Robustifying Reserving

Gary G. Venter, FCAS, MAAA, and Dumaria R. Tampubolon, Ph.D.

Abstract

Robust statistical procedures have a growing body of literature and in actuarial applications have been applied in loss severity fitting. Here an introduction of robust methods is made for loss reserving. In particular, following Tampubolon [1], reserve models for a development triangle are compared based on the sensitivity of the reserve estimates to changes in individual data points. This is then related to the generalized degrees of freedom used by the model at each point.

Keywords. Loss reserving; regression modeling; robust, generalized degrees of freedom.

All models are wrong, but some are useful.
Christian Dior (or maybe George E. P. Box)

0 INTRODUCTION

The idea of this paper is simple: for models of a loss development triangle, look at the derivative of the loss reserve with respect to each data point. All else being equal, models that are highly sensitive to a few particular observations are less preferred than ones that are not. This is supported by the fact that individual cells can be highly unstable. This general approach, based on Tampubolon [1], is along the lines of robust statistics, so some background into robust statistics will be the starting point. Published models on three data sets will be tested by this methodology. For two of them, unsuspected problems with the previously best-fitting models are found, leading to improved models.

The sensitivity of the reserve estimate to individual points is related to the power of those points to pull the fitted model towards them. This can be measured by what Ye [2] calls generalized degrees of freedom (GDF). For a model and fitting procedure, the GDF at each point is defined as the derivative of the fitted point with respect to the observed point. If any change in a sample point is matched by the same change in the fitted, the model and fitting procedure are giving that point full control over its fit, so a full degree of freedom is used. GDF does not fully explain the sensitivity of the reserve to a point, as the position of the point in the triangle also gives it more or less power to change the reserve estimate, but it adds some insight into that.

Section 1 provides some background into robust analysis and section 2 shows some previous

applications to actuarial problems. These help to place the current proposal into perspective in that literature. Sections 3, 4, and 5 apply this approach to some published development models. Section 6 concludes.

1 ROBUST METHODS IN GENERAL

Classical statistics takes a model structure and tries to optimize the fit of data to the model under the assumption that the data is in fact generated by the process postulated in the model. But in many applied situations, the model is a convenient simplification of a more complex process. In this case the optimality of estimation methods like maximum likelihood (MLE) may no longer hold. In fact a few observations that do not arise from the model assumptions can sometimes significantly distort the estimated parameters when standard techniques are used. For instance, Tukey [3] gives examples where even small deviations from the assumed model can greatly reduce the optimality properties. Robust statistics looks for estimation methods that in one way or another can insulate the estimates from such distortions.

Perhaps the simplest such procedure is to identify and exclude outliers. Sometimes outliers clearly arise from some other process than the model being estimated, and it may even be clear when current conditions are likely to generate such outliers, so that the model can then be adjusted. If the parameter estimates are strongly influenced by such outliers, and the majority of the observations are not consistent with those estimates, it is reasonable to exclude the outliers and just be cautious about when to use the model.

An example is provided by models of the US one-month Treasury bill rates at monthly intervals. Typical models postulate that the volatility of the rate is higher when the rate itself is higher. Often the volatility is proposed to be proportional to the pth power of the rate. The question is – what is p? One model, the CIR or Cox, Ingersoll, Ross model, takes $p = \frac{1}{2}$. Other models postulate p as 1 or even 1.5, and others try to estimate p as a parameter. An analysis by Dell'Aquila et al. [4] found that when using traditional methods, the estimate of p is very sensitive to a few observations in the 1979-82 period, when the US Federal Reserve bank was experimenting with monetary policy. Including that period in the data, models with $p=1.5$ cannot be rejected, but excluding that period finds that $p = \frac{1}{2}$ works just fine. That period also experienced very high values of the interest rate itself, so their analysis suggests that using $p = \frac{1}{2}$ unless the interest rate is unusually high would make sense.

A key tool in robust statistics is the identification of influential observations, using the influence

function defined by Hampel [5]. This looks at statistics calculated from a sample, such as estimated parameters, as functionals of the random variables that are sampled. The influence function for the statistic at any observation is a functional derivative of the statistic with respect to the observed point. In practice, analysts often use what is called the empirical influence. For instance, Bilodeau [6] suggests calculating that at each sample point as the sample size times the decrease (which may be negative) in the statistic from excluding the point from the sample. That is, the influence is *n* times [statistic with full sample minus statistic excluding the point]. If the statistic is particularly sensitive to a single or a few observations, that calls its accuracy into question. The gross error sensitivity (GES) is defined as the maximum absolute value of the influence function across the sample.

The effect on the statistic of small changes in the influential observations is also a part of robust analysis, as these effects should not be too large either. If each observation has a substantial randomness to it, the random component of influential observations would be having a disproportionate impact on the statistic. The approach used below in the loss reserving case is to identify observations for which small changes have large impacts on the reserve estimate.

Exclusion is not the only option for dealing with outliers. Estimation procedures that use but limit the influence of the outliers are also an important element of robust statistics. Also finding alternative models that are not dominated by a few influential points and estimating them by traditional means can be an outcome of a robust analysis. In the interest rate case, a model with one *p* parameter for October 1979 through September 1982 and another elsewhere does this. Finding alternative models with less influence from a few points is what we will be attempting in the reserve analysis.

2 ROBUST METHODS IN INSURANCE

Several papers on applying robust analysis to fitting loss severity distributions have appeared in recent years. For instance, Brazauskas and Serfling [7] focus on estimation of the simple Pareto tail parameter α assuming that the scale parameter *b* is known. In this notation the survival function is $S(x) = (b/x)^{\alpha}$. They compare several estimators of α, such as MLE, matching moments or percentiles, etc. One of their tests is the asymptotic relative efficiency (ARE) of the estimate compared to MLE, which is the factor which when applied to the sample size would give the sample size needed for MLE to give the same asymptotic estimation error. Due to the asymptotic efficiency of MLE, these factors are never greater than unity, assuming the sample is really from that Pareto

distribution.

The problem is, however, that the sample might not be simple Pareto. Even then, however, you would not want to identify and eliminate outliers: whatever process is generating the losses would be expected to continue, so no losses can be ignored.[1] The usual approach to this problem is thus to find alternative estimators that have low values of the GES and high values of ARE. Brazauskas and Serfling [7] suggest estimators they call generalized medians (GM). The kth generalized median is the median of all MLE estimators of subsets of size k of the original data. That can be fairly calculation-intensive, however, even with $k = 3$, 4, or 5.

Finkelstein et al. [8] define an estimator they call the probability integral transform statistic (PITS) which is quite a bit easier to calculate but not quite as robust as the GM. It has a tuning parameter t in (0,1) to control the trade-off between efficiency and robustness. Since $(b/x)^\alpha$ is a probability, it should be distributed uniform [0,1]. Thus $(b/x)^{t\alpha}$ should be distributed like a uniform raised to the t power. The average of these over a sample is known to have expected value $1/(t+1)$, so the PITS estimator is the value of β for which the average of $(b/x)^{t\beta}$ over the sample is $1/(t+1)$. This is a single-variable root-finding exercise. Finklestein et al. give values of the ARE and GES for the GM and PITS estimators, shown in Table 1. A simulation suggests that the GES for MLE for $\alpha = 1$ is about 3.9, and since its ARE is 1.0 by definition, PITS at 0.94 ARE is not worthwhile in this context. In general the generalized median estimators are more robust by this measure.

Other robust severity studies include Brazauskas and Serfling [9] who use GM estimation for both parameters of the simple Pareto, Gather, and Schultze [10] who show that the best GES for the exponential is the median scaled to be unbiased, but this has low ARE, and Serfling [11] who applies GM to the lognormal distribution.

Table 1: Comparative efficiency and robustness of two robust estimators of Pareto α

ARE	GM-k	PITS-t	GM-GES	PITS-GES
0.88	3	0.531	2.27α	2.88α
0.92	4	0.394	2.60α	3.54α
0.94	5	0.324	2.88α	4.08α

---

[1] A related problem is contamination of large losses by a non-recurring process. The papers on robust severity also address this, but it is a somewhat different topic than fitting a simple model to a complex process.

