ORIGINAL

BELLSOUTH TELECOMMUNICATIONS, INC.

DIRECT TESTIMONY OF ELLIS E. SMITH

BEFORE THE FLORIDA PUBLIC SERVICE COMMISSION

DOCKET NOS. 960833-TP, 960846-TP, 960757-TP, 971140-TP, 960916-TP

NOVEMBER 13, 1997

Q. PLEASE STATE YOUR NAME, ADDRESS AND OCCUPATION.

A. My name is Ellis E. Smith. My business address is 2514 Comanche Drive, Birmingham, Alabama. I am employed by and a part owner of Three Sigma, Inc., a scientific statistical sampling consulting firm.

Q. PLEASE GIVE A BRIEF DESCRIPTION OF YOUR EDUCATIONAL BACKGROUND AND WORK EXPERIENCE.

A. I attended the University of Alabama at Tuscaloosa, where I earned a Bachelor of Science degree as well as a Master of Arts degree in Mathematics. After joining South Central Bell in 1973, I completed a series of post graduate courses in statistics at the University of Alabama in Birmingham. While obtaining my Master of Arts degree, I also taught mathematics courses at the University of Alabama at Tuscaloosa.

During my 24 years with the AT&T and BellSouth companies (South Central Bell, BellSouth Services, Inc., and BellSouth

Telecommunications, Inc.) I spent 20 years as an internal statistical consultant handling scientific sample design, statistical analysis and mathematical analysis. After my retirement from BellSouth in December, 1996, I began my present employment with Three Sigma, Inc.

While I was with South Central Bell and with BellSouth Telecommunications, Inc., I regularly attended conferences and programs with other statisticians where topics relevant to my work were presented. In addition, I attended the basic two week course, and the more advanced one week course offered by AT&T, related to statistics and statistical sampling and successfully completed both courses.

Q. WHAT IS THE PURPOSE OF YOUR TESTIMONY?

A. The present proceeding is looking at certain cost studies that have been prepared and offered by BellSouth Telecommunications, Inc. One of those studies, examining the cost of a loop, was based in part on a statistical sample which I was instrumental in developing. The purpose of my testimony is to tell the Commission about statistical sampling, to explain what I did in connection with the loop sample I mentioned above, and to share with the Commission information about the precision of the sample and what it means.

Q. CAN YOU BEGIN BY GIVING A SHORT BACKGROUND ON THE USE OF STATISTICAL SAMPLING?

A. The best way to approach this may be with examples. If a person wanted to learn something about the average height of a group of 20 people, the easiest way would be to measure the height of every person in the group, add the results together and then divide by the number of people in the group. This would yield the average height of the group. Using this process to find out something about a limited number of objects, the "universe" in statistical terms, is relatively simple.

However, if the object were to find the average height of the total population of Jacksonville, a different process would be used. Specifically, you could take a "sample" of the relevant "universe," and if properly done, a measurement derived from that "sample" should fairly represent the same measurement for the "universe" as a whole.

To continue the example, if I wanted to find the average height of people in Jacksonville, I could identify every person in the city, get them to hold still while I measured them, sum the heights, divide by the number of people, and get a resulting average. Alternatively, I could determine a proper sample which would be representative of the entire population of Jacksonville, calculate the average height of the sample,

and reach, with certain levels of precision, an estimate of the average height of people in Jacksonville.

The concept of sampling is not a new one, and I am sure that it is familiar to everyone. The difficulty comes in selecting the sample.

Q. CAN YOU EXPLAIN WHAT YOU MEAN BY YOUR LAST COMMENT?

A. The issue, basically, is determining whether the sample that has been selected is actually representative of the "universe" that is being measured. If I walk up to a McDonald's restaurant in Jacksonville, and get the people there to stand still while I measure them and calculate an average height for that particular group, I would know their average height, but, absent pure chance, I would know nothing about the average height of the people in Jacksonville, because my sample probably would not be representative of the universe I am interested in measuring.

Q. HOW DO YOU DETERMINE A REPRESENTATIVE SAMPLE WHICH CAN BE USED TO MEASURE CHARACTERISTICS OF A UNIVERSE THAT IS TOO LARGE TO MEASURE DIRECTLY?

A. The appropriate way is to take a random sample of the objects in the universe which is large enough to allow us to estimate the size of the attribute or variable in which we are interested. An attribute is a

characteristic that is either present, or not present, for a sample item (i.e., agree/not agree, yes/no, on/off, etc.) so that the sample items with the characteristic can be counted. A variable sample measures some characteristic on a continuum, (i.e. height, weight, length, cost, etc.)

While I do not intend to teach a basic course in statistics, it is easy to see that my answer suggests that there are two things which must be present. First, the sample must be determined on a random basis and, second, the sample must be large enough to allow us to determine the result with the precision we seek.

The first task is the easier one to accomplish. Generally, where there is a defined body of objects to be studied, a sample can be selected by using a random number generator to determine the starting point, and then selecting objects at intervals calculated to give the required number of objects to achieve the precision that is desired.

To illustrate this point, assume that I have ten thousand people in a group, all lined up and numbered 1 through ten thousand, and the object is to determine the average height, a variable, of the people in the group. Further, assume that I have already decided that I want my sample size to be 50 people, a decision I will talk about more in a moment. The first thing I would do is calculate the sampling interval by dividing the universe by the sample size. Here, I would get an interval of 200. Then I would use a "random number generator," which is

nothing more than a computer program or a table, to give me the number between 1 and 200 to begin with. In this example, assume I use a random number generator and it tells me to begin (and again, this is completely at random; that is the point of the exercise) with the person having number 67. Since I have ten thousand people and I need a sample size of 50, I would begin with Person Number 67. I would then select every two hundredth person, so that when I was finished, I would have a group of 50 people. This is my random sample.

Q. HOW WELL WILL YOUR SAMPLE GROUP REPRESENT THE UNIVERSE?

A. That question takes us to the second part of my analysis. Simply stated, assuming the sample is in fact a random one, the size of the sample dictates the precision with which the sample represents the universe as a whole. The logic of this is inescapable. Obviously if I selected all ten thousand people and measured them, I could obtain the exact average height of the group. If I only measured 9,999 people, I could get pretty close to the actual average, but I could be off, although probably not by much. On the other hand, if I only selected one person out of the entire ten thousand, the likelihood that my sample actually matched the average of the group would be fairly minimal.

The common error that people make, however, is thinking that this is a linear relationship. In fact, a point is reached with sample sizes where increasing the sample size simply does not add significantly to the accuracy of the answer in a manner that is cost and time efficient.

This phenomenon is really well known to most of us, if we think about it. Who has not seen a televised Presidential Election night news report where, before the polls close, the television stations are predicting a winner, based on questions, an attribute, asked of a sample of 500 people as they left the polls! How, when 50 or 60 million people are voting (if we are lucky) can they predict the results of the election? The answer is in the rest of the information that the television news report gives. Normally, in small print, they will note that the results they are projecting are accurate within "plus or minus 3 (or a similar number) percentage points." That is, if Candidate A is selected as the winner because the television station is projecting that the candidate will win 60% of the votes cast, with a possible error of 3 percentage points, what the television station (or more accurately the pollster's statisticians) is really saying is that the actual vote that Candidate A will receive will fall between 57% of the vote and 63% of the vote, with 95% reliability.

This is nothing more than what a statistician calls a "confidence statement." Normally, the statistician would say "I am 95% confident that the real result will fall within 3 percentage points of the number that

I am reporting to you." A ninety-five percent confidence interval is the level normally used, although it can be lowered or increased.

The precision of the measurement, the "plus or minus 3 points" in my election example above, can be affected by sample size. If the pollsters for the television station had chosen to only interview 50 voters, they still would have been able to make a projection, but with 50 voters, they might have had to say "We think Candidate A will win with 60% of the vote, but the real result may vary within a range of plus or minus 20 percentage points." That is, the television station would have had to conclude that it thought Candidate A would get 60% of the vote, but it would have to admit that the real answer should fall between 40% of the vote and 80% of the vote. As you can see, this range isn't very helpful because you really cannot tell whether Candidate A is going to win by a landslide or lose!

