BOOTSTRAPPING METHODOLOGY
DOCKET NO. 02-057-02
DPU EXHIBIT 6.11

Bootstrapping can be used to estimate the uncertainty of a particular statistic such as, the sample mean or median. The uncertainty is measured as the "confidence" or probability interval for the statistic.

"For example, to estimate the uncertainty of the median from a dataset with [9] elements, we generate a subsample of [9] elements and calculate the median. This is repeated at least 500 times so that we have at least 500 values for the median. Although the number of bootstrap samples to use is somewhat arbitrary, 500 subsamples is usually sufficient. To calculate a 90% confidence interval for the median, the sample medians are sorted into ascending order and the value of the 25th median (assuming exactly 500 subsamples were taken) is the lower confidence limit while the value of the 475th median (assuming exactly 500 subsamples were taken) is the upper confidence limit."[1]

In the present case, I actually use 1000 subsamples of size 9. Each subsample is drawn from the original data set of 9 ROE estimates reported in Dr Williamson's testimony Exhibit QGC 5.3. The sampling is done with replacement based on a uniform distribution: on each draw from the data set, every value in the data set has an equal chance of being selected. After each subsample is selected, the subsamples mean and median are calculated and stored for later reference. At the completion of the 1000 subsamples, the confidence intervals for each statistic are calculated as described above, adjusting for the larger number of subsamples (i.e., the 50th and 950th values represent the lower and upper bounds respectively). A copy of the program, which is written in GUASS©, a matrix language statistical program, used to run the bootstrapping estimation is available upon request.

[1] *Engineering Statistics Handbook*, "Bootstrap Plot," http://www.itl.nist.gov/div898/handbook/eda/section3.