The Actuarial Profession
making financial sense of the future

GIRO Conference and Exhibition 2012

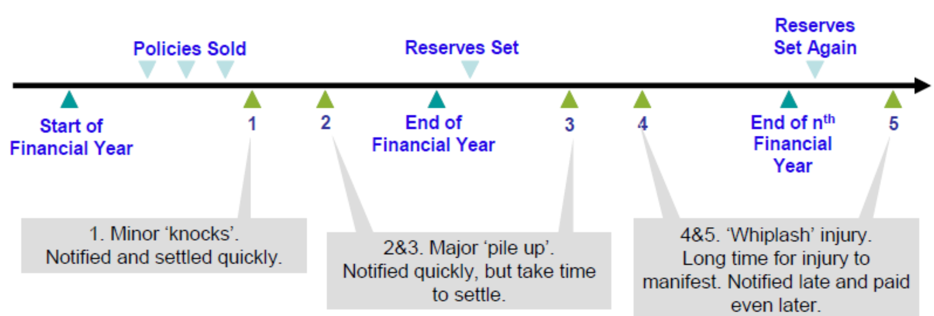
Adding prior knowledge to Double Chain Ladder

M.D. Martínez-Miranda - J.P. Nielsen

Outline

1. Introducing the problem: stochastic reserving
2. Motivating a statistical model for stochastic reserving: the double chain ladder model
3. Estimating the model from two run-off triangles: the double chain ladder method
4. Adding prior knowledge to double chain ladder with to purposes:
   - More stable estimates: Bornhuetter-Ferguson and double chain ladder
   - To consider more *general distributional models*: development severity inflation and zero-claims
5. Conclusions

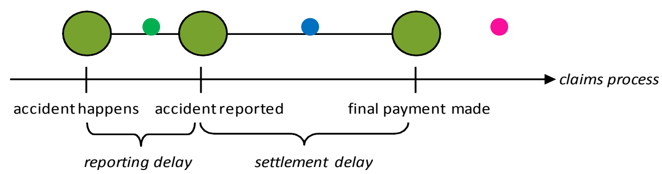
Introducing the problem: stochastic reserving



- The company needs to put **reserves** aside to fulfill his liabilities in the future, for both claims events that have already happened and also for claims that have not been fully settled yet
- A hard problem: much **uncertainty**, many **dependencies**…

The individual claims mechanism

- The life of an individual claim in the general claims process:



- Three categories of claim:
  - Incurred but not reported, **IBNR**
  - Reported but not settled, **RBNS**
  - Reported and paid

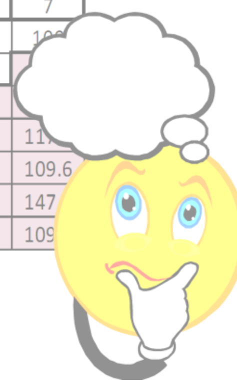
The problem: stochastic reserving

- Outstanding liabilities are impacted by two types of delay during the claims process:
  - **Reporting** delay
  - **Settlement** delay

- Main objectives in stochastic reserving:
  - Produce **point forecasts** for the outstanding reserve and cash flows
  - Produce accompanying **distributions**

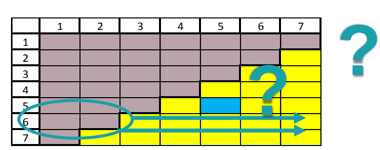
How to solve the problem?

	1	2	3	4	5	6	7
1	2200	1500	1000	650	300	150	10[illegible]
2	1900	1400	900	550	250	145	[illegible]
3	2300	1700	1200	750	400	175.9	[illegible]
4	3000	1800	950	500	369.9	183.4	11[illegible]
5	2700	1500	1000	641.8	345.8	171.4	109.6
6	3400	2200	1414.0	865.7	466.4	231.2	147[illegible]
7	2500	1629.0	1042.6	638.3	343.9	170.5	109[illegible]



The chain ladder method (CLM)

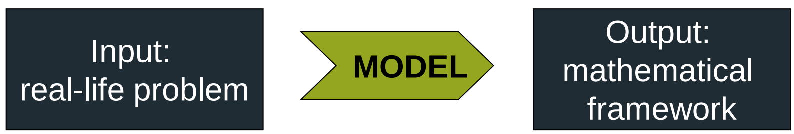
- CLM is the most popular **method** for calculating loss reserves:
  - simplicity and intuitive appeal
  - operates from one run-off triangle (payments, incurred)
- But CLM suffers from several drawbacks:
  - **Unstable** estimates
  - No information about the **tail**
  - Unable to separate **RBNS** and **IBNR** claims



- To address these **limitations** we need a **statistical model**
- We propose the **double chain ladder model**

The advantages of having a firm statistical model

- A model is a mathematical framework that completely describes a real-life problem



- It translates a real-life problem into a language which we, as mathematicians, can understand and work with

The model in practice

- The models are estimated from the data to provide solutions to the problem in practice
- Data requirements could make a model infeasible in practice
- Prior knowledge when it is available can be incorporated to:
  - provide more realistic and stable predictions
  - consider in practice more complicated models

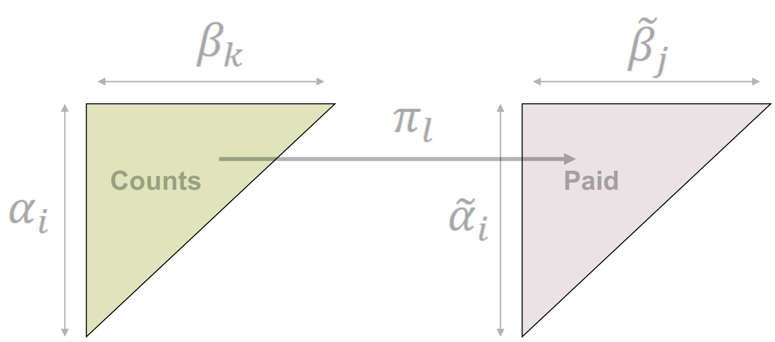
Summary

- The problem of stochastic reserving includes many dependencies
- These are implicit within the chain ladder method
- They will be made explicit in the double chain ladder model
- We need a method to estimate the model from the available information (data+prior knowledge)

