Institute of Actuaries of Australia

Transition Models Underlying Statistical Case Estimation

Prepared by Dion Oryzak

Presented to the Institute of Actuaries of Australia
16th General Insurance Seminar 9-12 November 2008
Coolum, Australia

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

Projected valuation finalised claim costs arising from statistical case estimation models are heavily dependent on the reliability of the projected 'states' (such as injury severity) at finalisation from which the cost is derived, as well as the date of finalisation itself. Long tail loss reserving is particularly dependent on the reliability of these transition models. Such transitions can arise from either a genuine change in the claim characteristics (such as a claimant deciding to seek legal representation) or the receipt of more reliable information. A claim's transitioning is complicated by the highly complex nature of long tail claims including the multitude of variables that may influence both the cost and transitioning behaviour, and the interaction of the variables.

This paper reviews the approaches to transition modelling underlying statistical case estimation of general insurance claims, particularly in long tail loss reserving. A CTP insurance portfolio is used in illustrating the approach of modelling transitioning as a multinomial distribution. It is hoped that this paper will, ultimately, lead to wider acceptance of statistical case estimation models that utilise more granular rating variables, whose transitioning behaviour may be seen as being too difficult to model.

Keywords: Statistical Case Estimation, transition, claim state, CTP, injury, severity, Markov, long tail, reserving

Introduction

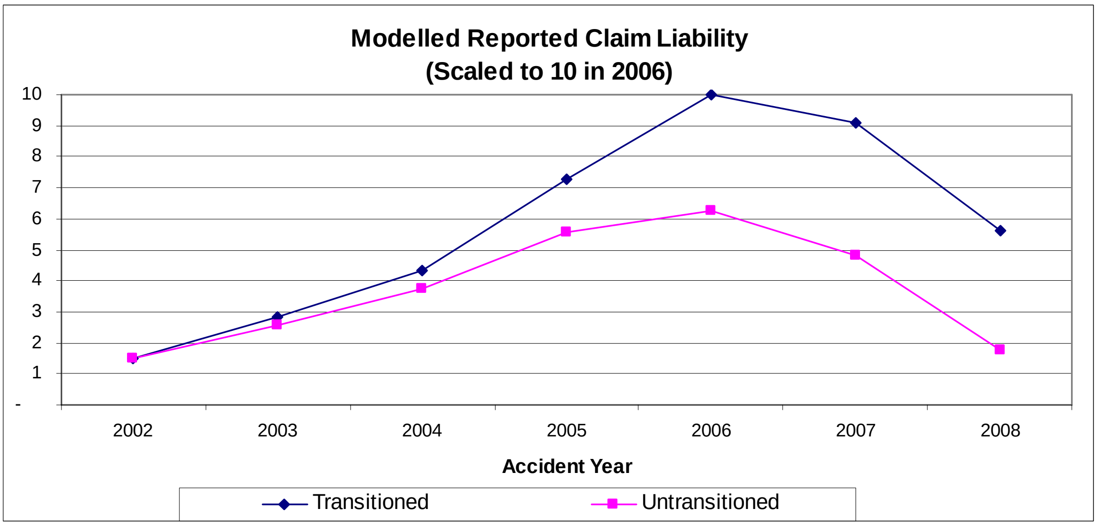
A statistical case estimation model estimates the value of an open claim as a function of the states that key variables take. A severe claim with catastrophic injuries and legal representation should carry a higher outstanding claim liability than a claim for a minor abrasion. A statistical case estimation model will need to be developed before the transition model for the predictor variables can be, so as to evaluate which are the variables that actually drive the modelled cost of claims and therefore which ones need to be included in the transition model. As explored in Brookes, R and Prevett, M (2004), CART® is a useful tool for determining which characteristics are strong predictors of claim cost, as well as the groupings or bandings that those predictors should take. Once the claim size model has been built and tested we are in a position to start building the transition model. The approach in this paper to transitioning assumes that the claim size model has already been built and the key claim characteristics that drive finalised cost have already been decided upon.

Projected finalised claim costs arising from statistical case estimation models are heavily dependent on the reliability of the projected 'states' (such as injury severity) at finalisation from which the cost is derived, as well as the date of finalisation itself. Long tail loss reserving is particularly dependent on the accuracy of these transition models. Such transitions can arise from either a genuine change in the claim characteristics (such as a claimant deciding to seek legal representation), the receipt of additional information or simply correcting incorrectly coded values. A claim's transitioning is complicated by the highly complex nature of long tail claims including the multitude of variables that may influence both the cost and transitioning behaviour, and the interaction of the variables.