The important point to remember is that while increasing the sample size can narrow the range within which the actual result is expected to fall, increasing the sample size may have limited benefits. For instance, narrowing a confidence interval of 10% to an interval of 5% would require quadrupling the sample size. To illustrate, go back to my example where I was trying to find the average height of a group of 10,000 people. If we picked a sample of 200 people, and after measuring them I found the average height was 5 feet, 11 inches, I might be able to say that the actual average of the group of people

would be within a range from 10% below that height to 10% above that height. If I wanted to decrease the interval so that I could say that the average height of the group fell in a range within 5% of the number I calculated from the sample, I would have to increase my sample size to 800. The question that persons employing statisticians have to ask is whether the additional accuracy is worth the cost of taking the larger sample. In my illustration regarding the Election Night results, the sample size was limited to 500 voters, where the universe was 50 or 60 million voters, because the television station felt that increasing the sample size simply would not improve the confidence level enough to warrant the extra time and cost that would be involved.

Q. IF THERE IS A POINT BEYOND WHICH A LARGER SAMPLE WILL ONLY MARGINALLY IMPROVE THE RESULTS, IS THERE A LIMIT BELOW WHICH THE SAMPLE SIZE SHOULD NOT GO AS WELL?

A. Yes. Although it is not an absolute rule, I try to keep my samples above thirty, because of various statistical tests that suggest that level.

Q WITH THIS BRIEF BACKGROUND, CAN YOU TELL US WHAT YOU DID IN CONNECTION WITH THE LOOP SAMPLE THAT YOU MENTIONED EARLIER IN YOUR TESTIMONY?

A. Yes. I was asked to develop a process which would allow the company to draw a sample of the loops which could be used to represent the universe of loops as defined by the company.

I expected, consistent with previous statistical studies in which I had participated, that we would want the sample to allow us to have a precision level between five and ten percent. That is, I intended to develop a loop sample where a measured characteristic or variable of the sample, such as the average loop investment, could be said to be within a range of 5 to 10 percent of the actual average loop investment of the universe of loops. Therefore, I had to take steps to insure that a random sample was drawn, and that the sample size was large enough to allow us to obtain the precision interval that I mentioned.

Q. DID YOU DO THAT?

A. Yes I did. The random sample was easy to pull. BellSouth's Customer Records Information System (CRIS) data base contains the identity of every loop that the company has, by telephone number. All I had to do was pick the numerical position of the beginning telephone number, using a random number generator and then have every succeeding working telephone number picked at a specified interval in order to obtain a sample of the size needed. In fact, this process was followed for each of the nine BellSouth states, since the cost study this was being done for was to be developed for all nine states.

Q. HOW DID YOU SELECT THE SAMPLE SIZE THAT WOULD BE NEEDED SO THAT YOU COULD ESTABLISH THE INTERVAL YOU MENTIONED EARLIER?

A. I had an advantage there because I had access to a BellSouth loop study done back in the 1980s. I could use the statistics calculated from that study, including the precision, mean and variance, and calculate an expected sample size for our study based on the desired precision results.

However, the earlier loop sample had cut across all types of loops and was not stratified in any way. Stratification is the grouping of a universe according to specific criteria. For instance, separating a loop universe into residence loops, business loops and pay telephone loops is a form of stratification. Then a sample is selected from each stratum. This will provide results for each stratum and these results can also be weighted together to get overall results. The earlier sample was not stratified in that manner. After looking at the earlier results, I concluded that a sample size of about 175 loops representing residence customers and about 175 loops representing business customers would probably be sufficient to give me the precision interval I was looking for in those strata.

