The Actuarial Profession
making financial sense of the future

Reserving Seminar
Joseph Lo, Aspen

Extending the Mack Bootstrap

June 2012

Aim of the Workshop

- The Mack model is a widely used tool to assess reserve uncertainty
  - original model gives prediction error, on ultimates that are based on the chain ladder
    - Mack, 1993 and after
    - formula driven, implementable simple spreadsheets
  - a key extension to the model gives simulations, that are consistent with the original model, via bootstrapping
    - England and Verrall, series of papers up to 2006
    - allows deployment in a Monte Carlo set up to give a full distribution, found popular for internal stochastic capital modelling

- By the end of the workshop, we shall be able to discuss:
  - potential key problems in implementing the Mack model
  - and possible practical solutions to them

Format of the Workshop

- Very brief introduction to how practitioners commonly implement the Mack model
- Consider potential problems
  - Data
  - Resampling – “bootstrapping” and “forecasting”
  - Scaling
- General discussions anticipated under each of the three topics
  - Brainstorming and sharing of experience with neighbours and with the wider workshop will be compulsory (or, at least, encouraged!)

A health warning about examples

- Examples are important in workshops: they are included in this workshop for the narrow purpose of communicating and discussing the technical concepts in relation to implementation of the Mack model
- For ease of reference, we shall be making use of publicly available claim development triangles
- In particular, it is outside the scope of this workshop to discuss and reach conclusions with regards to the organisations behind the examples: *reliance should not be placed on the contents of the slides or discussions in the workshop for such wider issues*

A common implementation of the Mack model (1: before resampling)

- Select a "representative" cumulative triangle, $\{C_{i,j}\}_{i=1,\ldots,I;\, j=1,\ldots,I+1-i}$
  - often with reference to the triangle(s) used to estimate the mean ultimates
  - possibly adjust data to remedy distorting features
- Estimate chain ladder factors, $\hat{f}_j$
  - possibly with curve fitting for smoothing and extending to the tail
- Estimate the variance parameters, $\hat{\sigma}_j$
  - key formula: $\mathrm{Var}\left(C_{i,j+1}\right) = {\sigma_j}^2 \cdot C_{i,j}$
  - if required, extending for the tail development
- Calculate residuals, $\{r_{i,j}\}_{i=1,\ldots,I;\, j=1,\ldots,I-i}$
  - from the observed development factors, $\{f_{i,j}\}_{i=1,\ldots,I;\, j=1,\ldots,I-i}$
  - using the estimated parameters, $\hat{f}_j$ and $\hat{\sigma}_j$
  - often with bias adjustments

A common implementation of the Mack model (2: after resampling)

- Repeatedly resample the $\{r_{i,j}\}_{i=1,\ldots,I;\,j=1,\ldots,I-i}$ into the triangle
  - in each instance, back out "pseudo development factors" $\tilde{f}_j$
  - possible curve fitting
  - "estimation error", "parameter error"

- In each instance, use a distribution to iteratively project into the bottom-half of the cumulative triangle
  - E.g. gamma distribution: giving only positive values
  - Two moments of $C_{i,j+1}$ for fitting the distributions: $\mathrm{E}\left(C_{i,j+1}\right) = \tilde{f}_j \cdot C_{i,j}$ and $\mathrm{Var}\left(C_{i,j+1}\right) = \hat{\sigma}_j{}^2 \cdot C_{i,j}$
  - "forecast error", "process error"

- Scaling and application of inflation volatility
  - Scaling is useful so that the means of the ultimate simulations match reserving's means
  - Scaling also used to adjust the volatilities of the resulting ultimate distributions that are deemed more reflective of the uncertainty
  - Inflation is often modelled separately, and applied to the outputs of the Mack bootstrap

Other implementations of the Mack model

- Use of the original model's formula to arrive at the prediction error and then fit overall distributions for the ultimate position
  - as suggested by Mack in his 1993 paper
  - simple and quick to run
  - but may underestimate the extreme tails (e.g. GIROC 2007)

- Use of Bayesian techniques
  - as discussed by England and Verrall in their 2006 paper
  - can get around issues surrounding negative development factors
  - has the ability to incorporate prior beliefs of parameter error
  - seemingly not widely deployed yet – why?
    - have we become comfortable with bootstrapping?
    - computational difficulties? "MCMC"

- Any comments at this stage? Are you doing something else?

