Two Approaches to Calculating Correlated Reserve Indications Across Multiple Lines of Business

Gerald Kirschner
Deloitte Consulting

Casualty Loss Reserve Seminar

September 2008

Presentation structure

- Correlation – a general discussion
- Methodologies
- Case study
  - Description
  - Results using correlation matrix approach
  - Results using bootstrap approach
- Conclusion

What this presentation is not

- Presentation is not going to discuss how to simulate ranges of reserves for a single line of business

- The discussion here is of methodologies that might be appropriate for reserve analysis, but are probably not appropriate for Enterprise Risk Management or Capital Adequacy analyses

Correlation

- Correlation vs. causality
  - Correlation is a a way of measuring the "strength of relationship" between two sets of numbers.
  - Causality is the relation between a cause (something that brings about a result) and its effect.
- Ideally would want to directly model effects of causality, but not always able to do so

Effects of correlation

- Suppose we have two lines, A & B, whose reserve indications exhibit correlation

- Strength of the correlation is irrelevant if we only care about the mean reserve indication for A + B:

  $$\text{mean }(A + B) = \text{mean }(A) + \text{mean }(B)$$

- Strength of correlation matters when we look towards the ends of the distribution of (A+B).

Effects of correlation: example 1

- 2 lines of business, N (100,25)
- 75^{th} percentile of A+B at different levels of correlation between A and B:

Correlation	Values at 75^{th} percentile	Ratio of Values at 75^{th} percentile
0.00	223.8	0.0%
0.25	226.7	1.3%
0.50	229.2	2.4%
0.75	231.5	3.4%
1.00	233.7	4.4%

Effects of correlation: example 2

- Same idea, but increase variability of distributions for lines A and B:

	Standard Deviation Value			
	25	50	100	200
Value for 0.00 correlation at the 75th percentile	223.8	247.7	295.4	390.8
Correlation	Ratio of values at 75th percentile			
0.25	1.3%	2.3%	3.8%	5.8%
0.50	2.4%	4.3%	7.3%	11.0%
0.75	3.4%	6.2%	10.4%	15.8%
1.00	4.4%	8.0%	13.4%	20.2%

Correlation methodologies

- Method 1: relies on the user to specify a correlation matrix that describes the relative strength of relationship between the lines of business by analyzed. Will use rank correlation technique to develop a correlated reserve indication.
- Method 2: uses the bootstrap process to maintain any correlations that might be implicit in the historical data. No other information is needed to develop the correlated reserve indication.

Rank correlation example

Index	A	B
1	155	154
2	138	125
3	164	100
4	122	198
5	107	128

Perfect Inverse Correlation			No Correlation			Perfectly Correlated		
Rank to Use			**Rank to Use**			**Rank to Use**		
A	**B**		**A**	**B**		**A**	**B**	
5	4		1	1		5	3	
4	1		2	2		4	2	
2	5		3	3		2	5	
1	2		4	4		1	1	
3	3		5	5		3	4	
Resulting Joint Dist.			**Resulting Joint Dist.**			**Resulting Joint Dist.**		
A	**B**	**A+B**	**A**	**B**	**A+B**	**A**	**B**	**A+B**
107	198	305	155	154	309	107	100	207
122	154	276	138	125	263	122	125	247
138	128	266	164	100	264	138	128	266
155	125	280	122	198	320	155	154	309
164	100	264	107	128	235	164	198	362
Range of Joint Dist.			**Range of Joint Dist.**			**Range of Joint Dist.**		
Low		264	**Low**		235	**Low**		207
High		305	**High**		320	**High**		362

Method 1 approach

1. Generate N reserve indications for each line of business on a stand-alone basis
2. For each line, sort the N reserve indications from low to high
3. Determine (through other means) the relative relationships between the lines of business being modeled – i.e. the correlations
4. Enter correlation information into a correlation matrix
   - Correlation matrix must be symmetric and positive definite
5. Using Cholesky decomposition, create the lower triangular matrix from the original correlation matrix

Method 1 approach continued

6. For each line, generate N random values from a Normal (0,1) distribution, where N = # reserve indications produced for each line on a stand-alone basis. Combining the random values across the lines of business produces N vectors of random numbers
7. Multiply the Cholesky decomposed lower triangular matrix with each of the N vectors of random numbers to produce N correlated normal vectors
8. For each line, rank the value of each correlated normal result.
9. Use relative positioning of the correlated Normal draws as the basis for pulling values from the sorted table of uncorrelated reserve indications to create correlated reserve indications across the lines of business

Symmetric and Positive Definite matrix

- Correlation Matrix: Create a square matrix where the i,jth entry is the correlation coefficient between the ith and jth lines of business. This will be a symmetric matrix, i.e. the Lower Triangle of the Matrix will be the mirror image of the Upper Triangle about its diagonal.
- Positive Definite (PD): The matrix should be positive definite, i.e. loosely speaking - the matrix will have a positive determinant. This will allow the Cholesky decomposition of this matrix.

