Performance Incentive Plan

The plan AT&T sets forth below is based upon its experience in attempting to develop effective consequence plans in several States, including most notably Texas and California. This plan also incorporates AT&T's experience in New York with the "Carrier-to-Carrier" proceeding and its real-world experience in attempting to place orders for local service with Bell Atlantic-New York (BANY). Overall, AT&T believes that this plan is substantially superior to the Performance Assurance Plan (PAP) proposed by BANY, because, among other things, it is significantly less complex, it is not subject to many layers of caps, and it does not apply artificial weightings to individual performance measures.

Introduction

It is well recognized that a meaningful system of self-enforcing consequences for discriminatory ILEC performance is critically important to the protection of the public's interest and the rapid and sustainable development of a competitive local telecommunications market. Incumbent LECs have strong business incentives and means to maintain their current monopolies through the delivery of inadequate and unlawful levels of operations support for CLECs. Thus, an appropriate system of self-enforcing consequences is absolutely necessary to assure that the competitive local telecommunications markets envisioned by the 1996 Act will be able to develop and survive.

In order to be effective, prompt enforcement of appropriate consequences must be assured. Because of the extensive delays inherent in the adjudication and appeals process, CLECs cannot rely solely upon the

legal/regulatory process to obtain appropriate remedies for discriminatory ILEC performance. Furthermore, the consequences must provide BANY with incentives that exceed the benefits it may derive by inhibiting competition, and such consequences must be immediately imposed upon a demonstration of poor BANY performance. The objective is to set the incentives in amounts that encourage BANY to take proactive steps to prevent its performance from becoming non-compliant and, when it does reach that level, to correct its performance failures promptly.

It is beyond dispute that any system of self-enforcing consequences must be based upon an underlying set of performance measurements that cover the full panoply of BANY activities upon which CLECs must rely to deliver their own retail service offerings. The Act requires that these activities, which touch upon every aspect of the business relationship between incumbents and CLECs, must be provided in a non-discriminatory manner. Thus, the interconnection agreements between incumbents and CLECs should ideally serve as a source for performance measurements. However, experience in New York and elsewhere has proven that CLECs have generally been unable to individually negotiate, or even arbitrate, a sufficiently robust set of performance measurements.[1] For that reason, the first step in constructing a system of self-enforcing consequences must include careful consideration of the adequacy of the underlying measurement set. At a minimum, the performance measurements must supply each CLEC with reliable data on the incumbent's performance for that CLEC. Such data must be sufficiently discrete (as to the processes monitored) and detailed (to isolate and compare only comparable conditions) so as to permit a CLEC to enforce the terms of its interconnection agreement with the incumbent. In addition, the

[1] As a starting point, the CLEC industry generally supports the measurement areas as documented in Local Competition Users Group (LCUG) – Service Quality Measurements (SQMs), Version 7.0, August 28, 1998.

underlying performance measurement system should demonstrate quality implementation of the following characteristics:

- A comprehensive set of comparative measurements that monitors all areas of support (i.e., pre-ordering, ordering, provisioning, maintenance & repair and billing) without preference to any particular mode of market entry
- Measurements and methodologies that are documented in detail so that clarity exists regarding what will be measured, how it will be measured and in what situations a particular event may be excluded from monitoring (such exclusions must also be tracked and reported)
- Sufficient disaggregation of results, so that only the results for similar operational conditions are compared and, particularly, so that the averaging of results will not mask discrimination[2]
- Pre-specified and pro-competitive performance standards exist. This includes identifying reasonably analogous performance delivered by the incumbent to its own operations[3] or, when such comparative standards are not readily identifiable, that absolute minimum standards for performance (benchmarks) are established[4]

[2] The importance of sufficient disaggregation is more fully discussed in Attachment A

[3] Analogous performance must be broadly interpreted and consider not only retail operations of the incumbent but also operations of affiliates. Often the incumbent's asserted lack of analogous performance relies upon very narrow (and inappropriate) interpretation of the term "analogous" to mean "precisely identical" rather than "similar in key aspects." Furthermore, if the incumbent delivers different levels of performance to an affiliate and its the retail operations, the CLEC experience should be compared to the better of the two.

[4] In all cases, benchmarks must provide an efficient competitor with a meaningful opportunity to compete.

- Sound quantitative methodology is used to compare CLEC experiences to analogous incumbent support[5]
- The overall performance measurement system is subject to initial and periodic validation, in order to assure that the performance results which form the foundation for all decisions regarding the quality of the performance delivered by the ILEC are correct representations of the CLECs' marketplace experience.[6]

It is critical that a performance measurement system incorporating all of the above characteristics exist before applying an incentive plan, because a robust and independently audited performance measurement system is a prerequisite to any effective system of self-enforcing consequences – whether it is this plan or the PAP. AT&T's recent experience in New York, both with UNE-Loop hot cuts and UNE-P orders, indicates that there are significant problems with the performance measures that have arisen from the Carrier-to-Carrier proceeding, Case No. 97-C-0139, and that such problems have led to misleading performance reports from BANY. These problems arise from at least four factors: (1) ambiguities in the performance measurement definitions that BANY has used in its own favor in reporting results; (2) BANY misinterpretations of performance measures that have had a similar effect; (3) measures that are still missing and that are needed to

[5] As a general rule, when benchmarks are employed, statistical comparisons of the measured result for the CLEC to the benchmark are not appropriate. Typically, the standards state a minimum performance level that is required to support effective competition and the minimum success level that must be demonstrated to attain the benchmark. Thus, the typical form of the standard is, for example, "95% installed within 3 days." Note that in the preceding example a 5% deviation from the benchmark is permitted and, as a result, the potential for random variation of the performance is fully addressed. Any further accommodation of variation, as would occur if statistical procedures were employed, would effectively "double count" forgiveness of variability.

[6] A sample plan for an initial audit of the performance measurement system is included as Attachment B.

develop a full view of BANY's performance for CLECs; and (4) extensive exclusion of data points occurs because of liberal of self-determined classification (e.g., customer not ready). [7]

Objectives of the Plan

A system of self-enforcing consequences must fully implement the following objectives:

- Consequences must be based upon the quality of support delivered on individual measures to individual CLECs
- Total consequences, in the aggregate, must have sufficient impact to motivate compliant performance without the need to apply a remedy repeatedly
- The imposition of financial consequences must be prompt and certain, and consequences should be self-executing so that opportunities for delay through litigation and regulatory review are minimized
- Consequences must escalate as the basis for concluding that a performance failure exists becomes more substantial and/or the performance repeatedly fails to meet the applicable standard
- Additional consequences must apply when non-compliant performance is provided to CLECs on an industry-wide basis

[7] For example, business rules for individual performance measurements may provide for automatic exclusions of data points from the calculation. If such provisions are made, however, the exclusions must be according to clearly defined rules and the number of data points excluded for each submeasurement and for each CLEC should be reported on a monthly basis.

- Exclusions from consequences must be minimized and the exclusions that are provided for must be monitored and limited to assure they do not mask discrimination
- Incumbents must have minimal opportunities to avoid consequences through such means as liability caps, offsetting credits, or a requirement that CLECs must demonstrate an ILEC's intent to harm
- Potential "entanglement" costs must be minimized so that, for example, access to mitigation measures for the incumbent does not become a means to revert to the legal/regulatory process and delay the application of consequences that should be self-enforcing

Structure of Consequences for Discriminatory BANY Performance

Consequences operating on two tiers are proposed. The first tier addresses the consequences for non-compliant performance delivered to an individual CLEC. The second addresses the consequences for non-compliant performance delivered to the CLEC industry as a whole. In general terms, Tier I provides a form of non-exclusive liquidated damages payable to individual CLECs. Tier II, by contrast, incorporates what can be characterized as regulatory fines that are necessary when BANY's performance affects the competitive market – and consumers -- as a whole.

The total amount of Tier I payments (which are only an estimate of the CLECs' actual damages) is unlikely to provide BANY with sufficient incentives to take the actions necessary to eliminate its monopoly. Rather, BANY may decide to treat such payments as the price for retaining its

monopoly and voluntarily incur them as a cost of doing business. Moreover, the harm that results when BANY provides discriminatory support for the CLEC industry in the aggregate has a major impact not only on CLECs but also on the operation of the competitive marketplace in general, which directly affects all New York consumers of telecommunications services. Thus, it is appropriate to establish incentives to prevent this type of harm from occurring (or continuing), and both Tier I and Tier II are necessary and complementary elements of an effective system of consequences. Together, they work in tandem to achieve the goals of the Act.

Tier I

A Tier I consequence should be payable to an affected CLEC whenever any performance result indicates support delivered by BANY to an individual CLEC fails to meet or exceed the applicable performance standard.[8]
The first step in establishing Tier I consequences is to define the rule for determining if performance for a particular period "passes" or "fails" and, if it fails, whether additional consequences are warranted. Defining "pass/fail" rules requires that the underlying measurements be mapped into one of two classes:

(1) those for which the performance standard is parity with analogous incumbent LEC performance results, and

[8] In the course of establishing Tier I consequences, the rights of an individual CLEC to pursue actual damages must be retained. However, if a CLEC sought to pursue a claim for actual damages, it would be reasonable to offset the damage award by any Tier I payments it received from the ILEC for the same time period and performance areas. In addition, a CLEC must retain the right to waive Tier I claims and pursue its individually negotiated contract remedies.

(2) those for which the performance standard is an absolute level of required performance (otherwise known as a benchmark)

The differentiation is important because when parity is the standard, statistical procedures are necessary to draw conclusions regarding compliance. In such situations (which should apply to the vast majority of cases), two separate data sets are compared – one for the CLEC and one for BANY. Each data set is characterized by a mean and standard deviation. Statistical tests are used to draw a conclusion regarding the likelihood that the data sets with the observed means and standard deviations were drawn from the same population (in this case a support process for CLECs with the same quality and/or timeliness as that employed for BANY). The proper test further allows quantification of the statistical confidence that the conclusion is correct (i.e., that parity does not exist), but it does not quantify "how far out of parity" the result is when parity is not indicated.[9]

In contrast, when a benchmark serves as the performance standard, it is possible to establish a performance failure directly and assess the degree to which performance departs from the standard. As explained below, the mechanism for determining a performance failure differs for each of these types of measurement standard, but the principle governing the application of the Tier I consequence is consistent: the consequence escalates with increasing evidence of non-compliant performance.

[9] Clearly, however, when all other factor are held constant, increased statistical confidence is directly correlated with larger differences in the two means being compared.

Tier I Business Rules for Parity Measurements

1. Use the Modified z-Statistic to Determine Compliance

The determination of whether performance is compliant (i.e., equal to or better than the appropriate standard) is based on the calculation of the modified z-statistic.[10] The calculated modified z-statistic is then compared to the cumulative normal distribution table to determine the statistical confidence implicit in a conclusion that parity does not exist.[11] For example, if the resulting modified z-static value is 3.00, then the statistical confidence of a "non-parity" conclusion is 99.87%.

2. Use Permutation Analysis for Small Samples

Permutation analysis is employed for small data sets (those with 30 or fewer observations in one of the data sets to be compared) to create a probability distribution as an alternative to the cumulative normal distribution.[12] By mutual agreement, permutation analysis can also be employed for larger data sets.

---

[10] See: Local Competition Users Group - Statistical Tests for Local Service Parity, February 6, 1998, Version 1.0 for documentation of the calculation and use of the modified z-statistic, which is included as Attachment C. BANY and the PAP concur that this test is appropriate.

[11] The modified z-statistic computation provides for the CLEC mean to be subtracted from the ILEC mean. Thus, a negative z-statistic critical value presumes that worse performance exists when the CLEC mean becomes larger than the ILEC mean. For example, worse performance exists when the order completion interval for the CLEC exceeds that for the ILEC. Thus a negative z-statistic critical value is appropriate. On the other hand, for a metric like "% completed within x days", worse performance for the CLEC occurs when the metric result is smaller for the CLEC vis-à-vis the ILEC. In this case a positive z-statistic critical value is appropriate.

[12] See Attachment D for a description of the procedural steps for performing permutation analysis. Again, BANY and the PAP generally concur that permutation analysis is appropriate for data sets of this size.

