

GUY CARPENTER



Robustifying Reserving

GARY G. VENTER



Robust Methods



Classical view



- Data is generated as a sample from model process being fitted
- Efficiency of methods like MLE come from this view

Problems



- Could be a more complex process that is generating the data and model is a convenient simplification
- Even a few points generated by a different process can throw off the estimated parameters

Responses



- Identify and exclude outliers
 - Try to understand when outliers arise and not use model in those circumstances
- Try to find models that are not so influenced by those points

GUY CARPENTER

1

Influence



Excluding points —»

- Look at change in parameters from leaving out observations
 - Done for each point
 - Called empirical influence function
- Sample size times change from excluding a point is called gross error sensitivity (GES)
- Look for estimators with low GES but close to efficiency of MLE

Changing points —»

- Look at change in parameters or predictions from changing a point
- E.g., take the derivative of the prediction with respect to each point
 - If the points have a lot of randomness, a point with strong effect will have strong effect from its random component

Reserving Application



Effect of changes —»

- Leaving out cells can be awkward so look at derivative of reserve wrt each point in triangle
 - Called impact of the cell on the reserve
- From Tampubolon PhD thesis
- Examples from previous CAS papers

Methodology —»

- Derivatives usually done numerically
 - Redo reserve estimate after small change in cell

GDFs —»

- Also look at generalized degrees of freedom
 - Change in fitted value for a cell wrt observed value
 - A better measure of degrees of freedom than just counting parameters when model is non-linear
- GDFs may help understand impacts

General Observations



Chain ladder —»

- All 3 corners of triangle have fairly high impact
 - Lower left
 - All development factors apply to it
 - Impact = cumulative factor
 - Upper right
 - Development factor applies to all accident years
 - Upper right
 - Increasing it reduces all development factors
 - Impact is thus negative and perhaps large

Reducing Impacts



Upper right —»

- Trending and averaging factors in the tail
- Using additive constants for the final lags
 - Both useful as individual factors rarely significant at the end

Lower left —»

- Consider alternatives to chain ladder
 - Cape Cod method models all accident years at same level
 - E.g. for on-level loss ratios
 - Intermediate models might have just a few accident year levels

A triangle

L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
11,305	18,904	17,474	10,221	3,331	2,671	693	1,145	744	112	40	13
8,828	13,953	11,505	7,668	2,943	1,084	690	179	1,014	226	16	616
8,271	15,324	9,373	11,716	5,634	2,623	850	381	16	28	558	
7,888	11,942	11,799	6,815	4,843	2,745	1,379	266	809	12		
8,529	15,306	11,943	9,460	6,097	2,238	493	136	11			
10,459	16,873	12,668	9,199	3,524	1,027	924	1,190				
8,178	12,027	12,150	6,238	4,631	919	435					
10,364	17,515	13,065	12,451	6,165	1,381						
11,855	20,650	23,253	9,175	10,312							
17,133	28,759	20,184	12,874								
19,373	31,091	25,120									
18,433	29,131										
20,640											

