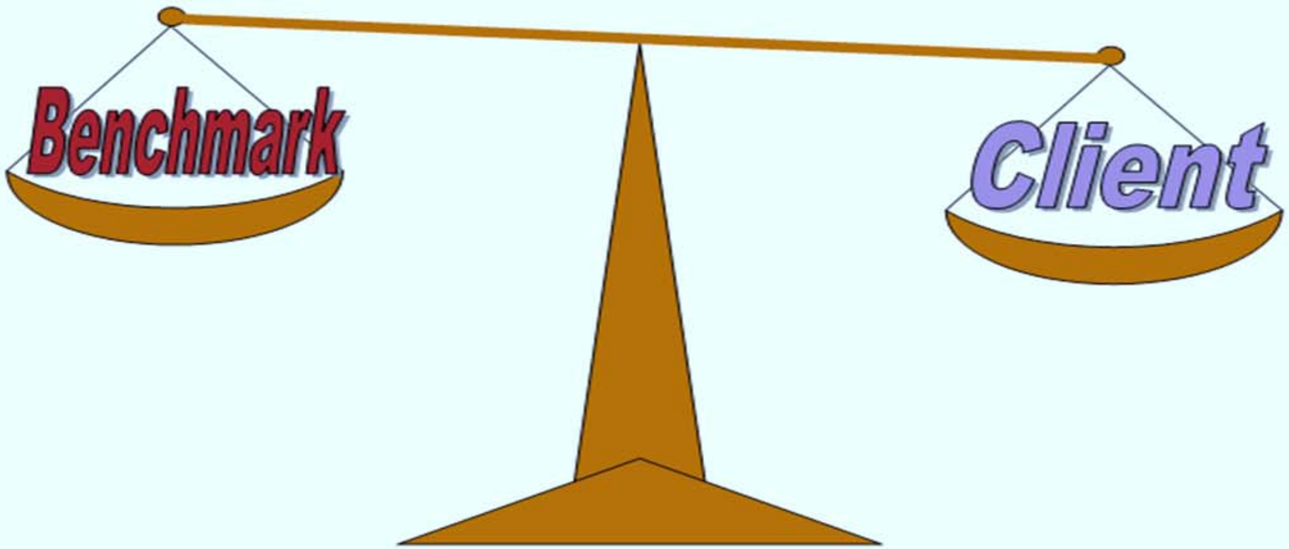
Antitrust Notice

Munich RE

The Casualty Actuarial Society is committed to adhering strictly to the letter and spirit of the antitrust laws. Seminars conducted under the auspices of the CAS are designed solely to provide a forum for the expression of various points of view on topics described in the programs or agendas for such meetings.

Under no circumstances shall CAS seminars be used as a means for competing companies or firms to reach any understanding – expressed or implied – that restricts competition or in any way impairs the ability of members to exercise independent business judgment regarding matters affecting competition.

It is the responsibility of all seminar participants to be aware of antitrust regulations, to prevent any written or verbal discussions that appear to violate these laws, and to adhere in every respect to the CAS antitrust compliance policy.



BAYESIAN LOSS DEVELOPMENT: CREATING AN INFORMATIVE PRIOR

CAS Casualty Loss Reserve Seminar – September 10-12, 2017

Dave Clark
Munich Reinsurance America, Inc

Munich RE

Agenda

1. Loss Development in Pricing and Reserving
2. The Bayesian Solution
3. Cross-Validation Concept
4. Example and Practical Implementation
5. Extending the Model

Loss Development Blending

Reinsurance pricing problem:

We have a loss development triangle from our client:

- May be sparse, not fully credible
- No tail beyond latest age in triangle

We have “benchmark” pattern from other sources:

- ISO / RAA / Reserving / Peer Companies
- Uncertain estimation and relevance for this client

Loss Development Blending

(numbers for illustration only)

Single Benchmark Example

	12	24	36	48	60	72	84	96
1990	73	262	469	528	536	591	604	606
1991	148	346	391	502	522	514	567	
1992	99	198	219	394	408	430		
1993	118	255	352	412	581			
1994	275	415	645	803				
1995	261	446	637					
1996	130	471						
1997	148							
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
1990	3.589	1.790	1.126	1.015	1.103	1.022	1.003	
1991	2.338	1.130	1.284	1.040	0.985	1.103		
1992	2.000	1.106	1.799	1.036	1.054			
1993	2.161	1.380	1.170	1.410				
1994	1.509	1.554	1.245					
1995	1.709	1.428						
1996	3.623							
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	
Avg ATA	2.168	1.412	1.271	1.115	1.047	1.060	1.003	

Loss Development Blending

(numbers for illustration only)

