

WillisTowersWatson IIIIIIII



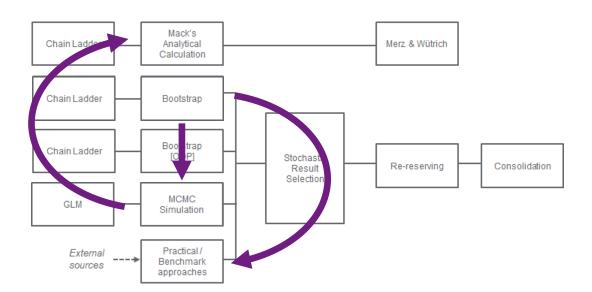
Antitrust notice

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.

Stochastic Models Beyond the Bootstrap...

...into what?



- **Part I:** The first half of the presentation will take a look at an alternative approach to analyzing reserve uncertainty using Bayesian models
- **Part II:** The second half will expand our focus to look at the wider reserving framework and some practical considerations when deploying these models

Stochastic Models Beyond the Bootstrap...

Part I: Introduction to Bayesian modeling

- Reserve Uncertainty in General
- From Mack and Bootstrapping to Bayesian MCMC Models
- Statistics 101
- Bayesian Modeling Steps
- Bayesian Modeling within the Reserving Context
- Simple Example No Simulations Needed
- Examples of Popular Sampling Techniques
- Conclusions

Part II: Deploying a stochastic modeling framework

- The link between assumptions in our deterministic framework and our uncertainty analysis
- Selection of models
- Use of multiple models
- Using benchmarks

Introduction to Bayesian modeling

Introduction to Bayesian modeling:

Reserve Uncertainty in General

Reserve Uncertainty

- The loss reserving process is critical to any company's operations. It requires an accurate estimate of all future claim payments associated with accidents that have already occurred.
- Since estimates are subject to uncertainty, actual future claim payments may develop different from expectations
- History has demonstrated that this difference can be material to a company's balance sheet.
- Quantifying the uncertainty associated with reserve estimates is crucial to a company's management.

Reserve Uncertainty (continued)

- The uncertainty associated with unpaid claims estimates impacts enterprise risk management calculations, capital allocations, and other business decisions.
- In addition, outside parties such as shareholders, regulators, and rating agencies require information on the risk associated with a company's reserves.
- All of these activities require an understanding and quantification of reserve uncertainty.

Introduction to Bayesian modeling:

From Mack and Bootstrapping to Bayesian MCMC Models

Mack's Model

The Mack model measures the standard error of the chain ladder unpaid claim estimate

Based on the following simple regression model:

$$C_{i,k+1} = f_k C_{i,k} + \sigma_k \varepsilon_{i,k} C_{i,k}^{\alpha}$$

- α = 1/2 the all year volume weighted average link ratio is the best linear unbiased estimator of the link ratio
- α = 1 the all year simple average link ratio is the best linear unbiased estimator of the link ratio

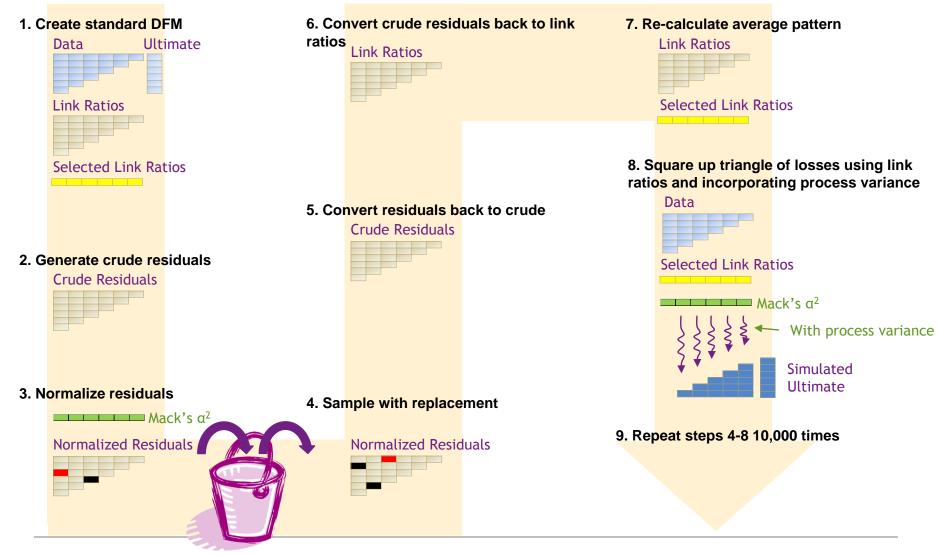
Analytical calculation is based on

- A "closed form" solution formula
- A "recursive" calculation

Bootstrapping in Reserving

- Bootstrapping is a generic process of simulating the estimation error (i.e. parameter variance), which can be applied to many different types of models:
 - Over-dispersed Poisson (ODP)
 - Lognormal
 - Gamma
 - Mack's Model
- Example: Bootstrapping Mack's Model

Bootstrapping Mack's Model



willistowerswatson.com

© 2017 Willis Towers Watson. All rights reserved. Proprietary and Confidential. For Willis Towers Watson and Willis Towers Watson client use only.

How Bayesian models are different

- The Mack and Bootstrap models assume that the loss generating process follows a prescribed analytical framework
- Bayesian models are different in three aspects
 - The probability density function associated with a loss generating process is based on fitting a GLM to the data and reflects actual loss emergence
 - The modeler can then make a "first guess" of the distributional format of the parameters underlying the probability density function
 - The final selection of the parameters of the probability density function employs both user judgment and statistical information of the actual data on hand

Introduction to Bayesian modeling:

Statistics 101

Statistics 101 – Bayes Theorem

- Bayes theorem indicates how prior subjective belief changes based on evidence
- $P(A/B) = \frac{P(B/A)P(A)}{P(B)}$, where
- P(A) is the prior belief
- P(A/B) is the posterior belief accounting for B
- P(B/A)/P(B) represents the support B provides to A

Statistics 101- Likelihood function

- Usually we think in terms of probabilities, i.e., the probability of an outcome X given a parameter Θ; P(X| Θ)
- The Likelihood instead is a function of Θ given an outcome, i.e. $L(\Theta|X)$
- With an observed outcome X the maximum likelihood principle chooses the parameter Θ that maximizes the P(X| Θ)
- Aggregate GLM models produce maximum likelihood functions

