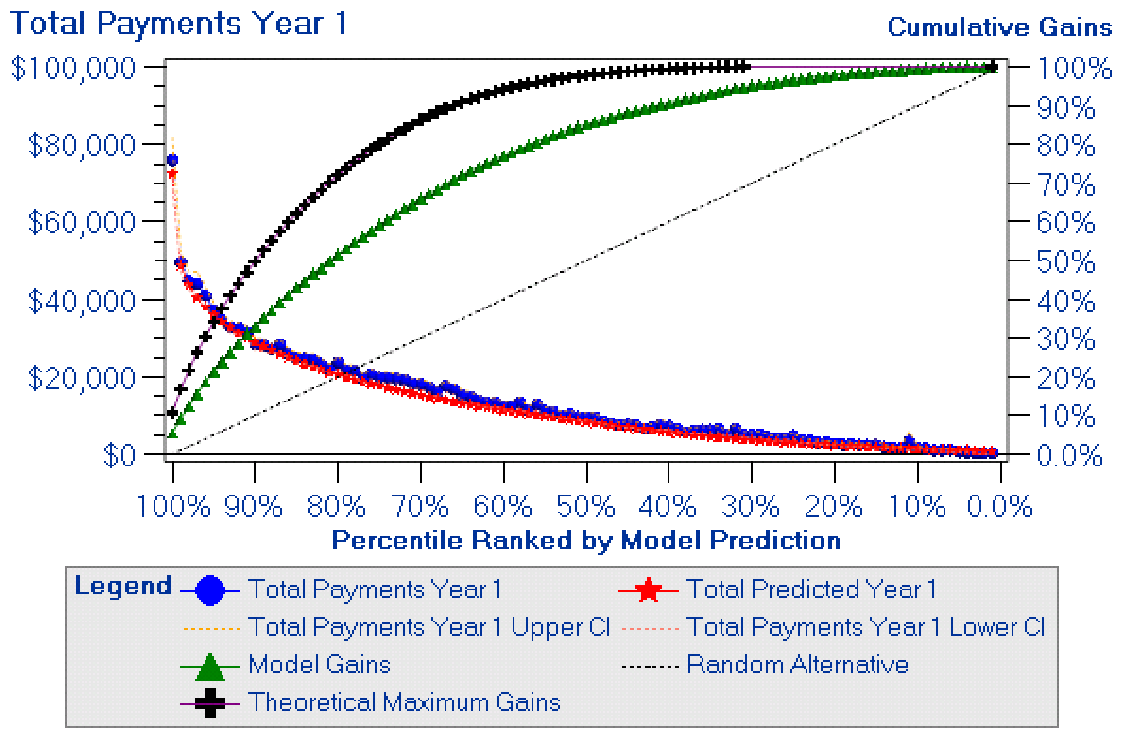


Model Performance – Recent Predictions

Predictions at 01 July 2002

- Evidence of superimposed inflation. Rehabilitation payments increased by almost 50% over this period and accelerated payments increase significantly.
- Superimposed inflation adjustments can be made for projection.
- Coefficients of Variation of 1.4 and 1.6 for total and weekly.
- CV of nearly 4 for medical (greater target variability).
- Predictions in the top percentile up to $75,000.
- A good match of actual and expected.

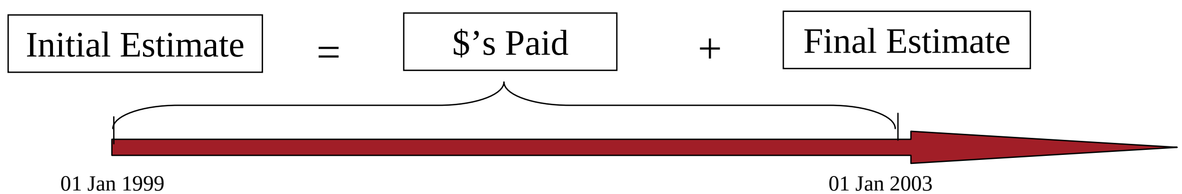
Payment Type	Payment Period	Mean Predicted	Root Average Square Error	Coefficient of Variation	R-Square
Total Net Payments	Year 1	13,664	19,503	1.43	49.4%
Weekly	Year 1	4,210	6,625	1.57	51.0%
Medical	Year 1	1,642	6,198	3.77	41.3%



Model Performance – SCEs vs Manual Case Estimates

Predictiveness

- SCEs predicted cash flows can be compared to actual cash flows. Can't do this for case estimates.
- We compared outstanding cost development.
- CVs and R-square values suggest SCEs are considerably more predictive than case estimates.

