CS-13: Simulation Modeling: Pitfalls and Best Practices

2009 CAS Seminar on Reinsurance

Thomas Hettinger

Managing Director North America

May 18th, 2009

Agenda

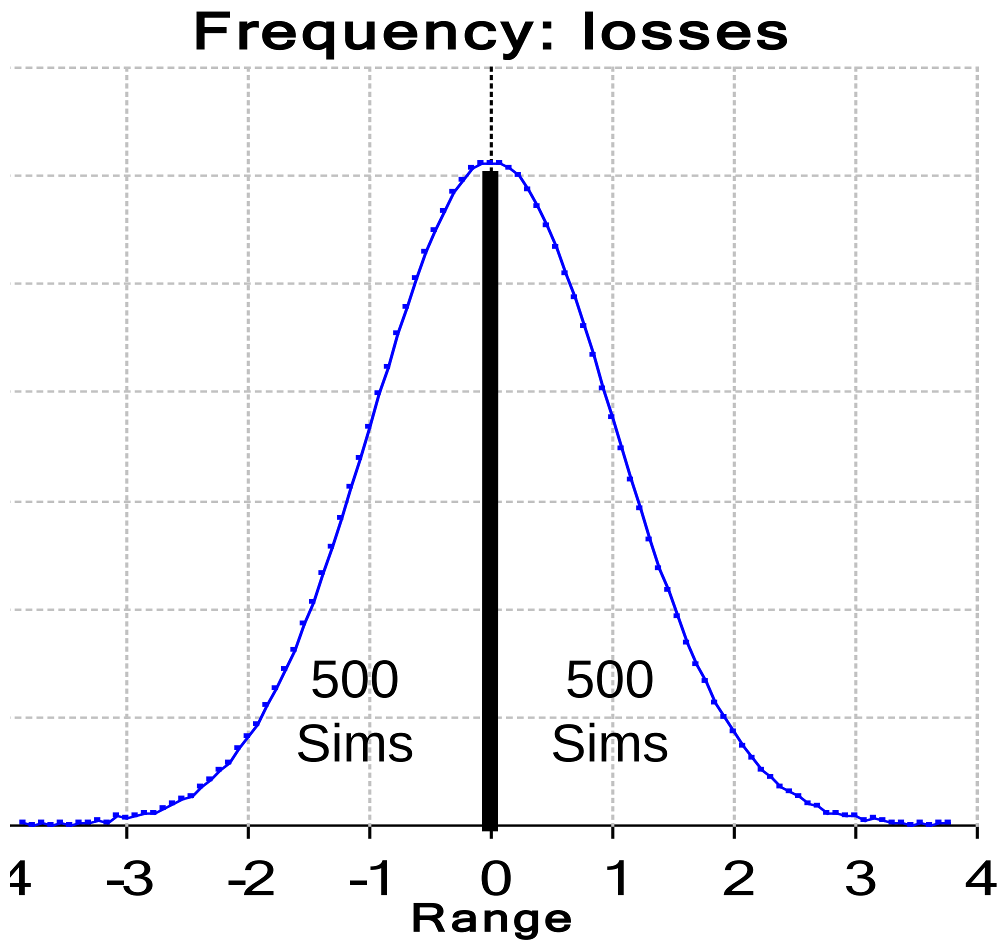
- Stratified Sampling
  - Basics
  - Benefits
  - Limitations
- Model Convergence
  - Metrics
  - Metric Weakness
  - Options
- Parameter Uncertainty
  - Include or Not
  - Consistency Issues
  - Practical Issues

Stratified Sampling - Basics

- Stratification is a technique to ensure that simulated values are distributed with a degree of evenness across the possible values for a distribution. The number of simulations in each base unit must be divisible by the number of strata in that base unit, so that each stratum has the same number of simulations within it. In the trivial case of 1 stratum, the random numbers generated are unconstrained. With n simulations and m strata, the simulations are constrained so that n/m simulations occur below the 100/m percentile of the distribution, n/m between the 100/m and 200/m percentiles, and so on.

