To: Ric Willard

From: Donna Glassbrenner

Date: October 19, 2011

Subject: Sample Size for CNG Bus Investigation

Per your October 6th request, I understand you have inspected two Thomas Built School Buses, finding one to have one of its five CNG tanks damaged due to abrasion with the tank's mounting cage. You are interested in whether inspecting 30 additional buses will determine to a reasonable degree of confidence whether the fleet of 1,045 Thomas Built School Buses has a tank-related problem.

It sounds like there are possibly three types of tank-related deficiencies you are interested in:

1. abraded tanks (from rubbing against the mounting cage),
2. tanks that have been abraded to the point of "damage", and
3. tanks that whose placement within the mounting cage is outside of a certain tolerance

CNG Tanks in a Mounting Cage, downloaded from ThomasBus.com

This memo addresses the second type of deficiency (tanks abraded to the point of damage). To address the first and third, I would need the number of abraded tanks among the ten inspected so far (assuming there were some abraded to a lesser degree than would be considered "damaged"), the hanging tolerance and any other tolerances you are looking at, and the number of tanks out of tolerance among the ten inspected so far.

For this analysis, I am assuming that tank deficiencies constitute independent events. That is, I assume that knowing that. e.g., Tank #2 on Bus #1 is deficient (abraded to the point of damage) would not impact your estimation of the likelihood that Tank #3 on the same bus is similarly damaged, nor your estimation of the likelihood that a tank on another bus is damaged. If this is not a reasonable assumption, I'll need to get information on the degree to which your estimation of these probabilities would change (e.g. would it double your estimation of the likelihood?).

The upshot of the analysis is as follows:

Conclusions from the Sample Size Analysis

Based on the two buses you have inspected so far, we can already conclude there is a problem, in that we can assert with high confidence (90%) that at 5% of the bus fleet has at least one damaged tank. However based on two buses, we can only estimate the percent of buses with at least one damaged tank to within an 89-point range (3% – 92%).

Inspecting 30 additional buses would allow us to narrow the estimated range (again using 90% confidence) for the proportion of buses with at least one damaged tank (what we presume to be the main statistic of interest) to at worst a 25-point range.

The remainder of this memo presents the statistical details underlying the analysis (which might not be of riveting interest). It also contains the 95% confidence results, in case you feel more comfortable having a higher level of confidence.

What we know from the two buses inspected to date – the details

First, it is worth noting that having found one of ten tanks damaged, we can be 90% confident that the percent of tanks in the fleet that are abraded to the point of damage is at least one percent.[1] With five tanks per bus, a 1% tank damage rate translates to 5% of the 1,045-bus fleet having at least one damaged tank. [2] That is, we can be 90% confident that at least 5% of the bus fleet is driving around with at least one tank abraded to the point of damage. Similar calculations give 90% confidence upper bounds of 34% for the tank damage rate and 87% for the percent of buses with at least one damaged tank, but these are less informative. To bound from both ends with the same confidence, we can be 90% confident that the tank damage lies between 0.5% and 39%. Each individual statement is true with 90% confidence, but combining statements (e.g. asserting that 1-34% of tanks are damaged) can only be asserted with lesser confidence than 90%.

What 30 additional bus inspections would tell us – the details

Inspecting 30 additional buses would bring the total tanks inspected to 160 tanks (including the 10 tanks inspected so far). Using similar calculations[3] yields the following estimation ranges with 90% confidence. I have also included the information we know from the two buses inspected so far, for comparison.

For instance, although the current inspections allow us to estimate the proportion of buses with at least one damaged tank to within only a disappointing 89-point range (3% – 92%), the additional 30 inspections will allow us to narrow this considerably to at most 25 points: The largest margin of error would occur if 19 tanks were found on the 30 buses (for a total of 20 damaged tanks, including the one found so far), in which case the estimated range is 35% – 60%, a 25-point range. This worst-case estimation interval, as well as the other 90% confidence intervals are depicted in the plot following the table.