CL predictions for payments

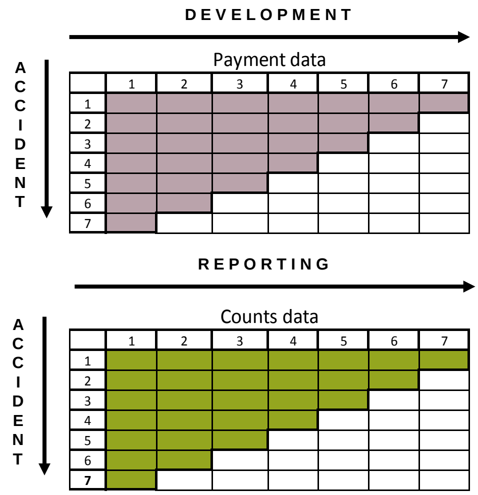
	1	2	3	4	5	6	7
1	2200	1500	1000	650	300	150	100
2	1900	1400	900	550	250	145	88.7
3	2300	1700	1200	750	400	175.9	112.5
4	3000	1800	950	500	369.9	183.4	117.3
5	2700	1500	1000	641.8	345.8	171.4	109.6
6	3400	2200	1414.0	865.7	466.4	231.2	147.9
7	2500	1629.0	1042.6	638.3	343.9	170.5	109.0

The double chain ladder model (DCL)



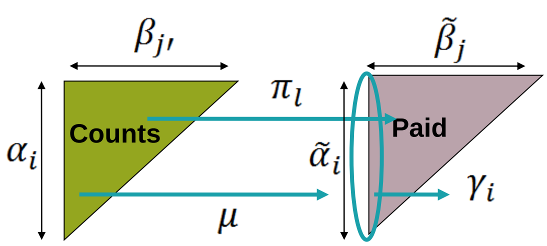
The modelled data: two run-off triangles

- We model annual/quarterly data triangles
  - Incremental aggregated payment data
  - Incremental aggregated counts data, which is assumed to have fully run off



12

The parameters involved in the model



Chain ladder parameters:

Ultimate claim numbers: α_i

Reporting delay: β_j,

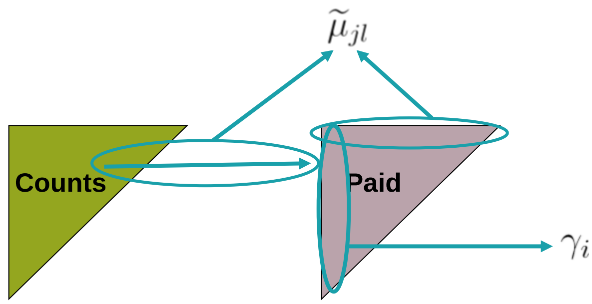
Development delay: $\tilde{\beta}_j$

Ultimate payment numbers: $\tilde{\alpha}_i$

Settlement delay π_l

13

The inflation parameters involved in the model



Inflation parameters:

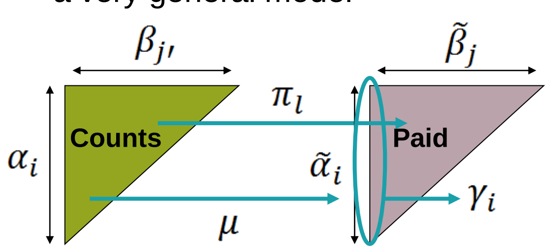
$\widetilde{\mu}_{jl}$ dependency on reporting delay and settlement delay

γ_i dependency on accident year

Individual payment mean = $\widetilde{\mu}_{jl}$ x γ_i

The DCL method to estimate the model

The DCL method allows to estimate a simple version of a very general model



Ultimate claim numbers: α_i

Reporting delay: $\beta_{j'}$

Settlement delay: π_l

Development delay: $\tilde{\beta}_j$

Ultimate payment numbers: $\tilde{\alpha}_i$

Severity inflation: γ_i

Individual payment mean in first period: μ

The DCL method: CLM twice

- Apply **CLM to counts and paid data** to get estimates

$$\widehat{\alpha}_i \quad \widehat{\beta}_j \quad \widehat{\widetilde{\alpha}}_i \quad \widehat{\widetilde{\beta}}_j$$

- Reminder:
  - Estimates of the row and column parameters in the Poisson chain ladder model $$\mathrm{E}[X_{ij}] = \widetilde{\alpha}_i \widetilde{\beta}_j \qquad \left(\sum_{j=0}^{m-1} \widetilde{\beta}_j = 1 \right)$$
  - Can be **calculated from development factors and ultimate claims**
  - They have a **clear interpretation** in the model:
  - $\widehat{\widetilde{\alpha}}_i$ (expected) total claim amount for the i[th] underwriting period
  - $\widehat{\widetilde{\beta}}_j$ proportion of total payments settled with *j* periods delay

The DCL method: a toy example

1. Apply CLM to counts data to get the estimates $\hat{\alpha}_i, \hat{\beta}_j,$

Estimated Counts

	1	2	3	4	5	6	7
1	227.30	104.53	38.63	10.48	2.52	1.57	1.00
2	208.40	95.84	35.42	9.61	2.31	1.44	0.92
3	195.00	89.69	33.14	8.99	2.16	1.34	0.86
4	280.40	128.98	47.66	12.93	3.11	1.93	[illegible]
5	236.20	108.64	40.15	10.89	2.62	1.63	1.04
6	287.70	132.32	48.90	13.26	3.19	1.98	1.27
7	240.00	110.38	40.79	11.06	2.66	1.65	1.06

Ultimate claims

$\hat{\alpha}$
386
353
331
476
401
489
408

$\hat{\beta}$

0.589	0.271	0.1	0.027	0.007	0.004	0.003

Development factors: F_j

1.4599	1.1164	1.0283	1.0066	1.0041	1.0026

$$\boxed{\widehat{\beta}_0 = \frac{1}{\prod_{l=1}^{m-1} F_l} \qquad \widehat{\beta}_j = \frac{F_j - 1}{\prod_{l=j}^{m-1} F_l}}$$

