Cass Business School
CITY UNIVERSITY LONDON

Reserving Using Paid and Incurred Data

Professor Richard Verrall

Cass Business School
CITY UNIVERSITY LONDON

Why Use Paid and Incurred Data?

- Paid data is "real"

BUT

- Paid claims data may be unstable
- Incurred data may include some useful information

"Munich chain-ladder"
Use paid claims and reported numbers

Cass Business School
CITY UNIVERSITY LONDON

Papers

Huijuan Liu and Richard Verrall: Bootstrapping MCL, to appear in Variance

Richard Verrall, Jens Perch Nielsen, Anders Jessen, Maria Martinez-Miranda: 2 papers in ASTIN on a new model using reported numbers and paid amounts

Cass Business School
CITY UNIVERSITY LONDON

Notation

We us the notation of Quarg and Mack, so that cumulative claims for accident year *i* and development year *j* are denoted by C_{ij}^P and C_{ij}^I .

$$F_{ij}^P = \frac{C_{i,j+1}^P}{C_{ij}^P} \qquad F_{ij}^I = \frac{C_{i,j+1}^I}{C_{ij}^I}$$

$$Q_{ij} = \frac{C_{ij}^P}{C_{ij}^I}$$

Cass Business School
CITY UNIVERSITY LONDON

Assumptions

$$E\left[F_{ij}^{P}\,\middle|\,P_i(j)\right]=f_j^{P} \qquad Var\left[F_{ij}^{P}\,\middle|\,P_i(j)\right]=\frac{\left(\sigma_j^{P}\right)^2}{C_{ij}^{P}}$$

$$E\left[F_{ij}^{I}\,\middle|\,I_i(j)\right]=f_j^{I} \qquad Var\left[F_{ij}^{I}\,\middle|\,I_i(j)\right]=\frac{\left(\sigma_j^{I}\right)^2}{C_{ij}^{I}}$$

$$E\left[Q_{ij}^{-1}\,\middle|\,P_i(j)\right]=q_j^{-1} \qquad Var\left[Q_{ij}^{-1}\,\middle|\,P_i(j)\right]=\frac{\left(\tau_j^{P}\right)^2}{C_{ij}^{P}}$$

$$E\left[Q_{ij}\,\middle|\,I_i(j)\right]=q_j \qquad Var\left[Q_{ij}\,\middle|\,I_i(j)\right]=\frac{\left(\tau_j^{I}\right)^2}{C_{ij}^{I}}$$

Cass Business School
CITY UNIVERSITY LONDON

Residuals

$$r_{ij}^{P}=\frac{F_{ij}^{P}-E\left[F_{ij}^{P}\,\middle|\,P_i(j)\right]}{\sqrt{Var\left[F_{ij}^{P}\,\middle|\,P_i(j)\right]}} \qquad r_{ij}^{I}=\frac{F_{ij}^{I}-E\left[F_{ij}^{I}\,\middle|\,I_i(j)\right]}{\sqrt{Var\left[F_{ij}^{I}\,\middle|\,I_i(j)\right]}}$$

$$r_{ij}^{Q^{-1}}=\frac{Q_{ij}^{-1}-E\left[Q_{ij}^{-1}\,\middle|\,P_i(j)\right]}{\sqrt{Var\left[Q_{ij}^{-1}\,\middle|\,P_i(j)\right]}} \qquad r_{ij}^{Q}=\frac{Q_{ij}-E\left[Q_{ij}\,\middle|\,I_i(j)\right]}{\sqrt{Var\left[Q_{ij}\,\middle|\,I_i(j)\right]}}$$

Cass Business School
CITY UNIVERSITY LONDON

Correlations

$$E\left[r_{ij}^{P}\left|B_i(j)\right.\right]=\rho^{P}r_{ij}^{Q^{-1}}$$

$$Cov\left[r_{ij}^{P},r_{ij}^{Q^{-1}}\left|P_i(j)\right.\right]=Corr\left[r_{ij}^{P},r_{ij}^{Q^{-1}}\left|P_i(j)\right.\right]=Corr\left[F_{ij}^{P},Q_{ij}^{-1}\left|P_i(j)\right.\right]=\rho^{P}$$

$$E\left[r_{ij}^{I}\left|B_i(\mathrm{j})\right.\right]=\rho^{I}r_{ij}^{Q}$$

$$Cov\left[r_{ij}^{I},r_{ij}^{Q}\left|B_i(\mathrm{j})\right.\right]=Corr\left[r_{ij}^{I},r_{ij}^{Q}\left|B_i(\mathrm{j})\right.\right]=Corr\left[F_{ij}^{I},Q_{ij}\left|B_i(\mathrm{j})\right.\right]=\rho^{I}$$

