Dependencies in Stochastic Loss Reserve Models

Glenn Meyers

Presentation to

ASTIN Colloquium 2016

The CAS Loss Reserve Database Created by Meyers and Shi

With Permission of American NAIC

- Schedule P (Data from Parts 1-4) for several US Insurers
  - Private Passenger Auto
  - Commercial Auto
  - Workers' Compensation
  - General Liability
  - Product Liability
  - Medical Malpractice (Claims Made)
- Available on CAS Website

  http://www.casact.org/research/index.cfm?fa=loss_reserves_data

Data Used in Study

- 50 Insurers from four lines of business
  - Commercial Auto (CA)
  - Personal Auto (PA)
  - Workers' Compensation (WC)
  - Other Liability (OL)
- 102 Pairs of triangles with these insurers
  - 29 CA-PA
  - 17 CA-WC
  - 17 CA-OL
  - 14 PA-WC
  - 15 PA-OL
  - 10 WC-OL

Illustrative Insurer – Paid Losses

		Cumulative Paid Losses										
Premium	AY/Lag	1	2	3	4	5	6	7	8	9	10	Source
5812	1988	952	1529	2813	3647	3724	3832	3899	3907	3911	3912	1997
4908	1989	849	1564	2202	2432	2468	2487	2513	2526	2531	2527	1998
5454	1990	983	2211	2830	3832	4039	4065	4102	4155	4268	4274	1999
5165	1991	1657	2685	3169	3600	3900	4320	4332	4338	4341	4341	2000
5214	1992	932	1940	2626	3332	3368	3491	3531	3540	3540	3583	2001
5230	1993	1162	2402	2799	2996	3034	3042	3230	3238	3241	3268	2002
4992	1994	1478	2980	3945	4714	5462	5680	5682	5683	5684	5684	2003
5466	1995	1240	2080	2607	3080	3678	4116	4117	4125	4128	4128	2004
5226	1996	1326	2412	3367	3843	3965	4127	4133	4141	4142	4144	2005
4962	1997	1413	2683	3173	3674	3805	4005	4020	4095	4132	4139	2006

Notation

- w = Accident Year $w = 1,\ldots,10$
- d = Development Year $d = 1,\ldots,10$
- $C_{w,d}$ = Cumulative paid loss

The Changing Settlement Rate (CSR) Model

- *logelr* ~ uniform(-5,0)
- $\alpha_1 = 0, \alpha_w$ ~ normal(0,$\sqrt{10}$) for *w* = 2,…,10
- β_{10} = 0, β_d ~ uniform(-5,5), for *d* = 1,…,9
- a_i ~ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as *d* increases
- $\mu_{w,d} = \log(\text{Premium}_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
- $C_{w,d}$ ~ lognormal($\mu_{w,d}$, σ_d)

CSR Model
Allow for Changing Loss Ratio

- *logelr* ~ uniform(-5,0)
- $\alpha_1 = 0, \alpha_w$ ~ normal$(0,\sqrt{10})$ for w = 2,…,10
- β_{10} = 0, β_d ~ uniform(-5,5), for d = 1,…,9
- a_i ~ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases
- $\mu_{w,d} = \log(\text{Premium}_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
- $C_{w,d}$ ~ lognormal$(\mu_{w,d}, \sigma_d)$

CSR Model
Changing σ_d

- *logelr* ~ uniform(-5,0)
- $\alpha_1 = 0, \alpha_w$ ~ normal($0,\sqrt{10}$) for w = 2,…,10
- β_{10} = 0, β_d ~ uniform(-5,5), for d = 1,…,9
- a_i ~ uniform(0,1)

- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases

- $\mu_{w,d} = \log(\text{Premium}_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
- $C_{w,d}$ ~ lognormal($\mu_{w,d}$, σ_d)

CSR Model
Speedup Interaction

- $logelr \sim$ uniform(-5,0)
- $\alpha_1 = 0, \alpha_w \sim$ normal$(0, \sqrt{10})$ for $w = 2,\ldots,10$
- $\beta_{10} = 0$, $\beta_d \sim$ uniform(-5,5), for $d = 1,\ldots,9$
- $a_i \sim$ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases
- $\mu_{w,d} = \log(Premium_w) + logelr + \alpha_w + \beta_d \cdot speedup_w$
- $C_{w,d} \sim$ lognormal$(\mu_{w,d}, \sigma_d)$

The $Speedup_w$ Parameters

- β_ds are almost always negative! ($\beta_{10} = 0$)
- $speedup_1 = 1$
- $speedup_w = speedup_{w-1} \bullet (1 – \gamma – (w-2) \bullet \delta)$
- Speedup rate = $\gamma + (w-2) \bullet \delta$
  - If positive, claim settlement speeds up
  - If negative, claim settlement slows down
  - Can change over time

