Parameterizing Collective Risk Models

CAS Seminar on Reinsurance 2013

R^2

Richard Rosengarten

6/7/2013

TOWERS WATSON

Overview

- Collective Risk Model (CRM) for multiple lines of business with correlation.

- Well-Trodden Ground:
  - Wang
  - Meyers and Collaborators
  - Mildenhall
  - Homer-Rosengarten
  - Many Others

- Correlation: By common shock method as found in several of the references above – with a twist.

- Along the way point out some underappreciated aspects of CRM.

- Actually parameterizing simulation *method* consistent with the *model*.

Overview

- Requirements:
  - **Efficient** as to runtime.
  - **Efficient** as to parameterization – relativity low number of parameters,
  - Simulate **small and large** losses – and reflect the appropriate dependency. Generate individual large losses and small losses in the aggregate.
  - Reflect **correlation** between lines/years.
  - **Consisten**t with underlying CRM.

CRM - Setup

- CV: For any random variable Y, the **coefficient of variation**, or CV is
  - $\boldsymbol{\nu}(Y) = \sqrt{Var(Y)}/E(Y)$

- CV is unit-less, makes for nice formulas.

- **Collective Risk Model**,
  - $Z = X_1 + \cdots + X_N$, X_i iid, X, N independent.

- Where $Z =$ aggregate losses, $X =$ severity, and the random variable N is the claim count, or "frequency"

- Independence of X, N could be violated by inhomogeneous data.

- Large/Small Losses – Threshold T such that (severity) losses $\geq T$ are "large", losses $< T$ are "small".

CRM – Contagion Factor, Moments

- Induced CRMs
  - $Z_L = X_{1,L} + \cdots + X_{N,L}$, $Z_S = X_{1,S} + \cdots + X_{N,S}$

- Contagion Parameter – Set $c = \boldsymbol{v}^2(N) - 1/E(N)$. Then c is invariant in the sense $c = c_L = c_S$ (follows from independence if X, N)

- Assume $c > 0$ (positive contagion).

- Moments of CRM:
  - $E(Z) = E(N)E(X)$
  - $\boldsymbol{v}(Z) = \sqrt{(\boldsymbol{v}^2(X) + 1)/E(N) + c}$

- It follows that $\boldsymbol{v}(Z) \to \sqrt{c}$ as $E(N) \to \infty$

CRM – Large, Small, Total Losses

- Correlation:

$$\boldsymbol{\rho}(Z_S, Z_L) = c/\big(\boldsymbol{\nu}(Z_S)\boldsymbol{\nu}(Z_L)\big)$$

(common shock based on identical mixing distributions)

- Total Variation:

$$E^2(Z)(\boldsymbol{\nu}^2(Z) - c) = E^2(Z_L)(\boldsymbol{\nu}^2(Z_L) - c) + E^2(Z_S)(\boldsymbol{\nu}^2(Z_S) - c).$$

CV Interval

- **Interval for $\boldsymbol{\nu}(Z)$:**

$$\sqrt{c+\frac{E^2(Z_L)}{E^2(Z)}(\boldsymbol{\nu}^2(Z_L)-c)} \leq \boldsymbol{\nu}(Z) \leq \sqrt{c+\frac{E^2(Z_L)}{E^2(Z)}(\boldsymbol{\nu}^2(Z_L)-c)+\frac{T}{E(Z)}\left(1-\frac{E(Z_L)}{E(Z)}\right)} \quad \text{(*)}$$

$$\sqrt{c} \leq \boldsymbol{\nu}(Z_S) \leq \sqrt{c+\frac{T}{E(Z_S)}}$$

- Inequality is sharp.

- **Proof :** Dividing the total variation equation by $E^2(Z)$ immediately gives the left-hand inequality in (*).

To prove the right-hand side, must show that

$\frac{E^2(Z_S)}{E^2(Z)}(\nu^2(Z_S)-c) \leq \frac{T}{E(Z)}\left(1-\frac{E(Z_L)}{E(Z)}\right)$, which reduces to

$$(\nu^2(Z_S)-c) \leq \frac{T}{E(Z_S)}$$

CV Interval

We use the following:

Fact: If Y is a non-negative random variable with support on $[0, T]$,then $Var(Y) \leq E(Y)\big(T - E(Y)\big)$.

