Half-Mack Stochastic Reserving

Frank Cuypers, Simone Dalessi

March 2013

Abstract

We suggest a stochastic reserving method, which uses the information gained from statistical reserving methods (such as the Mack Chain Ladder procedure), to fit the inverse cumulative loss development factors with loss saturation functions. These loss saturation functions straightforwardly yield the reserves and include an implicit tail factor.

The choice of the proper type of loss saturation function allows for actuarial or underwriting judgment based on appropriate experience with the analyzed line of business and jurisdiction. Moreover, the confidence intervals of the loss development factors determine the probability distribution of the reserves, which can be used *e.g.* for the purpose of modeling risk, capital and solvency.

1 Introduction

Since Thomas Mack published his seminal article on the standard error of the Chain Ladder reserves [1] a large portion of the actuarial reserving literature has gone beyond the mere calculation of best estimates and has further explored the reserves' variability. Some of the proposed methods are of statistical nature and estimate a finite number of moments of the probability distribution of the reserves. Mack's method [1] falls into this category. Other methods are of stochastic nature and generate a full probability distribution of the reserves. Prominent representatives of this approach are *e.g.* Chain Ladder bootstraps [2] and Markov chain Monte Carlos [3].

We suggest here a stochastic method, which makes only partial use of the information statistical methods yield, but stops short of using them for calculating the variance of the reserves. For this reason and in homage to the pioneer of the field, we call it *Half-Mack* Stochastic Reserving.

In the following Section we quickly summarize those outputs of the Mack procedure, which are useful in the context of Half-Mack Stochastic Reserving. Next we use these results to define the loss saturation function and describe how to generate Half-Mack reserve distributions. Finally we apply the Half-Mack Stochastic Reserving procedure to real data and compare its results with those of other statistical and stochastic methods.

2 Mack Chain Ladder

The Mack Chain Ladder procedure [1] is an example of a statistical reserving method: based on fairly simple and general assumptions, it estimates analytically the mean and the variance of the probability distribution of the reserves.

For the purpose of the Half-Mack Stochastic Reserving method, however, we are not interested in the moments of the reserves, but rather in those of the cumulative loss development factors (age-to-ultimate) $a2u$ at development year $t = 1 \ldots l$:

$$\mathbb{E}\left[a2u_t\right] = \prod_{k=t}^{l-1} \mathbb{E}\left[a2a_k\right] \tag{1}$$

$$\mathbb{V}\left[a2u_t\right] = \mathbb{E}\left[a2u_t\right]^2 \sum_{k=t}^{l-1} \frac{\mathbb{V}\left[a2a_k\right]}{\mathbb{E}\left[a2u_k\right]^2} \left(\frac{1}{L_k^{l-t}} + \frac{1}{\sum_{a=1}^{l-k-1} L_k^a} \right) \tag{2}$$

where L_t^a stands for the aggregate losses of accident year a and development year t, and the mean and variance of the incremental loss development factors (age-to-age) $a2a$ are given by

$$\mathbb{E}\left[a2a_t\right] = \frac{\sum_{a=1}^{l-t} L_{t+1}^a}{\sum_{a=1}^{l-t} L_t^a} \tag{3}$$

$$\mathbb{V}\left[a2a_t\right] = \frac{1}{l-t} \sum_{a=1}^{l-t} L_t^a \left(\frac{L_{t+1}^a}{L_t^a} - \mathbb{E}\left[a2a_t\right] \right) \tag{4}$$

Although for definiteness we focus here on the Mack Chain Ladder method, we could as well use other statistical reserving methods, like those based on generalized linear models [3]. Nevertheless, because of its widespread use and thanks to its simple analytic formulas, Mack's Chain Ladder is probably the likeliest candidate for this exercise.

3 Half-Mack Stochastic Reserving

We describe here a stochastic reserving method, which uses the confidence intervals of the cumulative loss development factors (1,2), but stops short of computing the Mack Chain Ladder reserves. However, it uses the information about the variance of these loss development factors to perform a stochastic fit of a loss saturation function, which in turn determines the full distribution of the reserves.