A variable used in the calculation of the liability of an outstanding claim, such as injury severity, will take a particular state at the time of data capture (today), but there is a non-zero probability that that variable will transition to a different state by finalisation. Since the statistical case estimate claim size model will typically be parameterised from finalised claims, using the finalised cost and claim characteristics at finalisation, open claims must be transitioned to the profile of expected characteristics at finalisation before the claim size model can be applied and a liability determined. The extended time frame over which long tail claims, such as CTP, will develop, makes any statistical case estimated values highly dependent on the accuracy of the transition model. This paper illustrates the importance of transitioning and the considerations necessary using claim data from a CTP scheme in Australia.

There is a trade-off between adopting a statistical case estimation model using a limited number of variables such as injury severity and operational time as described in Taylor, G. and McGuire, G. (2004), versus modelling finalised claim size using additional variables such as litigation level and suffering the increased model error and variability from having to also project the transitioning behaviour of these variables. It is hoped that this paper will spur debate on the merit of more 'detailed' statistical case estimation models which utilise additional claim information that is ignored in current modelling practice.

Claim characteristics tend, on average, to transition to more severe, and expensive, states. In our experience, when calculating an outstanding claims liability using individual claim characteristics approximately one third of the liability arises from the upward transitioning of claims. The graph below plots the outstanding claims liability on reported claims by accident year rescaling the 2006 liability to 10 (of unitless dimension) to disguise the total size of the liability. 'Transitioned' claims have had their claim characteristics simulated until finalisation, while 'Untransitioned' assumes that the claim characteristics at the date of data capture will remain unchanged until the date of finalisation, and the liability calculated accordingly. In both cases the finalisation date was still simulated. The effect of transitioning served to increase the liability by more than 220% for more recent accident years, reducing to almost no increase for accident years more than six years old.



The three traditional sources of error in actuarial models (model misspecification, parameter estimation and stochastic) can be estimated for the transition model, which may be useful for comparing the validity of alternative models or as an input to risk margin estimation.

Selection of Transition Variables

Claim states transition because claim characteristics at finalisation are different to the states as they exist at the date of data capture for one of three reasons:

1) An aspect of the claim genuinely changes state as a result of new action by either the claimant or the insurer. For example, legal representation may be sought between the date of data capture and the date of claim finalisation.
2) Greater completeness of information about the nature of the claim becomes available. For example, doctors' reports may become available that were not at the time of lodgement and injury recoding may occur as a result of a clearer picture of injuries at the time of the incident being painted.
3) Erroneous information may exist in the claim data. The possibility of coding error makes it necessary to model the transitioning of states even where the underlying claim state cannot change. For example, it may be necessary to model the transitioning of the gender variable from one state to another, or the age of the claimant at the date of the incident.

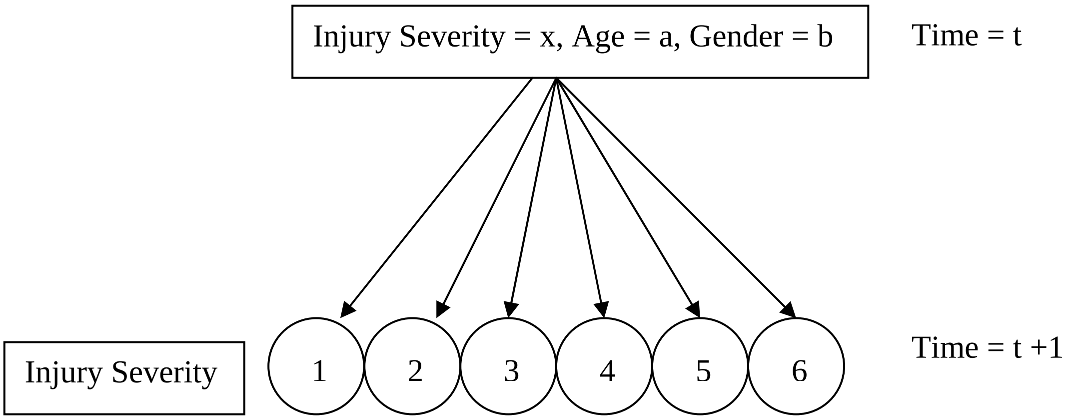
Transition modelling does not differentiate between these three sources of transitioning. For example, to the extent that incompleteness in the claim data at the time of lodgement is not actually identified by claims staff over the life of the claim, the transition model will not identify or project such transitions. Where incompleteness in the data is never corrected by claims staff, the claims cost model itself will adjust to compensate, effectively modelling finalised cost as a function of claim states recorded at lodgement. The statistically case estimated values will be less accurate and more volatile, but it will not necessarily introduce a bias in the estimates unless claim data recording practices change over time.