Statistics 101 – Posterior Probability

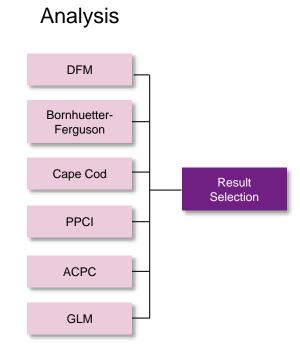
• Given prior belief $p(\theta)$ and observation x with likelihood $P(x / \theta)$ the posterior is:

$$P(\theta/x) = \frac{P(x/\theta)P(\theta)}{P(x)}$$
, i.e

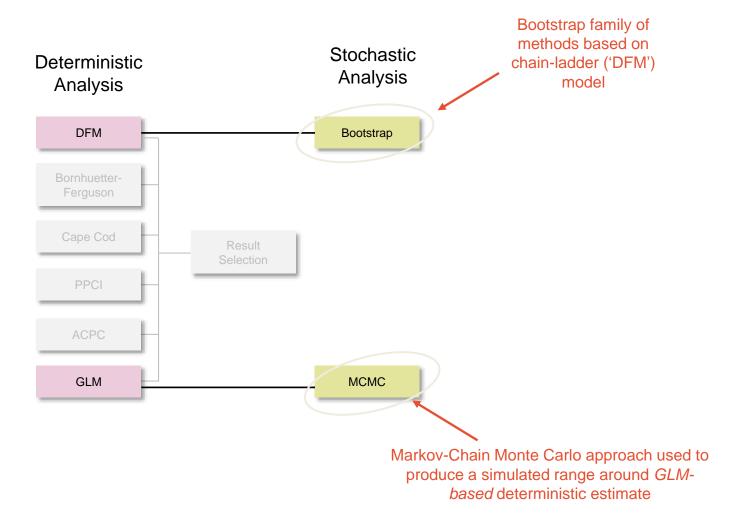
■ Posterior probability ∝ Likelihood x Prior Probability

Statistics 101 – MCMC compared to Bootstrapping

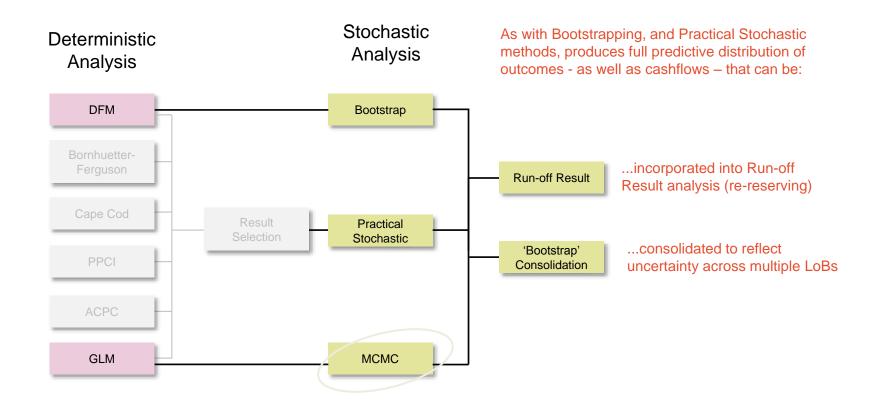
Deterministic



Statistics 101 – MCMC compared to Bootstrapping (cont'd)



Statistics 101 – MCMC compared to Bootstrapping (cont'd)



Statistics 101 – MCMC compared to Bootstrapping (cont'd)

Similarities

- Both are simulation methods
- They produce a full predictive distribution of outcomes including
 - Parameter risk
 - Process risk

Differences

- Bootstrapping
 - Samples with replacement the residuals from an actual versus expected comparison of historical development
 - Simulations are independent from one another
 - There is no convergence in the simulations
- MCMC
 - Samples the parameters of the resulting GLM likelihood function
 - Simulations are built through a Markov chain
 - The simulations converge into an equilibrium state

Introduction to Bayesian modeling:

Bayesian Modeling Steps

Bayesian modeling steps

Step 1: Specify probability distribution for the data given some unknown parameters (data distribution)

Step 2: Specify prior probability distribution for the parameters of the data distribution (prior distribution)

Step 3: Derive the likelihood function of the parameters, given the data (likelihood function)

Step 4: Combine prior distribution and likelihood function to derive posterior joint distribution of parameters(posterior distribution)

Step 5: Obtain parameters for posterior distribution

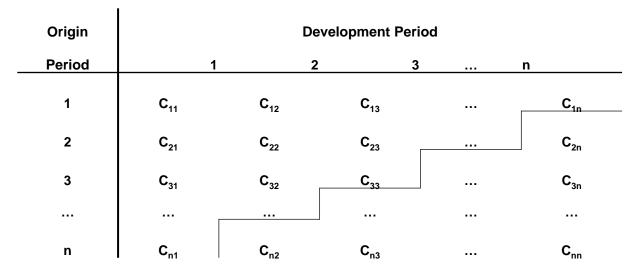
Step 6: Combine data distribution and posterior distributions to obtain forecast of predictive distribution

Introduction to Bayesian modeling:

Bayesian Modeling within the Reserving Context

Bayesian modeling within the reserving context

Reserving example



 C={Cij: i+j<n+1} is the upper-left triangle of observed payments, and the reserving problem attempts to estimate the unobserved values in the lowerright triangle

Bayesian modeling within the reserving context (cont'd)

- Step 1&2: Assumes C_{ij} follows a probability density distribution of f(C_{ij} / θ), where θ denotes parameters describing a particular claims generating process and π(θ) is the prior distribution function
 - An example of the probability density function is the ODP model
 - Prior distributions could assume some distributional phase (i.e. Lognormal, Gamma)
- Step 3: The likelihood function $L(\theta/\underline{c})$ for the parameters given observe data $L(\theta/\underline{c}) = \prod_{i+j \le n+1} f(C_{ij}/\theta)$

Bayesian modeling within the reserving context (cont'd)

- Step 4: Given the data distribution and the prior distribution, the posterior distribution f(θ/c) is proportional to the product of the likelihood and the prior: f(θ /c) ∝ L(θ /c) π(θ)
- Step 5: Parameters θ are obtained from the posterior distribution f(θ /c) and are used in Step 6
 - When the shape of the posterior distribution is not known, special statistical algorithms, like Gibbs MCMC, are employed

Bayesian modeling within the reserving context (cont'd)

Step 6: The known data C_{ij} for i + j ≤ n + 1 is used to predict unobserved values in the lower right triangle C_{ij} for i + j > n + 1 by means of the predictive distribution:

 $f(C_{ij}/\underline{c}) = \int f(C_{ij}/\theta) f(\theta/\underline{c}) d\theta$

- When the analytical format of the predictive distribution is not known generic sampling algorithms such as Adaptive Rejection Sampling (ARS) might be needed
- Predictive distribution can either be obtained in a closed form analytically or through a generic sampling algorithm instead