3.1 The Loss Saturation Function

When plotting the inverse cumulative loss development factors (1) of an aggregate loss triangle as a function of time (typically the development years), we obtain a series of points that tends to saturate towards an horizontal asymptote at 100%. We call a curve that reproduces this time dependence of the inverse cumulative loss development factors a *loss saturation function*.

Such loss saturation functions $F(t)$ describe in continuous time t how aggregate the losses (paid or incurred) L_t^a of any accident year a tend to develop over the discrete development years t and smoothly converge towards their ultimates:

$$L_\infty^a = \frac{L_t^a}{F(t)} \tag{5}$$

Most lines of business display a characteristic loss saturation function, whose shape can vary from one jurisdiction to another. In Figure 1 we display some typical generic behaviors of loss saturation functions. In general an experienced actuary or underwriter knows how the losses (paid or incurred) of a particular business tend to develop, and she will accordingly choose the appropriate family of loss saturation functions that should best reproduce the behavior of the inverse cumulative loss development factors.

If l is the number of observed accident years and development years of the observed aggregate loss triangle, we develop the last observed (diagonal) losses L_{l-a+1}^a of accident years $a = 1 \ldots l$ according to the value of the continuous loss saturation function $F(l - a + 1)$ at that particular point in time, and sum over all accident years to obtain the reserves

$$\begin{aligned} R &= \sum_{a=1}^{l} \left(L_\infty^a - L_{l-a+1}^a \right) \\ &= \sum_{a=1}^{l} L_{l-a+1}^a \frac{1 - F(l-a+1)}{F(l-a+1)} \end{aligned} \tag{6}$$

For the sake of definiteness and without claiming to be exhaustive, we focus in this paper on the following exponential family of loss saturation functions [4]:

$$F_{exp}(t) = \left(1 - \exp\left(-\frac{t-\tau}{\lambda}\right)\right)^\alpha \tag{7}$$

Prime Re Services AG
Lindenstrasse 2
CH-6340 Zug/Baar
+41 (41) 725 32 10
www.prs-zug.com

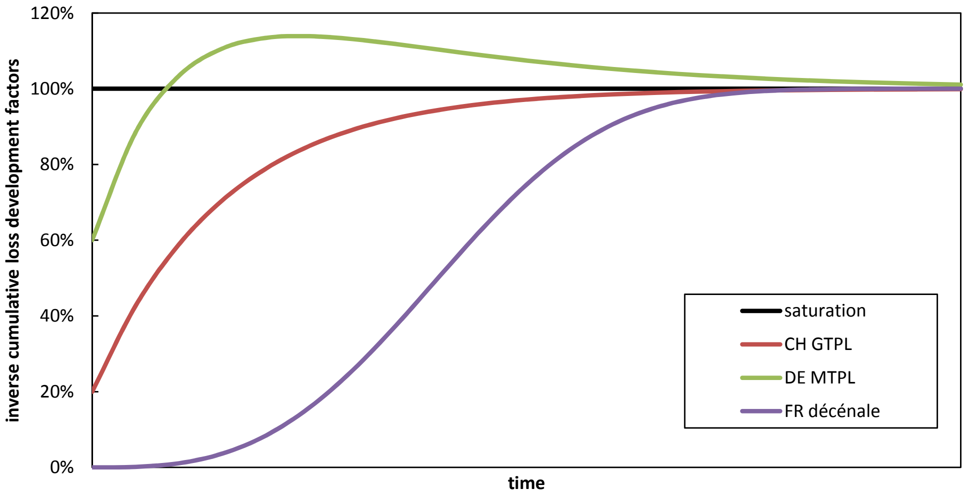


Figure 1: Generic patterns of loss saturation functions. The blue curve is typical for French Décénale business (construction insurance) and reflects the long reporting delay. The green curve is typical for German Motor business and reflects the market's tendency to over-reserve. The red curve is typical for Swiss Liability business. By definition a loss saturation function saturates on an asymptote at 100%.

where t is the continuous time of the development years, and τ, λ and α are respectively the location, scale and shape parameters.

Such loss saturation functions (7) can reproduce the behaviors of the French Décénale and Swiss Liability businesses depicted in Figure 1. However, they cannot reproduce over-reserving patterns like the one observed in the German Motor market.

