**Part 2 – Experiment and metrics design**

**Kellin Pelrine**

1. Measuring success of this experiment depends on the business goal(s) of encouraging driver partners to serve both cities, relative to the cost of reimbursing tolls. Just encouraging them to serve both cities in and of itself does not help the company. The business goal might be to increase customers, market share, profit… A clear understanding of the reasoning behind this experiment is needed to create an effective metric.

Therefore, my first step would be to acquire this understanding by communicating with the relevant colleagues, e.g. my boss, the managers of city operations who proposed the experiment, etc. If there is no consensus on the goal already, then I would work with colleagues to evaluate what we could gain here and determine the goals. After that, I would evaluate the costs (for example, the expected toll fees), estimating them as necessary. Finally, armed with both the goals and the cost, I would design a metric or metrics to evaluate success.

1. The design depends on the measure of success. In general, there are two options: compare metrics across time (before/after implementing the change), and compare metrics across groups (e.g. randomize which drivers are treated in the experiment, compare with control). This choice depends on what information we need to collect (for instance, if we only need information on drivers, comparing across groups could be a good choice), and downsides of changing policies broadly vs. for groups. Considering this may or may not be a policy we want to implement long-term, selecting a group of drivers to participate may be the best option, to avoid the risk changing full policies back and forth.

In general, that puts us in an RCT or perhaps A/B test setting. The exact tests depend on the metric(s) we’re evaluating, but there is no shortage of mathematically sound tests for these scenarios.

The key to interpreting the results will be accounting for potential confounders (e.g. pandemic changing customer patterns in the middle of the test). As long as that is possible and the metrics/experiment were well-designed according to the business goals, the experiment should deliver actionable results, especially whether to continue/stop/expand/shrink the experiment.