The DCL method: estimating the parameters

2. Apply CLM to the payment data to obtain the estimates $\widehat{\widehat{\alpha}}_i, \widehat{\widehat{\beta}}_j$

Estimated Payments

	1	2	3	4	5	6	7
1	2,292.8	1,494.0	956.2	585.4	315.4	156.3	100.0
2	2,033.8	1,325.3	848.2	519.3	279.8	138.7	88.7
3	2,579.7	1,681.0	1,075.8	658.7	354.9	175.9	112.5
4	2,689.4	1,752.4	1,121.6	686.7	369.9	183.4	117.3
5	2,513.7	1,638.0	1,048.3	641.8	345.8	171.4	109.6
6	3,390.6	2,209.4	1,414.0	865.7	466.4	231.2	147.9
7	2,500.0	1,629.0	1,042.6	638.3	343.9	170.5	109.0

Ultimate numbers

$\hat{\alpha}$

5900
5233
6638
6920
6468
8725
6433

$\hat{\beta}$

0.389	0.253	0.162	0.099	0.053	0.026	0.017

Development factors: F_j

1.6516	1.2525	1.1234	1.0592	1.0277	1.0172

$$\widehat{\beta}_0 = \frac{1}{\prod_{l=1}^{m-1} F_l} \qquad \widehat{\beta}_j = \frac{F_j - 1}{\prod_{l=j}^{m-1} F_l}$$

The DCL method: estimating the parameters

- The remaining parameters are **estimated so that the model can reproduce exactly the CLM** predictions:
- This is just a linear system

$\hat{\beta}$

0.589	0.271	0.1	0.027	0.007	0.004	0.003

$\hat{\alpha}$

386
353
331
476
401
489
408

$\widehat{\widehat{\beta}}$

0.389	0.253	0.162	0.099	0.053	0.026	0.017

$\widehat{\widehat{\alpha}}$

5900
5233
6638
6920
6468
8725
6433

$$\widehat{\alpha}_i \mu \gamma_i = \widehat{\widehat{\alpha}}_i \rightarrow \widehat{\gamma}_i \quad \widehat{\mu}$$

$$\sum_{l=0}^{j} \widehat{\beta}_{j-l} \pi_l = \widehat{\widehat{\beta}}_j \rightarrow \widehat{\pi}_l$$

The DCL method: estimating the parameters

- The resulting estimates:

	l= 1	2	3	4	5	6	7
$\hat{\pi}_l$	0.66	0.127	0.105	0.068	0.028	0.010	0.002

	i= 1	2	3	4	5	6	7
$\hat{\gamma}_i$	1	0.967	1.311	0.951	1.055	1.168	1.033

$\hat{\mu}$	15.28

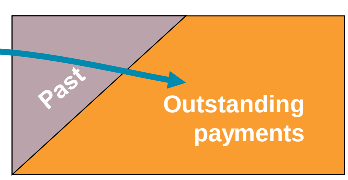
- Reminder:
  - π_l the proportion of claims settled *l* periods after reporting
  - γ_i the severity inflation in the *i*th accident period
  - μ the mean of individual payments in the first accident period

The DCL method: forecasting the reserve

Estimated Counts

	1	2	3	4	5	6	7
1	227.30	104.53	38.63	10.48	2.52	1.57	1.00
2	208.40	95.84	35.42	9.61	2.31	1.44	0.92
3	195.00	89.69	33.14	8.99	2.16	1.34	0.86
4	280.40	128.98	47.66	12.93	3.11	1.93	1.23
5	236.20	108.64	40.15	10.89	2.62	1.63	[illegible]
6	287.70	132.32	48.90	13.26	3.19	[illegible]	1.27
7	240.00	110.38	40.79	11.06	2.66	1.65	1.06

$\hat{\pi}_l \hat{\gamma}_i \hat{\mu}$



- We predict **RBNS and IBNR reserve** separately
- The **prediction formula** (the **mean** of future payments in the **model**)

$$E\left[X_{ij}\right] = \mu \alpha_i \gamma_i \sum_{l=0}^{j} \beta_{j-l}$$

Forecasting the RBNS claims

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1	227.3	104.5	38.63	10.48	2.523	1.565	1
2	208.4	95.84	35.42	9.606	2.313	1.435	
3	195	89.69	33.14	8.99	2.164		
4	280.4	129	47.66	12.93			
5	236.2	108.6	40.15				
6	287.7	132.3					
7	240						

$\widehat{\pi}_l \; \widehat{\gamma}_i \; \widehat{\mu}$

RBNS Estimations

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13
1								36	13	4	1	0	0
2							80	30	10	3	1	0	
3						58	98	35	11	2	0		
4					340	159	97	33	10	2			
5				526	296	130	79	26	7				
6			838	599	307	116	82	27					
7		479	397	259	107	36	43						

- RBNS claims contribute to cells to the right of the paid data
- We predict RBNS reserve using estimated parameters and estimated count data from the upper triangle.
- RBNS point prediction for cell (i,j): $$\widehat{X}_{ij}^{rbns} = \sum_{l=i-m+j}^{\min(j,d)} \widehat{N}_{i,j-l}\widehat{\pi}_l\widehat{\mu}\widehat{\gamma}_i$$

Worked example

- For illustration, we focus on payments in cell (1,11)

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1	227.3	104.5	38.63	10.48	2.523	1.565	1
2	208.4	95.84	35.42	9.606	2.313	1.435	
3	195	89.69	33.14	8.99	2.164		
4	280.4	129	47.66	12.93			
5	236.2	108.6	40.15				
6	287.7	132.3					
7	240						

RBNS Estimations

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2													
3													
4													
5													
6													
7													

- RBNS estimation for (1,11) comes from reported counts in the previous six years:
  - We have chosen a maximum delay of six years

Worked example

Estimated Count Data

$i \setminus j$	1	2	3	4	5	6	7
1	227.3	104.5	38.63	10.48	2.523	1.565	1
2	208.4	95.84	35.42	9.606	2.313	1.435	
3	195	89.69	33.14	8.99	2.164		
4	280.4	129	47.66	12.93			
5	236.2	108.6	40.15				
6	287.7	132.3					
7	240						