1. Data
brainstorming potential problems

Here are two examples of data for us to brainstorm potential problems with data when implementing the Mack model

Data: Axis Liability Reinsurance Incurred

Cumulative Triangle

	Dev Year						
UWY	1	2	3	4	5	6	7
2003	252	4,626	6,541	7,362	8,780	9,899	14,279
2004	5,290	16,791	19,230	25,196	27,390	30,290	
2005	7,376	23,607	34,388	35,671	40,478		
2006	12,899	31,034	39,681	47,118			
2007	17,758	37,132	47,463				
2008	21,838	40,483					
2009	18,206						
f-hat	235%	130%	116%	112%	111%	144%	100%
sigma^2-hat	16,037	268	329	37	3	0	0

Residuals

	Dev Year				
UWY	1	2	3	4	5
2003	220%	52%	-15%	120%	123%
2004	52%	-138%	137%	-117%	-70%
2005	63%	163%	-139%	44%	
2006	6%	-27%	41%		
2007	-30%	-30%			
2008	-63%				

Data: Axis Marine Insurance Incurred

Cumulative Triangle

	Dev Year							
AY	1	2	3	4	5	6	7	8
2002	23,087	29,866	35,051	34,675	33,947	33,393	33,515	33,225
2003	20,644	25,605	26,341	34,063	35,853	36,344	35,452	
2004	79,663	109,129	109,535	108,057	109,784	109,857		
2005	354,142	446,611	466,813	479,460	475,957			
2006	57,558	81,091	99,884	89,932				
2007	64,850	106,533	124,645					
2008	77,653	97,184						
2009	60,176							
f-hat	132%	108%	101%	100%	100%	99%	99%	
sigma^2-hat	1,524	847	881	54	8	14	8	

Residuals

	Dev Year					
AY	1	2	3	4	5	6
2002	-12%	61%	-16%	-58%	-131%	102%
2003	-33%	-31%	172%	156%	112%	-98%
2004	37%	-94%	-31%	88%	9%	
2005	-101%	-86%	40%	-68%		
2006	57%	163%	-132%			
2007	226%	111%				
2008	-55%					

1. Data: potential problems

- Volume of data
  - Do we have enough years?
- Are the small initial years distorting estimates or the residual set?
- Unusual events
  - Are they distorting the parameter estimates?
- Changing business mix / limit profiles
- Underwriting year cohorts
  - What's the problem?
- Unstable variance parameter estimates
- Claims not developed enough
- Trends and shocks
  - Is the data satisfying the independence assumptions?
  - Any origin period / calendar period trends or features?

1. Data: possible practical remedies
a. Isolating data and adjusting $\hat{\sigma}_j^{\,2}$

- *We sometimes have the option to put aside the Mack bootstrap in favour of other models:*
  - *this can be a realistic and practical option in some cases,*
  - *although it is outside the scope of this workshop*

- Are the small initial years distorting estimates or the residual set?
- Unusual events
  - Can take out distorting data points
  - And / or remove claims associated with the unusual events
  - Are they really distorting?
  - Special adhoc modelling would be required for claims that have been taken out

- Changing business mix / limit profiles
  - Can disaggregate the triangle into more homogenous triangles and perform Mack modelling on each
  - But would also need to calibrate dependencies
  - Another way is to estimate a set of variance parameters for each accident year
  - The simplest version is $\widehat{\sigma_{i,j}}^2 = (1+\gamma_i)\cdot\hat{\sigma}_j^{\,2}$, where γ_i are uplift factors separately calibrated – can also be useful for considering netting down for outwards reinsurance or for the effects of adjustment premiums
  - Another version is to calibrate variance parameters for sub triangles and then recombine them using weights

1. Data: possible practical remedies
b. Curve fitting to $\widehat{\sigma_j}^2$ and power parameter

- Underwriting year cohorts
- Unstable variance parameter estimates
- Claims not developed enough
  - Can consider fitting a curve through the $\widehat{\sigma_j}^2$ for smoothing and extrapolating
  - A candidate is the exponential curve – seems to work well in many cases (why!?)

