Minimizing Uncertainty in Property Casualty Loss Reserve Estimates

Chris G. Gross, ACAS, MAAA

The uncertain nature of property-casualty loss reserves

Property Casualty loss reserves are inherently uncertain. There are many potential outcomes of individual claims. There are claims that have not been reported. The claim handling, inflation, and judicial environment may all change in unexpected ways, impacting future loss payments.
While the uncertainty of reserves is inescapable, the *level* of this uncertainty is *not* fixed, but rather is a function of the strength of our models and judgment as actuaries. Less effective models of loss development will generally result in greater differences between predictions of future loss payments and the actual payments themselves. Superior judgment too, can reduce this predictive uncertainty. Shifts in the mix of heterogeneous claims within a body of claims data will tend to increase the uncertainty of the estimate. Greater amounts of information, if properly incorporated into an analysis will tend to decrease the uncertainty. At some point, however, the cost associated with obtaining and incorporating additional information is not warranted by the potential reduction in uncertainty, and a practical minimum of uncertainty is achieved.

In estimating any uncertain quantity two characteristics of the estimate are often discussed by statisticians, the bias of the estimate, and the variance of the estimate. Generally, it is accepted by actuaries that loss reserve estimates should be unbiased, (i.e. equal to the statistical expected value over the range of potential outcomes), although this view is not universally held[1]. Regardless, for a given level of bias it is desirable to have an estimate that has a low level of variance[2].

The effect of multiple estimates on uncertainty

It is common for actuaries to calculate multiple estimates of the same unpaid lost amount. The most common estimates are those based on multiplicative link ratios and on the Bornhuetter-Ferguson technique, both of these methods often applied to both paid and case-incurred (reported) losses. While actuaries may not necessarily speak of it explicitly, when multiple methods are used to arrive at an estimate of future loss payments, the goal is one of improving predictive accuracy/ reducing uncertainty. The actuary feels more comfortable with the selected estimate when it is consistent with multiple estimates. This is an implicit reflection of the reduction in variance that occurs from

---

[1] An example of advocating for other than the mean is the paper "Management's Best Estimate of Loss Reserves" by Rodney Kreps

[2] Note: This paper focuses on the variance (uncertainty) of estimates as measured by the difference between the estimate and the eventual *outcome* of the body of reserves being analyzed. While there is much discussion in the literature about variability of *central estimates* of reserves (i.e. variance of the mean of the distribution), the concept is not well defined. Each body of reserves is unique, and the distinction between parameter risk, process risk, and model risk is arbitrary, based on the model of development used. The variance between the estimate and the ultimate outcome, by contrast is well defined.

the addition of additional information. Selecting an estimate that is between estimates from various methods is an implicit weighted average of the estimates.

Formalization of the impact of using multiple methods

Formal consideration of weighting individual estimates together to achieve a combined estimate that minimizes uncertainty is instructive in understanding the process that happens often informally, and has the potential to bring new insights, and discipline to the process.

Notation and Assumptions:

n=number of estimates

P=the actual total of future payments (fixed, but unknown)

X_i=the estimate of P based on method i

$E(X_i) = P$ (X_i is unbiased)[3]

X = column vector $(X_1,X_2, \ldots X_n)$

$\sigma_{ii} = Var(X_i)=E((X_i-P)^2)$

$\sigma_{ij} = Covar(X_i,X_j) = E[(X_i-P)(X_j-P)]$

A = covariance matrix with elements $a_{ij}=\sigma_{ij}$

w = column vector of weights $(w_1, w_2,\ldots w_n)$ to be applied to the corresponding estimates to arrive at combined estimate Y

$Y=\mathbf{X'w}$

e = the column vector of length n containing values of 1 for all elements

Calculation of minimum variance weighting of estimates:

$\mathbf{e'w}=1$ (constraint that the weights sum to 1)

$Var(Y)=Var(\mathbf{X'w})=\mathbf{w'Aw}$

Our objective is to:

Minimize Var(Y) over **w**, subject to the constraint $\mathbf{e'w}=1$

This problem can be rewritten using a Lagrange multiplier as:

Minimize $\mathscr{L}$ over **w** and λ

where $\mathscr{L} = \mathbf{w'Aw} + \lambda(1-\mathbf{e'w})$

Taking the partial derivatives with respect to each element of **w** and setting them equal to zero gives the n equations:

$2\mathbf{Aw}=\lambda\mathbf{e}$

[3] In the case where it is believed that the estimate *is* biased, the estimate should be excluded entirely from this process or preferably the estimate should be adjusted to remove the perceived bias.

which can be rewritten:
$\mathbf{w}=\lambda\mathbf{A}^{-1}\mathbf{e}/2$

Substituting **w** into **e'w**=1:
$\mathbf{e}'\lambda\mathbf{A}^{-1}\mathbf{e}/2=1$
$\lambda=2/(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})$

and therefore:
$\mathbf{w}=\mathbf{A}^{-1}\mathbf{e}/(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})$

Stated in words, the weight to be applied to a given estimate is equal to the sum of the elements in the corresponding row (or column) of the inverted covariance matrix divided by the sum of all the elements of the inverted matrix.

The resulting variance can be calculated as well:
$\text{Var}(X_c)=(\mathbf{A}^{-1}\mathbf{e}/(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e}))'\text{A}\,(\mathbf{A}^{-1}\mathbf{e}/(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e}))$
$= (\mathbf{e}'\mathbf{A}^{-1}\mathbf{A}\mathbf{A}^{-1}\mathbf{e})/\,(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^2$
$= (\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})/\,(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})^2$
$=1/\,(\mathbf{e}'\mathbf{A}^{-1}\mathbf{e})$

Stated in words, the minimum variance is equal to the inverse of the sum of the elements of the inverted covariance matrix.

As an example, consider two estimates of total future loss payments with standard deviations of their errors \$1 million and \$2 million, with no correlation between the errors.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & .25 \end{bmatrix}$$

$$\boldsymbol{w} = \begin{pmatrix} \frac{1}{1.25} \\ \frac{.25}{1.25} \end{pmatrix} = \begin{pmatrix} .8 \\ .2 \end{pmatrix}$$

Obviously, considerably more weight is given to the estimate with lower uncertainty. The resulting standard deviation of the combined estimate is sqrt(1/1.25) = \$0.89 million, less than the best individual estimate, reflecting the benefit of the second method, even though it is inferior to the first.

Consider the case where the correlation coefficient between the two errors is estimated to be 0.1. Then:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 4 \end{bmatrix}$$

$$\boldsymbol{A^{-1}} = \begin{bmatrix} 1.0101 & -0.0505 \\ -0.0505 & 0.2525 \end{bmatrix}$$

$$w = \begin{pmatrix} \frac{0.9696}{1.1616} \\ \frac{0.2020}{1.1616} \end{pmatrix} = \begin{pmatrix} 0.8261 \\ 0.1739 \end{pmatrix}$$

The positive correlation between the errors of the two estimates resulted in more weight being given to the best individual estimate, since the other estimate provides less independent information than in the other example. The standard deviation of the combined estimate is now sqrt(1/1.1616) or $.93 million, a more uncertain result than the case with independent estimates.

Now consider the case where the correlation coefficient is equal to 0.5

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 4 \end{bmatrix}$$

$$\boldsymbol{A^{-1}} = \begin{bmatrix} 1.3333 & -0.3333 \\ -0.3333 & 0.3333 \end{bmatrix}$$

$$w = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In this case, all weight is given to the first estimate. The benefit of independent information from the second estimate is being exactly offset by the increased covariance with the first estimate. The standard deviation of the combined estimate in this case is of course $1 million.