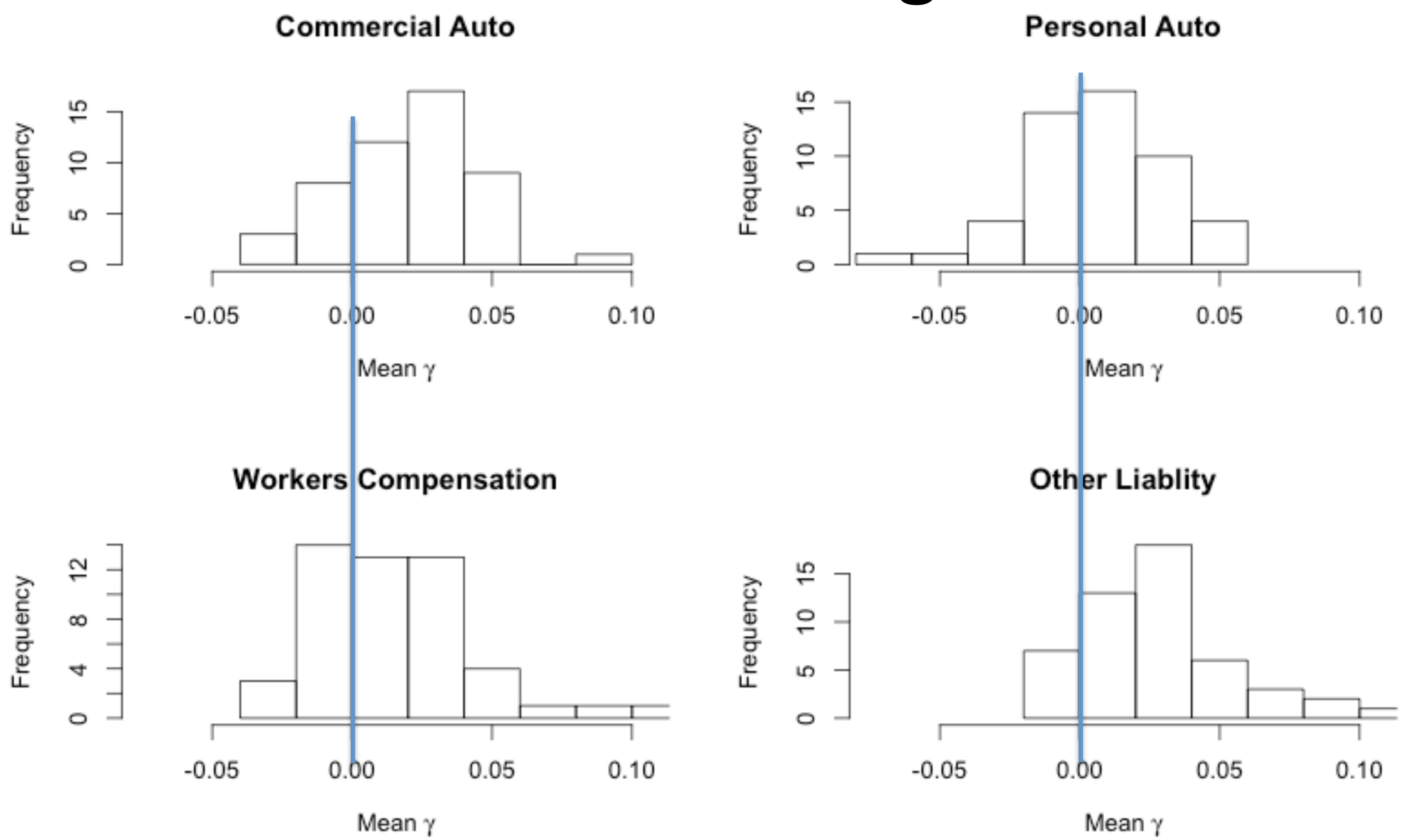
Bayesian MCMC Models

- Use R with “rstan” package
- Get a sample of 10,000 parameter sets from the posterior distribution of the model

  $$\{\alpha_w\}_{w=2}^{10}, \{\beta_d\}_{d=1}^{9}, \{\sigma_d\}_{d=1}^{10}, logelr, \gamma, \delta$$

- Use the parameter sets to get 10,000, $\sum_{w=1}^{10} C_{w,10}$, simulated outcomes
- Calculate summary statistics of the simulated outcomes
  - Mean
  - Standard Deviation
  - Percentile of Actual Outcome