If one were only interested in estimating the best estimate reserves it would now suffice to fit the parameters of a loss saturation function to observed data. This could *e.g.* be performed with a χ^2 fit, which assigns more weight to observations that come with a smaller confidence interval, and *vice versa*. Statistical reserving methods, like the Mack Chain Ladder procedure delineated in Section 2, yield just these confidence intervals for the cumulative loss development factors.

However, we prefer to go beyond the best estimate and probe the entire shape of the reserves' distribution, which we do in the next Section.

3.2 Sampling the Loss Saturation Function

One can determine the shape of the reserves by sampling the loss saturation functions stochastically, according to the probability distributions of the inverse cumulative loss development factors. Although these distributions are *a priori* unknown, their first moments can be estimated with a statistical reserving method, such as the the Mack Chain Ladder procedure described in Section 2. In this case their means and variances are given by Equations (1,2).

For the sake of definiteness we shall assume in Section 4 that the inverse cumulative loss development factors are distributed according to a normal distribution. This allows for a straightforward connection to the χ^2 procedure mentioned in Section 3.1. A possible alternative is to assume a lognormal behavior, which guarantees the generation of solely positive definite loss development factors.

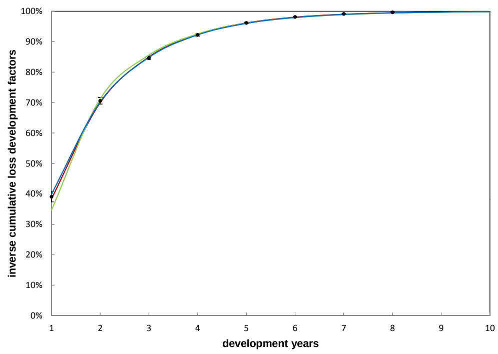
Having chosen for each inverse cumulative loss development factor a probability distribution whose moments match those given by the statistical reserving method, we bootstrap the reserves by sampling these distributions. Each realization generates a different sequence of loss development factors, to which we adjust (*e.g.* with an Ordinary Least Squares fit) a loss saturation function, which in turn uniquely determines the reserves associated with this realization (6). In this way a sufficiently large sample of independent realizations generates a full distribution of the reserves.

Anticipating on the results of Section 4, where we apply the half-Mack procedure to real data and compare it with other methods, we display in Figures 2 how 3 (out of 1'000) exponential loss saturation functions (7), osculate the data within their confidence intervals. The variability can be well observed in Figure 2f, where the confidence intervals are particularly large.

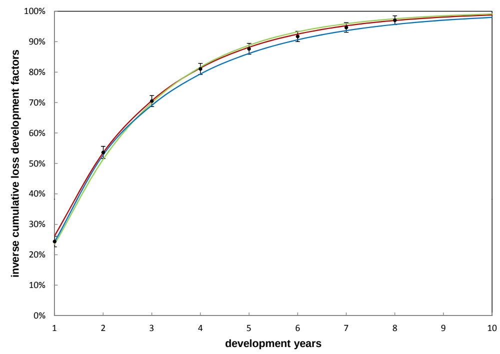
4 Application to Real Data

To test the practicability of the Half-Mack Stochastic Reserving method, we have applied it to the US long tail loss data published by the CAS [5]. This repository lists (among other things) for different lines of business the aggregate 10×10 loss squares of a large number of carriers for the accident years 1988 to 1997. We list these lines of business in Table 1.

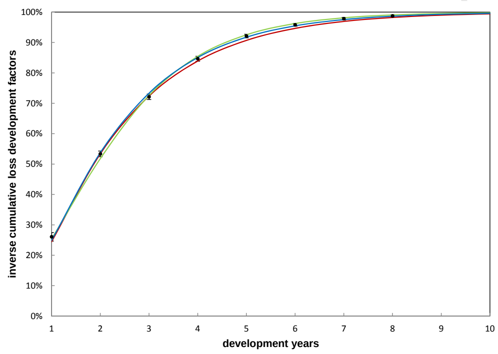
For the purpose of this study we have aggregated for each of the 6 lines of business listed in Table 1 the paid losses of all carriers, and we have examined how different stochastic methods develop the resulting upper left loss triangles to ultimate and generate a distribution of the reserves:



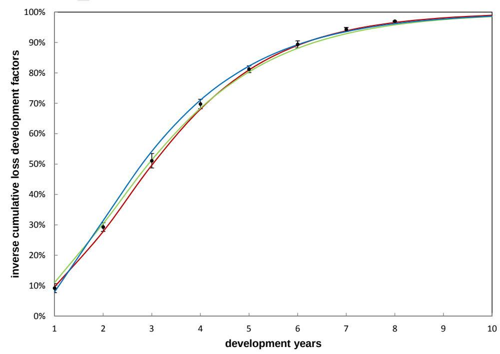
(a) Personal Auto



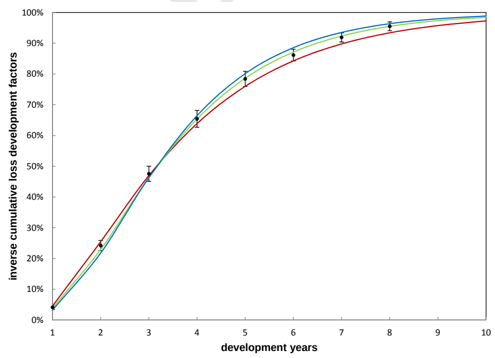
(b) Workers Compensation



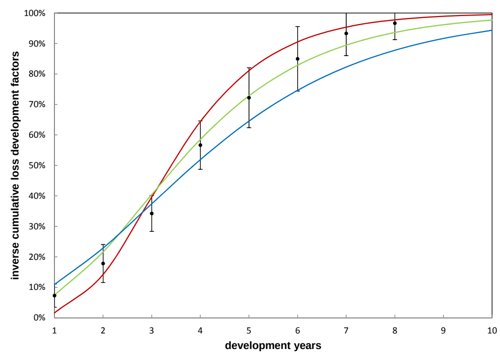
(c) Commercial Auto



(d) General Liability



(e) Medical Malpractice



(f) Product Liability

Figure 2: Inverse cumulative loss development factors and their confidence intervals (1,2). The continuous colored curves depict 3 of the 1'000 sampled loss saturation functions (7).

line of business	**volume**	χ^2**/dof**	τ	λ	α
personal auto	119'441'598	0.73	0.17	1.41	1.15
workers compensation	13'445'530	0.03	0.56	2.35	0.80
commercial auto	8'029'873	0.21	-0.91	1.53	3.93
general liability	4'889'532	0.30	-3.03	1.64	26.34
medical malpractice	3'242'744	0.04	0.35	1.95	2.53
product liability	1'360'268	0.23	-5.50	1.79	113.73

Table 1: US long tail lines of business [5] and their best fits to the loss saturation functions (7). The volumes (reported in thousand USD) are defined as the sum over all 10 observed accident years (from 1988 to 1997) of the paid losses observed in the 10^{th} development year, *i.e.*, the last column of the 10×10 aggregate loss square. The goodness of fit is indicated by the χ^2 statistic divided by the number of degrees of freedom $(8 - 3 = 5)$: a value exceeding 1.5 would indicate a poor fit.

Mack Chain Ladder: We assume the reserves are distributed lognormally with the Mack Chain Ladder moments [1].

Chain Ladder bootstrap: We bootstrap the Chain Ladder procedure [2] with 1'000 random samples.

Half-Mack Stochastic Reserving: We sample 1'000 random realizations of the loss saturation functions (7). In Figures 2 we display for each of the 6 lines of business 3 typical sampled loss saturation functions. Each realization generates a different reserve (6) and the full sample generates a distribution of the reserves.

In Table 1 we verify by means of a χ^2 fit that the choice of the exponential loss saturation function (7) is justified for each of the 6 lines of business.

In Figures 3 – 8 we display for each of the 6 lines of business listed in Table 1 how the cumulative probability distributions of the reserves generated by each of these methods compare with each other and the realized reserves. The latter are defined as the difference between the last column and the diagonal of the 10×10 aggregate loss square.