Introduction to Bayesian modeling:

Simple Example – No Simulations Needed

Simple example – no simulations needed

- Step 1&2: Assume the loss generating process follows a Poisson distribution with parameter and the parameter follows a Gamma distribution with some known parameters a and b
- $C_{ij}/\theta \sim \rho(\theta)$
- θ /a,b ~ *Gamma*(*a*,*b*)

Simple example – no simulations needed (cont'd)

- Step 3: The likelihood function is given by $L(\theta / \underline{c}) = \prod_{i=1}^{n} \theta^{x_i} e^{-\theta} / x_i!$
- Step 4: The posterior distribution is proportional to the product of the likelihood and the prior: $f(0/a + b) = \prod_{k=0}^{n} e^{\lambda t} e^{-\theta} / x + b^{2} / \sum_{k=0}^{n} e^{-\theta} = 0$

 $f(\theta/\underline{c},a,b) = \prod_{i=1}^{n} \theta^{x_i} e^{-\theta} / x_i! b^a / \Gamma(a) \theta^{a-1} e^{-b\theta} =$

~ Gamma (a + $\sum_{i=1}^{n} x_i$, b + n), i.e. posterior follows a Gamma distribution

In statistics it is a well known fact that the product of a Poisson distribution and a conjugate prior Gamma distribution results into a Gamma distribution

Simple example – no simulations needed (cont'd)

Step 6: The product of the posterior and the data distribution, i.e., the product of a Gamma and a Poisson distribution results in a negative binomial distribution:

NB(a +
$$\sum_{i=1}^n x_i$$
 , $rac{1}{1+b_1}$)

No need for complicated sampling here !

Introduction to Bayesian modeling:

Examples of Popular Sampling Techniques

What are the MCMC sampling techniques in general?

- MCMC methods are a class of algorithms for sampling from a probability distribution
 - This distribution is usually difficult to approximate with analytical functions
- MCMC constructs a random process that undergoes transition from one state to another, called *Markov Chain*
 - This process is *memoryless*, i.e. the next state is based only on the current state but not the sequence of the preceding states
 - The quality of the convergence to an equilibrium distribution improves with the number of steps employed in the process
 - The first few draws are usually thrown away (called *burn-in*) to ensure target is independent of starting point and improve convergence

Examples of popular sampling techniques Gibbs sampler

- Gibbs sampler avoids sampling from a complicated bivariate distribution f(x,y) by making random draws instead from univariate conditional distributions (f(x/y) and f(y/x))
- For two parameters and n iterations it produces an

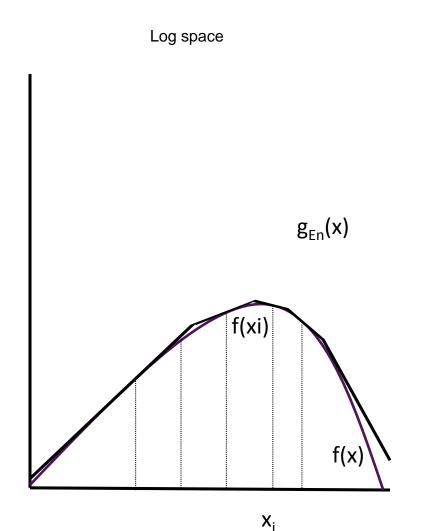
n x 2 table where x_0 is the initial value – Next steps:

$$y_1 \, \tilde{f}(y/x = x_0)$$

 $x_i \, \tilde{f}(x/y = y_{i-1})$
 $y_i \, \tilde{f}(y/x = x_i)$

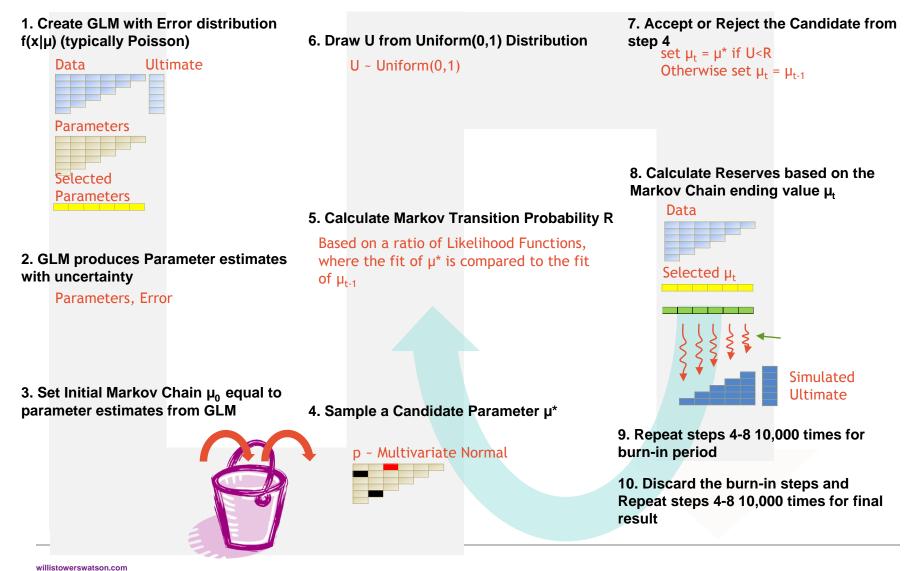
- Eventually (x_i,y_i) → (x,y)[~]f(x,y) for sufficient large number of iterations (so called burn-in sample)
- After burn-in, it is common to define a spacing between accepted points, maybe every m draws, to ensure independence of random draws

Examples of popular sampling techniques Adaptive Rejection Sampling



- ARS works with log concave densities f(x)
- An envelope function g_{En}(x) as an upper bound of the log density function is employed
- A random draw x_i from the x-axis is then sampled
- When the resulting g_{En}(x_i) is close to f(x_i) the envelope function remains unchanged
- When the resulting g_{En}(x_i) is much larger to f(x_i) the envelope function changes to incorporate a line that is tangent to f(x_i)

Examples of popular sampling techniques Metropolis-Hastings Algorithm



© 2017 Willis Towers Watson. All rights reserved. Proprietary and Confidential. For Willis Towers Watson and Willis Towers Watson client use only.