3 ROBUST APPROACH TO LOSS DEVELOPMENT

Omitting points from loss development triangles can sometimes lead to strange results, and not every development model can be automatically extended to deal with this, so instead of calculating the influence function for development models, we look at the sensitivity of the reserve estimate to changes in the cells of the development triangle, as in Tampubolon [1]. In particular, we define the impact of a cell on the reserve estimate under a particular development methodology as the derivative of the estimate with respect to the value in the cell. We do this for the incremental triangle, so a small change in a cell affects all subsequent cumulative values for the accident year. This seems to make more sense than looking at the derivative with respect to cumulative cells, whose changes would not continue into the rest of the triangle.

If you think of a number in the triangle as its mean plus a random innovation, the derivative with respect to the random innovation would be the same as that with respect to the total, so a high impact of a cell would imply a high impact of its random component as well. Thus models with some cells having high impacts would be less desirable. One measure of this is the maximum impact of any cell, which would be analogous to the GES, but we will also look at the number of cells with impacts above various thresholds in absolute value.

This is just a toe in the water of robust analysis of loss development. We are not proposing any robust estimators, and will stick with MLE or possibly quasi-likelihood. Rather we are looking at the impact function as a model selection and refinement tool. It can be used to compare competing models of the same development triangle, and it can identify problems with models that can guide a search for more robust alternatives. This is similar to finding models that work for the entire history of interest rate changes and are not too sensitive to any particular points.

To help interpret the impact function, we will also look at the generalized degrees of freedom (gdf) at each point. This is defined as the derivative of the fitted value with respect to the observed value. If this is near 1, the point's initial degree of freedom has essentially been used up by the model. The gdf is a measure of how much a point is able to pull the fitted value towards itself. Part of the impact of a point is this power to influence the model, but where it appears in the triangle also can influence the estimated reserve. Just like with the impact function, high values of the gdf would be a detriment.

For the chain ladder (CL) model, some observations can be made in general. All three corners of

the triangle have high impact. The lower left corner is the initial value of the latest accident year, and the full cumulative development applies to it. Since this point does not affect any other calculations, its impact is the development factor, which can sometimes be substantial. The upper right corner usually produces a development factor which, though small, applies to all subsequent accident years, so its impact can also be substantial. When there is only one year at ultimate, this impact is the ratio of the sum of all accident years not yet at ultimate, developed to the penultimate lag, to the penultimate cumulative value for the oldest accident year. The upper left corner is a bit strange in that its impact is usually negative. Increasing it will increase the cumulative loss at every lag, without affecting future incrementals, so every incremental-to-previous-cumulative ratio will be reduced. The points near the upper right corner also tend to have high impact, and those near the upper left tend to have negative impact, but the lower left point often stands alone in its high impact.

The GDFs for CL are readily calculated when factors are sums of incrementals over sums of previous cumulatives. The fitted value at a cell is the factor applied to the previous cumulative, so its derivative is its previous cumulative times the derivative of the factor with respect to the cell value. But that derivative is just the reciprocal of the sum of the previous cumulatives, so the gdf for the cell is its previous cumulative over the sum. Thus these GDFs sum down a column to unity, so each development factor uses up a total gdf of 1.0. Essentially each factor uses 1 degree of freedom, agreeing with standard analysis. The average gdf in a column is thus the reciprocal of the number of observations in that column. Thus the upper right cell uses 1 gdf, the previous column's cells use ½ each on average, etc. Thus the upper right cells have high GDFs and high impact.

We will use ODP to refer to the cross-classified development model in which each cell mean is modeled as a product of a row parameter and a column parameter, the variance of the cell is proportional to its mean, and the parameters are estimated by quasi-likelihood. It is well known that this model gives the same reserve estimate as CL. Thus if you change a cell slightly, the changed triangle will give the same reserve under ODP and CL. Thus the impacts of each cell under ODP will be the same as those of CL. The GDFs will not be the same, however, as the fitted values are not the same for the two models. The CL fitted value is the factor times the previous cumulative, whereas the ODP cumulative fitted values are backed down from the latest diagonal by the development factors, and then differenced to get the incremental fitted. It is possible to write down the resulting GDFs explicitly, but it is probably easier to calculate them numerically.

It may be fairly easy to find models that reduce the impact of the upper right cells. Usually the

development factors at those points are not statistically significant. Often the development is small and random, and is not correlated with the previous cumulative values. In such cases, it may be reasonable to model a number of such cells as a simple additive constant. Since several cells go into the estimation of this constant, the impact of some of them is reduced. Alternatively the factors in that region may follow some trends, linear or not, that can be used to express them with a small number of parameters. Again this would limit the impact of some of the cells.

The lower left point is more difficult to deal with in a CL-like model. One alternative is a Cape Cod-type model, where every accident year has the same mean level. This can arise, for instance, if there is no growth in the business, but also can be seen when the development triangle consists of on-level loss ratios, which have been adjusted to eliminate known differences among the accident years. In this type of model, all the cells go into estimating the level of the last accident year, so the lower left cell has much less impact. This reduction in the impact of the random component of this cell is a reason for using on-level triangles.

The next three sections illustrate these concepts using development triangles from the actuarial literature. The impacts and GDFs are calculated for various models fit to these triangles. The impacts are calculated by numerical derivatives, as are the GDFs except for those for the CL, which have been derived above.

4 A DEVELOPMENT-FACTOR EXAMPLE

4.1 Chain Ladder

Table 2 is a development triangle used in Venter [12]. Note that the first two accident years are developed all the way to the end of the triangle, at lag 11. Table 3 shows the impact of each cell on the reserve estimate using the usual sum/sum development factors. In the CL model an explicit formula can be derived for these impacts, but it is easier to do the derivatives numerically, simply by adding a small value to each cell separately and recalculating the estimated reserve to get the change in reserve for the derivative.

Table 2: Incremental Loss Development Triangle

L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
11,305	18,904	17,474	10,221	3,331	2,671	693	1,145	744	112	40	13
8,828	13,953	11,505	7,668	2,943	1,084	690	179	1,014	226	16	616
8,271	15,324	9,373	11,716	5,634	2,623	850	381	16	28	558	
7,888	11,942	11,799	6,815	4,843	2,745	1,379	266	809	12		
8,529	15,306	11,943	9,460	6,097	2,238	493	136	11			
10,459	16,873	12,668	9,199	3,524	1,027	924	1,190				
8,178	12,027	12,150	6,238	4,631	919	435					
10,364	17,515	13,065	12,451	6,165	1,381						
11,855	20,650	23,253	9,175	10,312							
17,133	28,759	20,184	12,874								
19,373	31,091	25,120									
18,433	29,131										
20,640											

Table 3: Impact of CL

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	-1.21	-0.34	0.04	0.39	0.73	1.10	1.48	1.85	2.46	3.35	4.61	7.31
AY1	-1.21	-0.34	0.04	0.39	0.73	1.10	1.48	1.85	2.46	3.35	4.61	7.31
AY2	-1.17	-0.29	0.08	0.44	0.78	1.14	1.53	1.89	2.51	3.39	4.66	
AY3	-1.15	-0.27	0.10	0.46	0.80	1.16	1.55	1.91	2.53	3.41		
AY4	-1.14	-0.27	0.11	0.46	0.80	1.17	1.56	1.92	2.54			
AY5	-1.10	-0.23	0.15	0.50	0.84	1.21	1.59	1.96				
AY6	-1.07	-0.20	0.18	0.53	0.87	1.24	1.62					
AY7	-1.03	-0.16	0.22	0.57	0.91	1.28						
AY8	-0.95	-0.08	0.30	0.65	0.99							
AY9	-0.73	0.14	0.52	0.87								
AY10	-0.31	0.57	0.95									
AY11	0.70	1.58										
AY12	4.95											

As discussed, the impacts are highest in the upper right and lower left corners, and the upper left has negative impact. The impacts increase moving to the right and down. The last four columns and the lower left point have impacts above 2, and six points have impacts above 4. Table 4 shows the GDFs for the chain ladder using the formula previous cumulative /sum previous cumulatives derived in Section 3. L0's GDFs are shown as identically 1.0. Like the impact function, these increase going to the right after lag 0. Within each column the sizes depend on the volume of the year.