3. Use a 15% Type I Error Rate (85% Confidence Level)

The threshold level to determine whether or not a performance failure exists is established by balancing Type I and Type II error.[13] This balance point is a function of the size of the CLEC data set (assuming the BANY data set is very large) and the extent to which the means for the two data sets differ (assuming that both data sets are normally distributed). Simulation comparing relatively small data sets (as would be likely for a CLEC) to a much larger data set (as would likely exist for BANY) demonstrates that the balancing of Type I and Type II error can reasonably be expected to occur in the range of 25% for "samples" with fewer than 100 data points but is about 5% for samples with 1000 data points.[14] To simplify the process and apply a compromise figure that can be universally used, 15% is proposed as the de facto balance point.

Thus, any modified z-statistic value that exceeds the value implicit in a parity violation conclusion at an 85% level of statistical confidence should constitute a performance failure for Tier I. When the standard normal distribution is utilized, this occurs when the z-statistic is less than –1.03686, assuming that a larger value of the CLEC performance measure (compared to the BANY result) indicates worse performance. No further demonstration

[13]The key consideration is balancing the probability of drawing erroneous conclusions -- either that performance is "bad" when it is actually "good" (Type I error) or that performance is "good" when it is actually "bad" (Type II error). The former error adversely impacts ILECs and the latter adversely impacts CLECs. Unfortunately, reducing the likelihood of one type of error increases the likelihood of the other type of error occurring. Thus the best means to create an equitable outcome for all parties is to balance the Type I and Type II error.

[14] See Response to Question 3 contained in AT&T Ex Parte filed in CC Docket 98-56 dated July 13, 1999.

need be made regarding a "non-compliance" determination for parity measurements.[15]

4. Increase Consequences as the Confidence in a "Non-Parity" Conclusion Increases

An appropriate means to take increased statistical confidence into consideration is to provide for higher amounts of monetary consequences as the statistical confidence in the "non-parity" conclusion increases. This is justified because (all other factors held constant) as the difference in the mean performance for the CLEC compared to the ILEC becomes larger, the absolute value of the modified z-statistic also becomes larger. Thus, it is appropriate that the performance consequence should escalate based upon the calculated value of the modified z-statistic.[16]

[15] See Attachment E for a further discussion of this position.

[16] See also MCI Worldcom and AT&T Joint Remedies Proposal Ex Parte filed in CC Docket 98-56, filed June 2, 1999, Table A of Attachment A. The PAP recognizes the soundness of this principle. However, the PAP provides for an arithmetic increase in consequences as the absolute value of the modified z score increases. This is incorrect as can be seen from examination of a cumulative normal distribution table (the basis for evaluating a z-statistic). For example, at z=1.00 (for a one-side test) the Type I error is 15.87%. At z=3.00 the Type I error is 0.13%. Breaking the interval into 20 equal part (similar to PAP proposal) illustrates the error in the PAP approach. Changing from 1.00 to 2.00 is a 50% change in the Z-statistic. But if the z-statistic is 2.00, the Type I error is 2.28% -- an 85.5% reduction in the risk of an incorrect decision [(.1587-.0228)/(.1587-.0013)]. Accordingly, rather than paying 50% of the increment, as proposed by the PAP, consistency would dictate no less than 85% of the increment be paid out. This treatment afforded by the PAP is simply another example of BANY's careful construction of the plan to minimize likelihood of paying meaningful consequences.

Table 1

Range of modified z-statistic value	Performance Designation	Applicable Consequence
greater than or equal -1.037	Compliant	$0
less than -1.037 to -1.645 (85% to 95% confidence)	Basic Failure	$2,500
less than -1.645 to -3.000 (95% to 99.87% confidence)	Intermediate Failure	$5,000
less than -3.000 (over 99.87% confidence)	Severe Failure	$25,000

Tier I Business Rules for Benchmark Measurements

1. Use a "Bright Line" Test for Benchmark Measurements

A benchmark is set to define the level of performance that is judged essential to permit competition to develop on a going-forward basis. As such, the benchmark level is at the lower range of what a viable competitive support process should be capable of delivering on a routine basis. Indeed, to assume otherwise would imply that the benchmark would not be achieved on a routine basis. In all events, because even the most tightly controlled process will produce performance outside the expected range, some margin of error is typically provided for the incumbent. Thus, the target performance is expressed as "X% meet or exceed the benchmark" where "X" is a figure set less than 100% in order to account for random variation consideration. Accordingly, a performance failure should be declared if the calculated performance is not equal to the "X%" level. For example, if the calculated result for a month was 94.5% of all orders completed within 3

days but the benchmark was 95% within 3 days, then a performance failure occurred. No subsequent application of a statistical test is appropriate.

2. Apply an Adjustment for Small Data Sets When Necessary

Because some measurement results may be calculated using small data sets, some adjustment is warranted. This need arises because the target for a particular measure with few underlying data points may be practically impossible to attain unless BANY always performs perfectly. The metric discussed in the prior paragraph can be used to illustrate the point: if only ten orders were completed in the month, then compliance would occur only if all 10 orders were (correctly) completed within three days. One order taking longer than 3 days would mean that, at best, the performance result would be 90% within 3 days, i.e., a failing performance level.

This situation is addressed through application of the following table which was calculated based upon a Type I error assumption of 15%[17]:

Table 2

CLEC Data Set Size	Benchmark Percentage Adjustments for Small Data Sets	
	90.0%	**95.0%**
5	80.0%	80.0%
6	83.3%	83.3%
7	85.7%	85.7%

[17] The table can be expanded to include all possible data set sizes from 1 to 30.

8	75.0%	87.5%
9	77.8%	88.9%
10	80.0%	90.0%
20	85.0%	90.0%
30	83.3%	90.0%
> 30	90%	95%

3. Increase Consequences for Increasingly Poor Performance

As with measurements that are judged against a parity standard, those compared to a benchmark standard should be subject to additional consequences as the performance becomes increasingly worse compared to the benchmark. The escalation is as follows (Note that "B" in Table 3, is the Benchmark Percentage as determined from Table 2):

Table 3

Range of Benchmark Result	Performance Designation	Applicable Consequence
Meets or exceeds B%	Compliant	$0
Meets or exceeds (1.5B-50)% but worse than B%	Basic Failure	$2,500
Meets or exceeds (2B-100)% but worse than (1.5B-50)%	Intermediate Failure	$5,000
Worse than (2B-100)%	Severe Failure	$25,000

Additional Tier I Business Rules Applicable to All Measurements

1. Increase Consequences for Chronic Performance Failures

Regardless of the type of measurement (parity or benchmark), if performance fails to achieve the Compliant level in consecutive reporting periods, then additional consequences should apply. The recommended treatment for chronic failures is to assess a chronic failure over-ride in the third consecutive month of non-compliant performance. When the chronic failure override applies, a consequence equal to a "severe failure" ($25,000 per chronic failure per month) should apply until such time as performance for the specific measurement result is again classified as Compliant.[18]

2. Any Additional Protection of BANY Should Be Limited To A Forgiveness Mechanism

Properly calibrated performance measures balance the risks of statistical errors and thus should eliminate any need for additional forms of protection for incumbents with respect to considerations of random variation.[19] Moreover, a procedural cap such as the one described below should allay any fears that additional protections are necessary for BANY. However, if the Commission were still to believe that additional protection for BANY were warranted (which it is not), such further measure should be limited to a system of forgivenesses. This mechanism provides for limited exemptions

[18] Alternatively, it is possible to institute consequences for repeated failures as early as the second consecutive month of failure. The amount of the consequence under such a structure would escalate more gradually. See Attachment A, Table A of MCI Worldcom and AT&T Joint Remedies Proposal Ex Parte filed in CC Docket 98-56, filed June 2, 1999.

[19] See Attachment F for further discussion of random variation and the inappropriateness of providing further mitigation if Type I and Type II error is balanced as recommended in this proposal.

from consequences that would otherwise be imposed under the terms of the above business rules. If a forgiveness mechanism is adopted (which AT&T does not advocate), it can only be reasonably applied measurements with "parity" standards.[20]

The forgiveness mechanism permits the waiver of an otherwise applicable basic consequence, subject to the following restrictions:

(1) One forgiveness is provided for each measurement result in a pre-specified period of time that reflects the assumed level of Type I error;[21]

(2) No more than a total of two forgivenesses may ever be accumulated for future application as an offset for an individual measurement result failure;[22]

(3) A forgiveness may only be used to offset a consequence applicable for the specific measurement result for which the forgiveness was originally granted;

(4) A forgiveness may only be used to offset a basic consequence (i.e., intermediate, severe and chronic consequences are always payable and not eligible for forgiveness); and

[20] Because the rationale for providing consequence offsets is the possibility of random variation, there is no justification for applying offsets to measurements that are monitored through the use of benchmarks. As explained above, random variability impacts are fully cared for in the structure of the benchmark standard, by permitting in advance a percentage of performance "misses."

[21] For example, if the Type I error is 15%, then in the long term, one random error would occur in a six month interval (computed by (1/Type I error) = 6.667). Thus, the first forgiveness would be available for each measure in the first month of the plan's operation. No additional forgiveness is then granted for any measure until another six months passes. Note that if Type I error were inappropriately set at 5%, then consistent treatment would dictate that a minimum of 20 (1/.05) months pass before the next forgiveness would be granted.

[22] Once BANY has earned two forgivenesses for a specific measure, it may not accumulate any more for that measure until it uses one of the accrued forgivenesses.

(5) Forgivenesses must be applied at the first opportunity, except that forgivenesses may not be applied to the same measurement in two consecutive months.

A forgiveness does not affect the performance designation for a month. Thus, BANY's performance is still registered as a "miss" for purposes of determining future chronic performance failures.[23]

3. A Small CLEC Adjustment Can Be Applied If Necessary

Some parties have expressed concern that "per measurement" consequences could provide a "windfall" for small CLECs. That is, if the incumbent's performance is only marginal (or worse), it is perceived that a per measurement system of consequences may create an incentive for a CLEC to enter a market for the primary purpose of receiving consequence payments, rather than serving customers. This thinking is off target for several critical reasons, including:

(1) The consequences payable to an individual CLEC are relatively small unless the incumbent's performance is extremely bad;
(2) A CLEC following such a strategy would generate relative low numbers of observations for any particular measurement, and the smaller the data set (all other things being equal) the more difficult it is to detect discrimination;
(3) Customers are not likely to remain with a CLEC that made little or no effort to provide quality service despite incumbent failings; and
(4) The largest consequences are paid under Tier II provision and are not, therefore, paid to the CLEC.

Nevertheless, if the Commission believes that a method for addressing a potential for "small CLEC windfalls" must be developed, the recommended approach is to have BANY determine the "per measurement" consequence due to individual CLECs and adjust that amount based on a table, such as the following:

CLEC size (based upon access lines)	**Adjustment Factor**
> 80% average CLEC size	1.0
> 60%, ≤ 80% avg CLEC size	0.9
> 40%, ≤ 60% avg CLEC size	0.7
> 20%, ≤ 40% avg CLEC size	0.5
> 10%, ≤ 20%avg CLEC size	0.3
≤ 10% avg CLEC size	0.1

Tier II

Tier II consequences are intended to enhance the BANY's incentives to provide performance that complies with its statutory obligations. Tier I consequences only compensate individual CLECs who actually receive discriminatory treatment from BANY. Tier II consequences are designed to counterbalance BANY's incentive to damage not just individual firms but the competitive marketplace itself. Thus, the two types of consequences are complementary, and both are necessary to achieve the intended results.

The applicability of Tier II consequences should be determined using the aggregate data for all CLECs within a particular measurement result and

[23] Appendix G sets forth the procedures for an alternate forgiveness plan that could be adopted in lieu of the one describer here.

disaggregation.[24] Except as noted below, identical business rules and measurements should be utilized as for Tier I. Thus, virtually the same data and computational processes can be utilized for both tiers. The differences are highlighted below and are due largely to a reduction of permissible Type I error to 5% at this level. The smaller Type I error is recommended for two reasons: (1) the data sets will be larger, so Type I and Type II error balance at a lower error rate; and, (2) higher consequences are proposed, so the confidence in the decision to apply a consequence should be greater.

Because Tier II consequences reflect harm to the public interest in a competitive marketplace, consequences under Tier II, unlike Tier I payments, should be paid to a public fund identified by the Commission and may be used for competitively neutral public purposes.[25]

Tier II Business Rules for Parity Measurements.