Introduction to Bayesian modeling: Conclusions

Conclusions

MCMC Bayesian stochastic reserving method has both

Advantages:

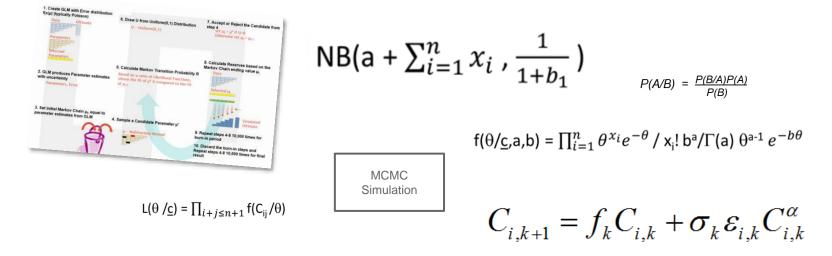
- Flexible not constrained by any "prescribed" assumption on the format of the loss generating process
- Allows the incorporation of user's judgment
- It is based on statistical characteristics of the data on hand

Disadvantages:

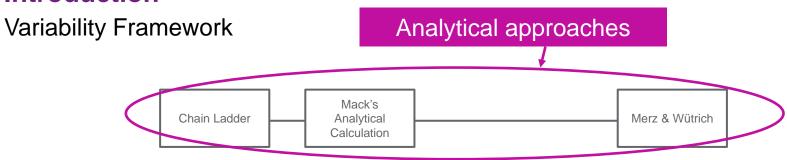
- More sophisticated mathematics
- Can be influenced by judgment
- Actuaries are "scared" of it

Part II: Deploying a stochastic modeling framework

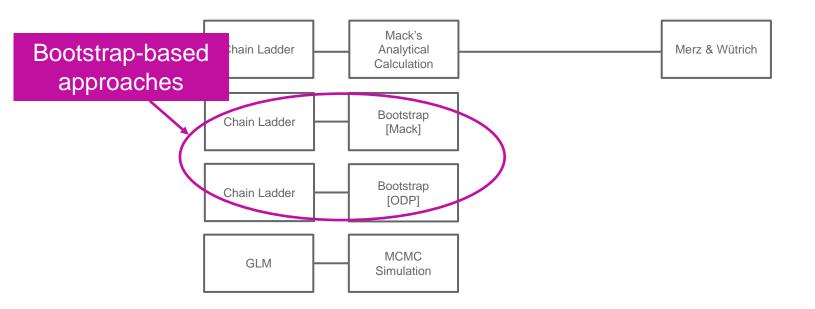
Manolis has provided an introduction to a number of stochastic approaches, focusing on Bayesian techniques

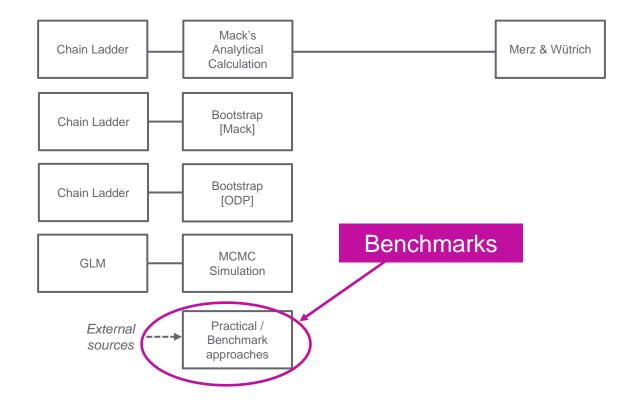


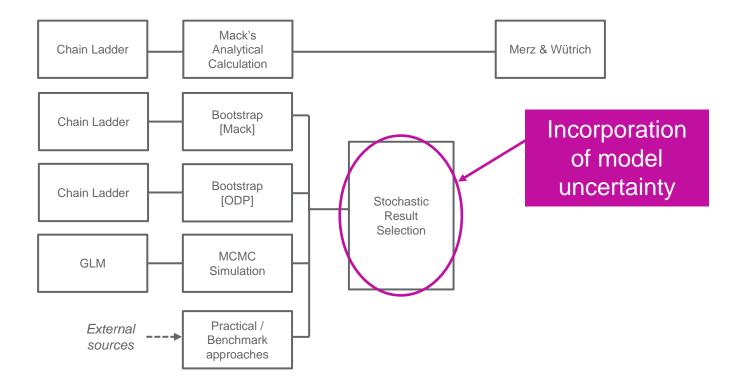
We're now going to shift our view to look at the wider array of variability approaches that are all widely available with the various reserving software applications currently on the market

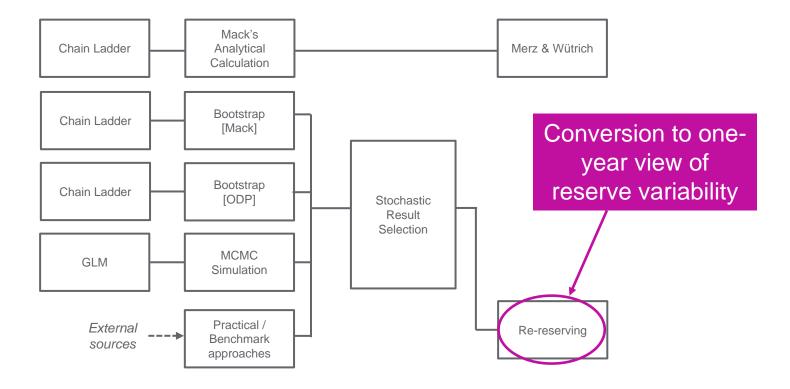


GLM		MCMC Simulation
-----	--	--------------------



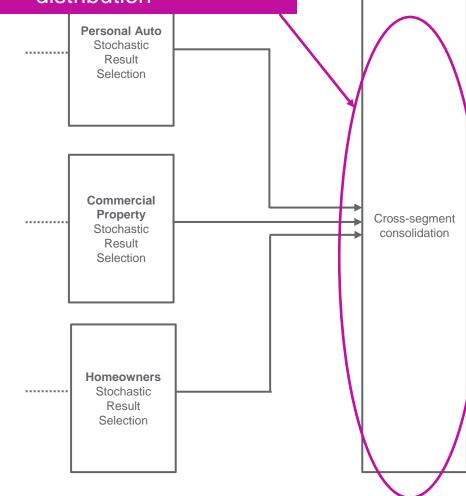


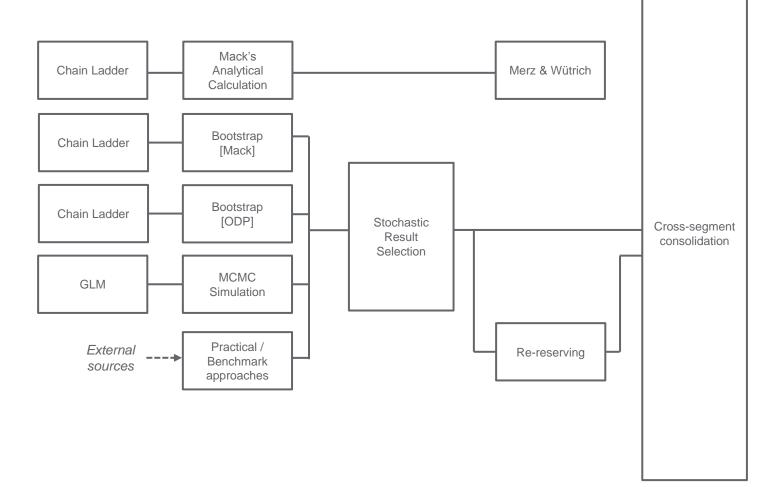




Introduction Variability Framework

Consolidation of simulations across multiple lines for aggregate view of distribution





Stochastic Models Beyond the Bootstrap...