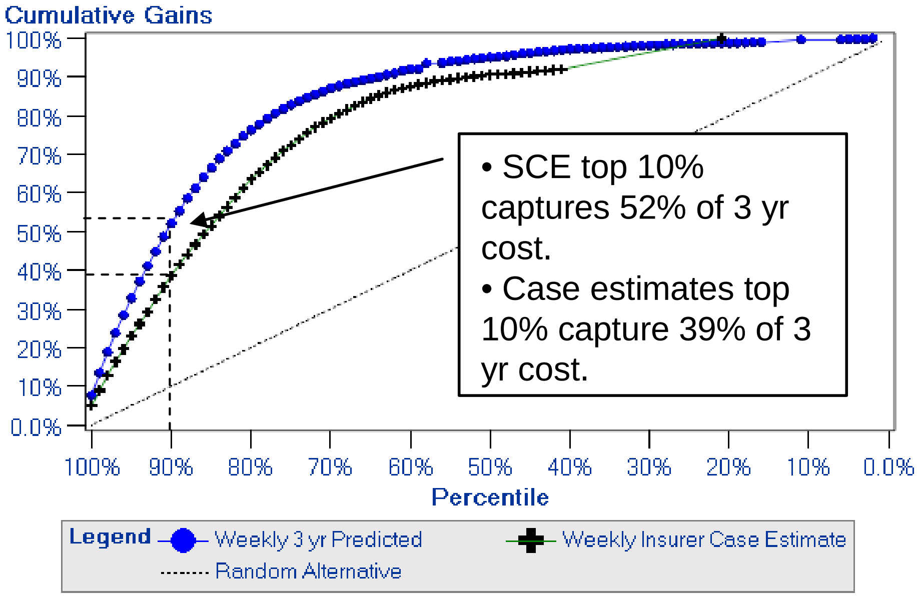


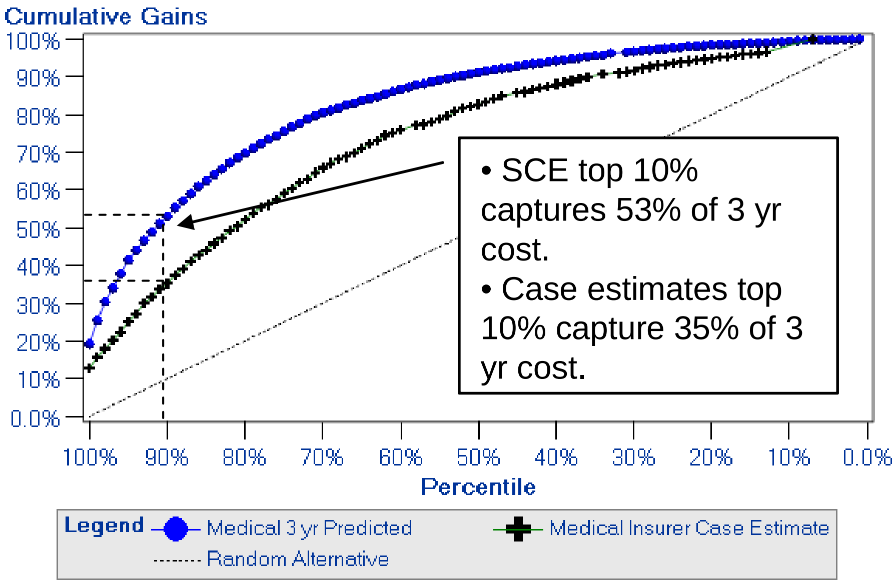
Payment Type	Estimate Type	Mean	Root Average Sq Error	Coefficient of Variation	R-square
Weekly	Case Estimate	14,712	51,865	3.53	18.9%
Weekly	SCE	8,564	21,095	2.46	49.0%
Medical	Case Estimate	4,206	36,860	8.76	28.2%
Medical	SCE	3,541	20,281	5.73	45.4%