Distribution of Mean γs for 200 Loss Triangles

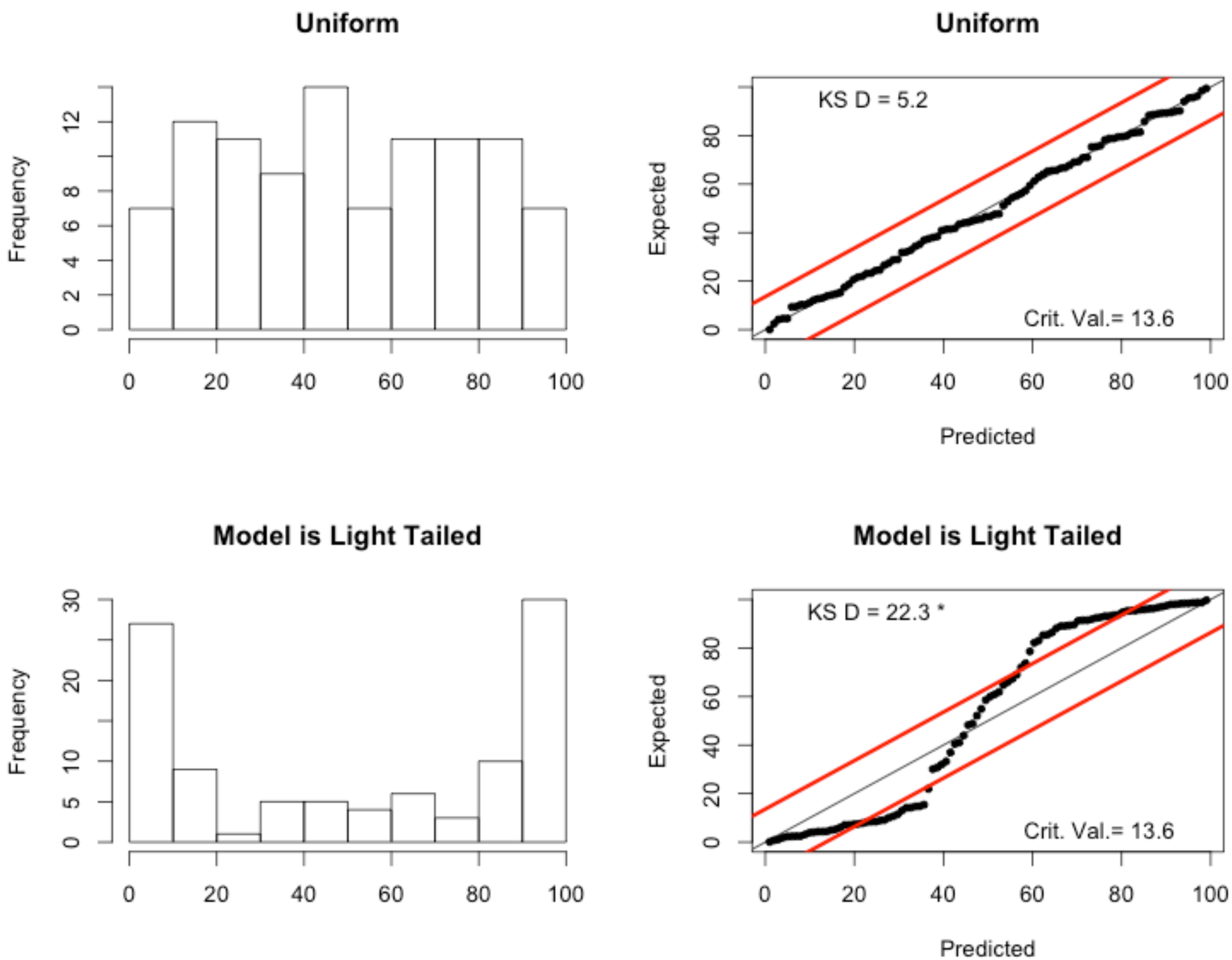


Speedup of claim settlement
for most insurers

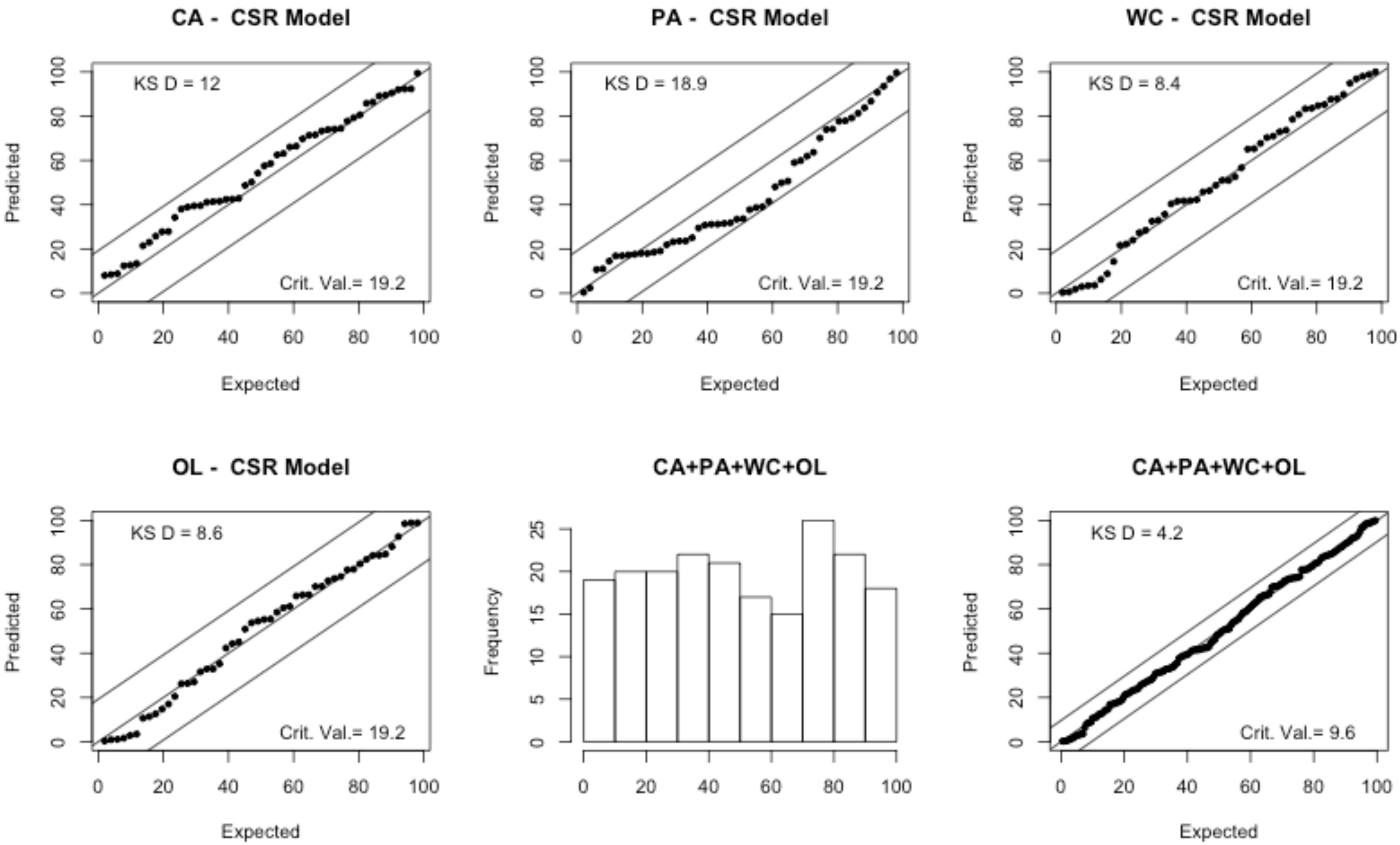
Criteria for a "Good" Stochastic Loss Reserve Model

- Using the predictive distributions, find the percentiles of the outcome data for several loss triangles.
- The percentiles should be uniformly distributed.
  - Histograms
  - PP Plots and Kolmogorov Smirnov Tests
    - Plot Expected vs Predicted Percentiles
    - KS 95% critical values = 19.2 for *n* = 50 and 9.6 for *n* = 200

Illustrative Tests of Uniformity



Test of CSR on Paid Data



Conclusion – Validates within KS Boundaries

In Passing - Result for Other Models
Software from "ChainLadder" package

- Incurred data
  - Mack model understates variability.
  - Fails to recognize dependencies between accident years – corrected with Correlated Chain Ladder (CCL) model.
- Paid data
  - Both Mack and Bootstrap ODP are biased upward.
  - Corrected by CSR model.