The same business rules apply under Tier II to the aggregate (or pooled) data of the individual CLECs as are employed for the individual CLEC data under Tier I, except that a 5% Type I error rate is used.[26] As a result, the applicable consequence table (Table 1 above) is modified as follows:

[24] Each occurrence counts equally in this calculation. Thus, the individual results for individual CLECs are not averaged together; rather the performance for all CLECs is pooled for each submeasurement result. Thus the pooled data analysis effectively creates a "super CLEC" for the purposes of determining Tier II consequences.
[25] Thus, under Tier II, individual CLECs are not compensated and there are no incentives for CLECs to enter the market for the purpose of receiving incentive payments.
[26] Alternative methodology exists for determining Tier II consequences. See, for example, the June 2, 1999 Joint AT&T and MCI ex parte filing made with the FCC in CC Docket 98-56.

Table 4

Range of modified z-statistic value	Performance Designation	Applicable Consequence
greater than or equal -1.645 (95% confidence)	Indeterminate	$0
less than -1.645 to -3.000 (95% to 99.87% confidence)	Market Impacting	n$5,000
less than -3.000 (over 99.87% confidence)	Market Constraining	n$25,000

Tier II Business Rules for Benchmark Measurements

The same business rules apply under Tier II to the aggregate (or pooled) data of the individual CLECs as are employed for the individual CLEC data under Tier I, except that consequences do not apply until the pooled CLEC performance results degrades to a point that is equivalent to an intermediate failure designation at the Tier I level. As with parity measures, the applicable consequences are adjusted to reflect the broader consequences of poor performance for the entire CLEC industry and the concomitant effects on the market and consumers.

Table 5

Range of Benchmark Result	Failure Designation	Applicable Consequence
Meets or exceeds (1.5B[27]-50)%	Indeterminate	$0
Meets or exceeds (2B-100)% but worse than (1.5B-50)%	Market Impacting	n$5,000
Worse than (2B-100)%	Market Constraining	n$25,000

Establishing the Value of "n" for Tier II

For both Tier II tables (Tables 4 and 5), the value for "n" should be determined based upon the most recent data for the state and company under consideration (in this case New York) relating to resold lines (Table 3.1) and UNE loops (Table 3.3) as reported in the most recent Report of Local Competition published by the FCC.[28] In effect, "n" is a multiplier for the Tier II consequence amount that takes into account, in general terms, the extent of competitive penetration within the state.[29]

[27] The value of B is determined by the departure permissible from the benchmark at Tier I levels for data sets >30. For example, if the benchmark provides a 95% level of attainment, the Tier II consequence does not apply until the pooled CLEC data indicates fewer than 92.5% of the observations met or exceeded the benchmark (that is, (1.5 * 95%)- 50%).

[28] If a company is not explicitly identified, then the aggregate result for the state would be utilized

[29] The calculation for a particular ILEC and state would be based on the most current data reported to the FCC and be as follows: (resold lines + UNE loops)/(total switched lines).

Table 6

Lines provided to CLECs/Total BANY and CLEC Lines	Value of "n"
more than 50%	0
more than 40% less than 50%	1
more than 30% less than 40%	2
more than 20% less than 30%	4
more than 10% less than 20%	6
more than 5% less than 10%	8
0% to less than 5%	10

Thus, as competition becomes established, the size of the applicable Tier II consequence is reduced to zero if BANY no longer provides a majority of the local lines to the CLECs in its serving area. Based upon current data, the current value of "n" for BANY is 10.

Other Considerations

1. Procedural Caps May Be Useful If Properly Implemented

In the course of early state consideration of consequence plans, regulators and incumbents expressed concern regarding the possible size of payments that an incumbent might be required to pay. In response, proposals were made to cap incumbents' potential liability. As a threshold matter, it should be noted that this concern reflects a tacit acknowledgement that the performance delivered by the incumbents has to date been largely non-complaint. Moreover, to the extent that any cap is considered at all, the very important difference between absolute and procedural caps must be

recognized. As shown below, if the Commission establishes any caps at all, they should be purely procedural and not place an absolute limit on the potential consequence payments due from BANY.[30]

The difference between procedural and absolute caps is significant. Absolute caps should be avoided entirely. First, such caps provide BANY with the means to evaluate the cost of market share retention through delivery of non-compliant performance. Second, absolute caps send the signal that once BANY's performance deteriorates to a particular level (i.e., reaching the absolute cap) then further deterioration is irrelevant.[31]

Procedural caps, on the other hand, establish a preset level at which BANY could seek regulatory review of the consequences that are due; however, the cap would not automatically absolve BANY of liability for a consequence. Procedural caps, therefore, avoid both of the problems of absolute caps. They do not provide BANY with the opportunity to evaluate the "cost" of retaining share through non-compliance. Likewise, they do not absolve BANY from consequences for unchecked performance deterioration.

To the extent a procedural cap is employed, it should be tailored to achieve the following:

[30] In this regard, it should be noted that the main purpose of any system of incentives is to have BANY accept its legal responsibility to perform at appropriate levels and not pay any consequences at all.

[31] Similarly, the use of weightings for individual performance measurements to determine the amount of consequences should also be avoided. Any weighting process is inherently subjective and thus arbitrary. Moreover, use of weightings may inappropriately influence the market entry mode selected by a particular CLEC. It is far superior to permit the market to determine which measures are most important by seeing what functions customers need from CLECs, and that CLECs in turn need from BANY.

(1) A meaningful level of consequences must be available before the procedural cap applies;

(2) The procedural cap should apply on a rolling twelve-month period and not to individual months;

(3) The procedural cap should not apply to Tier I consequences for the CLECs but only Tier II consequences.[32] No other caps should be applicable.

(4) To the extent that a procedural cap is exceeded, BANY must pay out consequences up to the procedural cap and put the amount in excess of the cap in an escrow account that earns a minimum interest rate as approved by the Commission;

(5) The Commission shall decide whether and to what extent the amount in excess of the procedural cap should be paid out. BANY should pay out any amount in excess of the cap, including accrued interest, according to Commission order.

The level of the procedural cap must be set high enough that meaningful incentives are immediately payable without intervention of the Commission. To permit otherwise would effectively prevent the performance consequences from being self-enforcing. It is reasonable to expect that any procedural cap should be proportionate to the size of the local market at issue. It is therefore recommended that, if a procedural cap is adopted, that it be no less than $1.00 per BANY line in the State prior to section 271 relief and at least twice that amount following section 271 relief.[33]

[32] As noted above, Tier I consequences principally act as a form of liquidated damages. Thus, there is no justification for capping such consequences whether for an individual CLEC or for the CLEC industry as a whole.

[33] SBC in Texas has agreed to a $120M annual limit for consequences where 9M lines are in service. In its PAP in New York, Bell Atlantic proposes an annual maximum exposure of $150M where 11M lines are in service. These prior amounts, expressed on the basis of a cap amount per line per month, are equivalent to $1.06 and $1.12, for Texas and New York, respectively.

2. Other Provisions Protect BANY From The Impact Of Extraordinary Events

The cut of a single cable may result in higher trouble rates and longer mean times to repair over a short period of time. This is referred to as clustering. While clustering may in fact occur, there is no particular reason to believe that any such events would result in disproportionate impacts on BANY or even the CLECs. Furthermore, there may be other events demonstrably beyond the control of BANY that may affect its service quality differently from the CLECs'. This condition does not argue that automatic exclusion should be provided for an otherwise applicable consequence. Nevertheless, BANY should not be denied protection from extraordinary impacts not anticipated in the construction of the consequence plan[34]. As a result, if such events occur, BANY should be permitted to pursue relief according to the following:

(1) BANY should notify the Commission and any potentially affected CLEC(s), using written and verifiable means of notice, of the intent to pursue an exception. Such notification must be provided before the applicable consequence is payable; otherwise BANY waives its rights.

[34] Root cause analysis should not defer payments of consequences. BANY must be liable to pay any consequences for poor performance. Completion of root cause analysis must not be a prerequisite for the delivery of payments to either the CLEC(s) or to the designated Tier II fund. Root cause analyses tend to be time consuming to conduct. While root cause analysis is desirable for long range performance improvement purposes, it is antithetical to self-enforcing consequences. Finally, the provisions set forth in the immediately preceding section provide a procedural mechanism available to BANY should after-the-fact root cause analysis indicate that a consequence was misapplied from BANY's perspective.

(2) All consequences not at issue under the exception petition must be immediately payable as provided for elsewhere in the plan. Those that are subject of the potential exemption shall be paid into an interest bearing escrow account no later than the due date applicable to the consequences that are at issue.

(3) No later than 15 calendar days following the due date of the consequences for which an exemption is sought, the incumbent shall submit to the Commission and all other affected parties all factual evidence supporting the exemption. To the extent BANY seeks proprietary protection of the information submitted, it shall employ a standard nondisclosure form, approved by the commission, before the plan is put into operation. BANY may not rely upon the lack of the proprietary form as a basis to delay the submission to the Commission, nor may the incumbent delay access to information by any CLEC that agrees to sign the standard nondisclosure form.

(4) By the later of 30 calendar days following notice by the incumbent or 15 calendar days following BANY's compliance with (3) above, interested CLECs shall file comments regarding the requested exemption. By mutual agreement, this period may be extended up to 15 calendar days.

(5) Following closure of the comment period provided in (4), if BANY and CLEC(s) have not reached a mutually agreeable settlement, the Commission shall either

(a) render a decision regarding the requested exemption, or

(b) seek further comment. The Commission shall render its decision regarding the exemption, which shall be binding on all parties,

within 90 calendar days of the payment due date of the consequences at issue.

(6) Payout of the consequences shall be according to Commission direction and liquidate the entire escrow account, including accrued interest. In addition, BANY should be responsible for reimbursing reasonably incurred legal fees of the CLECs. Such amounts should be reimbursed in the following proportion:

[1-(amount returned to the incumbent)]/total escrow balance at liquidation

As discussed in Attachment H, other steps may be taken to address potential measurement correlation issues once actual data has been gathered under the performance measurement system.

3. Additional Consequences Enforce the Operation of the Plan

Additional consequences should be applicable for other BANY failures related to performance reporting. At a minimum, consequences for the following areas of non-compliance are appropriate:

Late performance reports - If performance data and associated reports are not available to the CLECs by the due day, BANY should be liable for payments of $5,000 to a state fund for every day past the due date for delivery of the reports and data. BANY's liability should be determined based on the latest report delivered to a CLEC.

Incomplete or revised reports - If performance data and reports are incomplete, or if previously reported data are revised, then BANY should be

liable for payments of $1,000 to a state fund for every day past the due date for delivery of the original reports.

Inability to access detailed data - If a CLEC cannot access its detailed data underlying BANY's performance reports due to failures under the control of BANY, then BANY should pay the affected CLEC $1000 per day (or portion thereof) until such data are made available.

Interest on late consequence payments - If BANY fails to remit a consequence payment by the 15th business day following the due date of the data and the reports upon which the consequences are based, then it should be liable for accrued interest for every day that the payment is late. A per diem interest rate that is equivalent to BANY's rate of return for its regulated services for the most recent reporting year should apply.

Attachment A

Sufficient Disaggregation Is Essential to Permit Detection of Discrimination

A meaningful system of performance consequences cannot operate without a high-quality system of performance measurements. This requires not only a robust system of performance measurements that monitors all key aspects of market entry and ILEC support but also that the results derived from such measurements are sufficiently discrete to permit meaningful comparisons.[35]

Sufficient disaggregation is absolutely essential for accurate comparison of results to expected performance. This is true regardless of whether parity or a benchmark serves as the performance standard. Inadequate disaggregation of results means that not all key factors driving differences in performance results have been identified, which in turn interjects needless variability into the computed results. Such an outcome has two adverse effects. First, the ability to detect real differences is reduced for parity measures, because the modified z-statistic employs only the incumbent's variance in the denominator, which will increase with inappropriate averaging of dissimilar results (thus causing the calculated z-statistic to be smaller). Second, benchmark standards may be more permissive, both in terms of the absolute standard and the percentage "miss" accepted (to the extent it is factually supported at all), if the factual data underlying them are averages of widely divergent processes. Accordingly, inadequately disaggregated data impose

[35] Although some incumbents have raised vague concerns that sufficient disaggregation of results may over-burden regulators, those concerns are unfounded for two reasons. First, careful advance specification of disaggregation requirements will reduce, rather than increase, regulatory burden and permit superior quality decision making. Second, if fewer performance results are desired, statistical procedures for re-aggregating disaggregated results provide a superior approach to reliance upon overly aggregated measurement results.

very lenient targets that result in a very low probability that performance requirements will be missed.

