Applications of Reserve Ranges and Variability in Practice

Christopher Walker, FCAS, MAAA

Mark Littmann, FCAS, MAAA

Abstract

Motivation. Actuaries are faced with increased questions on reserve variability and "reasonable ranges." As consultants Mr. Littmann and Mr. Walker confront these questions frequently. We believe a survey of current uses of ranges, how actuaries may adequately address "reasonableness" in a way that is approachable to users, and how these methods may be bridged with more theoretical methods would be beneficial to actuaries and other interested parties.

Method. We present examples of sensitivity testing and how, using a consistent data set, illustrative conclusions on how range estimates may be derived. We also extend the examples with the application of the Mack technique for evaluating a distribution of possible outcomes and investigate the potential relationship between a range of reasonable estimates with distributions of possible outcomes.

Results. Our "results" are primarily illustrations that give the practitioner easy to develop ranges and may also provide a framework for application to a company's aggregate reserve position and "reasonable range."

Conclusions. "Actuarial judgment" will not satisfy the questions being asked of actuaries on ranges and variability. Actuaries need to have a structure in place to provide evidence/illustrations of variability.

Keywords. Reserve Variability; Reserve Ranges; Thomas Mack, sensitivity testing, distribution of possible outcomes.

1. Introduction

There has been significant time and effort spent by our colleagues in the CAS and elsewhere in the development of models and approaches for quantifying variability in loss reserve estimates. Many of these models, whether parametric or non-parametric, are quite complex and identify important theoretical issues, such as correlation between coverage lines, etc. However, practitioners are often constrained by data, time, and budget to apply these models. Furthermore, the contexts in which the reserve estimates are applied, such as financial reporting requirements of "best estimates," often limit or eliminate the usefulness of variability model outputs.

The goal of this paper is to emphasize the increasing importance of not only recognizing variability in reserve estimates, primarily through the assessment of a "reasonable range" around an actuarial central estimate, but also in providing understandable support to the variability, or range, assessment. We believe that the role of "actuarial judgment" in the construction of reserve ranges without specific support has been significantly diminished, as users seek more quantitative evidence in support of the asserted range.

This paper begins with a discussion of areas where the concept of reasonable variability and reasonableness is commonly encountered or may be emerging. We will also provide a brief discussion of two theoretical approaches to reserve variability presented in the literature, and also the concept of sensitivity testing that is not cited in the literature but is commonly used. Next, approaches commonly used by practitioners are discussed, along with advantages and disadvantages. We then provide some simple illustrations of the application of sensitivity testing to form a view on a range of reasonable estimates, and also apply a stochastic model to evaluate the potential relationship among reasonable ranges and distributions of outcomes. We close with observations on simple techniques to address variability when aggregations of reserve segments are considered.

2. Business Applications of Variability Concepts

For those familiar with property/casualty insurance, the inherent uncertainty in the ultimate settlement value of unpaid claim liabilities is self-evident. The uncertainty arises from the fact that all events and conditions affecting the ultimate settlement value of claims have not yet occurred, and for various coverages there may be a significant portion of unasserted claims at a particular point in time. Other key sources of uncertainty include future inflation rates on the costs of goods and services, future attitudes of claimants and juries to potential valuations of damages, and future law and legal rulings that may be retrospective in nature.

Thus, actuaries evaluate historical data to assess reporting and payment patterns in order to make projections of ultimate losses and the associated unpaid claim liabilities. The actuaries may generate a point-estimate, a range of reasonable estimates, or a distribution of possible outcomes, depending on the purpose and intended use of the actuarial analysis.

For financial reporting purposes, businesses do not report "ranges" of income or "ranges" of balance sheets. Loss reserves presented on a traditional Generally Accepted Accounting Standards (GAAP) basis or a Statutory Accounting Principles (SAP) basis are a unique number, not a range of estimates. However, users of such reports, such as the SEC and investors, increasingly require not just disclosure of the carried amounts but also the relative uncertainty in these estimates. This trend leads to the discussion of reserve variability and reserve ranges.

Examples of additional disclosures include the following:

2.1 Statements of Actuarial Opinion and Actuarial Opinion Summary

State regulators require a discussion of the business and its qualities that may introduce variability into the carried loss reserves being opined upon in the Statement of Actuarial Opinion (SAO). Regulators also require that an assessment be made of the risk of material adverse deviation (RMAD) in the recorded loss reserves. The variability issue also flows to the Actuarial Opinion Summary (AOS) where the Opining actuary has the option of listing a point estimate of the reserves or a range of reserves. In determining this range, actuaries are guided by Actuarial Standard of Practice (ASOP) No. 36, *Statements of Actuarial Opinion Regarding Property/Casualty Loss and Loss Adjustment Expense Reserves*, which states that an

> "actuary should consider a reserve to be reasonable if it is within a range of estimates that could be produced by an unpaid claim estimate analysis that is, in the actuary's professional judgment, consistent with both ASOP No. 43, *Property/Casualty Unpaid Claim Estimates*, and the identified stated basis of reserve presentation."

In turn, ASOP No. 43 describes considerations to be taken into account by the actuary in the analysis, aligned with the intended purpose. In this context, this "range of reserves" concept has no set definition, but it is generally approached from the standpoint of a range of reasonable estimates as opposed to a distribution of possible outcomes.

2.2 Securities and Exchange Commission filings

A fairly common disclosure in 10-K reporting is a discussion of the analysis that developed the carried reserve and the variability inherent in that estimate. For many years, the filings have included a table that shows the development of reserve estimates recorded at prior reporting dates, updated for the subsequent valuations that have occurred up to the latest reporting date. (This disclosure is comparable to the development of accident year ultimate loss and Defense & Cost Containment (DCC) estimates presented in Schedule P, Part 2, of the statutory basis Annual Statement.) The Securities and Exchange Commission (SEC) was particularly active in questioning insurers several years ago in light of filers posting large reserve redundancies/deficiencies related to prior accident years in their financial statements. Not only did the SEC focus on management's development of its "best estimate" but it also required discussion that would help investors better understand the risks and uncertainties that are inherent in that estimate and in the business as a whole. In response, the registrants have expanded their disclosures, both qualitatively and quantitatively. For example, one company with which we are familiar dedicates over twelve pages (out of 110 in total) of its 10-K filing to loss reserves and includes specific discussions on "Significant Risk Factors," "Determination of Best Estimate," and "Reserve Sensitivities." Within these categories, and from other companies filings, we have observed a variety of quantitative disclosures on variability such as confidence levels associated with low, reasonable, and high values, assessments of the impact of changes in key assumptions such as tail factors, frequency/severity, and/or inflation, and the prospective performance of reinsurance programs.

