CAS Monograph Series Stochastic Loss Reserving Using Bayesian MCMC Models

Glenn Meyers

May 18, 2015

May 17-20, 2015 • The Broadmoor • Colorado Springs, CO

2015 Spring Meeting

This Monograph is the Result of Bringing Together Two Concepts

- Retrospective testing with the CAS Loss Reserve Database
  - Hundreds of Schedule P Triangles
  - Includes outcomes obtained from subsequent NAIC Annual Statements
- Bayesian Markov-Chain Monte-Carlo MCMC
  - Truly a revolution is statistical modeling and computing.

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

Illustrative Insurer – Incurred Losses

		Cumulative Incurred Losses										
Premium	AY/Lag	1	2	3	4	5	6	7	8	9	10	Source
5812	1988	1722	3830	3603	3835	3873	3895	3918	3918	3917	3917	1997
4908	1989	1581	2192	2528	2533	2528	2530	2534	2541	2538	2532	1998
5454	1990	1834	3009	3488	4000	4105	4087	4112	4170	4271	4279	1999
5165	1991	2305	3473	3713	4018	4295	4334	4343	4340	4342	4341	2000
5214	1992	1832	2625	3086	3493	3521	3563	3542	3541	3541	3587	2001
5230	1993	2289	3160	3154	3204	3190	3206	3351	3289	3267	3268	2002
4992	1994	2881	4254	4841	5176	5551	5689	5683	5688	5684	5684	2003
5466	1995	2489	2956	3382	3755	4148	4123	4126	4127	4128	4128	2004
5226	1996	2541	3307	3789	3973	4031	4157	4143	4142	4144	4144	2005
4962	1997	2203	2934	3608	3977	4040	4121	4147	4155	4183	4181	2006

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

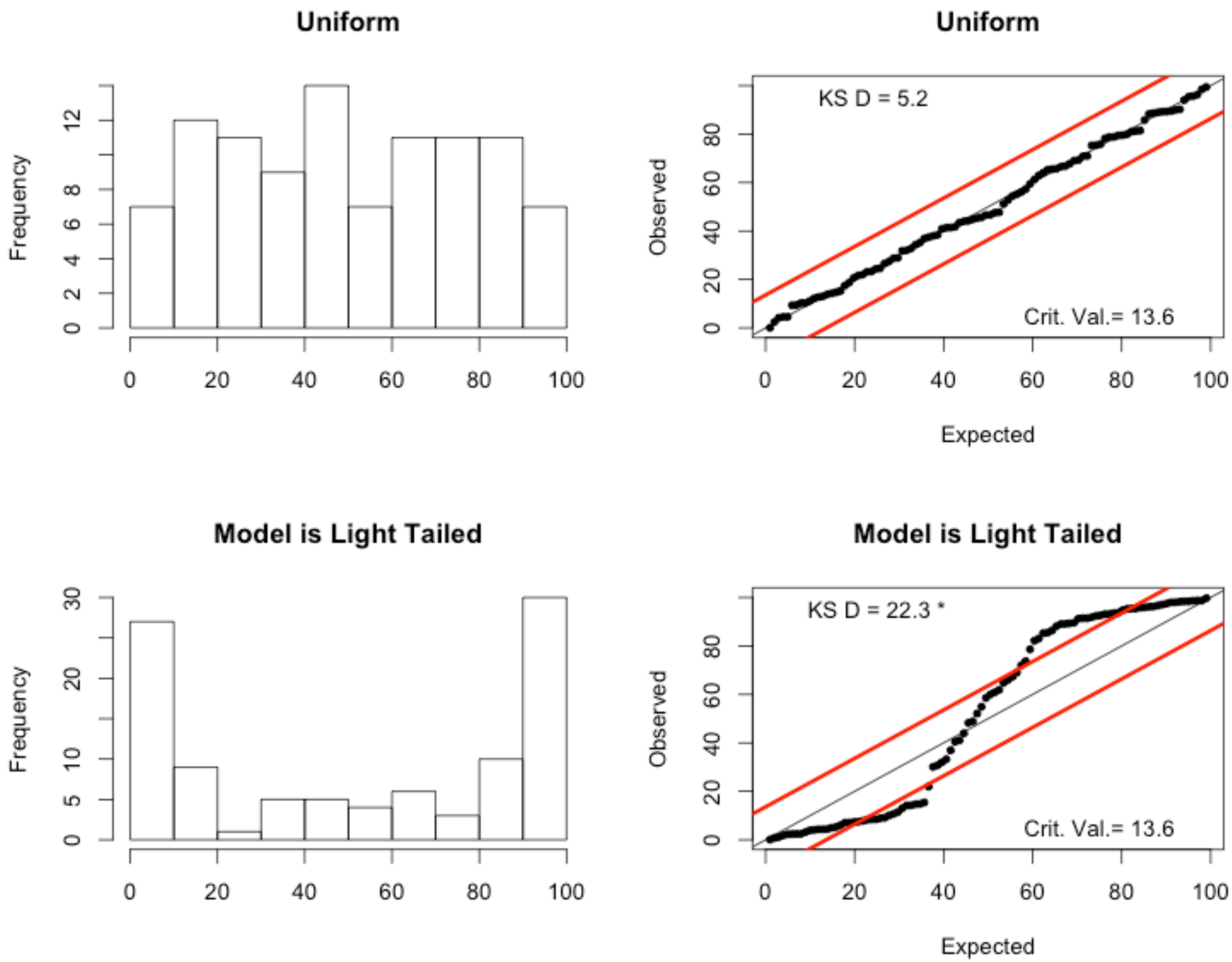
Criteria for a "Good" Stochastic Loss Reserve Model

- Using the upper triangle "training" data, predict the distribution of the outcomes in the lower triangle
  - Can be observations from individual (AY, Lag) cells
  - Can be sums of observations in different (AY,Lag) cells.

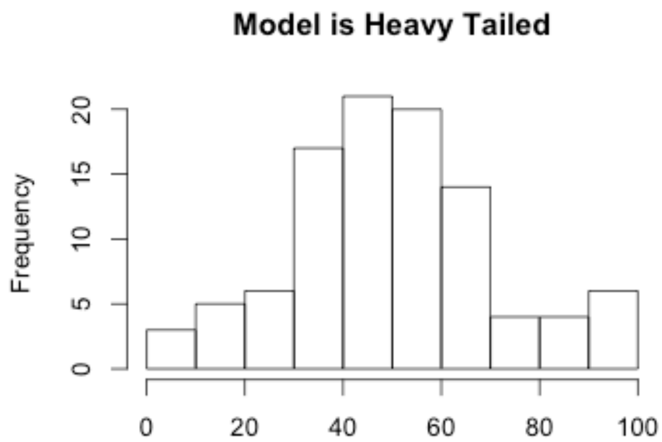
Criteria for a "Good" Stochastic Loss Reserve Model

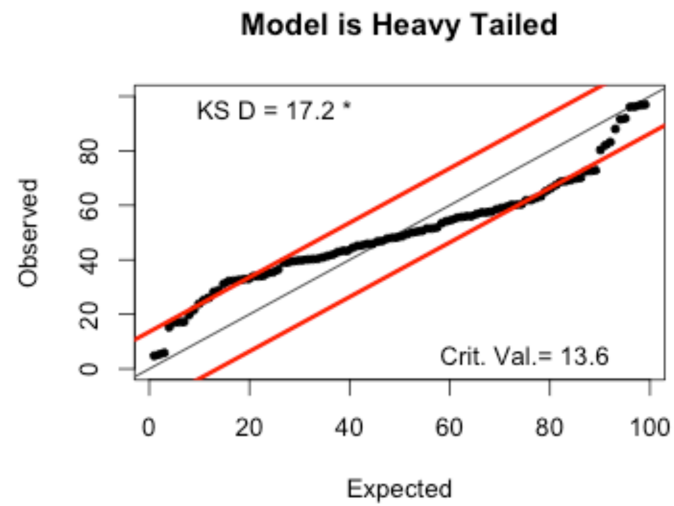
- Using the predictive distributions, find the percentiles of the outcome data for several loss triangles.
- The percentiles should be uniformly distributed.
  - Histograms
  - PP Plots and Kolmogorov Smirnov Tests
    - Plot Expected vs Predicted Percentiles
    - KS 95% critical values = 19.2 for *n*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