Cass Business School
CITY UNIVERSITY LONDON

MCL Adjusted development factors

$$E\left[F_{ij}^{P}\left|B_i(\mathrm{j})\right.\right]=E\left[F_{ij}^{P}\left|P_i(\mathrm{j})\right.\right]+\sqrt{\frac{Var\left[F_{ij}^{P}\left|P_i(\mathrm{j})\right.\right]}{Var\left[Q_{ij}^{-1}\left|P_i(\mathrm{j})\right.\right]}}Corr\left[F_{ij}^{P},Q_{ij}^{-1}\left|B_i(\mathrm{j})\right.\right]\left(Q_{ij}^{-1}-E\left[Q_{ij}^{-1}\left|P_i(\mathrm{j})\right.\right]\right)$$

$$\hat{\lambda}_{ij}^{P}=\hat{f}_{j}^{P}+\hat{\rho}^{P}\frac{\hat{\sigma}_{j}^{P}}{\hat{\tau}_{j}^{P}}\left(Q_{ij}^{-1}-\hat{q}_{j}^{-1}\right)$$

$$E\left[F_{ij}^{I}\left|B_i(\mathrm{j})\right.\right]=E\left[F_{ij}^{I}\left|I_i(\mathrm{j})\right.\right]+\sqrt{\frac{Var\left[F_{ij}^{I}\left|I_i(\mathrm{j})\right.\right]}{Var\left[Q_{ij}\mid I_i(\mathrm{j})\right]}}Cov\left[F_{ij}^{I},Q_{ij}\left|B_i(\mathrm{j})\right.\right]\left(Q_{ij}-E\left[Q_{ij}\left|I_i(\mathrm{j})\right.\right]\right)$$

$$\hat{\lambda}_{ij}^{I}=\hat{f}_{j}^{I}+\hat{\rho}^{I}\frac{\hat{\sigma}_{j}^{I}}{\hat{\tau}_{j}^{I}}\left(Q_{ij}-\hat{q}_{j}\right)$$

Cass Business School
CITY UNIVERSITY LONDON

Reserving and Bootstrapping

Define and fit statistical model

Obtain residuals and pseudo data
Re-fit statistical model to pseudo data

Obtain forecast, including process error

Any model that can be clearly defined can be bootstrapped

See England and Verrall: “Predictive Distributions of Outstanding Liabilities in General Insurance,” *Annals of Actuarial Science* 1, 2007

Cass Business School
CITY UNIVERSITY LONDON

Residuals

Residuals are $U_{ij} = \left\{ \left(r_{ij}^{P} \right), \left(r_{ij}^{Q^{-1}} \right), \left(r_{ij}^{I} \right), \left(r_{ij}^{Q} \right) \right\}$

$$r_{ij}^{P} = \frac{F_{ij}^{P} - \hat{f}_{j}^{P}}{\hat{\sigma}_{j}^{P}} \sqrt{C_{ij}^{P}} \qquad r_{ij}^{Q^{-1}} = \frac{Q_{ij}^{-1} - \hat{q}_{j}^{-1}}{\hat{\tau}_{j}^{P}} \sqrt{C_{ij}^{P}}$$

$$r_{ij}^{I} = \frac{F_{ij}^{I} - \hat{f}_{j}^{I}}{\hat{\sigma}_{j}^{I}} \sqrt{C_{ij}^{I}} \qquad r_{ij}^{Q} = \frac{Q_{ij} - \hat{q}_{j}}{\hat{\tau}_{j}^{I}} \sqrt{C_{ij}^{I}}$$

Bootstrap bias correction $\sqrt{(n-j) / (n-j-1)}$

Cass Business School
CITY UNIVERSITY LONDON

Bootstrap procedure

Resample grouped residuals, with replacement

$$U_{ij}^{B}=\left\{\left(r_{ij}^{P}\right)^{B},\left(r_{ij}^{Q^{-1}}\right)^{B},\left(r_{ij}^{I}\right)^{B},\left(r_{ij}^{Q}\right)^{B}\right\}$$