Munich RE

	Single Benchmark Example							
	12	24	36	48	60	72	84	96
1990	73	262	469	528	536	591	604	606
1991	148	346	391	502	522	514	567	
1992	99	198	219	394	408	430		
1993	118	255	352	412	581			
1994	275	415	645	803				
1995	261	446	637					
1996	130	471						
1997	148							
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
1990	3.589	1.790	1.126	1.015	1.103	1.022	1.003	
1991	2.338	1.130	1.284	1.040	0.985	1.103		
1992	2.000	1.106	1.799	1.036	1.054			
1993	2.161	1.380	1.170	1.410				
1994	1.509	1.554	1.245					
1995	1.709	1.428						
1996	3.623							
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	
Avg ATA	2.168	1.412	1.271	1.115	1.047	1.060	1.003	

Bayesian Philosophy

Munich RE

Bayes' Theorem:

$$\pi(\theta|X) = \frac{f(X|\theta) \cdot \pi(\theta)}{\int f(X|\theta) \cdot \pi(\theta)\, d\theta}$$

This formula has three components:

$\pi(\theta)$ A distribution representing "prior" knowledge of the parameters θ

$f(X|\theta)$ A likelihood function representing the probability of observing the actual data X given a certain parameter set.

$\pi(\theta|X)$ The "posterior" probability of the parameters, revised based on the data

Bayesian Philosophy

Tools for Evaluating the Mathematics:

1) Conjugate Families
2) Linear Approximation to Bayes Formula => Bühlmann-Straub
3) Numerical Approximation of the Formula
    a) Quadrature integration (old method)
    b) Simulation via MCMC (the new favorite)

Conjugate family has advantage of simple calculation and interpretability.

Conjugate Prior

When the prior distribution $\pi(\theta)$ and likelihood $f(X|\theta)$ are chosen such that the posterior distribution $\pi(\theta|X)$ has the same distribution form as the prior, then we have a *conjugate* relationship.

Common examples from the Exponential Family are:

$\pi(\theta)$	=>	$f(X\|\theta)$
Gamma	=>	Poisson
Beta	=>	Binomial
Dirichlet	=>	Multinomial
Normal	=>	Normal

Conjugate Priors - Interpretation

“Conjugate priors… have the desirable feature that prior information can be viewed as ‘fictitious sample information’ in that it is combined with the sample in exactly the same way that additional sample information would be combined.

“The only difference is that the prior information is ‘observed’ in the mind of the researcher, not in the real world.”

- Bayesian Econometric Methods; Koop, Poirier & Tobias

Conjugate Priors – Loss Development

For analysis of loss development patterns:

- Normal / Normal [Shi & Hartman (2014)]
- Dirichlet / Multinomial [Clark (2016), following Mildenhall (2006)]

Both of these conjugate models result in the same form that is easily implemented in practice.

Credibility Blending Formula

Munich RE

The credibility blending becomes a simple dollar-weighted average.

If you can calculate an age-to-age factor, then you can do a Bayesian model!

		Example of Blending Client and Benchmark Patterns						
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
ATA from Triangle								
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	-
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	-
ATA	**2.168**	**1.412**	**1.271**	**1.115**	**1.047**	**1.060**	**1.003**	
Benchmark Pattern								
Col. 1	1,419	2,027	2,546	2,933	3,383	3,633	3,717	3,042
Col. 2	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000
ATA	**2.819**	**1.973**	**1.571**	**1.364**	**1.182**	**1.101**	**1.076**	**1.315**
Blended Pattern								
Col. 1	2,523	3,949	4,622	4,769	4,849	4,738	4,321	3,042
Col. 2	6,393	6,713	6,639	6,047	5,535	5,171	4,606	4,000
ATA	**2.534**	**1.700**	**1.436**	**1.268**	**1.141**	**1.091**	**1.066**	**1.315**

All numbers for illustration only

Empirical Estimation of Credibility

- Subjective Bayes
  - Prior distributions selected based on expert judgment
  - Practical approach: reverse engineer based on implied credibility weights used by actuaries

- Empirical Bayes
  - Use other data to create a “prior”
  - Known as “regularization” in Predictive Analytics
  - We can use concept of Cross Validation

Note: We can compromise by estimating empirical credibility factors and allowing our experts to subjectively adjust them.

Empirical Estimation of Credibility

Munich RE

Aggregate Everything into One Triangle

============

- Potential bias due to missing changes in mix
- Bias when splitting into business units

Credibility Blending

A compromise solution is to find the best blending between these two extremes.