4.1 Personal Auto

We observe in Figure 2a that the large amount of homogeneous claims of this line of business yields extremely small confidence intervals on the data. Still, the exponential loss saturation function (7) fits these inverse loss development factors very well.

In Figure 3 we observe that all 3 considered reserving methods yield very similar reserve distributions. These distributions have a very moderate variance, which reflects the small size of the Mack Chain Ladder errors (2). This comes as no surprise, as this line of business is by nature quite homogeneous and the large business volume leaves little room for significant statistical fluctuations. However, all 3 methods predict their reserves to exceed with 99% probability the realized reserves. We have no explanation for this phenomenon.

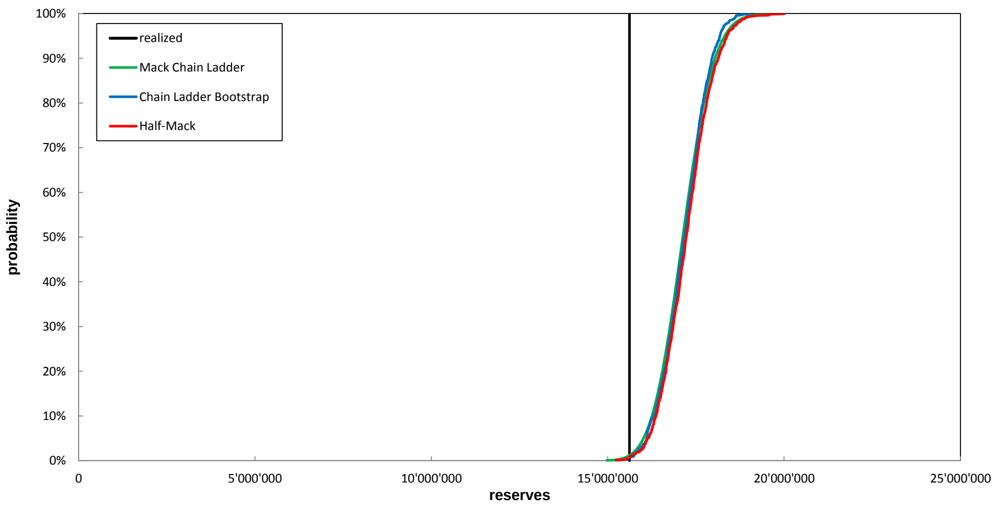


Figure 3: Reserve distributions for Personal Auto.

4.2 Workers Compensation

We observe in Figure 2b that the confidence intervals of this line of business are quite small, but the exponential loss saturation function (7) fits the inverse loss development factors very well.

In Figure 4 we observe that all 3 considered reserving methods yield similar reserve distributions with moderate variances. All 3 methods predict their reserves to exceed with high probability the realized reserves. We have no explanation for this phenomenon.

4.3 Commercial Auto

We observe in Figure 2c that the confidence intervals of this line of business are quite small, but the exponential loss saturation function (7) fits the inverse loss development factors very well.

In Figure 5 we observe that all 3 considered reserving methods yield similar reserve distributions with moderate variances. All 3 methods predict their reserves to exceed with high probability the realized reserves. We have no explanation for this phenomenon.