Proof of Fact:

$\frac{T^2}{4} \geq E\left(\left(\frac{T}{2} - Y\right)^2\right) = \frac{T^2}{4} - TE(Y) + E(Y^2) = \frac{T^2}{4} - TE(Y) + E^2(Y) + Var(Y)$, which gives the result.

Using fact:

$$(\nu^2(Z_S) - c) = \frac{E\left({X_S}^2\right)}{E(N_S)E^2(X_S)} = \frac{1}{E(N_S)}\left(1 + \frac{Var(X_S)}{E^2(X_S)}\right) \leq \frac{1}{E(N_S)}\left(1 + \frac{E(X_S)\big(T - E(X_S)\big)}{E^2(X_S)}\right)$$

$$= \frac{T}{E(N_S)E(X_S)} = \frac{T}{E(Z_S)}$$, as required.

CV Interval

For sharpness note that if we hold $E(Z_S)$ fixed while letting $E(N_S) \to \infty$, then $\nu^2(Z_S) \to c$, so that $\nu(Z) \to \sqrt{c + \frac{E^2(Z_L)}{E^2(Z)}(\nu^2(Z_L) - c)}$, which is the left-hand side of (*). Furthermore if we take X_S to be a 2-point distribution with masses at $X_S = 0$ and $X_S = T$ (with probability $p = \frac{E(Z_S)}{E(N_S)T}$), then equality holds for the right-hand side of (*).

Mixed Poisson CRM

- We now assume that the claim count r.v N is of *mixed Poisson* type, meaning $N \sim Poisson[E(N)G]$, where G is a r.v with mean 1.

- To draw from N:
  - 1. Draw g from G.
  - 2. Draw from $Poisson[E(N)g]$.

- $Var(G) = c$. Will use the notation $G[c]$

- N_L, N_S are also mixed Poisson with the same *mixing distribution* G.

- Example: $G \sim gamma$. Then $N \sim Negative\ Binomial$.

- Fact ("Severity is Irrelevant"): $Z/E(Z) \xrightarrow{D} G\ as\ E(N) \to \infty$

Simulation Method - CAD Algorithm with Frequency, "Severity" and Serial Common Shock

- Ref:Homer-Rosengarten (2011), Meyers-Klinker-LaLonde (2003)

- **Full Info CAD** (Have N, X)
  - Draw from N (i.e. draw from G and then from $Poisson[E(N)G]$)
  - Draw N_L from $Bin(N, q)$, where $q = 1 - \mathrm{CDF}_X(T)$. $N_S = N - N_L$.
  - Draw $X_{1,L}, \ldots, X_{N,L}$ large losses. $Z_L = X_{1,L} + \cdots + X_{N,L}$
  - Draw $\widetilde{Z_S}$ from Conditional Aggregate Distribution (eg, lognormal) matching $k \geq 2$ moments of $Z_S | N_S$.
  - $\tilde{Z} = \widetilde{Z_S} + Z_L$

- H-R Paper: $\tilde{Z}/E(Z), \widetilde{Z_S}/E(Z_S)\,(Z_L/E(Z_L)) \xrightarrow{D} G$. This generalizes the "severity is irrelevant" result. Also, the method generates the correct dependence between large and small losses

Simulation Method

- **Limited Info CAD** (Don't have N, X)
  - Draw from G only.
  - Draw N_L from $Poisson[E(N_L)G]$
  - Draw large losses as previously.
  - Draw $\widetilde{Z_S}$ from CAD matching first **two** moments of $Z_S|G$

- **Minimum Parameterization**: $G[c], E(N_L), X_L, E(Z), \boldsymbol{\nu}(Z)$

- Can then eliminate severity, N_S from equations for first two moments of $Z_S|G$.

- To wit, $E(Z_S|G) = GE(Z_S)$, $\boldsymbol{\nu}(Z_S|G) = \sqrt{(\boldsymbol{\nu}^2(Z_S) - c)/G}$

- **But,** it is not automatic that this minimal parameterization is consistent with CRM

Simulation Method

- To address, suppose we have all the minimal parameters except $\boldsymbol{\nu}(Z)$. We can then evaluate the lhs and rhs of inequality (*)

$$\sqrt{c+\frac{E^2(Z_L)}{E^2(Z)}(\boldsymbol{\nu}^2(Z_L)-c)} \le \boldsymbol{\nu}(Z) \le \sqrt{c+\frac{E^2(Z_L)}{E^2(Z)}(\boldsymbol{\nu}^2(Z_L)-c)+\frac{T}{E(Z)}\left(1-\frac{E(Z_L)}{E(Z)}\right)}$$

- Any choice for $\boldsymbol{\nu}(Z)$ within this interval is a) possible and b) consistent with MP CRM.