2.3 Financial Audits

Even when regulatory reporting is not a top consideration, such as in the financial reporting of privately held non-insurance companies, there is considerable focus by auditors on estimates, including loss reserves. A typical situation could be illustrated as a privately held manufacturer that may choose to self-insure its Product Liability exposure and is required to recognize a reserve for the unpaid losses incurred as of the accounting date. The auditors are faced with assessing the reasonableness of the reserve established. In performing such assessments, the auditors may be faced with a wide range of circumstances and assumptions that may or may not make sense for the situation. Auditors recognize that there may be differences in their point of view and those taken by the company; however, they are frequently faced with the dilemma of "how much difference is too much" and have generally applied formulaic materiality thresholds. Such thresholds, such as a -5%/+5% or -10%/+10% range around an independent reserve estimate, may or may not have a theoretical basis.

Auditors often rely on other measures, such as a balance's materiality to the financial statements as a whole, and this fact may render differences between the auditor's unpaid claim estimates and carried values moot; however, for insurance companies, loss and LAE reserves are usually the largest and most material liability on the balance sheet and, therefore, small differences between estimates and carried reserves may be highly material to the audit as a whole.

2.4 Mergers and Acquisitions

In both insurance and non-insurance transactions, there may be significant unpaid liabilities involved, some highly uncertain. For example, many insurers and manufacturers have legacy asbestos liabilities that continue to be a drain on current earnings. In these cases, an understanding and quantification of the potential for continued development may be a prime consideration as there may be a "true up" (evaluation several years after the close of the deal) involved or the purchase of third-party insurance/reinsurance as a condition of completing the transaction. In these cases, the idea of a reserve confidence interval or a range of possible outcomes, as opposed to a range of reasonable estimates, may be highly relevant metrics for pricing reinsurance and establishing horizons for "true ups."

2.5 Internal Revenue Service Considerations

The Internal Revenue Service (IRS) disallows explicit loss reserve margins for the calculation of insurance company federal income tax. The IRS may review various sources of documentation, such as the report of the appointed actuary or findings from the external auditor, in its audit to evaluate whether the recorded reserves included a margin, either explicit or implicit. The IRS also appears to be performing more detailed independent reviews of loss reserves even in the absence of explicit margins to support its audits. While we are not aware of definitive guidance, and as of this writing there are various on-going legal challenges, the quality of an analysis of a "reasonable range of estimates" could ultimately factor into IRS positions on company reserve redundancies.

3. Variability Concepts in the Literature

The literature of the CAS includes a broad range of papers and presentations on the topics of variability and distribution of reserve estimates. However, these generally do not offer any guidance for which portion of the distribution would constitute a reasonable range of estimates.

3.1 Thomas Mack Method

This approach is a "distribution free" technique to measure the variability of reserve estimates generated by a traditional application of the loss development (aka, chain ladder) method to a typical loss development triangle. The technique is relatively easy to apply, including being available in an Excel spreadsheet template that is publicly available on the CAS website or in commercially-available software. This approach yields an estimate for the estimated standard error (ESE) of a distribution of unpaid claim outcomes.

3.2 Boot Strapping

A basic premise of this approach is that the available data (typically a loss development triangle) is essentially "one" observation from a distribution of possibilities. Thus, the technique assumes there is a singular loss development pattern which is indicative of the "true" pattern, and views the data as random observations from this true pattern. Thus, the user prepares an estimate of the "true" pattern, often based on an average of the observed link ratios in the historical data, and then evaluates the residuals, being the differences between the actual observations and the observations that would be consistent with the true pattern. In this way, Boot Strapping is a method of re-sampling that allows the user to make inferences on the variability of mean values and distribution of possible outcomes. It is essentially a simulation process that requires many iterations (say, 1000) and the output is a distribution of possible outcomes. Similar to the Mack approach, there is no guidance as to how "a range of reasonable estimates" may compare to the derived distribution of possible outcomes produced via Boot Strapping.

3.3 Sensitivity Testing

While not a frequent subject described in the actuarial reserving literature, we believe that sensitivity testing[1] is one of the most prevalent approaches used to establish ranges of reasonable reserve estimates. This approach is not technically advanced, nor do we consider it to be a distinct method. Sensitivity testing essentially means that an actuary tests the effects of alternate judgments for the key parameters of the chosen method(s) in order to evaluate alternate low and high estimates of the unpaid claim liabilities. Thus, the relative ease by which the approach can be explained is a distinct advantage.

ASOP No. 36 and ASOP No. 43 implicitly acknowledge the concepts of sensitivity testing. ASOP No. 36 recognizes that an actuary may consider a reserve to be reasonable if it is within a range of estimates derived from appropriate methods and reasonable assumptions. ASOP No. 43 recognizes that an "actuarial central estimate" is an "expected value over a range of reasonably possible outcomes."

Considering that alternate methods may be considered appropriate for a particular unpaid claims analysis, and that alternate assumptions for the key parameters of the methods may be considered reasonable, sensitivity-testing is a natural indicator of reserve variability determined by replicating the collection of methods applied to various data sets and substituting high and low selections for the key parameters of the actuarial analysis. The most commonly-applied actuarial methods are the loss development (chain ladder) method and the Bornhuetter-Ferguson method, which is a blending of the loss development and expected loss method. These methods' key parameters are loss development link ratios, tail factors, and *a priori* expected loss ratios. These parameters can be modified to simulate the underlying drivers of reserve variability. For example, loss development factors/tail factors can be increased/decreased to represent inflation higher/lower than that represented in the underlying data. As another example, *a priori* loss ratios can be adjusted to reflect actual rate

[1] The International Actuarial Association recently published a paper on Stress Testing and Scenario Analysis by the Insurance Regulation Committee. While the paper highlights the role of stress testing and scenario analysis to enhance the risk culture of an organization, the framework may be useful for specific consideration of the variability of estimates of unpaid claims.

and schedule rating changes that may have proven to be different that contemplated in pricing or planning projections.