1.0000	0.9912	0.9000	0.9000
0.9912	1.0000	0.9000	0.9000
0.9000	0.9000	1.0000	0.9000
0.9000	0.9000	0.9000	1.0000

- $A_{ik} = A_{ki}$ for all i, k
- $A_{21} = 0.9912 = A_{12} = 0.9912$

- Determinant of the correlation matrix

= 0.000478 > 0 ➔ Correlation Matrix is PD

- Excel Application function: ***MDeterm*** calculates determinant of a matrix

Illustrative Example

Step 1: Generate N = 6 loss reserve indications for k = 3 lines of business

Reserve Indication				***Ranking***		
WC	***CAL***	***OL***	***Aggregate***	***WC***	***CAL***	***OL***
200	**2,000**	**60,000**	**62,200**	***2***	***2***	***6***
300	**5,000**	**30,000**	**35,300**	***3***	***5***	***3***
100	**6,000**	**40,000**	**46,100**	***1***	***6***	***4***
400	**3,000**	**20,000**	**23,400**	***4***	***3***	***2***
600	**1,000**	**50,000**	**51,600**	***6***	***1***	***5***
500	**4,000**	**10,000**	**14,500**	***5***	***4***	***1***

Step 2: For each line, sort the N = 6 reserve indications from low to high

Ranked Reserve Indication		
WC	***CAL***	***OL***
100	**1,000**	**10,000**
200	**2,000**	**20,000**
300	**3,000**	**30,000**
400	**4,000**	**40,000**
500	**5,000**	**50,000**
600	**6,000**	**60,000**

Illustrative Example *continued*

Step 3: Use extrinsic / intrinsic information to build a correlation matrix that defines the linear relationships among k = 3 lines. Assumed correlation matrix for this illustrative example is:

***R* =**

1.0000	0.9912	0.9000
0.9912	1.0000	0.9000
0.9000	0.9000	1.0000

Step 4: Symmetric and Positive Definiteness check of the Correlation Matrix (as described in slide 12)

Step 5: Cholesky's Decomposition: Creating a lower triangular matrix ***L** ; such that **L'L = R**.* We derive:

***L* =**

1.0000	0.0000	0.0000
0.9912	0.1324	0.0000
0.9000	0.0598	0.4318

Illustrative Example *continued*

Step 6 : For each line (k = 3), independently generate N = 6 random values from a *Normal(0,1)* distribution. Combine these realizations across the k lines of business to have N = 6 *uncorrelated* vectors.

Uncorrelated vector 1 →

.

.

.

Uncorrelated vector 6 →

Uncorrelated N(0,1)		
NORMINV(RAND(),0,1)		
Line 1	***Line 2***	***Line 3***
-0.7458	**0.5288**	**-0.9943**
-0.3670	**-0.0696**	**-2.0629**
0.2771	**-0.9564**	**-1.7082**
0.4380	**-1.1154**	**0.5344**
2.8119	**0.7469**	**-0.5076**
0.5754	**0.4771**	**0.7955**

Illustrative Example *continued*

Step 7 : Multiply the Cholesky decomposed lower triangular matrix with each of the N = 6 vectors of random numbers to produce N = 6 correlated normal vectors

L

1.0000	**0.0000**	**0.0000**
0.9912	**0.1324**	**0.0000**
0.9000	**0.0598**	**0.4318**

*

Z

NORMINV(RAND(),0,1)

-0.7458
0.5288
-0.9943

=

L*Z

MMULT

-0.7458
-0.6693
-1.0689

Uncorrelated N(0,1)

L *

-0.7458	-0.3670	0.2771	0.4380	2.8119	0.5754
0.5288	**-0.0696**	**-0.9564**	**-1.1154**	**0.7469**	**0.4771**
-0.9943	**-2.0629**	**-1.7082**	**0.5344**	**-0.5076**	**0.7955**

➡

Correlated N(0,1)

-0.7458	**-0.6693**	**-1.0689**
-0.3670	**-0.0696**	**-2.0629**
0.2771	**0.1480**	**-0.5454**
0.4380	**0.2864**	**0.5582**
2.8119	**2.8861**	**2.3563**
0.5754	**0.6335**	**0.8899**

Illustrative Example *continued*

Step 8: For each line, rank the value of each correlated normal realization.