Stratified Sampling – Simple Example

- Normal(0,1)
- 1000 Simulations
- 2 Strata



Stratified Sampling - Benefits

- Stratified sampling will result in lower sampling error
- Reduce variance of results
- Is very effective for highly skewed distributions
- Improves dependency simulation
  - Tail dependency

Stratified Sampling – Benefits Example

- 10,000 simulations
- Random seed reset 5 different times
  - Results ran with and without stratification
- Average

	Non Stratified					Stratified				
	1	2	3	4	5	1	2	3	4	5
Normal	1,000	998	1,009	990	1,005	1,000	1,000	1,000	1,000	1,000
Gamma	998	1,003	996	1,000	1,007	1,000	1,000	1,000	1,000	1,000
Lognormal	996	1,007	1,001	1,002	1,002	1,000	1,000	1,000	1,000	1,000
Pareto	1,715	1,715	1,717	1,718	1,719	1,719	1,718	1,718	1,718	1,718
GPD	1,699	1,697	1,702	1,699	1,697	1,700	1,700	1,700	1,700	1,700
Exponentia	1,711	1,701	1,673	1,710	1,709	1,700	1,700	1,700	1,700	1,700
Normal		0%	1%	-1%	1%		0%	0%	0%	0%
Gamma		0%	0%	0%	1%		0%	0%	0%	0%
Lognormal		1%	1%	1%	1%		0%	0%	0%	0%
Pareto		0%	0%	0%	0%		0%	0%	0%	0%
GPD		0%	0%	0%	0%		0%	0%	0%	0%
Exponential		-1%	-2%	0%	0%		0%	0%	0%	0%

Stratified Sampling – Benefits Example

- 10,000 simulations
- Random seed reset 5 different times
  - Results ran with and without stratification
- 99th Percentile

	Non Stratified					Stratified				
	1	2	3	4	5	1	2	3	4	5
Normal	2,190	2,176	2,188	2,152	2,164	2,163	2,164	2,165	2,165	2,163
Gamma	2,471	2,491	2,471	2,545	2,534	2,515	2,513	2,514	2,513	2,513
Lognormal	2,574	2,792	2,679	2,728	2,686	2,688	2,689	2,688	2,688	2,687
Pareto	3,261	3,211	3,278	3,294	3,294	3,288	3,289	3,288	3,289	3,290
GPD	2,772	2,806	2,796	2,788	2,732	2,783	2,781	2,781	2,782	2,783
Exponentia	7,918	7,615	7,549	7,874	7,874	7,838	7,831	7,840	7,831	7,841
Normal		-1%	0%	-2%	-1%		0%	0%	0%	0%
Gamma		1%	0%	3%	3%		0%	0%	0%	0%
Lognormal		8%	4%	6%	4%		0%	0%	0%	0%
Pareto		-2%	1%	1%	1%		0%	0%	0%	0%
GPD		1%	1%	1%	-1%		0%	0%	0%	0%
Exponential		-4%	-5%	-1%	-1%		0%	0%	0%	0%

Stratified Sampling – Benefits Example

- 10,000 simulations
- Random seed reset 5 different times
  - Results ran with and without stratification
- Window 99th VaR