- Shrinkage or Regularization* to statisticians
- Credibility to actuaries

Create Many Small Homogeneous Triangles

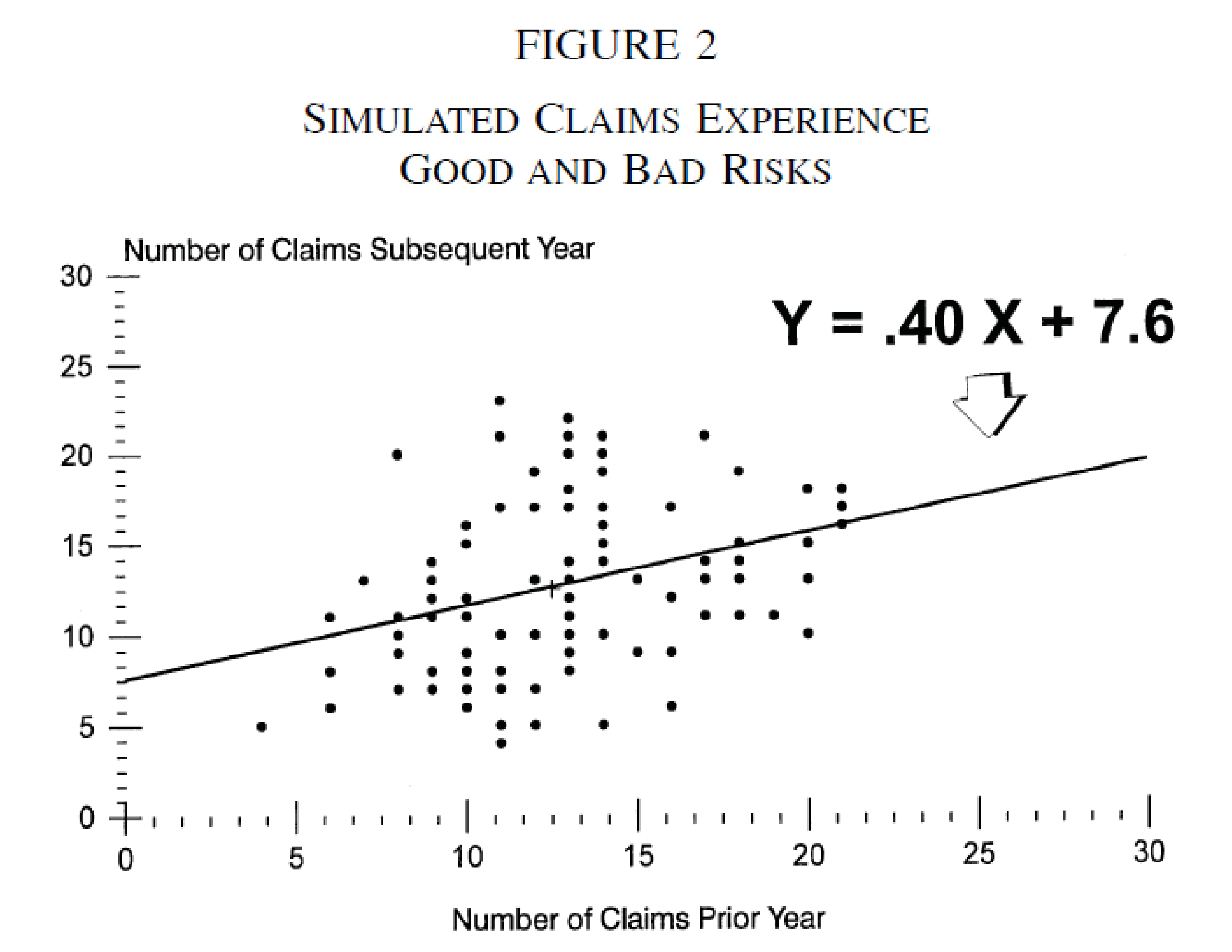
============

- High Variance
- Data is sparse
- "fit to noise"

*Andrew Gelman informally defines regularization as "a general term used for statistical procedures that give more stable estimates."