Variable Preparation

The broad approach to transition modelling is to capture data at regular intervals (eg monthly or quarterly) over the past year, say, and parameterise a multinomial probability distribution of the states that variable could transition to over each period. The transition model may or may not include interaction with other variables such as age or duration. For example, if the transitioning of the severity of the most serious injury (which can take values between 1 and 6) on a claim were being modelled, the multinomial transition model would take the form
$P(InjSev_{t+1}=y|\ InjSev_t=x,\ Age_t=a,\ Gender_t=b\ldots)$

Where:
$InjSev_t$ = Injury severity at time t

Age_t = Age at time t
$Gender_t$ = Gender at time t
x,y = 1,2,3,4,5,6



Once the probability of transitioning from state x to state y over the interval has been determined, the claim states can be simulated into the future as required. For example, the transitioning of injury severity between one of six states over the course of a quarter could be summarised in a 6x6 matrix like that below. This particular matrix assumes that a variable is path independent, i.e. that its state at time t+1 is dependent only on its state at time t and not on its states before time t or on the states of other variables at time t. In a particular simulation, once the claim has been projected to finalise, its characteristics would no longer transition.

$$\begin{matrix}1\\2\\3\\4\\5\\6\end{matrix}\begin{pmatrix} 0.8 & 0.1 & 0.05 & 0.02 & 0.02 & 0.01 \\ 0.06 & 0.79 & 0.06 & 0.04 & 0.03 & 0.02 \\ 0.02 & 0.06 & 0.8 & 0.07 & 0.03 & 0.02 \\ 0.01 & 0.08 & 0.1 & 0.75 & 0.05 & 0.01 \\ 0.01 & 0.02 & 0.03 & 01 & 0.8 & 0.05 \\ 0.01 & 0.01 & 0.01 & 0.03 & 0.04 & 0.9 \end{pmatrix}$$

Variables that are both correlated in their transitioning behaviour and strong drivers of claim cost should be combined and each unique combination treated as one state. For instance, each combination of the severity (taking values 1 to 6) and body region (taking values 1-9) of the most serious injury may be treated as a unique state, resulting in a 54x54 matrix. The complexity of such a transition model, as well as potentially low cell counts, may make it necessary to group some combinations if this interaction is allowed. For example, severities 2 and 3 might be grouped, or the body regions of neck and spine might be grouped. Our experience is that in modelling the transitioning of CTP claim characteristics, the probability of a transition is usually not dependent on the states taken before the current state (i.e. possesses the Markov property). However, the interaction between variables, both in claim cost behaviour and transitioning behaviour is highly complex., and best discovered by testing various combinations of models and variables and by consultation with claim staff.

Many not-so-obvious considerations apply to transition modelling. Extensive communication with claims staff can significantly accelerate the speed of development and reliability of the model compared to pure data mining. Claims staff are able to advise the modeller what to look for, are able to provide a causal explanation for behaviour exhibited in the data, and are able to detect new trends faster than will appear in the data. As transitioning behaviour is highly dependent on claims management practices, it is imperative that a continuous dialogue between the modellers and claims managers exist so that the modellers understand as best as possible why transitions occur. Below are some considerations specific to CTP transition modelling.

Behaviour of Claim Subsets

Certain CTP claims such as workers comp journey claims, psychological trauma claims or death claims may have a cost experience quite different to that of the rest. These should be separated out in the claim size modelling and modelled separately. However, they should still usually be included in the transition modelling as claims might transition into or out of these 'different' states before finalisation. Although modelling the transitioning into or out of a workers comp journey claim state, for instance, is still necessary, the unique nature of these claims may mean that workers comp journey claims should be excluded from the parameterisation of the rest of the transitioning probabilties, such as between litigation levels.

Absorbing States

Certain states, such as death or finalisation should be considered absorbing states. Finalisation should be an absorbing state not just the for the finalisation transition but for all other transition modelling. A claim's injury severity or litigation level usually cannot transition while a claim is in a finalised state. Claim reopenings will still need to be allowed for in the liability calculation, but if the dollar amounts involved are small it may be better to allow for reopenings using a simple percentage loading as a function of the time since finalisation. In this way, in the statistical case estimation and transition modelling, we are only interested in the date of first finalisation. This is suitable if the majority of settlement occurs at or before first finalisation, and reopenings only occur to make small payments or other minor changes.

Categorical vs. Continuous Variables

Categorical variables should be used where possible. Modelling transitioning of continuous variables is much more difficult and usually unhelpful. The bandings should be made in consultation with claims managers. For example, in NSW CTP general damages is payable when Whole Person Impairment (WPI) is greater than 10%, so rather than modelling WPI as a continuous variable, it should actually have only two states: less than 10% or greater than 10%. Similarly, since claims with contributory negligence greater than 25% are exempt from CARS, contributory negligence can be considered to have of two states: <25% or >25%.