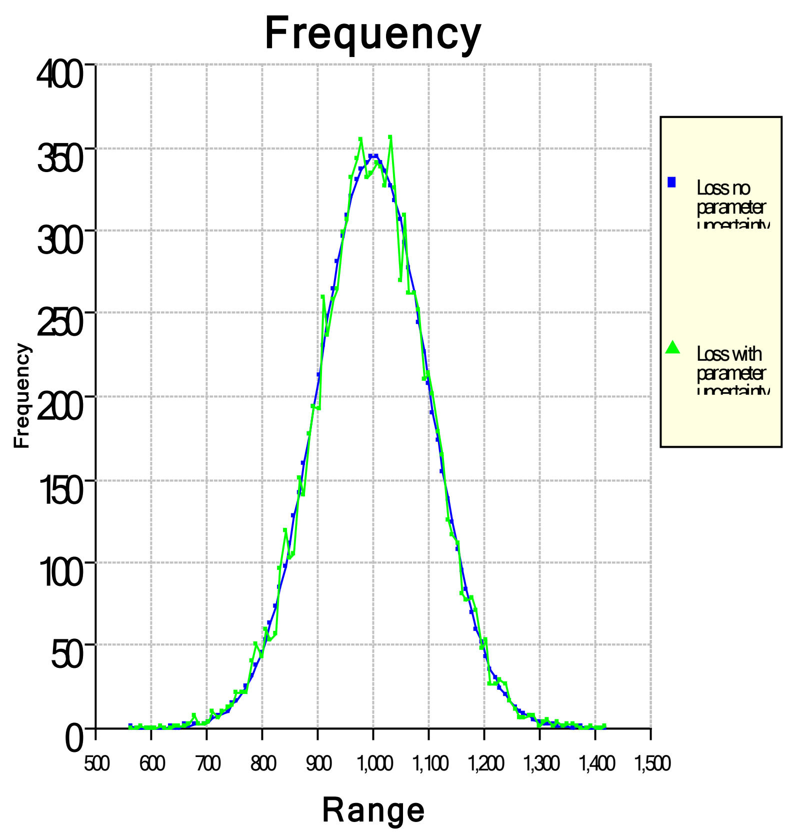
	Non Stratified					Stratified				
	1	2	3	4	5	1	2	3	4	5
Normal	2,197	2,176	2,188	2,162	2,173	2,171	2,171	2,171	2,170	2,171
Gamma	2,484	2,529	2,493	2,570	2,560	2,526	2,526	2,526	2,526	2,526
Lognormal	2,603	2,786	2,697	2,747	2,694	2,707	2,707	2,707	2,707	2,707
Pareto	3,309	3,234	3,319	3,331	3,324	3,320	3,320	3,320	3,320	3,320
GPD	2,786	2,814	2,804	2,791	2,756	2,794	2,794	2,794	2,794	2,794
Exponentia	7,974	7,770	7,652	7,983	7,949	7,905	7,903	7,905	7,904	7,905
Normal		-1%	0%	-2%	-1%		0%	0%	0%	0%
Gamma		2%	0%	3%	3%		0%	0%	0%	0%
Lognormal		7%	4%	6%	4%		0%	0%	0%	0%
Pareto		-2%	0%	1%	0%		0%	0%	0%	0%
GPD		1%	1%	0%	-1%		0%	0%	0%	0%
Exponential		-3%	-4%	0%	0%		0%	0%	0%	0%

Stratified Sampling – Limitations

- Need to be able to define the inverse distribution function
  - Have to know how to divide up the distribution
- Stratified sampling is slower
  - Generate stratified uniform(0,1) and then use the inverse distribution type for the specified distribution
- Some distributions can be slower than others
  - Anything that relies on a gamma distribution
  - InverseGaussian, Beta, Gamma, Chi Squared, F Distribution, and Student’s t
    - Speed difference depends on actual parameters of underlying distributions
    - In the grand scheme we have not seen a major impact on most models

Stratified Sampling – Limitations

- Care used when simulating with parameter uncertainty
  - For example if you pick a different catastrophe (of potentially different type) each simulation, which samples with a mean loss that varies by catastrophe type, then stratification won't help. You need deterministic parameters.
- Within Accept and Reject functions stratified sampling would create an infinite loop.
  - Accept(Normal(0,1), @>1.25)
    - consider using the Inv<distribution> function with a restricted range, e.g.
    - InvNormal(0,1,Uniform(0.8,1))



Model Convergence - Basics

- Closely related issue with stratification
- Less simulations needed for convergence when using stratification
  - Take the example of saverage(normal(0,1)).
    - With non stratified sampling, this will converge as 1/sqrt(nsims).
    - With stratified sampling it will go as 1/nsims.
    - The convergence for extreme values (e.g. 99th percentile) is not as good as pure averages, but is still much better than non-stratatification

Model Convergence - Considerations

- Considerations
  - Metric to measure
    - Weaknesses of metric
  - Consider individual distributions or end distribution
  - Tolerances