Beginning of Example – R^2 Ins Co.

Loss Parameters

Non-Cat

LoB	Premium	E(Z)	Loss Ratio	$\nu(Z)$	T	c	$E(N_L)$	$E(Z_L)$	X_L	$\nu(Z_L)$	$E(Z_S)$	$\nu(Z_S)$
GL	110,000,000	65,000,000	59.1%	0.2000	1,000,000	0.03	3.500	5,457,138	Empirical	0.7349	59,542,862	0.1940
WC	90,000,000	45,000,000	50.0%	0.2200	1,000,000	0.02	3.000	6,568,231	Empirical	0.7604	38,431,769	0.2065
CAL	40,000,000	22,000,000	55.0%	0.2750	1,000,000	0.04	0.250	512,500	Empirical	3.2929	21,487,500	0.2668
Umb	9,000,000	6,500,000	72.2%	0.5200	1,000,000	0.02	3.000	4,248,825	Empirical	0.7444	2,251,175	0.4525
PropNon-Cat	300,000,000	175,000,000	58.3%	0.1600	1,000,000	0.02	14.000	30,534,169	Empirical	0.3734	144,465,831	0.1513
Total Non-Cat	549,000,000	313,500,000	57.10%	0.1139			23.750	47,320,864		0.2877	266,179,136	0.1096
Cat												
SmallCat	549,000,000	40,000,000	7.3%	0.4300	2,000,000	0.16	10.000	4,000,000	Lognormal	0.4300	-	-
MajorCat (Net)	549,000,000	25,000,000	4.6%	1.9000	Inf	1.00	-	-	N\A	-	25,000,000	1.9000
Total Inc Cat	549,000,000	378,500,000	68.94%	0.1685			33.750	51,320,864		0.2847	291,179,136	0.195

R^2 Ins Co. – Mean, CV Parameters

Section C: If Model Choice a) = "Full Info"- ignore; b) = "Limited Info 1" - mean and cv of aggregate total losses; c) = "Limited Info 2" mean and cv of aggregate small losses.

For case b) and c), first parameterize large loss CRM in Section D below. Selection is then guided by Mu Min, CV Min, and CV Max.

Moments of:	All Losses	All Losses	All Losses	All Losses	All Losses	Small Losses	All Losses
Mean Mu Min[Yr]	5,457,138		512,500			0	
Mean Mu[Yr]	65,000,000	45,000,000	22,000,000	6,500,000	175,000,000	0	25,000,000
CV Nu Min[Yr]	**0.18331**	**0.17862**	**0.21423**	**0.49824**	**0.15375**	**0.40006**	**1.00007**
CV Nu Max[Yr]	**0.21841**	**0.22552**	**0.30043**	**0.54914**	**0.16845**	**99999.00006**	**6.40317**
CV Nu[Yr]	0.200000	0.220000	0.275000	0.520000	0.160000	0.430000	1.900000

Section D - "Full Info" - Total Loss CRM; otherwise Large Loss CRM

Parameters Of:	Large Loss CRM	Large Loss CRM	Large Loss CRM	Large Loss CRM	Large Loss CRM	Large Loss CRM	Large Loss CRM
E(Claim Count) Lambda [Yr]	3.5000	3.0000	0.2500	3.0000	14.0000	10.0000	0.0000
Severity Max[Yr]	999,999,999,999	999,999,999,999	999,999,999,999	############	999,999,999,999	999,999,999,999	############
Severity Distribution	Empirical	Empirical	Empirical	Empirical	Empirical	Lognormal	Empirical
See Below	0.0	0.0	0.0	0.0	0.0		0.0
shift						2,000,000.0	
mu						14.4	
sigma						0.5	
min						2,000,000.0	
max						##################	

Severtiy Values and Weights used for Empirical And Mixed Exponential severity disributions

Severity Values[Yr]	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	2,000,000	2,000,000
	1,250,000	1,250,000	2,000,000	1,250,000	1,250,000		
	1,500,000	1,500,000	10,000,000	1,500,000	1,500,000		

Common Shock Correlation

- Correlate LoBs modeled with MP CRM/CAD method.
- LoBs are organized into covariance groups. Only Lobs within the same covariance group co-vary with one another.
- Frequency, “severity”, and serial common shock.