Time Frame

The time frame over which the transitioning is projected, and parameters derived is critical. For CTP claims, with a typical dollar weighted duration of two to six years, it is reasonable to transition claims on a quarter by quarter basis. The length of historical period of data used in parameterisation needs to account for the trade-off between sufficiently populating all the transition state combinations, while ensuring that the transition experience is not too old and out-of-date, particularly in an evolving claims management environment. One or two years' worth of data may be optimal. If one year is being measured then five snapshots, each one quarter apart, are needed, recording the states at the beginning and end of each quarter.

Effect of Finalisation on Transitioning

The very act of finalising a claim may trigger claim managers to update variables at finalisation. A spate of transitioning in the finalisation period will be evidence of this, and if so, transition projections should explicitly allow for differences in transitioning behaviour in the finalisation period by including a flag indicating whether the claim finalised in that period as a predictor variable.

Effect of Duration on Transitioning

The transitioning of some variables are more dependent on duration than others. For instance, unlitigated claims that have been open for three years or more would not normally transition to litigation. Therefore it might be useful to model transition litigation separately for claims that have been open for less than three years and more than three years. The transitioning of other variables, such as injury, may not be as dependent on duration.

Legislative Change

The transitioning process should take account of known legislative changes in statutory schemes such as CTP. If the changes are significant enough that experience of claims in each legislative scheme are relatively independent, then the data may be subsetted by scheme before any analysis is undertaken and separate parameters (and possibly model structure) derived for each scheme. This is impossible in the early stages of a new scheme or regulatory change or where too few claims have arisen for claim states under the legislative scheme to be projected purely from the scheme's own experience. In these situations some statistical blending between the schemes is necessary. This may involve the scheme itself being a variable (i.e. scheme 1,2,3 etc), with some or full interaction between this scheme variable and other variables in transition modelling. The Motor Accidents Compensation Act (1999) (MACA) in NSW and the Civil Liability Act (2003) in QLD are examples of legislative changes far-reaching enough to warrant parameterising pre-MACA or pre-CLA and post-MACA or post-CLA claims separately.

The current lack of claims having been finalised under the Lifetime Care and Support scheme (LTCS) in NSW makes it difficult to model the transitioning of claims into or out of LTCS. Until sufficient claim experience exists, transitions into LTCS must be modelled using injury severity and other variables as a proxy. For instance, a claim will be deemed to transition into an LTCS claim if the injury severity, body injury region and/or some other variables are projected to transition to states that claim managers have already identified as sufficiently severe enough to be indicative of qualification for LTCS.

Transition Modelling

Once the claim size model is built, the modeller knows which variables are to be transitioned. But considerations which are not yet obvious are:

1) Which variables need to be transitioned together as a combined variable. The 'combined' variable can be decomposed into its constituent variables after each modelled transition;
2) Which other variables (which may not even appear in the claim size model) are useful predictors for the transitioning behaviour of the transition variables; and
3) Whether some levels of a transition variable need to be modelled separately. For example, injury severity transitioning behaviour may be very different for severities 1-3 compared to severities 4-6.

The LOGISTIC procedure in SAS is a useful tool for modelling multinomial transitioning behaviour. We explore the example of modelling the transitioning behaviour of the injury severity and body region of the two most severe injuries of a claim.

For our data (with non-ordinal responses), multinomial probabilities are best estimated using a generalised logit (glogit) link function, which takes the form:

$$\log\left(\frac{p_{ij}}{p_{ir}}\right) = \alpha_j + x'_i \beta_j$$

This models the log of the ratio of probability of transitioning to state j with the set of claim characteristics i over the probability of transitioning to state j with the referent category claim characteristics that constitute state r, as a linear combination of the characteristics that constitute state j.

In the interests of simplifying this example, we have collapsed the body region levels from 10 to 4 (labelled 0,1,2,7). Due to the highly interactive nature of injury characteristics, in the transition model we have grouped the severities of the first and second most serious injuries and the body region of the most serious injury together in the form:
(Injury 1 Severity) : (Injury 1 Region) : (Injury 2 Severity)

The value 4:1:3 indicates that the most serious injury is of severity 4 to region 1 and the second most serious injury is of severity 3. Combining variables in this manner to preserve complete interaction quickly leads to an almost unwieldy number of levels. However, this is necessary to prevent absurd results that would be possible if the variables were modelled independently of each other, such as the severity of the second most severe injury being more severe than the most serious injury.