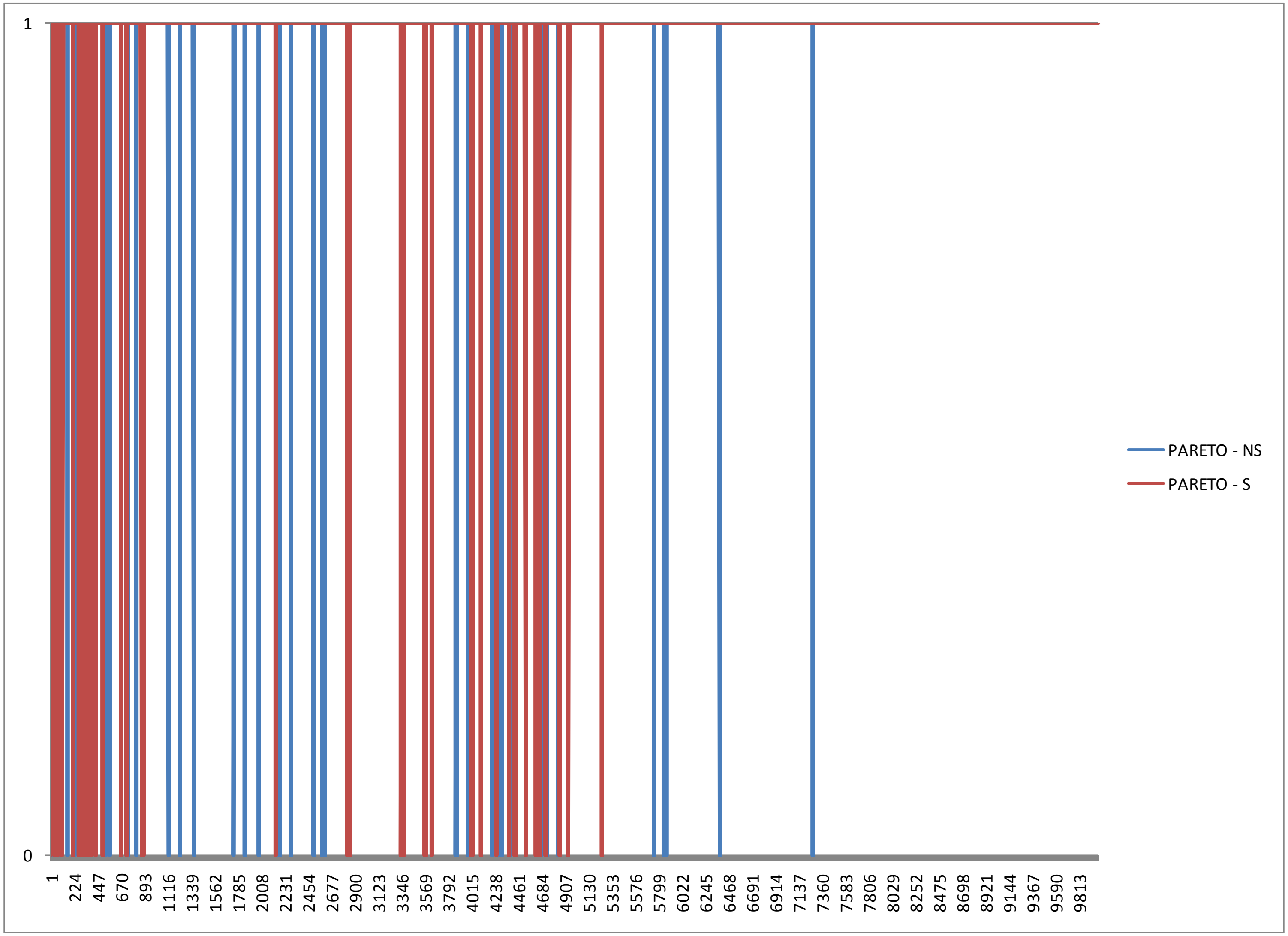
Model Convergence - Metrics

- What is the end goal?
  - RI pricing needs to consider extremes
  - May choose 99th percentile to assess convergence
  - One point is very sensitive to results
- Do you have the luxury to over simulate
  - Not always possible with large models
- Is a metric converged if the change in one period is less than tolerance?
  - Convergence of X# of sims may be more conservative