- Volume of data
  - Can consider using market or other representative triangles
  - Need to watch out for how the model translates market volatility to company specific volatility
  - In the process error, CoV is proportional to ${}^{1}/_{\sqrt{C_{i,j}}}$
  - So a market process CoV of 10% could translate to a potentially unrealistic company specific process CoV of 100% (if the company takes on around 1% of market share)
  - A possible solution is to deviate from the volume-weighted chain ladder method, and bring in a “power parameter” α, so that $\mathrm{Var}\left(C_{i,j+1}\right) = \sigma_j{}^2 \cdot C_{i,j}{}^{\alpha}$
  - The original model has 1 for the power parameter
  - As it tends to 2, the effect of the CoV increase would become less severe
  - The power parameter is also discussed briefly by Mack in a later paper

1. Data: possible practical remedies
c. Identifying trends and shocks

- Trends and shocks
  - Is the data satisfying the independence assumptions?
  - Any origin period / calendar period trends or features?
  - Hypothesis testing can be used to identify significant trends or shocks for further investigations
  - Recall that in the residual resampling step, we have backed out a high number of sets of pseudo development factors – one set for each instance of resampling
  - However, the resampled residuals could also give us distributions of the residual sample means for a particular subset of the triangle (e.g. a particular calendar period)
  - The observed sample means could then be tested against these distributions and the *p*-value estimated
  - Other statistics could be considered such as correlation coefficients between two adjacent development periods of residuals, or the gradient of the means of the residuals vs calendar year
  - Charts could also be used to identify trends and shocks
  - If trend is significant, can de-trend and then put in trend in the projection
  - Or can take out data that are contributing significantly to trends
  - If there are significant isolated features in the triangles, can take them out for special resampling
  - See *Extending the Mack Bootstrap* in the printed GIRO 2011 conference papers for more details
  - It discusses hypothesis testing and two resampling techniques in details, furnished with step-by-step examples
  - A few slides at the end of this workshop outlines and supplements the paper

2. Resampling brainstorming potential problems

Arch, 3rd Party Occurrence Insurance
Residuals from the Incurred Data

AY	1	2	3	4	5	6
2002	93%	37%	111%	119%	-32%	128%
2003	-15%	188%	-87%	-130%	-122%	-59%
2004	155%	70%	127%	92%	118%	
2005	-23%	-14%	34%	-18%		
2006	-32%	-94%	-114%			
2007	115%	-97%				
2008	-149%					

Arch, 3rd Party Claims Made Insurance
Reisudals from the Paid Data

AY	1	2	3	4	5	6
2002	48%	5%	-82%	-34%	-12%	-140%
2003	7%	-156%	-81%	184%	156%	22%
2004	-44%	-75%	72%	-1%	-74%	
2005	252%	-57%	-148%	-72%		
2006	10%	163%	99%			
2007	14%	11%				
2008	-43%					

XLI, Casualty Insurance
Residuals from the Incurred Data

AY	1	2	3	4	5	6	7	8
2000	120%	-169%	-92%	46%	-119%	-65%	-106%	98%
2001	165%	103%	198%	-100%	-14%	118%	133%	-102%
2002	-151%	57%	-122%	-22%	165%	-136%	-35%	
2003	45%	-33%	-74%	198%	-25%	57%		
2004	-51%	-19%	30%	-91%	87%			
2005	-46%	-36%	27%	-5%				
2006	46%	11%	9%					
2007	-121%	186%						
2008	-44%							

2. Resampling: potential problems
a. Calendar period correlations

Low Developments along a Calendar Period; Low Variability along another

XLI, Casualty Insurance

Residuals from the Incurred Data

AY	1	2	3	4	5	6	7	8
2000	120%	-169%	-92%	46%	-119%	-65%	-106%	98%
2001	165%	103%	198%	-100%	-14%	118%	133%	-102%
2002	-151%	57%	-122%	-22%	165%	-136%	-35%	
2003	45%	-33%	-74%	198%	-25%	57%		
2004	-51%	-19%	30%	-91%	87%			
2005	-46%	-36%	27%	-5%				
2006	46%	11%	9%					
2007	-121%	186%						
2008	-44%							