Only incumbents, such as BANY, have access to the highly detailed information regarding their retail performance necessary to determine the level of disaggregation that is required to permit apples-to-apples comparisons. Moreover, there are analytical procedures that allow factual conclusions to be made regarding how much disaggregation is "enough."[36] Yet it is the very parties that have access to both the data and the methodology who typically rely on bare assertions, rather than fact, to oppose anything more than the most minimal disaggregation of results. Indeed, in the limited instances where CLECs have been provided access to ILEC data and at least limited public disclosure of analysis was permitted, the facts showed both that ILECs have very detailed data and that very disaggregated results comparisons are necessary to avoid bias.[37] Until such time as the incumbents provide factual information regarding disaggregation, the level specified in LCUG SQM V7.0 should be employed so as to protect the public interest.

Establishing the appropriate level of disaggregation is not a "once-and-done" undertaking. Provision can be made to review, perhaps annually, the appropriateness of the disaggregation contained in BANY's performance measurement system. In this review process, BANY may demonstrate, through data it has collected pursuant to its performance measurement system, that the existing level of disaggregation is not providing any additional insight to an assessment of its performance quality and nondiscrimination. In that same review process, individual CLECs should

[36] For example, regression procedures may provide a workable methodology for establishing the extent of disaggregation required to make accurate comparisons.
[37] See AT&T Ex Parte filed July 20, 1999 in CC Docket 98-56.

also be permitted to request additional disaggregation.[38] The party requesting a change should have the burden of showing why the proposed change is appropriate provided that all parties have equal access to detailed data necessary to support the proposal.

There should not be any presumption that additional disaggregation creates a burden, for either BANY or this Commission. For BANY and all incumbents in general, additional disaggregation (once correct implementation is validated) simply involves repetitive computation – a task readily and quickly accomplished by today's computers. Such a small and largely one-time effort is a small price to pay for the vastly improved capability to protect the prospects for competition in New York.

[38] In such cases, the requesting CLEC should be required to make its request for further disaggregation to the incumbent LEC at least three months before initiation of the review process.

Attachment B

Sample Initial Audit Plan for A Performance Measurement System

Request For Proposal To Perform An Audit Of An Incumbents Performance Measurement Processes & Data

I. Overview

To effectuate its goal of opening all telecommunications markets to competition, the Telecommunications Act of 1996 requires local exchange carriers (ILECs) to permit interconnection of their networks with the networks of competing local telephone service providers(the CLECs), to offer their retail telecommunications services for resale at wholesale rates, and to provide nondiscriminatory access to elements within their networks on an unbundled basis("unbundled network elements") so that CLECs can use such elements to provide local phone service. In order to demonstrate compliance with Sections 251 and 271 of the Telecommunications Act of 1996("the Act"), an incumbent must establish that it offers non-discriminatory access and interconnection to its network and that it provides nondiscriminatory support for total services resale, use of unbundled network elements("UNEs"), and access to OSS. An incumbent must develop and implement a comprehensive, documented performance measurement methodology that demonstrates nondiscriminatory support and access as dictated by the Act.

An incumbent must demonstrate nondiscriminatory access and support through empirical evidence of sufficient quality and quantity. Memorandum and Opinion, In the Matter of Application of Ameritech Michigan Pursuant to Section 271 of the Communications Act of 1934, as amended, To Provide In-Region, InterLATA Services In Michigan, CC Docket No. 97-298,(Aug. 19, 1997) ("Ameritech Order"). The ILEC must produce actual measurement results demonstrating that it provides access and interconnection to its competitors that is at least equal to that which the ILEC provides itself. The FCC emphasized that "it is essential for [the FCC], as both fact-finder and decision-maker, to have the empirical evidence necessary to make a reasoned and informed decision." Ameritech Order ¶212.

The CLECs and the incumbent are seeking a vendor to conduct an audit of the incumbent's performance measurement processes, data and data retention associated with pre-ordering, ordering, provisioning, maintenance

and billing. This audit represents an initial independent validation of the incumbent's performance measurement data collection, analysis, and reporting. The CLECs are seeking to retain a vendor to assist it in assessing whether the incumbent's performance measurement results are an accurate and whether what is being reported and measured matches agreed upon performance definitions.

II. Purpose

The incumbent LEC, the Competitive Local Exchange Carriers (CLECs) and the Commission in this state are seeking a telecommunications auditor to (a) develop a comprehensive audit plan that will be used to review of the performance measurement processes, data, and data retention associated with pre-ordering, ordering, provisioning, maintenance and billing, and other support provided by the incumbent to CLECs and (b) to conduct a detailed audit of those processes and data based on the designed audit plan.

The project described in this proposal will be broken into two phases. In the first, the auditor will develop the audit plan, and in the second phase the auditor will assess whether the incumbent's Performance Measurement data generation, collection, analysis, retention and reporting procedures are sound and that data generation, collection, analysis, retention and reporting are timely, accurate and complete and adhere to documentation that faithfully reflects the content of the Service Quality Measures (SQM). The proposed schedule for each of the phases is outlined below. In the response, the auditor should provide a total fixed-price response for Phase I, an estimate statement of resources for Phase II of the project, and should also break out the price for Phase I and Phase II.

A. Phase I

The audit plan, developed in this phase, must be sufficient to allow the CLECs and the Commission (and its staff), reviewing the audit results, to determine whether the incumbent's performance measurement processes and data produce results that provide the Commission empirical, accurate evidence to make a reasoned, informed and valid determination whether the incumbent is providing nondiscriminatory access and interconnection to its network and that it provides nondiscriminatory support for total services resale, use of unbundled network elements and access to OSS. At a minimum, the audit plan should address documentation validation, end-to-end transaction tracking, report generation, data retention, and data extraction which are detailed in Appendix A, namely the Proposed Audit Program Steps. In developing the plan, the auditor will need to consult with the Commission, the incumbent and CLEC's planning to provide local services in this state, and any other appropriate organizations.

Appendix A provides a high-level outline of some program steps for auditing the performance measurement processes and data. While not intended as a comprehensive list, it provides a general background as to the types of factors that must be considered in developing an audit plan. The purpose for providing Appendix A is to give potential auditors a framework for understanding the factors that must be addressed in the audit plan. Once an auditing firm is selected, the Commission, the incumbent and CLECs will make their staff available as needed to provide supplemental information and explanation.

B. Phase II

In this phase, a detailed audit of those processes and data based on the designed audit plan will be conducted. This aspect of the evaluation is conducted to address the following:

1. To determine if procedures for initially documenting and maintaining performance measurement documentation exist and conform to reasonable levels of quality and quality control.
2. To determine what supporting documentation exists for performance measures, including calculations, exclusions, performance standards and disaggregation and that such documentation faithfully reflects Commission order(s) and meets reasonable standards for clarity and completeness.
3. To determine if data calculations comply with the documentation, including any provisions for exempting particular data from calculations and that adequate classification parameters (e.g. for disaggregation of results) are reflected.
4. To determine if data collection (including appropriate sampling) is comprehensive, that appropriate data ultimately is input to the performance measurement calculations and that data excluded from any result calculation is captured and stored with a designation of the reason for exclusion.
5. To determine if detailed documentation exists for procedures to extract data from relevant data stores, whether for the incumbent or CLECs, that operational procedures adhere to the documentation, and that change control procedures are reasonable and fully implemented.
6. To determine if the performance measurement process starts with complete and accurate data.
7. To determine if sufficient documentation exists, describing the data storage, back-up, retrieval, CLEC access and proprietary information protection procedures for both detailed data and the

results produced for performance measurement reporting and that operational procedures conform to such documentation.

8. To determine if stored and reported performance measurement results are an accurate reflection of the documented methodologies.
9. To determine if the Commission ordered methodology comparing CLEC monthly results, whether for individual CLEC or CLECs in aggregate, is complete, accurately reflects Commission order(s) and that the ordered methodology is correctly applied in drawing conclusions regarding conformance of the performance to designated performance standards
10. To determine if contents of results match the specified report details represented in the incumbent's SQM
11. To determine if those measures which the incumbent may deem to be "parity by design" are in fact "parity by design".

III. Scope

.

The major support categories addressed in this review are as follows:

- Pre-Ordering
- Ordering
- Provisioning
- Maintenance
- Billing
- Trunk Group Performance
- E911
- Collocation
- Operator Services and Directory Assistance

This project seeks to validate the incumbent's performance data collection, analysis and reporting systems/processes, not only to ensure that the performance reports are accurate but, more importantly, to evaluate the overall performance measurement system. Therefore, the audit plan should enable the verification of the following during the actual audit:

- Existence of procedural documentation specifying performance measurement definitions, calculations, performance standards, exclusions, disaggregation, data sources, data acquisition and data retention procedures
- Compliance of documentation with Commission order(s) and adherence and completeness of the implementation of data collection, calculation, and retention with the documentation relied upon by the incumbent

- Implementation of ordered statistical methodology for determining compliance of the incumbent with performance requirements
- Accuracy, timeliness and completeness of reported results including data retention and protection and raw data access afforded CLECs

The audit will be conducted for the performance reporting the incumbent provides for itself, its affiliates as well as for individual CLECs and the CLEC industry in aggregate.

IV. Specific Deliverables

A. Phase I

The vendor will be expected to provide an initial, detailed audit plan document, which shall provide a comprehensive plan to verify that the incumbent performance measurement processes and data produce accurate empirical data to make valid performance determinations of the incumbent's support of CLECs. The audit plan document should, at a minimum, address the full breadth of issues addressed in Appendix A and the additional details that may be provided to the auditor by the Commission, the incumbent & the CLECs once a vendor is selected.

Prior to delivery of the final test plan, the initial test plan document produced by the auditor will be given to the CLECs, the Commission and the incumbent for a one-week comment period. At the end of the comment period, the auditor will be expected to perform a revision to the audit plan, incorporating reasonable recommended changes and additions to the audit plan. The auditor will then be expected to deliver the final audit plan document. The CLECs, the incumbent and the Commission will have the right to delay the commencement of Phase II, or terminate Phase II, up until the time Phase II commences.

B. Phase II

The auditor will be expected to validate the incumbent's performance data collection, analysis and reporting systems/processes and, not only to ensure that the performance reports are accurate but, more importantly, to evaluate the overall performance measurement system. The auditor will be expected to perform the assessment in full compliance with the audit plan produced in Phase I.

At the end of the assessment, the auditor will be expected to provide a document that includes a report on the assessment results. This report

should provide results of the validation and should specifically provide details as to where the incumbent has met requirements specified in the audit plan. The report should also highlight any areas where the incumbent's processes or data inhibit the possibility of making accurate performance determinations.

The auditor is also expected to include supporting documents that describe the underlying approach of the assessment, describe the methodology used for validation, and list the results of each validation. This supporting documentation should provide sufficient detail to allow parties to fully understand how the results of the audit were derived.

In addition to the final report, the auditor is expected to provide the following deliverables:

1. Weekly documented status reports
2. Record of audit results and CLEC, Commission and the incumbent responses. The record should identify the standard used, the facilities audited, the person(s) on the audit team, the team's findings, the team's recommendations, the date of the audit and any other relevant information

V. Additional Services

In addition to the development and implementation of the audit plan, additional auditing services are required. It is essential that weekly status reports be prepared to convey the progression and roadblocks associated with the project. The status reports will be inputs for the weekly feedback sessions in which the auditors will present the status details. Audit findings would be key elements of the shared feedback. The auditors will gather needed input from the incumbent, the CLECs and the Commission during these sessions.

VI. Schedule

The following is proposed as a schedule for the implementation of Phases I and II. Auditor responses may provide their own proposed schedules for Phases I and II, if the selected firm feels for any reason that the schedule provided herein is not achievable. If its proposed vendor schedule in the response differs from the schedule herein, the vendor should provide a rational for any such difference.