As additional considerations, if the analysis includes a frequency/severity approach whereby component estimates for the number of claims and the average claim value are used to project ultimate claims costs, then varying the assumptions as to the trend rate in claim frequency and/or claim severity reflecting uncertainty in underlying loss cost drivers may be appropriate. In the case of a reserve analysis segment containing minimal historical data, incorporating different external benchmark parameters may serve as reasonable examples of sensitivity. We also consider the actuary's judgment to form a final point-estimate from among multiple preliminary projections to be a key consideration.

In essence, the application of sensitivity testing may require the actuary to perform an analysis three times reflecting low, central, and high estimates, and there may be many ways to reach each estimate. This labor-intensive feature of the approach may be considered a disadvantage. On the other hand, the approach is simple to apply, easy to understand, and, perhaps more importantly, is easily communicated to a third-party in light of central estimates.

4. Approaches in Practice

We observe that the range of approaches that are commonly used in the P&C industry to evaluate potential distributions of outcomes or to evaluate ranges of reasonable estimates is narrower than the range of approaches described in the literature. Quite simply, some of the methods in the literature, while being theoretically and conceptually sound, are difficult to apply in practice and perhaps even more difficult to explain, particularly in a financial reporting context, to the various stakeholders possessing varying degrees of analytical sophistication. We observe that the more technically-robust algorithms, such as development of specific loss distributions, are commonly applied to provide the inputs required for other applications, such as economic capital models.

Stochastic methods such as Boot Strapping or the Thomas Mack (Mack) technique may be used to evaluate distributions of possible outcomes, but we rarely observe these being used to describe ranges of reasonable estimates. These methods evaluate the variability of the historical data in the context of the chosen method for projecting ultimate claim values, while a range of reasonable estimates is more akin to a range of actuarial central estimates, or a range of expected means of the distributions given various parameter assumptions. Neither the Boot Strapping nor Mack approach can, by itself, respond to the question at the heart of financial reporting faced by practitioners: "To what portion of the distribution of outcomes does the range of reasonable estimates align?"

4.1 Judgment

In the context of reasonable ranges, judgment, or "support by experience," is often cited as the basis for an actuary's central estimate and a range of reasonable estimates. In some cases, actuaries or others may resort to "rules of thumb" or "arbitrary" judgments, such as plus or minus 5%, or plus or minus 10%. These judgments reflect merely an assumption as to the variability of the reserve estimates; in some circumstances, such as a financial reporting context, they may also reflect other metrics such as a certain proportion to shareholder equity (policyholder surplus) or net income.

We observe that ranges based on judgment alone are coming under increased scrutiny by external auditors as well as state and federal regulators. The use of "judgment" alone, without substantive analytical or qualitative evidence, is often considered a fallacious appeal to authority.

4.2 Sensitivity-Testing

This method can be used to derive ranges of reasonable estimates, though there is no common "standard" for performing sensitivity tests. However, we have observed some commonalities. For example, workers' compensation variability is often illustrated by adjusting tail factors to represent changing mortality, and property variability may be illustrated by adjusting claim severity to represent inflationary effects.

For situations where the substitution of alternate parameters in traditional actuarial methods may not be appropriate, we also see that illustrations of high and low estimates may reflect the inclusion/exclusion of high-valued events, such as policy limits Products Liability claims, in immature policy years.

5. Illustrations of Sensitivity Testing and Mack-Based Calculations

5.1 Sensitivity Testing

There are several levels at which sensitivity-testing within the framework of a typical analysis of unpaid claims estimates can be applied:

- Evaluate the dispersion of indications from one or more methods applied to one or more types of data. An actuary might elect to evaluate the dispersion of indications for all accident years combined, or for each accident year.
- Evaluate the effect of alternate judgments for the key elements of the methods as applied to the various sets of data, and generally keep the judgment about relative preferences among the methods the same.

Although we include illustrations of both approaches below, we would consider the second approach to be preferred.

The illustrations below are based on a data set that we consider indicative of personal automobile liability development and variability, but not associated with any actual company. The data consists of historical development of paid and reported losses by accident year at annual valuations. The loss development and Bornhuetter-Ferguson methods are applied to both types of data, generating four preliminary estimates of ultimate losses for each accident year. For simplicity, we keep the examples confined to the latest ten accident years, recognizing that actual company data may extend beyond ten years.

Consider the following illustrative preliminary ultimate loss projections shown in:

Table 1

	Projections of Ultimate Losses			
Accident Year (AY)	Loss Development on Paid	Loss Development on Reported	Bornhuetter-Ferguson on Paid	Bornhuetter-Ferguson on Reported
2003	1,127	1,157	1,127	1,157
2004	1,179	1,193	1,179	1,193
2005	1,089	1,119	1,090	1,119
2006	1,128	1,169	1,129	1,169
2007	1,608	1,634	1,603	1,634
2008	1,418	1,466	1,416	1,465
2009	1,430	1,463	1,430	1,463
2010	1,440	1,473	1,456	1,476
2011	1,800	1,782	1,693	1,739
2012	1,597	1,565	1,574	1,564

Using the minimum and maximum of the projections for each accident year for evaluating a potential range of reasonable estimates, the results are shown in Table 2:

Table 2

AY	Minimum	Mean	Maximum
	Projections of Ultimate Losses		
2003	1,127	1,142	1,157
2004	1,179	1,186	1,193
2005	1,089	1,104	1,119
2006	1,128	1,149	1,169
2007	1,603	1,620	1,634
2008	1,416	1,441	1,466
2009	1,430	1,447	1,463
2010	1,440	1,461	1,476
2011	1,693	1,753	1,800
2012	1,564	1,575	1,597
Sum	13,669	13,878	14,074
Inception-to-date Paid	11,690	11,690	11,690
Unpaid Claim Estimate	1,979	2,188	2,385
Difference to Mean	(209)		196
Difference as % Mean	-10%		9%

As shown in Table 2, if the actuary deems each of the projections to be reliable and is indifferent as to their relative merits, and the actuary considers that each year's estimate is independent of the next, the high and low projections are used to evaluate the end-points of a range of reasonable estimates. This approach yields a range of unpaid claim estimates that extends from 10% less than to 9% greater than the mean of the projections.