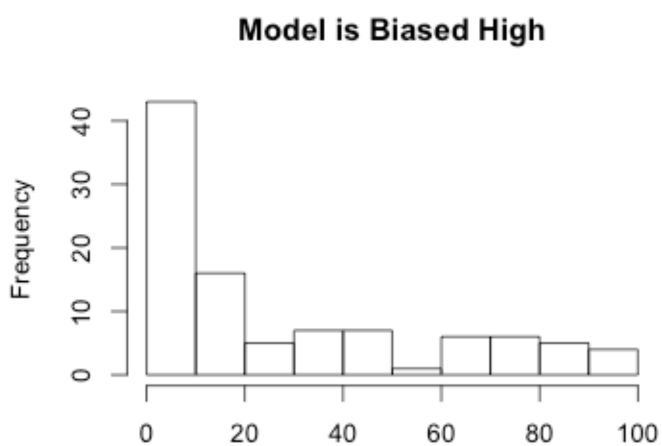
Illustrative Tests of Uniformity

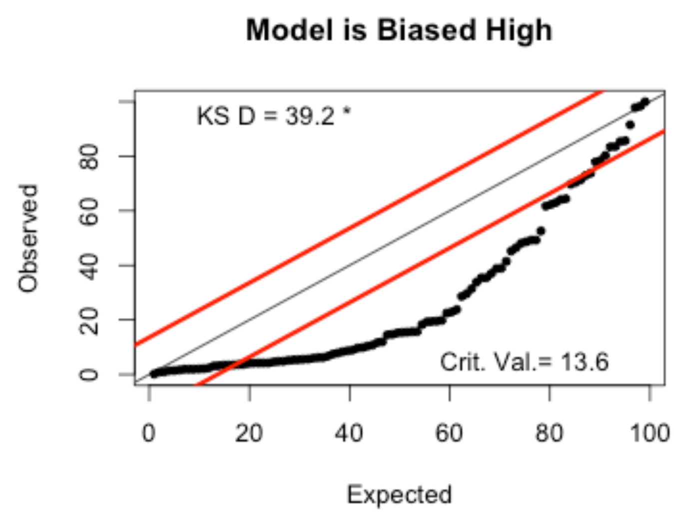


Illustrative Tests of Uniformity





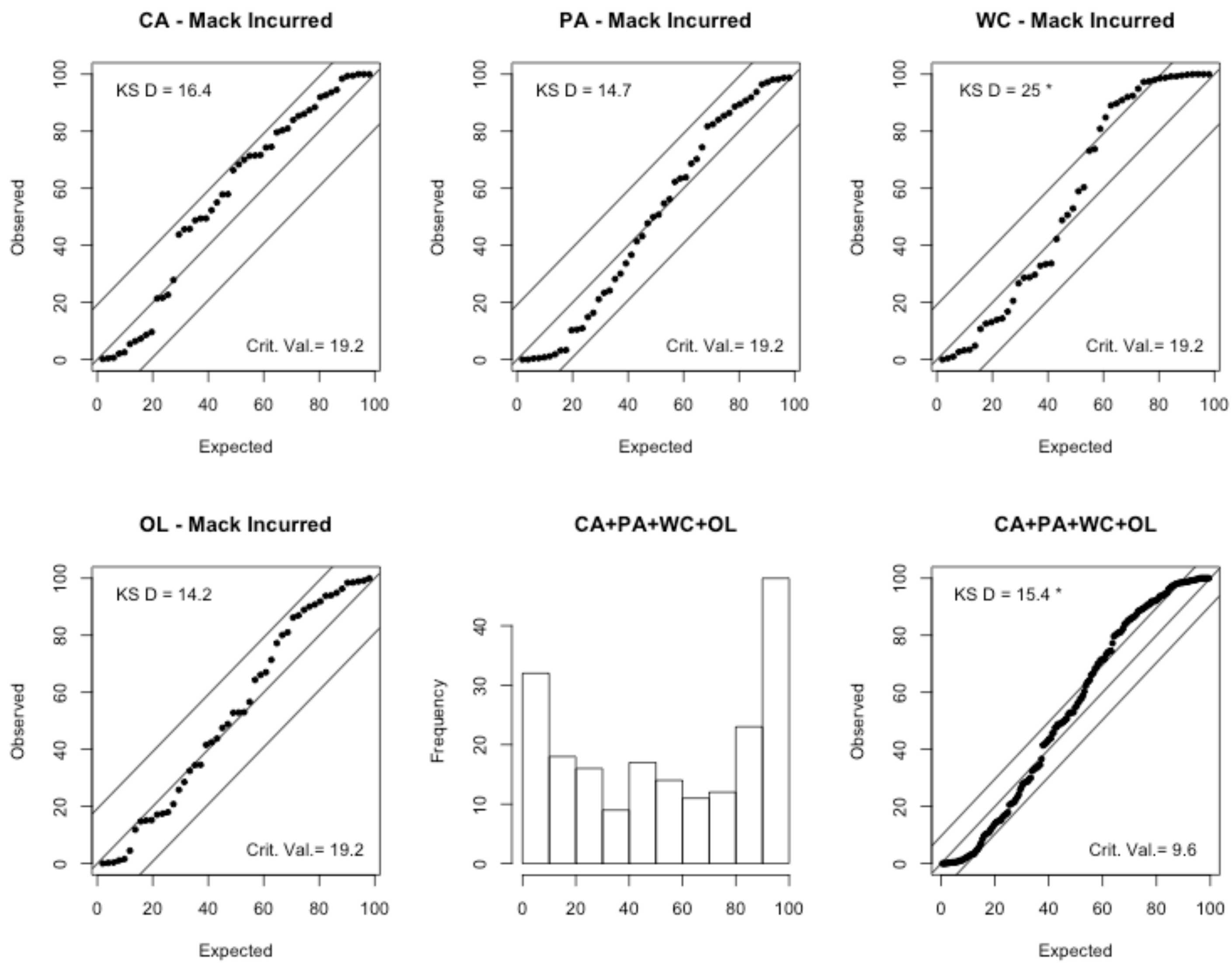




Data Used in Study

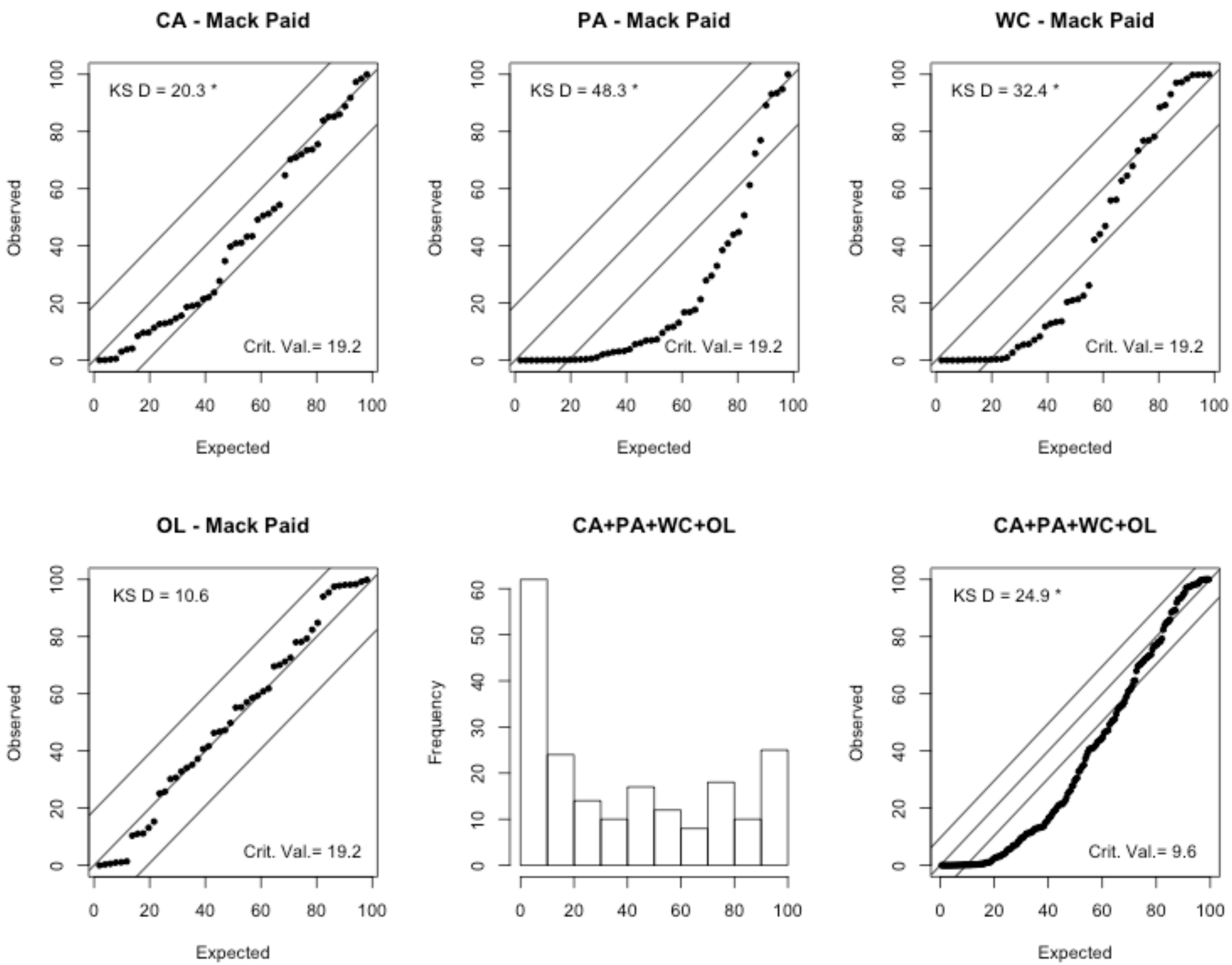
- List of insurers available in Appendix A.
- 50 Insurers from four lines of business
  - Commercial Auto
  - Personal Auto
  - Workers' Compensation
  - Other Liability
- Criteria for Selection
  - All 10 years of data available
  - Stability of earned premium and net to direct premium ratio
- Both paid and incurred losses

Test of Mack Model on Incurred Data



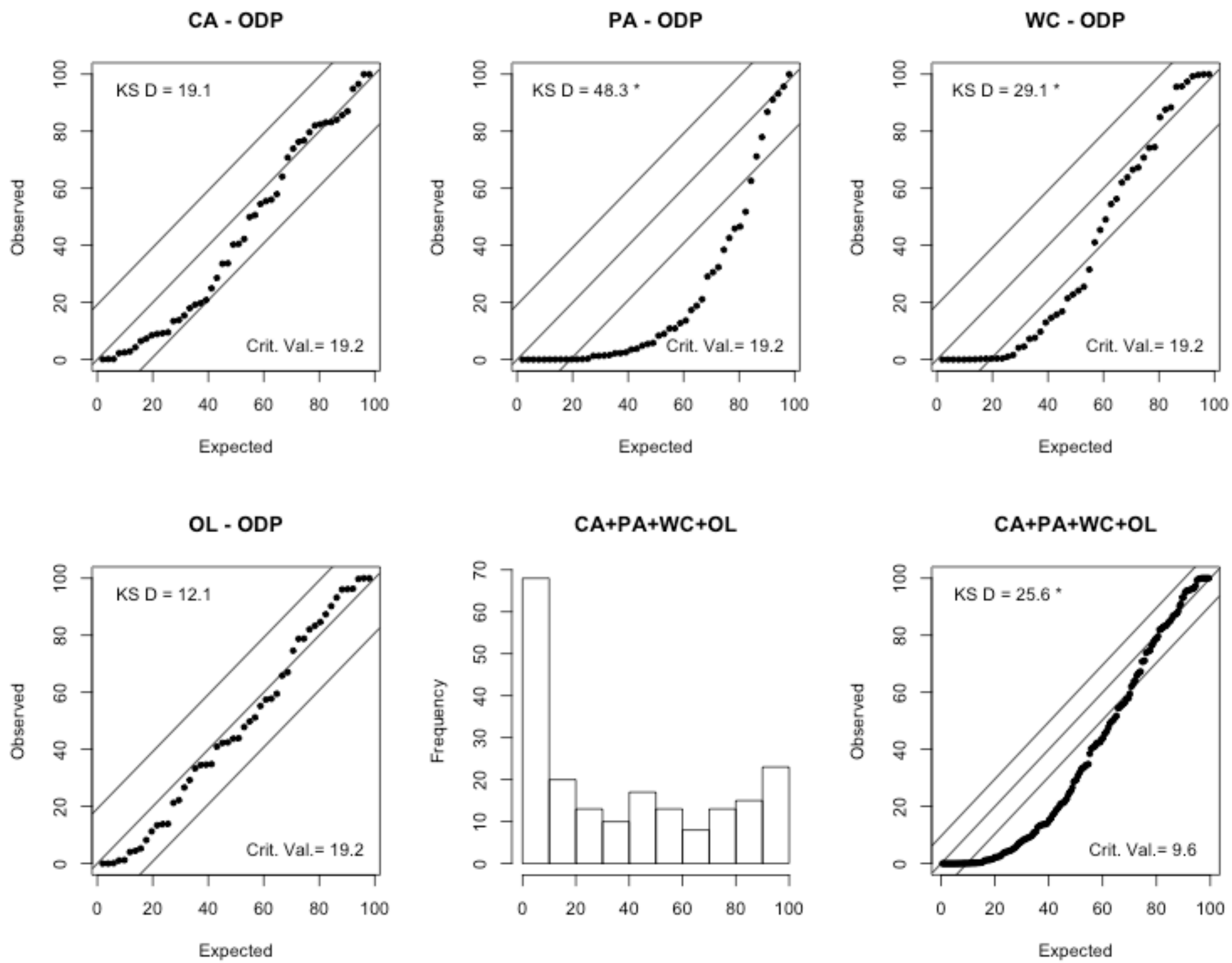
Conclusion – The Mack model predicts light tails.

Test of Mack Model on Paid Data



Conclusion – The Mack model is biased upward.

Test of Bootstrap ODP on Paid Data



Conclusion – The Bootstrap ODP model is biased upward.

Possible Responses to the Model Failures

- The “Black Swans” got us again!
  - We do the best we can in building our models, but the real world keeps throwing curve balls at us.
  - Every few years, the world gives us a unique “black swan” event.
- Build a better model.
  - Use a model, or data, that sees the “black swans.”

The Problem With Bayesian Analyses

Particularly Applicable to Loss Reserving

- Let θ be an *n*-parameter vector (e.g. development factors).
- Let X be a set of observations (e.g. a loss development triangle).

$$f(\theta | X) = \frac{f(X | \theta) \cdot \pi(\theta)}{\int_{\vartheta_1} \cdots \int_{\vartheta_n} f(X | \vartheta) \cdot \pi(\vartheta) \cdot d\vartheta}$$

- $f(X|\theta)$ is the likelihood of X given θ.
- $\pi(\theta)$ is the prior distribution of θ.
- $f(\theta|X)$ is the posterior distribution of θ.
- Calculating the *n*-dimensional integral is intractable.

A New World Order

- This impasse came to an end ~1990 when a simulation-based approach to estimating posterior probabilities was introduced.
  - (Circa the fall of the Soviet empire and Francis Fukuyama's "end of history")

Sampling-Based Approaches to Calculating Marginal Densities

ALAN E. GELFAND AND ADRIAN F. M. SMITH*

Journal of the American Statistical Association
June 1990, Vol. 85, No. 410, Theory and Methods

Markov Chains

- Let Ω be a finite state with random events

$$X_1, X_2, \ldots, X_t, \ldots$$

- A Markov chain *P* satisfies

$$\Pr\{X_t = y | X_{t-1} = x, \ldots, X_1 = x_1\} = \Pr\{x_t = y | x\} \equiv P(x,y)$$

  - The probability of an event in the chain depends only on the immediate previous event.
  - *P* is called a transition matrix

Markov Chains

The Markov Convergence Theorem

- There is a branch of probability theory, called Ergodic Theory, that gives conditions for which there exists a unique stationary distribution π such that

$$P^t(x,y) \rightarrow \pi(y) \text{ as } t \rightarrow \infty.$$

The Metropolis Hastings Algorithm
A Very Important Markov Chain

1. Time t=1: select a random initial position θ_1 in parameter space.
2. Select a **proposal distribution** $p(\theta|\theta_{t-1})$ that we will use to select proposed random steps away from our current position in parameter space.
3. Starting at time t=2: repeat the following until you get convergence:
   a) At step t, generate a proposal $\theta^* \sim p(\theta|\theta_{t-1})$
   b) Generate $U \sim$ uniform(0,1)
   c) Calculate $$R = \frac{f(\theta^* \mid X)}{f(\theta_{t-1} \mid X)} \cdot \frac{p(\theta_{t-1} \mid \theta^*)}{p(\theta^* \mid \theta_{t-1})}$$
   d) If $U < R$ then $\theta_t = \theta^*$. Else, $\theta_t = \theta_{t-1}$.

Dodging the Intractable Integral

$$R = \frac{f(\theta^* \mid X)}{f(\theta_{t-1} \mid X)} \cdot \frac{p(\theta_{t-1} \mid \theta^*)}{f(\theta^* \mid \theta_{t-1})}$$

$$R = \frac{\dfrac{f\left(X \mid \theta^*\right) \cdot \pi\left(\theta^*\right)}{\cancel{\int_{\vartheta_1} \cdots \int_{\vartheta_n} f\left(X \mid \vartheta\right) \cdot \pi\left(\vartheta\right) \cdot d\vartheta}}}{\dfrac{f\left(X \mid \theta_{t-1}\right) \cdot \pi\left(\theta_{t-1}\right)}{\cancel{\int_{\vartheta_1} \cdots \int_{\vartheta_n} f\left(X \mid \vartheta\right) \cdot \pi\left(\vartheta\right) \cdot d\vartheta}}} \cdot \frac{p(\theta_{t-1} \mid \theta^*)}{p(\theta^* \mid \theta_{t-1})}$$

The Metropolis Hastings Algorithm Restated

1. Time t=1: select a random initial position θ_1 in parameter space.
2. Select a **proposal distribution** $p(\theta|\theta_{t-1})$ that we will use to select proposed random steps away from our current position in parameter space.
3. Starting at time t=2: repeat the following until you get convergence:
   a) At step t, generate a proposed $\theta^* \sim p(\theta|\theta_{t-1})$
   b) Generate $U \sim$ uniform(0,1)
   c) Calculate $$R = \frac{f(X|\theta^*)\pi(\theta^*)}{f(X|\theta_{t-1})\pi(\theta_{t-1})} \cdot \frac{p(\theta_{t-1}|\theta^*)}{p(\theta^*|\theta_{t-1})}$$
   d) If $U < R$ then $\theta_t = \theta^*$. Else, $\theta_t = \theta_{t-1}$.

The Relevance of the Metropolis Hastings Algorithm

- Defined in terms of the conditional distribution

$$f(X|\theta)$$

and the prior distribution

$$\pi(\theta)$$

- The limiting distribution is the *posterior distribution*!
- Program $f(X|\theta)$ and $\pi(\theta)$ into a Markov chain and let it run for a while, and you have a large sample from the posterior distribution.

The Relevance of the Metropolis Hastings Algorithm

- The theoretical limiting distribution is the same, no matter what proposal distribution, $p(\theta|\theta_{t-1})$, is used.
- However, some proposal distributions will get a good sample much faster!
- There is no fundamental limit on the number of parameters in you model!
- The practical limit is within range of stochastic loss reserve models.

Metropolis Hastings in Practice

- "Tune" the proposal distribution, $p(\theta_t|\theta_{t\text{-}1})$, to minimize autocorrelation between θ_t and $\theta_{t\text{-}1}$.
- Convergence - Determine the interval *t* to *t*+*m* that contains a representative sample of the posterior distribution.
- There are several software packages for Bayesian MCMC that work with the R programming language.
  - WINBUGS
  - OpenBUGS
  - JAGS
  - Stan (New – In honor of Stanislaw Ulam)

The Situation Prior to My Retirement

- We have the CAS Loss Reserve Database
  - Hundreds of loss reserve triangles ***with outcomes***.
- The current "best practice" models do not correctly predict the distribution of outcomes.
- We have a game changing statistical model fitting methodology – Bayesian MCMC!

Begin with Incurred Data Models

Notation

- w = Accident Year w = 1,…,10
- d = Development Year d = 1,…,10
- $C_{w,d}$ = Cumulative (either incurred or paid) loss
- $I_{w,d}$ = Incremental paid loss = $C_{w,d} - C_{w-1,d}$

Bayesian MCMC Models

- Use R and JAGS (Just Another Gibbs Sampler) packages
- Get a sample of 10,000 parameter sets from the posterior distribution of the model
- 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

The Correlated Chain Ladder (CCL) Model

- $logelr \sim$ uniform(-5,0.5)
- $\alpha_w \sim$ normal(log(Premium$_w$)+$logelr$, $\sqrt{10}$)
- $\beta_{10} = 0$, $\beta_d \sim$ uniform(-5,5), for d=1,…,9
- $a_i \sim$ uniform(0,1)
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases
- $\mu_{1,d} = \alpha_1 + \beta_d$
- $C_{1,d} \sim$ lognormal($\mu_{1,d}$, σ_d)
- $\rho_d \sim$ uniform(-1,1)
- $\mu_{w,d} = \alpha_w + \beta_d + \rho_d \cdot (\log(C_{w-1,d}) - \mu_{w-1,d})$ for w = 2,…,10
- $C_{w,d} \sim$ lognormal($\mu_{w,d}$, σ_d)

Key Statements in the CCL Model

$$\mu_{w,d} = \alpha_w + \beta_d + \rho \cdot (\log(C_{w-1,d}) - \mu_{\mathrm{w-1,d}})$$

for $w = 2,\ldots,10$

$$C_{w,d} \sim \mathrm{lognormal}(\mu_{w,d}, \sigma_d)$$

The Correlated Chain Ladder Model Predicts Distributions with Thicker Tails

- Mack uses point estimations of parameters.
- CCL uses Bayesian estimation to get a posterior distribution of parameters.
- Chain ladder applies factors to last ***fixed*** observation.
- CCL uses ***uncertain*** "level" parameters for each accident year.
- Mack assumes independence between accident years.
- CCL allows for correlation between accident years,
  - $Corr[\log(C_{w-1,d}),\log(C_{w,d})] = \rho$

Predicting the Distribution of Outcomes

- Use JAGS software to produce a sample of 10,000 $\{\alpha_w\}$, $\{\beta_d\}$,$\{\sigma_d\}$ and $\{\rho\}$ from the posterior distribution.
- For each member of the sample
  - $\mu_{1,10} = \alpha_1 + \beta_{10}$
  - For w = 2 to 10
    - $C_{w,10} \sim \text{lognormal}\left(\alpha_w + \beta_{10} + \rho_d \cdot (\log(C_{w-1,10}) - \mu_{w-1,10})), \sigma_{10}\right)$
  - Calculate $\sum_{w=1}^{10} C_{w,10}$
- Calculate summary statistics, e.g. $E\left[\sum_{w=1}^{10} C_{w,10}\right]$ and $Var\left[\sum_{w=1}^{10} C_{w,10}\right]$
- Calculate the percentile of the actual outcome by counting how many of the simulated outcomes are below the actual outcome.

The First 5 of 10,000 Samples on Illustrative Insurer

Done in JAGS →

	MCMC Sample Number				
	1	2	3	4	5
α_1	8.2763	8.2452	8.2390	8.2591	8.2295
α_2	7.8226	7.7812	7.8008	7.8048	7.7810
α_3	8.2625	8.3200	8.2929	8.2883	8.2642
α_4	8.3409	8.3286	8.3539	8.3622	8.3159
α_5	8.2326	8.1166	8.1093	8.1855	8.1523
α_6	8.1673	8.0307	8.0491	8.1727	8.0470
α_7	8.6403	8.4776	8.4113	8.5815	8.4871
α_8	8.2177	8.2488	8.2708	8.0752	8.1763
α_9	8.3174	8.2007	8.2589	8.3744	8.2653
α_{10}	7.4101	8.0036	8.7584	8.4241	8.8420
β_1	-0.5125	-0.5180	-0.6504	-0.4947	-0.7384
β_2	-0.2756	-0.1014	-0.1231	-0.2138	-0.0844
β_3	-0.1271	-0.0313	-0.0622	-0.0758	-0.0498
β_4	-0.1013	-0.0090	0.0165	0.0439	0.0479
β_5	0.0518	-0.0109	0.0060	0.0034	0.0610
β_6	0.0180	0.0885	0.0139	0.0175	0.0709
β_7	0.0105	0.0583	0.0205	0.0427	0.0362
β_8	0.0400	-0.0090	0.0612	0.0444	0.0338
β_9	0.0005	0.0287	0.0419	0.0116	0.0333
β_{10}	0.0000	0.0000	0.0000	0.0000	0.0000
σ_1	0.3152	0.2954	0.3164	0.1895	0.2791
σ_2	0.2428	0.1982	0.2440	0.1858	0.1711
σ_3	0.1607	0.1632	0.2078	0.1419	0.1089
σ_4	0.1245	0.1133	0.0920	0.0842	0.0800
σ_5	0.0871	0.0830	0.0694	0.0747	0.0794
σ_6	0.0733	0.0649	0.0626	0.0508	0.0463
σ_7	0.0324	0.0281	0.0294	0.0368	0.0352
σ_8	0.0279	0.0247	0.0172	0.0270	0.0330
σ_9	0.0171	0.0239	0.0130	0.0267	0.0329
σ_{10}	0.0170	0.0237	0.0105	0.0241	0.0244
ρ	0.1828	0.4659	0.4817	0.1901	0.2155

Done in R ↓

$\mu_{1,10}$	8.2763	8.2452	8.2390	8.2591	8.2295
$C_{1,10}$	3917	3917	3917	3917	3917
$\widetilde{\mu}_{2,10}$	7.8221	7.7942	7.8172	7.8074	7.7904
$\widetilde{C}_{2,10}$	2520	2468	2480	2432	2453
$\widetilde{\mu}_{3,10}$	8.2643	8.3278	8.2924	8.2862	8.2674
$\widetilde{C}_{3,10}$	3893	4190	3939	4090	3802
$\widetilde{\mu}_{4,10}$	8.3414	8.3345	8.3474	8.3679	8.3107
$\widetilde{C}_{4,10}$	4229	4212	4233	4346	4075
$\widetilde{\mu}_{5,10}$	8.2341	8.1219	8.1109	8.1873	8.1527
$\widetilde{C}_{5,10}$	3761	3285	3269	3597	3676
$\widetilde{\mu}_{6,10}$	8.1670	8.0192	8.0400	8.1728	8.0593
$\widetilde{C}_{6,10}$	3450	3127	3120	3552	3196
$\widetilde{\mu}_{7,10}$	8.6365	8.4910	8.4140	8.5819	8.4893
$\widetilde{C}_{7,10}$	5488	4719	4441	5299	4765
$\widetilde{\mu}_{8,10}$	8.2129	8.2340	8.2634	8.0739	8.1720
$\widetilde{C}_{8,10}$	3652	3847	3933	3295	3708
$\widetilde{\mu}_{9,10}$	8.3156	8.2106	8.2655	8.3794	8.2752
$\widetilde{C}_{9,10}$	4112	3538	3949	4426	3914
$\widetilde{\mu}_{10,10}$	7.4112	7.9853	8.7659	8.4271	8.8414
$\widetilde{C}_{10,10}$	1613	3001	6511	4507	6763

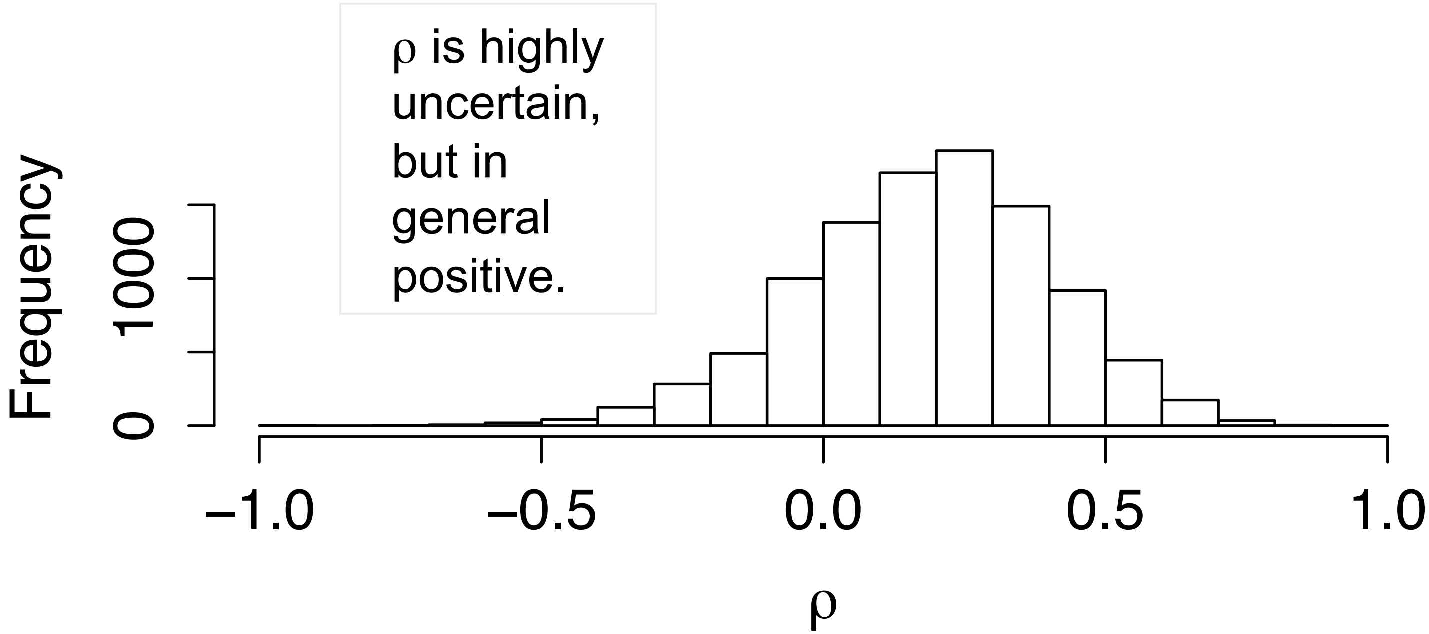
Results of “Tail Pumping”

	CCL - $\rho = 0$			Mack			
w	Estimate	Std. Dev.	CV	Estimate	Std. Dev.	CV	Outcome
1	3917	0	0.000	3917	0	0.000	3,917
2	2544	59	0.023	2538	0	0.000	2,532
3	4110	106	0.026	4167	3	0.001	4,279
4	4307	122	0.028	4367	37	0.009	4,341
5	3545	115	0.032	3597	34	0.010	3,587
6	3317	132	0.040	3236	40	0.012	3,268
7	5315	265	0.050	5358	146	0.027	5,684
8	3775	301	0.080	3765	225	0.060	4,128
9	4203	561	0.134	4013	412	0.103	4,144
10	4084	1157	0.283	3955	878	0.222	4,181
Total	39116	1551	0.040	38914	1,057	0.027	40,061
Percentile		76.38			86.03		

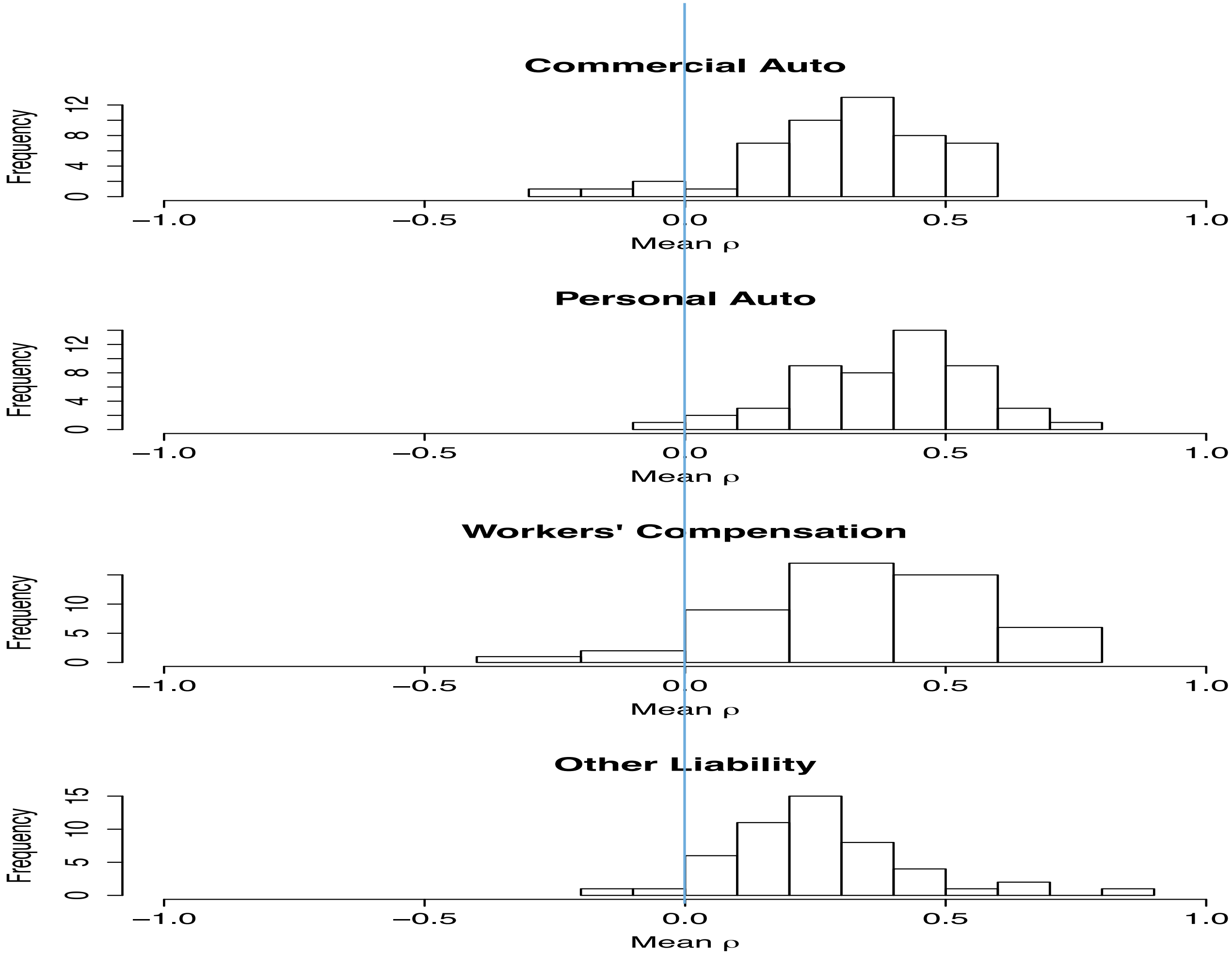
Results of "Tail Pumping"

	CCL			CCL - $\rho = 0$			
w	Estimate	Std. Dev.	CV	Estimate	Std. Dev.	CV	Outcome
1	3917	0	0.000	3917	0	0.000	3,917
2	2545	57	0.022	2544	59	0.023	2,532
3	4110	113	0.028	4110	106	0.026	4,279
4	4314	130	0.030	4307	122	0.028	4,341
5	3549	123	0.035	3545	115	0.032	3,587
6	3319	146	0.044	3317	132	0.040	3,268
7	5277	292	0.055	5315	265	0.050	5,684
8	3796	331	0.087	3775	301	0.080	4,128
9	4180	622	0.149	4203	561	0.134	4,144
10	4155	1471	0.354	4084	1157	0.283	4,181
Total	39161	1901	0.049	39116	1551	0.040	40,061
Percentile		73.72			76.38		

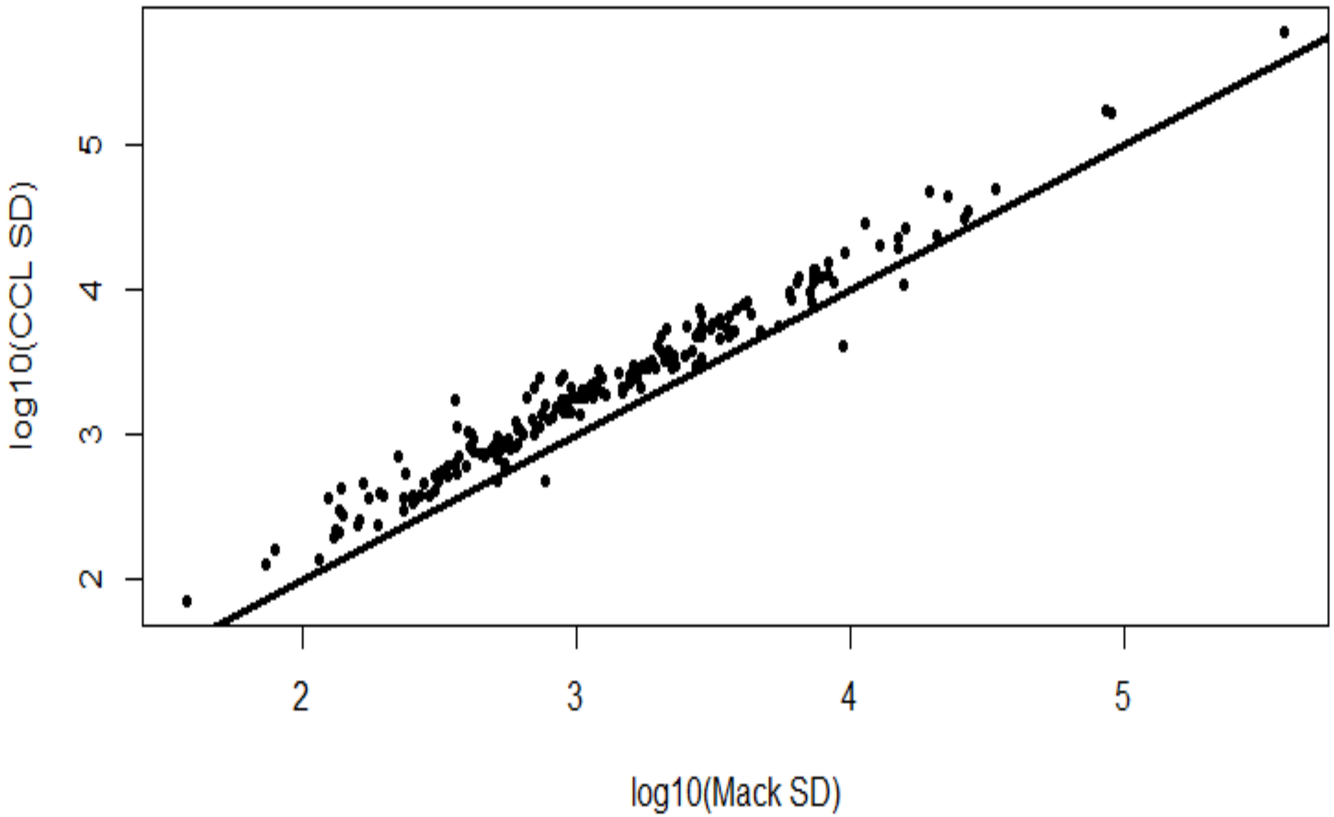
Posterior Distribution of ρ for Illustrative Insurer



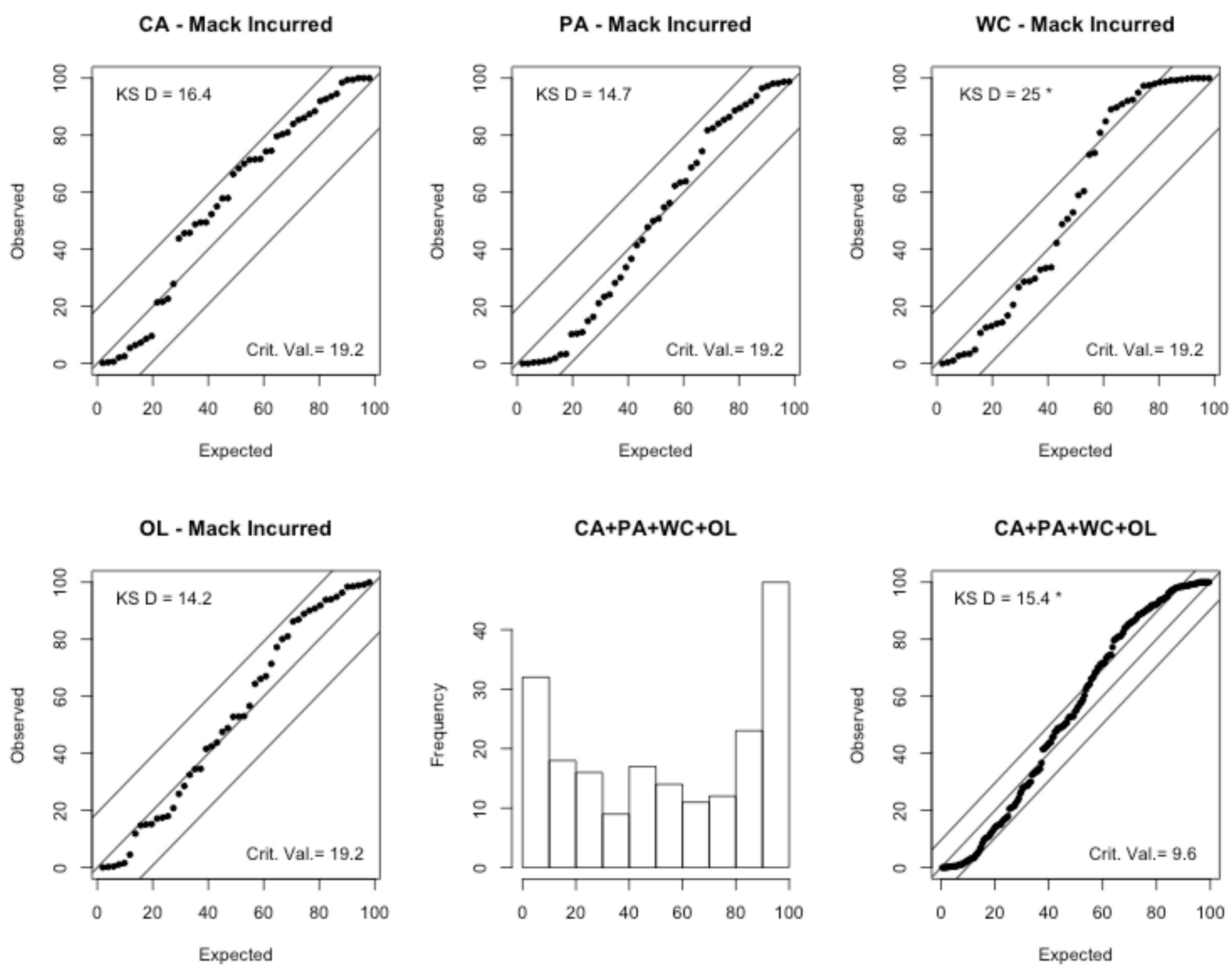
Generally Positive Posterior Means of ρ for all Insurers



Compare SDs for All 200 Triangles

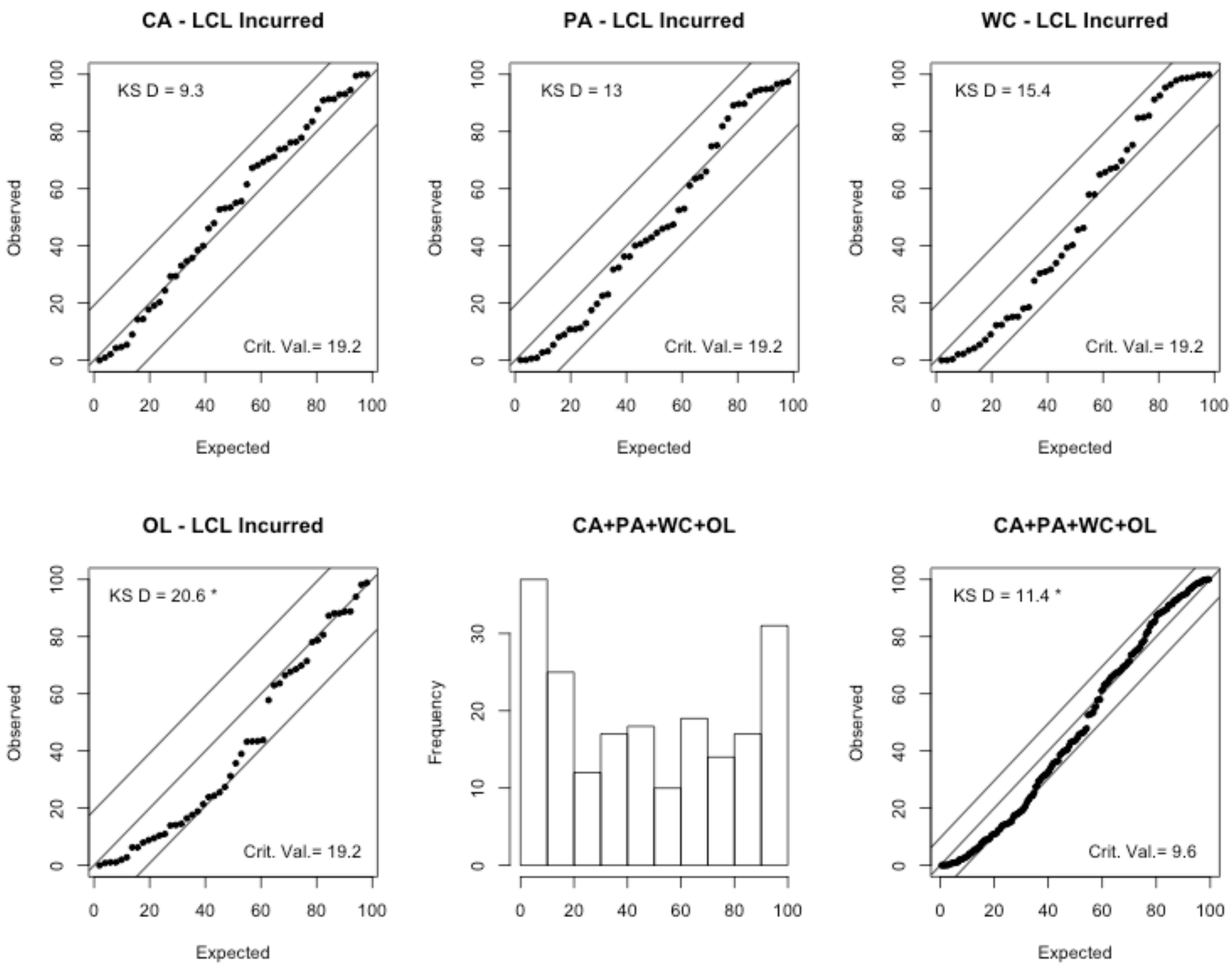


Test of Mack Model on Incurred Data



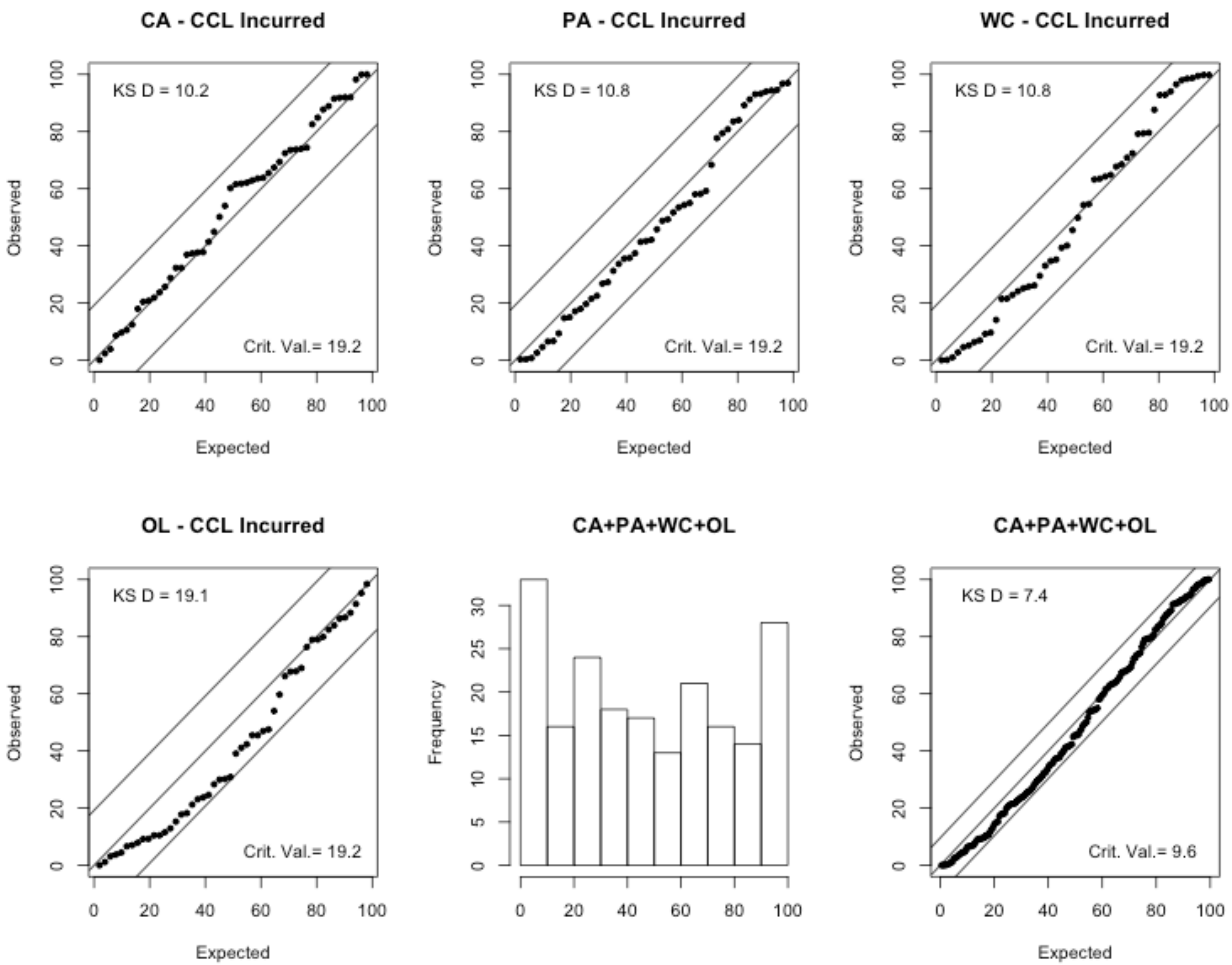
Conclusion – Predicted tails are too light