Frequency Common Shock

- General Idea: Common draw from mixing distribution.
- Need to allow that LoBs might have different mixing distributions.
- Solution is draw common uniforms and use these to invert the mixing distributions ($g = F_G^{-1}(u)$).
- Remaining problem is that this will tend to generate very high correlation.
- Usual solution is to assume that G is an independent product, ie
  - $G[c] = G_1[c_1]G_2[c_2]$
  - Then apply common shock only to G_1.
  - Note that $c = c_1 + c_2 + c_1 c_2$

Frequency Common Shock

- Variant is the "twisted product" $G[c] = G_1[c_1] \ltimes G_2\,[c_2]$ defined by $G = G_1 G_2[c_2/G_1]$.

- That is, to draw from G:
  - Draw g_1 from G_1.
  - Draw g_2 from $G_2[c_2/g_1]$.
  - $g = g_1 g_2$.

- Nice thing about twisted product is $c = c_1 + c_2$.

- **Parameter**: FrCoVarWt $= w, 0 \leq w \leq 1$. Varies by LoB.

- In twisted product set $c_1 = wc, c_2 = (1-w)c$ (where $G_i[0] \equiv 1$).

Serial Common Shock

- Bring in uniforms necessary to invert G_1's for frequency c.s. These vary by covariance group and year.

- Also bring in uniforms for G_2's – varying by LoB and year.

- Reason for G_2's is generate sufficient correlation between years but within LoB.

- Flip a weighted coin.

- For year $j, j \geq 2$, if coin flip comes up "heads" use the uniforms from year $j-1$. Otherwise use year j.

- **Parameter** – FrSerialCoVarWt – the weight for the coin flip. Can vary by covariance group or LoB. Usually by covariance group.

Serial Common Shock

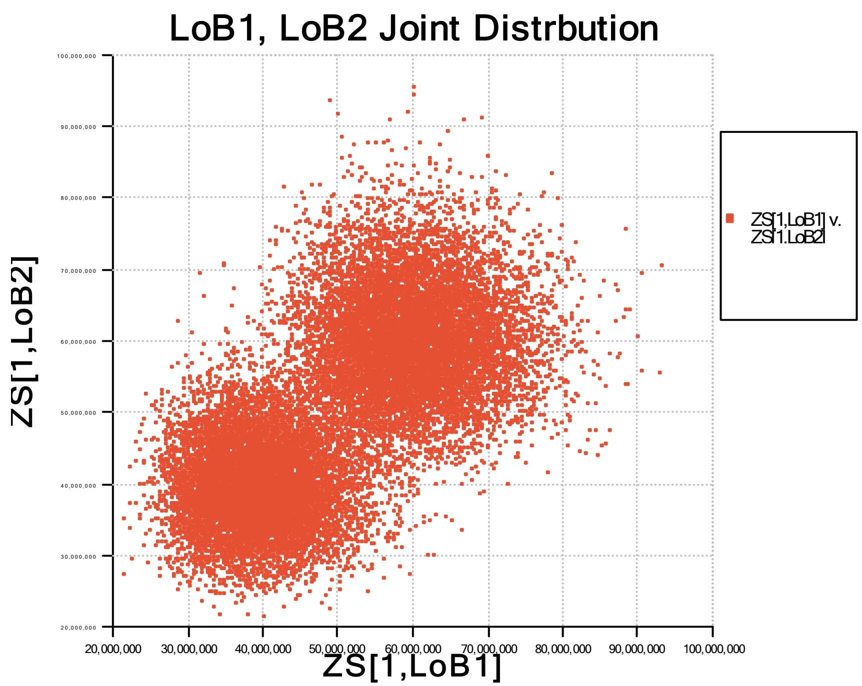
- Summary
  - G_1 correlates non-identical LoBs, both within-year and serially.
  - G_2 - serial correlation for identical LoBs.
  - Serial correlation decays by FrSerialCoVarWt.