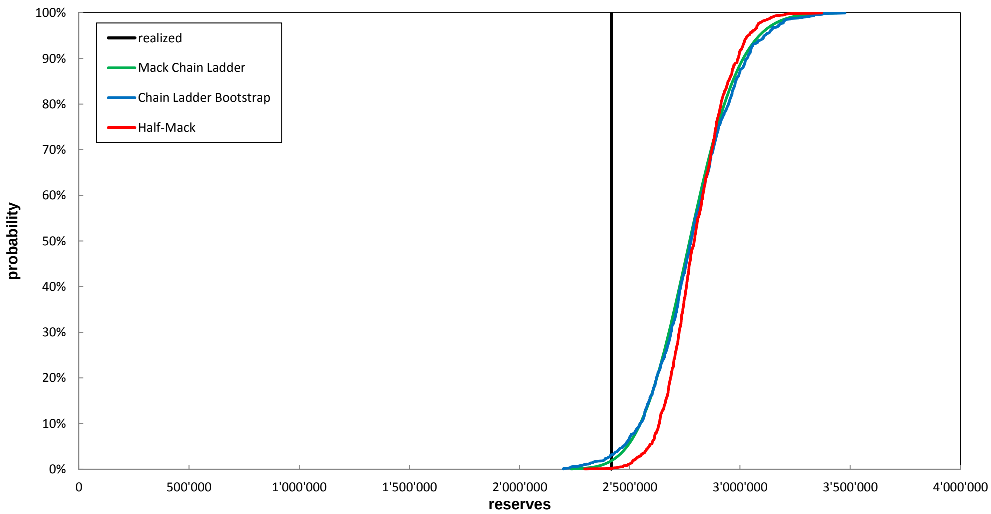


Figure 4: Reserve distributions for Workers Compensation.

4.4 General Liability

We observe in Figure 2d that the confidence intervals of this line of business are quite small, but the exponential loss saturation function (7) fits the inverse loss development factors very well.

In Figure 6 we observe that all 3 considered reserving methods yield similar reserve distributions with moderate variances. Here, however, all 3 methods yield a best estimate reserve, which matches the realized reserves well.

4.5 Medical Malpractice

We observe in Figure 2e that the confidence intervals of this line of business are quite small, but the exponential loss saturation function (7) fits the inverse loss development factors very well.

In Figure 7 we observe that all 3 considered reserving methods yield similar reserve distributions with moderate variances. All 3 methods predict their reserves to exceed with high probability the realized reserves. We have no explanation for this phenomenon.

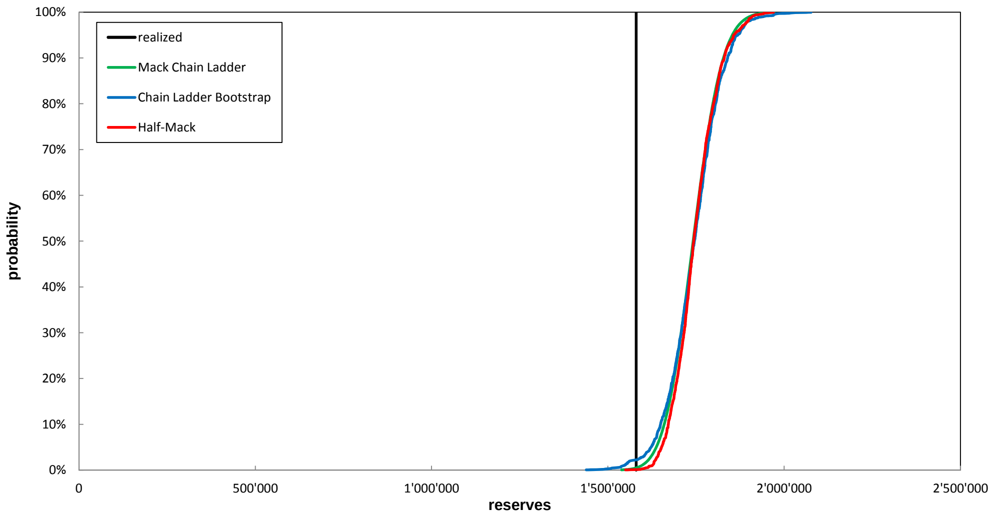


Figure 5: Reserve distributions for Commercial Auto.

4.6 Product Liability

We observe in Figure 2f that the confidence intervals of this line of business are significantly larger than for the other lines of business. It comes therefore as no surprise that the exponential loss saturation function (7) fits these inverse loss development factors very well.

In Figure 8 we observe that the Mack Chain Ladder and the Chain Ladder bootstrap yield similar reserve distributions, whose best estimate reserves match the realized reserves well. As expected from the large confidence intervals of the loss development factors, these distributions display relatively large variances.

In contrast, here Half-Mack Stochastic Reserving fails to generate a reasonable distribution of the reserves. Indeed, because the first inverse loss development factor has such a large confidence interval, the method samples a significant amount of exponential loss saturation functions (7) with very low initial values. These in turn generate very large reserves, which induce an absurdly fat distribution tail.