Empirical Estimation of Credibility

Munich RE



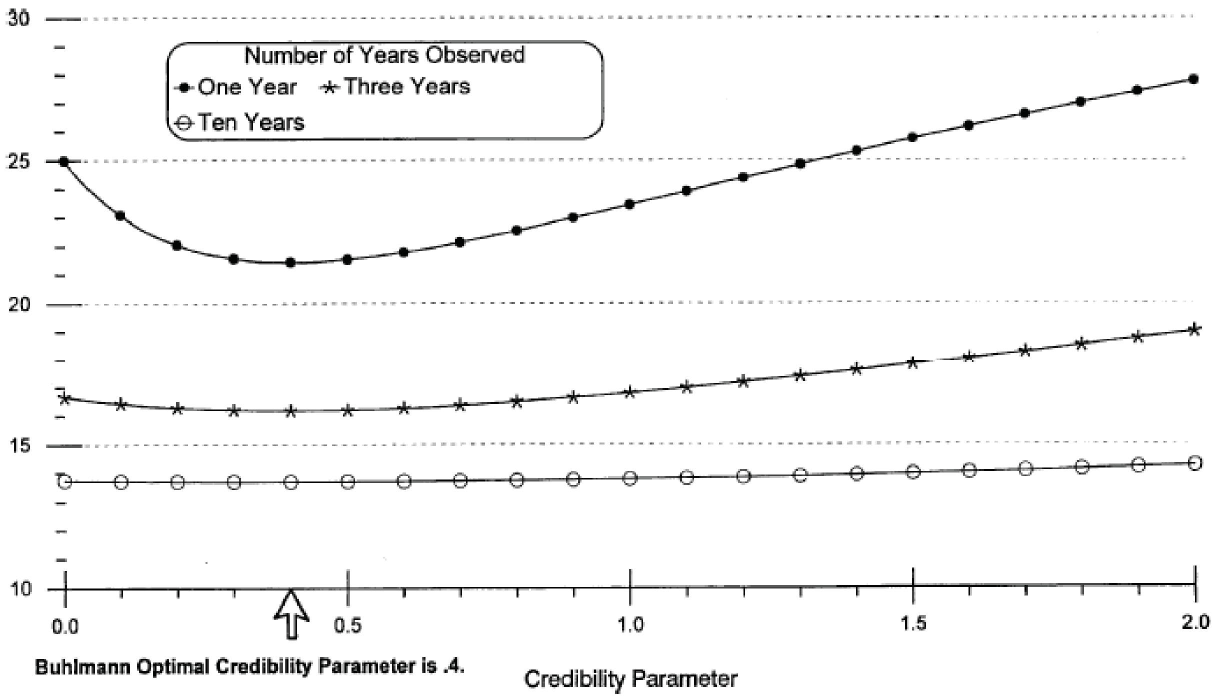
From Howard Mahler's 1998 paper, we are asking how predictive historical experience is for future claims.

Source: Mahler (1998)

Empirical Estimation of Credibility

FIGURE 15

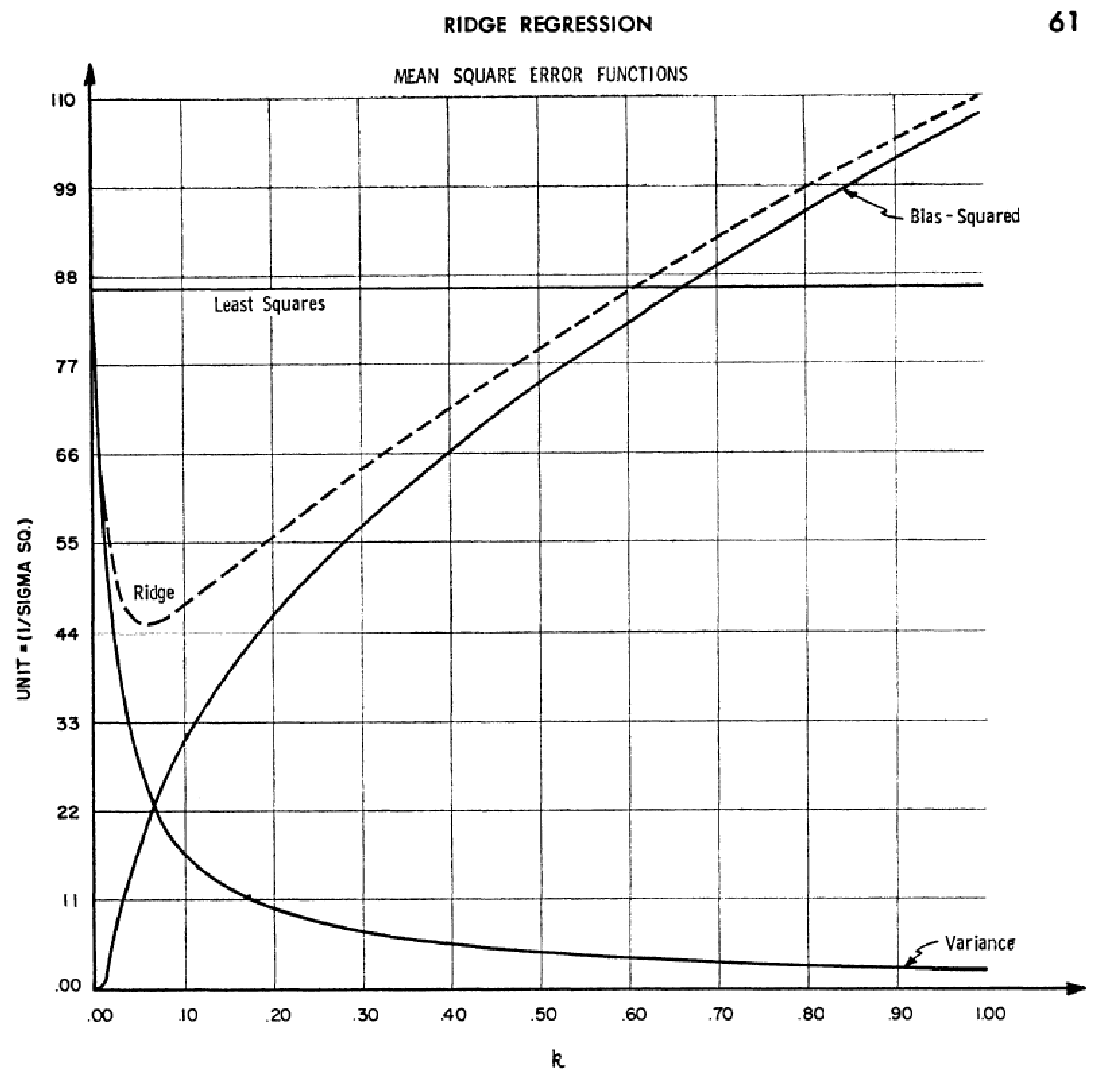
EXPECTED SQUARED PREDICTION ERRORS VS. CREDIBILITY PARAMETER USED TO DETERMINE WEIGHT GIVEN TO OBSERVED FREQUENCY
EXCELLENT, GOOD, BAD, AND UGLY RISKS