Dependencies Between Lines of Insurance

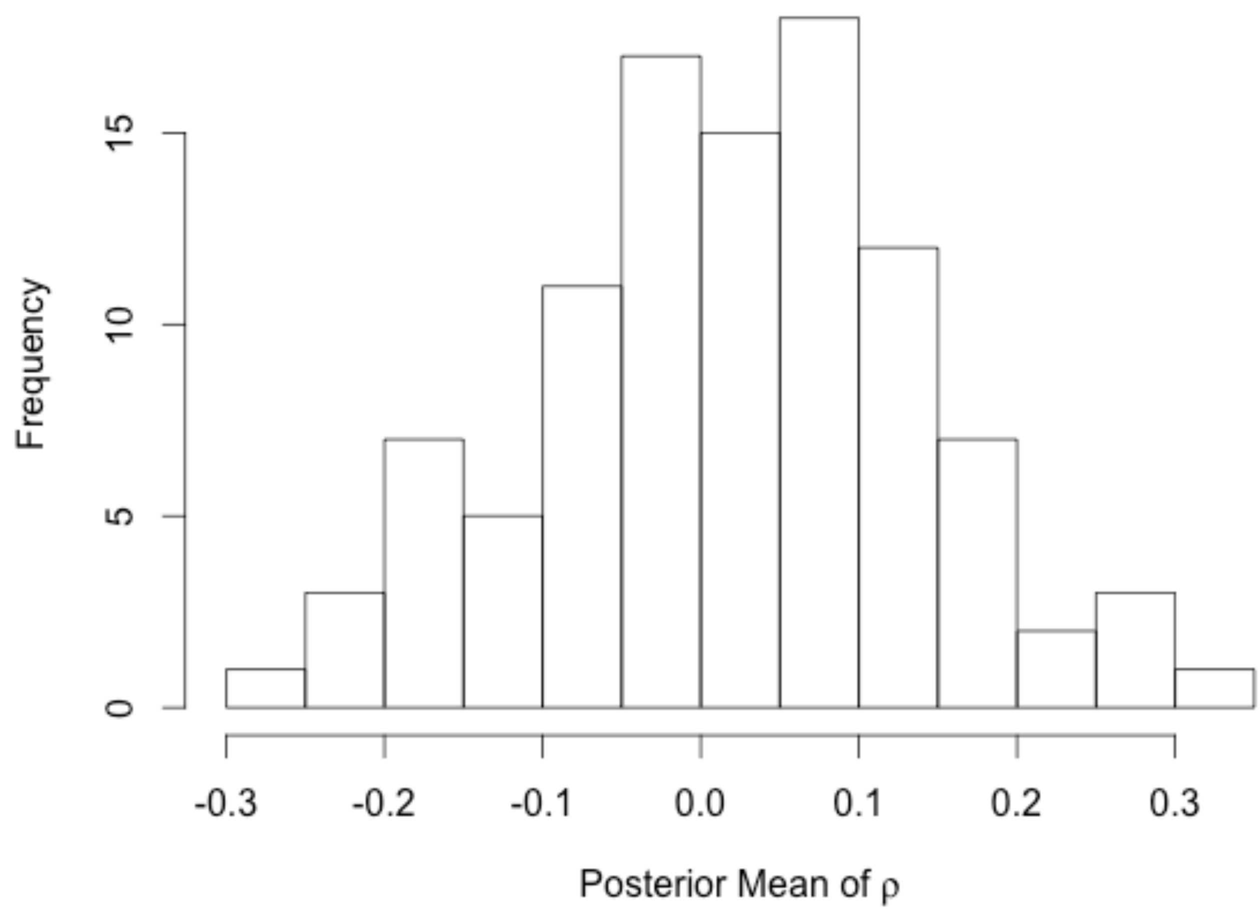
Joint Lognormal Distribution

$$\begin{pmatrix} \log\left(C^{X}_{wd}\right) \\ \log\left(C^{Y}_{wd}\right) \end{pmatrix} \sim \text{Multivariate Normal} \left(\begin{pmatrix} \mu^{X}_{wd} \\ \mu^{Y}_{wd} \end{pmatrix}, \begin{pmatrix} \left(\sigma^{X}_{wd}\right)^{2} & \rho\,\sigma^{X}_{wd}\sigma^{Y}_{wd} \\ \rho\,\sigma^{X}_{wd}\sigma^{Y}_{wd} & \left(\sigma^{X}_{wd}\right)^{2} \end{pmatrix} \right)$$

- Step 1 – Get univariate sample of 10,000 μ_{wd}s and σ_{d}s for each line X and Y = CA, PA, WC or OL
- Step 2 – For each parameter set in the univariate sample for each line, use MCMC to get a single ρ from the bivariate distribution of