On the other hand, the actuary might choose to evaluate the dispersion of the projections on an all-years basis for the four projections, as shown in Table 3:

Table 3

	Ultimate Loss Projections			
	Loss Development on Paid	Loss Development on Reported	Bornhuetter-Ferguson on Paid	Bornhuetter-Ferguson on Reported
AY's 2003 - 2012	13,816	14,021	13,698	13,978

	Minimum	Mean	Maximum
Ultimate Loss Projection	13,698	13,878	14,021
Inception-to-date Paid	11,690	11,690	11,690
Unpaid Claims Estimate	2,008	2,188	2,331
Difference to Mean	(181)		143
Difference as % Mean	-8%		7%

The indicated range of unpaid claim estimates based on the all-years approach extends from 8% less than the mean estimate to 7% greater than the mean estimate. Due to the feature that no method consistently generated the highest or lowest of the four projections for each accident year, the range is narrower than on an "each accident year" basis.

These two variations of evaluating a range of reasonable estimates do not, however, reflect the actuary's judgment for the relative reliability and/or predictive value of the various methods and data-types. For example, for the Products Liability line of business, use of the paid loss development method for relatively immature accident years may be inappropriate and subject to extreme variation over time.

Thus, we suggest that a deliberate analysis of low and high estimates using alternate yet reasonable assumptions and judgments is preferable to a rote derivation based on maximums or minimums, whether on an "each year" basis or "all-years" basis.

In the numerical examples that follow, we utilize the illustrative matrix of weights for each projection by accident year shown in Table 4, in order to form a blended point-estimate:

Table 4

AY	Weights to the Alternate Projections: Loss Development on Paid	Loss Development on Reported	Bornhuetter-Ferguson on Paid	Bornhuetter-Ferguson on Reported
2003	33%	67%	0%	0%
2004	33%	67%	0%	0%
2005	33%	67%	0%	0%
2006	33%	67%	0%	0%
2007	33%	67%	0%	0%
2008	32%	67%	0%	1%
2009	31%	66%	1%	2%
2010	29%	66%	2%	4%
2011	24%	62%	4%	10%
2012	14%	50%	8%	28%

The matrix reflects judgments that reported loss data provides more predictive reliability than the paid loss data, with the loss development projections assigned more weight than the BF projections, generally in proportion to the expected reported loss emergence pattern.

Applying the matrix of weights shown in Table 4 to the set of preliminary projections shown in Table 1 yields an ultimate loss estimate of $13,940 and a corresponding unpaid claims estimate of $2,250.

We extend the illustration with assumptions that will form a high-but-reasonable estimate. For simplicity's sake, we assume that the low-but-reasonable estimate is less than the point estimate by the same dollar amount as the high estimate is greater than the point estimate. In other words, we assume that the range of reasonable estimates would be symmetrical around the point estimate. (However, we do recognize that asymmetrical reasonable ranges are very often reported in practice.) We evaluate the effects on the estimate from alternate judgments for the key parameters of the loss development and Bornhuetter-Ferguson methods, namely the loss development factors (LDF's) and expected loss ratios (ELR's).

- Loss development factors: We considered the dispersion of various averages of the historical development factors as indicative of the potential variation of judgments that could be deemed reasonable. In simple terms, an actuary may deem the 5-year average link ratio to be indicative for ultimate loss projections. Another actuary may deem the 3-year or 7-year average to be indicative and reasonable. Alternate judgments may reflect assumptions for future inflation to be higher or lower than the levels embedded in the historical data, or for claim payment or reporting to be faster or slower than during the experience period. For reported losses, the baseline and alternate (high) link ratios and development factors to ultimate are shown in Table 5:

Table 5

Reported Loss Development – Link Ratios and Development Factors to Ultimate

	12 -24 Months	24 -36 Months	36 -48 Months	48 -60 Months	60 -72 Months	72 -84 Months	84 -96 Months	96 -108 Months	108 -120 Months	120 Months to Ultimate
Link Ratios										
Baseline	1.350	1.099	1.031	1.017	1.010	1.001	1.000	1.000	1.000	1.000
Alternate (High)	1.380	1.109	1.036	1.022	1.013	1.003	1.000	1.000	1.000	1.000

	12 Months to Ultimate	24 Months to Ultimate	36 Months to Ultimate	48 Months to Ultimate	60 Months to Ultimate	72 Months to Ultimate	84 Months to Ultimate	96 Months to Ultimate	108 Months to Ultimate	120 Months to Ultimate
Development Factors to Ultimate										
Baseline	1.573	1.165	1.060	1.028	1.011	1.001	1.000	1.000	1.000	1.000
Alternate (High)	1.646	1.193	1.076	1.038	1.016	1.003	1.000	1.000	1.000	1.000

Likewise, judgments were made for the baseline and alternate (high) link ratios for paid losses.

- Expected loss ratios. Different actuaries may have different judgments for ELR's for the Bornhuetter-Ferguson method, considering alternate sources of information. These sources include expected or target loss ratios from a company's pricing or business planning process, or peer company or industry external benchmark information. Alternatively, an actuary might adjust historical projected loss ratios for mature accident periods to current levels for loss trend and changes in pricing levels. A company's history of failing to achieve intended price changes may lead the actuary to select a higher ELR assumption as an alternative scenario. For our illustration, we considered the dispersion of alternate projections based on paid and reported development to be an indicator for alternate ELR judgments, as shown in Table 6:

Table 6

	2012	2011	2010	2009	2008	2007	2006	2005	2004	2003
Weighted LR from Loss Development Projections	68%	76%	62%	63%	63%	71%	52%	53%	59%	65%
Baseline ELR	67%	63%	65%	62%	59%	55%	59%	62%	65%	65%
Alternate (high) ELR	70%	65%	66%	62%	59%	55%	59%	62%	65%	65%

A comparison of the baseline and alternate (high) estimates based on the alternate judgments for LDF's and ELR's are shown in Table 7:

Table 7

	Estimated Ultimate		Unpaid Claims Estimate	
AY	Baseline	Alternate (High)	Baseline	Alternate (High)
2003	1,147	1,147	20	20
2004	1,188	1,188	11	11
2005	1,109	1,109	23	23
2006	1,155	1,155	35	35
2007	1,626	1,628	41	44
2008	1,451	1,457	92	99
2009	1,453	1,467	162	176
2010	1,464	1,487	286	309
2011	1,778	1,824	580	626
2012	1,570	1,646	1,000	1,076
Sum	13,940	14,108	2,250	2,418
			Difference	168
Difference as % Baseline Unpaid Claims Estimate				7%

Thus, based on this example, the high-but-reasonable unpaid claims estimate is $2,418, or $168 (7%) greater than the baseline estimate. In the context of this example, we consider this to be indicative of the high-end of a reasonable range of unpaid claims estimates. With our assumption of symmetry of the high and low estimates relative to the central estimate, the low estimate is $2,082.