The optimal credibility parameter minimizes the prediction error.

Source: Mahler (1998)

Empirical Estimation of Credibility



What Mahler (1998) called minimizing the prediction error is equivalent to what data scientists call optimizing the variance/bias trade-off.

Cross Validation is a popular tool for accomplishing this task.

Source: Hoerl & Kennard (1970)

Empirical Estimation of Credibility

Munich RE

Leave One Out Cross Validation (LOOCV):

Find the *Credibility Ballast* value that minimizes prediction error. For loss reserving, this may be a future age-to-age factor compared to the average age-to-age factor on the incomplete triangle averaged with a benchmark.

$$Prediction\ Error = \sum_{i=1}^{n} w_i \cdot \left(y_i - \hat{y}_{-i}(Cred\ Ballast)\right)^2$$

Example: sample data from the CAS database of Schedule P triangles.

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

Empirical Estimation of Credibility

Munich RE

Remember that it is this credibility ballast that we want to estimate.

		Example of Blending Client and Benchmark Patterns					
	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-Ult
ATA from Triangle								
Col. 1	1,104	1,922	2,076	1,836	1,466	1,105	604	-
Col. 2	2,393	2,713	2,639	2,047	1,535	1,171	606	-
ATA	**2.168**	**1.412**	**1.271**	**1.115**	**1.047**	**1.060**	**1.003**	
Benchmark Pattern								
Col. 1	1,419	2,027	2,546	2,933	3,383	3,633	3,717	3,042
Col. 2	4,000	4,000	4,000	4,000	4,000	4,000	4,000	4,000
ATA	**2.819**	**1.973**	**1.571**	**1.364**	**1.182**	**1.101**	**1.076**	**1.315**
Blended Pattern								
Col. 1	2,523	3,949	4,622	4,769	4,849	4,738	4,321	3,042
Col. 2	6,393	6,713	6,639	6,047	5,535	5,171	4,606	4,000
ATA	**2.534**	**1.700**	**1.436**	**1.268**	**1.141**	**1.091**	**1.066**	**1.315**

All numbers for illustration only

Empirical Estimation of Credibility

Munich RE

Using the Schedule P data by line of business, we use LOOCV to estimate the credibility constants for paid loss development.