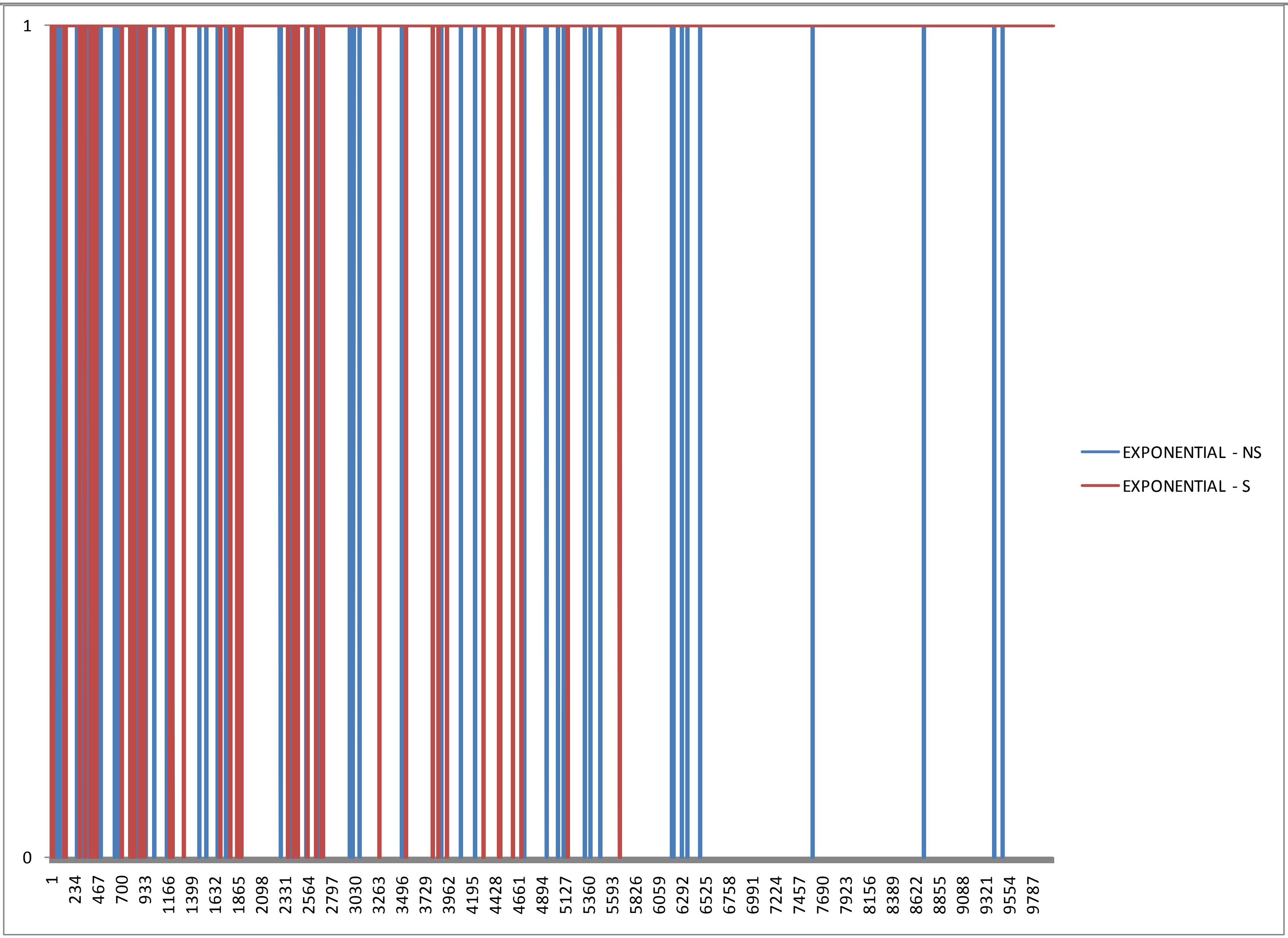
Model Convergence - Testing

- Over simulation and see convergence point
- Test variability in critical metrics for operations
  - Reset random number seeds
  - Up simulations and retest if convergence is not met
- Tolerances
  - What is the risk to the end use of the model?
  - What is the risk tolerance for the firm's ERM process?