Construct bootstrap samples

$$\left(F_{ij}^{P}\right)^{B}=\frac{\left(r_{ij}^{P}\right)^{B}\hat{\sigma}_{j}^{P}}{\sqrt{C_{i,j}^{P}}}+\hat{f}_{j}^{P} \qquad \left(Q_{ij}^{-1}\right)^{B}=\frac{\left(r_{ij}^{Q^{-1}}\right)^{B}\hat{\tau}_{j}^{P}}{\sqrt{C_{i,j}^{P}}}+\hat{q}_{j}^{-1}$$

$$\left(F_{ij}^{I}\right)^{B}=\frac{\left(r_{ij}^{I}\right)^{B}\hat{\sigma}_{j}^{I}}{\sqrt{C_{i,j}^{I}}}+\hat{f}_{j}^{I} \qquad \left(Q_{ij}\right)^{B}=\frac{\left(r_{ij}^{Q}\right)^{B}\hat{\tau}_{j}^{I}}{\sqrt{C_{i,j}^{I}}}+\hat{q}_{j}$$

Cass Business School
CITY UNIVERSITY LONDON

Bootstrap estimates

Calculate bootstrap parameter estimates (using MCL), and

$$\left(\hat{\lambda}_{ij}^{P}\right)^{B}=\left(\hat{f}_{j}^{P}\right)^{B}+\left(\hat{\rho}^{P}\right)^{B}\frac{\left(\hat{\sigma}_{j}^{P}\right)^{B}}{\left(\hat{\tau}_{j}^{P}\right)^{B}}\left(\left(Q_{ij}^{-1}\right)^{B}-\left(\hat{q}_{j}^{-1}\right)^{B}\right)$$

$$\left(\hat{\lambda}_{ij}^{I}\right)^{B}=\left(\hat{f}_{j}^{I}\right)^{B}+\left(\hat{\rho}^{I}\right)^{B}\frac{\left(\hat{\sigma}_{j}^{I}\right)^{B}}{\left(\hat{\tau}_{j}^{I}\right)^{B}}\left(\left(Q_{ij}\right)^{B}-\left(\hat{q}_{j}\right)^{B}\right)$$

Cass Business School
CITY UNIVERSITY LONDON

Predictive distributions

Bootstrapping recursive models, England and Verrall (2007)

$$X_{kl}^{P} \sim Normal\left(\left(\hat{\lambda}_{i,l-1}^{P}\right)^{B} \hat{X}_{k,l-1}^{P}, \left(\left(\hat{\sigma}_{l-1}^{P}\right)^{2}\right)^{B} \hat{X}_{k,l-1}^{P}\right)$$

$$X_{kl}^{I} \sim Normal\left(\left(\hat{\lambda}_{i,l-1}^{I}\right)^{B} \hat{C}_{k,l-1}^{I}, \left(\left(\hat{\sigma}_{l-1}^{I}\right)^{2}\right)^{B} \hat{X}_{k,l-1}^{I}\right)$$

Cass Business School
CITY UNIVERSITY LONDON

Example: Quarg and Mack data

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$
$i=1$	576	1804	1970	2024	2074	2102	2131
$i=2$	866	1948	2162	2232	2284	2348	
$i=3$	1412	3758	4252	4416	4494		
$i=4$	2286	5292	5724	5850			
$i=5$	1868	3778	4648				
$i=6$	1442	4010					
$i=7$	2044						

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$
$i=1$	978	2104	2134	2144	2174	2182	2174
$i=2$	1844	2552	2466	2480	2508	2454	
$i=3$	2904	4354	4698	4600	4644		
$i=4$	3502	5958	6070	6142			
$i=5$	2812	4882	4852				
$i=6$	2642	4406					
$i=7$	5022						

	Bootstrap		MCL		Mack	
	Paid	Incurred	Paid	Incurred	Paid	Incurred
$i=1$	0	43	0	43	0	43
$i=2$	37	94	35	96	32	97
$i=3$	109	131	103	135	158	88
$i=4$	277	321	269	326	331	277
$i=5$	299	296	289	302	407	191
$i=6$	657	651	646	655	919	465
$i=7$	5492	5646	5505	5606	4063	6380
Overall Total	6871	7182	6846	7163	5911	7540