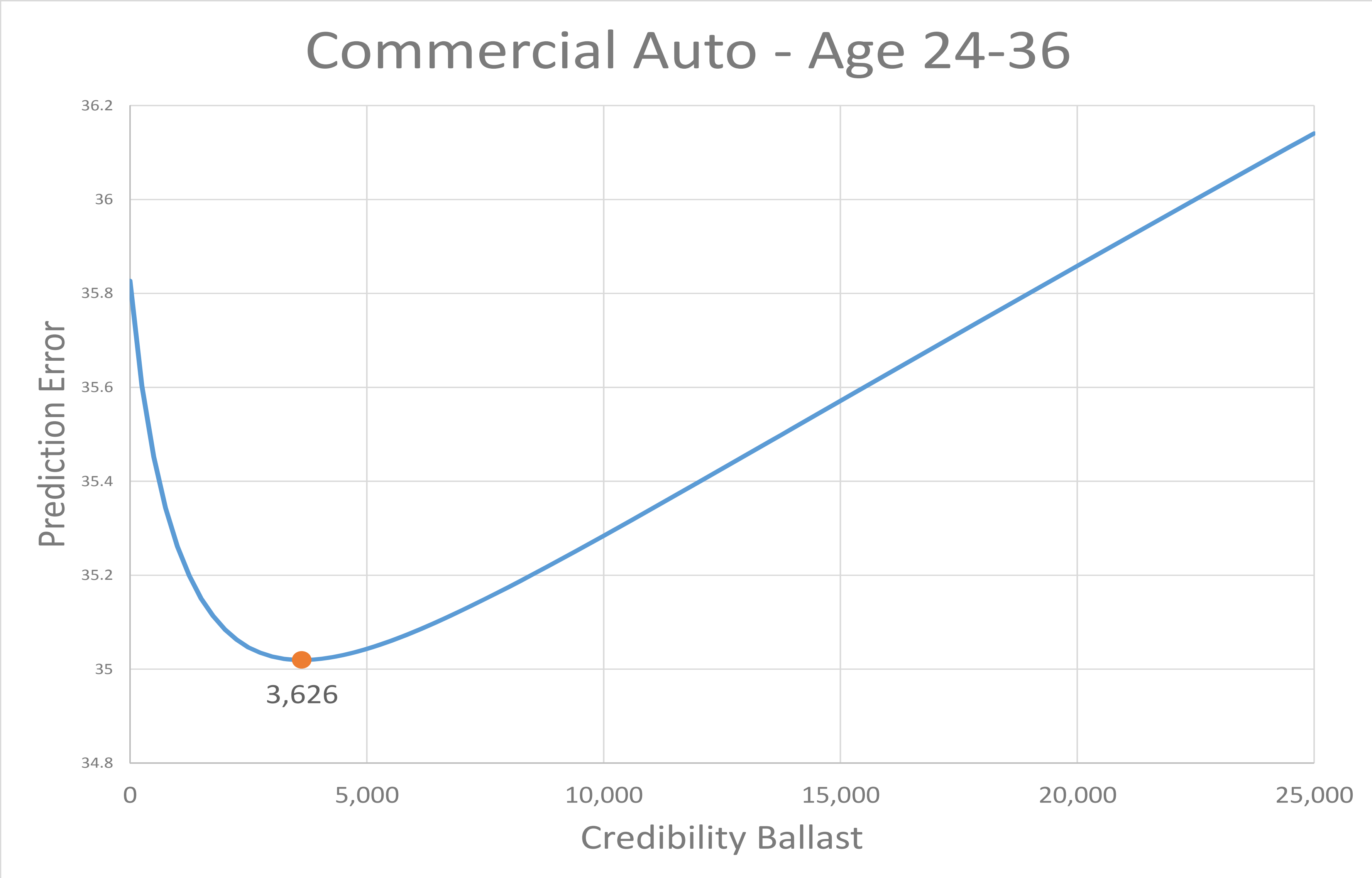
[these are estimates to replace the constant 4,000 in earlier slides]

	12-24	24-36	36-48	48-60	60-72	72-84	84-96	96-108	108-120
	Credibility Ballast from 1988-1997 Schedule P Database								
CAL	1,068	3,626	6,845	24,326	16,399	40,550	60,229	283,454	83,671
PPAL	3,933	9,324	6,069	6,924	12,366	28,015	31,809	66,268	80,551
GL	2,544	3,500	4,000	8,403	9,732	7,463	20,456	12,582	636,980
WC	9,861	7,875	9,651	12,431	20,674	24,052	37,642	999,999	15,103

[Excel example is provided to show how these numbers are calculated]