An interesting result occurs when we consider the case of a correlation coefficient equal to .75:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1.5 \\ 1.5 & 4 \end{bmatrix}$$

$$\boldsymbol{A^{-1}} = \begin{bmatrix} 2.2857 & -.8571 \\ -0.8571 & 0.5714 \end{bmatrix}$$

$$w = \begin{pmatrix} {}^{1.4286}/_{1.1429} \\ {}^{-0.2857}/_{1.1429} \end{pmatrix} = \begin{pmatrix} 1.25 \\ -.25 \end{pmatrix}$$

In this case greater than 100% weight is given to the first estimate and negative weight is given to the second estimate. In other words *the combined estimate is outside the individual estimates.* The standard deviation of the combined estimate is sqrt(1/1.1429) = $.94 million, an improvement over using the first estimate alone. While this result is somewhat counterintuitive at first glance, this can be explained by thinking about the second estimate as *indicating* the errors of the first estimate in advance. When the first estimate is too low, the second estimate will generally be too low by an even larger amount. When the first estimate is too high, the second estimate will generally be too high by an even large amount. So if the second estimate is less than the first estimate, this is seen as an indication that *both* estimates are probably too low. If the second estimate is greater than the first estimate, it is an indication that *both* estimates are probably too high. The greater the strength of this correlation, and the greater the difference between the uncertainty of the estimates, the greater this signaling effect will be. In fact, if the estimates are fully correlated, *any* difference in the uncertainty of the two estimates will imply an exact (i.e. certain) estimate[4]. In practical reality actuaries are hesitant to select an estimate that is outside of the range of their indications from individual methods. This would require putting a great deal of faith in the estimated correlation between methods. Cases where such an indication of negative weighting occurs can be useful information, however, potentially leading the actuary to looking for a common source between the methods for their errors, with the hope of explicitly adjusting methods to remove the error, or to consider with greater intensity situations where the estimates are very different from each other as to whether such signaling may be occurring.

If it is decided that negative weights are not allowed, the solution to the variance minimization problem may lie on one of the boundaries (i.e. where one or more of the weights equals zero). With the weight for a particular estimate equal to zero, the problem can be thought of as simply ignoring that estimate entirely (eliminating its corresponding row and column from the covariance matrix). By systematically removing estimates from the problem, and optimizing, all the potential solutions can be found, and the global minimum determined.

For the case of the two estimate example shown above, the non-negative solution is trivial (100% weight to the first estimate). An example of the technique will be revisited later in this paper with a four estimate example.

[4] The combined estimate in this case is equal to $(\alpha X_1 - X_2)/(\alpha - 1)$ where α = sqrt(σ22/σ11). This can not be calculated directly using the minimized solution formulas in this paper, but can be proven by taking the limit as the correlation coefficient approaches 1.

Estimating the uncertainty an individual reserve estimate

Until this point we have taken the covariance matrix of reserve estimates as a given. Estimating the uncertainty of such an estimate is not a trivial exercise. We could simply observe the errors of methods as of a particular point of loss development, compared to what eventually emerged when the body of claims is more or less fully developed. However, as with the general problem of developing the estimate itself, the complicating factor is that by the time a cohort of claims is considered fully developed, it can be quite old, with the potential for significant changes in the underlying environment and exposure since the time the claims were incurred or reported. For this reason, as well as for the reason of not wanting to exclude *any* meaningful data, immature data is typically used to assist in the estimation of the parameters of loss development (e.g. age-to-age factors combined together to create age-to-ultimate factors). Similarly, in order to fully use the immature data for the purposes of measuring the potential uncertainty around an estimate, the piecewise emergence of this uncertainty over time can be considered, and combined together to a single measure of uncertainty.

Measurement of such uncertainty has been discussed by a number of authors[5]. One complicating factor often left unaddressed is the potential for correlation between development periods. Using an example of loss development factor approaches, an example of negative correlation would be when an individual large claim emerges as an isolated incident. The observed development in the period where the claim emerges, relative to the previous period is larger than typical for that development period. Subsequent development, relative to losses at the previous period, is lower than typical. An example of positive correlation between development periods is where two groups of claims, one that develops more rapidly than another are combined together into a single triangle, where the mix between the two changes over time. The accident periods which contain more slow-developing losses will have relative development that is consistently greater for multiple development periods than those which have more rapidly developing losses.