GUY CARPENTER

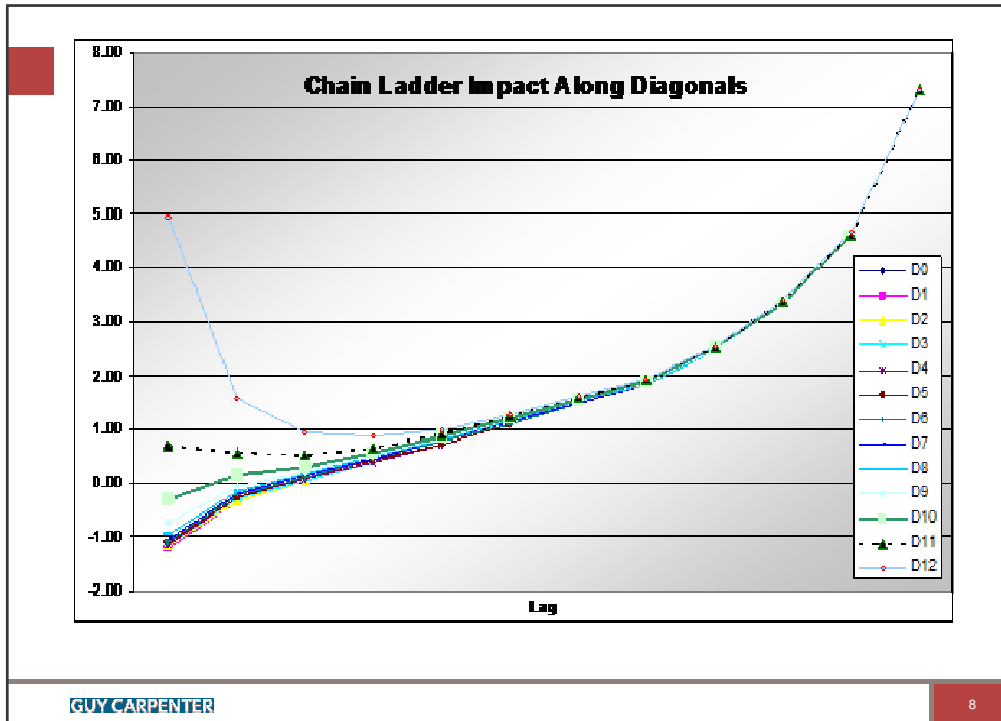
6

Chain Ladder Impacts

L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
-1.21	-0.34	0.04	0.39	0.73	1.10	1.48	1.85	2.46	3.35	4.61	7.31
-1.21	-0.34	0.04	0.39	0.73	1.10	1.48	1.85	2.46	3.35	4.61	7.31
-1.17	-0.29	0.08	0.44	0.78	1.14	1.53	1.89	2.51	3.39	4.66	
-1.15	-0.27	0.10	0.46	0.80	1.16	1.55	1.91	2.53	3.41		
-1.14	-0.27	0.11	0.46	0.80	1.17	1.56	1.92	2.54			
-1.10	-0.23	0.15	0.50	0.84	1.21	1.59	1.96				
-1.07	-0.20	0.18	0.53	0.87	1.24	1.62					
-1.03	-0.16	0.22	0.57	0.91	1.28						
-0.95	-0.08	0.30	0.65	0.99							
-0.73	0.14	0.52	0.87								
-0.31	0.57	0.95									
0.70	1.58										
4.95											

GUY CARPENTER

7



Regression model

Accident years —»

Lags —»

Diagonals —»

Residuals —»

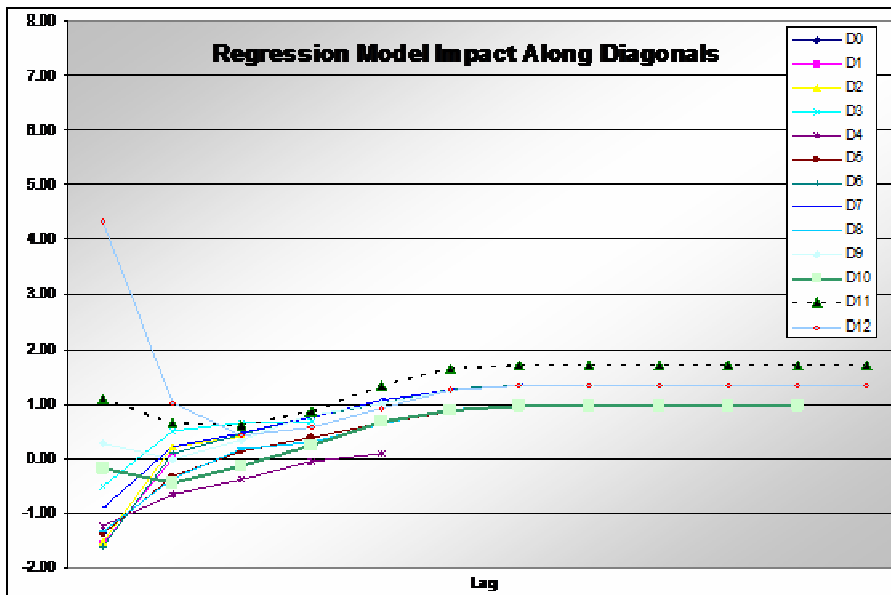
- All separate
- First 5 development factors
- Plus single additive constant for all cells
 - Picks up development after 5 also
- Effects included for 4th 5th 8th 10th and 11th diagonals
- IID normal
- Better fit than chain ladder