Cass Business School
CITY UNIVERSITY LONDON

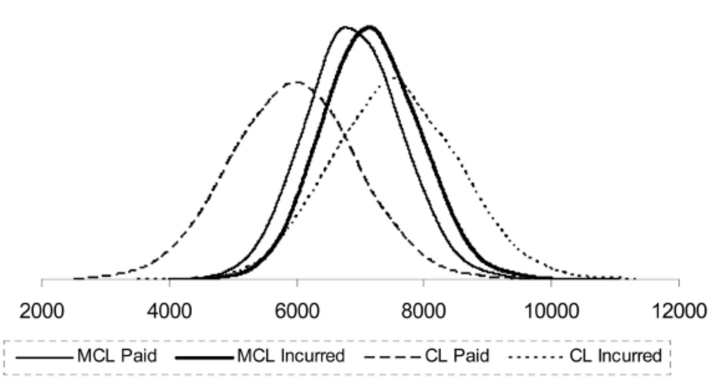
Example: Prediction errors

	MCL		Mack	
	Paid	Incurred	Paid	Incurred
$i = 1$	0	0	0	0
$i = 2$	5	5	15	9
$i = 3$	48	70	53	82
$i = 4$	61	86	68	105
$i = 5$	72	104	72	117
$i = 6$	215	208	289	216
$i = 7$	735	716	897	869
Overall Total	776	782	991	980

	MCL		Mack	
	Paid	Incurred	Paid	Incurred
$i = 1$	—	0%	—	0%
$i = 2$	14%	5%	45%	9%
$i = 3$	44%	53%	33%	93%
$i = 4$	22%	27%	21%	38%
$i = 5$	24%	35%	18%	61%
$i = 6$	33%	32%	31%	46%
$i = 7$	13%	13%	22%	14%
Total	11%	11%	17%	13%

Cass Business School
CITY UNIVERSITY LONDON

Example: Predictive distributions



Cass Business School
CITY UNIVERSITY LONDON

Modelling the claims process

Model payment numbers and amounts separately (if the data are available).
Should we use paid or incurred data?
Verrall, Nielsen and Jessen (to appear in ASTIN) assume we have reported numbers of claims and paid aggregate claims. We build a model using the methods set out in the literature by (for example):
Bühlmann, H., Schnieper, R. and Straub, E. (1980): Claims reserves in casualty insurance based on a probability model. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker.
Norberg, R., (1986): A contribution to modelling of IBNR claims. Scandinavian Actuarial Journal, 155-203.
Norberg, R., (1993): Prediction of outstanding liabilities in non-life insurance. Astin Bulletin, vol. 23, no. 1, 95-115.
Norberg, R., (1999): Prediction of outstanding claims: Model variations and extensions. Astin Bulletin, vol. 29, no. 1, 5-25.

Cass Business School
CITY UNIVERSITY LONDON

Modelling the claims process

We have reported numbers of claims N_{ij}. We build a payment delay model for the (latent) paid numbers of claims, by first considering

$N_{ijk}^{paid} \mid N_{ij}$ which we assume has a multinomial distribution. This is the payment delay.

The number of paid claims can be found by summing these:

$$N_{ij}^{paid} = N_{ij0}^{paid} + N_{i,j-1,1}^{paid} + \ldots + N_{i,0,j}^{paid}$$

The IBNR delay is considered by modelling the incurred number of claims.
The RBNS delay is considered through the model for $N_{ijk}^{paid} \mid N_{ij}$.

Cass Business School
CITY UNIVERSITY LONDON

Paid claims

Denoting paid claims by X_{ij},

$$X_{ij} = \sum_{k=1}^{N_{ij}^{paid}} Y_{ij}^{(k)}$$

$$E\left[X_{ij}\right] = E\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right] E\left[Y_{ij}^{(k)}\right]$$

$$V\left[X_{ij}\right] = E\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right] V\left[Y_{ij}^{(k)}\right] + V\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right]\left(E\left[Y_{ij}^{(k)}\right]\right)^{2}$$

Assuming claims are iid

$$E\left[X_{ij}\right] = E\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right]\mu$$

$$V\left[X_{ij}\right] = E\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right]\sigma^{2} + V\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right]\mu^{2}$$

Cass Business School
CITY UNIVERSITY LONDON

Paid claims

$$E\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right] = E\left[\sum_{k=0}^{\min\{j,d\}} N_{i,j-k,k}^{paid} \mid \mathbb{N}^{I}\right] = \sum_{k=0}^{\min\{j,d\}} E\left[N_{i,j-k,k}^{paid} \mid \mathbb{N}^{I}\right] = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k$$

$$V\left[N_{ij}^{paid} \mid \mathbb{N}^{I}\right] = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \left(1 - p_k\right)$$

$$E\left[X_{ij}\right] = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \mu$$

$$V\left[X_{ij}\right] = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \sigma^{2} + \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \left(1 - p_k\right)\mu^{2}$$

$$= \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \left(\sigma^{2} + \left(1 - p_k\right)\mu^{2}\right) \approx \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \left(\sigma^{2} + \mu^{2}\right)$$