Correlated Normal		
Line 1	*Line 2*	*Line 3*
-0.7458	-0.6693	-1.0689
-0.3670	-0.0696	-2.0629
0.2771	0.1480	-0.5454
0.4380	0.2864	0.5582
2.8119	2.8861	2.3563
0.5754	0.6335	0.8899

→

Ranking		
Line 1	*Line 2*	*Line 3*
1	1	2
2	2	1
3	3	3
4	4	4
6	6	6
5	5	5

Illustrative Example *continued*

Step 9: Use relative positioning of the correlated Normal draws as the basis for pulling values from the sorted table of uncorrelated reserve indications to create correlated reserve indications across the lines of business.

Ranked Uncorrelated Reserve Indications		
WC	***CAL***	***OL***
100	1,000	10,000
200	2,000	20,000
300	3,000	30,000
400	4,000	40,000
500	5,000	50,000
600	6,000	60,000

Correlated Normal Ranking		
Line 1	*Line 2*	*Line 3*
1	1	2
2	2	1
3	3	3
4	4	4
6	6	6
5	5	5

Correlated Simulations			
WC	*CAL*	*OL*	*Aggregate*
100	1,000	20,000	21,100
200	2,000	10,000	12,200
300	3,000	30,000	33,300
400	4,000	40,000	44,400
600	6,000	60,000	66,600
500	5,000	50,000	55,500

Bootstrap Correlation Method

– steps in a single-line Bootstrap process

Actual Cumulative Historical Data

Acc. Year	Development Age 12	24	36	48
1	1,000	1,500	1,750	2,000
2	1,200	2,000	2,300	
3	1,800	2,500		
4	2,100			
Ave Link Ratio	1.500	1.157	1.143	

Recast Cumulative Historical Data

Acc. Year	Development Age 12	24	36	48
1	1,008	1,512	1,750	2,000
2	1,325	1,988	2,300	
3	1,667	2,500		
4	2,100			

1. Keep current diagonal intact
2. Apply average link ratios to "back-cast" a series of fitted historical payments

Ex: 1,988 = 2,300÷1.157

Bootstrap Correlation Method

– steps in a single-line Bootstrap process

Actual Incremental Historical Data

Acc. Year	Development Age 12	24	36	48
1	1,000	500	250	250
2	1,200	800	300	
3	1,800	700		
4	2,100			

3. Convert both actual and fitted triangles to incrementals
4. Look at difference between fitted and actual payments to develop a set of Residuals (i.e. the variability parameters)

Recast Incremental Historical Data

Acc. Year	Development Age 12	24	36	48
1	1,008	504	238	250
2	1,325	663	312	
3	1,667	833		
4	2,100			

Residuals

= (Actual Incremental - Recast Incremental) / sqrt(Recast Incremental)

Acc. Year	Development Age 12	24	36	48
1	(0.259)	(0.183)	0.801	0.000
2	(3.437)	5.340	(0.699)	
3	3.266	(4.619)		
4	0.000			

Bootstrap Correlation Method
– steps in a single-line Bootstrap process

Random Draw from Residuals

Acc. Year	Development Age 12	24	36	48
1	1.462	(0.335)	5.963	1.462
2	9.749	(8.433)	(0.473)	
3	(1.275)	(6.275)		
4	9.749			

False History

Acc. Year	Development Age 12	24	36	48
1	1,055	497	330	273
2	1,680	445	304	
3	1,615	652		
4	2,547			

5. Create a "false history" by making random draws, **with replacement**, from the triangle of residuals. Note: this will be the key step in the Correlation process!
6. Combine the random draws with the recast historical data to come up with the "false history".

