Institute of Actuaries of Australia

Premium Rating for Personal Injuries Using Statistical Case Estimates

Prepared by
Sandra Chapman and
Rosi Winn

Presented to the Institute of Actuaries of Australia
XVth General Insurance Seminar 16-19 October 2005

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au Website: www.actuaries.asn.au

Abstract

Statistical case estimates are individual estimates of the future claim related costs arising from existing, open claims. These estimates are predicted for each individual claim, based on its risk characteristics.

This paper discusses the building of a statistical case estimate model for the New Zealand Accident Compensation Corporation and, in particular, how the resulting model was used as an allocation tool to set premium rates for work-related personal injuries.

Keywords: General Insurance, statistical case estimate models, premium rating, work-related personal injuries

1 Introduction

One of the unique aspects of New Zealand is that, since 1974, compensation for personal injuries has been provided on a universal no-fault basis. Whereas no-fault compensation from workplace and motor vehicle accidents is not uncommon, only the Accident Compensation Corporation (ACC) Scheme in New Zealand extends this to accidents from any cause.

The funding of the Scheme is from a variety of sources and part of the challenge of managing the Scheme is balancing the interests of the various stakeholders.

In particular the cost of workplace injuries is funded by a premium on employers and self-employed persons. This paper sets out how statistical case estimates (SCEs) have been used to refine the approach used to set premium rates for workplace injuries in 2005.

The rest of this paper is separated into the following sections

- A review of the main benefits provided by the ACC scheme and discussion on their funding approach.
- A brief review of the construction of the statistical case estimate model and discussion on the predictive ability of the model
- A review of the approach used for pricing work-related personal injuries before this SCE model was introduced and how this was amended to use the new SCE model.
- A discussion of the impact of the change and some of the issues faced in implementation of the new SCE model

2 A description of the benefits provided under the ACC Scheme

2.1 Eligibility

For a person to be eligible for benefits from the ACC scheme, they must have suffered an injury for which they have cover under the Injury, Prevention, Rehabilitation and Compensation (IPRC) Act 2001.

The personal injury must be of the following kind:

- Physical injuries suffered by a person; or
- Death; or
- Mental injury suffered by a person as a result of physical injuries or a criminal sexual offence performed by another person; or
- Damage to dentures or prostheses;

and the personal injury must be the result of:

- An accident to the person; or
- Injury suffered whilst receiving medical treatment; or
- Work related gradual process, disease or infection.

An accident to the person is defined as any of the following occurrences:

- Application of force or resistance external to the human body
- Inhalation or oral ingestion of a foreign substance (excluding virus / bacteria)
- Burns or exposure to rays of any kind (excluding exposure to the elements)
- Exposure to the elements that results in incapacity for a period exceeding 1 month, or in death.

2.2 Coverage by Account

For funding purposes, ACC claims are allocated to an account, depending on
- the cause of the accident resulting in personal injury, and
- the work status of the claimant.

Benefits provided **do not** vary by account. The following flowchart shows how claims are allocated to each account:

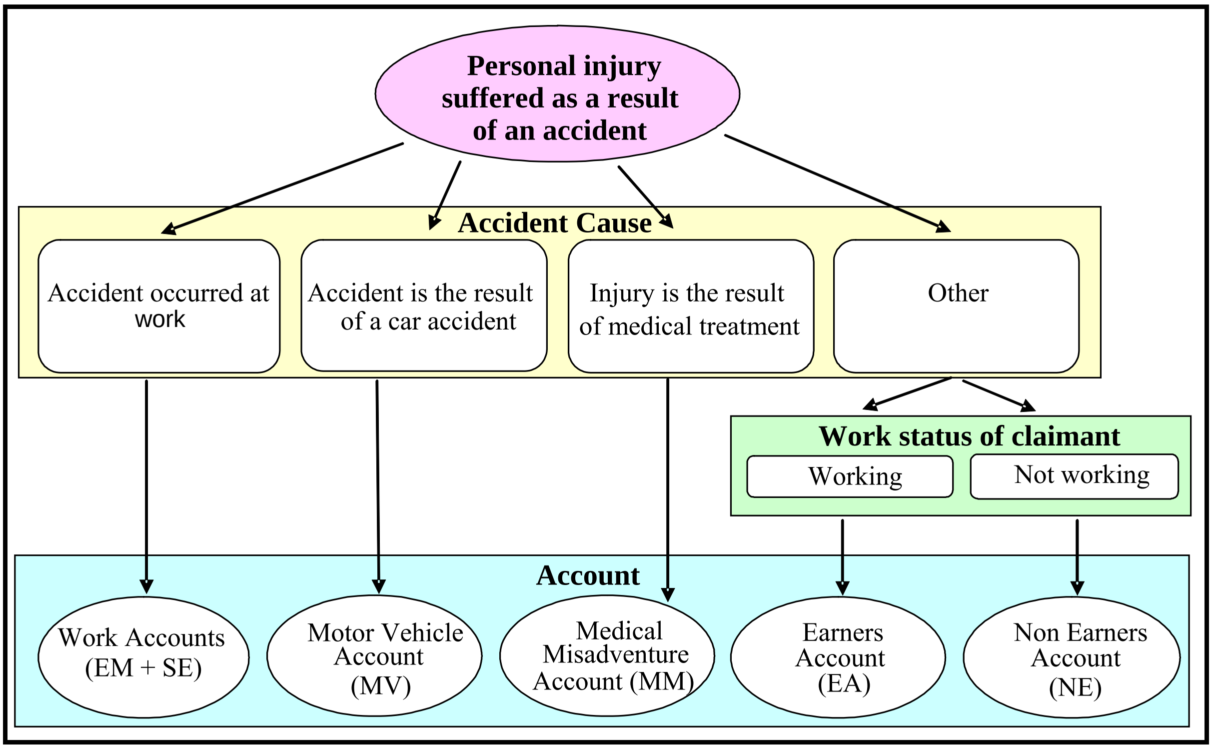

Figure 2.2.1: Allocation of claims to account

The Work Account is currently split into the Employers' Account (EM) which funds work-related claims made by employees and shareholder employees and the Self-Employed Work Account (SE) which funds work-related claims made by self-employed persons. The premiums for these two accounts are paid by employers on behalf of their employees and by self-employed persons respectively.