$\boxed{2.523} \times \hat{\pi}_6$

$= 2.523 \times 0.0011$
$= 0.0028$

- Consider the counts from six years ago – cell (1,5)
- Multiply by $\hat{\pi}_6$, which represents the proportion of claims for which a payment is made after six years
- Gives an estimate for the number of claims reported six years ago that contributes to our cell (1,11)

Worked example

Estimated Count Data

$i \setminus j$	1	2	3	4	5	6	7
1	227.3	104.5	38.63	10.48	2.523	1.565	1
2	208.4	95.84	35.42	9.606	2.313	1.435	
3	195	89.69	33.14	8.99	2.164		
4	280.4	129	47.66	12.93			
5	236.2	108.6	40.15				
6	287.7	132.3					
7	240						

$$\boxed{2.523} \times \hat{\pi}_6 + \boxed{1.565} \times \hat{\pi}_5 + \boxed{1} \times \hat{\pi}_4 = 0.046$$

- Proceed in the same way to find estimates for the number of claims reported four and five years ago that contributes to our cell (1,11)
- Sum to get the total estimate of the number of claims that contribute to (1,11)

Worked example

- We've estimated the total number of claims that contribute to (1,11) as 0.046
- Now we multiply by $\widehat{\mu}$ x $\widehat{\gamma}_1$, which represents the mean severity payment for claims which occurred in the first accident period
- This gives us our RBNS estimation for cell (1,11):

$$0.046 \times \widehat{\mu} \times \widehat{\gamma}_1 = 0.710$$

26

Estimating the IBNR claims

- Since the accidents are not reported yet, the IBNR reserves are derived from the lower triangle
- This fills in the paid triangle in the purple highlighted section:

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1							
2							0.9
3						1.3	0.9
4					3.1	1.9	1.2
5				10.9	2.6	1.6	1.0
6			48.9	13.3	3.2	2.0	1.3
7		110.4	40.8	11.1	2.7	1.7	1.1

Estimated Payment Data

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2							8.9	1.7	1.4	0.9	0.4	0.1	0.2
3						17.8	14.8	5.0	3.6	1.9	0.7	0.5	0.2
4					29.8	24.2	20.1	8.3	5.1	2.5	1.1	0.5	0.2
5				115.9	50.1	41.1	30.8	12.7	6.4	4.3	1.2	0.5	0.2
6			576.3	266.8	159.2	115.1	66.3	25.5	18.7	5.8	1.6	0.6	0.3
7		1,149.8	645.3	379.4	236.4	[illegible]	65.5	38.8	13.8	4.3	1.2	0.5	0.2

- IBNR point prediction for cell (i,j) : $\widehat{X}_{ij}^{ibnr} = \sum_{l=0}^{\min(i-m+j-1,d)} \widehat{N}_{i,j-l}\widehat{\pi}_l\widehat{\mu}\widehat{\gamma}_i$

27

Worked example

- For illustration, we focus on payments in cell (3,11)

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1							
2							0.9
3						1.3	0.9
4					3.1	1.9	1.2
5				10.9	2.6	1.6	1.0
6			48.9	13.3	3.2	2.0	1.3
7		110.4	40.8	11.1	2.7	1.7	1.1

Estimated Payment Data

i \ j	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2													
3													
4													
5													
6													
7													

- IBNR estimation for (3,11) comes from incurred but not reported counts in the previous six years:
  - We have chosen a maximum delay of six years

28

Worked example

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1							
2							0.9
3						1.3	0.9
4					3.1	1.9	1.2
5				10.9	2.6	1.6	1.0
6			48.9	13.3	3.2	2.0	1.3
7		110.4	40.8	11.1	2.7	1.7	1.1

$\boxed{3.1} \times \widehat{\pi}_6$

$= 3.1 \times 0.0011$
$= 0.0034$

- Consider the counts from six years ago – cell (3,5)
- Multiply by $\widehat{\pi}_6$, which represents the proportion of claims for which a payment is made after six years
- Gives an estimate for the number of claims reported six years ago that contributes to our cell (3,11)

29

Worked example

Estimated Count Data

i \ j	1	2	3	4	5	6	7
1							
2							0.9
3						1.3	0.9
4					3.1	1.9	1.2
5				10.9	2.6	1.6	1.0
6			48.9	13.3	3.2	2.0	1.3
7		110.4	40.8	11.1	2.7	1.7	1.1

$$\boxed{3.1} \times \widehat{\pi}_6 + \boxed{1.9} \times \widehat{\pi}_5 + \boxed{1.2} \times \widehat{\pi}_4 \quad = 0.056$$

• Proceed in the same way to find estimates for the number of claims reported four and five years ago that contributes to our cell (3,11)

• Sum to get the total estimate of the number of claims that contribute to (3,11)

30

Worked example

- We've estimated the total number of claims that contribute to (3,11) as 0.056
- Now we multiply by $\widehat{\mu}$ x $\widehat{\gamma}_3$, which represents the mean severity payment for claims which occurred in the third accident period
- This gives us our RBNS estimation for cell (1,11):

$$0.056 \times \widehat{\mu} \times \widehat{\gamma}_3 = 1.122$$

31

Rediscovering the chain ladder mean in our model

- Under the DCL model, the mean of the incremental payments for accident period *i* and development delay *j* is given by:

$$E[X_{ij}] = \mu\alpha_i\gamma_i \sum_{l=0}^{j} \beta_{j-l}$$

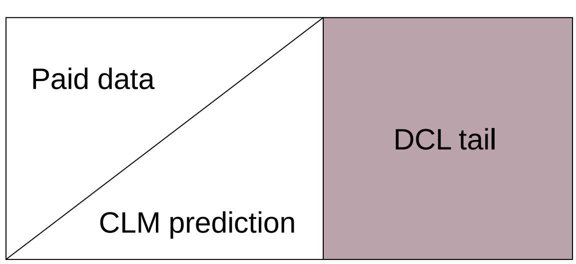
- But this is exactly the chain ladder mean

$$E[X_{ij}] = \tilde{\alpha}_i\tilde{\beta}_j$$

- DCL breaks down classical chain ladder in its components

Predicting the tail through DCL

- With CLM, when a triangle has not run-off one needs to fit a tail
- DCL provides the tail prediction as an intrinsic part of the model