In this case the choice of an unconstrained exponential loss saturation function (7) clearly does not correctly reflect the natural behavior of the initial inverse loss development factors. Either the family (7) should be constrained such as to avoid too small values at the origin, or another family of loss saturation functions should be used.

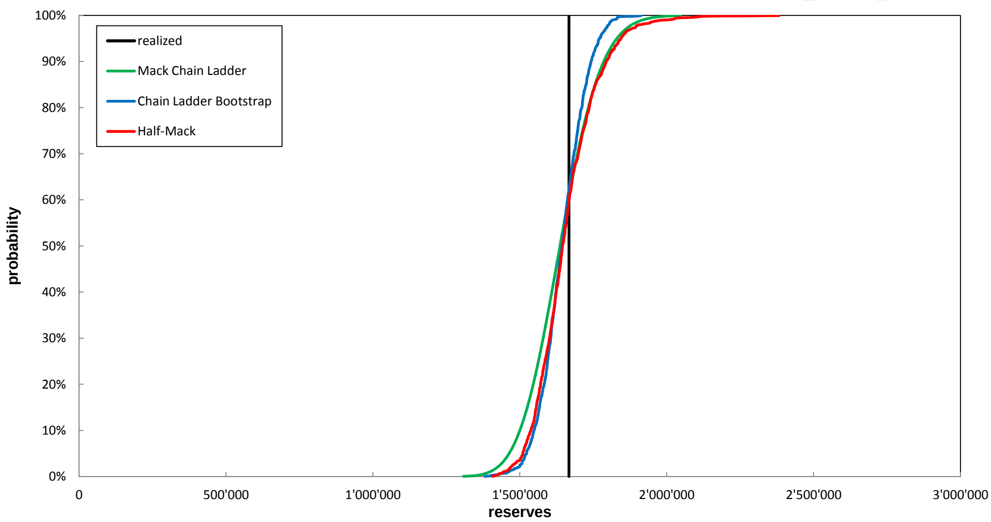


Figure 6: Reserve distributions for General Liability.

5 Conclusions

Half-Mack Stochastic Reserving is a straightforward technique to estimate the full probability distribution of loss reserves. It has the following advantages:

- It builds upon any standard deterministic aggregate multiplicative reserving technique. (Here we have focused on the Chain Ladder method.)
- It accounts for the statistical and systematic errors associated with the chosen reserving technique (Here we have focused on the Mack Chain Ladder confidence intervals [1]). and yields a full probability distribution of the reserves.
- It automatically smooths the loss development factors in a natural fashion. (Here we have focused on an exponential loss saturation function (7).)
- It automatically incorporates tail factors in a natural fashion. (Here we have focused on an exponential loss saturation function (7).)
- It allows for considering objective market circumstances. (Here we have focused on an exponential loss saturation function (7).)

However, the latest item may actually turn into a disadvantage of Half-Mack Stochastic Reserving, because there are situations where the method requires particularly sound actuarial or underwriting

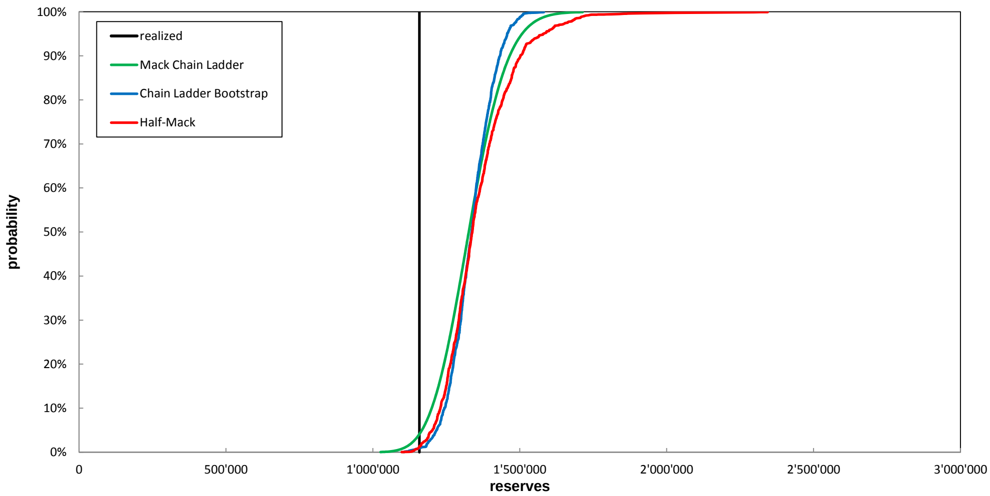


Figure 7: Reserve distributions for Medical Malpractice.