Vendor Selection

a. August 30, 1999	Issues RFP
b. September 7, 1999	Bidder Conference – answer questions
c. September 13, 1999	Vendor Proposals Due
d. September 14-17, 1999	Vendor Selection Meetings

e. September 24, 1999 Vendor Selected

Phase I

a. October 8, 1999 Initial Audit Plan Due
b. October 15, 1999 Comments on Audit Plan Due
c. October 22, 1999 Final Phase I Deliverable Due

Phase II

Phase II dates will be set upon the completion of Phase I, with the expectation that Phase II will be completed by December 1, 1999

Proposal Response

Vendors interested in responding to this RFP must submit 10 copies of the response to the Commission by the specified due date. Responses must provide a clear demonstration of the vendor's understanding of the objectives and deliverable of this engagement and illustrate the vendor's approach to meeting these objectives in a timely and comprehensive fashion. The proposal response should include the following:

a. Detailed description of the vendor's qualifications to perform Phases I and II of this engagement: Vendor should discuss its general experience in developing and implementing audit plans for clients in the telecommunications industry.

b. Detailed response on how vendor will meet each of the deliverables described for Phases I and II: The response should include some estimate of required vendor resources, as well as work break-down schedule for Phases I and II.

c. Details on the vendor's team: Vendor must provide name and credentials of the vendor team members who will be involved in both Phase I and Phase II.

d. Organization structure for vendor team: The vendor must provide the structure of its resources that will be involved in the project. If this structure differs for Phase I and Phase II, two organizational structures should be provided. The vendor should note which resources in this organizational structure will be dedicated to the project and which ones will be shared. Provide specific personnel that will work on each Phase of this project, their expected time commitment and role definition. These personnel should be available for pre-selection interviews. For any shared resources,

the vendor should specify what percentage of that resource's time will be allocated to the project.

e. Price proposal: The vendor shall provide a not-to-exceed cost in which the cost of professional services and out-of-pocket expenses are separately stated. The proposal must include the current professional fee rates for each individual. The bid shall provide a break-out of the price associated with Phase I work and the price associated with Phase II work. The vendor should detail any assumptions going into the price bid. Payments under the contract will be made according to a negotiated schedule of deliverables, with a significant portion of Phase I and II payments retained until completion of Phase II deliverables. Proposal should identify key milestones for payment.

f. Other work: The vendor shall identify each existing contract or other agreement that it has with the incumbent. or the incumbent's affiliates and shall describe any work that it or its affiliates are doing or have done for the incumbent in the past two years. The vendor shall also identify and describe any work that it or its affiliates are doing or have done for other telecommunications services providers in the past two years.

APPENDIX A – PROPOSED AUDIT PROGRAM STEPS

The following highlights some key steps that may be taken to effectively address the audit request:

Step 1 General CLEC/Incumbent Orientation

Participate in review session covering the following:

⇒ History
⇒ Impacted parties.
⇒ Audit goal and purpose
⇒ Critical timeframes
⇒ Key contacts
⇒ Available resources (e.g. office space, computer access….)

Step 2 General Understanding of the Processes

Obtain and Review the following documentation:

⇒ An incumbent's Service Quality Measures(SQM) including all Performance measurement Methods and Procedures relied upon by the incumbent (M&P's) and Supporting documentation for the measures prepared by the incumbent
⇒ Any CLEC User's Guide
⇒ Local Competition User's Group Statistical Tests for Local Service Parity Version 1.0
⇒ System documentation such as flowcharts, narratives, etc. for Pre-ordering, Ordering, Provisioning, Billing, Maintenance, Collocation, OS/DA employed by the incumbent
⇒ Any and all reports used to monitor the overall process
⇒ Data Extraction Methods & Procedures
⇒ Data Retention Methods & Procedures

Formulas overview including the following:

⇒ Define the formulas that are being used for all performance result calculations
⇒ Define the separate components (e.g. data elements) that make up each formula including their sources
⇒ Identify any and all exclusions and reasons for which they are to be considered exclusions, the basis for the exclusion, and the decision rule(s) for determining that a particular exclusion is applicable.
⇒ Describe start and end points for calculation

Step 3 General Understanding of OSS

System Information
⇒ Obtain and Review system flowcharts and narratives

Network Architecture
⇒ Interfaces to the incumbent systems
⇒ Overall architecture design

Change Management
⇒ Obtain and review change management policies and procedures (if company information is differentiated)
⇒ Obtain and review the latest relevant changes made to the system

Database Description

Step 4 End-To-End Transaction Tracking

Select Orders For Tracking

Obtain & Review The Following:
⇒ M&P's for ordering, preordering, provisioning, billing and maintenance
⇒ Review orders' progression through all applicable service areas up to the point of inclusion is data warehouse which represents the data source for performance measurement processes
⇒ Document results of tracking

Step 5 Report Generation

Obtain & Review the following:
⇒ Report generation M&P's
⇒ An incumbent SQM
⇒ Raw data associated with a specific report
⇒ Sampling methodologies

Recreate A Designated Performance Measurement Reports
⇒ Manually prepare report using acquired raw data
⇒ Compare prepared report to the incumbent version of report
⇒ Review and document results of testing

Step 6 Data Retention

Data Retention

⇒ Determine standards regarding record retention (i.e. which data elements are to be retained)

⇒ Determine sample type and size to be retained (i.e. statistical, random, judgmental, …etc.)

⇒ Review and document results of testing

Data Retention Policies and Procedures

⇒ Obtain and review record retention policies and procedures

⇒ On a sample basis test retention policies and procedures

⇒ Document results of retention policies and procedures testing

Other

⇒ Confirm data is stored in sufficient detail to permit subsequent independent review and analysis

⇒ Assess CLECs' ability to access detail in a timely and accurate manner (including security protections for individual CLEC data)

Step 7 Data Extraction

Data Extraction

⇒ Determine standards regarding record extraction (i.e. which data elements are to be extracted)

⇒ Determine sample type and size to be extracted

⇒ Review and document results of testing

Record Extraction Policies and Procedures

⇒ Obtain and review record extraction policies and procedures

⇒ On a sample basis test extraction policies and procedures

⇒ Document results of extraction policies and procedures testing

Step 8 Document findings and Issue Final Report

Issue Final Report and Findings

⇒ Document any open issues

⇒ Document adequacy of documentation, etc. (each of the eleven objectives of this audit should be addressed, including recommended corrective actions)

⇒ Document any potential claims(variance between documentation and practice) and gather documentation to support claim

⇒ Integrate report and finding into overall audit report, including any exception of the incumbent, any CLEC or the Commission

Attachment C

Local Competition Users Group

Statistical Tests for Local Service Parity

February 6, 1998
Membership: AT&T, Sprint, MCI, LCI, WorldCom

Version 1.0

EXECUTIVE SUMMARY **45**

INTRODUCTION **46**

PURPOSE 46
SERVICE QUALITY MEASUREMENTS 46
WHY WE NEED TO USE STATISTICAL TESTS 47

BASIC CONCEPTS AND TERMS **48**

POPULATIONS AND SAMPLES 48
MEASURES OF CENTRAL TENDENCY AND SPREAD 49
SAMPLING DISTRIBUTION OF THE SAMPLE MEAN 50
THE Z-TEST 51
TYPE 1 ERRORS AND TYPE 2 ERRORS 53
TESTS OF PROPORTIONS AND RATES 54

PROPOSED TEST PROCEDURES **54**

APPLYING THE APPROPRIATE TEST 54
TEST FOR PARITY IN MEANS 55
TEST FOR PARITY IN PROPORTIONS 56
TEST FOR PARITY IN RATES 57

Executive Summary

The Local Competition Users Group has drafted 27 Service Quality Measurements (SQMs) that will be used to measure parity of service provided by incumbent local exchange carriers (ILECs) to competitive local exchange carriers (CLECs). This set of measures includes means, proportions, and rates of various indicators of service quality. This document proposes statistical tests that are appropriate for determining if parity is being provided with respect to these measurements.

Each month, a specified report of the 27 SQMs will be provided by the ILEC, broken down by the requested reporting dimensions. The SQMs are to be systematically developed and provided by the ILECs as specified. Test parameters will be calculated so that the overall probability of declaring the ILEC to be out of parity purely by chance is very small. For each SQM and reporting dimension reported, the difference between the ILEC and CLEC results is converted to a *z*-value. Non-parity is determined if a *z*-value exceeds a selected critical value.

Introduction

Purpose

The Local Competition Users Group (LCUG) is a cooperative effort of AT&T, MCI, Sprint, LCI and WorldCom for establishing standards for the entry of new companies (competitive local exchange carriers, or CLECs) into the local telecommunications market. A key initiative of the LCUG is to establish measures of parity for services provided by incumbent local exchange carriers (ILECs). In short, parity means that the support ILECs provide on behalf of the CLECs is no lesser in quality than the service provided by the ILECs to their own customers.

The LCUG has drafted a document listing service quality measurements (SQMs) that must be reported by the ILECs to insure that CLECs are given parity of suppport. The SQM document has been submitted to the FCC and made available to PUCs in all 50 states and is pending approval by many of these regulatory agencies. This document has been drafted to describe statistical methodology for determining if parity exists based on the measurements defined in the SQM document.

Service Quality Measurements

The LCUG has identified 27 service quality measurements for testing parity of service. These are:

Category	ID	Description
Pre-Ordering	PO-1	Average Response Interval for Pre-Ordering Information
Ordering and Provisioning	OP-1	Average Completion Interval
	OP-2	Percent Orders Completed on Time
	OP-3	Percent Order Accuracy
	OP-4	Mean Reject Interval
	OP-5	Mean FOC Interval
	OP-6	Mean Jeopardy Interval
	OP-7	Mean Completion Interval
	OP-8	Percent Jeopardies Returned
	OP-9	Mean Held Order Interval
	OP-10	Percent Orders Held > = 90 Days
	OP-11	Percent Orders Held > = 15 Days
Maintenance and Repair	MR-1	Mean Time to Restore
	MR-2	Repeat Trouble Rate
	MR-3	Trouble Rate

	MR-4	Percentage of Customer Troubles Resolved Within Estimate
General	GE-1	Percent System Availability
	GE-2	Mean Time to Answer Calls
	GE-3	Call Abandonment Rate
Billing	BI-1	Mean Time to Provide Recorded Usage Records
	BI-2	Mean Time to Deliver Invoices
	BI-3	Percent Invoice Accuracy
	BI-4	Percent Usage Accuracy
Operator Services and Directory Assistance	OSDA-1	Mean Time to Answer
Network Performance	NP-1	Network Performance Parity
Interconnect / Unbundled Elements and Combos	IUE-1	Function Availability
	IUE-2	Timeliness of Element Performance

The Service Quality Measurements document describes the importance of each measure as an indicator of service parity. The SQM document also describes reporting dimensions that will be used to break each measure out by like factors (*e.g.*, major service group).

Why We Need to Use Statistical Tests

The Telecommunications Act of 1996 requires that ILECs provide nondiscriminatory support regardless of whether the CLEC elects to employ interconnection, services resale, or unbundled network elements as the market entry method. It is essential that CLECs and regulators be able to determine whether ILECs are meeting these parity and nondiscriminatory obligations. In order to make such a determination, the ILEC's performance for itself must be compared to the ILEC's performance in support of CLEC operations; and the results of this comparison must demonstrate that the CLEC receives no less than equal treatment compared to that the ILEC provides to its own operations. Where a direct comparison to analogous ILEC performance is not possible, the comparative standard is the level of performance that offers an efficient CLEC a meaningful opportunity to compete.

When making the comparison of ILEC results to CLEC results, it is necessary to employ comparative procedures that are based upon generally accepted statistical procedures. It is important to use statistical procedures because all of the ILEC-CLEC processes that will be measured are processes that contain some degree of randomness. Statistical procedures recognize that

there is measurement variability, and assist in translating results data into useful decision-making information. A statistical approach allows for measurement variability while controlling the risk of drawing an inappropriate conclusion (*i.e,* a "type 1" or "type 2" error, discussed in the next section).

Basic Concepts and Terms

Populations and Samples

Statistical procedures will permit a determination whether the support that the ILECs provide to CLECs is indistinguishable from the support provided by the ILECs to their own customers. In statistical terms, we will determine whether two "samples", the ILEC sample and the CLEC sample, come from the same "population" of measurements.