5.2 Mack-based Calculations

Continuing with the same sample data set, we supplemented the sensitivity-testing by applying the Mack approach for evaluating a measure of variation in the projections. Table 8 shows the estimated standard error (ESE) of the ultimate loss projection for each accident year and all years combined based on applying the Mack technique to the historical paid and reported loss development data with the same baseline loss development factors as used in the sensitivity testing above. The amount of the ESE of the ultimate loss projection is the same as the ESE of the unpaid claim estimate since the difference (the amount of the known inception-to-date claim payments) is a constant. We observe that the ESE calculated by the Mack approach does not incorporate the variability of any tail development beyond the oldest maturity of the historical data.

Table 8

	Estimated Standard Error	
AY	Paid Data	Reported Data
2003		
2004	1	1
2005	1	1
2006	1	2
2007	2	2
2008	21	23
2009	30	33
2010	31	33
2011	109	76
2012	166	132
All Years	219	175

We observe that the ESE based on the paid loss development data is greater than the ESE based on reported loss development data. This is a feature of the sample data set and not necessarily indicative that ESE's based on paid development data are always greater than the ESE's based on reported loss development data.

In order to generate a distribution of possible outcomes for the unpaid claims amounts, we chose an ESE of $197, based on the average of the two indicated ESE's. The chosen ESE was equivalent to 9% of the baseline unpaid claim estimate of $2,250.

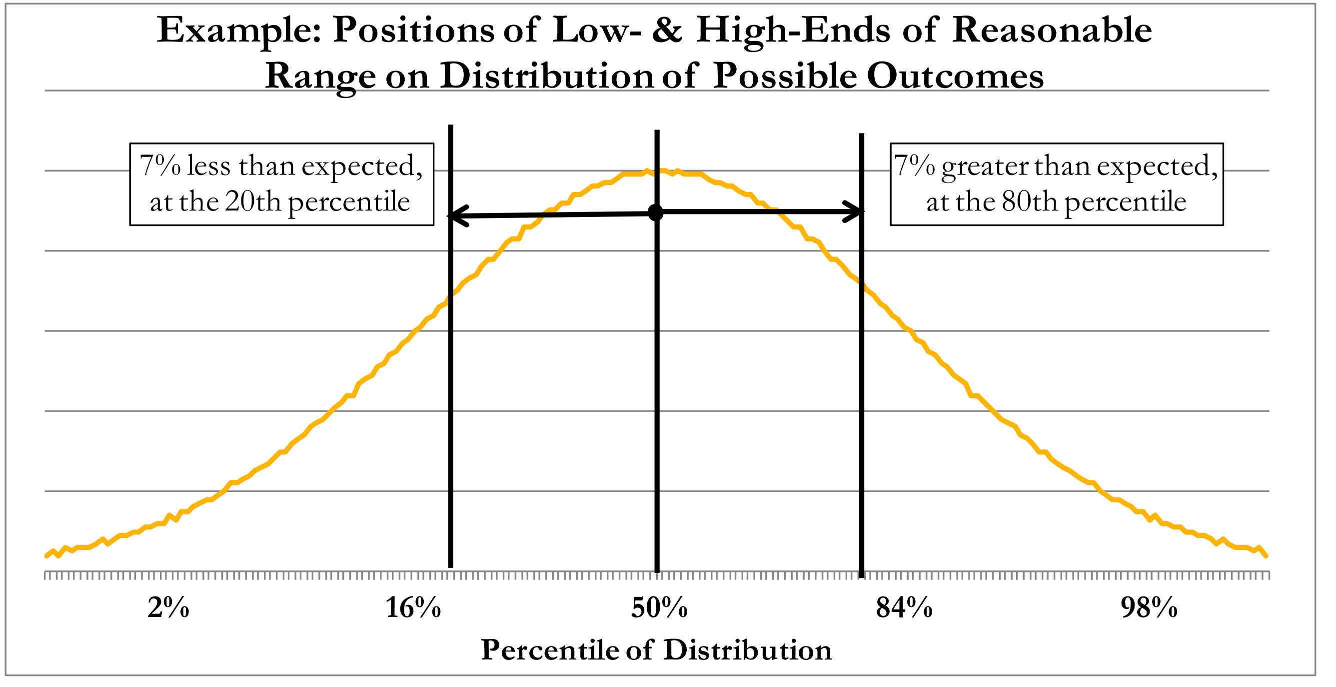
In keeping with the spirit of the non-technical nature of this review, we elected to assume a normal distribution to characterize the dispersion of possible outcomes of unpaid claim amounts. One of our goals with this paper was to describe a framework for connecting information about a reasonable range of estimates based on sensitivity testing to information about a distribution of possible outcomes based on a stochastic approach such as the Mack technique. To that end, we evaluated the end-points of a confidence interval around the mean from a normal distribution with a standard deviation based on the selected ESE from the Mack technique, where the confidence interval would align with the range of estimates generated by the sensitivity testing. The results are shown in Table 9:

Table 9

Percentiles of Distribution		Unpaid Claim Estimate	
Low	High	Low	High
20%	**80%**	2,082	2,418

In this case, the range based on the sensitivity testing extended from 7% less than to 7% greater than the baseline estimate. The 20th and 80rd percentiles of the distribution of outcomes based on our analysis with the Mack technique aligned with this range. This relationship is illustrated in Chart 1:

Chart 1



The chart illustrates the normal distribution, by the familiar bell-shaped curve, with x-axis markers at the 2nd, 16th, 50th, 84th, and 98th percentiles of the distribution.