Test of CCL ($\rho = 0$) Model on Incurred Data



Conclusion – Predicted tails are too light

Test of CCL Model on Incurred Data



Conclusion – Plot is within KS Boundaries

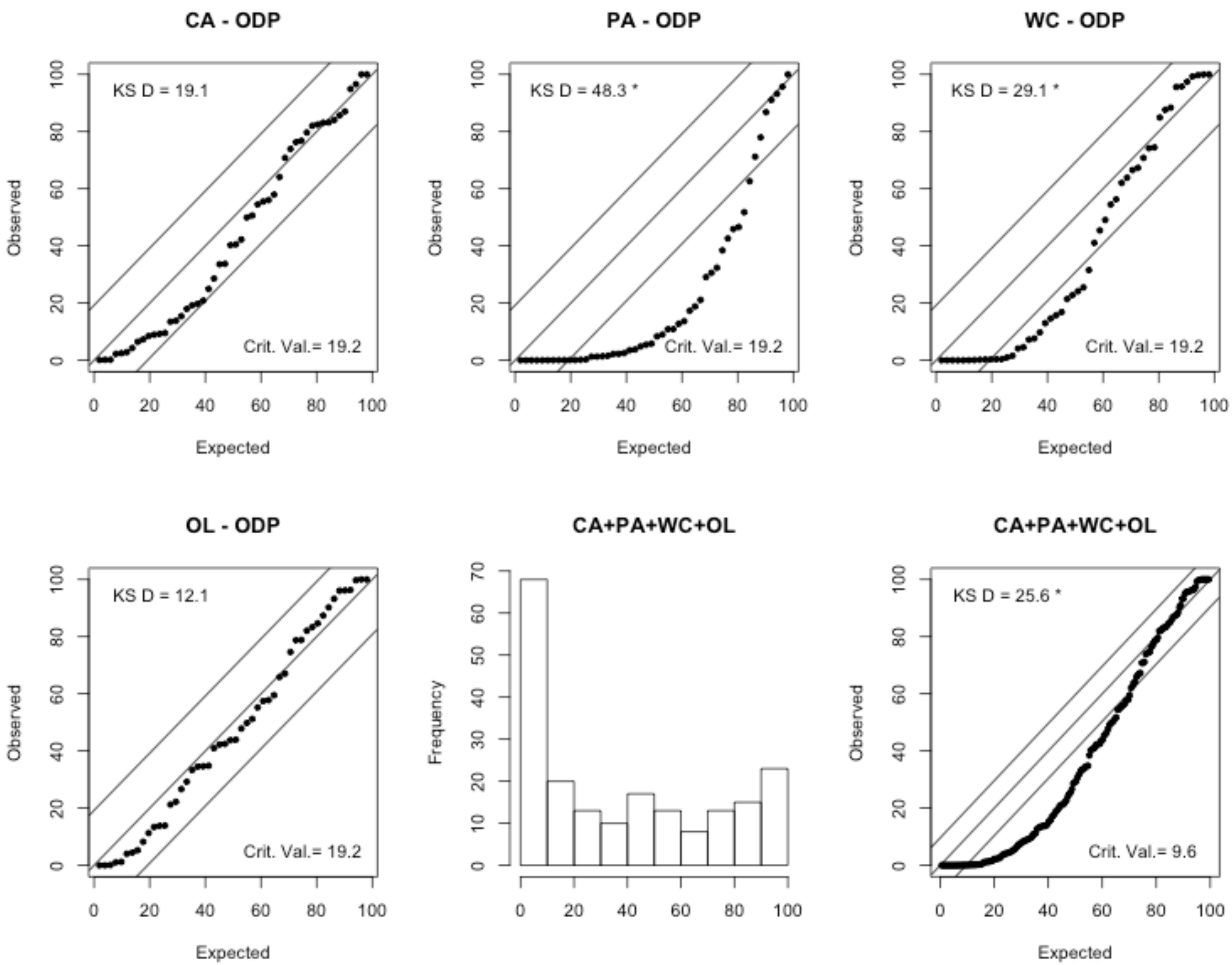
Identified Improvements with Incurred Data

- Accomplished by “pumping up” the variance of Mack model.

What About Paid Data?

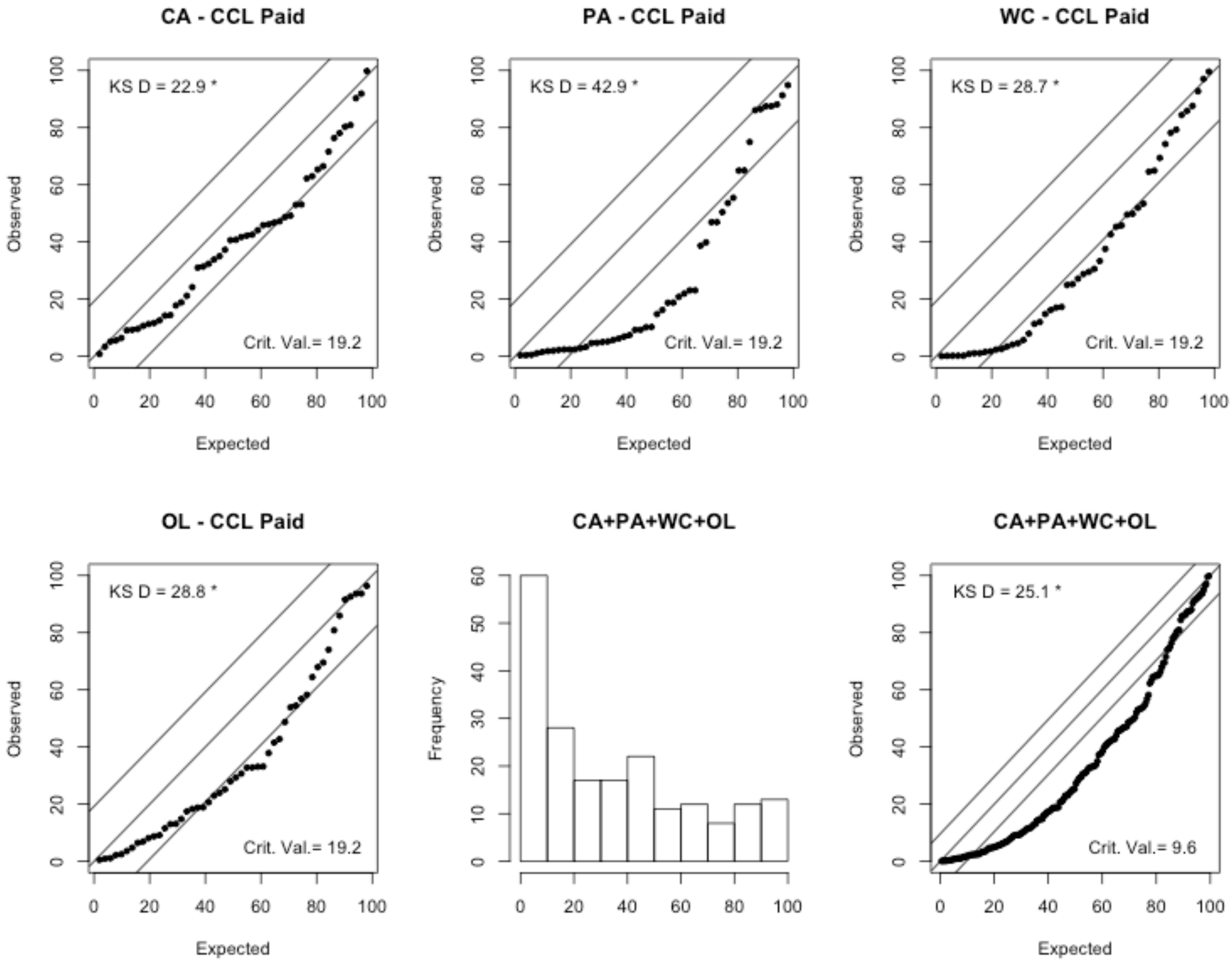
- Start by looking at CCL model on cumulative paid data.

Test of Bootstrap ODP on Paid Data



Conclusion – The Bootstrap ODP model is biased upward.

Test of CCL on Paid Data



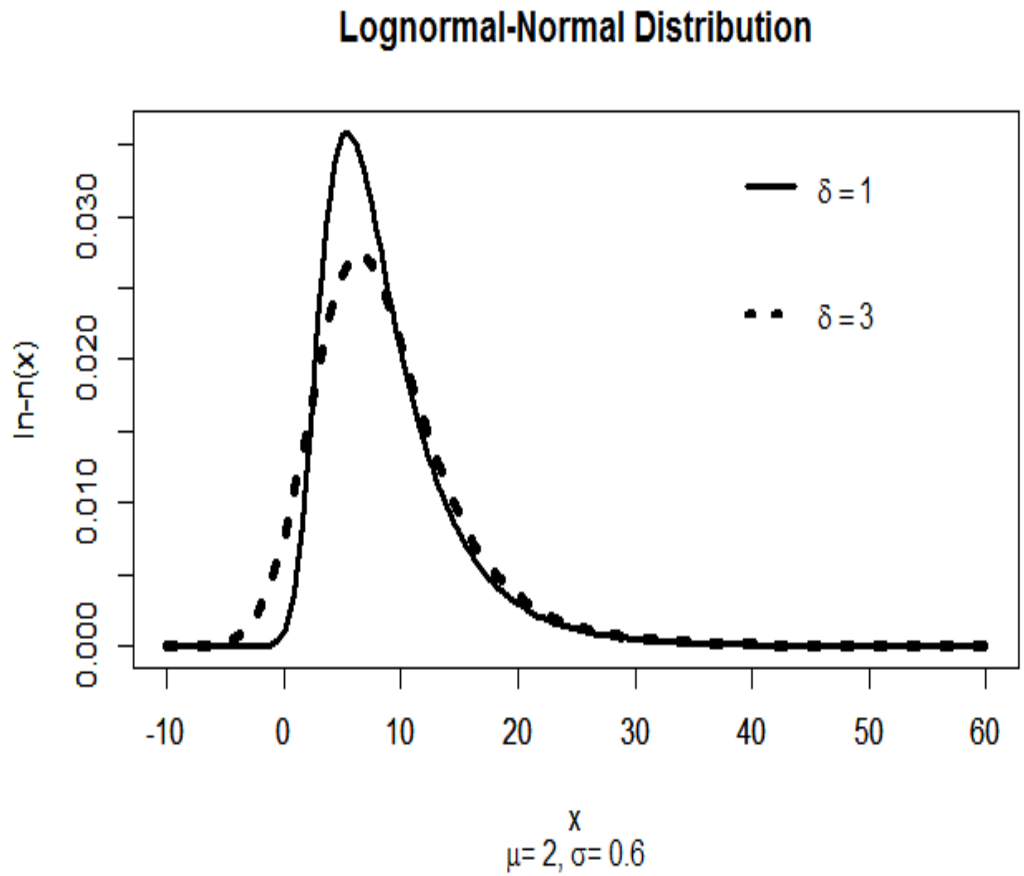
Conclusion – Biased upward.

How Do We Correct the Bias?

- Look at models with payment year trend.
  - Ben Zehnwirth has been championing these for years.
- Payment year trend does not make sense with cumulative data!
  - Settled claims are unaffected by trend.
- Recurring problem with incremental data – Negatives!
  - We need a skewed distribution that has support over the entire real line.

The Lognormal-Normal (ln-n) Mixture

$Z \sim \text{Lognormal}(\mu,\sigma)$, $X \sim \text{Normal}(Z,\delta)$



The Correlated Incremental Trend (CIT) Model

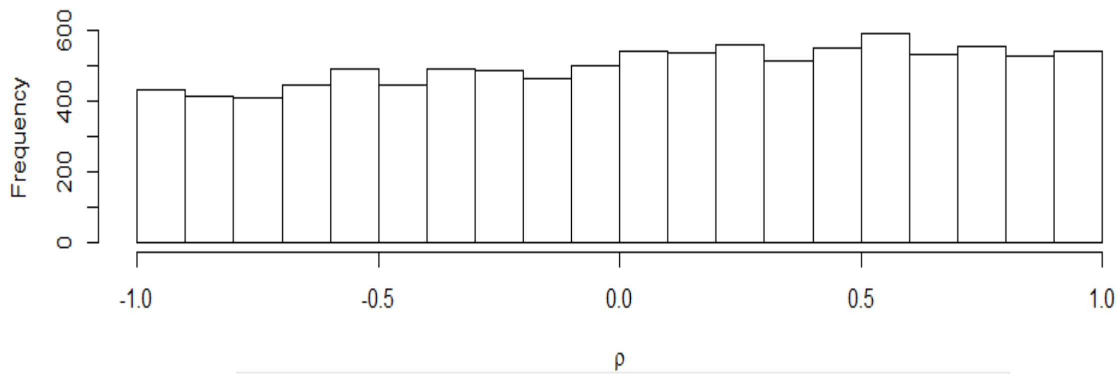
- $\mu_{w,d} = \alpha_w + \beta_d + \tau \cdot (w + d - 1)$
- $Z_{w,d} \sim \text{lognormal}(\mu_{w,d}, \sigma_d)$ subject to $\sigma_1 < \sigma_2 < \ldots < \sigma_{10}$
- $I_{1,d} \sim \text{normal}(Z_{1,d}, \delta)$
- $I_{w,d} \sim \text{normal}(Z_{w,d} + \rho \cdot (I_{w-1,d} - Z_{w-1,d}) \cdot e^{\tau}, \delta)$
- Estimate the distribution of $\sum_{w=1}^{10} C_{w,10}$
- "Sensible" priors
  - Needed to control σ_d
  - Interaction between τ , α_w and β_d.

CIT Model for Illustrative Insurer

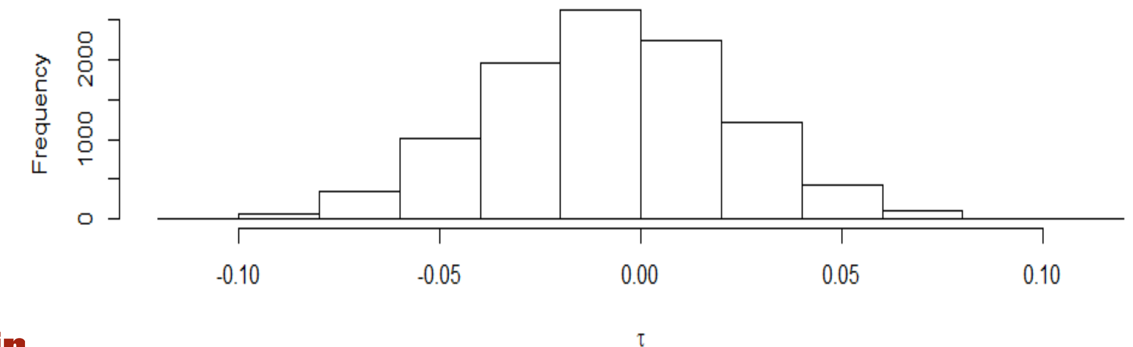
	CIT				CCL				Outcome
w	$C_{w,10}$	SD	CV		$C_{w,10}$	SD	CV		$C_{w,10}$
1	3912	0	0		3912	0	0.0000		3912
2	2536	5	0.002		2563	110	0.0429		2527
3	4175	11	0.0026		4153	189	0.0455		4274
4	4378	29	0.0066		4320	224	0.0519		4341
5	3539	35	0.0099		3570	207	0.0580		3583
6	3043	105	0.0345		3403	255	0.0749		3268
7	5037	114	0.0226		5207	465	0.0893		5684
8	3501	556	0.1588		3649	467	0.1280		4128
9	3980	710	0.1784		4409	895	0.2030		4144
10	4661	1484	0.3184		5014	2435	0.4856		4139
Total	38763	1803	0.0465		40200	3070	0.0764		40000
Percentile		81.87				51.24			

