**Experiment and Metrics Design**

1. What would you choose as the key measure of success of this experiment in encouraging driver partners to serve both cities, and why would you choose this metric?

I would calculate the statistical significance of the difference in coverage between the experiment window prior periods using a p-value. The p-value would have to be less than 0.05. This would be a straight-forward and reliable method.

1. Describe a practical experiment you would design to compare the effectiveness of the proposed change in relation to the key measure of success. Please provide details on:

One practical method would be a Toll Amnesty Week. This would be well publicized to drivers in advance of the experiment and would involve reimbursement of tolls during weekdays for a routine week with no holidays, etc.

Since the goal is to encourage driver partners to serve both cities, we would calculate a ratio for drivers. Using trip origination data during a period, we can calculate ratios of trip originations in each city for each day.

1. how you will implement the experiment

The null hypothesis would be that driver’s ratios of trip origination cities remains the same. The alternative hypothesis would be that the ratios decrease (flatten).

The prior week or weeks would be a valid control comparison.

1. what statistical test(s) you will conduct to verify the significance of the observation

Using a permutation test we calculate the likelihood that improved results were possible considering our null hypothesis that there was no effect. If we have greater than 95% confidence that the results were not random, (p-value < .05) then we have significance.

1. how you would interpret the results and provide recommendations to the city operations team along with any caveats.

We would present significance measures and a summary of any financial impact and customer satisfaction impact observed. These measures could also have permutation tests applied.

Any predictions would include the stated caveat of the significance evaluation and well as an explanation that additional factors may come into play over time.

Overall, this ought to be a straight-forward opportunity to explore and quantify the cost and potential benefit of one way to achieve Ultimate’s goal.