2.3 Benefits Provided

The following benefits are available for ACC claims:

Weekly compensation non fatal – Weekly income replacement benefits to the value of 80% of pre-injury earnings (benefit capped at NZ$77,295 per annum for year ending 30 June 2006) are payable for all persons deemed unfit to work. People who have the ability to work but are unable to find suitable work are not entitled to receive income replacement benefits. A stand down period of 1 week applies before benefits commence, which must be covered by the employer for a work injury claim.

Weekly compensation fatal – Weekly compensation is payable to the surviving spouse and dependants of a fatally injured claimant, subject to a maximum of 100% of weekly income replacement benefits to which the fatally injured person would otherwise have been entitled. This benefit is payable to the surviving spouse for a maximum of 5 years or until they cease to have dependents.

Social Rehabilitation – costs for treatment provided to assist in restoration of the claimant's independence, in particular, to restore the claimant's life skills and abilities that are essential for every day functioning. This includes provision of aids and appliances, attendant care, child care, home help, education support, home and vehicle modifications and training for independence. These benefits can be grouped under the sub-categories of capital, care, assessment and other.

Treatment – Medical, ambulance & hospital costs, vocational rehabilitation, hospital rehabilitation and other rehabilitation.

Lump Sums – Awarded to claimants who are in excess of 10% permanently impaired according to American Medical Association (AMA) guides (including mental impairment). Lump sums are awarded according to a predetermined scale based on impairment level, with a maximum of NZ$100,000. Between 1 July 1992 and 31 March 2002, lump sums were replaced by a regular quarterly independence allowance payment.

There is no facility for common law in New Zealand and weekly benefits make up a large proportion of the costs.

3 The Statistical Case Estimate (SCE) Model

3.1 What is a Statistical Case Estimate?

SCE's are individual estimates of the future claim related costs arising from existing, open claims. These estimates are predicted for each individual claim, based on its risk characteristics such as:

- 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

3.2 Why was the SCE Model built?

The SCE model is a tool for calculating the outstanding liability for individual claims. However its' main ability is to distinguish the *relativity* between small groups of claims, however the groups are defined. Our experience with SCE's is that they can react to changing experience in a volatile fashion, making it difficult to analyse and project underlying trends. The SCE model may not give an accurate picture of the total liabilities since it will not react to overall payment trends in the same way as an aggregate scheme valuation. Therefore, the SCE model operates better as a tool for *allocating* the liabilities as measured by the main valuation to various groups.

Potential applications of the SCE include:
- Setting premium rate relativities.
- Providing "drill downs" of valuation results to subgroups of claims.
- Providing estimates of ultimate claims cost by branch/case management team as an input to KPIs.
- Formulation of lead indicators of outstanding claims cost
- Providing input to claims management. The SCE will allow a greater focus on financially significant claims and ensure that management programs are in place for them.
- Providing supporting analysis for the main valuation
- Ad hoc analysis and costings.

3.3 SCE Model Construction

The SCE model was built for weekly compensation non-fatal and social rehabilitation payment types only.

We have used a data mining approach in the construction of the SCE. This approach moves away from the analysis of aggregate data and the identification of trends at an overall level. Instead the modelling process uses individual case information and builds models which are trained to predict a target variable representing the payments over the next few periods.

The basic form of the modelling is to find the characteristics in the input dataset which best predict the target variable in the output dataset. This gives a set of primary sub-models which apply for a period after the valuation date. We then separately model the pattern of payments within that period and the run-off of payments over the remaining lifetime of the claim. Once the quarterly cash flows have been predicted for the future lifetime of the claim these can be discounted back to the current date and summed to give the statistical case estimate. This part of the modelling is similar to that done for many actuarial projects.

The data mining approach requires that the modelling dataset be randomly partitioned into two groups of "learn" and "test". Models are built using the learn and then evaluated on the test data. This is needed because any modelling process may produce some parameters that fit to random features of the data used to parameterise them. This makes the model less useful when it is applied to a different dataset. By evaluating the models on an independent part of the data we select those which are better general models.

3.3.1 Characteristics of a good SCE model

There are several fundamental requirements of a good SCE model:

- A good differentiation between high cost claims and lower cost claims
- A good ranking of claims from high cost to low cost
- A close "fit" so that in any subgroup of claims the average SCE is close to the actual, eventual mean cost of those claims.

There are various methods for evaluating these properties, such as gains charts and actual versus expected tables and charts. We have used these techniques as well as comparisons to the aggregate valuation to assess the accuracy and reasonableness of the models.

3.4 Data

3.4.1 Modelling Period

In order to construct the SCE model using a data mining approach we required information on the cases to be modelled as at a given point in time (modelling date) and a target variable of the payments over the next few periods to model for those cases.

We selected a modelling date of 30 June 2001 which allowed us a three year modelling period. The data used to construct the SCE model consists of:

- Information for each case as at 30 June 2001.
- The quarterly payments for each payment type for each case over the 3 year period from 1 July 2001 to 30 June 2004

3.4.2 Which cases were included in the modelling dataset?

The starting point for the cases to include in the modelling dataset is all cases which were open at the modelling date of 30 June 2001. The ACC staff suggested that the open/closed flag is not robust for some cases and indeed is not populated until claims are fully registered. As a result of this we also considered cases which had payment activity during the 3 month period leading up to the modelling date.

3.4.3 Data Sources and Variables

An SCE model makes use of a large amount of information for each individual case. The use of the data mining approach means that there is little reason to restrict the input data as the modelling process will identify those variables which are of value as predictors. The variables considered for inclusion in the dataset included identifier and possible predictor variables for each case. The possible predictors included variables covering: claimant characteristics, employer characteristics, claim status, claim characteristics and claim payment history.

3.5 Modelling

3.5.1 Overview

The four main stages in the modelling are:

1. Decision Tree Modelling
2. Regression Modelling
3. Pattern Fitting
4. Tail Modelling

The decision tree and regression modelling are used to build models to predict the expected payments over a three year time horizon.

The payment pattern stage then allocates the cost across the individual quarters within this period and the tail modelling extends the predictions to the full lifetime of the claim. The actual benefit structure is overlayed on to modelled results by making explicit adjustments for mortality and retirement as the individual claimants get older.

