Pt. 2 Experiment and metrics design

The neighboring cities of Gotham and Metropolis have complementary circadian rhythms: on weekdays, Ultimate Gotham is most active at night, and Ultimate Metropolis is most active during the day. On weekends, there is reasonable activity in both cities.
However, a toll bridge, with a two way toll, between the two cities causes driver partners to tend to be exclusive to each city. The Ultimate managers of city operations for the two cities have proposed an experiment to encourage driver partners to be available in both cities, by reimbursing all toll costs.
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?
2) 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:

a) how you will implement the experiment

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

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

Note: The two cities of Gotham and Metropolis are not in the provided dataset; however, you do not need this information to answer Part 2.

Response:

1. This is a problem of supply/demand. There is an imbalance of supply/demand in the two cities during their respective busy times. The goal is to create an equilibrium to even out this imbalance. The primary metric would be: Does a rider's wait time decrease in the busy city and a driver's wait time decrese in the slow city during the experiemnt? Overall we would also expect to see the total ride count increse (when controlled for average ride time/distance).

- There is also a more broad scope measure: Is revenue minus toll expense greater during experimental times than non-experimental times?

2. A) We would need to implement an A/B test, over a long enough period of time to control for random variables that effect supply/demand like special events, weather, and even day of the week. I would rotate every other week meaning 1 week of offering drivers (in the slow city) the reimbursement initiative and the next week, not. This could go on for 2-3 months. We would calculate the average wait time in each city during busy hours before and after the experiemnt to compare. Wait time for a rider is pick up time minus request time. Wait time for a driver is request time minus the last drop off time (assuming the driver is actively looking for a rider).

B) Our null hypothesis would be that there won't be a change in wait time for riders in the busy city or drivers in the slow city during the 'on' weeks. With a 95% confidence interval, we would have a P-value of 5% (.05), meaning that if the wait times changed in more than 5% of instances of the experiment, we would reject the null hypothesis and assume our incentive did have an effect.

C) Some caveats: A 'change' in wait time isn't necessarily good. We would want to specifically see a *decrease* in wait time during the experimental times. Another potential issue could be if too many drivers accepted the initiative and went to the busy city, creating a surplus of drivers(supply) and not achieving an equilibrium between the cities. If this was the case we could adjust the percentage of drivers that are offered the incentive until a balance is reached. If wait times of riders/drivers decreased in both cities during their respective busy hours, we would expect revenue to increase. If revenue minus the toll expense was higher than the pre-incentive era, the recommendation would be to implement the incentive full time.