2nd percentile	amount that is 2 standard deviations less than the mean
16th percentile	amount that is 1 standard deviation less than the mean
50th percentile	the mean amount
84th percentile	amount that is 1 standard deviation greater than the mean
98th percentile	amount that is 2 standard deviations greater than the mean

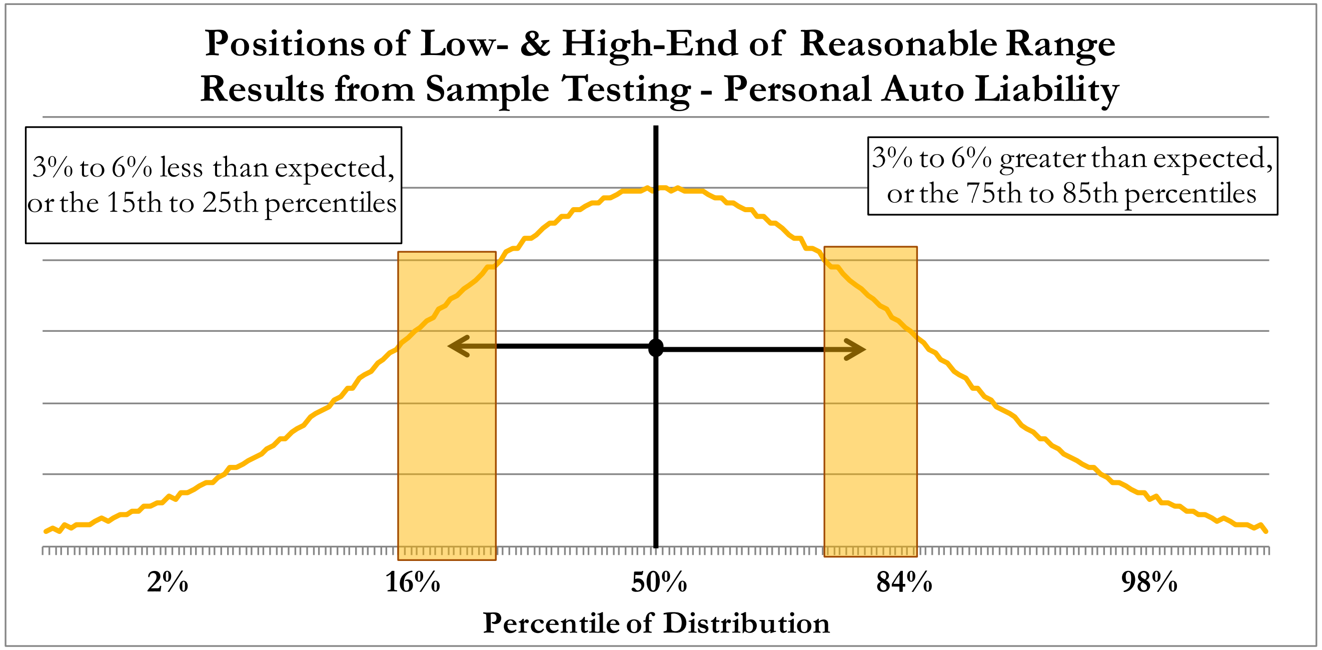
The three vertical lines correspond with the low, central, and high estimates from the example. Stated another way, the 7% differential from the baseline mean unpaid claim estimate to the high estimate based on sensitivity testing was equivalent to 0.85 of the ESE (also known as standard deviation) from the Mack-based distribution analysis.

5.3 Exploring a Potential Relationship between Sensitivity-based Ranges and Mack-based Distributions

We applied the approaches described above (supporting Tables 4 to 9) to a set of publicly-available data for Personal Auto Liability, Homeowners, and General Liability – Occurrence coverage data for 10 insurance companies. The findings shared herein are intended to be indicative of the application of the framework for integrating metrics from the sensitivity approach and a stochastic approach in order to help establish a potential connection between a range of reasonable estimates and a distribution of possible outcomes. These are not intended to be construed as "the definitive statement" on the relationship between the two approaches.

The results from our sample testing for Personal Auto Liability are summarized in Chart 2:

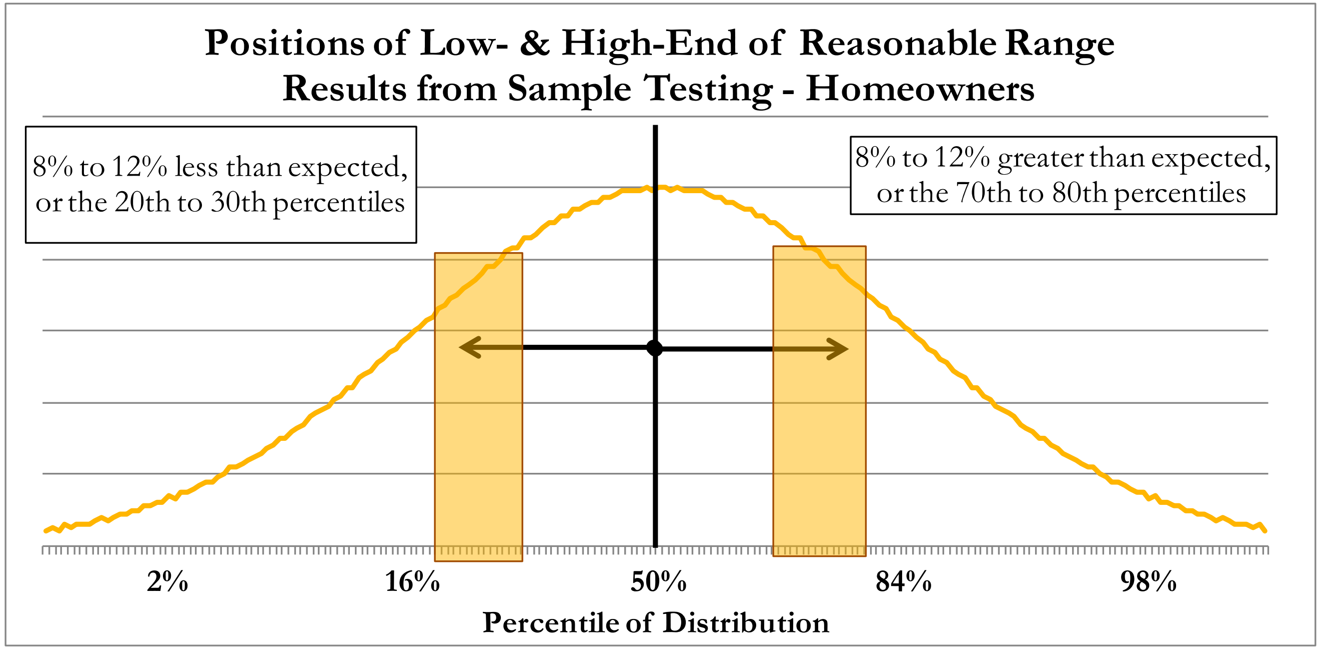
Chart 2



Based on the sample testing for Personal Auto Liability, we observe that high-ends of the reasonable estimate reserve ranges were generally 3% to 6% above the baseline estimate, and these high-ends tended to correspond with distribution percentiles at the high-end of confidence intervals that were generally in the 75% to 85% interval. With our assumption of a symmetrical range of reasonable estimates and distribution of outcomes, the low-ends of the reasonable range tended to correspond with distribution percentiles at the low-end of confidence intervals that were generally in the 15% to 25% interval.