[5] A few of the papers on this topic include: Thomas Mack, "The Standard Error Of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor", *ASTIN Bulletin* Vol. 29. No. 2, 1999, pp 361-366; Daniel M. Murphy, "Chain Ladder Reserve Risk Estimators", *CAS E-Forum* Summer 2007; Pinheiro, Andrade e Silva, & Centeno, "Bootstrap Methodology in Claim Reserving".

Consider the following triangle of cumulative paid loss development:

	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	67,280	176,011	297,430	370,321	460,803	492,379	525,096	547,416	551,408	554,232
Acc Yr 2	46,242	175,038	300,846	433,181	577,144	663,648	704,047	719,313	782,765	
Acc Yr 3	36,758	182,030	372,771	547,660	651,400	714,248	784,677	842,825		
Acc Yr 4	76,464	302,499	578,500	796,966	888,212	1,088,212	1,209,348			
Acc Yr 5	85,376	319,303	563,223	835,815	1,125,921	1,237,492				
Acc Yr 6	64,905	387,337	781,353	1,131,217	1,352,403					
Acc Yr 7	47,353	229,982	466,113	741,354						
Acc Yr 8	38,509	220,996	404,076							
Acc Yr 9	41,188	217,770								
Acc Yr 10	52,371									

	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-ult
Selected Age-Age Factors:	4.967	1.914	1.496	1.229	1.113	1.106	1.050	1.030	1.020	1.040
Age-Ultimate Factors	24.685	4.970	2.597	1.736	1.412	1.269	1.147	1.093	1.061	1.040

"Selected" age to age factors have been included to illustrate that estimated future development is not necessarily a specific simple or weighted average of observed development, but often reflect significant professional judgment on the part of the actuary.

The indicated ultimate loss for Accident Year 10 is 32,371 * 24.685 = 1,292,769. What is the uncertainty surrounding this estimate as of Age 1? We cannot directly observe this uncertainty. We have ten observed Accident Years as of Age 1, but we do not have any "observations" of the ultimate loss for these same accident years. The closest we have is an observation at Age 10 (Accident Year 1). However, similar to the link ratio development method itself, we can observe the behavior of our errors piece-wise over time.

Starting with the loss at Age 1 for each Accident Year, predictions of cumulative paid loss at each age of development can be made by chaining through the appropriate selected age-to-age factors.

	Predictor	Predictions								
	Age 1	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Ultimate
Acc Yr 1	67,280	334,180	639,620	956,872	1,175,995	1,308,883	1,447,624	1,520,005	1,565,606	1,596,918	1,660,794
Acc Yr 2	46,242	229,684	439,615	657,664	808,269	899,604	994,962	1,044,710	1,076,051	1,097,572	1,141,475
Acc Yr 3	36,758	182,577	349,452	522,781	642,498	715,100	790,900	830,445	855,359	872,466	907,364
Acc Yr 4	76,464	379,797	726,931	1,087,489	1,336,523	1,487,551	1,645,231	1,727,493	1,779,317	1,814,904	1,887,500
Acc Yr 5	85,376	424,063	811,656	1,214,237	1,492,297	1,660,927	1,836,985	1,928,834	1,986,700	2,026,434	2,107,491
Acc Yr 6	64,905	322,383	617,041	923,094	1,134,482	1,262,679	1,396,523	1,466,349	1,510,339	1,540,546	1,602,168
Acc Yr 7	47,353	235,202	450,177	673,465	827,689	921,218	1,018,867	1,069,810	1,101,904	1,123,942	1,168,900
Acc Yr 8	38,509	191,274	366,099	547,684	673,103	749,164	828,576	870,004	896,104	914,027	950,588
Acc Yr 9	41,188	204,581	391,568	585,785	719,930	801,282	886,218	930,529	958,445	977,614	1,016,718
Acc Yr 10	52,371	260,127	497,883	744,832	915,399	1,018,839	1,126,836	1,183,178	1,218,673	1,243,047	1,292,769

The cumulative error terms can be now be observed by subtracting the corresponding observation of actual cumulative paid loss, where it is known.