The final stage is to calculate the inflated discounted case estimates by combining the projected cash flows with the economic assumptions from the main scheme valuation.

3.5.2 Decision Trees

Decision trees were used as the first stage in the modelling. In our experience these provide a robust tool which is relatively simple to use and often achieve good results in terms of capturing most of the predictive ability of the input variables.

Decision trees assign cases with similar characteristics into ‘nodes’ based on a series of decision ‘rules’. This is done by selecting the best ‘splitter’ variable at each stage, partitioning the dataset using that splitter and repeating.

The advantage of this approach is given a large range of inputs, decision trees narrow down the important cost drivers and give an indication of their predictive ability.

3.5.3 Regression Modelling

A hybrid regression model was used to refine the decision trees and produce the final 3-year payment models. This approach allows functional relationships for some of the key predictors to be combined with the decision tree results.

Expressing this all mathematically:

Decision trees approach

$$y = \beta_0 + \beta_1 NODE_1 + \beta_2 NODE_2 + \ldots \beta_K NODE_K$$

Where NODEi is a dummy variable for the ith decision tree node

Regression approach

$$y = \beta_0 + \beta_{R1} X_1 + \beta_{R2} X_2 + \beta_{R3} X_3 + \ldots \beta_R X_R + \varepsilon$$

Where Xi is the ith predictor variable or transformation of the ith predictor variable

Hybrid approach

$$y = \beta_0 + \beta_1 NODE_1 + \beta_2 NODE_2 + \ldots + \beta_R NODE_R + \beta_{R1} X_1 + \beta_{R2} X_2 + \beta_{R3} X_3 + \ldots \beta_R X_R + \varepsilon$$

- The first part of this formula replicates the form of the decision tree results (note the parameters will shift as other variables are added to the model)
- Further variables added as main effects will capture effects common across all nodes
- Common effects are undetectable within terminal nodes (i.e. extending the tree) because the effect within each terminal node may be masked by noise. They become apparent when looking across nodes
- Effects detected across terminal nodes are likely to be weak as the strong effects will already be captured by the tree. However a number of weak effects can be very significant

The hybrid regressions smooth the distribution of predictions from a series of steps to a continuous distribution. This provides the potential for cases to have predicted values closer to the actual values. It also allows for a better differentiation of the higher cost claims.

3.5.4 Payment Patterns

The payment pattern stage of the modelling develops patterns which can be applied to the predicted three year targets to produce quarterly cash flows for the three year period.
This step is needed as a forerunner to extending the modelling to the tail and because the cash flows are needed in order to apply the appropriate inflation and discounting factors to get the final statistical case estimates.

3.5.5 Tail Modelling

The final stage in the modelling is the prediction of the tail liability. This is done by projecting the cash flows beyond the initial three year model period. A survival model approach was used to project the future cash flows. We analysed the hazard rate over the payment pattern period and used this to project into the tail:

Hazard Rate Theory and Application

The hazard rate reflects the rate of fall off of the quarterly payments in the tail. Thus the payments (*PQtr*) can be calculated from those in the previous quarter as follows:

$$PQtr_{(t)} = PQtr_{(t-1)} \exp(-\lambda)$$

Where λ is the hazard rate.

This has been generalised by varying the hazard rates for the different payment pattern groups and by duration. We have also added in the effects of mortality and retirement on a case by case basis.

3.5.6 The Final Model

Our final model produces a series of expected cash flows for each claim, by payment type. This has been derived via 3 separate model components:

- A 3 year payments model.
- Payment patterns for the first 12 quarters.
- Tail hazard rate model

The final inflated discounted SCEs were calculated from the cash flows modelled by the data mining process, combined with the economic assumptions used for the main scheme valuation.

4 SCE Construction Example: Weekly Compensation

4.1 Data

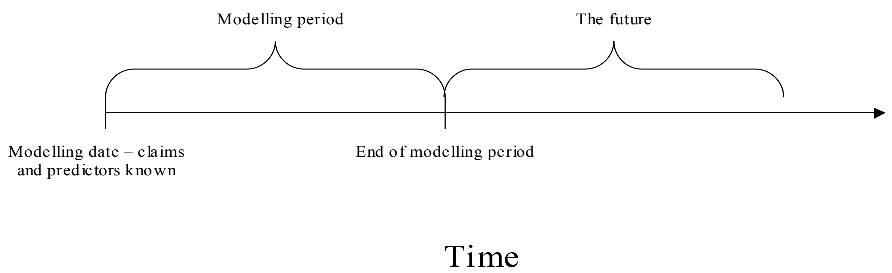

Figure 4.1.1 Illustration of Modelling Data

The model was based on a selected claim "universe" of open and recently active closed claims with the data extracted from the systems to reflect the known characteristics at 30 June 2001 and the experience for each claim over the three years to 30 June 2004.

This dataset included 644,000 claims; these were split into two groups with 70% of the data used to build the model (the "learn" data) and the remaining 30% used to independently test the models (the "test" data). This dataset included approximately 280 potential predictor variables.

4.1.1 Weekly Compensation Targets

The targets for weekly compensation represent the total amounts paid over the three year period from 1 July 2001 to 30 June 2004. The distribution of this target variable is shown in the figure below. The table shows the full distribution of the target values; the graph shows a distribution for the cases where the target is greater than zero.

Number in full dataset	644,148
Mean	1,766
Standard Deviation	12,683
Skewness	13
Kurtosis	460
Quantiles	
100% Max	1,319,110
99%	68,315
95%	559
90%	0
75% Q3	0
50% Median	0
25% Q1	0
10%	0
5%	0
1%	0
0% Min	-321,826
Percentage Negative	0.1%
Percentage Equal Zero	94.4%
Percentage Positive	5.5%

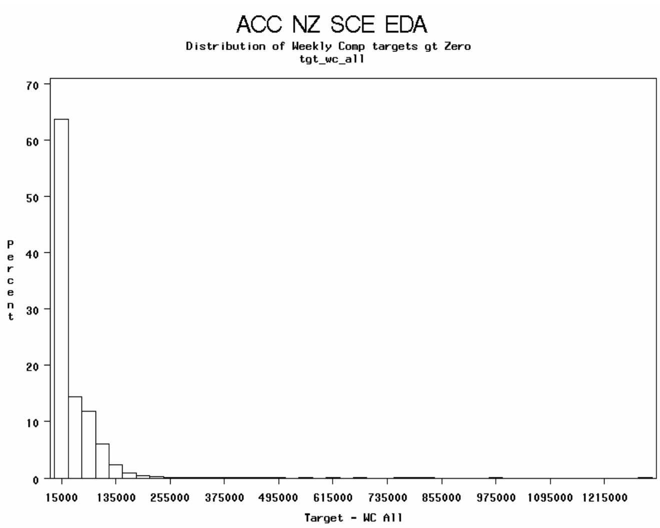

Figure 4.1.2 Distribution of weekly compensation targets

As one can see the distribution is extremely skewed with the large majority of cases having no future payment but some individual cases having payments over $1,000,000. The average 3 year weekly payment is $1,766.