The procedures described in this paper are based on the following assumption: *When parity is provided, the ILEC data and CLEC data can both be regarded as samples from a common population of possible outcomes.* In other words, if parity exists, the measured results for a CLEC should not be distinguishable from the measured results for the ILEC, once random variability is taken into account. Figure 1 illustrates this concept. On the right side of the figure are histograms of two samples. In this illustration, the ILEC sample contains 200 observations (data values) and the CLEC sample contains 50. Note that the two histograms are not exactly alike. This is due to sampling variation. The assumption that parity exists implies that both samples were drawn from the same population of values. If it were possible to observe this population completely, the population histogram might appear as shown on the left of the Figure. If the samples were indeed taken from this population, histograms drawn for larger and larger samples would look more and more like the population histogram. Figure 1 shows that even when parity is being provided, there will be differences between the samples due to sampling variability. Statistical tests quantify the differences between the two samples and make proper allowance for sampling variability. They assess the chance that the differences that are observed are due simply to sampling variability, if parity is being provided.

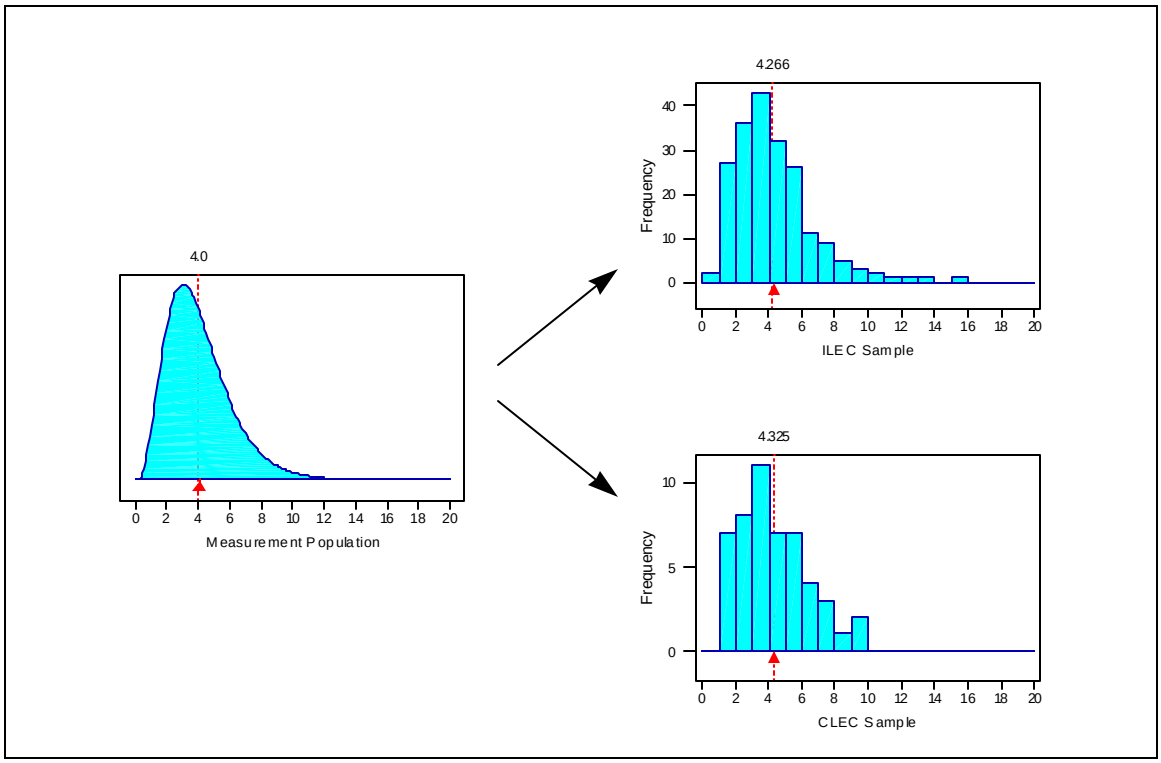

Figure 1.

Measures of Central Tendency and Spread

Often, distributions are summarized using "statistics." For the purpose of this paper, a "statistic" is simply a calculation performed on a sample set of data. Two common types of statistics are known as measures of "central tendency" and "spread."

A measure of central tendency is a summary calculation that describes the middle of the distribution in some way. The most common measure of central tendency is called the "mean" or "average" of the distribution. The mean of a sample is simply the sum of the data values divided by the sample size (number of observations). Algebraically, this calculation is expressed as

$$\bar{x} = \frac{\sum x}{n},$$

where x denotes a value in the sample and n denotes the sample size. The mean describes the center of the distribution in the following way: *If the histogram for a sample were a set of weights stacked on top of a flat board placed on top of a fulcrum (a "see-saw"), the mean would be the position along the board at which the board would balance.* (See Figure 1.) The mean in Figure 1 is indicated by the small triangle at approximately the value "4" on the horizontal axis.

A measure of spread is a summary calculation that describes the amount of variation in a sample. A common measure of spread is a called the "standard deviation" of the sample. The standard deviation is the typical size of a deviation of the observations in the sample from their mean value. The standard deviation is calculated by subtracting the mean value from each observation in the sample, squaring the resulting differences (so that negative and positive differences don't offset), summing the squared differences, dividing the sum by one less than the sample size, then taking the square root of the result. Algebraically, this calculation is expressed as

$$\sigma = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n - 1}}.$$

While the notion of mean and standard deviation exists for populations as well as samples, the mathematical definition for the mean and standard deviation for populations is beyond the scope of this paper. However, their interpretation is generally the same as for samples. In fact, for very large samples, the sample mean and sample standard deviation will be very close to the mean and standard deviation of the population from which the sample was taken.

Sampling Distribution of the Sample Mean

In Figure 1 we showed the positions of the means of the population and the two samples with triangular symbols beneath the distributions. If we sample over successive months, we will get new ILEC samples and new CLEC samples each and every month. These samples will not be exactly like the one for the first month; each will be influenced by sampling variability in a different way. In Figure 2, we show how sets of 100 successive ILEC means and 100 successive CLEC means might appear. The ILEC means can be thought of as being drawn from a population of sample means; this population is called the "sampling distribution" of these ILEC means. This sampling distribution is completely determined by the basic population of measurements that we start with, and the number of observations in each sample. The sampling distribution has the same mean as the population.

Figure 2 illustrates two important statistical concepts:

1. The histogram of successive sample means resembles a bell-shaped curve known as the Normal Distribution. This is true even though the individual observations came from a skewed distribution.
2. The standard deviation of the distribution of sample means is much smaller than the standard deviation of the observations themselves. In fact, statistical theory establishes the fact that the standard deviation on

the population of means is smaller by a factor $\sqrt{n}$, where *n* is the sample size. This effect can be seen in our example: the distribution of the CLEC means is twice as broad as the distribution of the ILEC means, since the ILEC sample size (200) is four times as large as the CLEC sample size (50).

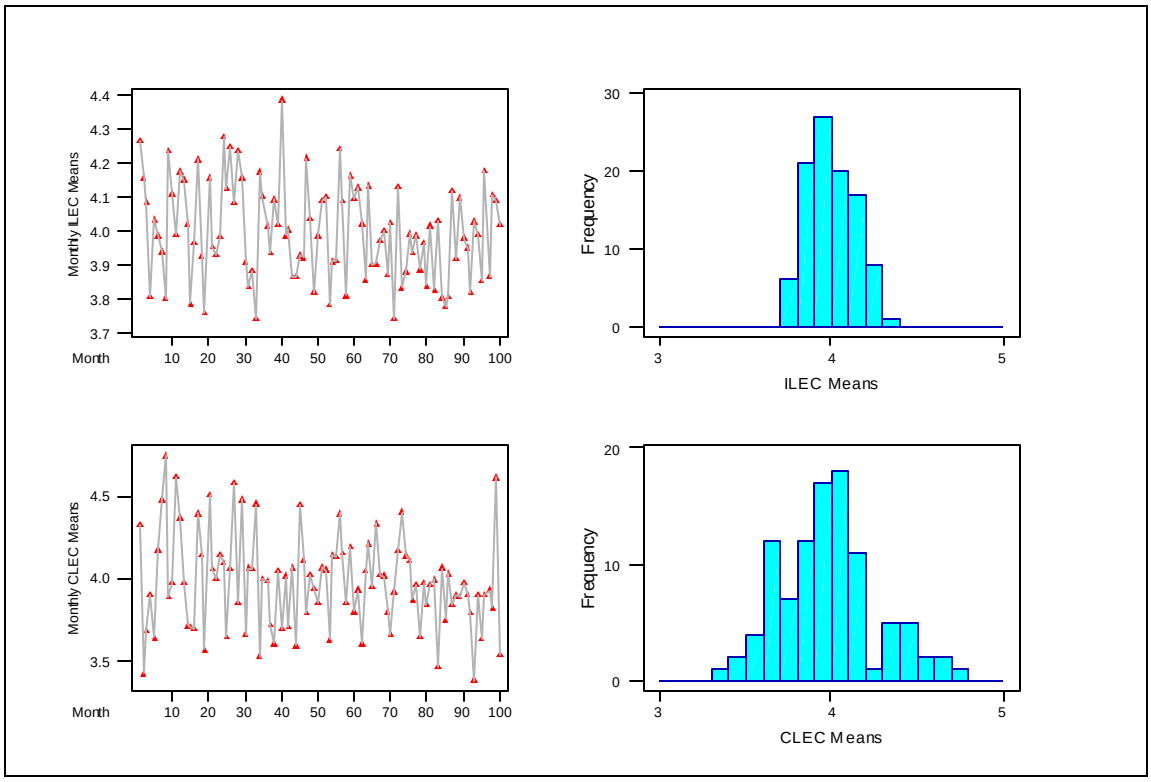

Figure 2.

It is common to call the standard deviation of the sampling distribution of a statistic the "standard error" for the statistic. We shall adopt this convention to avoid confusion between the standard deviation of the individual observations and the standard deviation (standard error) of the statistic. The latter is generally much smaller than the former. In the case of sample means, the standard error of the mean is smaller than the standard deviation of the individual observations by a factor of $\sqrt{n}$.

The Z-test

Our objective is to compare the mean of a sample of ILEC measurements with the mean of a sample of CLEC measurements. Suppose both samples were drawn from the same population; then the difference between these two sample means (*i.e.*, *DIFF* = $\bar{x}_{\text{CLEC}} - \bar{x}_{\text{ILEC}}$) will have a sampling distribution which will

(i) have a mean of zero; and

(ii) have a standard error that depends on the population standard deviation and the sizes of the two samples.

Statisticians utilize an index for comparing measurement results for different samples. The index employed is a ratio of the difference in the two sample means (being compared) and the standard deviation estimated for the overall population. This ratio is known as a z-score. The z-score compares the two samples on a standard scale, making proper allowance for the sample sizes.

The computation of the difference in the two sample means is straightforward.

$$DIFF = \bar{x}_{\text{CLEC}} - \bar{x}_{\text{ILEC}}$$

The standard deviation is less intuitive. Nevertheless, statistical theory establishes the fact that

$$\sigma^2_{\text{DIFF}} = \frac{\sigma^2}{n_{\text{CLEC}}} + \frac{\sigma^2}{n_{\text{ILEC}}},$$

where is the standard deviation of the population from which both samples are drawn. That is, the squared standard error of the difference is the sum of the squared standard errors of the two means being compared.[39]

We do not know the true value of the population because the population cannot be fully observed. However, we can estimate given the standard deviation of the ILEC sample ($_{\text{ILEC}}$).[40] Hence, we may estimate the standard error of the difference with

$$\sigma_{\text{DIFF}} = \sqrt{\frac{\sigma^2_{\text{ILEC}}}{n_{\text{CLEC}}} + \frac{\sigma^2_{\text{ILEC}}}{n_{\text{ILEC}}}} = \sqrt{\sigma^2_{\text{ILEC}}\left[\frac{1}{n_{\text{CLEC}}} + \frac{1}{n_{\text{ILEC}}}\right]}$$

If we then divide the difference between the two sample means by this estimate of the standard deviation of this difference, we get what is called a "*z*-score".

$$z = \frac{DIFF}{\sigma_{\text{DIFF}}}$$

[39] Winkler and Hays, *Probability, Inference, and Decision*. (Holt, Rinehart and Winston: New York), p. 370.

[40] Winkler and Hays, *Probability, Inference, and Decision*. (Holt, Rinehart and Winston: New York), p. 338.