Posterior Distribution of μ and τ for Illustrative Insurer

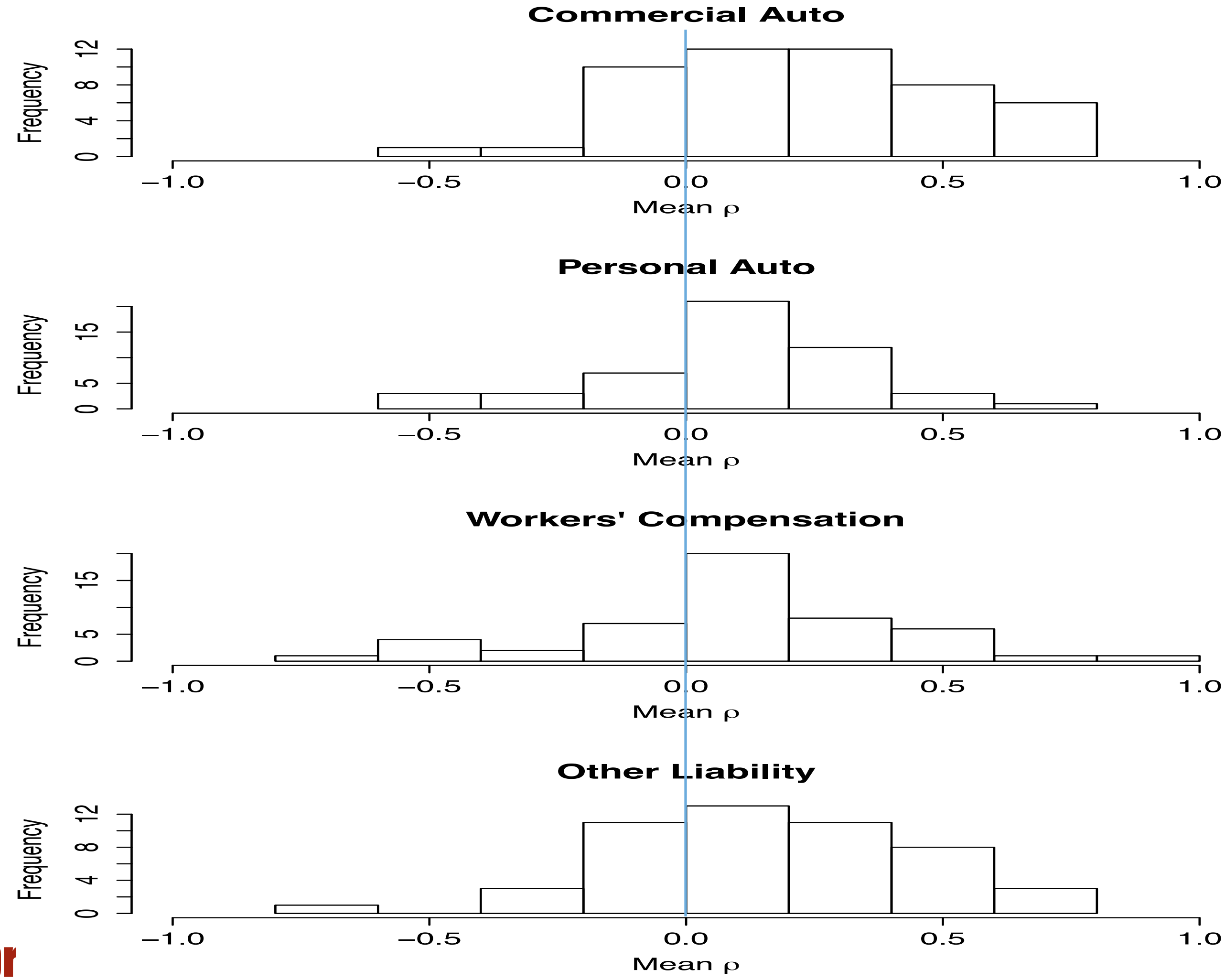
Should we allow ρ in the model?