I am sure that some one might question how I could use "judgment" and get the "right" sample size, but that is not the issue. I could have simply picked any sample size, and we could have done the analysis I have been describing. If we did it with 50 loops, we would then test the precision level, just as I illustrated with my Election Night example above, and if the precision interval was too large, we would just have to expand the size of the sample, by adding additional randomly selected loops. The problem is that this adds cost, since it is very time consuming and expensive to keep analyzing loops time after time. Therefore, what I did was try to use prior information regarding sample size to estimate the sample size that I thought, based on my experience, would bring us within the desired precision intervals on the first try. In fact, I asked that 25% more, or approximately 220, loops be pulled so that the sample size could be increased if necessary to obtain the necessary precision level.

Q. WAS THE SAMPLE OF LOOPS FOR RESIDENCE AND BUSINESS LOOPS CREATED AS YOU DESCRIBED?

A. Yes, and I was then given the data associated with the loops so that I could analyze the sample information in order to determine whether the sample represented the universe within the precision levels that I mentioned earlier. The loops were identified, the detailed records were pulled and reviewed and the data from the loops in the overall sample was provided to me. I then analyzed the sample loop data, determined

the mean investment as well as the variance around the mean, and reached a conclusion, using standard statistical tools, as to the precision interval for the sample.

Q. CAN YOU GIVE US THOSE RESULTS?

A. Yes. The characteristic that we were examining was the loop investment. We were trying to determine, among residential and business loops, the average investment required for each. I determined, with a confidence level of 95%, that the actual average investment in residential loops in the universe represented by our sample fell within a range of 5.8% above or below the average investment derived from the residential sample. Similarly, I determined, with a confidence level of 95%, that the actual average investment in business loops in the universe represented by our sample fell within a range of 5.2% of the average investment determined from our business sample.

Q. WHAT COULD YOU HAVE DONE IF THE RESULTS FELL OUTSIDE OF THE PRECISION INTERVAL THAT YOU WERE SEEKING?

A. I would have simply increased the sample size, first by using the extra loops that were initially selected to see if this would have put us in the desired range. However, you should recall from my earlier example that improving the precision interval does not involve a linear

relationship, and if I had been wrong, I might have had to increase the sample size considerably more than these additional loops in order to appreciably decrease my precision level. Doing this is not without a tremendous cost, that is, the cost of having an additional number of loop records pulled, examined, recast if necessary, and run through a process to determine the investment in the additional loops.

Q. WHAT DO YOU MEAN BY YOUR COMMENT ABOUT RECASTING THE LOOPS?

A. What we are trying to do here is not only select a sample that will represent the existing universe of loops, but which will also represent the universe of loops as it will exist in the future. As I understand what we are doing, we are attempting to determine the cost of a loop using forward looking, most efficient technology. I also understand that one impact of this is that certain assumptions regarding the makeup of these forward looking loops are made, such as one that says that all loops beyond 12,000 feet in length will be carried on fiber instead of copper. If loops in the sample were more than 12,000 feet in length, but were carried on copper, the loop would have to be recast to treat it as if it were actually carried on fiber, which it presumably would be in the future.

Q. DO SUCH ADJUSTMENTS AFFECT THE REPRESENTATIVE NATURE OF THE LOOP SAMPLE?

A. No. Remember, what we are trying to do is to find a sample that represents the universe of loops under study. The universe we are trying to measure consists of loops which are built using forward looking, most efficient technology. The samples we selected, adjusted for the assumptions necessary to make them meet these criteria, represented this forward looking universe within the parameters that I have previously described in detail in almost every situation.

Q. PLEASE SUMMARIZE YOUR TESTIMONY.

A. I was asked to develop a sampling procedure to estimate the average investment for a loop in Florida. I decided that a stratified systematic sampling procedure would be an appropriate process to estimate the investment for both residence and business loops, and would also allow the weighting for a combined result in most cases. I used a previous loop study to estimate an overall sample size and then decided that a sample of about 175 loops for residence and about 175 loops for business should be adequate for current purposes. The sample was selected, recast, and the data was developed and provided to me. I analyzed these data and concluded that for almost every case the sample fell within the 5% - 10% precision range that had been the original design criterion.

Q. DOES THIS CONCLUDE YOUR TESTIMONY?

A. Yes it does.