$$(\log(C^{X}_{wd}), \quad \log(C^{Y}_{wd}))$$

Got Samples of ρ for 102 Pairs of Triangles in CAS Database

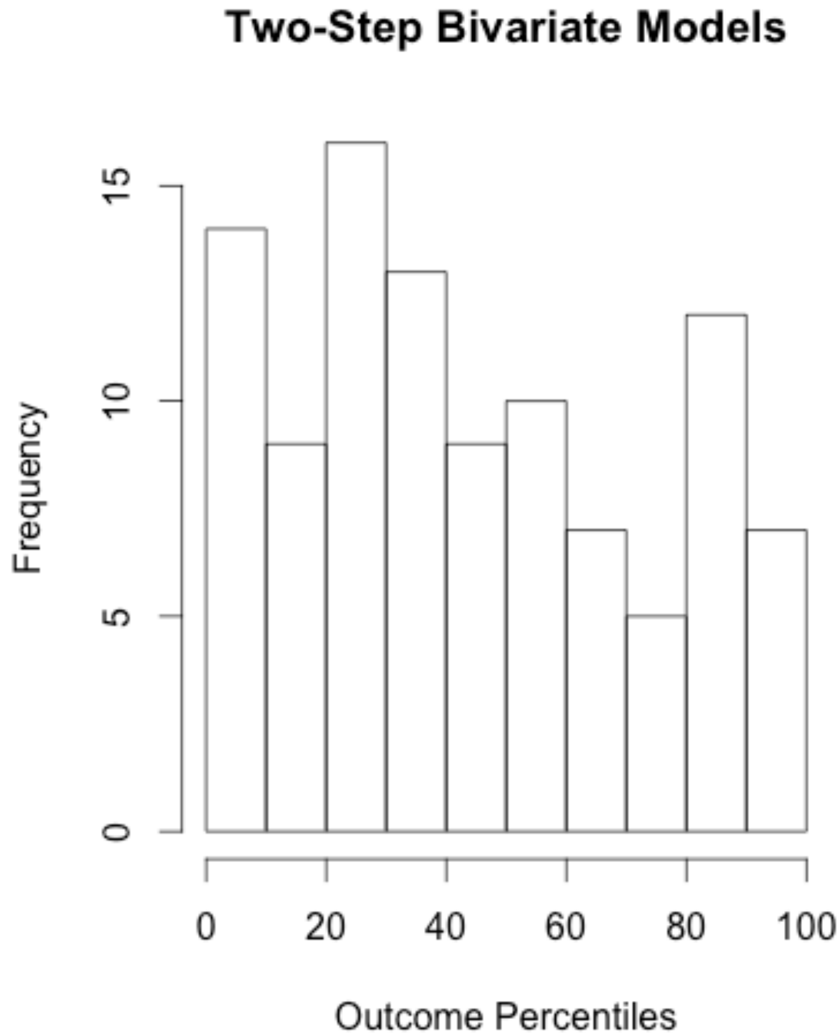


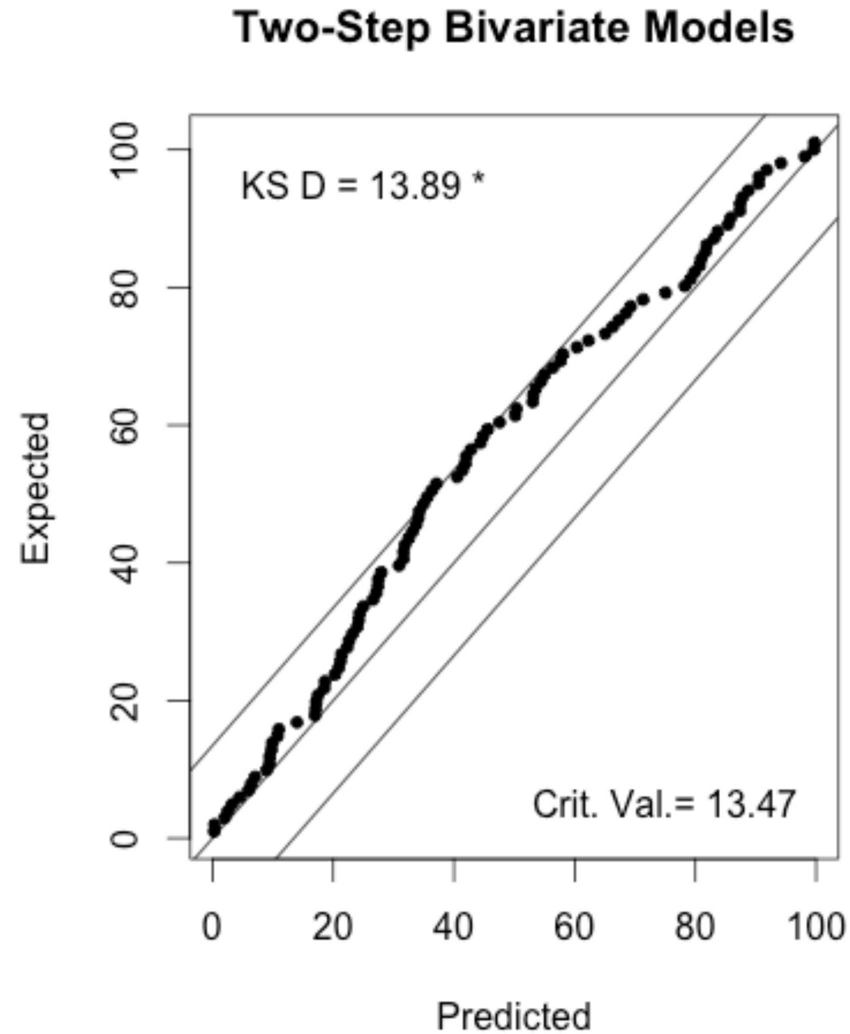
Distribution of the Sum of Losses for Two Lines of Insurance