Combining variables can lead to a large number of levels both in the transitioned variables and also in the predictor variables. Low cells counts in some transition combinations, and the resulting model instability, means that, contrary to intuition, it is not always best to even include the state of the transition variable combination at time t in predicting its state at time t+1. For instance, in predicting the Inj1Sev:Inj1Reg:Inj2Sev combination at time t+1 the best compromise was, rather than use the Inj1Sev:Inj1Reg:Inj2Sev combination at time t, the following variables were used:

- A flag indicating if the severity of the most serious injury is zero or not (ie there are no genuine injuries recorded) at time t.
- A flag indicating if the second most serious injury has severity 1 or not at time t.
- A flag indicating if the claim is litigated at time t.
- Two spline-based transformations of the claimant's age at accident.
- Two spline-based transformations of the development period (ie number of quarters between lodgement and data capture) at time t.
- The reporting delay.
- The number of injuries recorded at time t.

As an example, the following SAS code illustrates how to perform the multinomial logistic regression above with 1:7:1 as the base level in the GLM. Each record in the input dataset 'data_in' contains the values at time t, and the value of inj1sev_inj1reg_inj2sev at time t+1. Thus, if we are examining x claims transitioning over y quarters, there will be x*y records.

```
proc logistic data=data_in outmodel=params;
class inj1sev_0flag inj2sev_1flag litig_flag /param=ref;
model    inj1sev_inj1reg_inj2sev(ref='1:7:1')=    inj1sev_0flag
inj2sev_1flag  litig_flag  age1  age2  period1  period2  delay2
numinjy /link=glogit aggregate;

output out=data_out predprobs=(individual);
run;
```

Model Validation

The model described above produced a reasonable fit, but the heterogeneity of the higher and lower severities at first led to questionable model validity. As a result of the questionable validity, the data was split into two subsets and modelled separately. The first group is claims where the most serious injury takes values of 0, 1, 2 or 3, and the second group is claims where the most serious injury takes values of 4, 5 or 6. All the predictor variables were retained in the low severity model (0,1,2,3), while only the predictor variables of one of the delay splines and the number of injuries were retained in the high severity model (4,5,6) due to smaller cell counts at the higher injury severities. The AIC and log likelihood summary statistics supported the chosen models over alternatives tested.

Misclassification tables for all three models are shown below. The tables compare the projected injury combination counts (horizontal axis) to the actual counts (vertical axis). Ideally each columns total will be close to the corresponding row total, and the diagonal will be well populated compared to the cells off the diagonal. The tables have been truncated due to the large number of combinations.

All Severities

		Predicted														
		0:0:0	1:0:0	1:1:0	1:1:1	1:2:0	1:2:1	1:7:0	1:7:1	2:1:0	…	6:1:2	6:1:3	6:2:0	6:2:1	Total
Actual	0:0:0	3714	58	76	51	4	4	141	223	27	…	0	0	5	0	4581
	1:0:0	50	733	21	3	3	0	37	9	13	…	0	0	46	0	934
	1:1:0	77	7	237	22	6	2	600	78	192	…	0	0	134	0	1745
	1:1:1	52	24	21	385	1	28	44	1182	16	…	0	0	11	0	2768
	1:2:0	4	0	6	1	0	0	19	3	5	…	0	0	3	0	50
	1:2:1	4	3	2	28	0	2	5	91	1	…	0	0	1	0	229
	1:7:0	145	14	595	53	17	5	2541	269	487	…	0	0	270	0	5682
	1:7:1	227	69	75	1167	3	91	257	5485	64	…	0	0	35	0	12147
	2:1:0	27	4	192	18	5	2	487	67	168	…	0	0	105	0	1455
	….	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	6:1:2	0	0	0	0	0	0	0	0	0	…	0	0	0	0	4
	6:1:3	0	0	0	0	0	0	0	0	0	…	0	0	0	0	4
	6:2:0	5	5	142	10	3	1	290	34	113	…	0	0	131	0	907
	6:2:1	0	0	0	0	0	0	0	0	0	…	0	0	0	0	1
	Total	4579	953	1742	2766	50	229	5676	12140	1453	…	5	5	907	1	

Severities 0,1,2,3 Only

		Predicted														
		0:0:0	1:0:0	1:1:0	1:1:1	1:2:0	1:2:1	1:7:0	1:7:1	2:1:0	…	3:7:0	3:7:1	3:7:2	3:7:3	Total
Actual	0:0:0	3740	58	77	52	4	4	143	226	28	…	10	14	4	2	4581
	1:0:0	51	736	32	3	3	0	54	11	21	…	2	1	0	0	934
	1:1:0	78	5	268	24	7	2	659	85	217	…	31	6	9	2	1745
	1:1:1	52	24	23	390	1	29	48	1196	18	…	2	73	2	0	2768
	1:2:0	4	0	7	1	0	0	20	3	6	…	1	0	0	0	50
	1:2:1	4	3	2	28	0	3	5	92	1	…	0	7	0	0	229
	1:7:0	146	9	649	57	18	5	2693	287	530	…	117	22	37	7	5682
	1:7:1	228	68	82	1180	4	92	273	5543	70	…	15	413	14	2	12147
	2:1:0	28	2	215	20	5	2	532	74	188	…	30	6	11	2	1455
	….	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	3:7:0	9	0	33	3	1	0	116	18	30	…	10	2	7	3	353
	3:7:1	14	2	7	73	0	7	23	413	7	…	2	45	2	0	1006
	3:7:2	4	0	6	3	0	0	33	18	8	…	9	2	32	22	529
	3:7:3	1	0	2	1	0	0	11	4	2	…	3	1	21	24	382
	Total	4581	934	1745	2768	50	229	5682	12147	1455	…	353	1006	529	382	