Errors:	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	158,169	342,190	586,551	715,192	816,504	922,528	972,589	1,014,198	1,042,686
Acc Yr 2	54,646	138,769	224,483	231,125	235,956	290,915	325,397	293,286	
Acc Yr 3	547	(23,319)	(24,879)	(8,902)	852	6,223	(12,380)		
Acc Yr 4	77,298	148,431	290,523	448,311	399,339	435,883			
Acc Yr 5	104,760	248,433	378,422	366,376	423,435				
Acc Yr 6	(64,954)	(164,312)	(208,123)	(217,921)					
Acc Yr 7	5,220	(15,936)	(67,889)						
Acc Yr 8	(29,722)	(37,977)							
Acc Yr 9	(13,189)								

The loss that would have been predicted for Accident Year 1 at Age 2 using the loss amount at Age 1, and the selected development factor from Age 1 to Age 2 is larger than the actual observed loss at Age 2 by 158,169. By Age 10, the error (based on the prediction at Age 1) has grown to 1,042,686.

It is useful to view these error terms relative to a base amount to reflect changes in volume. A natural choice is the value of the predictor variable, in this case the loss at Age 1.

Relative Errors:	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	2.35	5.09	8.72	10.63	12.14	13.71	14.46	15.07	15.50
Acc Yr 2	1.18	3.00	4.85	5.00	5.10	6.29	7.04	6.34	
Acc Yr 3	0.01	(0.63)	(0.68)	(0.24)	0.02	0.17	(0.34)		
Acc Yr 4	1.01	1.94	3.80	5.86	5.22	5.70			
Acc Yr 5	1.23	2.91	4.43	4.29	4.96				
Acc Yr 6	(1.00)	(2.53)	(3.21)	(3.36)					
Acc Yr 7	0.11	(0.34)	(1.43)						
Acc Yr 8	(0.77)	(0.99)							
Acc Yr 9	(0.32)								
RMSE (-1)	1.176	2.778	4.906	6.178	7.082	8.711	11.369	15.074	
RMSE (all)	1.114	2.622	4.574	5.804	6.711	8.064	9.284	11.564	15.498
RMSE Growth		2.230	1.646	1.183	1.086	1.139	1.066	1.017	1.028

Three summarizing rows have been added to the bottom of this table. The first shows the root mean squared error (RMSE) across all accident years with the exception of the most recent year. The second row shows the RMSE for all years. Organizing the data in this way allows us to calculate a RMSE growth factor from age to age (row three). For example, the RMSE at Age 2 for Accident Years 1 through 8 was 1.176. The RMSE at Age 3 for the same group of accident years grew to 2.622, implying a growth factor of 2.622/1.176 = 2.230. By using these growth factors we can take advantage of all the information in the triangle (including immature accident periods), and also reflect correlation *across development periods*.

The factor to apply to the loss at Age 1 in order to arrive at the estimated RMSE of the Ultimate loss is calculated as:

1.114 * 2.230 * 1.646 * … * 1.028 * 1.046 = 6.976

This was arrived at by multiplying the Age 2 RMSE factor for all years (1.114) by the growth factors, including a growth factor to take us from Age 10 to ultimate (1.046). This final factor was selected based on the size of the selected tail factor of 1.04. Since this tail factor is greater than the size of the 9-10 factor, but less than the size of the 8-9 factor times the 9-10 factor, we multiplied the RMSE growth factors for ages 8-9 and ages 9-10 to arrive at this factor.

This same procedure can be applied for each of the points of development:

- Project all future loss amounts based on a common point of development.
- Subtract observed amounts from the predicted amounts.
- Relate the error terms back to the predictor variable.

- Calculate the RMSE and growth in RMSE from age to age.
- Combine RMSE and RMSE growth factors to arrive at a factor to the predictor variable.