All numbers for illustration only

Empirical Estimation of Credibility



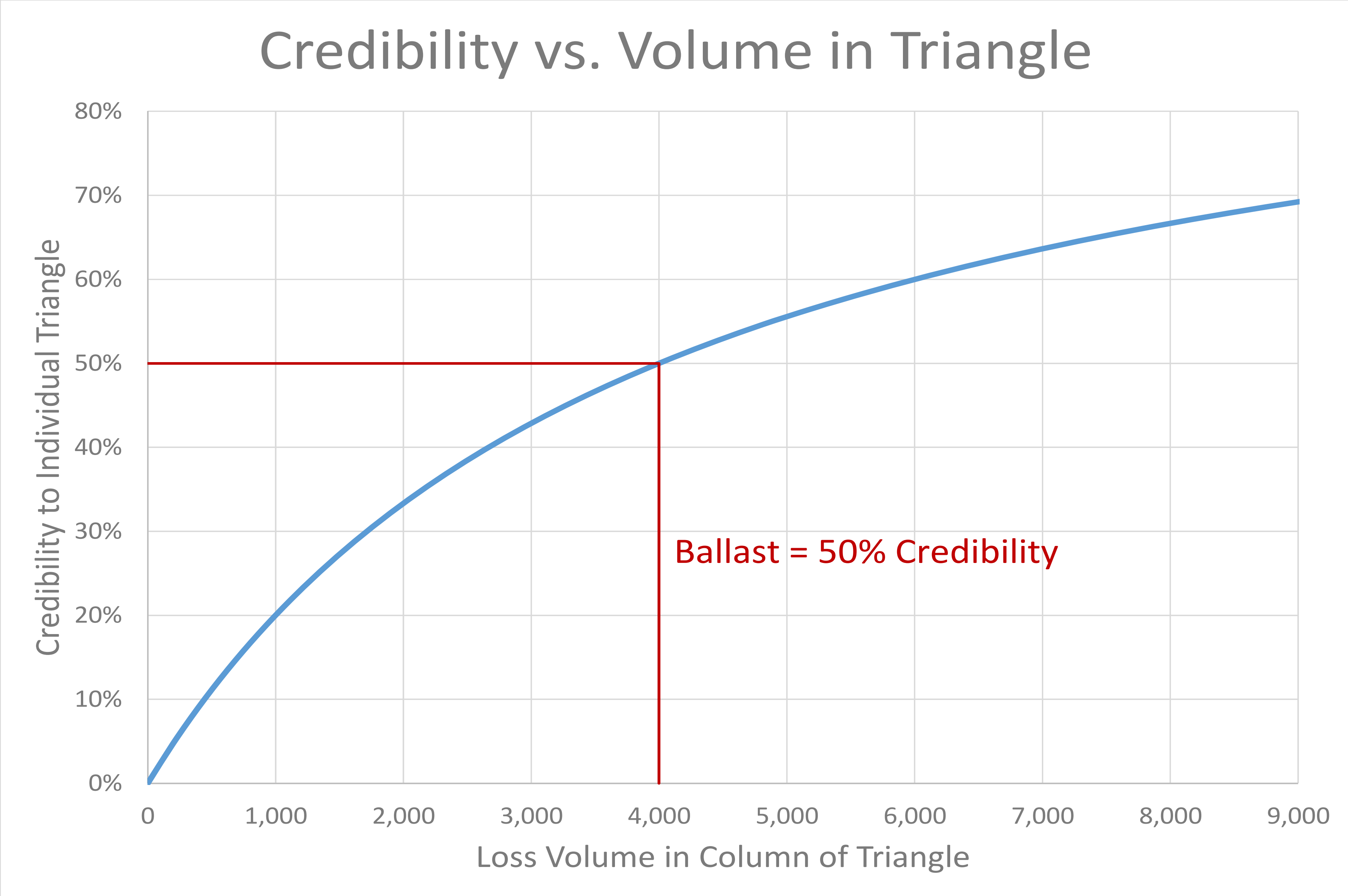
Empirical Estimation of Credibility

Munich RE

Comments on Cross-Validation:

- The credibility parameter is estimated from the data and is therefore subject to estimation error. A larger data set for this estimation is always better, but the final number may still need to be “smoothed” judgmentally.
- The estimation depends upon the selection of the error structure to be minimized (I am using Chi-Square error term).
- For the CAS Schedule P data: the credibility “ballast” is calculated separately for each development age.

Subjective Estimation of Credibility



Munich RE

EXTENDING THE MODEL

Extending the Model

Munich RE

The model discussed so far has assumed that each age of development is to be blended individually.

We can extend this by allowing a dependence structure between ages.

- Shi & Hartman (2014) use a Normal/Normal model that allows for a correlation matrix to be included in the multivariate Normal distribution
- Clark (2016) uses a finite mixture distribution
- Alternative approaches allow for fitted curves (e.g., Sherman inverse power) instead of individual age-to-age factors

Extending the Model

Munich RE

We can extend this model further by including mixtures of prior distributions.

Perhaps we know that companies are naturally grouped into Fast, Medium, or Slow payment patterns. But we do not know to which group our client belongs.

		Cumulative Loss Development Factors						
	12	24	36	48	60	72	84	96
Fast	14.014	4.930	2.607	1.759	1.406	1.263	1.191	1.155
Medium	21.950	7.787	3.946	2.512	1.842	1.558	1.415	1.315
Slow	49.240	15.860	7.407	4.163	2.706	2.057	1.750	1.567

All numbers for illustration only

Extending the Model

Munich RE

We assign initial weights to each of the groupings (perhaps 33%/33%/33%) and then apply Bayes' theorem to update the weights.

			Bayesian Updating of Probabilities			
			Difference	Relative	Original	Revised
		LogLikelihood	in LL	Likelihood	Weights	Weights
		A	B=A-Max(A)	C=exp(B)	D	E=C*D/Avg(C)
	Slow	-4.61	-0.77	0.464	33.33%	20.41%
	Baseline	-4.06	-0.21	0.810	33.33%	35.61%
	Fast	-3.84	0.00	1.000	33.33%	43.98%
				0.758	100.00%	100.00%

This allows us to adjust our "tail" based on which group is closest to our client's data.