Table 4: GDFs of CL

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	1.0	0.080	0.093	0.114	0.133	0.151	0.177	0.201	0.245	0.306	0.394	0.581
AY1	1.0	0.063	0.070	0.082	0.097	0.110	0.128	0.145	0.174	0.221	0.285	0.419
AY2	1.0	0.059	0.073	0.079	0.103	0.124	0.147	0.167	0.202	0.250	0.321	
AY3	1.0	0.056	0.061	0.076	0.089	0.106	0.128	0.148	0.177	0.223		
AY4	1.0	0.061	0.073	0.086	0.104	0.126	0.149	0.168	0.202			
AY5	1.0	0.074	0.084	0.096	0.113	0.130	0.149	0.170				
AY6	1.0	0.058	0.062	0.077	0.089	0.106	0.123					
AY7	1.0	0.074	0.086	0.098	0.123	0.146						
AY8	1.0	0.084	0.100	0.134	0.149							
AY9	1.0	0.122	0.141	0.158								
AY10	1.0	0.138	0.156									
AY11	1.0	0.131										
AY12	1.0											

Figure 1 graphs the impacts by lag along the diagonals of the triangle. After the first four lags, the impacts are almost constant across diagonals.

Figure 1: Impact of Chain Ladder by Diagonal
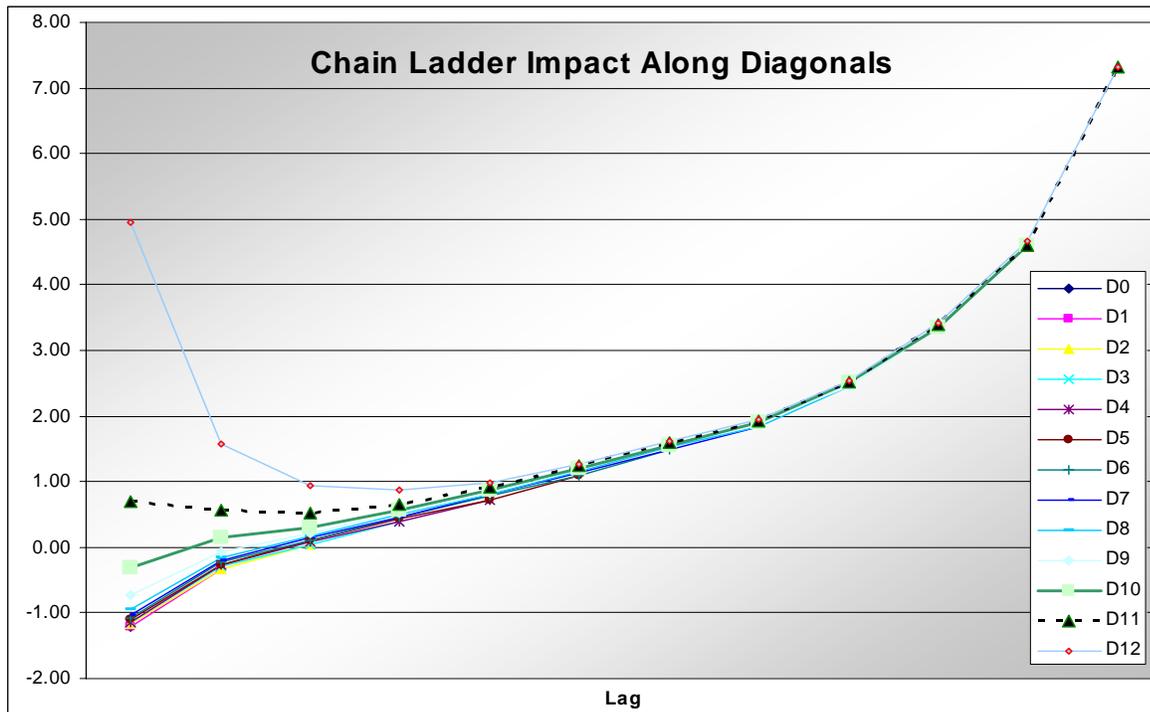
4.2 Regression Model

Venter [12] fit a regression model to this triangle, keeping the first five development factors but including an additive constant. The constant also represents development beyond lag 5. By stretching out the incremental cells to be fitted into a single column **Y**, this was put into the form of

a linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, which assumes a normal distribution of residuals with equal variance (homoscedasticity) across cells. $\mathbf{X}$ has the previous cumulative for the corresponding incrementals, with zeros to pad out the columns, a column of 1's for the constant. There were also diagonal (calendar year) effects in the triangle. Two diagonal dummy variables were included in $\mathbf{X}$, one with 1s for observations on the 4th diagonal and 0 elsewhere, and one equal to 1 on the 5th, 8th, and 10th diagonals, -1 on the 11th diagonal, and 0 elsewhere. The diagonals are numbered starting at 0, so the 4th is the one beginning with 8,529 and the 10th starts with 19,373. The variance calculation used a heteroscedasticity correction. This model with eight parameters fit the data better than the development factor model with 11 parameters. Here we are only addressing the robustness properties, however.

Table 5 gives the impact function for this model. It is clear that the large impacts on the right side have been eliminated by using the constant instead of factors to represent late development. The effects of the diagonal dummies can also be seen, especially in the right of the triangle. Now only 1 point has impact above 2, and above 4.

Table 5: Impact of Regression Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	-1.36	0.02	0.42	0.67	0.10	0.87	1.35	1.35	0.97	1.35	0.97	1.73
AY1	-1.56	0.22	0.66	-0.04	0.67	1.28	1.35	0.97	1.35	0.97	1.73	1.35
AY2	-1.53	0.52	-0.39	0.38	1.02	1.27	0.97	1.35	0.97	1.73	1.35	
AY3	-0.51	-0.64	0.15	0.78	1.07	0.90	1.35	0.97	1.73	1.35		
AY4	-1.24	-0.31	0.45	0.76	0.64	1.27	0.97	1.73	1.35			
AY5	-1.38	0.11	0.47	0.32	1.00	0.89	1.73	1.35				
AY6	-1.61	0.22	0.18	0.80	0.68	1.66	1.35					
AY7	-0.89	-0.36	0.35	0.24	1.34	1.25						
AY8	-1.34	0.00	-0.12	0.87	0.94							
AY9	0.29	-0.44	0.61	0.57								
AY10	-0.18	0.66	0.43									
AY11	1.11	1.04										
AY12	4.31											

Table 6 shows the GDFs for the regression model. For regression models the GDFs for the observations in the $\mathbf{Y}$ vector are known to be calculable as the diagonal of the "hat" matrix, where hat $= \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$, e.g., see Ye [2]. However in development triangles, changing an incremental value also changes subsequent cumulatives, so the $\mathbf{X}$ matrix is a function of lags of $\mathbf{Y}$. This requires the derivatives to be done numerically. The total of these, excluding lag 0, is 8.02, which is a bit above

the usual number of parameters, due to the exceptions to normal linear models. Compared to the CL, the GDFs are lower for lag 6 onward, but are somewhat higher along the modeled diagonals. They are especially high for diagonal 4, which is short and gets its own parameter.