CY	2005	2006
Mean	-85%	-40%
SD	41%	25%

2. Resampling: potential problems
b. Correlations in other dimensions

Positive Correlations between Successive Development Periods

Arch, 3rd Party Occurrence Insurance

Residuals from the Incurred Data

AY	1	2	3	4	5	6
2002	93%	37%	111%	119%	-32%	128%
2003	-15%	188%	-87%	-130%	-122%	-59%
2004	155%	70%	127%	92%	118%	
2005	-23%	-14%	34%	-18%		
2006	-32%	-94%	-114%			
2007	115%	-97%				
2008	-149%					

Correlation between the 3rd and 4th periods: 98%

Positive Correlations between Successive Development Periods; High Developments on an Origin Period

Arch, 3rd Party Occurrence Insurance

Residuals from the Incurred Data

AY	1	2	3	4	5	6
2002	93%	37%	111%	119%	-32%	128%
2003	-15%	188%	-87%	-130%	-122%	-59%
2004	**155%**	**70%**	**127%**	**92%**	**118%**	
2005	-23%	-14%	34%	-18%		
2006	-32%	-94%	-114%			
2007	115%	-97%				
2008	-149%					

Mean of the 2004 AY: 112%

2. Resampling: potential problems
c. Significant skewness in a subset of residuals

Positive Skewness in a Developemnt Period

Arch, 3rd Party Claims Made Insurance

Reisudals from the Paid Data

AY	1	2	3	4	5	6
2002	48%	5%	-82%	-34%	-12%	-140%
2003	7%	-156%	-81%	184%	156%	22%
2004	-44%	-75%	72%	-1%	-74%	
2005	252%	-57%	-148%	-72%		
2006	10%	163%	99%			
2007	14%	11%				
2008	-43%					

Skewness of the 1st period: 208%

2. Resampling: possible remedies
a. Estimation error

- Those features of the triangles that contradict the model assumptions are “exceptions”
- Here the assumptions are that all the residuals are independent against one another

- A way to remedy this is to extend the model so that in the new extended model, the features would no longer appear exceptional
- Two techniques are:
  - **Sieve resampling**
  - **Exception resampling**

- We could perform hypothesis testing on the extended model:
  - To identify new exceptions for further investigations
  - To verify if the old exceptions had been accommodated

2. Resampling: possible remedies
a.i. Sieve resampling

- A.k.a. “partition resampling” or “constrained resampling” by practitioners
- The triangle is partitioned into two or more parts
- And the residuals are constrained to be resampled their respective parts

- Useful when a subset of the triangle contains residuals that come from significantly different distributions
  - E.g. significant positive skewness in the first development periods

- Impact to the overall volatility is typically low
  - Performing such resampling of residuals to the Arch 3rd Party Claims Made paid triangle decreases estimation error by around 2%

2. Resampling: possible remedies
a.ii. Exception resampling

- Here, the exceptional features (e.g. the 2005 calendar period in the XL casualty insurance incurred triangle) are sampled onto randomly selected calendar periods:
  - E.g. in any simulation, any calendar period could independently have the 2005 calendar period residuals (and *only* have them)

- The extended model recognises that the exceptional feature is an instance of a dynamic that every so often produces such a feature

- The impact to estimation error is around 10% with this example

- Exception resampling could also be performed with origin periods, with pairs of development periods, etc.

