


*WP1: Testing Loss Reserving Methods
and Assumptions -
Loss Simulation Model Working Party*



Presented by Robert A. Bear
Consulting Actuary and Arbitrator
RAB Actuarial Solutions, LLC
2008 Casualty Loss Reserve Seminar

Why develop a simulation tool?

- Mistakes are a part of being human. Appreciate your mistakes for what they are: precious life lessons that can only be learned the hard way. Unless it's a fatal mistake, which, at least, others can learn from.

- [Al Franken](#)

Loss Simulation Model Working Party

- Sponsored by DRMC in 2005, the LSMWP began work in 2006.
- Purpose: creation of a simulation model that will generate claims that can be summarized into loss development triangles and complete rectangles.
- Deliverables: Open source program available to CAS members, seminars, and a CAS Working Party paper documenting work.
- Time Frame: Complete testing and refinement of APL prototype in 2008. Develop and test alternate Visual Basic, R or J version in 2009.

Goals of LSMWP

- Goals: To generate triangles by layer, by different type of claim information (e.g., paid, incurred, Salvage and Subrogation, claim counts, etc.), by hazard, by line of business, etc.
- The working party will not be focusing on actual testing of reserving methods and models (including tail factor methods), but will focus on creating the simulated data sets for future research related to testing.

Goals of LSMWP (continued)

- Accordingly, a primary criterion for judging the quality of this model will be to evaluate the simulated data to make sure that it is realistic - i.e., it cannot be distinguished statistically from real data sets.
- Establish procedure to review and test modifications proposed by model users.

LSMWP Organization

- Co-Chairpersons: Bob Bear and Mark Shapland
- Group A: Literature & Test Criteria
 - Led by Curt Parker and Glenn Meyers.
 - Survey existing literature and prepare bibliography.
 - Develop testing criteria for determining “realism” of simulated data.
 - Can simulated data be statistically distinguished from actual data?
 - Ultimate test: “DRM Challenge”

LSMWP Organization (continued)

- Group B: Data, Parameters & Testing - led by Joe Marker.
 - Identify data sources and parameterize model.
 - Test model & data using criteria from Group A.
- Group C: Model Development - led by Dick Vaughan.
 - Evaluate modeling options and develop simulation model in at least two software environments. Open source for enhancements.
 - Refine and enhance model as a result of feedback from Group B.

LSMWP Organization (continued)

- Group C has chosen to model individual losses and transactions rather than aggregate triangles and statistics.
 - Don't need to choose in advance the triangle's time intervals.
 - Actuaries may use reserve estimators based on individual loss data rather than triangle data.
 - Aggregate simulation models are vulnerable to criticism that “Of course that model predicts that simulated data well, the simulated data is based on that model!”
 - A model of the loss process will be much easier to test against real data than a model of the triangles derived from the loss process.

Status Report - Groups A

- Group A has surveyed the literature on loss simulation and use of simulation to test reserving methods.
- It has developed an approach to test the “realism” of simulated triangles. This is a key test of the quality of the simulation model.
- The approach that is recommended is summarized in a 2007 ASTIN paper by Glenn Meyers entitled “Thinking Outside the Triangle” and in a 2006 CAS Fall Forum paper entitled “Estimating Predictive Distributions for Loss Reserve Models.”
- The Group A report “Literature and Test Criteria Subcommittee Report” is available on the CAS web site at www.casact.org/research/lsmwp.

Status Report - Groups B

- Group B has worked closely with a Data Source that has provided data for testing purposes.
- Ball State University students have done the necessary database work and have estimated parameters for the prototype model.
- The Group B “Parameterizing the Loss Simulation Model” (Ball State University Research Course) is completed and is also available on the CAS web site.
- Group B still needs to test Group C’s loss simulation model using the statistical test developed by Group A and to document these tests. Group B will also create the “DRM Challenge” whereby CAS members will be asked to distinguish actual from simulated loss triangles.

Status Report - Group C

- Group C has developed a prototype model in the APL programming language that has been reasonably tested.
- Both the run time and source code versions of the APL prototype are available on the CAS web site together with user instructions.
- Work on a Visual Basic, R or J version is being planned.
 - The interface and model features will differ from the APL version due to differences in software capabilities.
 - We expect that this version will be developed, tested and documented in 2009.
 - R is currently used by Group A in its statistical test and Group B in parameter estimation.

Basic Model Underlying Prototype

- *(1) Observation period:* Assume that relevant loss process involves accidents or occurrences between dates t_0 and t_1 . Simulator tracks transactions until accidents are settled.
- *(2) Time intervals:* Assume that parameters are constant throughout calendar months but may change from one month to next. Lags are measured in days.
- *(3) Exposures:* The user may specify a measure of exposure for each month. By default, the system assumes constant unit exposure. The purpose of the exposure parameter is to allow the user to account for a principal source of variation in monthly frequencies.