The results from our sample testing on Homeowners data is shown in Chart 3:

Chart 3



Based on the sample testing for Homeowners multi-peril coverage, the reasonable reserve range high-ends were generally 8% to 12% above the baseline estimate, corresponding with high-end percentiles that were in the 70% to 80% range.

We also performed our testing on a sample of company data for General Liability – Occurrence coverage. The results are shown in Chart 4.

Chart 4

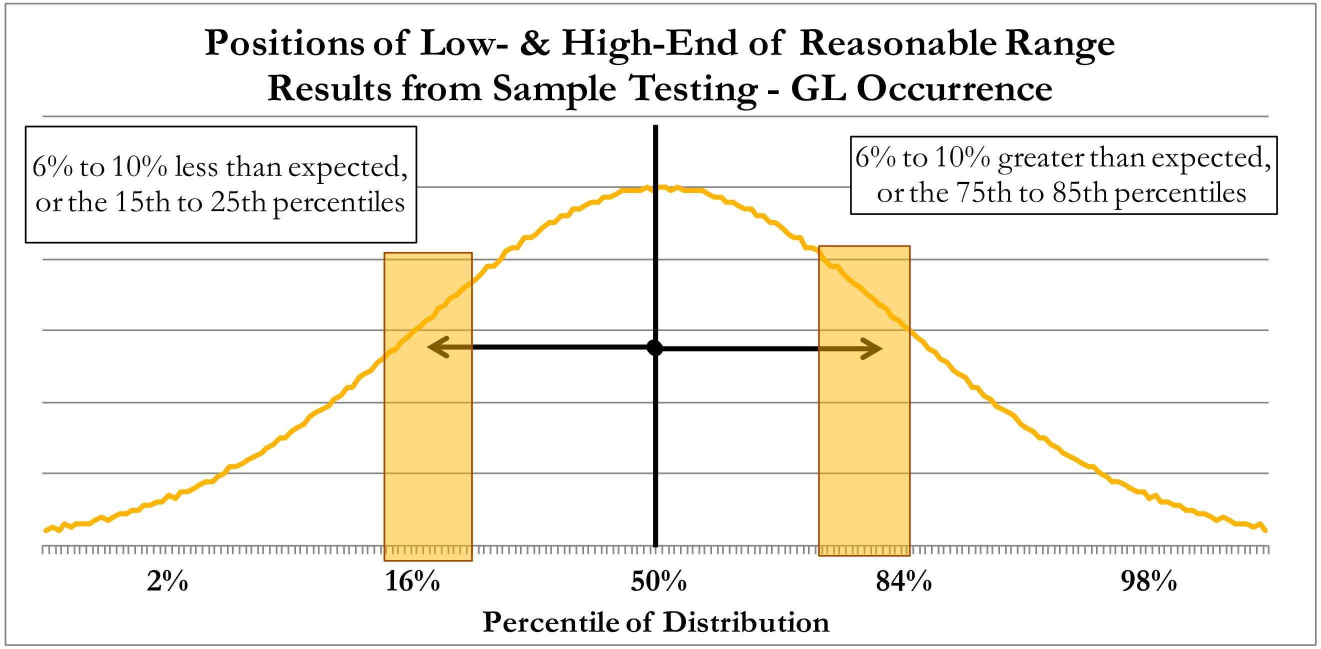


Table 10 summarizes our observations from our testing for the three lines.

Table 10

	High-End Reasonable Range as % Reserves	Percentiles of Distribution aligning with High-End of Reasonable Range	# Std Dev's from Mean to High-End of Reasonable Range	Estimated Standard Deviation of Distribution
Personal Auto Liability	3% to 6%	75th to 85th	0.7 to 1.0	3% to 7%
Homeowners	8% to 12%	70th to 80th	0.6 to 0.9	12% to 16%
GL Occurrence	6% to 10%	75th to 85th	0.7 to 1.0	6% to 12%

The breadth of the range (difference from the reasonable range high-end to the mean/baseline estimate of unpaid claims) expressed in relation to the unpaid claim estimate for Homeowners tended to be larger than for GL-Occurrence and Personal Auto Liability. We believe this is a consequence of Homeowners having a greater proportion of the ultimate loss estimates being paid for a given accident year maturity than the other lines, and thus the measure of uncertainty/range was relatively greater in proportion to the unpaid claim estimate. We also observe that actual reserves for these lines recorded by insurance companies may have subsequently developed within these indicated reasonable ranges or beyond the end-points (low or high) of the ranges. Tracking the actual development of recorded amounts may be another way to consider an evaluation of ranges of reasonable estimates, but that would incorporate an element of hindsight testing, while we are considering the reasonability of estimates based on information available at a point in time.

6. Considerations of Ranges on an Aggregate Basis[2]

While an evaluation of a range of reasonable estimates for an individual business segment has an inherent degree of difficulty, the challenge is elevated for an evaluation of a reasonable range or a distribution of outcomes on an aggregated basis. The higher degree of difficulty is, in part, due to the need to consider and reflect potential correlations among reserve segments. Nevertheless, since actuarial opinions are primarily given for a company, for which there are generally multiple analysis segments, or for a corporate group, comprised of multiple

[2]In this section we explore considerations for an aggregate range of reasonable estimates only. The CAS literature contains a variety of papers describing approaches for evaluating aggregate distributions of possible outcomes. Our primary focus throughout this Call paper has been on ranges of reasonable estimates.

companies, the actuary must consider how to approach the analysis of ranges on an aggregated basis. We believe that there are merits in a "bottom-up" approach and a "top-down" approach as discussed below.

6.1 Bottom-Up Approach

Under a bottom-up approach, the actuary would first evaluate ranges for individual reserve segments, and then form an aggregate range. An analysis of the ranges for individual lines, coverages, or other attributes by which the data are organized or the business is managed can provide management with insights on the relative certainty of estimates of ultimate losses and the associated unpaid claims. The fundamental issue in the aggregation is the consideration of potential correlations among the various reserve segments. If all segments are deemed to be independent of each other, than a simple "square root of the sum of the squares" may be practical and sufficient in evaluating an aggregate range. If all segments are deemed to be fully (and positively) correlated with each other, then the sum of the high and low ends of the individual ranges would be indicative of an aggregate range.