judgment based on a solid knowledge of the considered line of business and applicable jurisdiction. For instance, we observed for Product Liability in Section 4.6 that the exponential loss saturation function (7) does not suitably reproduce the behavior of the claims payments, in spite of yielding a perfectly acceptable fit to the data.

Regarding the data we confirm a phenomenon already observed by Meyers and Shi [6]: using paid data, the predictive distributions obtained from several different stochastic reserving methods fail to reproduce the eventually realized outcomes. In contrast, the reserves actually posted by the carriers in their original statements turn out to provide a much better prediction. To cite Glenn Meyers [7]:

> One has to wonder what the insurers saw that we did not see in the data.

6 Acknowledgments

We are very thankful to Glenn Meyers for pointing out to us the US long tail loss data repository [5].

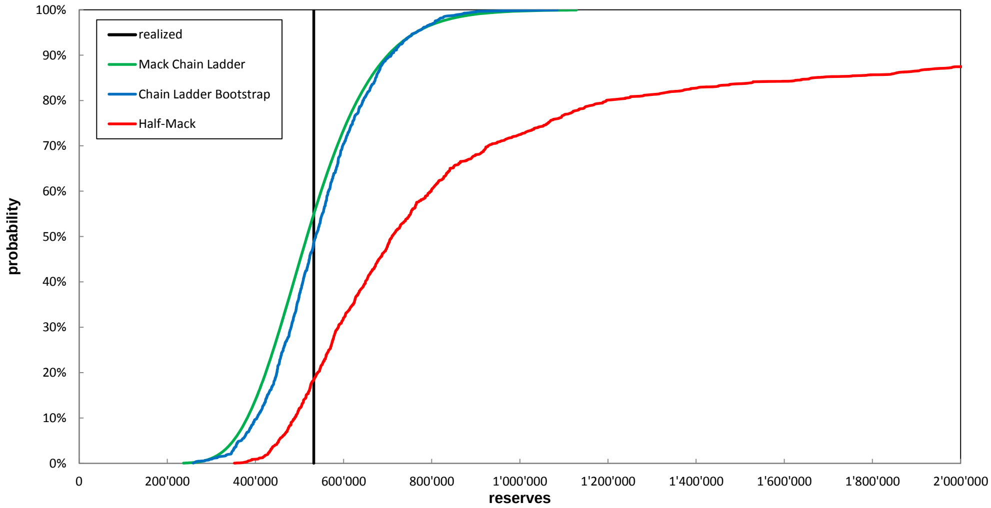


Figure 8: Reserve distributions for Product Liability.

References

[1] Thomas Mack.
Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Estimates.
ASTIN Bulletin 23:2 213–225, 1993.

[2] Peter England and Robert Verral.
Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving.
Insurance: Mathematics and Economics 25:3 281–293, 1999.

[3] Mario V. Wüthrich and Michael Merz.
Stochastic Claims Reserving Methods in Insurance.
Wiley Finance, 2008.

[4] Rameshwar D. Gupta and Debasis Kundu.
Generalized Exponential Distributions.
Austral. & New Zealand J. Statist. 41:2 173–188, 1999.

[5] Glenn G. Meyers and Peng Shi.
Loss Reserving Data pulled from NAIC schedule P.
`http://www.casact.org/research/index.cfm?fa=loss_reserves_data`.

[6] Glenn Meyers and Peng Shi.

The Retrospective Testing of Stochastic Loss Reserve Models.
Casualty Actuarial Society E-Forum, Summer 2011.

[7] Glenn Meyers.
The Leveled Chain Ladder Model for Stochastic Loss Reserving.
Casualty Actuarial Society E-Forum, Summer 2012.