- **The key driver of the increase is that there is now correlation between the simulated $\tilde{f}_j$'s**

2. Resampling: possible remedies
b. Process error

- Exceptional features could also be resampled in the forecasting (i.e. in the projection of the bottom-half of the triangle)

- We present an approach here using "calendar period drivers". It simulates calendar period exceptional features into the future.
  - This could be generalised for other dimensions (see slides at end of GIRO workshop)

- Secondary dependencies could also be imposed

2. Resampling: possible remedies
b.i. Calendar period drivers

- Recall that:
  - In the common approach, each cell in the bottom-half of the triangle is simulated using a (e.g.) gamma distribution
  - When simulating, it is typical to simulate uniform random variables and then apply the inverse CDFs to obtain a simulation from the distribution

- We can mimic the exceptional calendar period feature by simulating the uniforms with biased weights
  - E.g. the XL casualty insurance incurred triangle has 2005 calendar year residuals that have a significantly low mean
  - Whenever the 2005 calendar period is simulated in a future calendar year, we can obtain low developments by sampling from uniforms with appropriately calibrated low weights
  - In this instance, doing so gives an increase in process error of around 8%

2. Resampling: possible remedies
b.ii. Other possible dependencies

- Using this technique, it is possible to impose further dependencies
  - Between (the means of) calendar periods
  - Between origin periods in the same calendar periods

- E.g. for XL casualty insurance incurred data,

  - with 10% correlation between the means of adjacent calendar period drivers (the observed data gives *minus* 40%)
  - and 10% correlation between adjacent origin periods in the same calendar period (the observed data gives a wide range of *minus* 60% to *plus* 20%)
  - we obtain a further increase in prediction error of around 3%

  - *In total, with the 2005 calendar year used for exception resampling (a.ii.), and then for calendar period driver (b.i.), together with the dependencies above, we obtain a total increase in the prediction error of around 17%*

3. Scaling: brainstorming potential problems

- Going back to the XL Insurance Casualty Incurred example:
  - The chain ladder method (with no further tail) gives an IBNR of 1,048
  - The (original) Mack bootstrap gives a mean IBNR of 1,048 and SD of 433
  - The published IBNR was 2,811

3. Scaling: potential problems

- Should we aim to preserve SD or CoV in scaling?
- Should we adjust the CoV in response to a different mean?
- What should we do if we achieve negative reserves (or even negative ultimates!) for many simulations after scaling?
- Is the Mack model still valid if there is a large amount of scaling?
- How should we scale claim emergence?

3. Scaling: possible remedies

- Standard deviations and CoVs behave in two ways under the Mack model for different levels of latest amounts
  - The estimation error, being a product of the latest amounts with the volatility of the pseudo development factors, $\widetilde{f}_j$'s, preserves CoVs on scaling
  - The process error CoV, as suggested previously, is proportional to ${}^{1}/_{\sqrt{C_{i,j}}}$

- A potentially good way to scale is to scale the $\widetilde{f}_j$'s themselves, start the iteration from the *actual* latest amounts, and let the mechanics of the model produce the combined volatility
  - Would require different sets of pseudo development factors for different origin cohort
  - The scaling would likely naturally give full correlations between the different sets
  - A possible way of scaling the $\widetilde{f}_j$'s for an accident year is first scale the estimation error ultimate distribution (by preserving the estimation error CoV), and then scale the $\widetilde{f}_j$'s so that the scaled development factors would take the latest amounts to the scaled estimation error ultimates
  - The scaling of the $\widetilde{f}_j$'s can be done in the log scale
  - If the actual latest amounts are very small compared with what one expects it should be, then the volatility of the next step would be underestimated.
  - Ideas from "Robust Estimation of Reserving Risk" could be used to obtain an "as-if" latest amount to apply the formula $\mathrm{Var}\left(C_{i,j+1}\right) = {\sigma_j}^2 \cdot C_{i,j}$
  - Scaling in this way would help to model claim emergence – and so could contribute to claim experience monitoring and one-year modelling
  - Finally, negative reserves / ultimates would become less of an issue – although not eliminated entirely

Literature

A great deal has been (and is still being) published on the Mack model. Here is a selection. – the first one is an example of a recent paper, refining the model for unstable data; the second one discusses, among other topics, bootstrapping on the Mack model and other implementation possibilities; the third one is the original paper on the subject.