Because we assumed that both samples were in fact drawn from the same population, this *z*-score has a sampling distribution that is very nearly Standard Normal, *i.e.*, having a mean of zero and a standard error of one. Thus, the *z*-score will lie between ± 1 in about 68% of cases, will lie between ± 2 in about 95% of cases, and will lie between ± 3 in about 99.7% of cases, always assuming that both samples come from the same population. Therefore, one possible procedure for checking whether both samples come from the same population is to compare the *z*-score with some cut-off value, perhaps + 3. For comparisons where the values of z exceed the cutoff value, you reject the assumption of parity as not proven by the measured results. This is an example of a statistical test procedure. It is a formal rule of procedure, where we start with raw data (here two samples, ILEC measurements and CLEC measurements), and arrive at a decision, either "conformity" or" violation".

Type 1 Errors and Type 2 Errors

Each statistical test has two important properties. The first is the probability that the test will determine that a problem exists when in fact there is none. Such a mistaken conclusion is called a type one error. In the case of testing for parity, a type one error is the mistake of charging the ILEC with a parity violation when they may not be acting in a discriminatory manner. The second property is the probability that the test procedure will not identify a parity violation when one does exist. The mistake of not identifying parity violation when the ILEC is providing discriminatory service is called a type two error. A balanced test is, therefore, required.

From the ILEC perspective, the statistical test procedure will be unacceptable if it has a high probability of type one errors. From the CLEC perspective, the test procedure will be unacceptable if it has a high probability of type two errors.

Very many test procedures are available, all having the same probability of type one error. However the probability of a type two error depends on the particular kind of violation that occurs. For small departures from parity, the probability of detecting the violation will be small. However, different test procedures will have different type two error probabilities. Some test procedures will have small type two error when the CLEC mean is larger than the ILEC mean, even if the CLEC standard deviation is the same as the ILEC standard deviation, while other procedures will be sensitive to differences in standard deviation, even if the means are equal. Our proposals below are designed to have small type two error when the CLEC mean exceeds the ILEC mean, whether or not the two variances are equal.

Tests of Proportions and Rates

When our measurements are proportions (*e.g*. percent orders completed on time) rather than measurements on a scale, there are some simplifications. We can think of the "population" as being analogous to an urn filled with balls, each labeled either 0(failure) or 1(success). In this population, the fraction of 1's is some "population proportion". Making an observation corresponds to drawing a single ball from this urn. Each month, the ILEC makes some number of observations, and reports the ratio of failures or successes to the total number of observations; the ILEC does the same does the same for the CLEC. The situation is very similar to that discussed above; however, rather than a wide range of possible result values, we simply have 0's (failures) and 1's (successes). The "sample mean" becomes the "observed proportion", and this will have a sampling distribution just as before. The novelty of the situation is that now the population standard deviation is a known function of the population proportion[41]; if the population proportion is *p*, the population standard deviation is $\sqrt{p(1-p)}$, with similar simplifications in all the other formulas.

There is a similar simplification when the observations are of rates, *e.g.*, number of troubles per 100 lines. The formulas appear below.

Proposed Test Procedures

Applying the Appropriate Test

Three z-tests will be described in this section: the "Test for Parity in Means", the "Test for Parity in Rates", and the "Test for Parity in Proportions". For each LCUG Service Quality Measurement (SQM), one or more of these parity tests will apply. The following chart is a guide that matches each SQM with the appropriate test.

Measurement (Corresponding LCUG Number)	*Test*
Preordering Response Interval (PO-1)	**Mean**
Avg. Order Completion Interval (OP-1)	**Mean**
% Orders Completed On Time (OP-2)	**Proportion**
% Order (Provisioning) Accuracy (OP-3)	**Proportion**
Order Reject Interval (OP-4)	**Mean**
Firm Order Confirmation Interval (OP-5)	**Mean**
Mean Jeopardy Interval (OP-6)	**Mean**
Completion Notice Interval (OP-7)	**Mean**
Percent Jeopardies Returned (OP-8)	**Proportion**
Held Order Interval (OP-9)	**Mean**

[41] Winkler and Hays, *Probability, Inference, and Decision*. (Holt, Rinehart and Winston: New York), p. 212.

% Orders Held ≥ 90 Days (OP-10)	**Proportion**
% Orders Held ≥ 15 Days (OP-11)	**Proportion**
Time To Restore (MR-1)	**Mean**
Repeat Trouble Rate (MR-2)	**Proportion**
Frequency of Troubles (MR-3)	**Rate**
Estimated Time To Restore (MR-4)	**Proportion**
System Availability (GE-1)	**Proportion**
Center Speed of Answer (GE-2)	**Mean**
Call Abandonment Rate (GE-3)	**Proportion**
Mean Time to Deliver Usage Records (BI-1)	**Mean**
Mean Time to Deliver Invoices (BI-2)	**Mean**
Percent Invoice Accuracy (BI-3)	**Proportion**
Percent Usage Accuracy (BI-4)	**Proportion**
OS/DA Speed of Answer (OS/DA-1)	**Mean**
Network Performance (NP-1)	**Mean, Proportion**
Availability of Network Elements (IUE-1)	**Mean, Proportion**
Performance of Network Elements (IUE-2)	**Mean, Proportion**

Test for Parity in Means

Several of the measurements in the LCUG SQM document are averages (*i.e.*, means) of certain process results. The statistical procedure for testing for parity in ILEC and CLEC means is described below:

1. Calculate for each sample the number of measurements (n_{ILEC} and n_{CLEC}), the sample means ($\bar{x}_{\text{ILEC}}$ and $\bar{x}_{\text{CLEC}}$), and the sample standard deviations ($_{\text{ILEC}}$ and $_{\text{CLEC}}$).

2. Calculate the difference between the two sample means; if *larger* CLEC mean indicates possible violation of parity, use *DIFF* = $\bar{x}_{\text{CLEC}} - \bar{x}_{\text{ILEC}}$, otherwise reverse the order of the CLEC mean and the ILEC mean.

3. To determine a suitable scale on which to measure this difference, we use an estimate of the population variance based on the ILEC sample, adjusted for the sized of the two samples: this gives the standard error of the difference between the means as

$$\sigma_{\text{DIFF}} = \sqrt{\sigma^2_{\text{ILEC}}\left[\frac{1}{n_{\text{CLEC}}} + \frac{1}{n_{\text{ILEC}}}\right]}$$

4. Compute the test statistic

$$z = \frac{DIFF}{\sigma_{\text{DIFF}}}$$

5. Determine a critical value *c* so that the type one error is suitably small.

6. Declare the means to be in violation of parity if $z > c$.

Example:

c:	3.58	Critical value for the test

ILEC			CLEC			Test	
n	mean	variance	n	mean	variance	z	Violation
250	4.038	1.9547	50	5.154	23.2035	5.15	YES!

Test for Parity in Proportions

Several of the measurements in the LCUG SQM document are proportions derived from certain counts. The statistical procedure for testing for parity in ILEC and CLEC proportions is described below. It is the same as that for means, except that we do not need to estimate the ILEC variance separately.

1. Calculate for each sample sample sizes (n_{ILEC} and n_{CLEC}), and the sample proportions (p_{ILEC} and p_{CLEC}).

2. Calculate the difference between the two sample means; if *larger* CLEC proportion indicates worse performance, use *DIFF* = $p_{CLEC} - p_{ILEC}$, otherwise reverse the order of the ILEC and CLEC proportions.

3. Calculate an estimate of the *standard error for the difference* in the two proportions according to the formula

$$\sigma_{DIFF} = \sqrt{p_{ILEC}(1 - p_{ILEC})\left[\frac{1}{n_{CLEC}} + \frac{1}{n_{ILEC}}\right]}$$

4. Hence compute the test statistic

$$z = \frac{DIFF}{\sigma_{DIFF}}$$

5. Determine a critical value *c* so that the type one error is suitably small.

6. Declare the means to be in violation of parity if $z > c$.

Example:

c:	3.58	Critical value for the test

ILEC			CLEC			Test	
num	den	p	num	den	p	z	Violation
5	250	2.00%	7	40	17.50%	6.50	YES!

Test for Parity in Rates

A rate is a ratio of two counts, *num/denom*. An example of this is the trouble rate experience for POTS. The procedure for analyzing measurements results that are rates is very similar to that for proportions.

1. Calculate the numerator and the denominator counts for both ILEC and CLEC, and hence the two rates $r_{ILEC} = num_{ILEC}/denom_{ILEC}$ and $r_{CLEC} = num_{CLEC}/denom_{CLEC}$.

2. Calculate the difference between the two sample rates; if *larger* CLEC rate indicates worse performance, use $DIFF = r_{CLEC} - r_{ILEC}$, otherwise take the negative of this.

3. Calculate an estimate of the *standard error for the difference* in the two rates according to the formula

$$\sigma_{DIFF} = \sqrt{r_{ILEC}\left[\frac{1}{denom_{CLEC}} + \frac{1}{denom_{ILEC}}\right]}$$

4. Compute the test statistic

$$z = \frac{DIFF}{\sigma_{DIFF}}$$

5. Determine a critical value *c* so that the type one error is suitably small.

6. Declare the means to be in violation of parity if $z > c$.

Example:

c:	3.58	Critical value for the test

ILEC			CLEC			Test	
num	den	rate	num	den	rate	z	Violation
250	610	0.409836	34	30	1.133333	6.04	YES!

Attachment D
Permutation Analysis Procedural Steps

Permutation analysis is applied to calculate the z-statistic using the following logic:

1. Choose a sufficiently large number T.

2. Pool and mix the CLEC and ILEC data sets

3. Randomly subdivide the pooled data sets into two pools, one the same size as the original CLEC data set (n_{CLEC}) and one reflecting the remaining data points, (which is equal to the size of the original ILEC data set or n_{ILEC}).

4. Compute and store the Z-test score (Z_S) for this sample.

5. Repeat steps 3 and 4 for the remaining T-1 sample pairs to be analyzed. (If the number of possibilities is less than 1 million, include a programmatic check to prevent drawing the same pair of samples more than once).

6. Order the Z_S results computed and stored in step 4 from lowest to highest.

7. Compute the Z-test score for the original two data sets and find its rank in the ordering determined in step 6.

8. Repeat the steps 2-7 ten times and combine the results to determine P = (Summation of ranks in each of the 10 runs divided by 10T)

9. Using a cumulative standard normal distribution table, find the value Z_A such that the probability (or cumulative area under the standard normal curve) is equal to P calculated in step 8.

10. Compare Z_A with the desired critical value as determined from the critical Z table. If Z_A > the designated critical Z-value in the table, then the performance is non-compliant.

Attachment E

Statistical Demonstrations of Non-Parity are Sufficient

Some incumbents have proposed that, when comparing the CLEC data set to the ILEC data set for a particular performance measurement result, a lack of parity should not be declared unless both the performance difference is statistically significant and the difference has "competitive or economic significance."[42] This notion is contrary to FCC's interpretation of the terms of the 1996 Act (the Act). The FCC has found that the term "nondiscriminatory" as used in the Act is a more stringent standard than the "unjust and unreasonable discrimination" standard set forth in other provisions of the Communications Act.[43] Thus, the term "nondiscriminatory access" means that: (1) the quality of performance must be equal among all carriers requesting the support, and (2) where technically feasible, the support must no less in quality and timeliness than that which the incumbent provides to itself.[44]

---

[42] Although the PAP has no such requirement, BANY has recently begun to argue that its performance scores are not directly relevant to whether it has met the requirements of Section 271 the Act, claiming that regardless of its non-parity performance in many key areas, the New York market is irreversibly open to competition. There is clearly no legal basis for this assertion.

[43] See FCC Docket No. 96-98, Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, First Report and Order released August 8, 1996, ¶ 217, 859 ("Local Competition Order").

[44] Local Competition Order, ¶315 (access must be provided on terms that are "equal to the terms and conditions under which the incumbent LEC provisions such elements to itself"); Second Order on Reconsideration, Implementation of the Local Competition Provisions of the Telecommunications Act of 1996, CC Docket No. 96-98 (released December 13, 1996) ¶9 (OSS access "must be equal to" the access that the ILEC provides to itself); FCC CC Docket No. 97-137, In the Matter of Ameritech Michigan Pursuant to Section 271 of the Communications Act of 1934, as amended, To Provide In-Region InterLATA Services in Michigan, Memorandum Opinion and Order released August 19, 1997 ("Ameritech Michigan Order"),¶139 ("BOC must provide access to competing carriers that is equal to the level of access that the BOC provides to itself . . . in terms of quality, accuracy and timeliness"); ¶166 (ILEC "must provide competing carriers access to such OSS function equal to the access that it provides to its retail operations").