Part I: Introduction to Bayesian modeling

- Reserve Uncertainty in General
- From Mack and Bootstrapping to Bayesian MCMC Models
- Statistics 101
- Bayesian Modeling Steps
- Bayesian Modeling within the Reserving Context
- Simple Example No Simulations Needed
- Examples of Popular Sampling Techniques
- Conclusions

Part II: Deploying a stochastic modeling framework

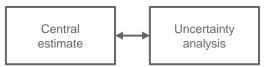
- The link between assumptions in our deterministic framework and our uncertainty analysis
- Selection of models
- Use of multiple models
- Using benchmarks

Deploying a stochastic modeling framework:

The link between assumptions in our deterministic framework and our uncertainty analysis

From central estimate to range

We often find that a 'variability' study is undertaken as a separate analysis to the 'central estimate'

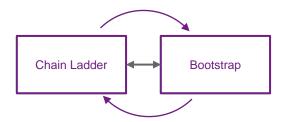


However, the approach to estimating the uncertainty in either a prediction or an outcome should be closely associated with the approach used in determining the single estimate

This is true both in a general sense of the analytical approach that is used, and also in the assumptions that are used

From central estimate to range

An example of this is the intrinsic link between bootstrap-based approaches and the underlying chain- or loss development factor models...



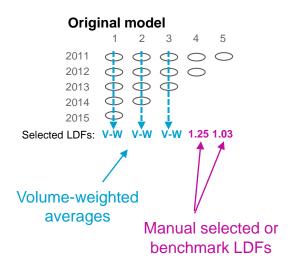
... which is basically just doing a bunch of simulated chain-ladders

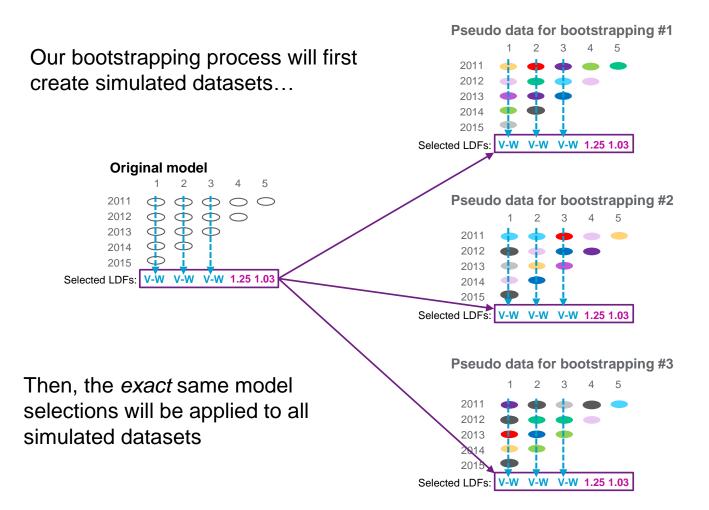
Therefore, when reviewing or performing a bootstrapped-based variability analysis, we must remember to look not just at the simulation and variance settings, but also at the underlying model that is being applied

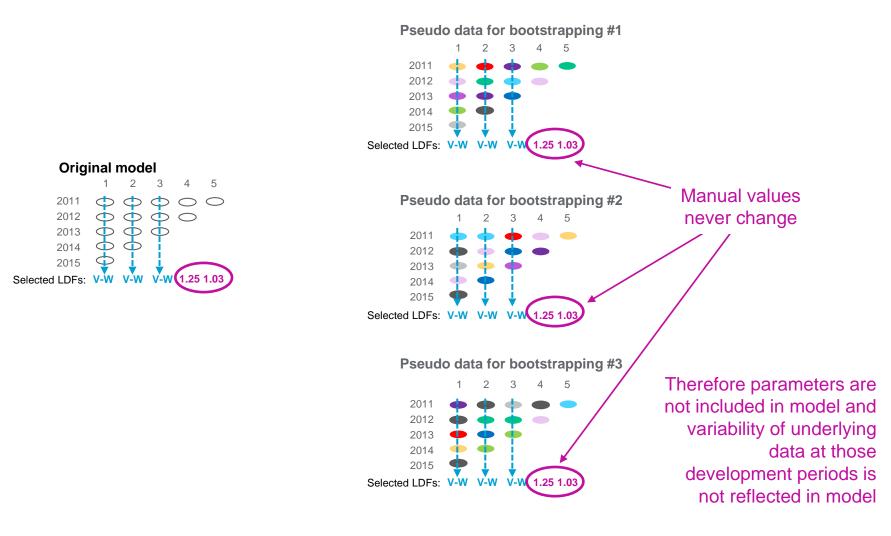
For example....

Imagine our loss development factor model uses:

- Volume-weighted ("vw") averages for first 3 development periods
- Manually-selected ldfs for development periods 4 and 5







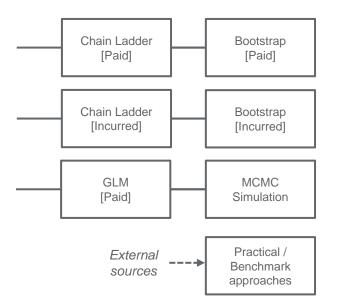
Simulation approaches such as bootstrapping or MCMC are only as robust as the projection method and the data upon which they are based

The user should be aware of the impact of any changes away from 'standard' assumptions

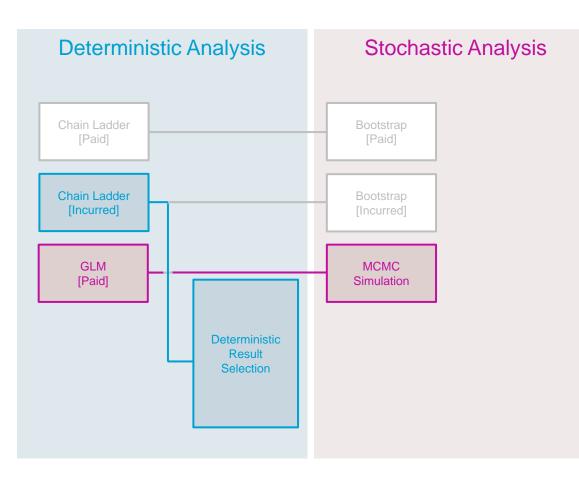
- Does data / model used in reserve risk model match that which drives deterministic estimate?
- Is underlying model (e.g. chain-ladder or GLM) appropriate?
 - Is there sufficient data?
 - Is the underlying model appropriate for simulation?
 - What if the data has changed over time?

