Part 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?

   The key measure of success would be the increase in the number of drivers serving both cities within a given timeframe, compared to before the toll reimbursement.

   This key measure reflects whether the policy incentivized drivers to cross the toll bridge and operate in both markets — which is the ultimate goal of the experiment.

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:

   I would run an experiment tracking the number of unique drivers who operate in both cities over a 4-week period before and after the toll reimbursement policy goes into effect.

   a) how you will implement the experiment

   I would first collect baseline data on unique drivers who serve both cities, total cross-city trips, and time-of-day coverage for a 4-week control period. Then, I would launch the toll reimbursement and continue monitoring the same metrics during a 4-week post-intervention period. If possible, I would tag drivers who crossed the toll bridge to ensure we are isolating the behavior shift caused by the incentive.

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

I would use a paired t-test if a control group is available, such as drivers in a nearby city without toll changes.
If not, a time-series analysis with change point detection would help determine if the reimbursement caused a statistically significant shift in cross-city activity.

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

I would look for a statistically significant increase in cross-city driver behavior post-policy. If the effect is strong, I would recommend continuing or expanding the reimbursement program, especially during off-peak hours in both cities.

Caveats include potential confounding factors, such as weather, local events, surge pricing, uneven day-of-week behavior, and driver churn that may influence the result. I would also recommend segmenting results by time of day and weekday vs weekend to assess whether the incentive supports the complementary demand patterns between the cities.