Model Performance – SCEs vs Manual Case Estimates

Predictiveness and Ranking

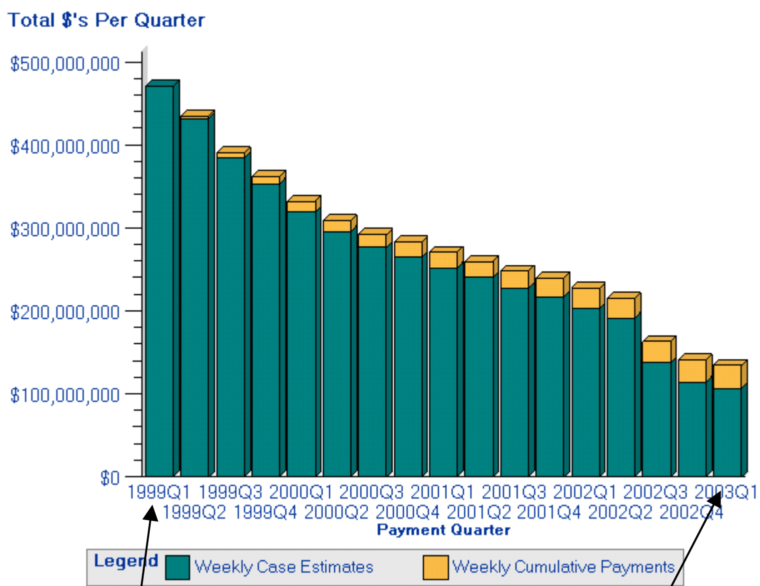
- Gains charts can be used to compare the relative ranking of claims by the 2 estimates.
- The top decile ranked by weekly SCE captures 52% of the next 3 year weekly cost while the top decile for manual case estimates only captures 39%.
- For Medical, SCEs capture 53% while case estimates capture 35%.

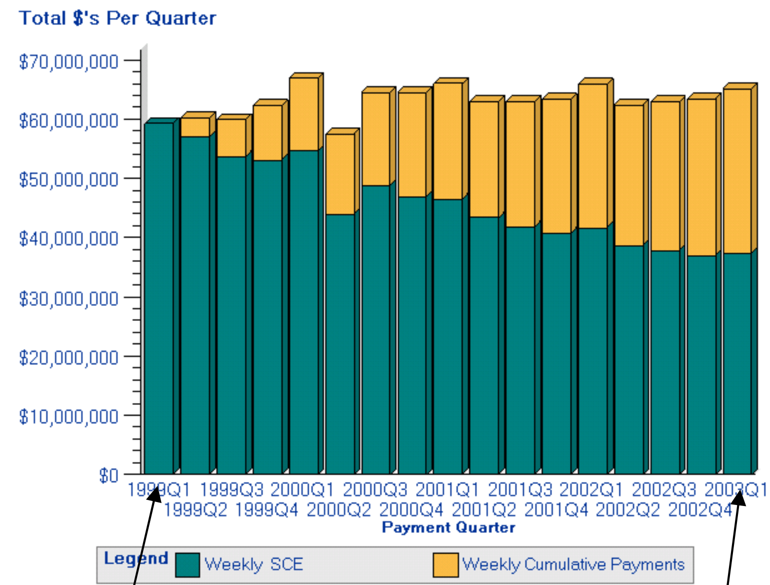




Model Performance – Development of High Value Estimates

- Top 1% of claims by weekly manual case estimate as at 01 January 1999.
- Development of case estimates (left) demonstrates significant early over estimation.
- SCEs (right) do not appreciably increase or decrease as they develop.





- $471m case estimates
- 874 open claims

- $107m case estimates
- $28m paid
- 503 open claims

- $59m SCEs
- 874 open claims

- $37m SCEs
- $28m paid
- 503 open claims
- Inflation 4% greater than anticipated over period.

What Next?

Potential Applications

- Pricing
- Benchmarking of performance
- Monitoring for sub-groups of claims
- Supporting tool for the main valuation
- Formulation of lead indicators
- Input to claim management

Potential Model Refinements

- Technical issues
  - Payment patterns
  - Tail extrapolation

Comparatively minor

- Better data
  - More robust data

More effect

- More relevant data, especially
  - Health status via standard instruments
  - Psycho-social and attitudinal factors
  - Diagnostic flags

Most significant