Summarizing the results for predicting at Age 2:

	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10	Ultimate
Acc Yr 1	0.224	0.759	0.901	1.119	1.349	1.438	1.552	1.630	
Acc Yr 2	0.195	0.389	0.222	0.125	0.310	0.439	0.213		
Acc Yr 3	(0.134)	(0.145)	(0.059)	(0.007)	0.021	(0.082)			
Acc Yr 4	0.002	0.229	0.583	0.319	0.334				
Acc Yr 5	0.150	0.246	(0.007)	0.041					
Acc Yr 6	(0.103)	(0.057)	0.028						
Acc Yr 7	(0.113)	(0.360)							
Acc Yr 8	0.086								
RMSE (-1)	0.147	0.380	0.491	0.585	0.799	1.063	1.552		
RMSE (All)	0.141	0.377	0.448	0.524	0.712	0.870	1.108	1.630	
RMSE (Growth)		2.556	1.180	1.067	1.216	1.088	1.042	1.050	1.094
Factor to Predictor	0.141	0.361	0.426	0.455	0.553	0.602	0.627	0.658	**0.720**

Note that the RMSE growth factors are *not* the same as the growth factors when looking at predictions from Age 1, even for the same points of development.

Continuing this method for each of the development ages gives:

	Cumulative Paid Loss	Age-Ult Loss Development Factor	Indicated Ultimate Loss	Indicated RMSE Factor	Selected RMSE Factor	Estimated RMSE	Reserve	CV of Reserve	CV of Ultimate
Acc Yr 1	554,232	1.040	576,401		0.100	55,423	22,169	2.50	0.10
Acc Yr 2	782,765	1.061	830,357		0.120	93,932	47,592	1.97	0.11
Acc Yr 3	842,825	1.093	920,891	0.210	0.150	126,424	78,066	1.62	0.14
Acc Yr 4	1,209,348	1.147	1,387,431	0.112	0.170	205,589	178,083	1.15	0.15
Acc Yr 5	1,237,492	1.269	1,570,209	0.052	0.200	247,498	332,717	0.74	0.16
Acc Yr 6	1,352,403	1.412	1,909,926	0.166	0.250	338,101	557,523	0.61	0.18
Acc Yr 7	741,354	1.736	1,286,731	0.404	0.300	222,406	545,377	0.41	0.17
Acc Yr 8	404,076	2.597	1,049,197	0.368	0.500	202,038	645,121	0.31	0.19
Acc Yr 9	217,770	4.970	1,082,265	0.720	1.000	217,770	864,495	0.25	0.20
Acc Yr 10	52,371	24.685	1,292,769	6.976	6.000	314,226	1,240,398	0.25	0.24

The actual RMSE factors to be applied to the cumulative paid loss amounts were *selected* based on the indicated RMSE factors. Just as in estimating the mean, judgment can play an important role. It is intuitive that the RMSE factor should decrease with the age of the accident period. It is also intuitive that the coefficient of variation (CV) of reserves should increase as the age of the accident period increases and that the CV of ultimate losses should decrease.

Combining accident year estimates of uncertainty, including correlation effects

In addition to correlation between development periods within an accident period, correlation can exist regarding future loss payments between accident periods. Note that we are not talking about correlation between losses that have already occurred, but only the estimates of the amounts *yet to be paid*. These correlations can be brought about by changes in claim handling or other environmental impacts (judicial environment, inflation, etc.).

One way to estimate such correlations is by observing the errors in incremental amounts across the triangle:

Prediction as of the prior period

	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	334,180	336,885	444,955	455,125	512,874	544,571	551,351	563,838	562,436
Acc Yr 2	229,684	335,023	450,066	532,379	642,361	733,995	739,249	740,892	
Acc Yr 3	182,577	348,405	557,665	673,074	725,008	789,958	823,911		
Acc Yr 4	379,797	578,983	865,436	979,471	988,580	1,203,562			
Acc Yr 5	424,063	611,146	842,582	1,027,217	1,253,150				
Acc Yr 6	322,383	741,363	1,168,904	1,390,266					
Acc Yr 7	235,202	440,186	697,305						
Acc Yr 8	191,274	422,986							
Acc Yr 9	204,581								

Predictive Error

	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	158,169	39,455	74,634	(5,678)	20,495	19,475	3,935	12,430	8,204
Acc Yr 2	54,646	34,177	16,885	(44,765)	(21,287)	29,948	19,936	(41,873)	
Acc Yr 3	547	(24,366)	10,005	21,674	10,760	5,281	(18,914)		
Acc Yr 4	77,298	483	68,470	91,259	(99,632)	(5,786)			
Acc Yr 5	104,760	47,923	6,767	(98,704)	15,658				
Acc Yr 6	(64,954)	(39,990)	37,687	37,863					
Acc Yr 7	5,220	(25,927)	(44,049)						
Acc Yr 8	(29,722)	18,910							
Acc Yr 9	(13,189)								