Cass Business School
CITY UNIVERSITY LONDON

Model for Paid claims

Hence, with this approximation, we can use an ODP model for paid claims: similar to chain-ladder, except there is more going on in the parameters.

$$V\left[X_{ij}\right] \approx \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} p_k \left(\sigma^2 + \mu^2\right) = \varphi E\left[X_{ij}\right]$$

$$\varphi = \frac{\sigma^2 + \mu^2}{\mu^2}$$

Cass Business School
CITY UNIVERSITY LONDON

Applying the model

Apply a chain-ladder (Poisson or ODP) to the reported numbers of claims. The projected numbers of reported claims will allow us to analyse IBNR claims (separately).
Apply an ODP model to the triangle of paid aggregate claims. The mean is

$$E\left[X_{ij}\right] = \sum_{k=0}^{\min\{j,d\}} N_{i,j-k} \psi_k$$

$$p_k = \frac{\psi_k}{\sum_{k=0}^{d} \psi_k} \qquad \mu = \sum_{k=0}^{d} \psi_k \qquad \sigma^2 = \varphi \sum_{k=0}^{d} \psi_k - \left(\sum_{k=0}^{d} \psi_k\right)^2$$

Cass Business School
CITY UNIVERSITY LONDON

Data: Paid claims

i\j	0	1	2	3	4	5	6	7	8	9
1	451288	339519	333371	144988	93243	45511	25217	20406	31482	1729
2	448627	512882	168467	130674	56044	33397	56071	26522	14346	
3	693574	497737	202272	120753	125046	37154	27608	17864		
4	652043	546406	244474	200896	106802	106753	63688			
5	566082	503970	217838	145181	165519	91313				
6	606606	562543	227374	153551	132743					
7	536976	472525	154205	150564						
8	554833	590880	300964							
9	537238	701111								
10	684944									

Cass Business School
CITY UNIVERSITY LONDON

Data: Incurred counts

i\j	0	1	2	3	4	5	6	7	8	9
1	6238	831	49	7	1	1	2	1	2	3
2	7773	1381	23	4	1	3	1	1	3	
3	10306	1093	17	5	2	0	2	2		
4	9639	995	17	6	1	5	4			
5	9511	1386	39	4	6	5				
6	10023	1342	31	16	9					
7	9834	1424	59	24						
8	10899	1503	84							
9	11954	1704								
10	10989									

Cass Business School
CITY UNIVERSITY LONDON

Delay functions

j	Development Factor	IBNR Delay
0		0.8752
1	1.1353	0.1184
2	1.0038	0.0038
3	1.0009	0.0009
4	1.0003	0.0003
5	1.0003	0.0003
6	1.0002	0.0002
7	1.0001	0.0001
8	1.0003	0.0003
9	1.0004	0.0004

The average IBNR delay is 0.14 years.
The average RBNS delay is 1.52 years.

j	0	1	2	3	4	5	6	7
p	0.3637	0.2881	0.1134	0.0852	0.0661	0.0358	0.0255	0.0222

Cass Business School
CITY UNIVERSITY LONDON

Estimated outstanding claims

i	IBNR	RBNS	TOTAL	CHAIN LADDER
2	628	605	1,233	1,685
3	1,350	4,514	5,863	29,379
4	1,510	43,623	45,133	60,638
5	1,967	94,526	96,493	101,158
6	2,579	171,633	174,212	173,802
7	3,168	299,136	302,304	249,349
8	5,349	509,334	514,684	475,992
9	14,280	852,144	866,423	763,919
10	254,499	1,135,678	1,390,177	1,459,860
Total	285,329	3,111,192	3,396,521	3,315,779

Cass Business School
CITY UNIVERSITY LONDON

Conclusions

Many methods, including chain-ladder, use just one triangle of aggregated data. Using a little more information, you can get more information out. This is what Munich chain-ladder tries to do, but it starts from the same point as the chain-ladder technique (aggregate triangles) rather than considering the way the data is generated.

When considering scenarios for capital modelling, it is better to be able to look at quantities that have real meaning.