Situations in-between these two extremes can be tricky. The practitioner can make judgments for the correlations, or may perform advanced calculations in an attempt to quantify correlations in development among the different pairs of segments. As a simplified alternative, the actuary could assume 100% and 0% correlations to calculate the two aggregate indications, and form an aggregate view on correlation in order to construct a weighted-average of the two aggregate indications.

In practice, we often observe actuaries simply summing up the "low" ends and, similarly, the "high" ends, to development a range of reserves in the aggregate.

6.2 Top-Down Approach

An alternative to a bottom-up approach to evaluate a range of reasonable estimates at an aggregate level would be to evaluate the potential variation in central estimates by applying sensitivity testing or the Mack technique to aggregated data. We do not generally advocate an analysis of aggregated data for evaluating a point estimate, but consider it potentially useful to perform sensitivity testing or stochastic analysis in order to assess an aggregate range of reasonable estimates. We observe that the mix of the underlying coverages should be relatively stable over the experience period for such an analysis of aggregate data; to the extent that there are substantial shifts of the mix of business (for instance, relative proportion of long and short tail business), we would caution against this approach. When the underlying data is satisfactory for this purpose, the top-down approach has a relative advantage of implicitly handling correlation among the underlying business segments.

The illustration presented in section 5.1 above yielded an estimated range of reasonable estimates of the unpaid claims that extended from 7% less than to 7% greater than the point-estimate of $2,250; this segment will be referenced as Line 1. We performed a similar analysis for Line 2, for which the estimated range of reasonable estimates of unpaid claims extended from 9% less than to 9% greater than the point-estimate of $1,000. We also performed a similar analysis on the combined data for the two lines, for which the range extended from 6% less than to 6% greater than the point estimate of $3,168[3]. The illustrative results are summarized in Table 11.

Table 11

	Unpaid Claims Estimate (UCE)	High-end of Reasonable Range minus UCE	High-end of Reasonable Range as % Reserves
Line 1	2,250	168	7%
Line 2	1,000	91	9%
Combined	3,168	191	6%

If the two lines were 100% correlated, then the difference from the central estimate to the high-end of the reasonable range for the combined data would be the sum of the two lines' differences, or $259. If the two lines were deemed independent of each other, the difference from the combined central estimate to the high-end could be reasonably approximated as the square root of the sum of the squares of the lines' metrics, or $191. As the difference to the high-end of a reasonable range based on the combined data was evaluated at $191 greater than the point estimate for the combined data, we infer that the two lines have an approximate 0% correlation.

From our testing on the Personal Auto Liability and Homeowners data for five companies in our sample, we observed implied correlations between the reserve ranges for the two lines ranging from (0.3) to +0.8. The implied correlations were highly sensitive to alternate judgments around the reasonable range on the combined data; thus, we do not believe the reader should take away any particular "rule of thumb" on correlations.

7. Conclusion

We wrote this Call paper with the goal being to describe a variety of practical approaches that we have observed for assessing variability of unpaid claim estimates and to present illustrations of the application of chosen methods for evaluating and comparing ranges of

---

[3] We acknowledge that the sum of the point-estimates for the two lines is $3,250, which is slightly greater than the point-estimate based on the combined data. LDF's for the analysis of aggregate data were calibrated based on the parameters for the two lines; the small difference arose from small differences in the ELR's and the weights applied to the various projections to form the point-estimate. We do not consider the differences significant in the context of our discussion of the framework of the analysis.

reasonable estimates and distributions of possible outcomes. We believe the framework described herein is practical and can be reasonably explained to the variety of stakeholders who seek insights and opinions from actuaries on point-estimates and the associated uncertainty.

In the course of preparing this paper, we discovered an apparent relationship that the illustrative ranges of reasonable estimates for the three lines reviewed tended to align with portions of the distribution of outcomes that extend up to one standard deviation above and below the mean. While the estimated standard errors for each segment reflected the inherent nature of the line and the company's claims development experience, the ranges of reasonable estimates tended to be subject to similar degrees of variability. This should be an area of further and more robust research.

Just as there is uncertainty and judgment inherent in the process for determining a central estimate of unpaid claim liabilities, these attributes are inherent in evaluating a range of reasonable estimates. While the accuracy of a point estimate will ultimately be known when all subject claims are settled and paid, expressions of a range of reasonable estimates are much more tenuous and cannot be tested with hindsight; therefore, such expressions primarily serve as indications of the effects of plausible differences in assumptions. We believe that the days of expressions of reasonable ranges based solely on judgment or rules of thumb are over, as stakeholders seek a more-reasoned response to questions regarding the basis of a stated range.

Acknowledgment

The authors acknowledge the contributions of Barbara Jackie and Sun Sun in preparing the analyses of the data summarized in this paper.

8. References

[1] Bornhuetter, Ronald L and Ferguson, Ronald E, "The Actuary and IBNR," *PCAS* **1972**, Vol. LIX, 181-195.

[2] Mack, Thomas. "Measuring the Variability of Chain Ladder Reserve Estimates," *Casualty Actuarial Society Forum, Spring 1994,* 101-182.

Abbreviations and notations

ESE, estimated standard error

LDF, loss development factor

SEC, Securities & Exchange Commission

GAAP, generally accepted accounting principles

SAP, statutory accounting principles

Biographies of the Authors

Christopher Walker is a principal in the Actuarial & Insurance Management Solutions practice of PricewaterhouseCoopers LLP. His primary technical responsibilities include loss reserve, ratemaking, and regulatory assistance for dozens of insurance carriers and self-insurers. Particular experience is in personal lines, primary/excess workers' compensation, warranty, and insurance company valuation. He is a Fellow of the CAS and a Member of the American Academy of Actuaries.

Mark Littmann is a principal in the Actuarial & Insurance Management Solutions practice of PricewaterhouseCoopers LLP. He leads actuarial teams supporting external financial audits of insurance organizations and corporations with self-insured exposures and providing consulting services regarding reserving

valuations and processes and other analytical applications. He has a degree in Mathematics and Economics from Valparaiso University. He is a Fellow of the CAS and a Member of the American Academy of Actuaries. He is a member of the program planning committee for the Casualty Loss Reserve Seminar.