$$\sum_{w=1}^{10} C_{w,10}^{X} + \sum_{w=1}^{10} C_{w,10}^{Y}$$

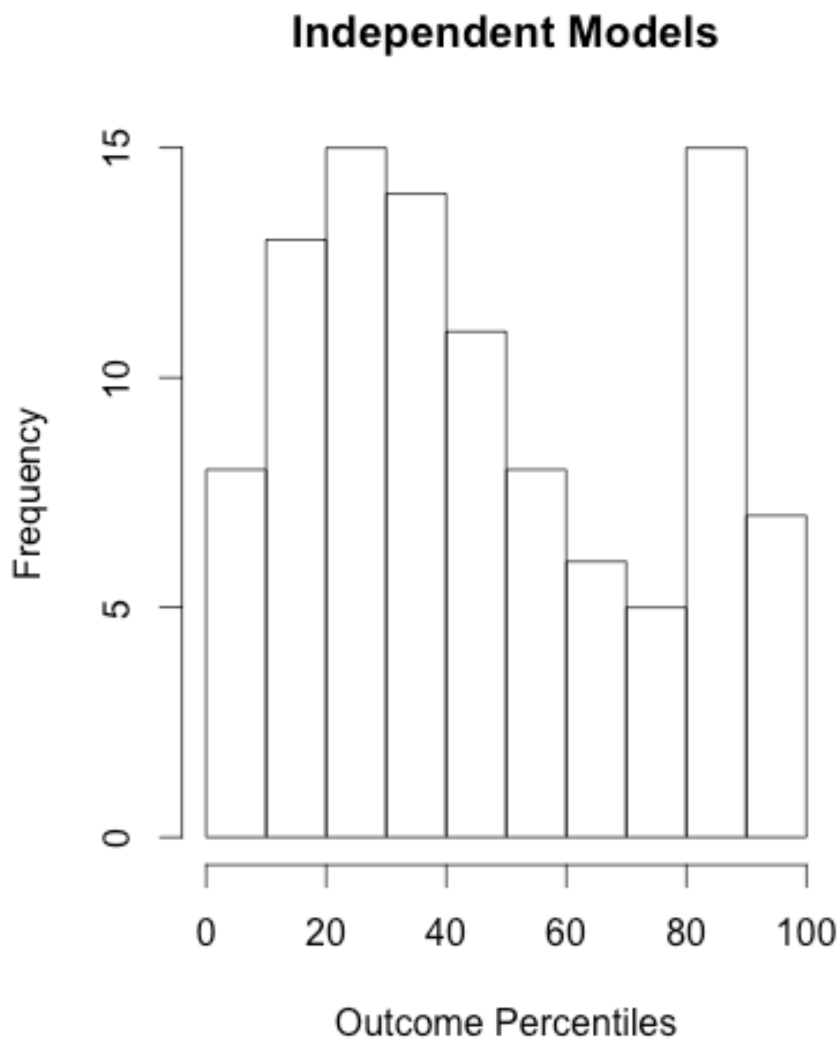
- From the 2-step bivariate model
- From the independent model formed as a random sum of losses from the univariate models.

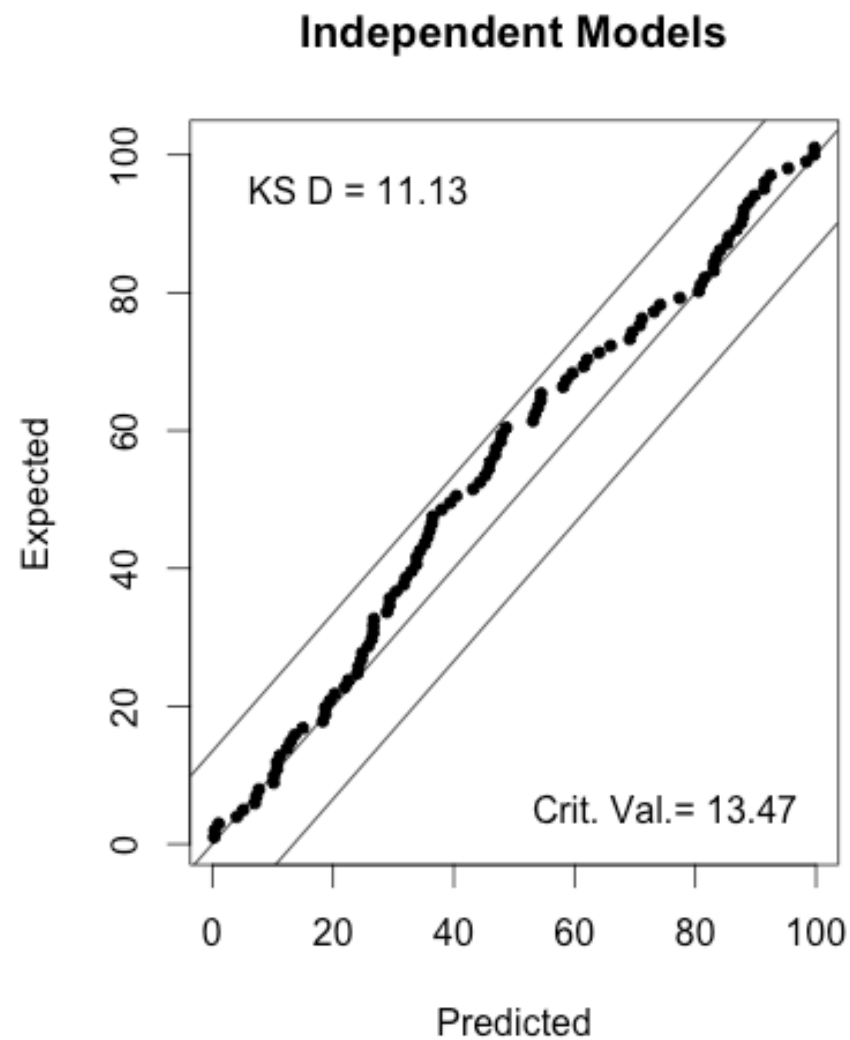
Test 2-Step Bivariate Model on 102 Pairs of Lines in CAS Database





Test Independent Model on 102 Pairs of Lines in CAS Database





Model Selection

Choosing between 2-Step and Independent

- If we fit model, *f*, by maximum likelihood define

$$AIC = 2 \cdot p - 2 \cdot L\left(x \mid \hat{\theta}\right)$$

- Where
  - p is the number of parameters
  - $L\left(x \mid \hat{\theta}\right)$ is the maximum log-likelihood of the model specified by *f*.
- Lower AIC indicates a better fit
  - Encourages larger log-likelihood
  - Penalizes for increasing number of parameters