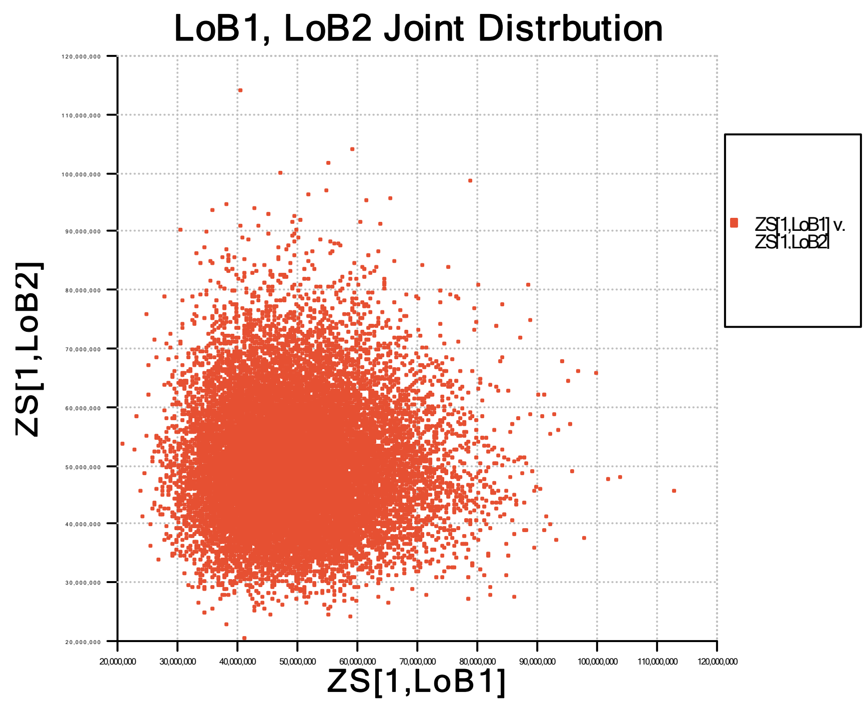
"Severity" Common Shock

- Really it's c.s. applied to the conditional aggregate distribution generating $\widetilde{Z_S}$.
- By H-R, the particular distribution family used doesn't matter.
- Assume lognormal, with $Mu, Sigma$ the conditional parameters.
- **Parameters**: ZSCoVarWt, ZSSerialCoVarWt.
- Express CAD as a product of Lognormals
- $CAD = \mathrm{logn}\left[.5Mu, Sigma\sqrt{ZSCovarWt}\right]\mathrm{logn}\left[.5Mu, Sigma\sqrt{1-ZSCovarWt}\right]$
- Play same game as previously.

Why do we need ZSCoVarWt?

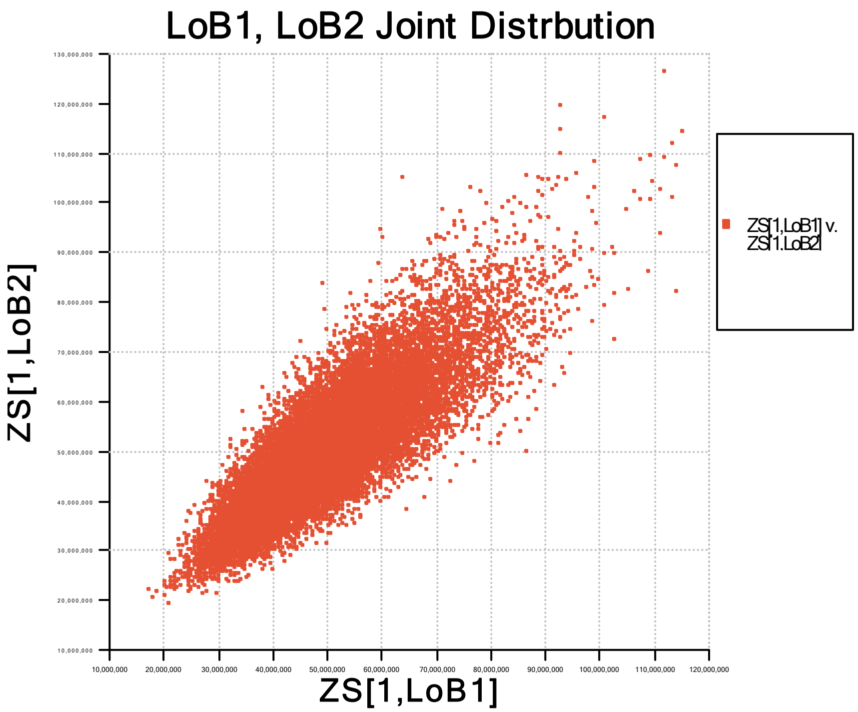
- Example: Identical LoBs LoB1, LoB2
- $FrCoVarWt = .85$, $ZSCoVarWt = 0$, $G_1 = 1 \pm \sqrt{c}$, with probabiltiy .5.
- $c = O(\boldsymbol{v}^2)$ - High Correlation $c = 0\ (\boldsymbol{v}^2 \gg c)$ - No Correlation





