Robustness in Reserving  
B. Avanzi, M. Lavender, G. Taylor, B. Wong  

Key words: Reserving, Chain Ladder, Mack model, Outliers, Robustness, Influence functions

Purpose of your paper:

We consider robust (relatively immune to outliers) reserving. In particular, we focus on the Mack model, and extend existing statistical techniques to allow for typical properties of insurance data, such as skewness.

Synopsis:

Often, in actuarial applications, data sets display values that are potential outliers. Reserving is no exception (see, e.g., Verdonck et al., 2009 or van Wouwe et al., 2009, Verdonck and van Wouwe, 2011). It is sometimes impossible to determine whether those data are real outliers or mistakes. A robust method is one that balances the two extreme approaches of not taking the datum into account, or conversely, taking it into account as a normal datum. In other words, robustness refers to the ability of a model or estimation procedure to not be overtly influenced by outliers in the dataset under investigation and/or deviations from the underlying assumptions of the model.

Robust statistical techniques have existed for a long time, but recent developments encouraged us to revisit and extend the (relatively scarce) literature on robust reserving. To somewhat formalise what a robust procedure should strive for we refer to the three fundamental features outlined by Huber and Ronchetti (2009)

1. Efficiency: Optimal or nearly optimal efficiency under the assumptions of the chosen model.
2. Stability: Small divergences from the model should only have a minor effect on its performance.
3. Breakdown: Moderately greater divergences should not lead to a disaster.

The first point means that to achieve an effective robust procedure we should not detract significantly from efficiency of the assumed model. The following two points go to the heart of the robustness issue and explain how the procedure should behave in the presence of outliers or deviations. It is this intuition that will be carried through this paper as we explore the robustness of classical reserving models and propose robust counterparts.

The chain ladder technique is highly susceptible to outliers due to the well known non-robustness of the mean (Verdonck et al., 2009). Verdonck and Debruyne (2011) highlight the vulnerability of these approaches by calculating the influence functions of the parameters, future claim estimates and reserves with respect to incremental claims under the Poisson generalised linear model (GLM) specification of the chain ladder approach. Importantly, these influence functions are shown to be unbounded. An unbounded influence function means that a small contamination of a single data point can have an arbitrarily large effect on the estimate. This result comes from the use of the maximum likelihood (and quasi-likelihood) estimation procedure of which the susceptibility to outliers has been shown by numerous authors (see for example Pregibon, 1982, Künsch et al., 1989). A robust procedure should have a bounded influence function. Similar results regarding robustness of the chain-ladder techniques are found in Venter and Tampubolon (2010), who explore the impact of incremental claims on final reserve estimates under a range of models and also look at the sensitivity of each fitted value to the corresponding initial observed value.

Importantly, if applied blindly, reserving models may lead to severely inaccurate reserve estimates and in turn final conclusions. Hence the extent to which the models employed to forecast future reserves are robust deserves considerable attention. We will initially give such attention to exploring the robustness of some classical reserving techniques.

We will then evaluate some robust multivariate loss reserving techniques and address some identified shortcomings. The issue of robustness in multivariate reserving has been somewhat considered in the literature. Verdonck and van Wouwe (2011) have put forward a robust bivariate chain-ladder technique that employs two techniques to detect and adjust outliers. One of these techniques is based
on the minimum covariance determinant (MCD, see Rousseeuw, 1984) robust estimation of the location and scale of a data set. From here a robust distance measure (Mahalanobis distance) of each data point is calculated and those beyond a certain threshold are considered outliers. This approach relies on the assumption of elliptical symmetry of the underlying data which can be problematic when claims observations are skewed as is often the case in reality. The other technique is based on the bagplot (Rousseeuw et al, 1999) and is purely graphical. Furthermore, both these techniques rely heavily on the arbitrary selection of tuning parameters. We will discuss these issues.

The theoretical results and practical techniques presented in this paper will facilitate an understanding of the robustness of reserving techniques as well as provide viable methods to handle features of loss data that are known to occur in reality. Results will be illustrated with simulated and real data.

References: