BEFORE THE

FLORIDA PUBLIC SERVICE COMMISSION

REBUTTAL TESTIMONY OF

ROBERT M. BELL, PH.D.

ON BEHALF OF

AT&T COMMUNICATIONS OF THE SOUTHERN STATES, INC.
WORLDCOM, INC.
DIECA COMMUNICATIONS COMPANY D/B/A COVAD COMMUNICATIONS COMPANY
NEW SOUTH COMMUNICATIONS CORP.
MPOWER COMMUNICATIONS CORP.
E.SPIRE COMMUNICATIONS, INC.
ITC^DELTACOM COMMUNICATIONS, INC.
RHYTHMS LINKS INC.

DOCKET NO. 000121-TP

MARCH 21, 2001

Q. PLEASE STATE YOUR NAME AND BUSINESS ADDRESS.

A. My name is Robert M. Bell. My business address is AT&T Labs-Research, 180 Park Avenue, Florham Park, New Jersey 07932.

Q. WHAT IS THE PURPOSE OF YOUR TESTIMONY?

A. My testimony responds to certain portions of the direct testimony of Dr. Edward Mulrow and describes errors in Exhibit DAC-6 filed by Mr. David Coon. The greatest part of my testimony discusses problems with an example that Dr. Mulrow uses to illustrate how the parameter delta should be specified.

Q. HOW SHOULD THE COMMISSION EVALUATE ALTERNATIVE VALUES FOR THE PARAMETER DELTA REQUIRED TO SET A BALANCING CRITICAL VALUE?

A. Any particular value of the parameter delta implies a specific degree of departure from parity between the service received by ALEC customers and BellSouth's retail customers. Deciding on the appropriateness of a particular value of delta for a measure should be based on a judgment about the impact that the corresponding disparity of service would have on the competitive environment. Delta should define the minimum degree of disparity that causes a material impact on competition. Thinking about real measures is the best way to make these judgments.

Q. ON PAGES 18-19 OF HIS DIRECT TESTIMONY, DR. MULROW PRESENTED AN EXAMPLE TO ILLUSTRATE THE EFFECTS OF TWO DIFFERENT VALUES OF DELTA FOR A SPECIFIC MEASURE. PLEASE SUMMARIZE THIS EXAMPLE?

A. Dr. Mulrow performs calculations for a measure, time to provision a dispatched retail order, which he assumes has a mean of 5.0 days and standard deviation of 0.5 days for BellSouth's customers. He states that a delta of 1.0 implies that the minimum difference between the ALEC and BellSouth means that would be material equals one-half the BellSouth standard deviation (0.25 days or 6 hours). He writes that a delta of 1 "means that as long as the average time taken to provide the relevant service to the ALECs did not exceed the BellSouth mean (5 days) plus one-half of the standard deviation I mentioned (half a day), the difference would not be material. That is, if the mean for the ALECs for this period were 5.25 days or less, the difference would not be material." (p. 18, lines 15-19). This leads to the question, "Is it material that BellSouth took 6 hours longer over a five-day period on average to provide service to the ALEC than to its own retail services?" (p. 19, lines 6-8). A corresponding calculation for delta equals 0.5 led to the question "Is it material that BellSouth took 3 hours longer, on average?" (p. 19, lines 8-9).

Q. IS DR. MULROW'S EXAMPLE USEFUL?

A. No. There are two problems with the example that result in the statement of misleading questions about material impact. First, Dr. Mulrow incorrectly includes a factor of one half in his calculation of the difference implied by any value of delta. Second, he assumes an implausibly small value for BellSouth's standard deviation for the time-to-provision measure. Consequently, he understates the resulting disparity by a factor of 20 or more.

Q. WHAT WAS THE FIRST PROBLEM WITH THE EXAMPLE AND WHY IS IT IMPORTANT?

A. Dr. Mulrow argues that the parameter delta should be set so that the minimum material difference equals 0.5 x delta x BellSouth's standard deviation. Including the factor of one-half violates the balancing principle because balancing occurs when the true difference in means equals delta x BellSouth's standard deviation. The Louisiana joint statistician's report implicitly defines materiality in terms of the alternative hypothesis, "If a standard of materiality is set by stating a specific alternative hypothesis for the test, ...then a critical value can be determined so that the two error probabilities are equal." (Exhibit EJM-1, page 9 of 39). That is, a material difference must be defined as delta x BellSouth's standard deviation (the difference between the BellSouth mean and the ALEC mean under the alternative hypothesis). If delta is set incorrectly, so that a difference of

one-half that size is material, then proper balancing does not occur. The probability of a Type II error when there is a difference corresponding to one-half delta remains at 50 percent, no matter how low the Type I error falls.

Q. WHAT WAS THE SECOND PROBLEM WITH THE EXAMPLE AND WHY IS IT IMPORTANT?

A. Dr. Mulrow's example assumed that BellSouth's standard deviation was 0.5 days—only one-tenth the average time to provision. Because distributions for waiting times tend to have long tails (i.e., some customers may take 30 days or more to provision), these measures would be expected to have standard deviations that exceed their means. For example, in a later example, Dr. Mulrow assumes that the standard deviation of the interval for providing service to BellSouth's retail analog is 4 days, compared with a mean of only 3 days. Real data demonstrate the same relationship (see "Qwest Performance Results (ROC 271 PID 2.0," December 21, 2000). Page 54 of the report shows monthly summary data for the measure OP-4A, "Installation Interval (Average Days) – Dispatches within MSAs" for residences. For nine of the ten reported months (January to October 2000), the CLEC standard deviation exceeds the CLEC mean (the report does not report standard deviations for Qwest customers). The ratios of the standard deviations to the means range

from 0.91 to 1.66 with a median of 1.19. Similar results hold for dispatches outside MSAs (measure OP-4B).

Consider the consequence of using a more realistic, but still conservative, standard deviation of 5 days in Dr. Mulrow's example. In that cases, a delta of 1.0 would imply a difference of 5 days between the ALEC and BellSouth means—20 times larger than the 6 hours calculated by Dr. Mulrow. Based on the improved calculation, delta equal to 1.0 implies that as long as the average time taken to provide the relevant service to the ALECs did not exceed 10 days, the difference would not be material.

The following table shows the correct degree of disparity associated with various values of delta (assuming BellSouth's standard deviation equals 5 days). For example, delta = 0.50 (second row from bottom) implies a difference of 2.5 days, so that balancing occurs with for an alternative hypothesis that the ALEC mean equals 7.50 days. Note that the disparities stated in Dr. Mulrow's testimony actually correspond to delta values of 0.025 and 0.05.

Degree of Disparity Associated with

Selected Values of Delta

Delta	Disparity (in days)	ALEC Mean (in days)
0.025	0.125	5.125
0.05	0.25	5.25
0.10	0.50	5.50
0.25	1.25	6.25
0.50	2.50	7.50
1.00	5.00	10.00

Disparity = difference in means implied by Delta (Delta x 5 days)

ALEC Mean = ALEC Mean under alternative hypothesis (BellSouth mean + Disparity)

Q. DR. MULROW REJECTS A FLOOR FOR THE BALANCING CRITICAL VALUE CONCLUDING, "BELLSOUTH WOULD BE PAYING A PENALTY EVEN THOUGH THE FOUR-DAY THRESHOLD THAT ACTUALLY REPRESENTS A MATERIAL DIFFERENCE HAS NOT BEEN MET IN THE FOURTH SET OF OBSERVATIONS." (PP. 21-22). DO YOU AGREE WITH HIS CONCLUSION?

A. No. In the line that Dr. Mulrow cites, the balancing critical value is –12.35. Consequently, unequivocal evidence of non-parity with z scores of -8, -10, or even -12 would fail to trigger a remedy. This decision would only be justified if we could be assured that the observed difference did not hinder competition. However, setting delta is not an exact science. There is no magic point at which disparities suddenly become material. Even if there were, we could not identify that point with any certainty. Consequently, a floor on the balancing critical value provides some protection against failing to trigger a remedy in the face of unequivocal, material disparity for measures with large samples, when delta has been set too high. If delta is set at 0.50 or higher, this risk is clear enough that a floor on the balancing critical value should be used. Although the same danger theoretically exists for delta = 0.25, the value used in the joint ALEC plan, the danger is sufficiently small, at current samples sizes, so that I do not anticipate a floor would be necessary. Even at current sample sizes, however, a delta of 0.50 or 1.00 would be problematic.

Q. DOES TRUNCATED Z PREVENT PARITY SERVICE IN SOME CELLS FROM CONCEALING DISCRIMINATION IN OTHER CELLS?

A. No. The truncation step, setting $Z_j^* = \min(0, Z_j)$, is designed to keep a single cell where the ALEC's customers receive much better than parity service from canceling out poor service in other cells. However, it does not prevent parity, or better, service in a large number of cells from

concealing very poor service in other cells. Suppose that BellSouth provides very poor service in a few cells (e.g., modified z scores extreme enough to rule out random variation as the explanation) and parity service in other cells being aggregated. The more parity cells that are included the greater the chance is that truncated z will not be significant. The reason is that each cell that is in parity tends to increase the expected value of the truncated z statistic (high values are take as evidence of parity). In addition, each cell that is in parity decreases the balancing critical value that truncated z must fall below to be judged significant. Similarly, parity service in just a few large cells can conceal very poor service in much smaller cells because the truncated z statistics weights the modified z scores according to sample sizes in the cells. This feature of truncated z is not a flaw in the procedure, but it can result in unintended consequences if very heterogeneous cells are aggregated.

Q. ARE THE CALCULATIONS ILLUSTRATING THE SEEM REMEDY PROCEDURE, ON PAGES 4-6 OF EXHIBIT DAC-6, CORRECT?

A. No. The ILEC sample sizes for cells 1-10, which are not provided, would be required to validate exactly the modified z, truncated z and balancing critical values. However, there is enough information available to prove that the balancing critical values shown in the tables are wrong by as much as a factor of 70. The tables all report balancing critical values of –0.21. However, for Order Completion Interval (p. 5), if the total ILEC

sample size of 50,000 is divided equally among the ten cells, the correct balancing critical value (BCV) is -14.58. If, instead, the ILEC sample is divided in proportion to the CLEC sample, the correct BCV is -14.67. Even if each ILEC cell size were only 10 (for a total of ILEC sample of 100), the correct BCV would be -4.75. Under any of these three scenarios for the correct BCV, a truncated z of –1.92 would not even approach the BCV, and no payout would be made. Consequently, all three tables give a distorted impression of the SEEM remedy procedure.

Q. DOES THAT CONCLUDE YOUR TESTIMONY?

A. Yes.