When reviewing or performing a variability analysis, *look beyond* the selections made in your stochastic models, and ensure the implication of the selections in your underlying deterministic models is understood

Deploying a stochastic modeling framework: Selection of models

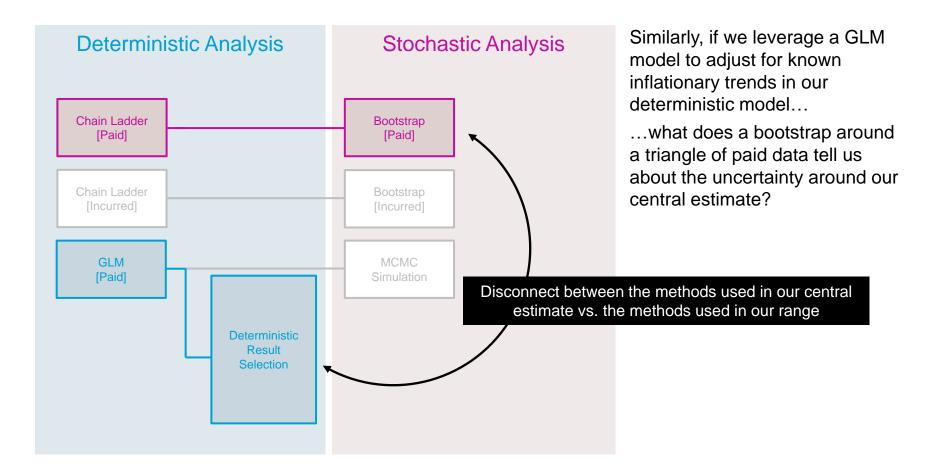


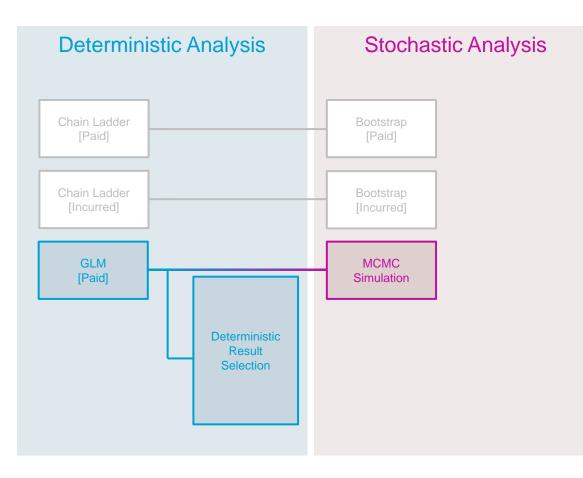
We're now going to take a step back and look not at the *selections* used in the model, but which models/data are used in our analysis to begin with...



Imagine if our deterministic estimate is based primarily on incurred chain-ladder model

While performing a MCMC type simulation applied to a GLM model might sound good, does it actually tell us anything useful about the variability around our selected estimate?





Far more insightful to use a consistent set of models in both our deterministic and stochastic frameworks

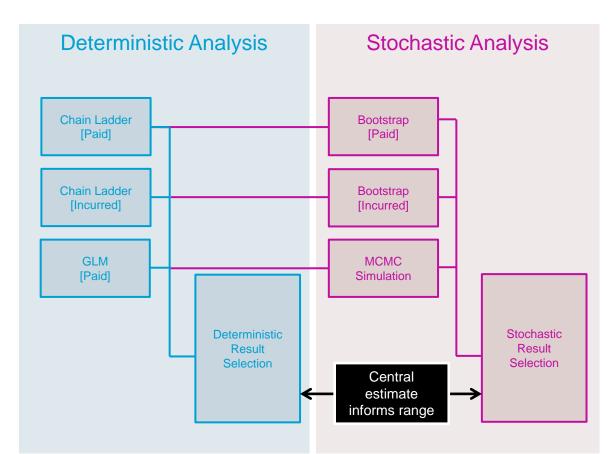
Summary

When selecting an model for developing a range, *look beyond* the paid bootstrap and use an approach that is most consistent with that used for selecting a central estimate

- If loss development-based models are primarily used in best-estimate modeling, then use loss development-based models in uncertainty analysis
- Similarly, if you typically rely on reported loss models for your deterministic analysis, then using paid loss models in your uncertainty models may not be reflective of the actual or potential volatility

What if we *do* use multiple deterministic models to inform our central estimate using weighting that reflects our confidence in each model...?

Deploying a stochastic modeling framework: Using multiple models

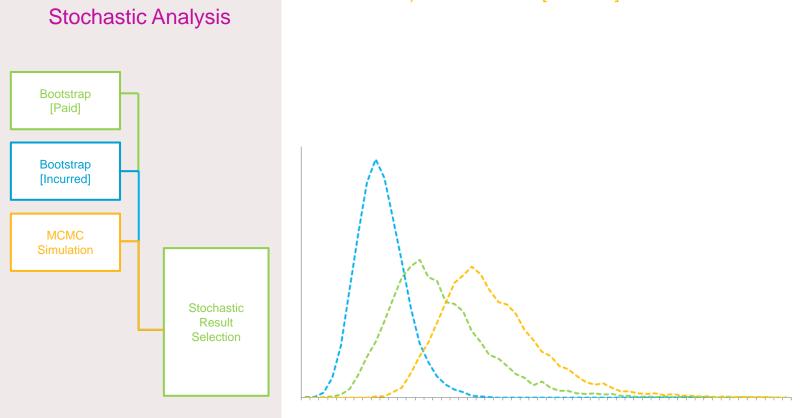


What if we *do* use multiple deterministic models to inform our central estimate using weighting that reflects our confidence in each model?

