Understanding Qwest's 271 Statistical Reports

Reporting Format

Qwest generates its service provisioning performance reports organized by checklist item[1] for each state, comparing CLEC aggregate performance to Qwest retail performance where appropriate, or to a fixed standard in the case of a benchmark measurement.

The reports show the following detailed data for a rolling twelve-month period:

CLEC numerator – the numerator of the expression producing the CLEC result, e.g., number of orders completed on time, total number of days to complete all orders, number of trouble tickets.

CLEC denominator – the denominator of the expression producing the CLEC result, e.g, total orders completed, total tickets completed, total installed lines or circuits.

CLEC result – the CLEC performance measurement, i.e., the CLEC numerator divided by the CLEC denominator.

Std Dev – the CLEC standard deviation. For measured variables (e.g., average intervals), this is the sample standard deviation, the square root of the sum of squared deviations around the mean divided by n-1, where n is the denominator. For counted variables (i.e., proportions, reported as percentages), this is the standard deviation, which is equal to the square root of p*(1-p), where p is the sample proportion.

Qwest numerator – the numerator of the expression producing the Qwest result.

Qwest denominator – the denominator of the expression producing the Qwest result.

Qwest result – the Qwest performance measurement, i.e., the Qwest numerator divided by the Qwest denominator.

Mod Z Scr – the modified z-score, defined and discussed below.

Parity Scr – the parity score, defined and discussed below, incorporating the permutation test results for small samples, indicating whether there is parity or not.

Within each state, Qwest breaks out the results geographically:

---

[1] Local Interconnection; collocation, OSS Gateway availability, etc. Qwest also generates the report in PID order for other purposes.

A. MSA dispatch – results for specified services where a technician/installer was dispatched in Metropolitan Statistical Areas (MSA).

B. Non-MSA dispatch – results for specified services where a technician/installer was dispatched in Non-MSA areas.

C. Non-dispatch – results for specified services where no dispatch was involved, in both MSA and non-MSA areas.

D. Interval Zone 1 – results for specified services in areas designated by Qwest as Interval Zone 1 areas (generally, urban areas).

E. Interval Zone 2 – results for specified services in areas designated by Qwest as Interval Zone 2 areas (generally, rural areas).

Note that categories A, B, and C typically apply to non-designed services (i.e., those not requiring design work), while D and E typically apply to designed services (i.e., those requiring attention by a design specialist to complete order processing).

In addition to the numerical results, Qwest provides a graph of the CLEC performance, along with the benchmark or Qwest retail performance for visual comparison. The graph includes an arrow pointing in the "good" performance direction.

Types of Measures – Benchmarks and Parity Comparisons

In general, the Performance Indicator Definitions ("PIDs") specify that CLEC performance is compared to Qwest retail performance where a retail analogue is available. Where a retail analogue is not available, a benchmark standard is specified. For some measurements, the PID specifies "diagnostic" [2] or "parity by design"[3] as the standard, indicating that no established benchmark or retail comparison that applies.

Benchmark measurements compare Qwest's performance for the CLECs to a benchmark using a "stare and compare" method. Qwest does not perform

[2] A standard of "diagnostic" is used for a variety of reasons, including (a) the measurement is considered supplemental to another measurement against which a specific standard is applied; (b) the measurement does not reflect Qwest's performance (e.g., PO-4, percent rejected LSRs, or MR-10, customer-caused repair misses); (c) the measurement or, in some cases, certain products listed for reporting in a measurement, are too new to assign a retail analogue or benchmark; (d) the measurement overlaps to some degree another measurement that has a standard specified.

[3] "Parity by design" asserts that the processes or systems underlying the performance being measured are inherently non-discriminatory, such that Qwest has no ability to render different service to CLECs in comparison to retail customers. The performance measurements audits are designed to verify this assertion.

statistical tests of benchmark measurements. Random variation in the measurements is assumed to be accounted for in the establishment of the benchmarks.

Parity measurements reflect Qwest's ability to perform for CLECs as it does for retail customers. Qwest uses statistical methods to assess that performance.

The purpose of statistical testing is to account for random variation in the performance results, so that true differences in service can be distinguished from random differences, with specified levels of confidence. It is vitally important that "false positives" – finding a difference in service where no difference truly exists – be controlled. If Qwest is continually confronted with false positives, then it will be unable to implement the process improvements necessary to ensure parity of service.

Qwest follows the standard hypothesis testing procedure:

1. Specify a null hypothesis, H_0, that the performance is equivalent,
2. Specify an alternative hypothesis, H_a, that there is a difference in performance between CLEC and Qwest retail,
3. Select a test statistic, e.g., modified z,
4. Select a critical value for the test statistic, z*.