Model Selection with the WAIC Statistic

- If we have an MCMC model with parameters $\left\{\theta_i\right\}_{i=1}^{10,000}$

$$WAIC = 2 \cdot \hat{p}_{WAIC} - 2 \cdot \overline{\left\{L\left(x \mid \theta_i\right)\right\}_{i=1}^{10,000}}$$

- Where
  - $\hat{p}_{WAIC}$ is the ***effective*** number of parameters
    - Decreases as the prior distribution becomes more "informative" i.e. less influenced by the data.
  - $\overline{\left\{L\left(x \mid \theta_i\right)\right\}_{i=1}^{10,000}}$ = Average log-likelihood of the model

WAIC Calculations

- Done with R package "loo"
- LOOIC - Another model selection statistic similar to WAIC
  - Pareto Smoothed Importance Sampling
  - Leave one out
  - PSIS-LOO
  - Included in paper

Model Selection

Choosing between 2-Step and Independent

- WAIC and LOOIC statistics indicate that the independent model is preferred

for ALL 102 pairs of lines!

- Counterintuitive to many actuaries.
  - Inflation affects all claims
  - Cyclic effects
- I think I owe an explanation.

The Changing Settlement Rate (CSR) Model

- *logelr* ~ uniform(-5,0)
- **$\alpha_1 = 0, \alpha_w$ ~ normal(0,$\sqrt{10}$) for *w* = 2,…,10**
- β_{10} = 0, β_d ~ uniform(-5,5), for *d* = 1,…,9
- a_i ~ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as *d* increases
- $\mu_{w,d}$ = log(Premium$_w$)+*logelr* **$+ \alpha_w + \beta_d \cdot speedup_w$**
- $C_{w,d}$ ~ lognormal($\mu_{w,d}$, σ_d)

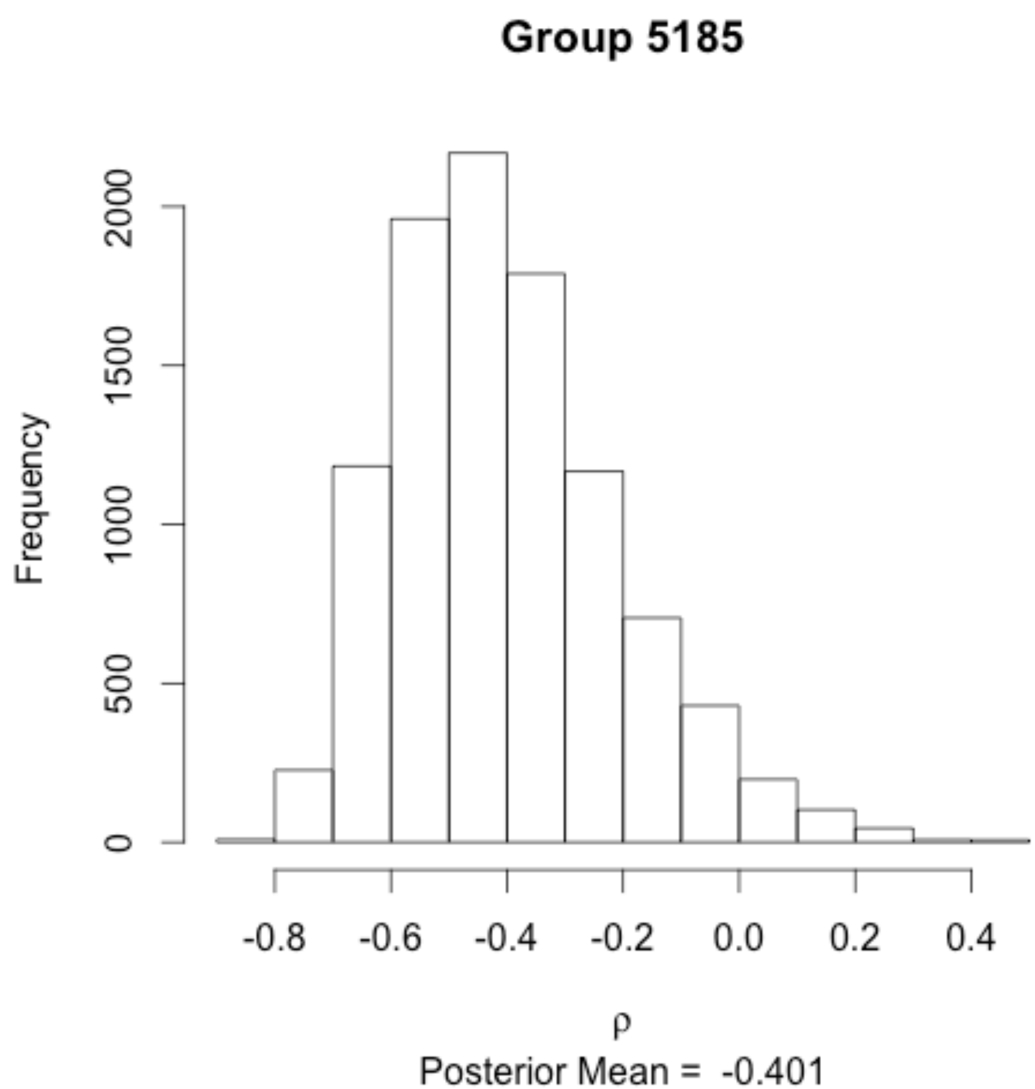
The Stochastic Cape Cod (SCC) Model

- *logelr* ~ uniform(-5,0)
- ~~$\alpha_1 = 0, \alpha_w$ ~ normal(0,$\sqrt{10}$) for w = 2,…,10~~
- $\beta_{10} = 0$, β_d ~ uniform(-5,5), for d = 1,…,9
- a_i ~ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases
- $\mu_{w,d} = \log(\text{Premium}_w) + logelr + \beta_d$
- $C_{w,d}$ ~ lognormal($\mu_{w,d}$, σ_d)

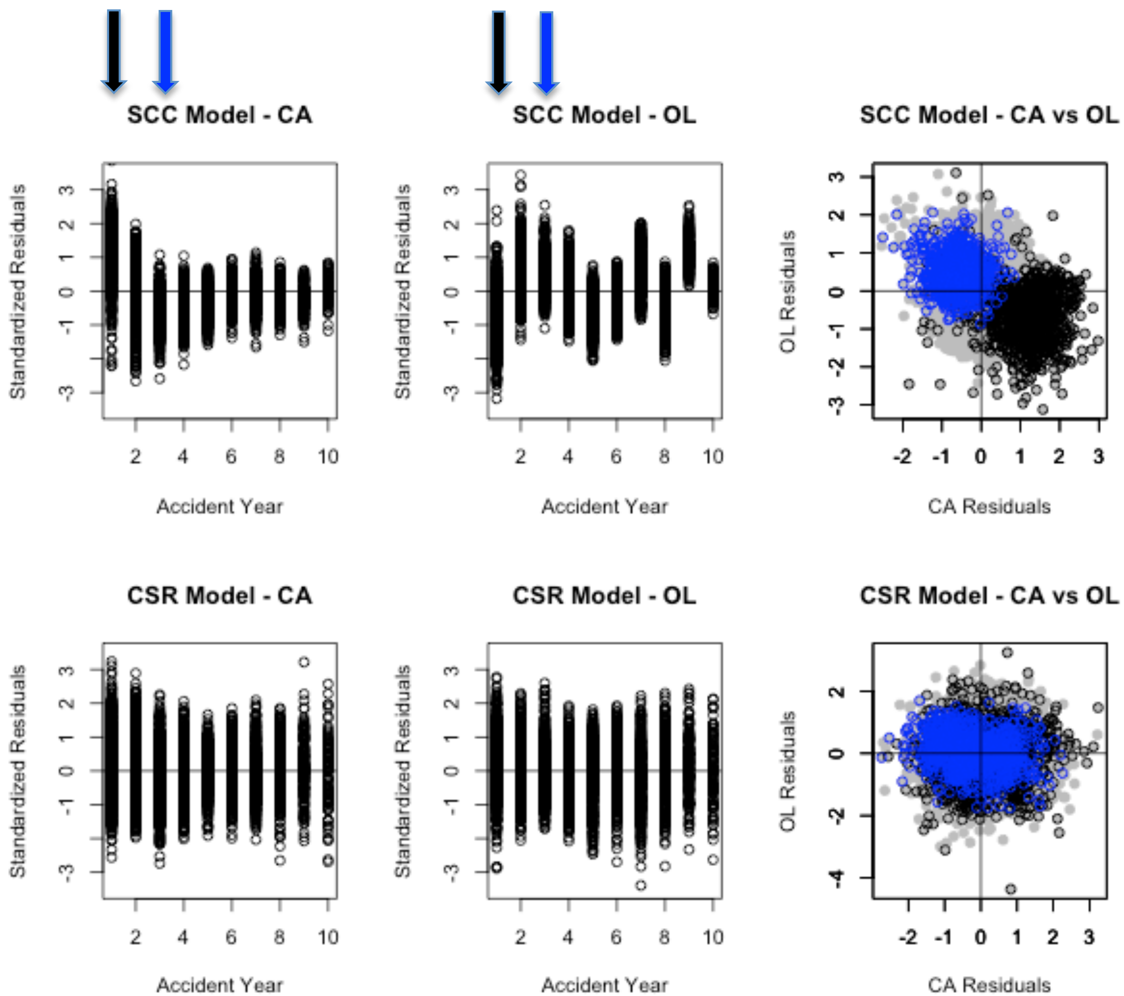
The Stochastic Cape Cod (SCC) Model

- Simpler than the CSR model
- Resembles an industry standard
  - Bornhuetter Ferguson with a constant ELR
    - Source Dave Clark and Jessica Leong in the references
- 2-Step SCC model is preferred for some insurers by the WAIC statistic.
- Look at a sample of standardized residual plots
- Insurer 5185 for CA and OL favors 2-Step
  - Picked as an illustration

Posterior Distribution of ρ



Sample of Standardized Residual Plots Insurer 5185 - CA and OL 2-Step



AY 1 borders are black

AY 3 borders are blue

In general, SCC residuals tend to find their own corner. If many are in the NW-SE corner, we see a negative mean ρ.

Implications of Independence

- Cost of capital risk margins should have a "diversification" credit.
- As an example, the EU Solvency II adds risk margins by line of business – implicitly denying a diversification credit.
- With a properly validated MCMC stochastic loss reserve model, one can get 10,000 stochastic scenarios of the future and calculate a cost of capital risk margin, and reflect diversification.
- I am preparing a paper on risk margins.
  - Session at the 2016 CLRS

A Proposed "Law" for Dependency Modeling

- Using the 2-Step procedure, we can fit multivariate distributions.
- We can compare the 2-Step model to a model that assumes independence with WAIC statistics.

The Law

- If your dependent bivariate model is "better" than the independent model, you should look for something that is missing from your model.

References by Glenn Meyers

- "Stochastic Loss Reserving Using Bayesian MCMC Models" CAS Monograph Series
  http://www.casact.org/pubs/monographs/index.cfm?fa=meyers-monograph01

- "Dependencies in Stochastic Loss Reserving Models" CAS eForum, Winter 2016. *This is a working paper, an updated paper has been submitted to Variance and the 2016 ASTIN Colloquium*
  http://www.casact.org/pubs/forum/16wforum/Meyers.pdf