Bootstrap Correlation Method

– steps in a single-line Bootstrap process

Cumulated False History

Acc. Year	Development Age 12	24	36	48
1	1,055	1,551	1,881	2,154
2	1,680	2,125	2,429	
3	1,615	2,267		
4	2,547			
Ave Link Ratio	1.367	1.172	1.145	

Squaring of the Cumulated False History

Acc. Year	Development Age 12	24	36	48
1	1,055	1,551	1,881	2,154
2	1,680	2,125	2,429	**2,782**
3	1,615	1,615	**1,893**	**2,168**
4	2,547	**3,480**	**4,080**	**4,673**

6. Calculate link ratios from the data in the cumulated false history triangle
7. Use the link ratios to square the false history data triangle
8. Note…there are several additional steps to follow, but these can be found in the Appendix to the paper
9. Repeat process N times to get N different reserve indications.

Bootstrap Correlation Method

– variability parameters in each triangle

Variability Parameters Calculated from Original Data

Triangle A

AY	Development Year 1	2	3	4
1	$(1_A,1_A)$	$(1_A,2_A)$	$(1_A,3_A)$	$(1_A,4_A)$
2	$(2_A,1_A)$	$(2_A,2_A)$	$(2_A,3_A)$	
3	$(3_A,1_A)$	$(3_A,2_A)$		
4	$(4_A,1_A)$			

Triangle B

AY	Development Year 1	2	3	4
1	$(1_B,1_B)$	$(1_B,2_B)$	$(1_B,3_B)$	$(1_B,4_B)$
2	$(2_B,1_B)$	$(2_B,2_B)$	$(2_B,3_B)$	
3	$(3_B,1_B)$	$(3_B,2_B)$		
4	$(4_B,1_B)$			

Calculated Variability Parameters

- As part of the bootstrap process, a triangle of residuals (the variability parameters) are created for each line of business being analyzed.
- These residuals are used to create the variations in possible ultimate losses that come out of the bootstrap process.

Bootstrap Correlation Method

– uncorrelated versus correlated Residual value selections

Uncorrelated Bootstrapping - Random reshuffling of variability parameters for Triangle B is independent of the reshuffling in Triangle A

AY	Development Year 1	2	3	4
1	$(2_A,1_A)$	$(3_A,2_A)$	$(1_A,3_A)$	$(3_A,1_A)$
2	$(2_A,2_A)$	$(2_A,3_A)$	$(1_A,2_A)$	
3	$(3_A,1_A)$	$(1_A,1_A)$		
4	$(1_A,1_A)$			

AY	Development Year 1	2	3	4
1	$(2_B,2_B)$	$(3_B,2_B)$	$(1_B,3_B)$	$(2_B,2_B)$
2	$(3_B,1_B)$	$(2_B,3_B)$	$(2_B,2_B)$	
3	$(1_B,3_B)$	$(1_B,1_B)$		
4	$(1_B,2_B)$			

Randomly Selected Variability Parameters to be used in the creation of one possible pseudo-history

Correlated Bootstrapping - Reshuffling of variability parameters in Triangle B is identical to the reshuffling in Triangle A

AY	Development Year 1	2	3	4
1	$(2_A,1_A)$	$(3_A,2_A)$	$(1_A,3_A)$	$(3_A,1_A)$
2	$(2_A,2_A)$	$(1_A,2_A)$	$(2_A,3_A)$	
3	$(3_A,1_A)$	$(1_A,1_A)$		
4	$(1_A,1_A)$			

AY	Development Year 1	2	3	4
1	$(2_B,1_B)$	$(3_B,2_B)$	$(1_B,3_B)$	$(3_B,1_B)$
2	$(2_B,2_B)$	$(1_B,2_B)$	$(2_B,3_B)$	
3	$(3_B,1_B)$	$(1_B,1_B)$		
4	$(1_B,1_B)$			

Randomly Selected Variability Parameters to be used in the creation of one possible pseudo-history

Note the difference in the selection of the residuals in Triangle B in the uncorrelated versus the correlated situations

Pros / Cons of each approach

Correlation Matrix Pros

- More flexible - not limited by observed data

Correlation Matrix Cons

- Requires modeler to do additional work to quantify the correlations between lines