4.1.2 Data segmentation

Further investigation of the targets revealed that the claims population could be divided into two distinct groups with quite different characteristics. The ACC Scheme receives a large number of claims which are dealt with very quickly without being fully registered on the systems. These are labelled non-escalated claims. The remaining claims have a much higher average cost and remain active for longer. The characteristics of the two groups are shown in the table below.

Figure 4.1.3 Characteristics of escalated and non-escalated claims

	Non-Escalated Cases	Other Cases
Number of Cases	287,000	357,000
Average 3 year target	$21	$3,075

Furthermore, the predictor variable information that is available for the non-escalated cases is a significantly reduced subset of that available for the other cases. Based on these differences the claims universe was segmented into these two groups for modelling.

4.2 Modelling

This section discusses the models built for the escalated group of claims. This group captures the majority of the total cost. A similar approach was used to construct models for the non-escalated claims although the final models adopted for this group were relatively simple since they are less material.

4.2.1 Decision Trees

The decision tree model built for weekly compensation had 113 terminal nodes. There is a considerable differentiation in the predicted future payment with the lowest cost node predicting $26 and the highest cost node predicting $159,066.

Key Predictors

The most important predictors are shown in the table below, together with the number of decision tree splits that are based on each of these predictors.

Figure 4.2.1 Weekly Compensation Predictors

Variable	Importance	Number of Rules
Cumulative Total Non-Fatal WC payments	100	4
Normal WC payments paid to date	99.7	3
Cumulative Total payments	97.6	10
Total Non-Fatal WC payments last qtr	95.0	3
Normal WC payments last qtr	94.6	1
Incapacity days between start and end date	93.8	5
Total Non-Fatal WC payments last mth	92.7	5
Normal WC payments last mth	92.2	2
Total payments last qtr	91.3	1
Total payments last mth	91.2	2
Days in last incapacity period	87.8	0
Days since lodgement	37.5	0
Days since entry	37.5	1
Days since escalation	36.8	0
Days since accident	36.6	2
Days since outcome	35.6	6
Days since first WC payment	35.2	0
Days since start of incapacity	34.8	1
Abated WC payments last qtr	33.5	0
Annual Wage based on WC amount	29.5	3

We can see that past payments are the most important predictors, with cumulative non-fatal payments and normal weekly compensation payments to date being the most important. This is interesting since it indicates that the past management of an individual claim is more important in predicting the outstanding cost that the actual nature of the injury to the claimant.

Model Fit

We used gains charts and tables of actual and expected results to evaluate the model fit. Gains charts plot the cumulative total cost captured by a model. Claims are ranked by predicted payment size. The total payments within each percentile of the cases is summed and divided by the total cost of all cases. A good model will capture a large proportion of the total cost in the upper percentiles.

Gains plots for the modelled predictions are compared to two other benchmarks:

- **No model** – in this situation we would be randomly selecting a proportion of the cases from the population and by general reasoning the same proportion of the total cost would be captured. A completely random selection of cases will always produce this result and as such will lead to a gains plot which is a straight line from 0% to 100%.

- **Perfect model** – here the model ranks cases from highest to lowest exactly, producing the theoretical best model possible. To determine the gains we would achieve from such a model we can simply rank the cases by their actual values.

In the plot below, the cumulative percentages for the gains plots are on the right hand side axis and the band percentiles on the horizontal axis. The green line is the gains plot of the predictions; the purple line shows the cumulative costs of the actual payments. A good model is as close as possible to the purple line and far above the diagonal line. By itself, the gains chart does not indicate a great deal about the goodness-of-fit of the model. However, it becomes very useful when comparing two competing models (where higher gains means a better model). We also use the root average squared error (RASE) as a measure of the goodness-of-fit of the model. This statistic is similar to the standard deviation or spread of the residuals and hence a lower RASE indicates a better model.

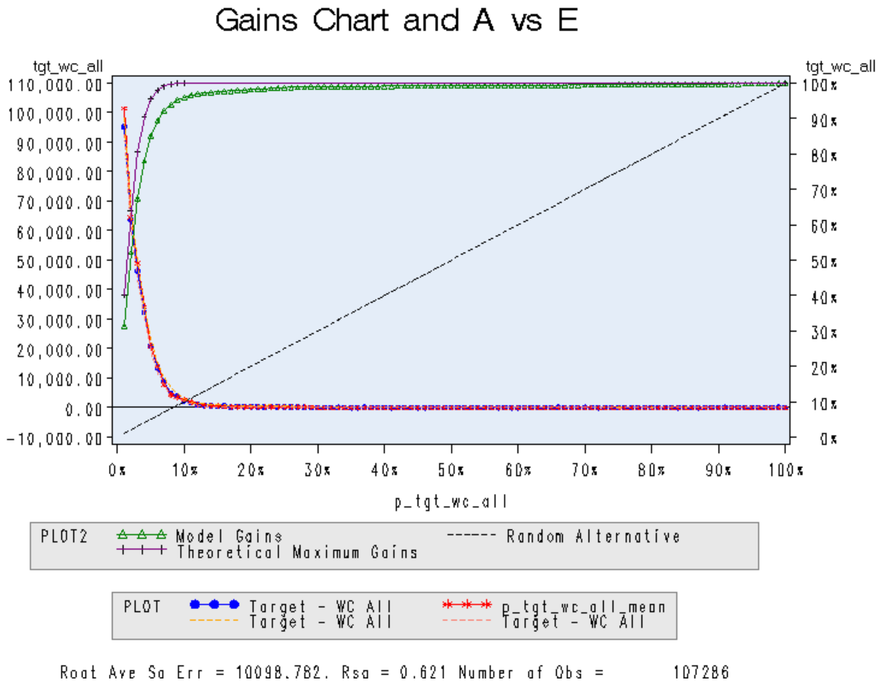

Figure 4.2.2 Weekly Compensation Decision Tree Gains Chart for Test Data

From the plot we can see that there is no consistent bias in actual and expected and the model has achieved significant gains whilst falling a little short of the perfect model. The actual and expected values deviate slightly in the top percentile illustrating that our decision tree model tends to underestimate the largest claims.