Model Convergence - Testing



Model Convergence - Testing

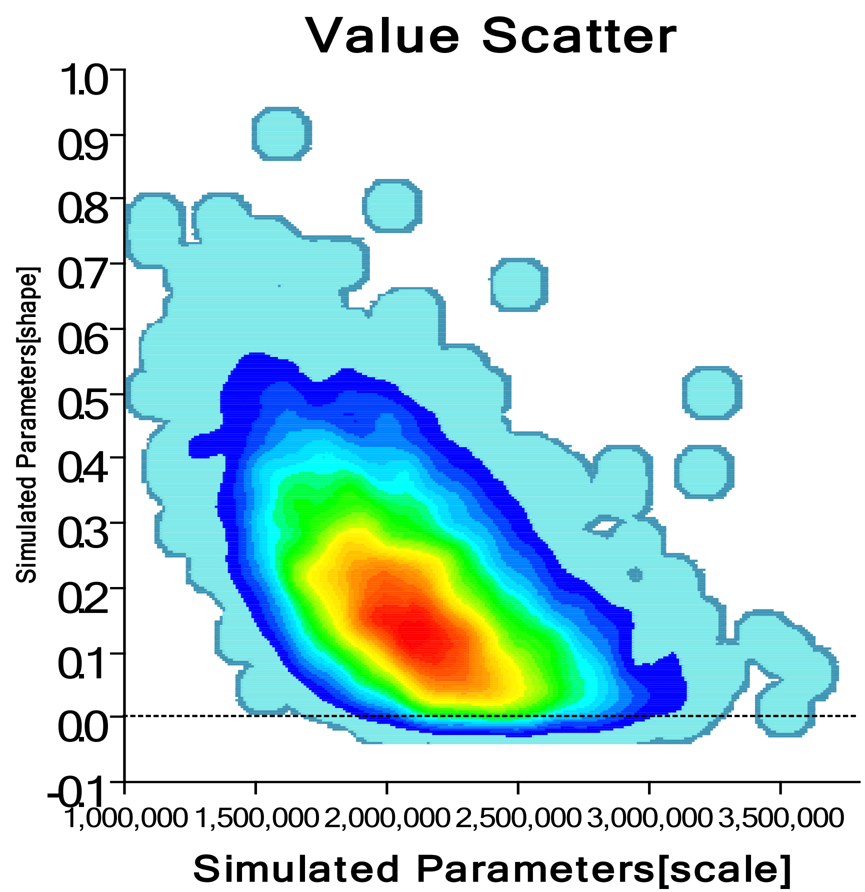


Model Convergence - Options

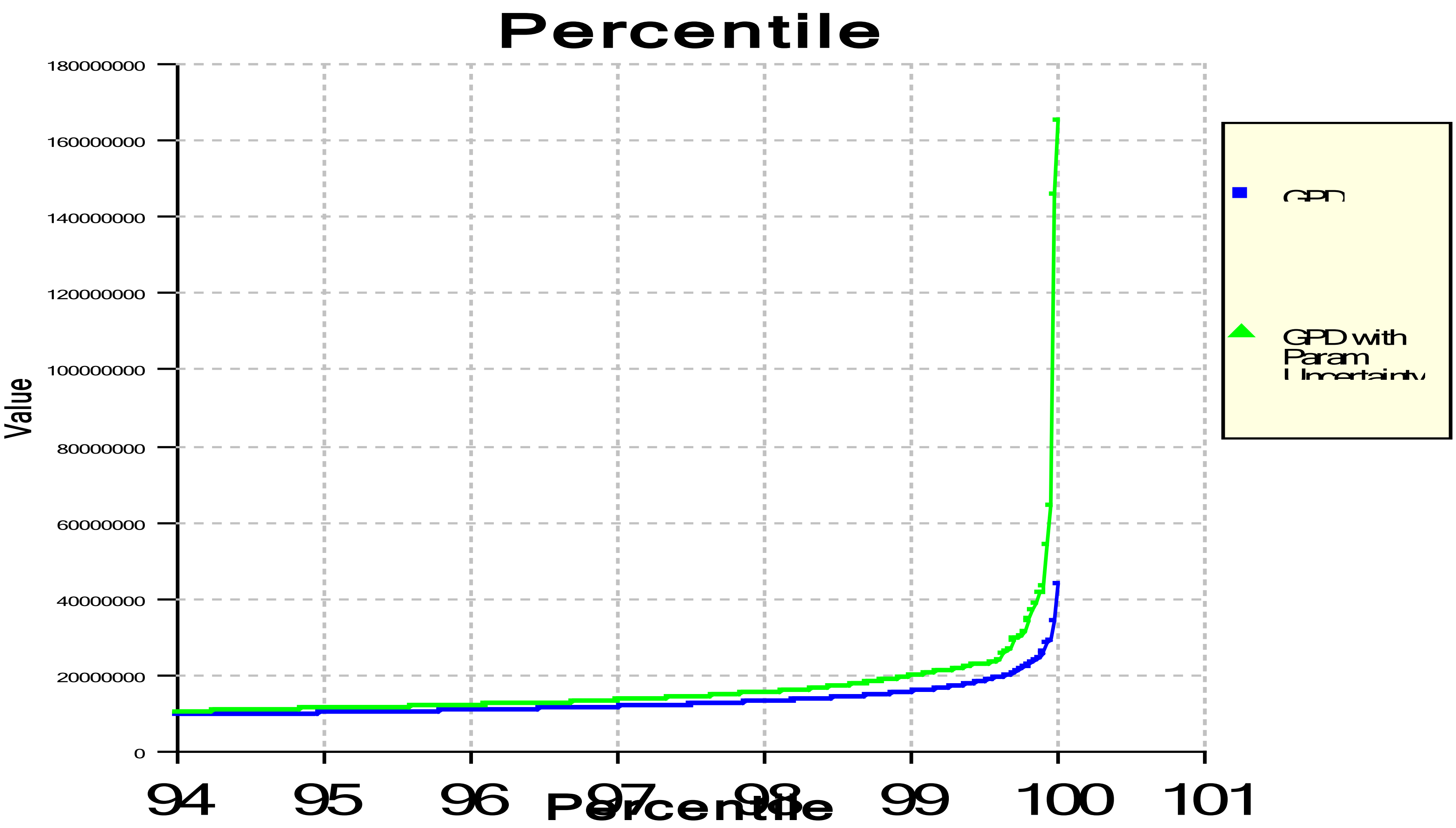
- Pick stable metrics
  - VaR can produce a lot of variability
  - Window VaR, TVaR
- Over simulate
  - Not always possible
  - Space, time, and memory issues
- Break simulations up into smaller run sizes with different seeds
  - Average results across runs
  - Not ideal but may be only choice

Parameter Uncertainty

- Should it be included
  - Working with limited data as is
  - Can have a major impact
- Considerations
  - Dependency in parameters
  - Is there model risk too?
    - Choice of distribution
    - Possibility of different extremes



Parameter Uncertainty



Parameter Uncertainty - Consistency

- What parts of your ERM or RI pricing models already include parameter uncertainty:
  - CAT models
  - Reserve risk
  - Is it appropriate to mix and match models?
- Should risk transfer analyses include?
  - How can you be 100% confident in 10 points or less in some cases?

Parameter Uncertainty - Practical Issues

- Have we over smoothed the data?
- Did our assumptions in our data adjustments have inherent volatility in them?
  - Can you feed other models to create stochastic adjustment factors before fitting data?
    - LDF's – reserve risk models
    - Inflation – ESG outputs
    - Other
- What if we don't want to use parameter uncertainty?
  - Smoothed data may be fine for mean selection
  - Should we use unsmoothed data in selecting volatility?