Tendency for negative trends

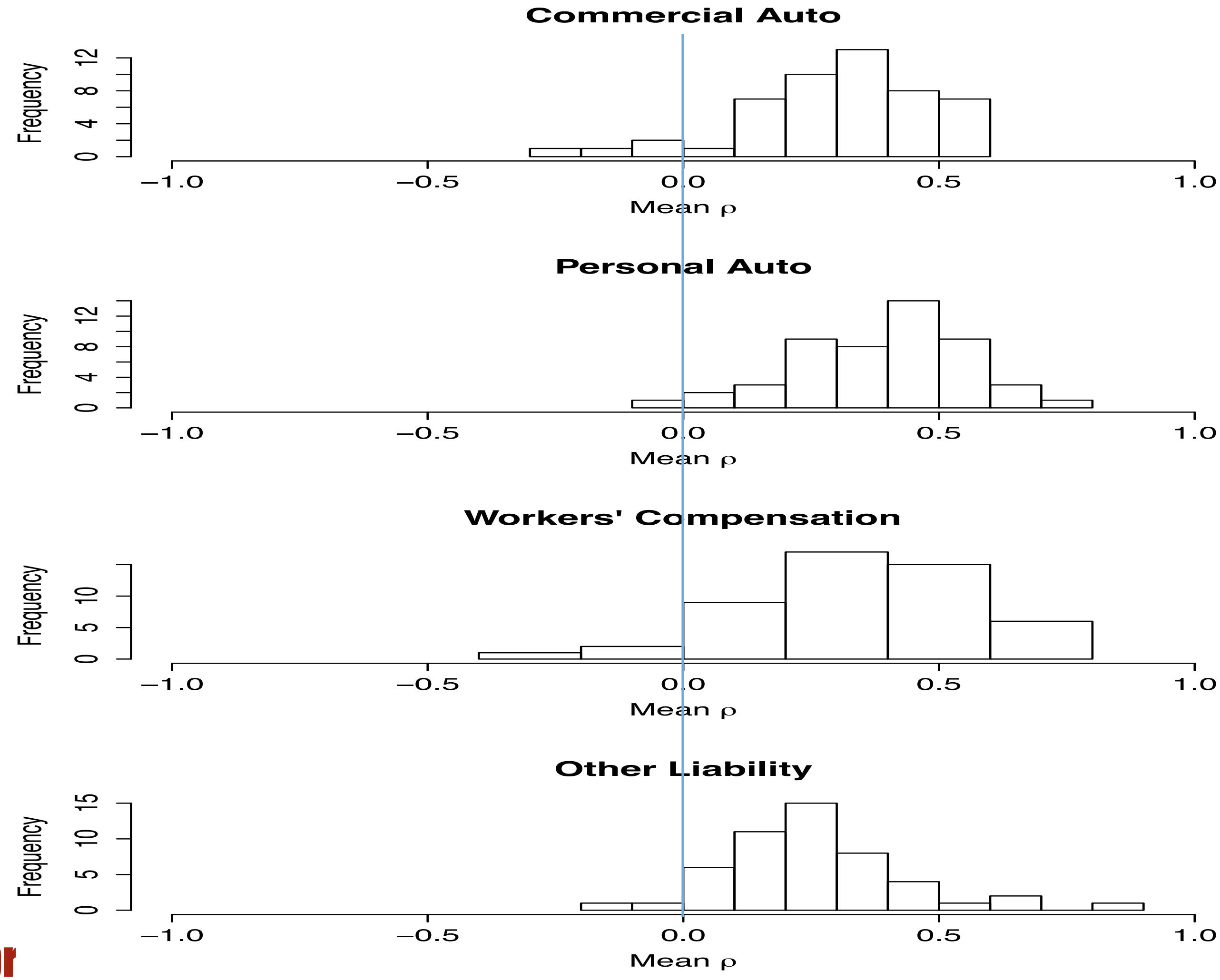


Posterior Mean ρ for All Insurers For CIT On Paid Data



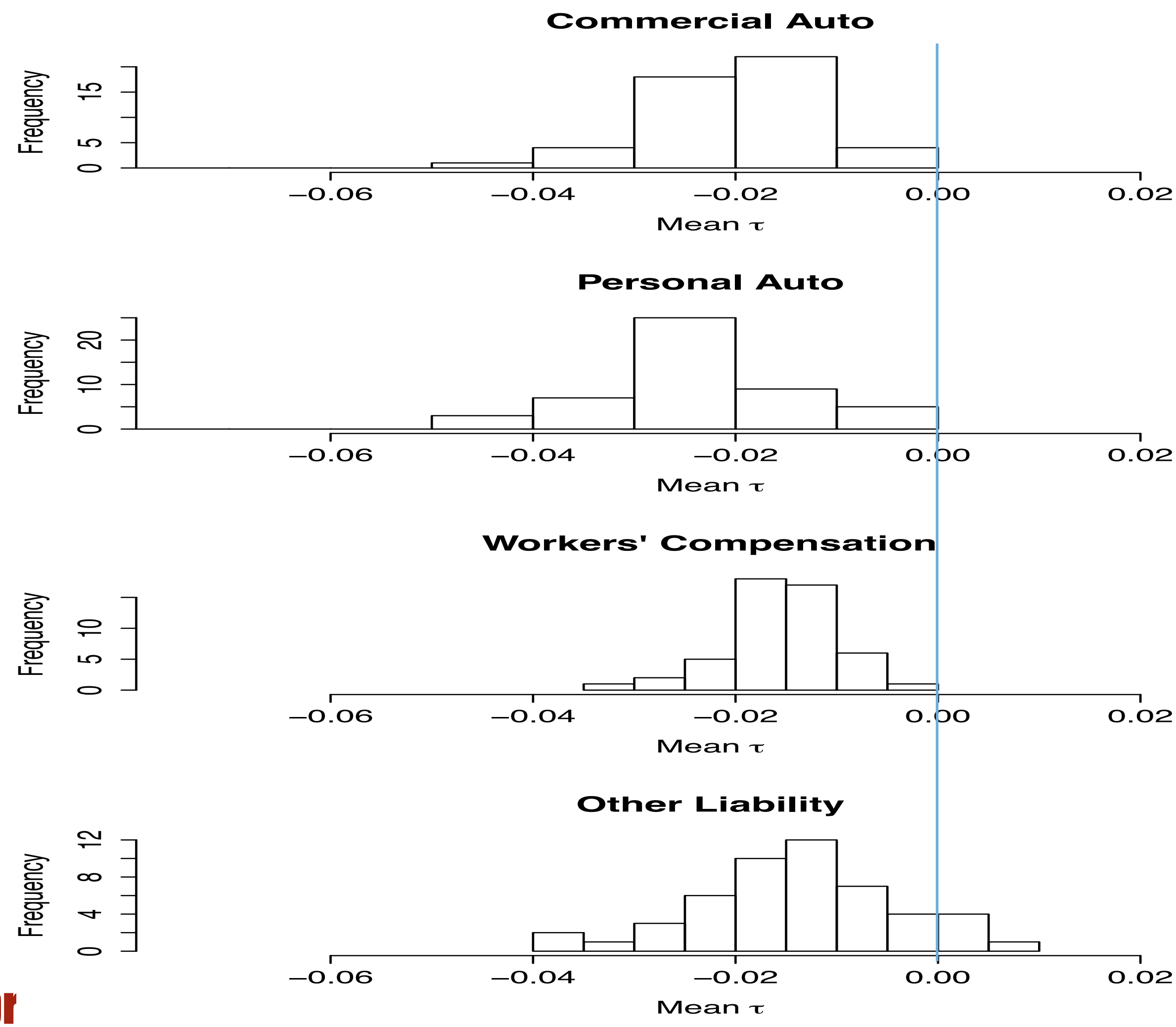
2015 Spr

Posterior Mean ρ for All Insurers For CCL On Incurred Data



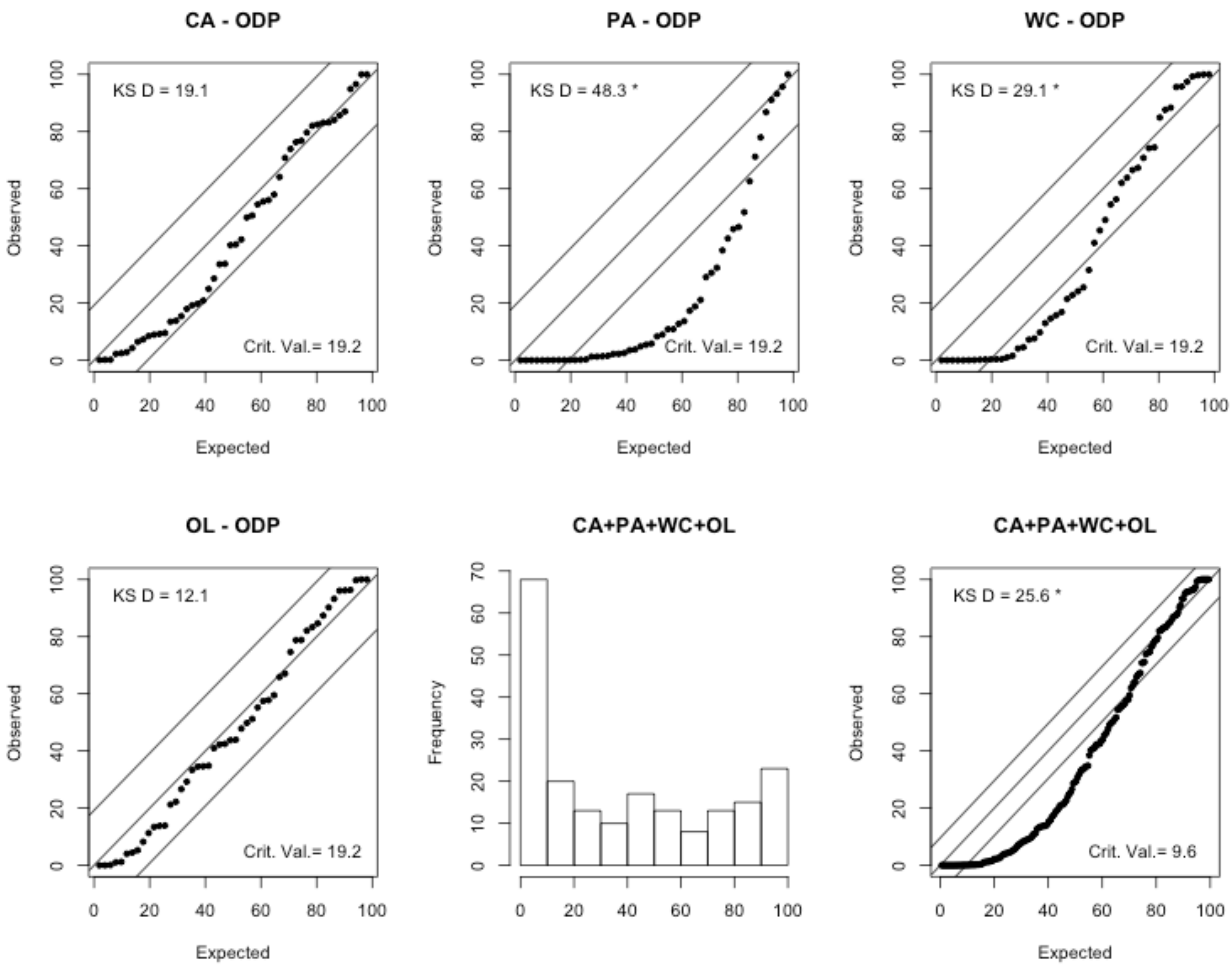
2015 Spr

Posterior Mean τ for All Insurers



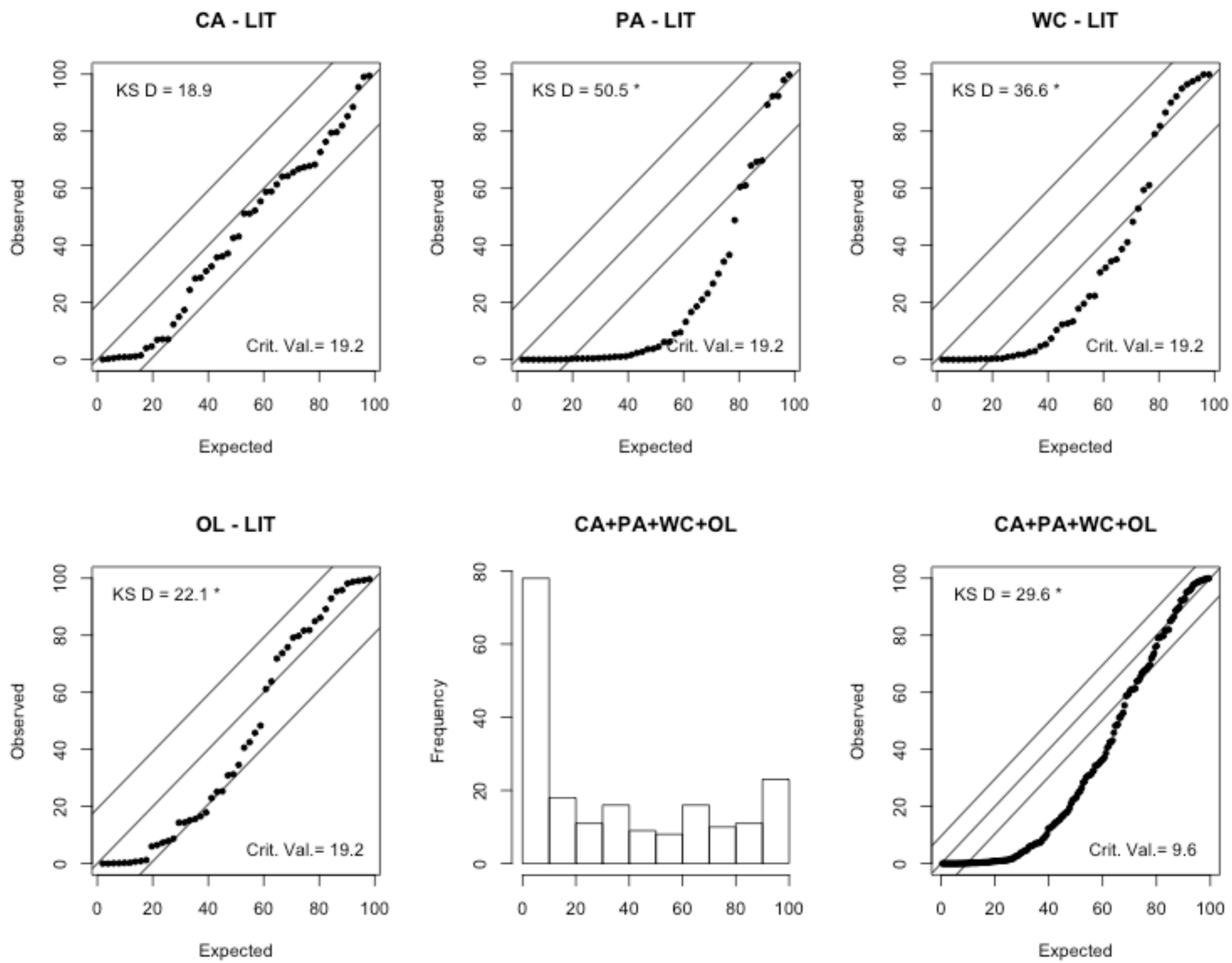
2015 Spr

Test of Bootstrap ODP on Paid Data



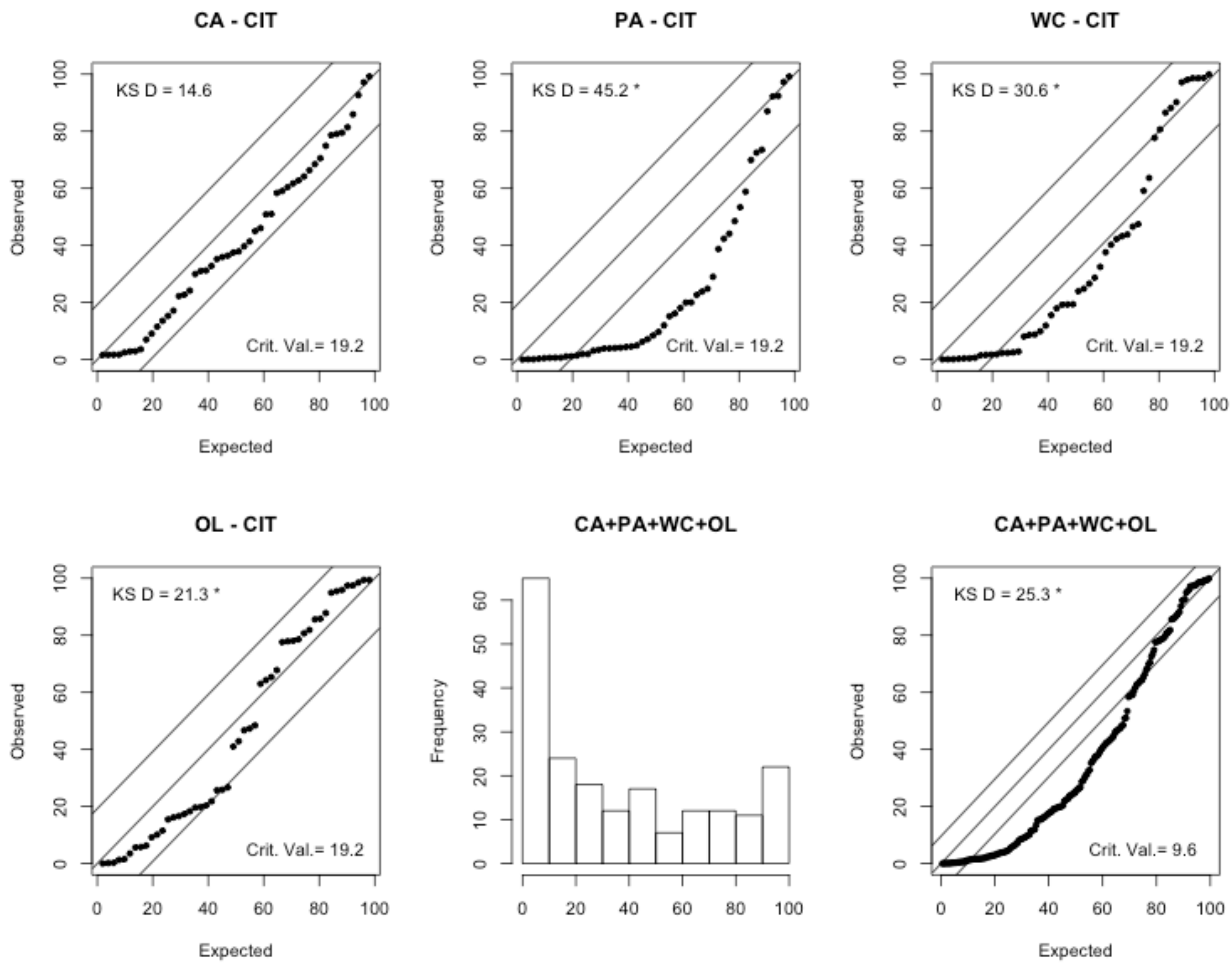
Conclusion – Biased upward.

Test of CIT with ρ = 0 on Paid Data



Not as good as Bootstrap ODP

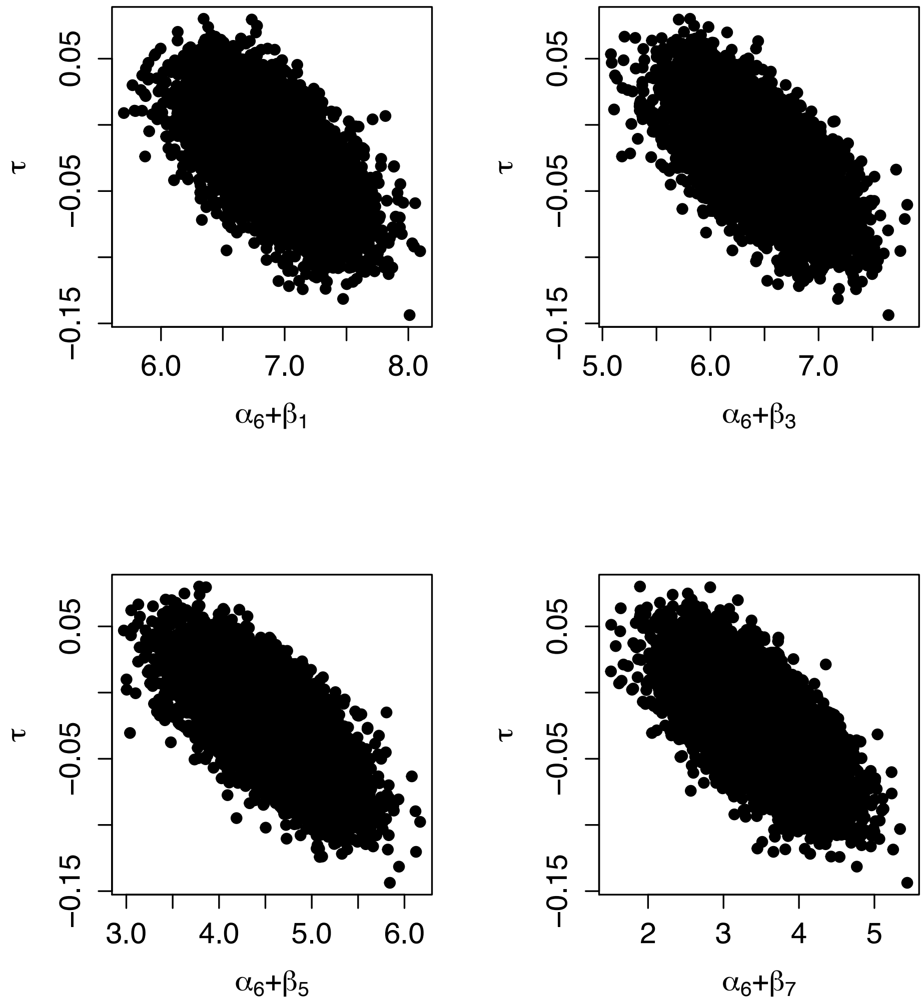
Test of CIT on Paid Data



Comparable to Bootstrap ODP - Still Biased

Why Don't Negative τs Fix the Bias Problem?

Low τ offset by Higher $\alpha + \beta$



Original Monograph Submission

- Stopped here – Concluded that we should use incurred losses.
- An anonymous referee commented that claim settlement was speeding up in recent years.
- Challenged the idea that the CIT model would correct for the claim settlement speedup.

The Changing Settlement Rate (CSR) Model

- *logelr* ~ uniform(-5,0)
- α_w ~ normal(log(Premium$_w$)+*logelr*,$\sqrt{10}$)
- $\beta_{10} = 0$, β_d ~ uniform(-5,5), for $d = 1,\ldots,9$
- a_i ~ uniform(0,1)

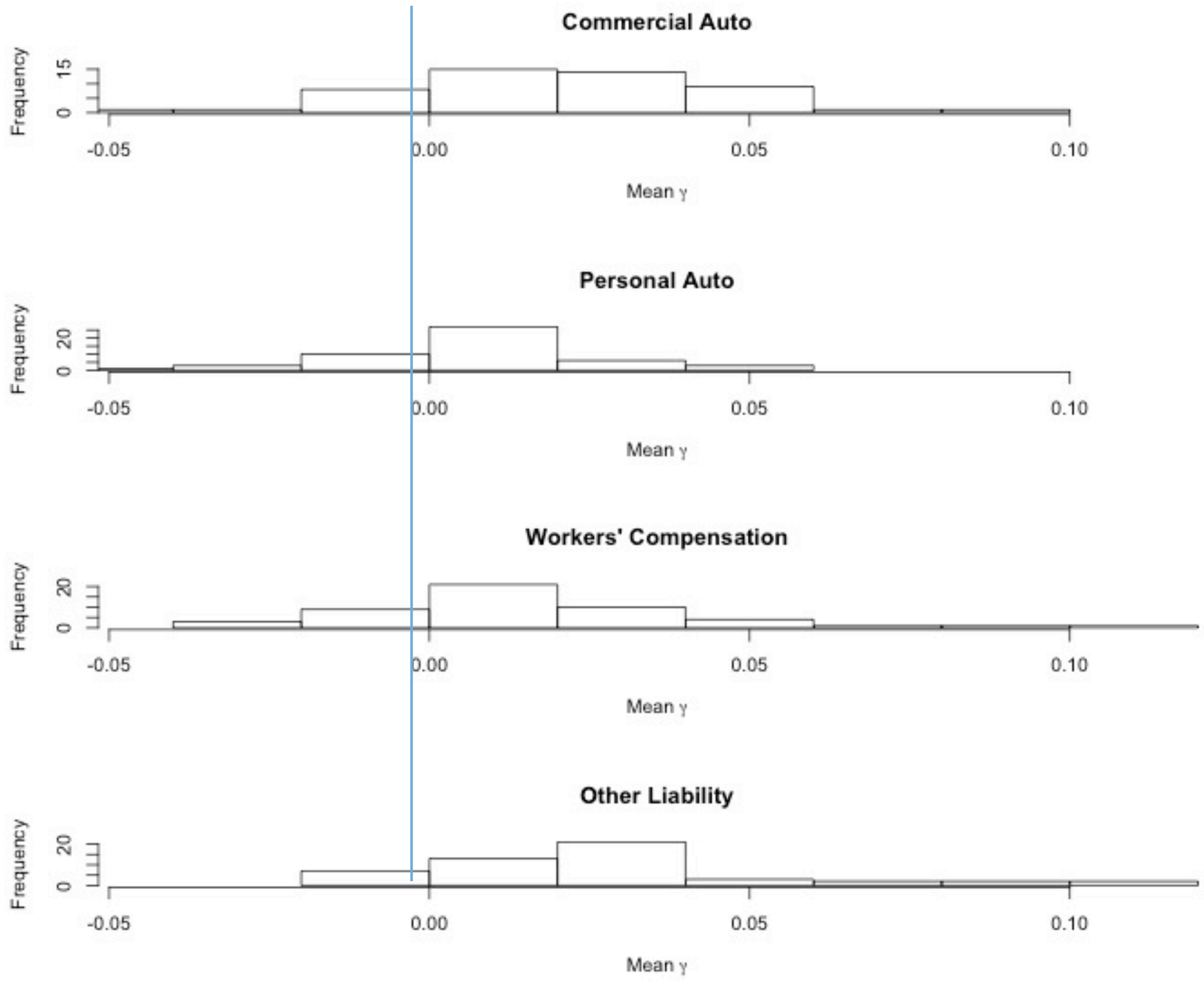
- $\sigma_d = \sum_{i=d}^{10} a_i$ Forces σ_d to decrease as d increases

- $\mu_{w,d} = \alpha_w + \beta_d \cdot (1-\gamma)^{(w-1)}$ γ ~ Normal(0,0.05)
- $C_{w,d}$ ~ lognormal($\mu_{w,d}$, σ_d)

The Effect of γ

- $\mu_{w,d} = \alpha_w + \beta_d \cdot (1 - \gamma)^{(w - 1)}$
- βs are almost always negative! ($\beta_{10} = 0$)
- Positive γ – Speeds up settlement
- Negative γ – Slows down settlement
- Model assumes speed up/slow down occurs at a constant rate.

Distribution of Mean γs

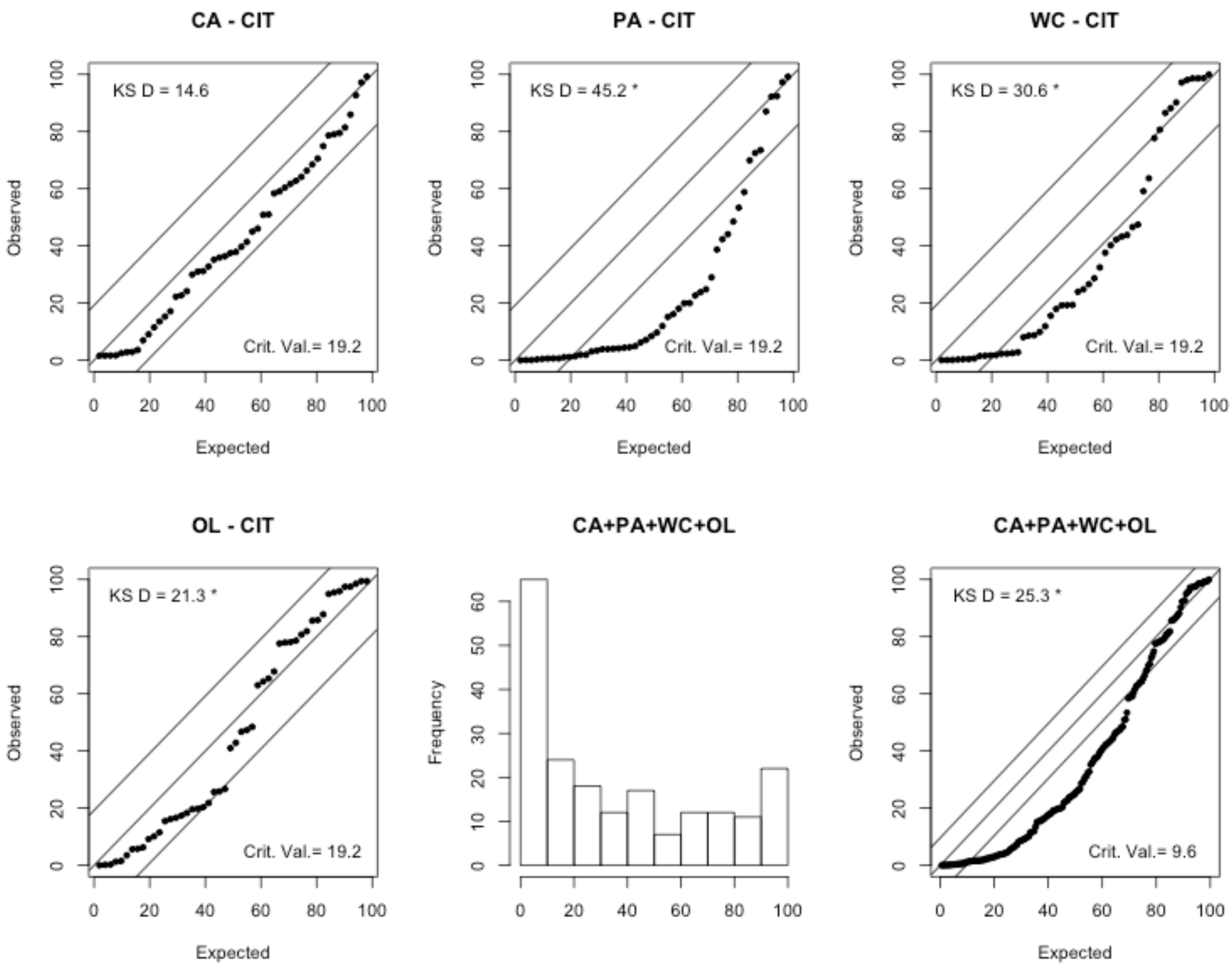


Predominantly Positive γs

CSR Model for Illustrative Insurer

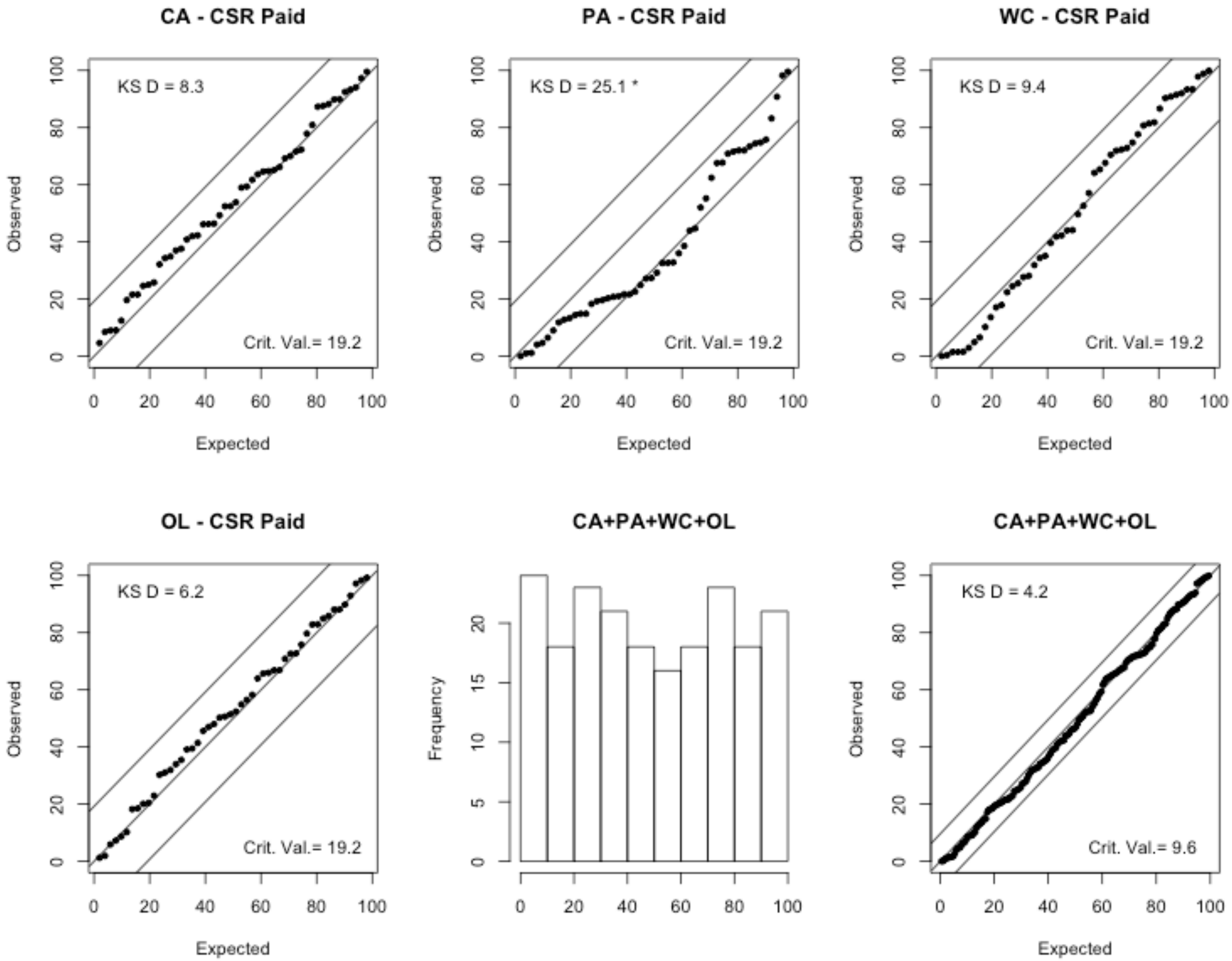
	CIT			CSR			Outcome
w	$C_{w,10}$	SD	CV	$C_{w,10}$	SD	CV	$C_{w,10}$
1	3912	0	0	3912	0	0	3912
2	2539	9	0.0035	2559	103	0.0403	2527
3	4183	21	0.0050	4135	173	0.0418	4274
4	4395	40	0.0091	4285	198	0.0462	4341
5	3553	42	0.0118	3513	180	0.0512	3583
6	3063	101	0.0330	3317	216	0.0651	3268
7	5062	123	0.0243	4967	404	0.0813	5684
8	3512	514	0.1464	3314	402	0.1213	4128
9	4025	707	0.1757	3750	734	0.1957	4144
10	4698	1482	0.3155	3753	1363	0.3632	4139
Total	38942	1803	0.0463	37506	2247	0.0599	40000
Percentile		79.04			87.62		

Test of CIT on Paid Data



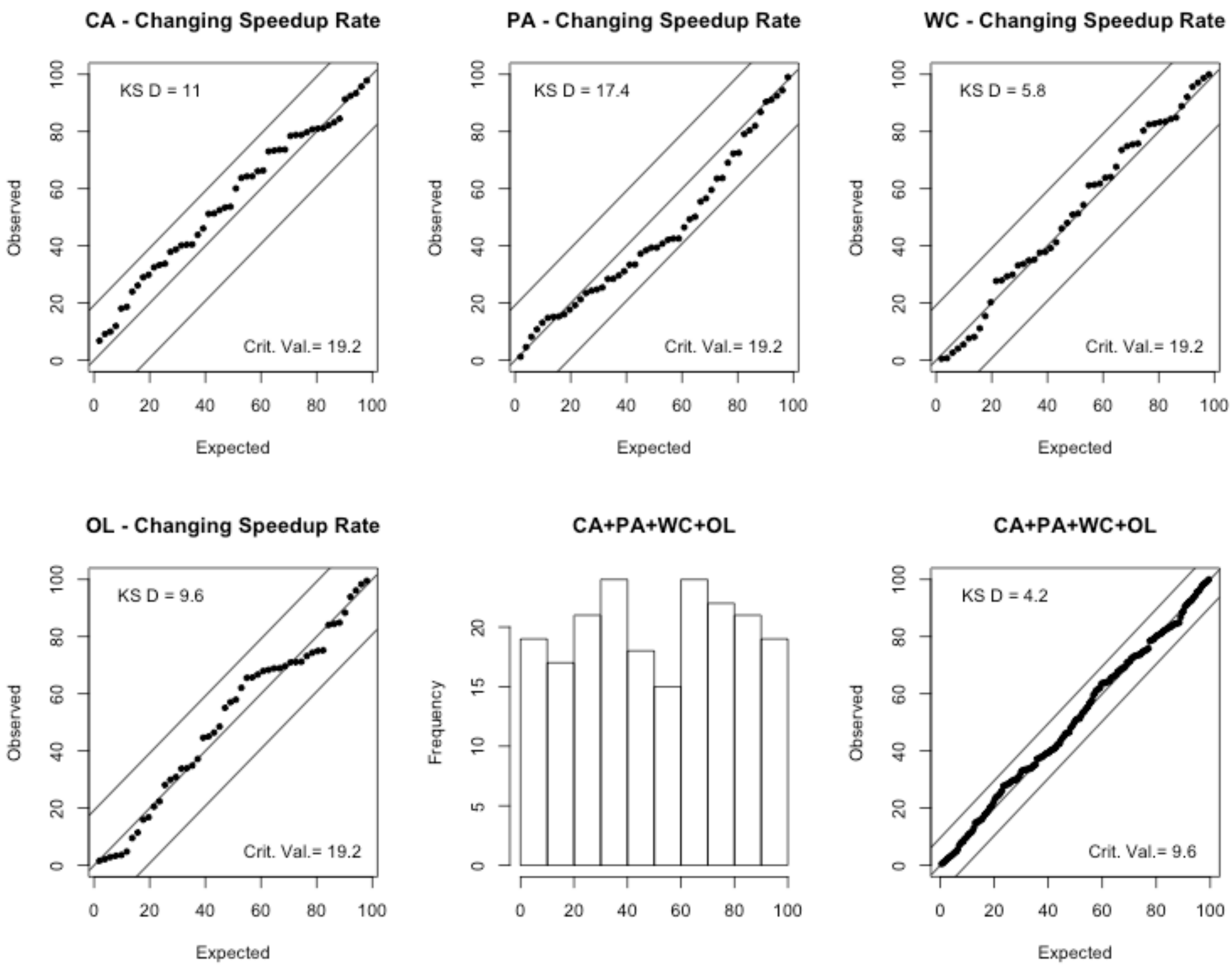
Comparable to Bootstrap ODP - Still Biased

Test of CSR on Paid Data



Conclusion - Much better than CIT

New - Allow a Changing Speedup Rate



Got PA within Critical Values

A Recent Development

"Predicting Multivariate Insurance Loss Payments Under a Bayesian Copula Framework"

by Yanwei (Wayne) Zhang – FCAS

and Vanja Dukic

Awarded the 2014 ARIA Prize by CAS

The General Idea in Zhang/Dukic

Given X_1 ~ MCMC Model 1

Given X_2 ~ MCMC Model 2

Fit the joint (X_1, X_2) ~ MCMC Models 1 and 2

A Current Project I am Working On

Given X_1 ~ MCMC Model 1

Given X_2 ~ MCMC Model 2

Fit the joint (X_1,X_2) ~ MCMC Models 1 and 2

- X_1 = CA Paid Loss, Model 1 = CSR Model
- X_2 = PA Paid Loss, Model 2 = CSR Model

Joint Lognormal Distribution

$$\begin{pmatrix} \log\left(C_{wd}^{CA}\right) \\ \log\left(C_{wd}^{PA}\right) \end{pmatrix} \sim \text{Multivariate Normal}\left(\begin{pmatrix} \mu_C \\ \mu_P \end{pmatrix}, \begin{pmatrix} \sigma_C^2 & \rho\,\sigma_C\sigma_P \\ \rho\,\sigma_C\sigma_P & \sigma_P^2 \end{pmatrix}\right)$$

- Preliminary results show that ρ is frequently negative!
- Stay tuned.

Short Term Conclusions Incurred Loss Models

- Mack model prediction of variability is too low on our test data.
- CCL model correctly predicts variability at the 95% significance level.
- The feature of the CCL model that pushed it over the top was between accident year correlations.

Short Term Conclusions
Paid Loss Models

- Mack and Bootstrap ODP models are biased upward on our test data.
- Attempts to correct for this bias with Bayesian MCMC models that include a calendar year trend failed.
- Models that allow for changes in claim settlement rates work much better.
- *Claims adjusters have important information!*

Long Term Recommendations
New Models Come and Go

- Transparency - Data and software released
- Large scale retrospective testing on real data
  - While individual loss reserving situations are unique, knowing how a model performs retrospectively on a large sample of triangles should influence one's choice of models.
- Bayesian MCMC models hold great promise to advance Actuarial Science.
  - Illustrated by the above stochastic loss reserve models.
  - Allows for judgmental selection of priors.