4.2.2 Regression Model

The regression model was constructed using the hybrid modelling approach described above. The model was constructed in SAS and was built using the most predictive continuous variables identified in the decision tree stage as initial inputs together with identifier variables for each of the decision tree nodes.

The modelling was done using a two stage approach where the total cost was modelled as:

		"Binary model"		*"Size Model"*
Expected future payment	*x*	*Probability of having a non-zero payment*	*x*	*Size of payment given that it is non-zero*

Separate regression models were built for the probability of payment and the expected size.

We refined the models by including interactions between the different predictor variables and also forcing additional key variables into the model where there was a relationship which was significant for a subgroup of cases but which became insignificant when evaluated across all cases.

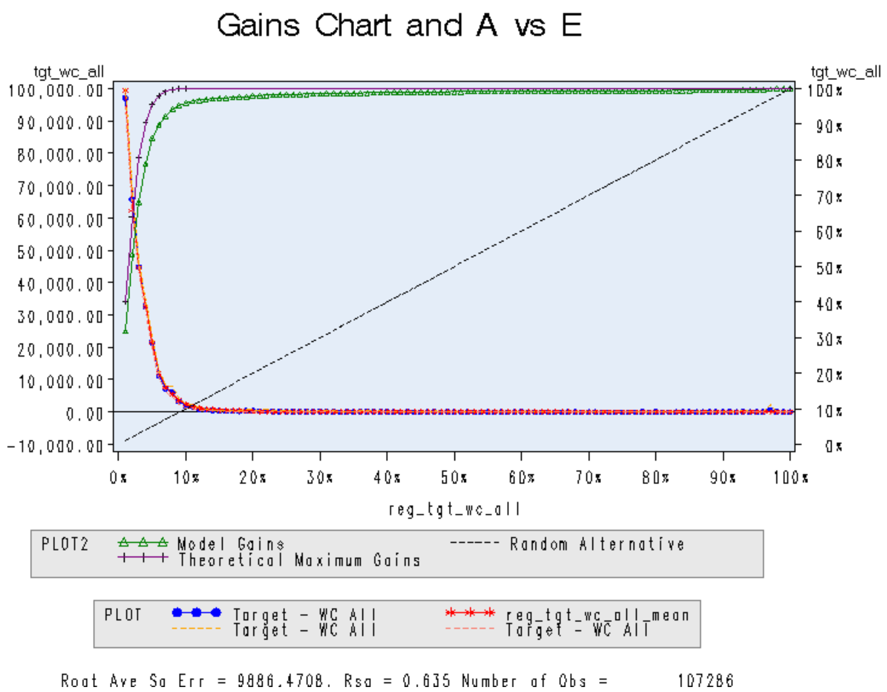

Figure 4.2.3 Weekly Compensation Regression Gains Chart for Test Data

Again we can see that there is no consistent bias in actual and expected and the model has achieved significant gains whilst falling a little short of the perfect model. We can also see that the regression modelling has refined the decision tree model, this can be seen in:

- The slight increase in R-Square value (Rsq)
- The reduction in root average squared error (Root Ave Sq Err)
- A smoother match between actual and expected
- The improved the fit for the largest claims, since the actual and expected values have moved closer in the top percentile.

4.2.3 Payment Pattern and Tail Modelling

The decision tree and regression models produce an estimate of the total payments in the three year period following the model run date. The payment pattern and tail phases split this predicted three year payment into quarterly cash flows and extend the cash flows into the tail.

For the weekly compensation payment type the claims were split into those from serious and non-serious injuries and then banded by predicted value. The payment patterns were fitted by smoothing the observed patterns for each group using regression techniques.

We used the same bands for modelling the tail. For each band we identified the payment pattern and then analysed the hazard rate applying over the third year of this pattern. Our initial estimates of the tail were calculated by using the assumed hazard rate to project future payments and then adjusting this on a case by case basis for expected retirements. This was refined by comparing the results for different accident years to those produced by the main valuation. An illustration of the shape of the projected cash flows is shown below.

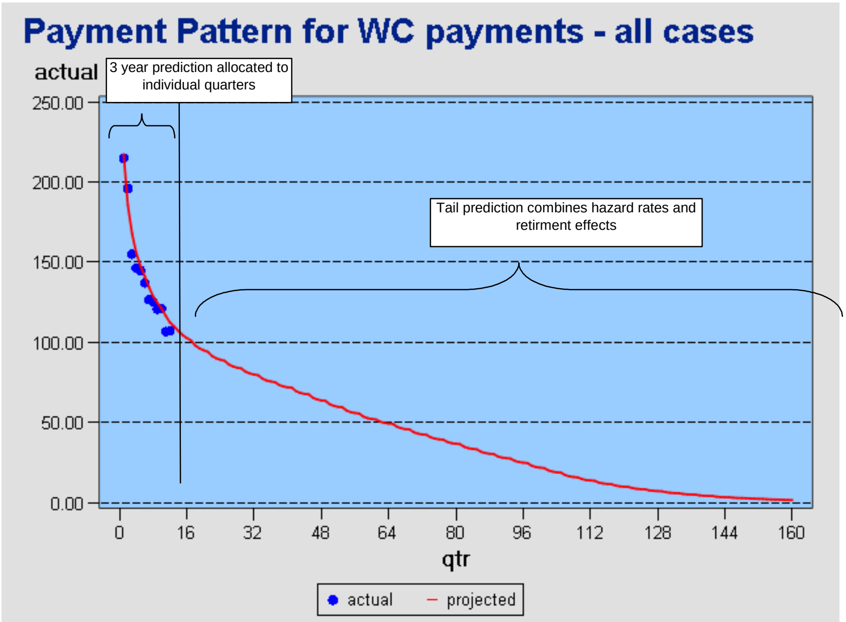

Figure 4.2.3 Weekly Compensation Payment Pattern and Tail Fitting Illustration

4.3 Results

It is a challenge to know how good the SCE is since by definition we do not know the future of individual cases. We have done a number of checks on the model to establish whether the results are reasonable.