The predicted reserves trough DCL

RBNS Predictions

i \ j	1	2	3	4	5	6	7
1							
2							79.8
3						158.1	97.8
4					340.1	159.1	97.2
5				525.9	295.7	130.3	78.8
6			837.7	598.9	307.2	116.1	81.6
7		479.2	397.3	258.9	107.5	36.2	43.5

IBNR Predictions

i \ j	1	2	3	4	5	6	7
1							
2							9.9
3						14.4	12.0
4					29.8	24.2	20.1
5				115.9	50.1	41.1	30.8
6			576.3	266.8	159.2	115.1	66.3
7		1,149.8	645.3	379.4	236.4	134.3	65.5

….Or by accident year

Total Predictions

i \ *j*	1	2	3	4	5	6	7
1							
2							89.7
3						173.9	112.5
4					369.9	183.4	117.3
5				641.8	345.8	171.4	109.6
6			1,414.0	865.7	466.4	231.2	147.9
7		1,629.0	1,042.6	638.3	343.9	170.5	109.0

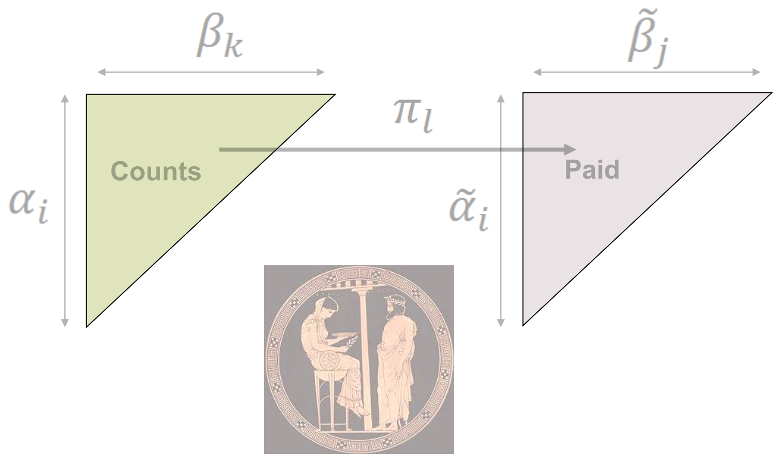
Cash flow by calendar year

….Or the total reserve

Summary

- DCL is a firm statistical model which breaks down the chain ladder estimates into their individual components
- Practical advantages:
  - Intrinsic tail estimation
  - Separates RBNS and IBNR reserves
  - among others…
- Two **limitations** of DCL:
  - DCL suffers from the same **instability** as classical CLM
  - From 2 triangles can estimate a **simple distributional model** (severity only changes in the accident year)
- Next we overcome these limitations **adding prior knowledge**

Adding prior knowledge to DCL to correct the instability



The DCL method with real data

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	52	513	748	555	426	212	213	16	9	37	0	0	0	0	0	0	0	0	0
2	144	1006	910	736	593	766	615	245	116	15	36	165	15	67	6	11	0	0	
3	346	1467	1292	1237	1127	779	392	845	94	230	12	11	21	84	10	17	0		
4	408	1875	1810	1860	1806	1422	762	307	110	140	53	37	0	7	0	0			
5	712	3254	2696	2593	3377	2101	923	435	124	30	23	0	59	31	12				
6	941	3615	3274	4479	3841	2033	1242	472	120	59	5	0	9	0					
7	1221	5814	5905	7112	5321	2426	857	197	134	40	12	66	99						
8	1685	8164	7609	7722	6298	1981	830	580	198	124	64	29							
9	2253	9480	7697	8260	5872	2340	1099	363	147	44	14								
10	2043	8792	9169	7864	5895	1978	722	245	60	-1									
11	1570	9962	9670	8024	6121	2392	618	98	71										
12	1456	9182	8262	8374	4995	1886	883	241											
13	1129	7676	8515	6467	4505	1502	461												
14	1381	11548	8890	7964	4951	1980													
15	2196	12381	10391	7516	4969														
16	2068	14179	11164	7740															
17	1747	11600	8808																
18	3295	15210																	
19	4664																		

CLM	DCL
-	-
-	-
-	-
-	-
17	17
35	35
138	138
245	245
352	352
394	394
552	552
684	684
1,050	1,050
2,536	2,536
5,737	5,737
14,089	14,089
21,006	21,006
44,688	44,688
98,972	98,972
190,496	**190,496**

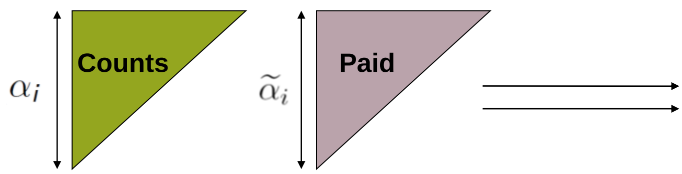
- For recent accident periods, we wish to improve our estimates using prior knowledge. To do this, we turn to the Bornhuetter-Ferguson method.