Severities 4,5,6 Only

		Predicted														
		4:1:0	4:1:1	4:1:2	4:1:3	4:2:0	4:2:1	4:2:2	4:2:3	4:7:0	…	6:1:2	6:1:3	6:2:0	6:2:1	Total
Actual	4:1:0	0	0	1	1	1	1	1	0	0	…	0	0	9	0	19
	4:1:1	0	2	6	8	3	4	6	9	0	…	0	0	4	0	65
	4:1:2	1	6	18	32	8	14	20	37	0	…	0	0	13	0	226
	4:1:3	1	8	31	89	10	24	39	108	0	…	1	1	8	0	511
	4:2:0	1	3	8	11	5	7	9	11	0	…	0	0	14	0	96
	4:2:1	1	4	14	26	6	11	16	29	0	…	0	0	9	0	176
	4:2:2	1	5	19	41	8	15	23	48	0	…	0	0	10	0	263
	4:2:3	0	9	35	108	10	28	46	132	0	…	1	1	9	0	608
	4:7:0	0	0	0	0	0	0	0	0	0	…	0	0	10	0	14
	….	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
	6:1:2	0	0	0	1	0	0	0	1	0	…	0	0	0	0	4
	6:1:3	0	0	0	1	0	0	0	1	0	…	0	0	0	0	4
	6:2:0	11	6	15	4	17	12	11	3	10	…	0	0	749	0	907
	6:2:1	0	0	0	0	0	0	0	0	0	…	0	0	0	0	1
	Total	19	65	226	511	96	176	263	608	14	…	4	4	907	1	

The parameters can be then applied to any dataset with the predictor variables using the following SAS code. The modeller can then actually perform the transition simulation or test the validity of the model using the ‘hold-out’ method, k-fold-cross validation or any other statistical validation technique.

```
proc logistic inmodel=params;
      score data=validationdata out=data_out;
   run;
```

Error Estimation

The three traditional sources of error in actuarial models are:

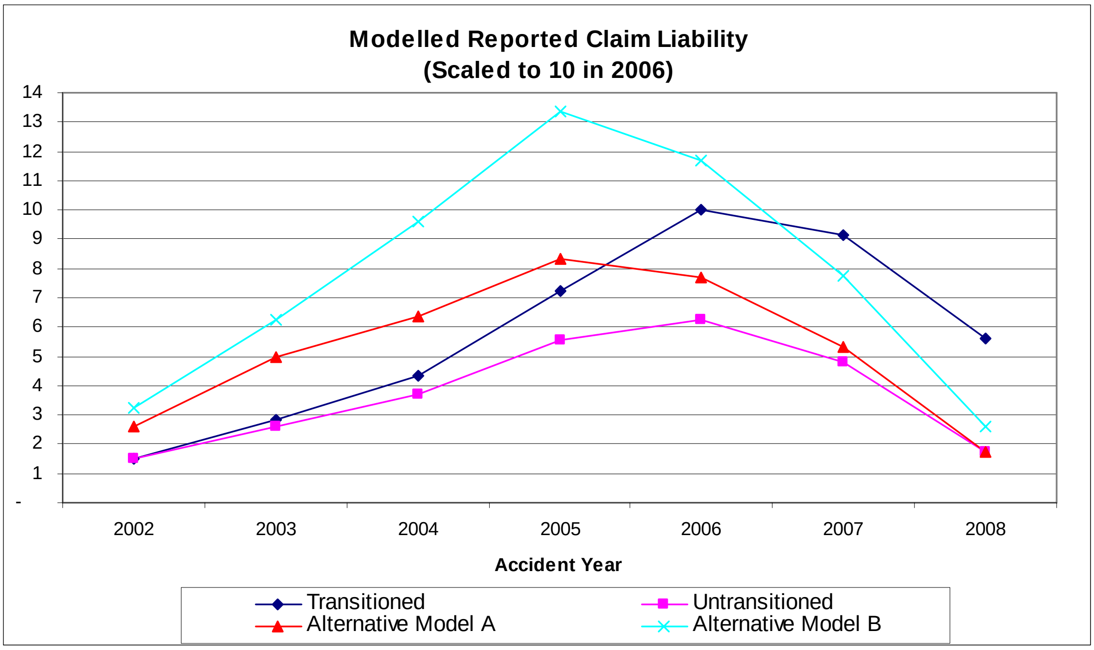
1) Model misspecification error;
2) Parameter error;
3) Process error