Table 6: GDFs of Regression Model

|	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	1.0	0.071	0.089	0.127	0.352	0.195	0.034	0.034	0.055	0.034	0.055	0.076
AY1	1.0	0.047	0.056	0.305	0.107	0.099	0.034	0.055	0.034	0.055	0.076	0.034
AY2	1.0	0.046	0.299	0.084	0.098	0.123	0.055	0.034	0.055	0.076	0.034	
AY3	1.0	0.297	0.067	0.058	0.074	0.107	0.034	0.055	0.076	0.034		
AY4	1.0	0.064	0.056	0.072	0.120	0.128	0.055	0.076	0.034			
AY5	1.0	0.062	0.073	0.110	0.118	0.149	0.076	0.034				
AY6	1.0	0.040	0.067	0.061	0.095	0.140	0.034					
AY7	1.0	0.082	0.075	0.118	0.182	0.172						
AY8	1.0	0.077	0.134	0.212	0.207							
AY9	1.0	0.198	0.239	0.246								
AY10	1.0	0.245	0.253									
AY11	1.0	0.192										
AY12	1.0											

Figure 2: Impact of Regression Model by Diagonal
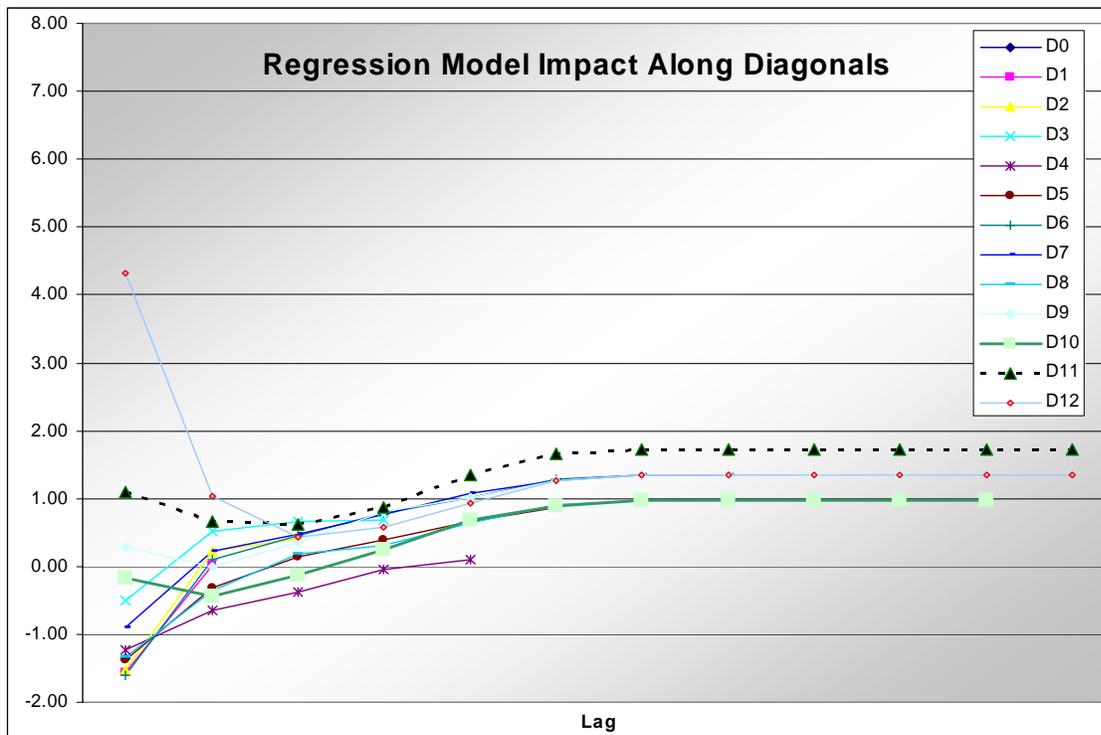
Figure 2 graphs the impacts. Note that due to the diagonal effects, diagonal 11 has higher impact than diagonal 12 after the first two lags.

4.3 Square Root Regression Model

As a correction for heteroscedasticity, regression courses sometimes advise dividing both **Y** and **X** by the square root of **Y**, row by row. This makes the model $\mathbf{Y}^{\frac{1}{2}} = (\mathbf{X}/\mathbf{Y}^{\frac{1}{2}})\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where the $\boldsymbol{\varepsilon}$ are IID mean zero normals. Then $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Y}^{\frac{1}{2}}\boldsymbol{\varepsilon}$, so now the variance of the residuals is proportional to **Y**. This sounds like a fine idea, but it is a catastrophe from a robust viewpoint. Table 7 shows the impact function. There are 12 points with impact over 2, 7 with impact over 4, 5 with impact over 10, and 3 with impact over 25.

Table 7: Impact of Square Root Regression Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	-0.94	-0.08	0.16	0.68	0.15	0.56	0.01	0.00	0.01	0.38	4.57	15.61
AY1	-1.06	-0.10	0.28	-0.30	2.19	1.86	0.01	0.15	0.00	0.15	10.21	0.01
AY2	-0.58	0.12	-0.09	0.20	0.68	0.39	0.01	0.03	28.26	3.09	0.02	
AY3	-0.20	-0.50	0.13	0.66	0.69	0.27	0.00	0.11	0.00	32.67		
AY4	-0.90	-0.15	0.33	0.41	0.59	0.56	0.03	0.14	37.14			
AY5	-1.28	-0.36	0.17	0.37	2.05	2.87	0.00	0.00				
AY6	-1.20	-0.09	0.01	0.77	0.71	2.34	0.02					
AY7	-1.02	-0.18	0.36	0.23	0.76	1.97						
AY8	-0.86	-0.07	-0.01	1.23	0.46							
AY9	-0.91	-0.06	0.59	1.02								
AY10	-0.45	0.48	0.89									
AY11	0.50	1.46										
AY12	4.56											

Part of the problem is that the equation $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Y}^{\frac{1}{2}}\boldsymbol{\varepsilon}$ is not what you would really want. The residual variance should be proportional to the mean, not the observation. This setup gives the small observations small variance, and so the ability to pull the model towards them. But the observations might be small because of a negative residual, with a higher expected value. So this formulation gives the small values too much influence.

Table 8 shows the related GDFs. It is unusual here that some points have GDFs greater than 1. A small change in the original value can make a greater change in the fitted value, but due to the non-linearity the fitted value is still a ways from the data point. The sum of the GDFs is 13.0, which is sometimes interpreted as the implicit number of parameters.

4.4 Gamma-*p* Residuals

Venter [13] fits the same regression model, but by maximum likelihood with gamma-*p* residuals. The gamma-*p* is a gamma distribution, but each cell is modeled to have the variance proportional to

the same power p of the mean. This models the cells with smaller means as having smaller variances, but the effect is not as extreme as in the square root regression, where the variance is proportional to the observation, not its expected value.

Table 8: GDFs of Square Root Regression Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	1.0	0.082	0.078	0.129	0.697	0.074	0.000	0.000	0.000	0.010	0.356	1.102
AY1	1.0	0.071	0.076	0.230	0.227	0.175	0.000	0.004	0.000	0.011	0.720	0.000
AY2	1.0	0.053	0.287	0.031	0.074	0.042	0.000	0.001	2.199	0.218	0.000	
AY3	1.0	0.201	0.045	0.081	0.064	0.025	0.000	0.008	0.000	0.906		
AY4	1.0	0.051	0.077	0.061	0.066	0.061	0.002	0.010	1.030			
AY5	1.0	0.076	0.102	0.089	0.252	0.322	0.000	0.000				
AY6	1.0	0.073	0.045	0.104	0.072	0.221	0.001					
AY7	1.0	0.069	0.103	0.053	0.106	0.249						
AY8	1.0	0.075	0.051	0.246	0.068							
AY9	1.0	0.117	0.193	0.208								
AY10	1.0	0.145	0.166									
AY11	1.0	0.144										
AY12	1.0											

Table 9: Impact of Gamma-p Residual Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	-0.59	-0.07	0.24	0.59	-0.03	1.47	1.37	1.37	1.23	1.25	-1.45	7.97
AY1	-0.90	-0.05	0.28	0.10	0.90	1.11	1.30	0.77	1.37	0.91	6.73	1.36
AY2	-0.46	0.08	-0.07	0.56	0.94	1.43	1.22	1.45	-5.62	4.33	1.35	
AY3	-0.29	-0.58	0.21	0.47	1.31	1.37	1.21	0.98	1.47	0.10		
AY4	-0.68	-0.15	0.19	0.51	0.94	1.48	1.24	1.96	0.02			
AY5	-1.04	-0.18	0.20	0.49	0.96	1.07	1.43	1.38				
AY6	-1.00	0.09	0.22	0.45	1.28	1.13	1.41					
AY7	-1.02	-0.18	0.50	0.50	0.95	1.17						
AY8	-0.71	-0.12	0.12	0.66	0.96							
AY9	-0.85	-0.02	0.80	0.86								
AY10	-0.44	0.48	0.88									
AY11	0.46	1.45										
AY12	4.43											

In this case, p was found to be 0.71. The impacts are shown in Table 9 and graphed in Figure 3. It is clear that these are not nearly as dramatic as the square root regression, but worse than the regular regression, and perhaps comparable to the chain ladder. Diagonals 10 and 11 can be seen to have a few significant impacts. These are at points with small observations that are also on modeled diagonals. Even with the variance proportional to a power of the expected value, these points still have a strong pull. The GDFs are in Table 10.