All numbers for illustration only

Extending the Model

Other potential extensions of this model:

- Alternative variance structures
- More refined benchmark patterns (e.g., a collection of benchmarks dependent upon company type or case reserving practice)
- Inclusion of frequency/severity, exposure bases and expected loss ratios (Mildenhall 2006)
- Use of the model for selection of reserve ranges

Conclusions

- Credibility in Loss Development pattern selection has benefits
  - Stability in estimation – can break data into small homogeneous pieces
  - Consistency in pricing
  - Even very sparse data from a client can update the benchmark

- We can use Cross Validation to estimate starting points for the Bayesian credibility constants

- The Bayesian framework can be extended for ever more realistic assumptions

Selected References

Munich RE

Bolstad, William M., “Introduction to Bayesian Statistics,” Wiley, 2007.

Clark, David R., “Introduction to Bayesian Loss Development,” *CAS Forum* 2016.

http://www.casact.org/pubs/forum/16sforum/Clark.pdf

Hoerl, Arthur E., and Robert W. Kennard, “Ridge Regression: Biased Estimation for Nonorthogonal Problems”, *Technometrics*, Vol. 12, No.1, February 1970.

Mahler, Howard C., “A Graphical Illustration of Experience Rating Credibilities”, *Proceedings of the Casualty Actuarial Society* 1998: LXXXV 654.

http://www.casact.org/pubs/proceed/proceed98/980654.pdf

Selected References

Munich RE

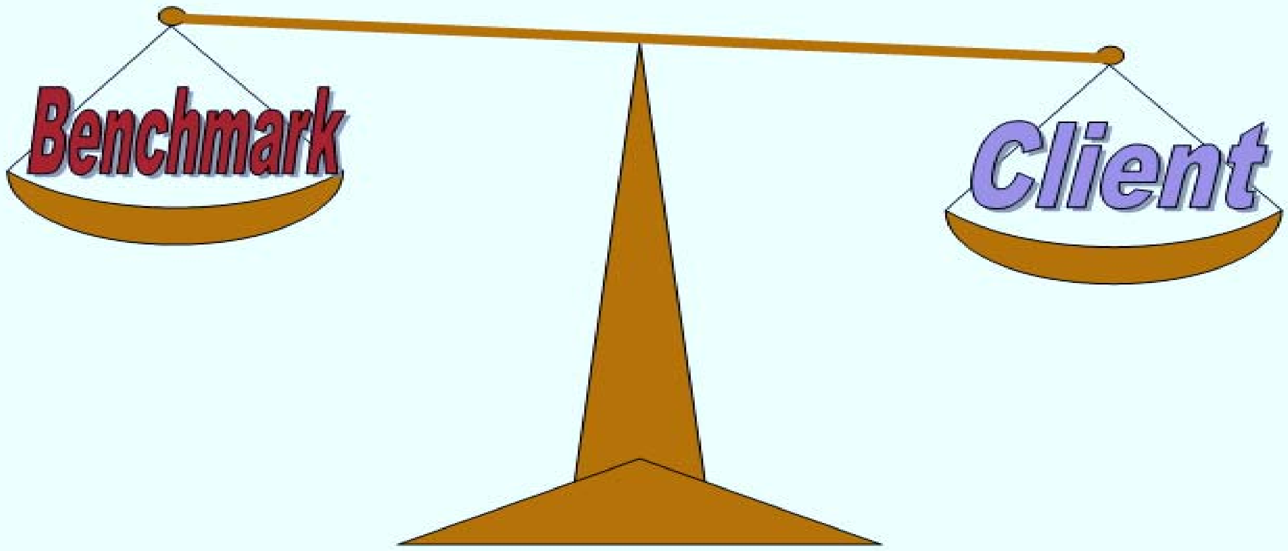
Mildenhall Stephen J., "A Multivariate Bayesian Claim Count Development Model with Closed Form Posterior and Predictive Distributions," *CAS Forum* Winter 2006, 451-493.

http://www.casact.org/pubs/forum/06wforum/06w455.pdf

Shi, Peng and Brian M. Hartman, "Credibility in Loss Reserving", *CAS Forum* 2014: Summer, Vol. 2, 1-21.

http://www.casact.org/pubs/forum/14sumforumv2/Shi_Hartman.pdf

Wong, Tzu-Tsung, "Generalized Dirichlet Distribution in Bayesian Analysis," *Applied Mathematics and Computation* 97 (1998), 165-181.



Dave Clark
daveclark@munichreamerica.com
https://twitter.com/DaveclarkR

Munich RE