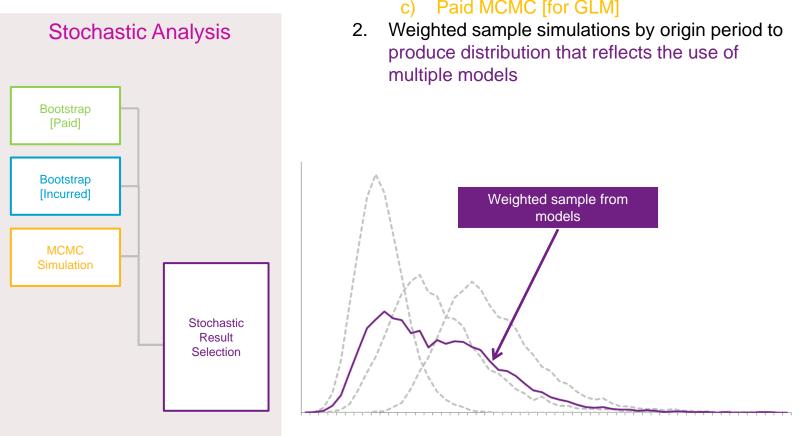
In this, case, no one variability model in isolation will be able to reflect the uncertainty around our central pick

Our approach for deriving the uncertainty in our variability estimate should similarly reflect our reliance on multiple underlying models

This further ensures a coherent relationship between central- and range- estimates

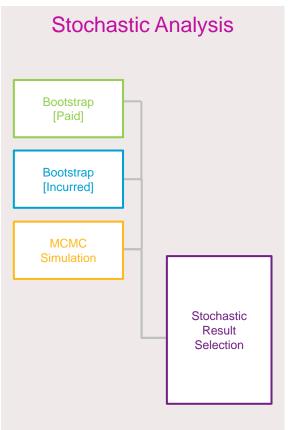


- 1. Produce distributions around each model used in deterministic result:
 - a) Paid Bootstrap [for CL]
 - b) Incurred Bootstrap [for CL]
 - c) Paid MCMC [for GLM]

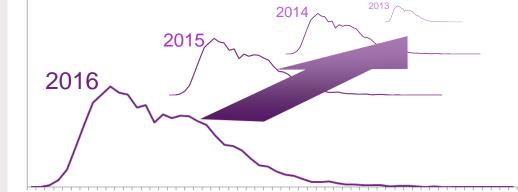


- Produce distributions around each model used in 1. deterministic result:
 - a) Paid Bootstrap [for CL]
 - b) Incurred Bootstrap [for CL]
 - c) Paid MCMC [for GLM]

willistowerswatson.com © 2017 Willis Towers Watson. All rights reserved. Proprietary and Confidential. For Willis Towers Watson and Willis Towers Watson client use only.



- 1. Produce distributions around each model used in deterministic result:
 - a) Paid Bootstrap [for CL]
 - b) Incurred Bootstrap [for CL]
 - c) Paid MCMC [for GLM]
- 2. Weighted sample simulations by origin period to produce distribution that reflects the use of multiple models
- 3. Aggregate origin year results (using correlation matrix) to derive Total Reserve distribution



'Traditional' models for calculating reserve uncertainty are typically based on a single underlying deterministic model and include **parameter** and **process** components of uncertainty

Weighted sampling simulations from a range of variability methods allows us to reflect in our variability estimate our reliance on multiple **models**



Summary

Models used in uncertainty analysis should reflect the models used in deterministic analysis, even if that means incorporating multiple models

Weighted-sampling is a relatively simple way to leverage the output from multiple simulation-based models

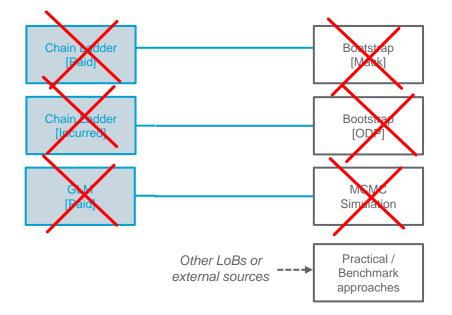
Sampling is ideally performed on an individual origin-year basis that reflect the deterministic weights use for that year

Care should be taken to correlate the individual year results in a manner that is appropriate for derivation of a range around the overall (i.e. 'all year') result

When developing a framework for investigating uncertainty, *look beyond* any single model and incorporate the indications from multiple models – just as you would in your central estimate

Deploying a stochastic modeling framework: Using benchmarks

Using benchmarks



What happens when you can't use a chain-ladder or a GLM?

- Incomplete data?
- Insufficient development history?
- Etc.

What happens if can't apply a bootstrap or an MCMC type approach?

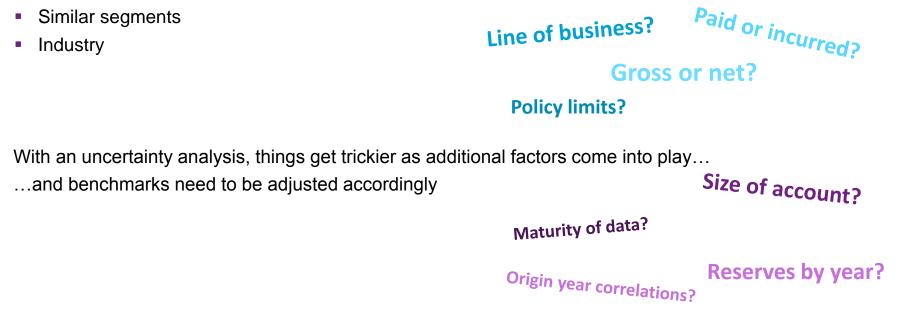
- Not enough data points?
- Didn't rely on CL or GLM for central estimate?
- Etc.

As with deterministic projections, even when the data fails, we need to provide some kind of estimate This is where benchmarks come in useful

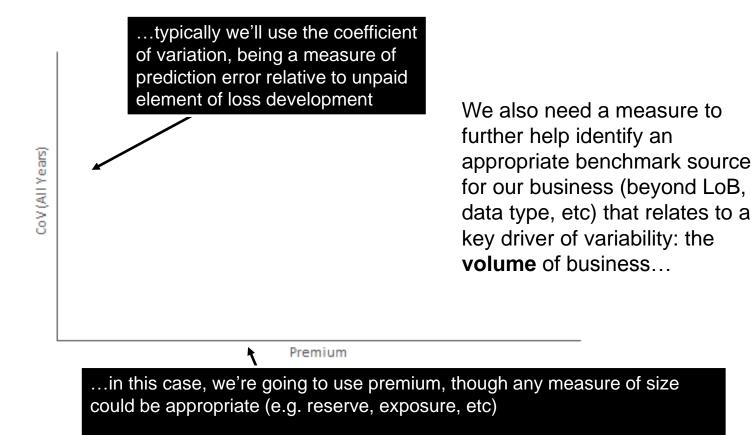
Benchmarking

What happens when the data available is simply not sufficient for modeling?