Figure 3: Impact of Gamma-*p* Residual Model

Table 10: GDFs of Gamma-*p* Residual Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	1.0	0.082	0.078	0.129	0.697	0.074	0.000	0.000	0.000	0.010	0.356	1.102

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY1	1.0	0.071	0.076	0.230	0.227	0.175	0.000	0.004	0.000	0.011	0.720	0.000
AY2	1.0	0.053	0.287	0.031	0.074	0.042	0.000	0.001	2.199	0.218	0.000	
AY3	1.0	0.201	0.045	0.081	0.064	0.025	0.000	0.008	0.000	0.906		
AY4	1.0	0.051	0.077	0.061	0.066	0.061	0.002	0.010	1.030			
AY5	1.0	0.076	0.102	0.089	0.252	0.322	0.000	0.000				
AY6	1.0	0.073	0.045	0.104	0.072	0.221	0.001					
AY7	1.0	0.069	0.103	0.053	0.106	0.249						
AY8	1.0	0.075	0.051	0.246	0.068							
AY9	1.0	0.117	0.193	0.208								
AY10	1.0	0.145	0.166									
AY11	1.0	0.144										
AY12	1.0											

Again this is less dramatic than for the square root regression, but the small points on the modeled diagonals still have high GDFs. The total of these is 11.3, which is still fairly high. This is somewhat troublesome, as the gamma-*p* model fit the residuals quite a bit better than did the

standard regression. The fact that the problems center on small observations on the modeled diagonals suggests that additive diagonal effects may not be appropriate for this data. They do fit into the mold of a generalized linear model, but that is not too important when fitting by MLE anyway. As an alternative, the same model but with the diagonal effects as multiplicative factors was fit. The multiplicative diagonal model can be written:

$$\mathbf{EY} = \mathbf{X}[,1:6]\beta[1:6]*\beta[7]^{X[,7]}*\beta[8]^{X[,8]},$$

which means that the first six columns of **X** are multiplied by the first six parameters, which includes the constant term, and then the last two diagonal parameters are factors raised to the power of the last two columns of **X**. These are now the diagonal dummies, which are 0, 1, or –1. Thus the same diagonals are higher and the same lower, but now proportionally instead of by an additive constant. It turns out that this model actually fits better, with a negative loglikelihood of 625, compared to 630 for the generalized linear model. This solves the robustness problems as well. The impacts are in Table 11, the GDFs in Table 12, and the impacts are graphed in Figure 4.

Table 11: Impact of Gamma-*p* Multiplicative Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	-0.94	-0.03	0.22	0.58	0.09	1.16	1.43	1.43	1.42	1.36	0.55	2.31
AY1	-1.02	0.00	0.32	0.17	0.56	1.02	1.43	1.26	1.43	1.30	2.14	1.43
AY2	-0.74	0.15	-0.46	0.39	0.98	1.30	1.42	1.42	-0.78	1.82	1.42	
AY3	-0.25	-0.50	-0.02	0.46	0.97	1.26	1.43	1.33	1.43	0.69		
AY4	-0.68	-0.39	0.23	0.51	0.83	1.26	1.39	1.50	0.64			
AY5	-1.09	-0.10	0.33	0.26	0.93	0.69	1.43	1.43				
AY6	-1.02	0.05	0.00	0.45	0.79	1.12	1.42					
AY7	-0.72	-0.37	0.31	0.29	1.11	1.07						
AY8	-0.81	-0.01	-0.21	0.92	0.99							
AY9	-0.76	-0.25	0.85	0.88								
AY10	-0.58	0.56	0.94									
AY11	0.35	1.50										
AY12	4.34											

Diagonal 11 still has more impact than the others, but this barely exceeds 2.0 at the maximum. The sum of the GDFs is 8.67. There are eight parameters for the cell means but two more for the gamma-*p*. It has been a question whether or not to count those two in determining the number of parameter used in the fitting. The answer to that from the gdf analysis is basically to count each of those as 1/3 in this case. Here the robust analysis has uncovered a previously unobserved problem with the generalized linear model, and lead to an improvement.

Table 12: GDFs of Gamma-*p* Multiplicative Model

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11
AY0	1.0	0.079	0.087	0.125	0.323	0.136	0.034	0.033	0.038	0.040	0.093	0.074
AY1	1.0	0.063	0.069	0.191	0.210	0.132	0.034	0.048	0.033	0.046	0.066	0.034
AY2	1.0	0.053	0.410	0.079	0.085	0.068	0.038	0.035	0.175	0.050	0.034	
AY3	1.0	0.361	0.105	0.070	0.071	0.063	0.033	0.044	0.031	0.101		
AY4	1.0	0.107	0.070	0.067	0.111	0.084	0.040	0.034	0.106			
AY5	1.0	0.079	0.094	0.158	0.185	0.276	0.030	0.033				
AY6	1.0	0.066	0.106	0.081	0.104	0.117	0.035					
AY7	1.0	0.143	0.093	0.127	0.108	0.200						
AY8	1.0	0.080	0.200	0.220	0.102							
AY9	1.0	0.355	0.281	0.208								
AY10	1.0	0.316	0.196									
AY11	1.0	0.163										
AY12	1.0											

Figure 4: Impact of Gamma-*p* Multiplicative Model
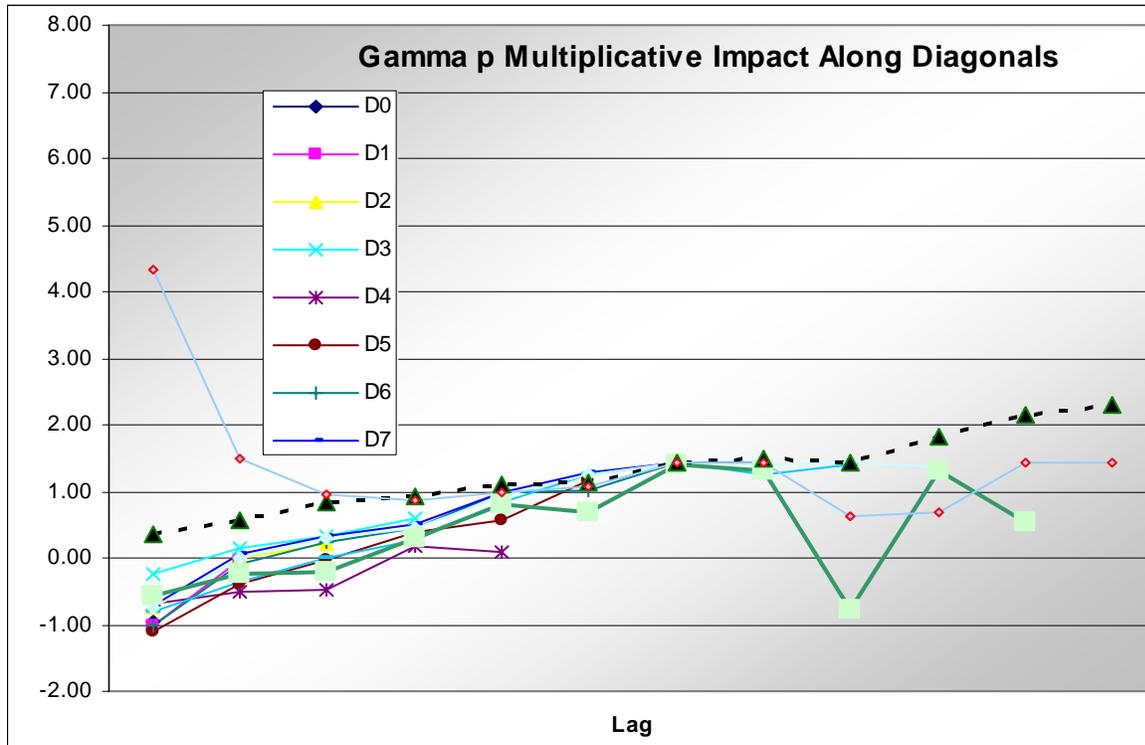
5 A MULTIPLICATIVE FIXED-EFFECTS EXAMPLE

A multiplicative fixed-effects model is one where the cell means are products of fixed factors from rows, columns, and perhaps diagonals. The most well-known is the ODP model discussed in

section 3, where there is a factor for each row, interpreted as estimated ultimate, a factor for each column, interpreted as fraction of ultimate for that column, and the variance of each cell is a fixed factor times its mean. This model if estimated by MLE gives the same reserve estimates as the chain ladder and so the same impacts for each cell, but the GDFs are different, due to the different fitted values.