Why ZSCoVarWt?

- $FrCoVarWt = 0, ZSCoVarWt = .85, c = 0$



Why ZSCoVarWt?

- For Identical LoBs:

	FrCoVarWt=1 ZSCoVarWt=0	FrCoVarWt=0 ZSCoVarWt=1
$\mathbf{v}^2 \to c$	$\boldsymbol{\rho} \to 1$	$\boldsymbol{\rho} \to 0$
$\mathbf{v}^2 \gg c$	$\boldsymbol{\rho} \to 0$	$\boldsymbol{\rho} \to 1$

More Tricks

- Can increase skewness by adding shift to mixing distributions.
  - Shifted lognormal: $G = s + Logn\left[ln\left(\frac{(1-s)^2}{\sqrt{c+(1-s)^2}}\right), \sqrt{ln\left(1+\frac{c}{(1-s)^2}\right)}\right]$
  - Skewness $= \frac{\sqrt{c}}{(1-s)}\left(3+\frac{c}{(1-s)^2}\right)$
- Can use discrete mixing distribution to create a mass at 0, for example.

R^2 Ins Co. – Correlation Parameters, Mixing Distributions

CAD Large Loss/Small Loss Simulation - Elements that vary by year are indicated with [Yr]

Section A - Setup

Clases	GL	WC	CAL	Umb	PropNon-Cat	SmallCat	MajorCat
Large Loss Threshold[Yr]	1,000,000	1,000,000	1,000,000	1,000,000	1,000,000	2,000,000	999,999,999
ModelChoice	Limited Info1	Limited Info1	Limited Info1	Limited Info1	Limited Info1	Limited Info2	Limited Info1
CoVarGroup	1	1	1	1	2	3	4
FrCoVarWT	0.5000	0.3500	0.2500	1.0000	0.2500	0.5000	0.6108
ZSCoVarWT	1	0.075	0.075	1	0	0	0
FrSerialCoVarWT	0.3				0	0	0
ZSSerialCovarWeight	0.3				0	0	0

Section B - Parameterize Mixing Distribution. Mixing Dist of the form G = G1 compound with G2. G has mean 1 and variance c (G1 mean 1 and variance c1).

Note that c1=FrCoVarWt*c So, choice of FrCoVarWt=1 (and p=1 for Wtd Sum) means c1=c and G=G1; Choice of FrCoVArWt=0 (and p=0 for Wtd. Sum) means c2=c and G=G2.

Note that the G1's by class co-vary within the defined covariance groups.

Compounding Type	TwistedProduct	TwistedProduct	TwistedProduct	TwistedProduct	TwistedProduct	TwistedProduct	StraightProduct
	0.5	0.5	0.5	0.5	0.5	0.5	0.5
c	0.03000	0.02000	0.04000	0.02000	0.02000	0.16000	1.00000
G1Choice	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	DiscreteUniform
shift	0	0	0	0	0	0	
p							0.3892
m							1
G2Choice	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal
shift	0	0	0	0	0	0	0.75

R^2 Ins Co. – Correlation Parameters, Mixing Distributions

- Casualty lines co-vary. (Covariance group 1)
- Non-Cat Property, Cats are independent (CoVar groups 2-4).
- Mixing Distributions all of form $\mathrm{G} = logn \ltimes logn$ except Major Cat
- Major Cat (Net of Cat XoL)
  - From Cat modelling know $Prob(0), E(Major\ Cat), \boldsymbol{\nu}, \gamma = skewness$**.**
  - Modeling Solution: $\mathrm{G} = Discrete\ Uniform * (shift\ logn), c = 1$
  - Parameters of Discrete Uniform (inluding $c_1 = FrCoVarWt$) set up to match probability mass at 0.
  - Shifted lognormal set up to match skewness.
- Umbrella: parameters set up to give higher correlation with GL than other casualty lines.