The claim state transitioning model described above allows for analysis of each of these sources of error. Such error estimation is useful in the selection of competing models, or as part of the risk margin calculation. It is important to note that in the methodology described below, the statistical claim size model is taken as given, and the variability is due to the transition model alone. The claim size model will carry its own model misspecification error, parameter error and process error which must be estimated separately. Interactions of error between the two models can also be explored.

Model Misspecification Error

An infinite number of alternative transitioning models exist. Different predictor variables or even different model structures such as neural networks could be used. The modeller can gain some insight into the effect of model misspecification by trying numerous reasonable models and observing the differences in results between them. However, the notion of what constitutes 'reasonable' by the modeller could lead to underestimation of model misspecification error. For example, a modeller may feel that only multinomial models are reasonable and their estimation of model misspecification may be limited to observing the differences between a number of multinomial models with different predictor variables and interaction combinations but neglect other model structures under the belief that they are 'unreasonable'. It may be useful therefore to also test some 'unreasonable' models. The graph below compares the reported outstanding claims liability against two alternative models with different predictor variables for the transitioning of the dynamic variables.

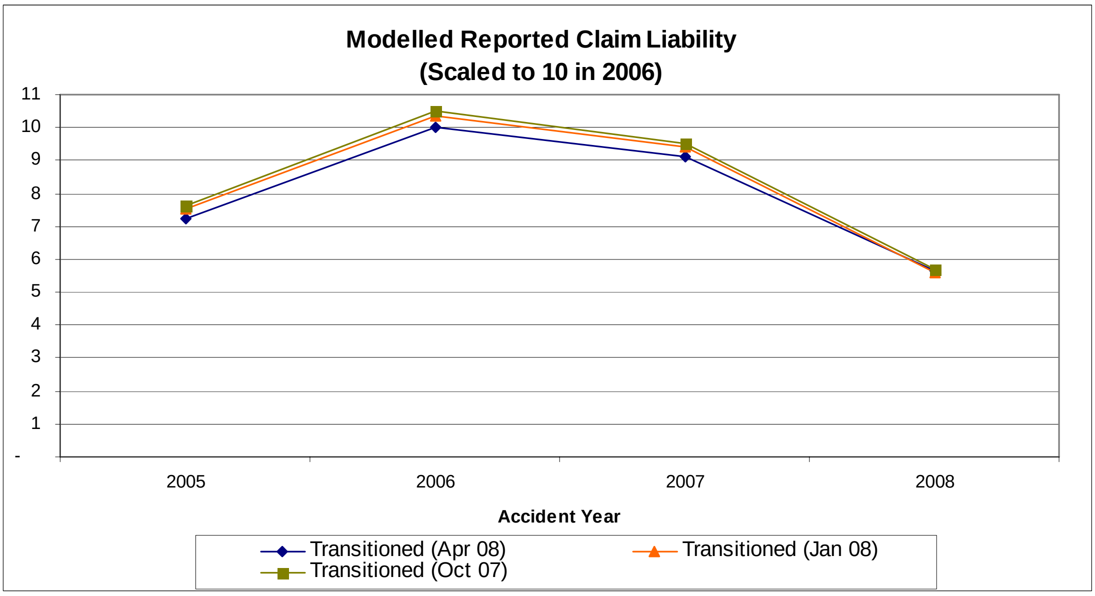
Alternative Model A models the Inj1Sev:Inj1Reg:Inj2Sev combination at time t+1 using only Inj1Sev:Inj1Reg:Inj2Sev combination at time t rather than the multitude of other predictors described in the previous section. This model was highly unstable, did not converge properly, and produced counter-intuitive results including the negligible transitioning of claims in recent accident years compared to higher transitioning of older claims. Alternative Model B modelled the transitioning Inj1Sev, Inj1Reg and Inj2Sev independently. This model was also unstable, did not converge properly and the lack of interaction produced a liability much larger than under other models.