The triangle for this example comes from Taylor-Ashe (1983) and is shown in Table 13. The CL = ODP impacts are in Table 14 and are graphed in Figure 5.

Table 13: Incremental Triangle Taylor-Ashe (1983)

Lag 0	L1	L2	L3	L4	L5	L6	L7	L8	L9
357,84	766,940	610,542	482,940	527,32	574,39	146,34	139,95	227,22	67,94
352,11	884,021	933,894	1,183,28	445,74	320,99	527,80	266,17	425,04	
290,50	1,001,79	926,219	1,016,65	750,81	146,92	495,99	280,40		
310,60	1,108,25	776,189	1,562,40	272,48	352,05	206,28			
443,16	693,190	991,983	769,488	504,85	470,63				
396,13	937,085	847,498	805,037	705,96					
440,83	847,631	1,131,39	1,063,26						
359,48	1,061,64	1,443,37							
376,68	986,608								
344,01									

Table 14: Impact of CL = ODP on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	-3.11	-1.62	-1.01	-0.45	0.01	0.51	1.16	2.27	4.54	12.59
AY1	-2.87	-1.38	-0.77	-0.20	0.25	0.76	1.40	2.51	4.78	
AY2	-2.43	-0.93	-0.33	0.24	0.69	1.20	1.85	2.95		
AY3	-2.21	-0.72	-0.11	0.45	0.91	1.41	2.06			
AY4	-1.95	-0.46	0.15	0.71	1.17	1.67				
AY5	-1.67	-0.18	0.43	0.99	1.45					
AY6	-1.25	0.25	0.85	1.42						
AY7	-0.14	1.35	1.96							
AY8	2.07	3.57								
AY9	13.45									

Figure 5: Impact of CL = ODP on TA
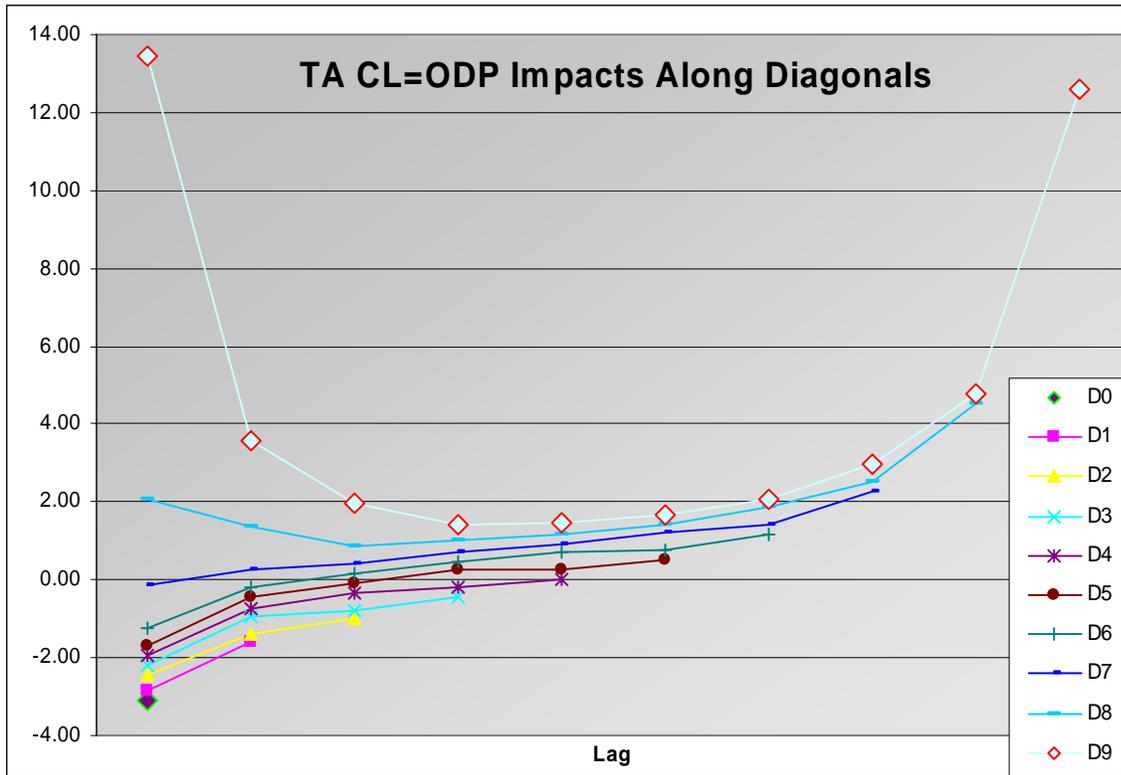
Because the development factors are higher, the impacts are higher than in the previous example. Even though it is a smaller triangle, 14 points have impacts with absolute values over 2, 4 are over 4, and 2 are over 12. The CL GDFs are in Table 15. These sum to 9, excluding the first column, and are fairly high on the right where there are few observations per column. The ODP GDFs are in Table 16. These sum to 19, and are fairly high near the upper right and lower left corners.

Table 15: GDFs of CL on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	1.0	0.108	0.110	0.115	0.120	0.153	0.208	0.272	0.423	1.0
AY1	1.0	0.106	0.121	0.144	0.182	0.211	0.258	0.365	0.577	
AY2	1.0	0.087	0.126	0.147	0.175	0.222	0.259	0.363		
AY3	1.0	0.093	0.138	0.146	0.204	0.224	0.275			
AY4	1.0	0.133	0.111	0.141	0.157	0.189				
AY5	1.0	0.119	0.130	0.145	0.162					
AY6	1.0	0.132	0.126	0.161						
AY7	1.0	0.108	0.139							
AY8	1.0	0.113								
AY9	1.0									

Table 16: GDFs of ODP on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	0.154	0.261	0.273	0.295	0.229	0.224	0.253	0.301	0.459	1.0
AY1	0.186	0.295	0.308	0.333	0.276	0.281	0.325	0.400	0.612	
AY2	0.187	0.300	0.312	0.338	0.278	0.282	0.324	0.398		
AY3	0.188	0.304	0.317	0.344	0.280	0.282	0.323			
AY4	0.184	0.309	0.322	0.348	0.275	0.271				
AY5	0.197	0.331	0.346	0.374	0.293					
AY6	0.221	0.375	0.391	0.423						
AY7	0.284	0.498	0.519							
AY8	0.370	0.747								
AY9	1.0									

The GDFs can be used to allocate the total degrees of freedom of the residuals of $n - p$. The n is allocated 1 to each observation, and the p can be set to the gdf of each observation. This would give a residual degree of freedom to each observation which could be used in calculating a standardized residual that takes into account how the degrees of freedom vary among observations.

Venter [12] looked at reducing the number of parameters in this model by setting parameters equal if they are not significantly different, and using trends, like linear trends between parameters. Also diagonal effects were introduced. The result was a model where each cell mean is a product of its row, column, and diagonal factors. There are six parameters overall. For the rows there are three parameters, for high, medium, and low accident years. Accident year 0 is low, year 7 is high, year 6 is the average of the medium and high levels, and all other years are medium. There are 2 column factors: high and low. Lags 1, 2, and 3 are high, lag 4 is an average of high and low, lag 0 and lags 5 to 8 are low, and lag 9 is 1 minus the sum of the other lags. Finally there is one diagonal parameter c. Diagonals 4 and 6 have factors 1+c, lag 7 has factor 1 – c, and all the other diagonals have factor 1.