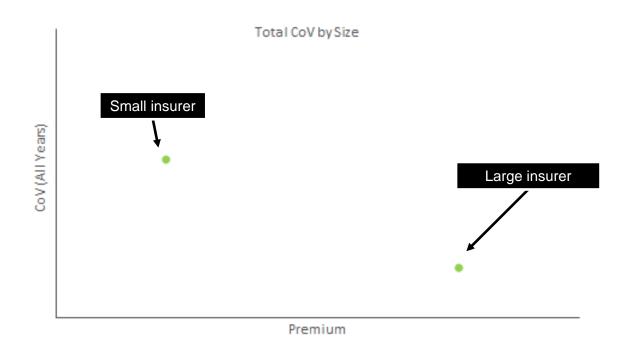
As with in a deterministic framework, we can look to leverage benchmarks (development patterns, expected loss ratios, etc), based either on:

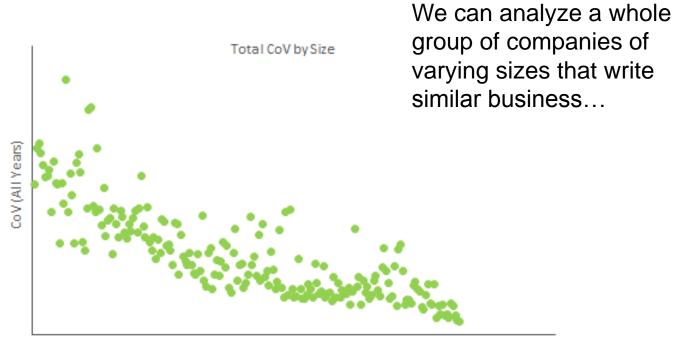


Just as with a deterministic approach, we need to select a measure to use as our benchmark...

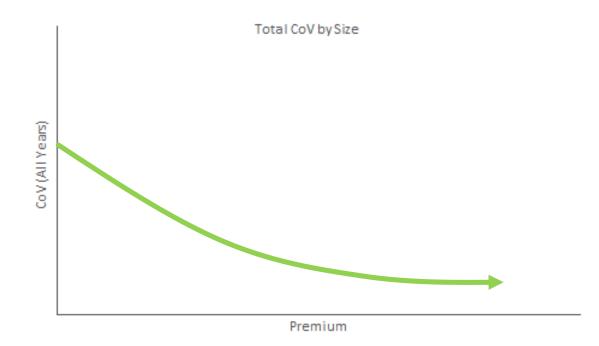


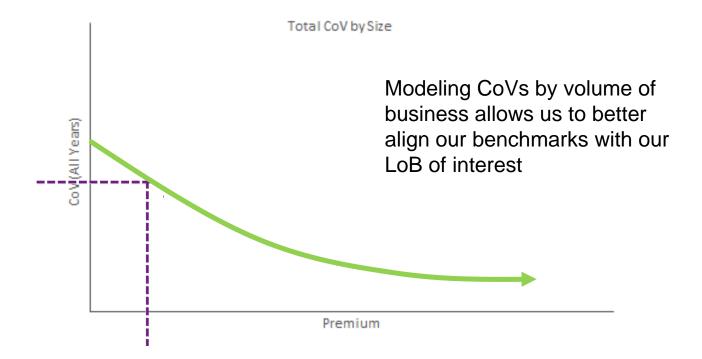
Where would we expect a larger insurer to lie on the graph?





Premium





Summary

Just as in a deterministic framework, selecting an appropriate variability benchmark depends on appropriately identifying the key properties that drive the uncertainty:

- Which segment?
- Gross or net?
- Paid or incurred?

Additionally - for variability benchmarks - we also must consider:

- Size of account?
- Maturity of the data
- How are origin years correlated?

If your data does not support a robust simulation-based approach, *look beyond* your own historical development data and incorporate benchmarks, adjusting accordingly

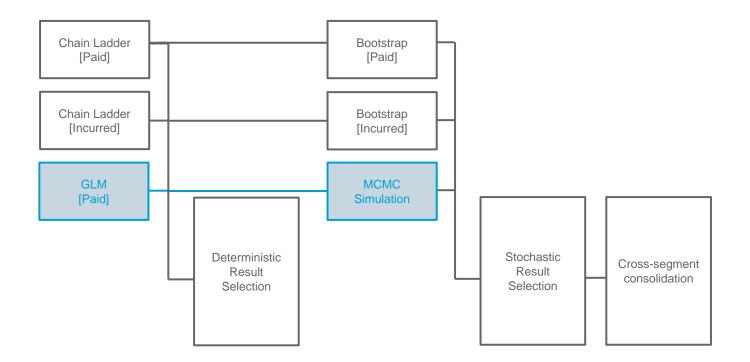


Selected benchmark

Deploying a stochastic modeling framework: Summary

Summary

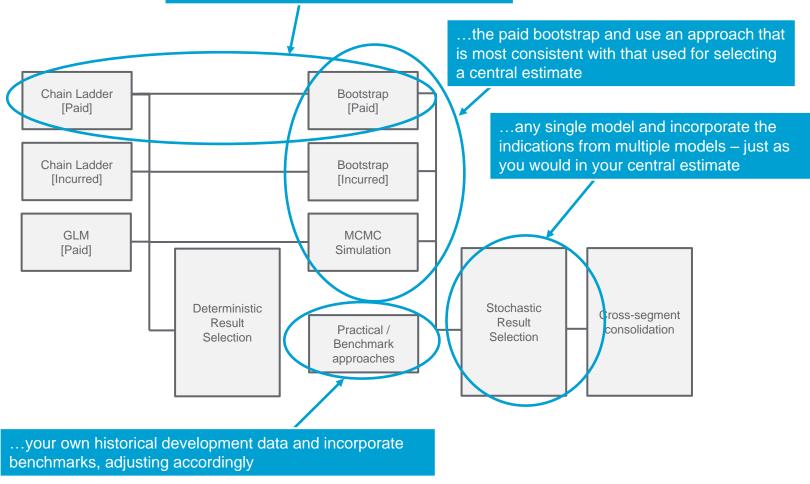
It doesn't matter how statistically-robust your chosen model is...if you're careless with how you deploy, adjust or combine the output



Stochastic framework should be an *extension* of your deterministic process – not a separate analysis

Look beyond...

...the selections made in your stochastic models, and ensure the implication of the selections in your underlying deterministic models is understood



Wrap-up

Stochastic Models Beyond the Bootstrap...

Part I: Introduction to Bayesian modeling

- Reserve Uncertainty in General
- From Mack and Bootstrapping to Bayesian MCMC Models
- Statistics 101
- Bayesian Modeling Steps
- Bayesian Modeling within the Reserving Context
- Simple Example No Simulations Needed
- Examples of Popular Sampling Techniques
- Conclusions

Part II: Deploying a stochastic modeling framework

- The link between assumptions in our deterministic framework and our uncertainty analysis
- Selection of models
- Use of multiple models
- Using benchmarks

Thank you