The error terms here are calculated using the prior period's loss amount as the predictive variable. We can now compare correlation of neighboring accident years at the same point in time. For example we can pair the error from Acc Yr 2 in Age 2 with the error from Acc Yr 3 in Age 3 since they are subject to common development environments (i.e. same calendar period). In this way we can look at many pairs:

54,646	39,455
547	34,177
77,298	(24,366)
104,760	483
(64,954)	47,923
5,220	(39,990)
(29,722)	(25,927)
(13,189)	18,910
34,177	74,634
(24,366)	16,885
483	10,005
47,923	68,470
(39,990)	6,767

Etc.

We can then calculate a correlation coefficient. In this case, comparing adjacent accident periods only, we arrive at a correlation coefficient of 0.100. With only 36 pairs of observations, this result is not statistically significantly different from zero[6]. We can similarly compare accident periods with those two periods away, three periods away, etc.

Acc Period Difference	Measured Correlation	Number of Pairs	Significance Level	Selected Correlation
1	0.1001	36	0.5497	0
2	-0.1296	28	0.4947	0
3	-0.0876	21	0.6911	0
4	0.1991	15	0.4435	0
5	0.1174	10	0.7164	0
6	**-0.7491**	6	**0.0324**	0
7	0.1125	3	0.8571	0

The only correlation that was measured as statistically significant from zero was the correlation between accident periods 6 away from each other. Intuitively, there is little justification for this correlation given that the nearer period comparisons show no correlation.

We subjectively determined, for this data, that the correlations should

- Not be negative
- Not be larger than a correlation for a nearer accident period
- Only be used if statistically significant (otherwise set to zero)

The RMSE of our estimated ultimate loss, in total, can now be estimated by using the correlation matrix and accident year specific RMSE estimates:

σ = sqrt(**x'Cx)**

In this case the correlation matrix is simply the identity matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

And the total RMSE is therefore $694,376 which is simply the square root of the sum of the squared values for each accident period for this specific case.

[6] Using a t-test where t = ρ*sqrt(n)/sqrt(1-ρ^2) with n degrees of freedom.

Generalization of this Technique

This technique can be generalized to be used with other loss development methods. The generalization to using case-incurred data instead of paid data is trivial. The potential for case savings does not pose problems, since no distributional form is assumed. The generalization for the Bornhuetter-Ferguson method is relatively straightforward. The loss development from a particular point to a particular point is simply a factor to earned premium that is a function of the seed loss ratio and the portion of the overall development expected to occur over that time period. The logical base for the RMSE factor is the premium for the year times the seed loss ratio for the year.

For this example, we have calculated the following estimates of reserve standard deviation under each of the four common methods of loss development:

Paid LDF	694,376
Inc LDF	1,280,230
Paid B-F	1,273,936
Inc B-F	707,665

Correlation between the Methods

Similar to the method for looking at accident year correlation, we can estimate the correlation between the various reserve estimates by looking at corresponding pairs of errors, in this case matching up by accident and development period for two methods to be compared.

Predictive Error - Paid LDF

	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	158,169	39,455	74,634	(5,678)	20,495	19,475	3,935	12,430	8,204
Acc Yr 2	54,646	34,177	16,885	(44,765)	(21,287)	29,948	19,936	(41,873)	
Acc Yr 3	547	(24,366)	10,005	21,674	10,760	5,281	(18,914)		
Acc Yr 4	77,298	483	68,470	91,259	(99,632)	(5,786)			
Acc Yr 5	104,760	47,923	6,767	(98,704)	15,658				
Acc Yr 6	(64,954)	(39,990)	37,687	37,863					
Acc Yr 7	5,220	(25,927)	(44,049)						
Acc Yr 8	(29,722)	18,910							
Acc Yr 9	(13,189)								