With just six parameters this model actually provides a better fit to the data than the 19 parameter model. The combining of parameters does not degrade the fit much, and adding diagonal effects improves the fit. An improved fit over that in Venter [12] was found by using a gamma-p distribution with $p = \frac{1}{2}$ so the variance of each cell is proportional to the square root of its mean. The impacts and GDFs of this model are shown in Tables 17 and 18, and the impacts are graphed in Figure 6, this time along accident years.

Table 17: Impact of 6-Parameter Gamma-½ on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	0.65	-0.82	-1.08	-2.07	-0.87	0.97	-0.32	0.33	0.53	12.06
AY1	1.45	-0.02	0.68	0.60	-0.25	1.90	1.40	1.61	1.57	
AY2	1.64	0.75	-0.19	0.84	0.90	1.93	1.66	1.36		
AY3	1.26	0.43	-0.21	0.97	-0.36	1.70	1.71			
AY4	1.62	0.08	0.67	0.37	0.63	1.35				
AY5	1.19	-0.11	0.57	0.51	1.17					
AY6	2.56	1.19	0.91	1.13						
AY7	2.18	1.27	1.49							
AY8	1.72	0.92								
AY9	1.59									

Table 18: GDFs of 6-Parameter Gamma-½ on TA

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9
AY0	0.046	0.152	0.211	0.288	0.150	0.017	0.248	0.095	0.082	0.938
AY1	0.044	0.031	0.057	0.155	0.014	0.115	0.018	0.051	0.043	
AY2	0.055	0.041	0.134	0.027	0.102	0.114	0.046	0.049		
AY3	0.045	0.078	0.062	0.028	0.181	0.052	0.064			
AY4	0.078	0.057	0.026	0.119	0.011	0.037				
AY5	0.037	0.147	0.083	0.032	0.026					
AY6	0.254	0.200	0.095	0.100						
AY7	0.111	0.527	0.250							
AY8	0.047	0.031								
AY9	0.047									

Figure 6: Impact of Gamma-½ on TA
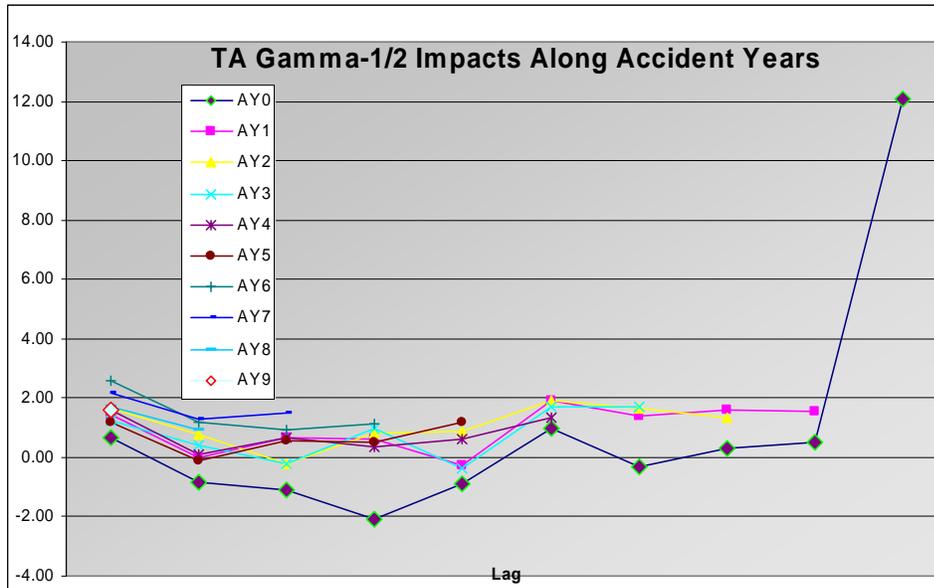
The impacts are now all quite well contained except for one point – the last point in AY0. Possibly because AY0 gets its own parameter, lag 9 influences the level of the other lags' parameters, and this is a small point with a small variance, this model only slightly reduces the high level of impact that point has in ODP. The same thing can be seen in the GDFs as well, where this point has slightly less than a whole gdf. The points on AY7 and the modeled diagonals also have relatively high GDFs, as do some small cells. The total of the GDFs is 6.14. There are six parameters affecting the means, plus one for the variance of the gamma. That one can affect the fit slightly, so counting it as $1/7^{th}$ of a parameter seems reasonable.

In an attempt to solve the problem of the upper-right point, an altered model was fit: lag 9 gets half of the paid in the low years. This can be considered a trend to 0 for lag 10. Making the lags sum to 1.0 now eliminates a parameter, so there are five. The NLL is slightly worse, at 722.40 vs. 722.36, but that is worth saving a parameter. The robustness is now much better, with only two impacts above 2.0, the largest being 2.35.

6 PAID AND INCURRED EXAMPLE

Venter [15], following Quarg and Mack [16], builds a model for simultaneously estimating paid and incurred development, where each influences the other. The paid losses are part of the incurred losses, so the separate effects are from the paid and unpaid triangles, shown in Tables 19 and 20.

First the impacts on the reserve (7059.47) from the average of the paid and incurred chain ladder reserves is calculated, where the paids at the last lag are increased by the incurred-to-paid ratio at that lag. Tables 21 and 22 show the impacts of the paid and unpaid triangles, and Tables 23 and 24 show the GDFs.

Table 19: Quarg-Mack Paid Increments

	L0	L1	L2	L3	L4	L5	L6
AY0	576	1228	166	54	50	28	29
AY1	866	1082	214	70	52	64	
AY2	1412	2346	494	164	78		
AY3	2286	3006	432	126			
AY4	1868	1910	870				
AY5	1442	2568					
AY6	2044						

Table 20: Quarg-Mack Unpaid

	L0	L1	L2	L3	L4	L5	L6
AY0	402	300	164	120	100	80	43
AY1	978	604	304	248	224	106	
AY2	1492	596	446	184	150		
AY3	1216	666	346	292			
AY4	944	1104	204				
AY5	1200	396					
AY6	2978						

Table 21: Average Reserve Impact of Paid

	L0	L1	L2	L3	L4	L5	L6
A0	-0.68	-.02	0.32	0.86	2.32	5.95	13.99
A1	-0.45	0.20	0.54	1.08	2.54	6.17	
A2	-0.41	0.24	0.58	1.12	2.59		
A3	-0.36	0.30	0.64	1.18			
A4	-0.32	0.33	0.67				
A5	-0.20	0.46					
A6	1.37						

Table 22: Average Reserve Impact Unpaid

	L0	L1	L2	L3	L4	L5	L6
A0	-0.29	-0.15	-0.26	-0.72	-1.76	-4.01	14.99
A1	-0.29	-0.15	-0.26	-0.72	-1.76	3.57	
A2	-0.29	-0.15	-0.26	-0.72	1.77		
A3	-0.29	-0.15	-0.26	1.08			
A4	-0.29	-0.15	0.82				
A5	-0.29	0.68					
A6	0.84						

Table 23: Average Reserve GDF of Paid

	0	L1	L2	L3	L4	L5	L6
A0	1	0.068	0.109	0.140	0.233	0.476	1
A1	1	0.102	0.117	0.153	0.257	0.524	
A2	1	0.167	0.227	0.301	0.509		
A3	1	0.271	0.319	0.406			
A4	1	0.221	0.228				
A5	1	0.171					
A6	1						

Table 24: Average Reserve GDF Unpaid

	0	L1	L2	L3	L4	L5	L6
A0	1	0.067	0.106	0.139	0.232	0.464	1
A1	1	0.126	0.129	0.160	0.269	0.536	
A2	1	0.198	0.219	0.306	0.499		
A3	1	0.239	0.300	0.395			
A4	1	0.192	0.246				
A5	1	0.180					
A6	1						

The impacts of the lower left are not great, mostly because the development factors are fairly low in this example. The impacts on the upper right of both paid and unpaid losses are quite high, however. The unpaid losses not on the last diagonal have a negative impact, because they lower subsequent incurred development factors, but do not have factors applied to them. The GDFs are similar to CL in general.