Understanding the problem through DCL

- The model breaks down the chain ladder estimates into their individual components

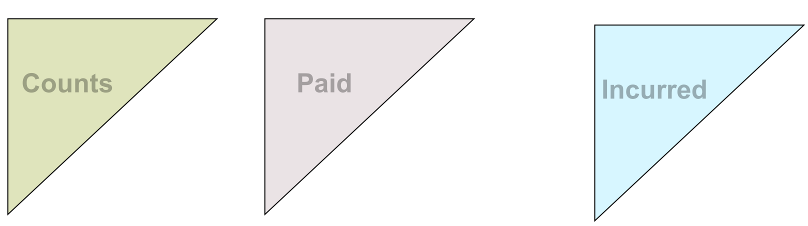
$$\tilde{\alpha}_i = \mu \alpha_i \gamma_i$$

- The instability comes from the estimation of the severity inflation in the underwriting period

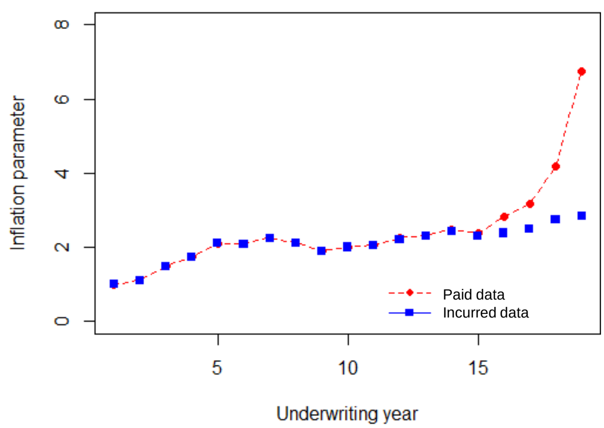


$$\boxed{\gamma_i = \frac{\tilde{\alpha}_i}{\mu \alpha_i}}$$

Bornhuetter-Ferguson and Double Chain Ladder



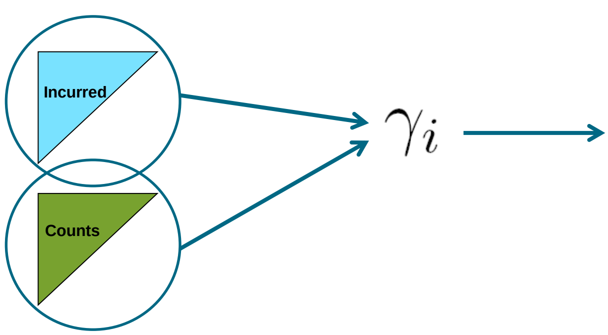
Estimating the inflation using incurred data



- The instability comes from the estimates of the inflation (see the last 4 years)
- The estimates from the incurred data are much more stable
- First correction: estimate the inflation trough DCL from incurred data: **BDCL**

The BDCL method

- The BDCL method takes a more realistic estimation of the inflation parameter from the incurred data



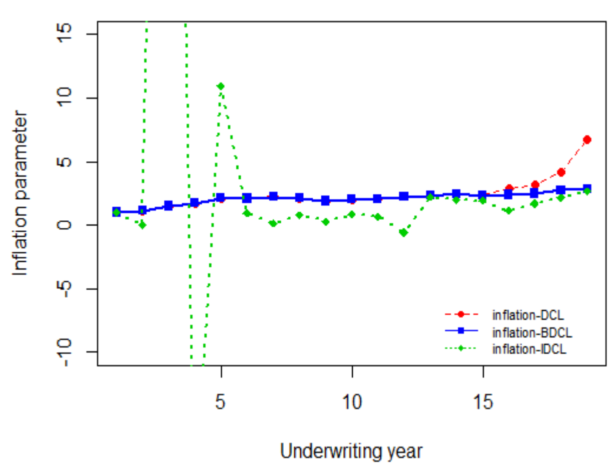
CLM	DCL	BDCL
-	-	-
-	-	-
-	-	-
-	-	-
17	17	17
35	35	35
138	138	138
245	245	244
352	352	349
394	394	392
552	552	547
684	684	668
1,050	1,050	1,046
2,536	2,536	2,503
5,737	5,737	5,564
14,089	14,089	11,848
21,006	21,006	16,469
44,688	44,688	29,435
98,972	98,972	41,844
190,496	**190,496**	**111,098**

The BDCL method: comparisons

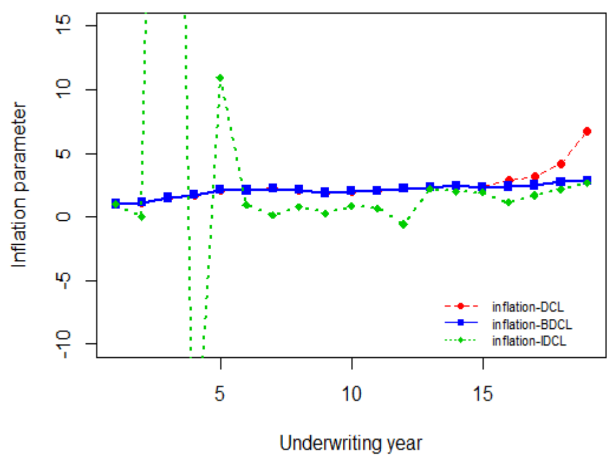
paid CLM	DCL	BDCL	inc CLM
-	-	-	-
-	-	-	-
-	-	-	4
-	-	-	-10
17	17	17	20
35	35	35	14
138	138	138	9
245	245	244	91
352	352	349	53
394	394	392	156
552	552	547	175
684	684	668	-158
1,050	1,050	1,046	930
2,536	2,536	2,503	2,029
5,737	5,737	5,564	4,432
14,089	14,089	11,848	6,255
21,006	21,006	16,469	11,684
44,688	44,688	29,435	23,119
98,972	98,972	41,844	39,152
190,496	**190,496**	**111,098**	**87,956**

- The total reserve is significantly lower using the incurred data to estimate the inflation
- BDCL does not give us the same reserve as CLM with incurred.
- BDCL is a method which seems to give results in the middle
- We next define a third method to reproduce CLM with incurred: the IDCL method

The IDCL method for our model



The IDCL method providing CLM with incurred



inc

CLM	IDCL
-	-
-	-
4	3
-10	-30
20	20
14	14
9	9
91	91
53	53
156	156
175	175
-158	-158
930	930
2,029	2,029
4,432	4,432
6,255	6,255
11,684	11,684
23,119	23,119
39,152	39,152
87,956	**87,934**

The IDCL method: forecasting RBNS/IBNR reserve

Reserve by accident year:

RBNS	IBNR	Total	inc CLM
-	-	-	-
-	-	-	-
3	-	3	4
-30	-	-30	-10
20	-	20	20
13	1	14	14
9	0	9	9
89	2	91	91
52	1	53	53
151	5	156	156
169	6	175	175
-151	-7	-158	-158
889	42	930	930
1,953	76	2,029	2,029
4,319	113	4,432	4,432
6,146	110	6,255	6,255
11,410	274	11,684	11,684
22,127	992	23,119	23,119
29,311	9,842	39,152	39,152
76,478	**11,456**	**87,934**	**87,956**