The checks include:

- Checks on first 3 years of payments
- Short term comparisons of actual experience emerging versus the expected experience at different model run dates
- Comparisons to aggregate valuation
- Tracking the runoff of the SCE results

An example of the tracking of SCE results is shown in the graph below.

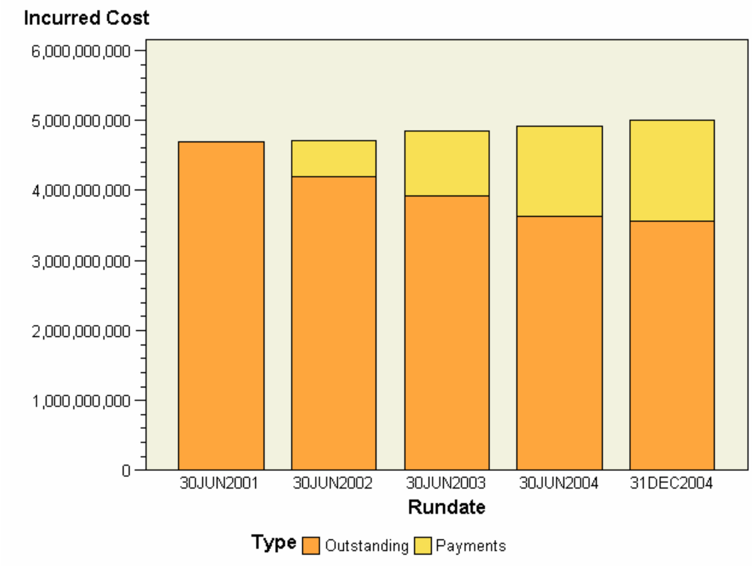

Figure 4.3.1 Tracking of SCE results at 30 June 2001

The outstanding claims estimated by the model run at 30 June 2001 are shown at the left hand side of the graph. At each subsequent model run the cost is re-estimated as the payments made plus the remaining outstanding estimate. If the total incurred cost is stable this indicates that the experience is emerging as expected by the model.

4.3.1 Model Limitations

The SCE model has a number of limitations.

The first of these is that it is an individual claim model and hence makes no allowance for IBNR. This means that adjustment is needed if a total estimate of the outstanding cost of the scheme is needed.

A further weakness is that the predictive ability is comparatively weak for some groups of claims. A particular example of this is for the most recently reported cases. There is less information available for these cases and little history of past claims management to inform the estimate.

It was apparent from the model construction process that not all the relevant data is available for modelling. An improved data collection might enable the models to be improved.

5 The Approach Used for Pricing the Work Accounts

5.1 Background

The premium rates for the two work accounts are risk-rated by the industry group in which the employer or self-employed person is operating. There is no individual experience rating applied (as per legislation) which means that the premium rates that are determined for each group are, in the main, the final premium rates individual employers and self-employed persons pay.

The scheme valuation report is prepared each year as at 31 March with a roll-forward produced as at 30 June. Twelve different payment types are modelled, mostly using PPAC models, giving estimates of the aggregate outstanding claims liability for each accident year and account. Figure 5.1.1 below shows the outstanding claims liability as at 30 June 2004 by account and payment type.

Figure 5.1.1: Outstanding Claims Liability as at 30 June 2004

($ Million)	30 JUNE 2004 TOTAL	WORK ACCOUNTS (EM + SE)	MOTOR VEHICLE ACCOUNT (MV)	MEDICAL MIS-ADVENTURE ACCOUNT (MM)	EARNERS' ACCOUNT (EA)	NON-EARNERS ACCOUNT (NE)
REHABILITATION						
Medical treatment	466	159	68	17	116	106
Social Rehabilitation	2,946	573	1,139	230	236	769
Miscellaneous	962	335	182	31	218	195
	4,374	1,067	1,389	278	570	1,070
COMPENSATION						
Income maintenance	3,955	1,944	965	133	820	93
Impairment benefits	541	95	106	32	92	216
	4,496	2,039	1,071	165	913	309
Present value of the claims liability	8,871	3,106	2,460	443	1,483	1,380
Present value of the operating costs of meeting these claims	444	155	123	22	74	69
Bulk billed costs	33	6	5	1	6	15
Total present value of the claims liability (as at 30 June 2004)	9,347	3,267	2,588	465	1,563	1,464

To set premium rates at the risk group level we need to find some way of allocating the account level liability across the risk groups.

Claims that are allocated to the work accounts are required to have the employer or self-employed person to whom the claim is attributed and the individual classification unit (CU) that identifies the industry identified at registration.

The classification units used by ACC are based on the Australian and New Zealand Industrial Classification System (ANZSIC) and we are currently working through the process of updating to ANZSIC 2006 next year. The CUs are grouped into Levy Risk Groups (LRGs) on the basis of historical claim size relative to payroll with the proviso that all CUs in a LRG must belong to the same industry group (Level 1 of ANZSIC).

So individual claims that have been reported can be grouped into LRGs and the task remains to allocate a portion of the account claims liability to the claims in each LRG.

5.2 Previous Model

Prior to the completion of the new SCE model, ACC did have a model that estimated the ultimate liability for individual claims. However this old SCE model was developed by external actuarial consultants in 1998/99 with little input from internal staff. The model was never documented and was developed using the APL programming language, which is unfamiliar to ACC staff.

Consequently the previous SCE model operated much like the proverbial “black box” with no ability to review and maintain the model, or test the parameters for appropriateness in light of new experience.

The other main issue with the previous SCE model is that it only modelled payments for non-fatal weekly compensation and only for claims in the work accounts.