The model in Venter [15] used generalized regression for both the paid and unpaid triangles, where regressors could be from either triangle or from the cumulative paid and incurred triangles. Except for the first couple of columns, the previous unpaid losses provided reasonable explanations of both the current paid increment and the current remaining unpaid. The paid and unpaid at lags 3 and on were just multiples of the previous unpaid, with a single factor for each. That is, expected paids were 33.1%, and unpaids 72.3%, of the previous unpaid. Since these sum to more than 1, there is a slight upward drift in the incurred. The lag 2 expected paid was 68.5% of the lag 1 unpaid. The best fit to the lag 2 expected unpaid was 9.1% of the lag 1 cumulative paid. For lag 1 paid, 78.1% of the lag 0 incurred was a reasonable fit. Lag 1 unpaid was more complicated, with the best fit being a regression, with constant, on lag 0 and lag 1 paids. There were also diagonal effects in both models. The residuals were best fit with a Weibull distribution. Tables 25 – 28 show the fits.

Table 25: Weibull Model Impact of Paid

	L0	L1	L2	L3	L4	L5	L6
A0	0.09	-0.18	-1.58	4.38	0.38	7.67	5.45
A1	0.04	0.26	0.59	1.90	2.75	2.32	
A2	-.37	0.33	0.42	0.57	-0.28		
A3	-.13	0.17	0.67	1.26			
A4	-.02	0.20	0.31				
A5	-.94	0.70					
A6	1.25						

Table 26: Weibull Model Impact Unpaid

	L0	L1	L2	L3	L4	L5	L6
A0	0.06	0.67	-1.02	-1.45	-1.82	0.51	4.14
A1	-0.17	-0.44	-1.80	-0.73	0.52	2.56	
A2	-0.20	-0.16	0.47	-1.17	3.63		
A3	-0.09	-0.32	-1.17	2.51			
A4	-0.10	-0.34	1.89				
A5	-0.32	1.47					
A6	0.65						

Table 27: Weibull Model GDF of Paid

	0	L1	L2	L3	L4	L5	L6
A0	1	0.938	0.725	0.235	0.268	0.125	.143
A1	1	0.451	0.057	0.052	0.066	0.065	
A2	1	0.192	0.347	0.377	0.290		
A3	1	0.137	0.250	0.145			
A4	1	0.094	0.277				
A5	1	0.269					
A6	1						

Table 28: Weibull Model GDF Unpaid

	0	L1	L2	L3	L4	L5	L6
A0	1	0.824	0.172	-0.058	0.015	.072	.054
A1	1	0.357	0.700	-0.044	0.115	.052	
A2	1	0.152	0.465	-0.173	0.113		
A3	1	0.050	0.136	0.123			
A4	1	0.507	0.089				
A5	1	0.687					
A6	1						

The two highest impacts for the average of paid and incurred are 14 and 15. For the Weibull they are 7.7 and 5.5. The average has two other points with impacts above 5, whereas the Weibull has none. Below 5 the impacts are roughly comparable. Since the Weibull has variance proportional to the mean squared, small observations have lower variance, and so a stronger pull on the model and higher impacts. In total, excluding the first column, the GDFs sum to 9.9, but including the diagonals (see Venter [15] for details) there are 12 parameters plus two Weibull shape parameters. The form of the model apparently does not allow the parameters to be fully expressed. The Weibull model still has more high impacts than would be desirable, but it is a clear improvement over the average of the paid and incurred. The reserve is quite a bit lower for the better-fitting Weibull model as well: 6255 vs. 7059.

7 CONCLUSION

Robust analysis has been introduced as an additional testing method for loss development models. It is able to identify points that have a large influence on the reserve, and so whose random components would also have a large influence. Through three examples, customized models were found to be more robust than standard models like CL and ODP, and in two of the examples, even better models were found as a response to the robust analysis.

REFERENCES

[1] Tampubolon, Dumaria R. 2008. "Uncertainties in the estimation of outstanding claims liability in general insurance," PhD thesis, Macquarie University, Sydney, Australia.

[2] Ye, Jianming. 1998. "On measuring and correcting the effects of data mining and model selection," *Journal of the American Statistical Association*, March 1998, 93(441): 120-131.

[3] Tukey, John W. 1960. "A survey of sampling from contaminated distributions," in I. Olkin (Ed.), *Contributions to Probability and Statistics*, (pp. 448-485) (Palo Alto: Stanford University Press).

[4] Dell'Aquila, Rosario, Elvezio Ronchetti, and Fabio Troiani. 2003. "Robust GMM analysis of models for the short rate process," *Journal of Empirical Finance* 10:373-397.

[5] Hampel, Frank R. 1968. "Contributions to the Theory of Robust Estimation." PhD Thesis, University of California, Berkeley.

[6] Bilodeau, Martin. 2001. Discussions of Papers Already Published. *North American Actuarial Journal* 5(3): 23–28. (online version included in Brazauskas, Vytaras, and Robert Serfling, 2000)

[7] Brazauskas, Vytaras, and Robert Serfling. 2000a. "Robust and Efficient Estimation of the Tail Index of a Single-Parameter Pareto Distribution," *North American Actuarial Journal* 4(4):12–27.

[8] Finkelstein, Mark, Howard G. Tucker, and Jerry Alan Veeh. 2006. "Pareto Tail Index Estimation Revisited," *North American Actuarial Journal* 10(1):12–27.

[9] Brazauskas, Vytaras and Robert Serfling. 2000b. "Robust Estimation of Tail Parameters for Two-Parameter Pareto and Exponential Models via Generalized Quantile Statistics," *Extremes* 3(3) Sept: 231-249

[10] Gather, Ursula and Verena Schultze, 1999. "Robust Estimation of Scale of an Exponential Distribution," *Statistica Neerlandica* 53(3):327-341

[11] Serfling, Robert 2002. "Efficient and Robust Fitting of Lognormal Distributions," *North American Actuarial Journal* 6(4):95-116.

[12] Venter, Gary G., 2007a. "Refining Reserve Runoff Ranges," CAS *E-Forum*, Summer 2007.

[13] Venter, Gary G., 2007b. "Generalized Linear Models beyond the Exponential Family with Loss Reserve Applications," CAS *E-Forum*, Summer 2007.

[14] Taylor, Greg C. and Frank R. Ashe, 1983. "Second Moments of Estimates of Outstanding Claims," *Journal of Econometrics* 23:37-61.

[15] Venter, Gary G., 2008. "Distribution and Value of Reserves Using Paid and Incurred Triangles," CAS Reserves Call Papers, Casualty Actuarial Society *E-Forum*, Fall 2008, http://www.casact.org/pubs/forum/ (to be posted August 2008).

[16] Quarg, Gerhard and Thomas Mack, 2004. "Munich chain-ladder—A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses," *Blätter DGVFM* Heft 4, 597–630.

Biographies of the Authors

Gary G. Venter is managing director at Guy Carpenter, LLC. He has an undergraduate degree in philosophy and mathematics from the University of California and an MS in mathematics from Stanford University. He has previously worked at Fireman's Fund, Prudential Reinsurance, NCCI, Workers Compensation Reinsurance Bureau and Sedgwick Re, some of which still exist in one form or another. At Guy Carpenter, Gary develops risk management and risk modeling methodology for application to insurance companies. He also teaches a graduate course in loss modeling at Columbia University.

+1.917.937.3277   **gary.g.venter@guycarp.com**

Dumaria R. Tampubolon is a lecturer at the Faculty of Mathematics and Natural Sciences at Institut Teknologi Bandung (Bandung Institute of Technology), Bandung, Indonesia. She has a "Sarjana S1" (similar to Honors) degree in mathematics, majoring in statistics, from Institut Teknologi Bandung; an M.Sc degree in statistics from Monash University, Melbourne, Australia; and a PhD degree (2008) in actuarial studies from Macquarie University, Sydney, Australia. Her research interests are in the area of general insurance.

+62.22.250.2545 ext.101   **dumaria@math.itb.ac.id, drtampubolon@gmail.com**