The testing procedure involves comparing the test statistic, z, with the critical value, z*. If z is greater than or equal to z*, then there is a significant difference in service.

For parity measurements, Qwest uses the modified z-test.[4] In contrast to some states, including New York, Qwest structures the z-tests so that a positive z-score indicates inferior performance provided to the CLECs.

The modified z-test is:

$$z = \frac{\bar{X}_{CLEC} - \bar{X}_{Qwest}}{S_D}$$

where:

[4] This test was initially proposed by the Local Competition Users Group (LCUG), a group of CLECs.

$$s_D = s_{Qwest}\sqrt{\frac{1}{n_{ILEC}} + \frac{1}{n_{CLEC}}}$$

$s_{Qwest} = \text{Qwest standard deviation}$

$n_{Qwest} = \text{Qwest sample size}$

$n_{CLEC} = \text{CLEC sample size}$

$\bar{x}_{CLEC} = \text{CLEC mean (or proportion)}$

$\bar{x}_{Qwest} = \text{Qwest mean (or proportion)}$

For z*, Qwest uses the value 1.645. This corresponds to a 5% significance level, or alternatively a 95% confidence level.

The 95% confidence level is commonly used in this sort of hypothesis testing. At this level, one test has a 5% probability of finding a difference due to chance alone, even when no true difference exists (a Type I error). This level of significance also provides a fair compromise between Type I and Type II errors (the probability of failing to detect a true difference).[5] Sample size is the best way to reduce Type II error. However, since these results do not come from a controlled study, sample size can not be controlled.

The modified z-test gives a z-score that is then compared to a critical z-value, z*. If z is greater than or equal to z*, the difference is statistically significant at the level indicated by z*, e.g., 95%.

Qwest presents the comparison of z with z* in the parity score.

Parity Scores

The parity score indicates whether there is numerical parity or disparity between CLEC and Qwest retail results. For large sample sizes, the parity score is calculated directly from z and z*. For small sample sizes, an additional step is performed to calculate the parity score.

The z-test assumes large sample sizes. When sample sizes are small, or the data are skewed, the z-test may give a result that is more likely to be in error.

To overcome these problems, Qwest uses a permutation testing procedure. The permutation test is a non-parametric test, which means that it does not make any assumptions about the distribution of the data, e.g., normality, or number of observations. It constructs its own distribution based on the data available. It does this by randomly generating 1000 samples, each dividing the data between CLEC and retail. If 50 samples out of the 1000 have a mean difference larger

[5] See Appendix B of the FCC's order in BANY-NY, FCC99-404, p.9.

than the actual mean difference, then we have a non-significant result (at α=0.05). If not, we have a significant result.[6]

Qwest calculates the actual probability of obtaining a result at least as extreme as the actual result. This is called the p-value, and takes a value from zero to 100. If the p-value is .05 or less, the result is statistically significant.

Qwest then converts the p-value to a z-score using a standard normal distribution function, and calculates the parity score from this z-score.

Qwest performs permutation tests on all samples where the z-score is greater than zero and where the CLEC or Qwest sample size is 100 or less.

For proportion type measures, Qwest uses an exact probability test, which is also a form of a permutation test.[7] The results of the exact test are reflected in the parity score. Qwest performs the exact probability test for all proportion type measures where there is a positive difference between the CLEC and retail proportions and where the CLEC or Qwest sample size is less than or equal to 500.

The parity scores indicate whether there is parity between a wholesale measurement and its retail comparative, taking into account the permutation testing results. The following table indicates how to interpret parity scores:

Parity Score	Meaning
< -1.0	Wholesale is "better" than retail comparative
= -1.0	Wholesale and retail appear to be exactly equal
> - 1.0 and < 0.0	Retail appears to be "better" than wholesale, but the difference is not statistically significant
>= 0.0	Retail is better than wholesale and the difference is statistically significant.

The parity score is calculated from the z-score and z*, the z critical value.[8]

[6] Joint Proposed Rule Language for Statistical Analysis Process by Drs. Michael Carnall and Colin Mallows, Colorado Public Utilities Commission 97R-153T, December 9, 1998.

[7] This test is based on Fisher's exact probability test, based on a hypergeometric distribution, but does not restrict the marginals. Sidney Siegel, Nonparametric Statistics, New York: McGraw-Hill, 1956, p. 96.

[8] The critical z* value is from a standard table of critical z* values. Qwest uses a z* that sets α=.05 for one test, or z* = 1.645. This corresponds to a 95% level of confidence.

$$Score = \frac{z - z^*}{z^*}$$

where:

z = the calculated z-score, and

z* = the critical z-value (1.645).

The parity score represents the relationship of the calculated z-score to the critical z-value. If it is zero or greater, there is a statistically significant difference at the nominal significance level of 5%.