These limitations meant that ACC was not able to:

- Enhance and update the model
- Provide case estimates for existing claims by all payment types and injury types
- Use the model in the premium setting process for other accounts
- Drill down into the valuation aggregate liability to provide insight into the drivers of results across payment types and accounts
- Support cost benefit analysis of Injury Prevention initiatives
- Provide information on the ultimate claim costs for sub-groups such as employers participating in the optional self-insurance Partnership Programme

The method used to estimate outstanding costs for other payment types was to inflate the actual payments in the same ratio as the outstanding to actual for the weekly compensation payment type modelled. It is fairly obvious that the pattern of payment for other payment types can be quite different to that for the weekly compensation payment type, resulting in the potential for estimates of outstanding costs that are significantly different to actual values.

5.3 The New SCE Model

The new SCE model separately models social rehabilitation payments and weekly compensation payments which, together make up approximately 80% of total claim costs. The social rehabilitation payment type is broken down into payment sub-types of capital, care, assessment and other and each sub-type is split into serious and non-serious injuries. Each of these 8 groups is modelled separately by the new SCE model.

For each individual claim in the work account with a valid CU allocation, we can use the SCE model to estimate ultimate costs for weekly compensation and social rehabilitation payments.

There are two payment types that are modelled separately from the SCE model, being weekly compensation payments to spouse and dependants for fatal claims and independence allowance/lump sum payments made for permanent impairment. These two payment types are both subject to potential commutation and to strict criteria for payment, with little chance of improvement in the condition for which the claim was made. These aspects make the cash flows for these payment types suitable for an annuity model approach.

5.4 Calibration process

The individual case estimates can be grouped at any level but there is no allowance for IBNR claims and only some for re-opened claims. To allow for these we spread the aggregate valuation results across all the open cases at the lowest level in the valuation.

The cases are grouped by accident year, account and payment type and the sum of the case estimates compared to the corresponding valuation liability. A calibration factor is determined for each group and applied to each case in the group. This spreads the IBNR and re-opened claim reserves across the open claims.

Case estimates for the payment types not modelled by the SCE or annuity models (vocational rehabilitation, medical treatment, hospital rehabilitation and other rehabilitation) are calculated through the process of calibrating these payment types to the valuation.

Vocational Rehabilitation: These payments are provided to claimants to enable them to return to work and, as such, are closely correlated with the levels of non-fatal weekly compensation paid. The estimate for an individual case of the outstanding costs for vocational rehabilitation (VR_{case} O/S) is calculated by multiplying the estimate of outstanding costs for non-fatal weekly compensation (NWC_{case} O/S), determined from the SCE model, by the ratio of outstanding vocational rehabilitation costs to outstanding non-fatal weekly compensation costs for the account (A/C) and accident quarter (AQ) of the case, as determined in the valuation.

$$VR_{case}\ O/S = NWC_{case}\ O/S \times \Sigma_{A/C,\ AQ}\,(VR_{Valn}\ O/S)\ /\ \Sigma_{A/C,\ AQ}\,(NWC_{Valn}\ O/S)$$

Lump Sums and Independence Allowance: These payments cannot intuitively be linked to either weekly compensation or social rehabilitation individually and the number of cases already receiving these payments is too small to share the costs of IBNR and re-opened cases. Lump Sums have only been in existence since 1 April 2002 giving a limited amount of data on payment history.

For these reasons, we apportion the valuation outstanding amount at the individual case level in proportion to estimates of ultimate independence allowance costs for the CU the case is attributed to.

The first step is to calculate an allocation of the ultimate independent allowance costs to the CUs (ACU). This is done by dividing the total ultimate independence allowance costs for the last three accident years prior to 1 April 2002 that are attributable to the CU (Ult IA_{CU}), as calculated in the separate Lump Sum/Independence Allowance model, by the total estimated ultimate costs for the last three accident years prior to 1 April 2002 that are attributed to the CU for payment types non-fatal weekly compensation, fatal weekly compensation and social rehabilitation (Ult$\{NWC_{CU}, FWC_{CU}, SR_{CU}\}$).

$$ACU = \Sigma_{3\ AY} (\text{Ult } IA_{CU}) / \Sigma_{3\ AY} (\text{Ult}\{NWC_{CU}, FWC_{CU}, SR_{CU}\})$$

This is then used to calculated an uncalibrated outstanding amount for the Lump Sum/Independence Allowance payment type at the individual case level (LS_{case} O/S) by multiplying the CU allocation by the total estimated ultimate costs for non-fatal weekly compensation, fatal weekly compensation and social rehabilitation for the individual case (Ult$\{NWC_{case}, FWC_{case}, SR_{case}\}$).

$$\text{Uncalibrated } LS_{case} \text{ O/S} = ACU \text{ x } (\text{Ult}\{NWC_{case}, FWC_{case}, SR_{case}\})$$

These outstanding amounts are then calibrated to the valuation at account and accident quarter level.

$$LS_{case} \text{ O/S} = \text{Uncalibrated } LS_{case} \text{ O/S x } \Sigma_{A/C,\ AQ} (LS_{Valn} \text{ O/S}) / \Sigma_{A/C,\ AQ} (LS_{case} \text{ O/S})$$

Other Rehabilitation, Hospital Rehabilitation, Medical Treatment: These payment types are treated in a similar way to the lump sums/independence allowance payment type except the attribution to CUs (ACU) is done on the basis of paid to date costs (PTD) in proportion to paid to date costs of non-fatal weekly compensation and social rehabilitation. The three year summation is done over the last three payment years (PY).

For Other Rehabilitation:

$$ACU = \Sigma_{3\ PY} (\text{PTD } OR_{CU}) / \Sigma_{3\ PY} (\text{PTD}\{NWC_{CU}, SR_{CU}\})$$

$$\text{Uncalibrated } OR_{case} \text{ O/S} = ACU \text{ x } (NWC_{case} \text{ O/S} + SR_{case} \text{ O/S})$$

$$OR_{case} \text{ O/S} = \text{Uncalibrated } OR_{case} \text{ O/S x } \Sigma_{A/C,\ AQ} (OR_{Valn} \text{ O/S}) / \Sigma_{A/C,\ AQ} (OR_{case} \text{ O/S})$$

Hospital Rehabilitation and Medical Treatment are calibrated in the same way.

5.5 Capping of Individual Case Estimates

The result of the calibration process is an estimate of the ultimate claims cost for all claims in the work accounts, the sum of which are in agreement with the valuation. These individual cases can now be grouped into the CUs and LRGs for the purposes of setting a relativity with which to calculate the premium rate for the individual LRGs.

However there are still some situations where one case with a high estimate of ultimate cost can have a large effect on a LRG relativity. To remove the effects of isolated high cost cases, the individual case estimates are capped at a maximum cost of $250,000. The capping level was chosen so that the excess of ultimate claims above this value would be approximately 10%.

Cases in the work accounts are capped at a maximum of $250,000 and the excess is spread across the industry in the account and accident year in which it falls in proportion to the ultimate claim amount. If there are insufficient cases to spread the excess then the capping level is ignored.

6 Impact of the New SCE Model

6.1 SCE Comparison: New Vs. Old

Variances in individual case estimates using the new SCE model compared to case estimates using the old SCE model were expected and were apparent. However when the individual case estimates are grouped into statistically significant groups (for premium setting, defined as >250 escalated claims per annum or >$400m in annual payroll), the differences in the relativities are small. We use a hierarchical credibility weighted approach for setting the relativities for each LRG, meaning that groups that are not statistically significant are credibility weighted with larger groups.

For LRGs with large numbers of claims and/or large amount of payroll there was little change in the relativities determined using the new SCE model. The biggest changes in the actual amount of estimated ultimate claim costs occurred in LRGs with small numbers of claims and low annual payroll. These groups have always had large variation in estimates from year to year and the credibility weighting is used to counteract just such a situation.

The graph below shows the comparison between the ratio of the case estimates to the payroll as determined by the new and the old SCE models, for self-employed people operating in LRG 120 – Livestock Farming, excluding Poultry. This group makes up a large proportion of the Self-Employed Work Account and is considered fully credible with approximately $1.3 billion of payroll and over 2,000 escalated claims each year. As figure 6.1.1 shows, there are only small differences between the two SCE models for an LRG of this size.

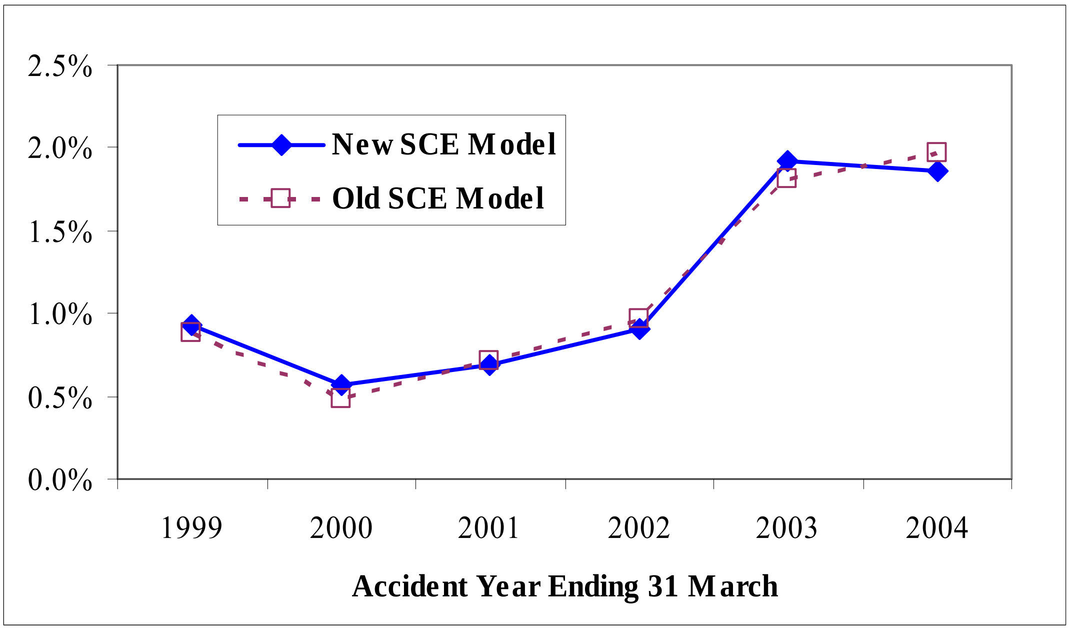

Figure 6.1.1: Comparison of ratio of case estimates to payroll for SE in LRG 120

Where significant differences did occur, this gave the ability to go back to the calibration process and the individual SCE results for a reality check. Frequently the reason for this was a small number of cases over which to spread the total ultimate cost and often this would lead to a refinement in the process that improved the overall allocation.

6.2 Benefits of implementing the SCE Model for Premium Setting

The major benefit from implementing the SCE model was the ability to understand how the case estimates were determined and what assumptions have significant impact on the results. The ability of internal staff to re-parameterise the model and run for different dates and times cannot be underestimated. This means that we can react to changes in experience and legislation that impact the claim costs for different groups of claims in different ways.

The ability to model more than one payment type allows the case estimates to more closely reflect the expected costs given by the scheme valuation. Currently we are modelling nearly 80% of the scheme costs using non-fatal weekly compensation and social rehabilitation payments. In the future there is the potential to model other payment types, giving not just better case estimation, but the ability to analyse changes in observed payment patterns.

There is also the potential to use the SCE model to analyse claims in other accounts, not just the work accounts. In particular, the Motor Vehicle Account is viewed as an important focus for the future.

Robustness and stability of the SCE model is an important advantage in the premium setting process. It reduces the impact of fluctuations from year to year due to the modelling of claim costs.

6.3 Implementation Issues

Like all models, the result is very dependant on the quality of the data used for input. One of the discoveries from the process of building the SCE model was that there is a lot of data stored in our file systems that has little practical use. The reasons for this are varied, but one of the main culprits is the small number of cases that have a particular field entered. For example, a parameter that could potentially be a good predictor is rendered worthless because only 5% of claims have the field populated.

Another issue with the data is the dynamic nature of it. Changes in legislation and treatments available can cause data to become out of date very quickly. The lump sum/independence allowance payment type is an example of this situation.

Building an SCE model involves a large up-front investment in both time and monetary resources. There will also be an ongoing requirement to maintain the model and, at different times, re-parameterise. It is anticipated that these costs will be more than compensated for by the potential areas of the business to which the SCE model can be applied.