9

Regression Model Impacts

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10
AY0	-1.36	0.02	0.42	0.67	0.10	0.87	1.35	1.35	0.97	1.35	0.97
AY1	-1.56	0.22	0.66	-0.04	0.67	1.28	1.35	0.97	1.35	0.97	1.73
AY2	-1.53	0.52	-0.39	0.38	1.02	1.27	0.97	1.35	0.97	1.73	1.35
AY3	-0.51	-0.64	0.15	0.78	1.07	0.90	1.35	0.97	1.73	1.35	
AY4	-1.24	-0.31	0.45	0.76	0.64	1.27	0.97	1.73	1.35		
AY5	-1.38	0.11	0.47	0.32	1.00	0.89	1.73	1.35			
AY6	-1.61	0.22	0.18	0.80	0.68	1.66	1.35				
AY7	-0.89	-0.36	0.35	0.24	1.34	1.25					
AY8	-1.34	0.00	-0.12	0.87	0.94						
AY9	0.29	-0.44	0.61	0.57							
AY10	-0.18	0.66	0.43								
AY11	1.11	1.04									
AY12	4.31										

Regression Model Impact Along Diagonals



Problem of IID Normal Residuals



In general —»

- Not supported by data
- Not likely anyway

Alternatives tried —»

- Regression on square root of incremental values
- Gamma residuals with variance $\sim \text{mean}^{0.71}$.
 - Both had problems with high impacts

What worked —»

- Gamma with multiplicative diagonals
 - Before they were additive
- Gave better fit without problem of high impacts
- Impacts similar to model with IID normal residuals but with more realistic distribution of residuals
- Robust analysis showed weakness of alternatives

Impacts of CL and ODP on Taylor-Ashe Triangle

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	-3.11	-1.62	-1.01	-0.45	0.01	0.51	1.16	2.27	4.54	12.59
AY1	-2.87	-1.38	-0.77	-0.20	0.25	0.76	1.40	2.51	4.78	
AY2	-2.43	-0.93	-0.33	0.24	0.69	1.20	1.85	2.95		
AY3	-2.21	-0.72	-0.11	0.45	0.91	1.41	2.06			
AY4	-1.95	-0.46	0.15	0.71	1.17	1.67				
AY5	-1.67	-0.18	0.43	0.99	1.45					
AY6	-1.25	0.25	0.85	1.42						
AY7	-0.14	1.35	1.96							
AY8	2.07	3.57								
AY9	13.45									

Regression model



- Accident years** —»
 - Three levels: high, medium, low, plus average of high and medium
- Lags** —»
 - High and low levels of % of ultimate paid in cell
 - Average of high and low, and 1 – sum of others also used
- Diagonals** —»
 - Effects included for 4th 6th 7th diagonals
- Residuals** —»
 - Gamma with variance \propto mean²
 - Better fit than chain ladder or ODP

Impacts of Regression Model on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	0.65	-0.82	-1.08	-2.07	-0.87	0.97	-0.32	0.33	0.53	12.06
AY1	1.45	-0.02	0.68	0.60	-0.25	1.90	1.40	1.61	1.57	
AY2	1.64	0.75	-0.19	0.84	0.90	1.93	1.66	1.36		
AY3	1.26	0.43	-0.21	0.97	-0.36	1.70	1.71			
AY4	1.62	0.08	0.67	0.37	0.63	1.35				
AY5	1.19	-0.11	0.57	0.51	1.17					
AY6	2.56	1.19	0.91	1.13						
AY7	2.18	1.27	1.49							
AY8	1.72	0.92								
AY9	1.59									

TA Regression



Remaining problem —»

- Upper right

Alternate model —»

- Lag 9 gets half the % paid as low level
 - Consider as a trend to 0% for lag 10
- Still force lag factors to sum to 1.0
- Largest impact now 2.35, and only 2 above 2

Summary and Extensions

- Robust analysis looks for observations with high impact on result
- Problem in that random component would have high impact
- Derivative of reserve wrt each cell used as impact measure
- Add to list of model checks
- Led to finding improved models in example cases
- Possible extension: multiply impact by modeled standard deviation of cell estimate
 - Would combine impact of a small change with degree of change likely