- Busse, M., Mueller, U., & Dacorogna, M. (2010). Robust estimation of reserve risk. *Astin Bulletin, 40*(2), 453-490.
- England, P., & Verrall, R. (2006). Predictive distributions of outstanding liabilities in general insurance. *Annals of Actuarial Science, 1*(II), 221-270.
- Mack, T. (1993). Distribution-free calculation of the standard error of chain ladder reserve estimates. *Astin Bulletin, 23*(2), 213-225.

We have only briefly discussed exceptional features in this workshop. Further details can be found in the first paper, which is supplemented by the appendix of the second GIRO workshop reference here:

- Lo, J. (2011 August). Extending the Mack bootstrap, hypothesis testing and resampling techniques. *GIRO 2011 Conference Papers*, 29-79.
- Lo, J. (2011 October). Implementing the Mack model. *GIRO 2011 Conference Workshop A9.*

Scaling also only had limited time today. Further details have been illustrated in the following presentation.

- Koslover, P., & Lo, J. (2012). Reserve uncertainty in the London Market. *London Market Actuaries Group lecture, March 2012*.

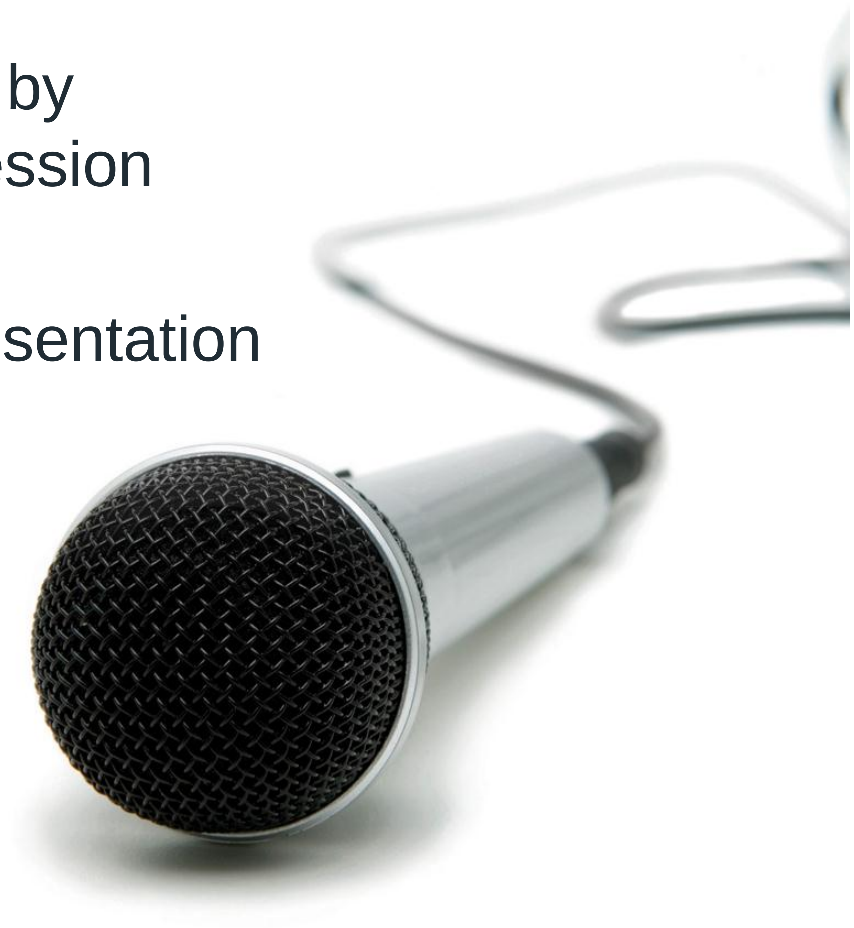
A technical survey of stochastic reserving is

- Wüthrich, M. V., & Merz, M. (2008). *Stochastic claims reserving methods in insurance.* Wiley.

Finally, here is a paper that discusses practical issues in bootstrap modelling (although not on the Mack model)

- Shapland, M., & Leong, J. (2010). Bootstrap modeling: beyond the basics. *CAS E-Forum, 2010 Fall*.

Questions or comments?



Expressions of individual views by members of The Actuarial Profession and its staff are encouraged.

The views expressed in this presentation are those of the presenter.