Predictive Error - Incurred LDF

	Age 2	Age 3	Age 4	Age 5	Age 6	Age 7	Age 8	Age 9	Age 10
Acc Yr 1	63,856	64,318	51,033	4,630	10,993	23,402	(10,667)	3,778	75
Acc Yr 2	65,030	122,853	441	(50,411)	22,416	10,133	(18,730)	(3,184)	
Acc Yr 3	(76,509)	46,664	(28,013)	25,992	29,096	(44,162)	30,221		
Acc Yr 4	3,836	33,019	18,569	23,737	(50,475)	10,238			
Acc Yr 5	(64,145)	(50,898)	(47,898)	(6,952)	(11,110)				
Acc Yr 6	(116)	(39,177)	82,700	2,547					
Acc Yr 7	70,916	5,857	(77,161)						
Acc Yr 8	1,150	3,682							
Acc Yr 9	(63,861)								

Each of the tables above is based on predictions based on loss as of the prior period. Matching up pairs of corresponding values (ex. Paid Error for AY 1, Age 2 with Incurred Error for AY 1, Age 2) we measure a correlation coefficient of 0.345.

We calculate the full correlation matrix using the standard four methods as:

	Paid LDF	Inc LDF	Paid BF	Inc BF
Paid LDF	1.000	0.345	0.604	0.388
Inc LDF	0.345	1.000	0.145	0.669
Paid BF	0.604	0.145	1.000	0.581
Inc BF	0.388	0.669	0.581	1.000

The result is fairly intuitive. Paid LDF and Paid BF methods are highly correlated because they both rely on paid development factors. Incurred LDF and Incurred BF methods are similarly correlated. The two Bornhuetter-Ferguson methods are correlated with each other because they rely on the same seed loss ratio parameters. Already you can quickly see that it is somewhat false comfort if all of your indications are aligned. If they are highly correlated, you are simply seeing slightly different variations on the same theme.

Optimized Weights

The covariance matrix between the four methods is:

	Paid LDF	Inc LDF	Paid BF	Inc BF
Paid LDF	4.82E+11	3.06E+11	5.34E+11	1.91E+11
Inc LDF	3.06E+11	1.64E+12	2.37E+11	6.06E+11
Paid BF	5.34E+11	2.37E+11	1.62E+12	5.24E+11
Inc BF	1.91E+11	6.06E+11	5.24E+11	5.01E+11

The inverted matrix is:

3.949E-12	-1.036E-12	-1.611E-12	1.433E-12
-1.036E-12	1.591E-12	9.097E-13	-2.483E-12
-1.611E-12	9.097E-13	1.767E-12	-2.335E-12
1.433E-12	-2.483E-12	-2.335E-12	6.900E-12

Which suggests the following optimal weights:

Paid LDF	69%
Inc LDF	-26%
Paid BF	-32%
Inc BF	89%

This weighting gives more than 100% weight to a combination of Paid LDF and Inc BF estimates, with negative weights to the other two methods. This is due to the lower estimated variability of the Paid LDF and Inc BF methods and the fact that the estimates with higher variability are highly correlated with the other estimates. The resulting estimated standard deviation is $502,340, which is lower than for the best single-method estimate.

The optimal non-negative weighting is given by:

Paid LDF	52%
Inc LDF	0%
Paid BF	0%
Inc BF	48%

About equal weight is given to the Paid LDF and Incurred BF estimates with this weighting, with no weight given to the other two methods. The interpretation is that giving any weight to the other two methods would decrease the predictive accuracy because of their added volatility.

This is the result for a specific example. Other data would result in different weights. Different factor or seed loss ratio selection would result in different weights as well.

<u>Summary</u>

When considering the weight to be given to different indications of property casualty loss reserve estimates, one should consider reducing uncertainty to be an important goal. Estimating this uncertainty for each method, as well as correlations between methods provides the opportunity to calculate weights that result in a combined estimate with a minimized variance. Understanding *why* such optimized weights are indicated is of value for understanding the ramifications of choosing between various methods, even if such weights are not actually relied upon.