Parameter Error

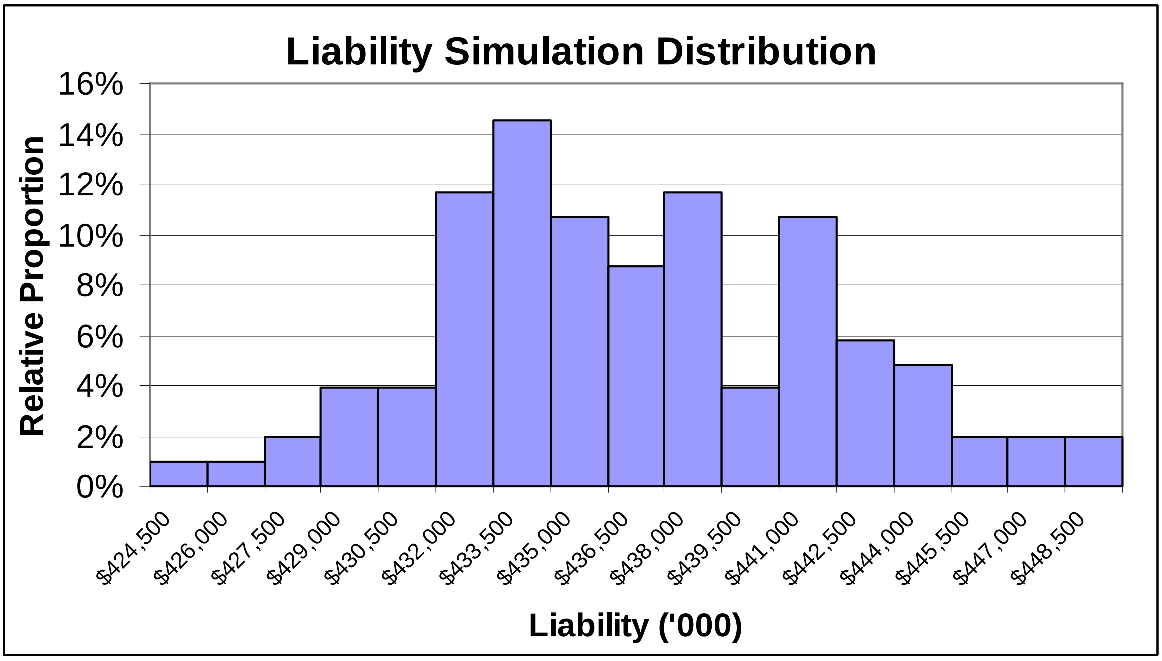
Parameter error can estimated by a number of methods including estimating the parameters using subsets of the data or data from different periods and observing the effect on the results. A model that exhibits stable results regardless of subsetting the data in the parameter estimation process or the period of data from which the parameters were estimated can reasonably be believed to have lower parameter error.

The models described above each calculated the transitioning model parameters using the eight quarterly intervals from 31 January 2006 to 30 April 2008. The graph below shows how the outstanding claim liability shifts if the parameters were estimated from either the eight quarterly intervals from 31 October 2005 to 31 Jan 2008, or 31 July 2005 to 31 October 2007. The difference between each of these periods is essentially only a shift of the eight-period moving average by one period, yet made more than 5% difference to the liability. A sampling period longer than eight quarters would give more stable results, but potentially at the risk of confounding the model with older, irrelevant experience.



Process Error

The liabilities shown above are the average of 100 simulations. Process error can be estimated by plotting a histogram of the results of the simulation, taking the model and parameters as given, so that only the randomly generated numbers underlying the distribution are allowed to vary. The graph below shows the simulated distribution of the liability (rescaled to disguise the size). As stated before, the apparent variability is that due to the transition model only, so the process error from the claim size model will need to be analysed separately.



Conclusion

This paper has aimed to impart a greater appreciation for the issues involved in simulating the transitioning of individual claims characteristics for the purpose of using a wider range of claim information in statistical case estimation. It is hoped that this may spur debate on the trade-off between using more granular claims data versus the potential difficulty of modeling the transitioning of such characteristics, and ultimately lead to greater acceptance of granular statistical case estimation among valuation modeling practitioners.

References

Brookes, R and Prevett, M (2004) Statistical Case Estimation Modelling – An Overview of the NSW WorkCover Model. IAA Accident Compensation Seminar 2004

Chapman, S and Winn, R (2004) Premium Rating for Personal Injuries Using Statistical Case Estimates. IAA XVth General Insurance Seminar 2005

McGuire, G (2007) Individual claim modelling of CTP data. IAA Accident Compensation Seminar 2007

McGuire (2007) Building a reserving robot. IAA Biennial Convention 2007

Mulquiney, P (2004) Application of Soft-Computing Techniques in Accident Compensation. IAA Accident Compensation Seminar 2004

Taylor, G. and McGuire, G. (2004) Loss reserving with GLMs: A Case Study.
IAA Accident Compensation Seminar 2004

Taylor, G, McGuire, G and Sullivan, J (2006) Individual claim loss reserving conditioned by case estimates. Institute of Actuaries UK