Some ILECs have also argued that, as the number of data points underlying the computed performance result increases (all other factors held constant), smaller differences in means will be statistically significant. This statement is true; nevertheless, the statistical test has achieved its exact purpose by identifying unequal performance. Furthermore, the term "discriminatory" under the Act should not be confused with direct and provable competitive injury. The language of the Act does not permit the incumbent to discriminate against a CLEC by showing that no specific competitive harm was experienced by the CLEC.[45] Moreover, as a theoretical matter, it is impossible to state how large a difference in means could be permitted before the difference is considered "competitively significant." Such a proviso would immediately convert a self-enforcing plan into a regulatory quagmire. Classic economic theory supports the conclusion that *any* discriminatory performance will increase the CLEC's costs and/or degrade service and thereby impair the CLEC's ability to compete in the marketplace.

[45] Indeed, requiring a CLEC to demonstrate the specific anticompetitive consequences of an ILEC performance failure would effectively render these new protections into mere reiterations of Section II of the Sherman Act. Long experience under antitrust law shows how difficult and protracted such a requirement is in practice.

Attachment F

Mitigation for Potential Impacts of Random Variation is Unnecessary When Type I and Type II Error is Balanced

Random variation is differences in the expected output (or result) of a process that cannot be entirely explained as a result of differences in the inputs to the process. Said another way, running the very same process multiple times using exactly the same key inputs may not (and likely will not) produce exactly the same outcomes. The differences in the outcomes are "explained" as random variation.

There is little debate that the support processes that incumbents utilize to support CLECs tend to be complex and that a variety of factors influence the quantity and quality of the support delivered. As a result, provided the necessary steps have been taken to disaggregate measurement results sufficiently to account for factors correlated with different outcomes, random variation should be accommodated. In doing so, a reasonable balance needs to be struck between (1) protecting BANY from consequences that are a result of random variation, and (2) protecting competitors from the adverse effects of discrimination by BANY.

As discussed above, the first step in mitigating the effects of random variation is to minimize the risk of making an incorrect decision. In this situation, the two potential incorrect decisions are (1) declaring performance compliant when it is actually discriminatory and (2) declaring performance non-compliant when it is actually within acceptable limits. If these two risks are balanced, then, the consequences for "false" failures exactly offset the consequences for undetected failures.

Some regulators have expressed concerns, in light of what they consider to be sizable consequences necessary to motivate compliant ILEC performance and the inability to precisely balance risk, that additional mitigating factors should be instituted. Unfortunately, virtually all the mechanisms discussed are designed to protect the incumbent at the expense of the protecting the competitive process. The following mechanisms have been proposed, but each suffer from serious flaws.

a. Credits for "Better than Required" Performance Permit Gaming

This approach to mitigation is misguided and has the potential to cause extreme harm with little upside potential. In this flawed approach to mitigation, consequences for failed performance could be negated if the incumbent provides "better than required" performance at a different time (or for a different measurement) and thus earns a "credit." For example, the incumbent could deliver bad performance in one area and offset the consequence through performance credits "earned" in a separate but unrelated area or through credits for compliant performance previously (or subsequently) delivered. In all cases, such credits provide incumbents extensive opportunities to "game the system." Credits give ILECs the opportunity to deliver highly variable results that swing between very good and extremely poor performance and still be absolved of any consequence. Likewise, incumbents have the opportunity to temporarily provide compliant performance and then discriminate with impunity. In either case, the CLECs' position in the marketplace compared to the incumbent is harmed. Moreover, because CLECs only learn of "better" performance after the fact (in a performance report), they cannot take practical advantage of such

performance. Thus they get no benefit that offsets the real harm they and their customers have actually suffered.

b. Absolute Caps On Liability Are Unwarranted

There is no logical or practical basis to set an absolute limit on any incumbent's liability under any consequences plan, especially for Tier I type consequences. Such consequences are intended to compensate CLECs for actual harm they have sustained as a result of documented poor performance. Thus, there should never be a limit on this type of consequence. Moreover, to the extent that Tier II consequences become especially large, it may be appropriate to establish a procedural cap to provide an opportunity to assess whether the calculated consequence for an incumbent's market-affecting behavior should be limited.

Attachment G
Alternate Forgiveness Procedures

An alternative forgiveness mechanism is to calculate the number of performance measurements that would be expected to fail in a particular month due solely to possibility of random variation and to forgive the ILEC for up to that number of failures in calculating its liability for than month. The number of parity measurements (k) that would be expected to fail due to random variation is a function of the number of measurement results having a parity standard and the Type I error rate used in determining compliance. The following should apply to any such process:

(1) A basic failure for an individual measurement may be excluded no more than one time in any continuous six-month period;
(2) Only basic failures for individual performance results with a parity standard are eligible for exemption. Thus, intermediate, severe and chronic consequences are not eligible for exclusion, regardless of whether or not the allowable level of "k" exclusions is reached through exclusion of basic consequences; and
(3) Up to a maximum of "k" failure exemptions are permitted. Exemptions are granted to eligible measurement failures in ascending order of the consequence that would be applicable if the exemption were not provided (i.e., failed measurement consequences are exempted from the smallest to largest dollar amount until "k" is reached or all eligible failures are "excluded", whichever occurs first).

Forgiveness of a consequence under the "k" exemption methodology does not affect the performance designation for a month.

The following methodology is recommended for determining the permissible number of exclusions in a month (k):

The number k_1 of allowed individual violations, and the Type I error of each of the individual tests[46], α_1, are determined so that the probability of falsely claiming overall violation is controlled at a known level[47], which we call α.

Suppose we are aggregating N individual tests. Let K_1 be the number of these tests that indicate violations this month, and let K_2 be the number of tests that have shown violations in each of the past three months. Our proposed procedure is to claim overall violation if either (i) K_1 exceeds some number k_1, or (ii) K_2 exceeds zero. We show how k_1 and the type I error α_1 of each individual test can be determined so that the Type I error of the overall procedure is held at some desired level α.

To determine k_1 and α_1 when we know N, (the number of tests to be aggregated), and α, we proceed as follows. Throughout this calculation, we are assuming that the ILEC is fully in compliance, so that for each individual test the probability of (falsely) indicating non-parity is α_1.

a) Choose a tentative value for α_1. We start with $\alpha_1 = \alpha$. This value of α_1 will be adjusted (downwards) later.

b) Determine k_1 to be the largest number such that the probability that the overall procedure indicates violation[48] (is greater than α).

---

46 Also referred to as the size of the individual test.

47 Also referred to as the size of the overall aggregated test.

48 This probability is: $1 - (1 - \alpha_1^3)^N * P(k, N, p)$ where P(,,) is the cumulative probability of the binomial distribution. That is, P(k, N, p) is the probability that

c) Decrease α_1 until the probability of overall violation using the value of k_1 that was determined in step b), is exactly α.

The resulting value of α_1 (and the corresponding critical value on the z-score scale) is to be used in each of the individual tests. Then non-compliance is indicated if any series fails the test in three successive months, or if more than k_1 fail in any single month.

The following table provides an example of how k_1 is determined for the values N = 100 and α = 5%. As shown, the value of k_1 = 8 is the largest value of k that corresponds to a probability of no less than 5% of being exceeded. In this case, the probability of claiming an overall violation is 7.40%.

Table 1

Determination of k_1 for N = 100, α = 5%

k	Prob{$K_1 > k, K_2 > 0$} = 1 - (1 - α_1^3)^N * P(k, N, p)	
5	38.95%	
6	24.17%	
7	13.76%	
8	7.40%	← select this k for k_1
9	3.99%	

the number of false parity test failures is <= k when the probability of an individual false parity test failure is p, and where p = (α_1- α_1^3)/(1 - α_1^3).

10	2.36%

The next step is to iteratively decrease α_1 and recompute the overall probability of violation, with k_1 held at 8, until we arrive at a value for α_1 for which this probability is .05. In this case, that value of α_1 is .04601.

Now we can use the t-tables (or permutation distribution calculations) to determine the appropriate critical values for each individual test. The following Table 2 provides k_1, α_1, and critical values (assuming large sample sizes for each test) for $\alpha = .05$ and a number of values of N.

Table 2

Determination of k_1 and α_1 for a range of N

where k_1 satisfies $1 - (1 - \alpha_1^3)^N * P(k_1, N, p) = .05$

N	k_1	k_1 as a % of N	α_1	Critical Value (c)
70	6	8.57%	.0465	1.6803
80	6	7.50%	.0408	1.7411
90	7	7.78%	.0437	1.7096
100	8	8.00%	.0460	1.6849
120	9	7.50%	.0442	1.7038
140	10	7.14%	.0430	1.7170
160	12	7.50%	.0462	1.6825
180	13	7.22%	.0452	1.6937
200	14	7.00%	.0443	1.7026
250	17	6.80%	.0441	1.7046
300	20	6.67%	.0440	1.7060
400	26	6.50%	.0437	1.7095
500	32	6.40%	.0431	1.7155
600	38	6.33%	.0423	1.7247
700	44	6.29%	.0412	1.7374
800	49	6.13%	.0397	1.7543
900	55	6.11%	.0384	1.7696
1000	60	6.00%	.0371	1.7851

Attachment H

Addressing Measurement Overlap And Correlation

In advance of actual data collection, it is difficult to consider fully the issue of measurement overlap. Measurement overlap occurs when one or more measurements effectively measure the same performance. If two measurements overlap, then consequences should attach to only one of them. Note, however, a measurement addressing timeliness and a measurement addressing quality for the same area of performance do not overlap. Also, it should be noted that, given the care taken in defining measurements in LCUG SQM Version 7.0, there are no obvious areas of significant measurement overlap

Measurement correlation is different from measurement overlap. Measurement correlation occurs when one or more measurement results move at the same time. The direction of movement need not be the same. That is, one may improve (e.g., quality) while another deteriorates (e.g., timeliness). As such, measurement correlation does not automatically argue for adjustment to the measurements eligible for consequences. Indeed, an incumbent that is intentionally and pervasively discriminating would be capable of showing a high degree of correlation among all measurement results both within and across months – all results would be deteriorating.

If there are reasons to believe that measurements are somewhat overlapping and correlation is suspected, the solution is not to immediately eliminate one or both measurements. Rather the superior approach is to create “families” for the purpose of applying consequences. Each measurement “family” would be eligible for only a single consequence. Whether and to what degree a family is eligible for a consequence would be determined by the

worst performing individual measurement result within the family for the month under consideration. Thus, use of measurement families eliminates the possibility of consequence "double jeopardy"[49] without making any advance value judgement regarding the usefulness of individual measurements.

Use of measurement families has the potential for significant harm for an otherwise effective consequence plan due because: (1) inappropriate grouping can mask areas of discrimination by placing non-overlapped measurements in the same family; and, (2) by reducing eligible measurements, without adjusting the per measurement consequence, the overall plan incentives are diminished. As a result, establishment of measurement families must be approached with extreme caution and sparingly used. At least the following conditions must be imposed.

(1) measurements that address separate support functionality may not be placed in the same family;
(2) measurements that address different modes of market entry may not be placed in the same family;
(3) measurement families may not be used as a means to avoid disaggregation detail;
(4) measurements that address (a) timeliness, (b) accuracy, and (c) completeness may not be placed within the same family;
(5) measurement families, to the extent used, must be identical across all CLECs;

[49] If the measurements in the family are truly overlapping and correlated they point to the same conclusion (incidents of failure and severity). Measurement families thus treat the incumbent preferentially: either the measurements are effectively the same and only one consequence applies or they were inappropriately grouped and the incumbent avoids one or more consequences that should have been incurred.

(6) even if correlation can be demonstrated, measurement families must not be used to combine otherwise independent measurements of a deficient process; and,

(7) establishment of measurement families must not reduce the maximum consequence payable by more than 10% without an offsetting increase in the basic, intermediate, and severe consequence payable per failed measurement.

To the extent new measurement families are proposed or a proposal is set forth to eliminate or modify and existing family, the advocate of the change should bear the burden of demonstrating compliance with the above minimum requirements. The consideration should be in a public forum where all interested parties participate, and in the event of a disagreement, the Commission should decide based upon the record established. Prospective changes of measurement families should not affect any prior determinations regarding consequences.

No proposal to establish measurement families should be considered until the consequence plan has been operational and produced at least six months of independently verified data.