Charge #6 -Expert Witness References

2.4.2.2 Choosing the Parametric CER Distribution Shape and Point Estimate Location

Regardless of the CER method employed, even if the CER inputs (independent variables) are known precisely, the CER will return a result that is not certain. Depending on how the CER is developed, the error is either additive or multiplicative (a factor of the CER result). In cost estimating, we expect the potential error of the CER to scale with the CER result thus making multiplicative error terms the preferred approach for modeling the CER uncertainty (see **Appendix A.8.4**).

Two critical decisions when applying uncertainty to CERs are: selecting the uncertainty shape and defining where the point estimate falls within the distribution. Both of these decisions should be based upon an understanding of the regression method used to develop the CER. It is also possible to fit a distribution to the residuals (see **Section 2.4.3.8**). Point estimate location considerations for the most common regression methods are:

- **OLS Linear**: A premise of the OLS method is that the errors will be normally distributed in fit space. The OLS linear CER result is the center (mean, median, mode) of the normal distribution.
- **OLS Log Space**: Applying OLS in log space yields a multiplicative lognormal uncertainty in unit space. The point estimate result is the **median**, not the mean of the lognormal uncertainty.
- **MUPE**: The MUPE CER delivers the mean; it has zero proportional error for all points in the CER. Goodness-of-fit measures can be derived to judge the quality of the model if the CER error is assumed to be normal (a common assumption).
- **ZMPE**: The ZMPE method also delivers the mean and zero proportional error for all the data points in the CER. Distribution shape is arbitrary; however, some analysts prefer using lognormal.

Two critical decisions: Select the uncertainty shape and define where the point estimate falls.

2.4.2.3 Calculating the Prediction Interval

CER error terms provide an average error for the equation. This average error needs to be adjusted to account for the number of sample points, degrees of freedom, and where the estimate is located in the data set. Most statistical software packages will do this calculation for OLS CERs and some will address MUPE. For details on how to calculate simple OLS CER prediction intervals manually, see **Appendix A.8.5** and **A.8.8.** If the CER was generated using a statistical tool, or if the detailed statistical results are available, this information should be used to calculate the prediction interval.

The prediction interval is the preferred source of distribution bounds.

2.4.2.4 Estimating the Prediction Interval

When obtaining a prediction interval is not possible, the CER's standard error (SE, also called standard error of the estimate SEE, see **Appendix A.1.9**) may be used to obtain objective uncertainty. This may arise in the absence of the data and the time to create the CER and its prediction interval, or the lack of detailed statistical results from a published CER. Also known as the CER standard deviation, the SE alone is generally not sufficient to define CER uncertainty. An adjustment is required to account for the location of the point estimate within the dataset used to generate the CER. As the estimate moves away from the center of the CER dataset, the spread of the prediction interval will increase.