Basic Model Underlying Prototype

- (4) *Events*: Each claim may be described by the dates and amounts of the events it triggers: the accident date, the report date and an initial case reserve, zero or more subsequent valuation dates and case reserves changes, zero or one payment date and amount, and zero or one recovery date and amount.
- (5) *Distributions*: For most variables, the user may specify a distribution and associated parameters.
- (6) *Frequency*: Monthly claim frequency is assumed to have a Poisson distribution with mean proportional to earned exposure, or a Negative Binomial distribution with mean and variance proportional to earned exposure.

Basic Model (continued)

- (7) *Report lag*: The lag between occurrence and reporting is assumed to be distributed Exponential, Lognormal, Weibull, or Multinomial. The Multinomial distribution allows the user to define proportions of claims reporting within one month, two months, and so on.
- (8) *The lags between reporting and payment, between one valuation date and the next, and between payment and recovery or adjustment, are also assumed to be distributed Exponential, Lognormal, Weibull, or Multinomial.*
- (9) *Size of loss*: The actual size of the loss to the insured, independent of responsibility for payment, is distributed Lognormal, Pareto, or Weibull.

Basic Model (continued)

- (10) *Case reserve factor*: Case reserves are assumed to equal the actual size of loss, adjusted for the minimum, the maximum, the deductible, and the probability of closure without payment, all multiplied by an adequacy factor. This factor is assumed to be distributed Lognormal. The user may specify the mean factor at four points in time between the report and payment dates.
- (11) *Fast-track reserve*: A value may be assigned to each loss at first valuation, independent of regular case reserves and case reserve factor.

Basic Model (continued)

- (12) Initial payment factor: The initial payment of each loss not closed without payment is assumed to equal the actual size of loss, adjusted for the minimum, the maximum, the deductible, multiplied by a payment adequacy factor (PAF). The PAF determines the size of any subsequent adjustment or recovery.
- (13) Second-level distributions: The LSMWP models the drift in parameter values that may take place for many reasons but chiefly because of business turnover. It has developed an autoregressive model to reflect parameter drift.

Basic Model (continued)

- (14) Monthly vectors of parameters: For nearly all distributional parameters, the user may specify a single value or a vector of values.
- (15) *Frequency Trend and Seasonality*: The user may specify monthly trend and seasonality factors for frequency that are applied to means.

Basic Model (continued)

- *(16) Severity Trend:* The user may specify monthly trend factors for severity.
 - The “main” trend is allowed to operate up to the accident date and a fraction of this trend, defined by Butsic’s “alpha” parameter, is allowed to operate between accident and payment dates.
 - Case reserves before the adequacy factor are centered around the severity trended to the payment date.

Basic Model (continued)

- *(17) Lines and Loss Types:* The prototype model recognizes that loss data often involves a mixture of coverages and/or loss types with quite different frequencies, lags, and severities. Therefore, it allows the user to specify a two-level nested hierarchy of simulation specifications, with one or more “Lines” each containing one or more “Types”.
 - A typical Line might be “Auto,” typical Types within that Line might be “APD”, “AL-BI”, and “AL-PD.”
 - This hierarchy allows the user to set up any reasonable one or two level classification scheme.
 - Accident frequencies are modeled at the Line level and loss counts per accident are distributed among Types using a discrete distribution.

Basic Model (continued)

- *(18) Lines and Loss Types Example:* An Automobile occurrence might give rise to a single APD claim with probability 0.4, to a single AL-PD claim with probability 0.2, to a single APD and a single AL-PD claim with probability 0.2, to a single AL-BI claim with probability 0.1, to two AL-BI claims with probability 0.05, etc.
- *(19) Correlations:* The prototype model makes it possible to request correlated samples of certain variables without fully specifying their joint distribution. For the moment these variables are (a) the mean frequencies across Lines and (b) the size of loss and report lag within a Type.

Basic Model (continued)

- (20) *Clustering*: The prototype simulator allows a selectable fraction of loss sizes and a selectable fraction of case reserves to be rounded to two significant digits, imitating clustering around round numbers frequently observed.
- (21) *Output*: The prototype simulator produces output as tab-delimited text files or by launching an instance of Excel and populating it with worksheets. In both cases, the possible output tables include claim and transaction files (together displaying the complete loss history), all the usual triangles, a table of large losses, a summary of the simulation specifications, and a summary of the frequency derivation by month.

Summary



- The LSMWP has made considerable progress in developing a model that we hope will become a valuable tool in researching reserving methods and models. Stay tuned!
- We hope that actuaries will use this model to:
 - Better understand the underlying loss process.
 - Determine what methods and models work best in different reserving situations.