Bootstrap Correlation Pros

- Do not need to make assumptions about underlying correlations

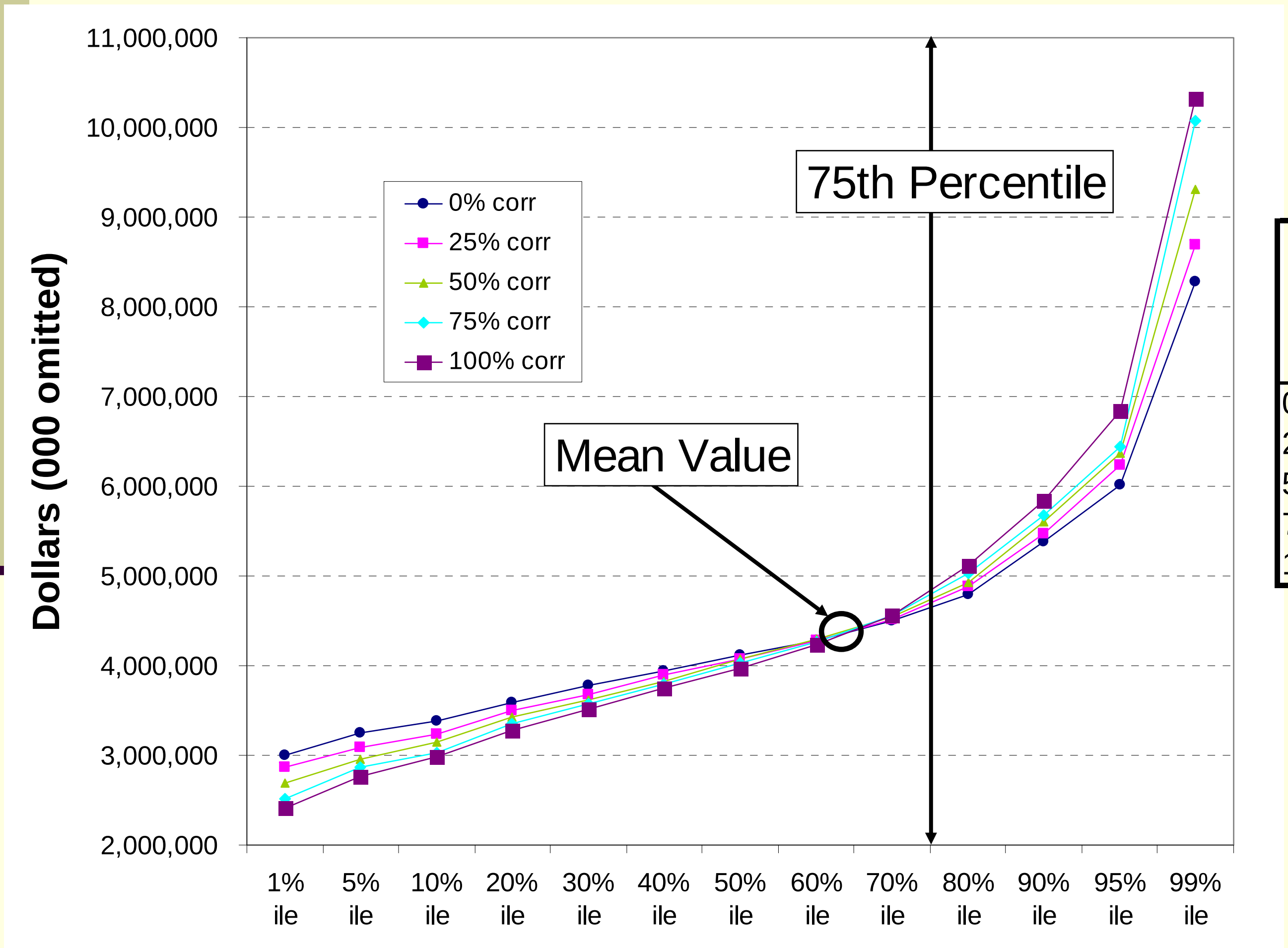
Bootstrap Cons

- Results reflect only those correlations that were in the historical data

Case Study

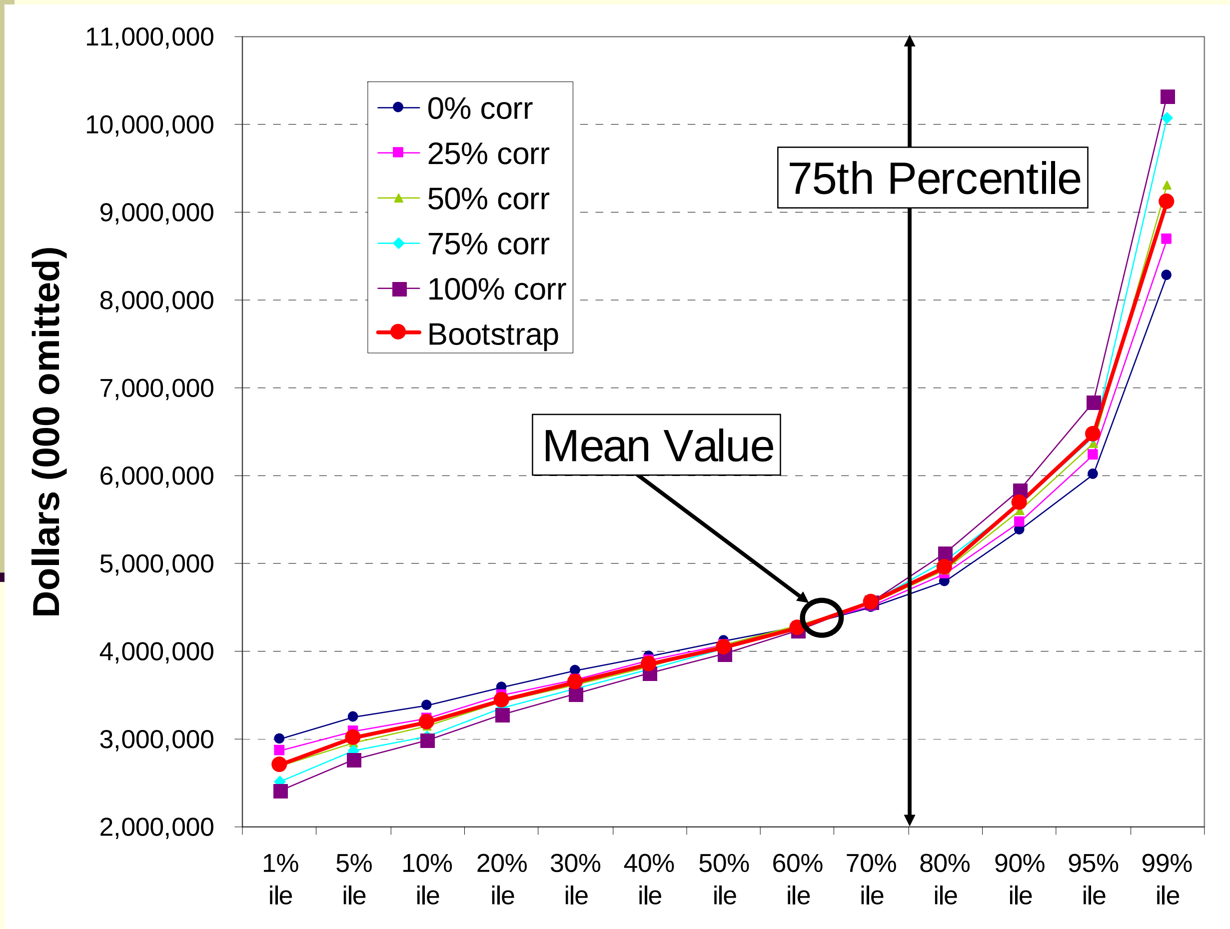
- Three lines of business
- All produce approximately the same mean reserve indication, but with different levels of volatility around the mean
- Run a 5,000 iteration simulation exercise for each line
- Examine the results for the aggregated reserve indication at different percentiles of the aggregate distribution

Case Study 1: Rank correlation results



	Estimated 75th Percentiles	Percent Change from Zero Percentile
0% corr.	4,640,039	n/a
25% corr.	4,697,602	1.2%
50% corr.	4,739,459	2.1%
75% corr.	4,794,767	3.3%
100% corr.	4,836,166	4.2%

Case Study 1: Add bootstrap results



	Estimated 75th Percentiles	Percent Change from Zero Percentile
0% corr.	4,640,039	n/a
25% corr.	4,697,602	1.2%
50% corr.	4,739,459	2.1%
75% corr.	4,794,767	3.3%
100% corr.	4,836,166	4.2%
Bootstrap	4,755,952	2.5%

Case Study conclusions

- Mean aggregated reserve = 4.33B
- Reserves at the 75^{th} percentile range from 4.64B to 4.84B
- Bootstrap tells us that there does appear to be correlations in the underlying data

General conclusions

- To calculate an aggregate reserve distribution, must understand and be able to quantify the dependencies between underlying lines of business
- Correlation is probably not an important issue for lines of business with non-volatile reserve ranges, but might be important for ones with volatile reserves, especially as one moves further towards a tail of the aggregate distribution