Correlation Matrix

CorrZ								
		1						
		GL	WC	CAL	Umb	PropNon-Cat	SmallCat	MajorCat
1	GL	1.0000	0.2981	0.2923	0.2755	-0.0110	-0.0017	0.0004
	WC	0.2981	1.0000	0.1654	0.1453	-0.0033	-0.0075	-0.0079
	CAL	0.2923	0.1654	1.0000	0.1345	-0.0057	0.0034	-0.0007
	Umb	0.2755	0.1453	0.1345	1.0000	-0.0003	0.0056	0.0030
	PropNon-Cat	-0.0110	-0.0033	-0.0057	-0.0003	1.0000	0.0015	-0.0028
	SmallCat	-0.0017	-0.0075	0.0034	0.0056	0.0015	1.0000	-0.0072
	MajorCat	0.0004	-0.0079	-0.0007	0.0030	-0.0028	-0.0072	1.0000
2	GL	0.2672	0.0780	0.1010	0.0834	-0.0053	0.0006	0.0072
	WC	0.0846	0.2368	0.0494	0.0395	-0.0069	-0.0132	0.0011
	CAL	0.0814	0.0482	0.2754	0.0488	-0.0032	-0.0095	0.0130
	Umb	0.0882	0.0287	0.0435	0.0459	0.0051	0.0147	0.0008
	PropNon-Cat	-0.0005	0.0212	0.0058	0.0234	-0.0030	0.0012	-0.0033
	SmallCat	-0.0004	-0.0045	0.0075	0.0067	-0.0131	-0.0134	0.0002
	MajorCat	0.0084	0.0070	-0.0042	0.0087	0.0046	0.0143	-0.0020
3	GL	0.0768	0.0318	0.0309	0.0234	0.0007	-0.0099	-0.0002
	WC	0.0293	0.0655	0.0235	0.0082	-0.0045	-0.0057	0.0029
	CAL	0.0343	0.0322	0.0967	0.0113	-0.0020	-0.0107	0.0092
	Umb	0.0324	0.0167	0.0165	0.0167	-0.0039	-0.0065	-0.0004
	PropNon-Cat	-0.0053	0.0029	0.0085	-0.0010	0.0080	-0.0027	-0.0002
	SmallCat	0.0015	-0.0011	0.0074	0.0028	0.0067	0.0025	0.0021
	MajorCat	0.0042	0.0108	0.0054	0.0026	0.0054	0.0008	-0.0037

Reinsurance Cover for R^2 Insurance Co.

- Aggregate Stop-Loss – Term: 2 years
- Subject Losses:
  - 100% of Non-Cat losses limited to $\$1m$ per risk
  - 50% of $\$1m\ xs\ \$1m$ per risk
  - 100% of Cat losses limited to 20% of Subject Premium $(\$549m)$ per year.
- Subject Loss Ratio: 64.5%
- Coverage: $15\%\ xs\ 75\%$ of SP
- Premium: 5% of SP (33% RoL), 30% of which is margin, the remainder to an experience account.
- Profit Commission: 100% of residual EA

Reinsurance Cover Results

- NPV basis (Have also developed payout patterns by LoB).
- Low parameter model allows for efficient sensitivity testing.
- Key Stats (Reinsurer PoV):

Key Stats w\Sensitivity Testing				
	Base	**c's, CVs Up**	**CoVar Wts Up**	**Skewness Up**
NPV(Profit/Loss)	12,851,332	9,773,105	12,281,670	12,776,752
Prob(Econ. Loss)	11.59%	17.51%	12.32%	11.73%
TVaR(95)	(35,406,559)	(50,849,025)	(41,391,711)	(35,944,163)
TVaR(97.5)	(47,754,569)	(66,138,266)	(56,187,593)	(48,446,883)
RoRaC(95)	9.46%	6.00%	8.08%	9.34%
RoRaC(97.5)	7.53%	5.08%	6.48%	7.45%
ERD	-4.40%	-8.00%	-5.23%	-4.52%