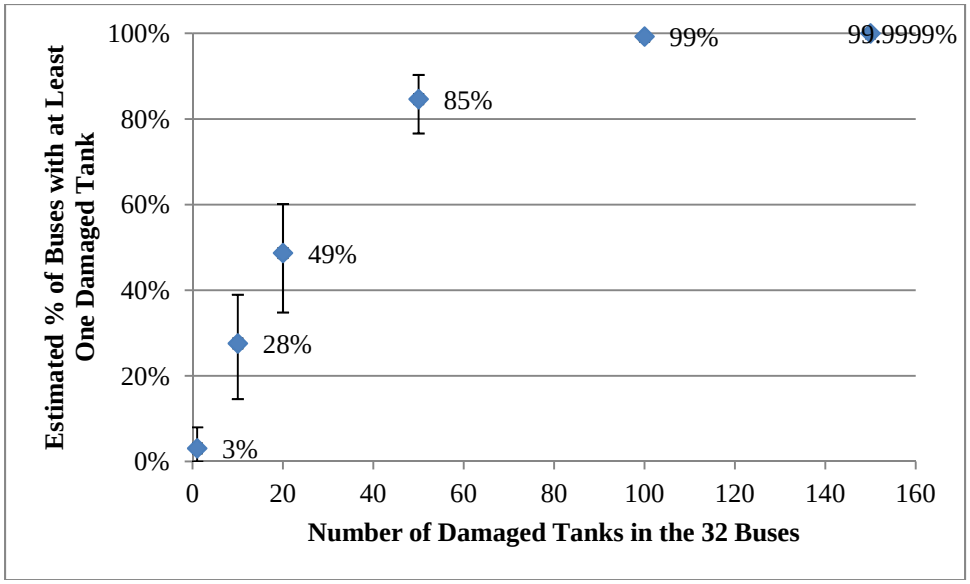
90% Confidence Bounds

Number of tanks inspected	Number of damaged tanks	Estimated tank damage rate	Confidence Ranges for the Tank Damage Rate				Confidence Ranges for the Proportion of Buses with at Least One Damaged Tank				Confidence Ranges for the Number of Buses with at Least One Damaged Tank			
			One-Sided Confidence Ranges		Two-Sided Confidence Interval		One-Sided Confidence Ranges		Two-Sided Confidence Interval		One-Sided Confidence Ranges		Two-Sided Confidence Interval	
			lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound
10	1	10%	1%	34%	0.5%	39%	5%	87%	3%	92%	54	914	26	957
160	1	0.6%	0%	1.4%	0%	1.6%	0%	7%	0%	8%	0	72	0	83
	10	6.3%	4%	9%	3%	9%	18%	37%	15%	39%	184	382	152	407
	20	13%	9%	16%	8%	17%	38%	58%	35%	60%	398	604	364	628
	50	31%	27%	36%	25%	37%	79%	89%	77%	90%	822	932	801	944
	100	63%	58%	67%	56%	69%	99%	99.6%	98%	99.7%	1,031	1,041	1,028	1,042
	150	94%	91%	96%	91%	97%	99.9995%	99.99999%	99.999%	99.999997%	1,045	1,045	1,045	1,045

[1] The probability of finding a damaged tank (at least one) in a ten-tank sample is $1 - (1-d)^{10}$, where d is the proportion of damaged tanks among the 5,225 tanks in the fleet This quantity is less than 10% when $d < .01$, meaning we can be 90% confident that at least 1% of the 5,225 tanks in the fleet are damaged from abrasion.

[2] The probability of a bus having at least one damaged tank when 1% of tanks are damaged (and damages are independent events) is: $1 - (1-0.01)^5 = 0.05$, or 5%.

[3] Although the calculations are similar in spirit, the large sample size (160 tanks) forces us to use a normal approximation to the binomial distribution, which fortunately, is very accurate for a sample this large.



90% Confidence Intervals for the Percent of Buses with at Least One Damaged Tank, as a Function of the Number of Damaged Tanks Seen in the Two Buses Inspected So Far, Plus the 30 Additional Inspections

In the event you feel more comfortable having a stronger degree of confidence, namely 95% confidence, the next table provides these. For instance, based on the 10 tanks inspected so far, we can be 95% confident that at least 2.5% of buses have one or more damaged tanks. If the inspection of 30 additional buses turns up e.g. 19 damaged tanks, we would be 95% confident that the percent of buses with a damaged tank is in the range 32% – 62%.

95% Confidence Bounds

Number of tanks inspected	Number of damaged tanks	Estimated tank damage rate	Confidence Ranges for the Tank Damage Rate				Confidence Ranges for the Proportion of Buses with at Least One Damaged Tank				Confidence Ranges for the Number of Buses with at Least One Damaged Tank			
			One-Sided Confidence Ranges		Two-Sided Confidence Interval		One-Sided Confidence Ranges		Two-Sided Confidence Interval		One-Sided Confidence Ranges		Two-Sided Confidence Interval	
			lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound
10	1	10%	0.5%	39%	0.3%	45%	2.5%	92%	1%	95%	26	957	13	992
160	1	0.6%	0%	1.6%	0%	1.8%	0%	8%	0%	9%	0	83	0	93
	10	6.3%	3%	9%	2%	10%	15%	39%	12%	41%	152	407	124	428
	20	13%	8%	17%	7%	18%	35%	60%	32%	62%	364	628	333	649
	50	31%	25%	37%	24%	38%	77%	90%	75%	91%	801	944	781	953
	100	63%	56%	69%	55%	70%	98%	99.7%	98%	99.8%	1,028	1,042	1,026	1,042
	150	94%	91%	97%	90%	98%	99.999%	99.999997%	99.999%	99.999999%	1,045	1,045	1,045	1,045