Reserve by calendar year:

RBNS	IBNR	Total
28,692	562	29,254
19,487	3,005	22,493
14,062	2,295	16,357
7,705	2,252	9,957
2,853	1,675	4,528
1,401	721	2,122
708	400	1,108
367	209	576
350	93	443
232	73	306
259	33	291
184	43	226
138	32	170
22	40	62
36	7	43
-16	16	-
1	-1	-
-1	1	-
76,478	**11,456**	**87,934**

The IDCL method providing the tail

Tail reserve:

RBNS	IBNR	Total
40	-	40
79	-	79
115	-	115
153	-	153
275	-	275
46	68	114
11	16	27
62	129	191
24	58	82
36	176	212
15	164	179
13	-133	-120
-157	557	400
-288	741	453
-327	809	481
-244	559	
-203	549	346
-277	749	-
-	591	591
-628	**5,034**	**3,619**

Total reserve:

Total
40
79
115
153
20,434
13,739
8,841
91,523
52,789
155,779
175,386
-158,193
930,688
2,029,406
4,432,856
6,255,693
11,683,760
23,119,121
39,152,239
87,964,447

Summary

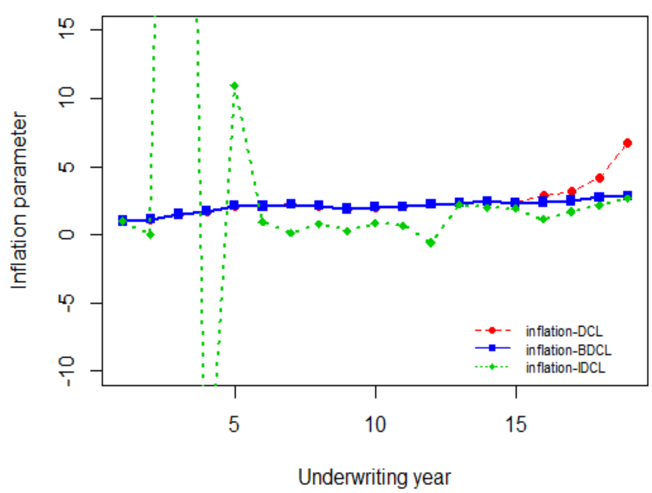
- Double Chain Ladder Method (**DCL**)
  - 2 run-off triangles – counts and paid
  - Reproduces paid CLM results
  - Unstable, particularly regarding data from recent accident years
- Bornhuetter-Ferguson Double Chain Ladder Method (**BDCL**)
  - 3 run-off triangles – counts, paid and incurred
  - Very stable
  - Does not produce the same reserve as CLM with incurred

Summary

- Incurred Double Chain Ladder Method (**IDCL**)
  - 3 run -off triangles – counts, paid and incurred
  - Reproduces incurred CLM reserve by accident year
  - Allows predictions to be analysed by calendar year

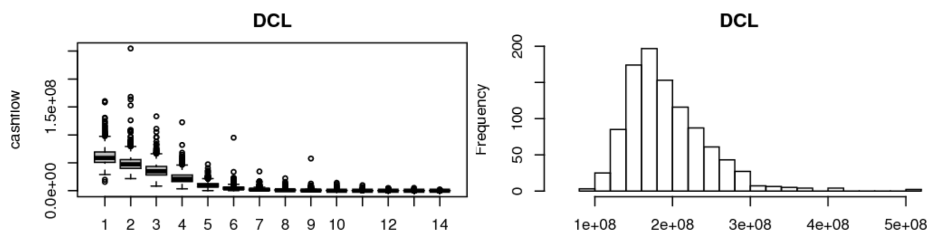
- All 3 methods allow us to evaluate the tail, and to separate IBNR and RBNS claims and perform a proper validation of results
- Also from them we can derive the distribution of the possible cash flows

Summary



- But how to choose among DCL, BDCL and IDCL?
- Apply validation techniques
  - Work in progress with RSA

The distribution of cash flows



Deriving the distribution of cash flows

- Point predictions only requires general assumptions about the mean
- We have shown that chain ladder mean is very rich: the DCL model shows that it allows many dependencies
- When it comes to the estimation the available information is that restrict the generality of the model
- From two triangles (counts and payments) we can estimate the simplest DCL model
- Next we go through the model and try to generalize it incorporating prior knowledge

A full statistical model: distributional assumptions

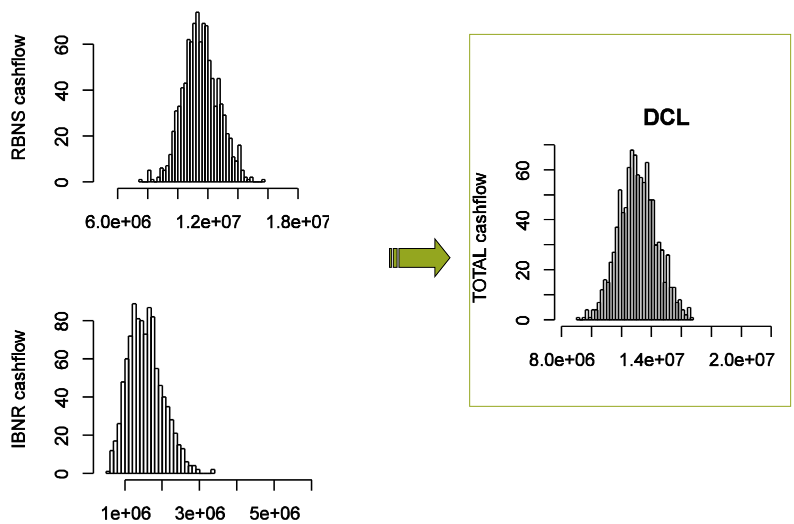
- The simplest DCL model assumes that the mean of individual payments only depends on the accident year: $\mu \times \gamma_i$
- To derive the distribution it only introduces a single new parameter: the variance of the individual payments
- The following statistical distributions are assumed for each of the components in the model:

Component	Distribution
Count data	Poisson
RBNS delay	Multinomial
Individual payments	Gamma

Bootstrap methods to provide the distribution

- The variance is estimated using over-dispersion arguments
- The distribution of cash flows are provided simulating the future payments from the estimated distributions: **parametric bootstrap**
- Thus we can derive empirical distributions of:
  - The cash flows
  - The total reserve
  - Separating RBNS and IBNR claims and including the tail

Empirical illustration



54

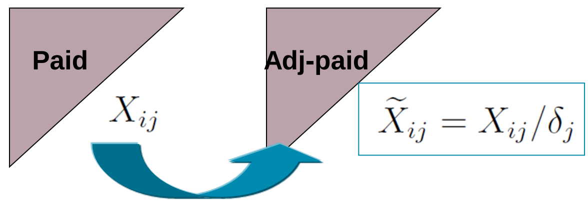
More general distributional model

- We assume that the individual payments have mean depending also on development severity inflation. It involves new parameters for each development period: $\boxed{\delta_j}$
- Also the distribution of the severity is a mixed distribution with probability of zero-claims for each accident period $\boxed{Q_i}$
- The **severity mean** is: $\boxed{\mu\gamma_i\delta_j(1-Q_i)}$

and the variance: $\gamma_i^2\delta_j^2(1-Q_i)\left(\sigma^2+Q_i\mu^2\right)$

Adding prior information about severity inflation

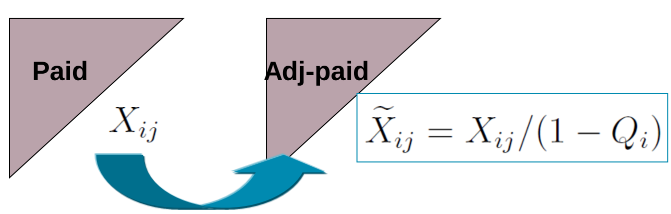
- Assume prior knowledge: $\delta_0, \delta_1, \ldots, \delta_{m-1}$
- To incorporate this information into DCL:
  1. Adjust the incremental payments by this inflation



  2. Apply DCL to the adjusted triangle
  3. Revert the adjustment on the predictions (point and distribution)

Adding prior information zero-claims

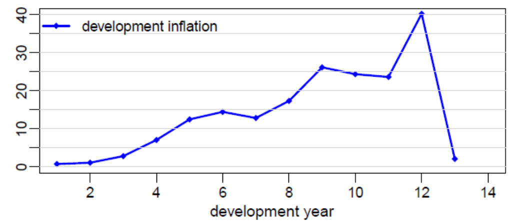
- Assume prior knowledge: $Q_1, Q_2 \ldots, Q_m$
- To incorporate this information into DCL:
  1. Adjust the incremental payments by this inflation



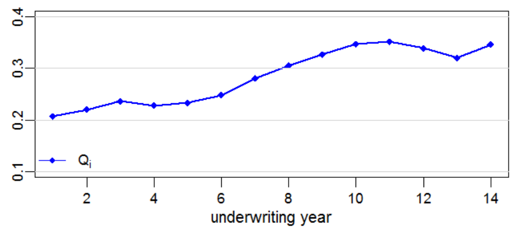
  2. Apply DCL to the adjusted triangle
  3. Simulate the distribution using the prior knowledge
  4. Revert the adjustment on the predictions (point and distribution)

Empirical illustration

- Prior A: we know that the severity development inflation is

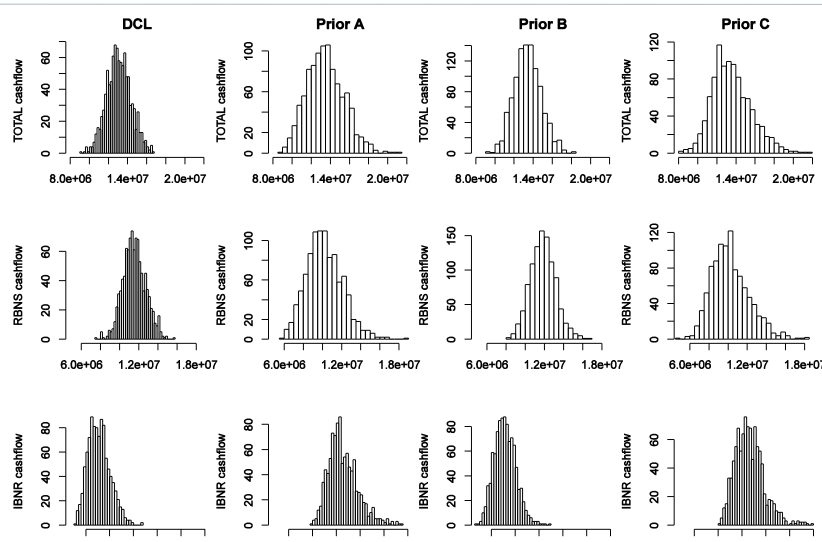


- Prior B: we know that probability of zero claims for each underwriting year is



- Prior C: we incorporate both A and B at the same time

Distribution of the total reserve (RBNS/IBNR)



Some comments about the results

- Adding these two types of prior knowledge does not change the best estimate very much– much like double chain ladder itself
- However, it has consequences on the distribution
- In this example the distribution around the best estimate is affected with a tendency towards thicker tails, when adding in this extra information
- Further extensions:
  - Other kind of prior information can be added in a similar way such as settlement delay inflation or calendar year effect
  - The prior knowledge about severity inflation and zero-claims can be estimated in practice using the DCL method on additional triangles such as the number of payments

Concluding remarks

- Double chain ladder (DCL) is a fully consistent statistical model that is able to replicate exactly classical chain ladder on paid and incurred data.
- Provides a natural method for introducing prior knowledge in a consistent way
- Using prior knowledge from incurred we have solved the instability of chain ladder: Bornuetter-Fergurson
- Claims inflation and zero claims probabilities can be added to DCL as prior knowledge in a surprisingly straightforward way

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.