Part 2: 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 will assume for the purpose of this question, that Ultimate's goal is to maximize revenue. One easy way of seeing if revenue has been maximized is to see if a price equilibrium has been reached in both cities. In other words, if the cost per unit distance that Ultimate charges is the same in both Gotham and Metropolis, then Ultimate is maximizing its revenue. A simple example shows why.

If most of Ultimate's drivers are in Gotham, then it might be the case that a trip in Gotham costs $5 per unit distance. Meanwhile, because there is a reduced supply of drivers in Metropolis, Ultimate might be able to charge $10 per unit distance there to eliminate unmet demand. In order to reap more revenue, Ultimate would want some of its drivers in Gotham to move to Metropolis so they can take advantage of the higher returns there. Doing this would also reduce the supply of drivers in Gotham, thereby also allowing the drivers who remain in Gotham to earn more.

Presumably, this arrival at equilibrium should happen on its own. However it might be the case that the presence of the toll distorts the market in such a way that drivers think it to be better for them to remain in one city. So, from an economist's point of view, the toll might cause drivers to behave irrationally.

It is not hard to see why. Returning to the example above, if demand leads to $10 per unit distance in Metropolis and only $5 per unit distance in Gotham, drivers should leave Gotham for Metropolis until demand evens out. However, if we introduce a $20 toll that drivers must pay, then it is not worth it for Gotham drivers to cross into Metropolis, at least for the short term. By removing this cost as a consideration for drivers, Ultimate allows drivers to think only about meeting demand. It removes the burden for drivers of calculating whether it is worth crossing to earn the extra cost per unit distance. Of course, Ultimate itself should do research into whether drivers earn enough money by crossing to offset the toll

cost. So the calculation would be, does demand remain high enough in Metropolis and do drivers complete enough trips in Metropolis to warrant crossing into Gotham (we can also add to this the extra revenue earned by the drivers remaining in Gotham). But this is beyond the scope of this problem. Here we are assuming that Ultimate has already done all this research and wants drivers to cross when a demand imbalance exists.

So in my mind, the best metric for success would be the percentage of drivers who cross the bridge from the city with less demand to the city with greater demand when there is a demand imbalance in a given time period. To perform the experiment properly, we would want to know with high confidence what percentage of drivers cross the bridge when Ultimate doesn't pay the toll so we can see how much this increases when Ultimate pays the toll. We would want this increase to be as large as possible.
If we add a little more complexity to the question, and a lot more verbosity, in order to have a more accurate answer from the revenue maximization point of view, the metric we might want to maximize would be 'the percentage of drivers who cross the bridge such that Ultimate believes that the extra revenue gained per trip by crossing more than offsets the cost of crossing'.

A secondary, related metric we could use is the time it takes to even out demand imbalances. We would expect it to be smaller when we pay the tolls than when we don't.

A number of assumptions were made in the above answer. If Ultimate has the primary goal to have drivers crossing over often, then the best measure of success would be the most obvious one: the number of times drivers cross the bridge per whatever time frame we want. This is equivalent to the expenditures Ultimate must make for paying the toll. In this case, Ultimate wants to maximize trips made and its expenditures on tolls.

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.

As mentioned in the first answer, before implementing the proposed change (paying drivers' toll fees) we would want to get a really strong estimate of what percentage of drivers already cross the bridge and pay the toll in order to take advantage of the higher demand in the other city. We would then want to see what that number changes to when we pay the toll fees. In terms of practical implementation, this should not be difficult to implement assuming Ultimate has an app that tracks drivers locations as well as the current average demand in each city (as given by price/unit distance).

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

We would use a hypothesis test comparing population proportions where the proportion in question is the percentage of drivers who cross from the city with less demand to the city with greater demand in a given frame of time. Our null hypothesis would be that the proportion is the same irrespective of our proposed change.

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

If we saw that the change in proportion (or percentage) of drivers who cross the bridge from the city with greater demand to the city with less demand in a given period was large, then we would know that paying the drivers' toll fees was an effective intervention, at least in terms of quickly meeting the demand. If we saw that this percentage change was small, then we would know that the proposed change was not effective, and we would need to provide additional incentives to drivers to cross the bridge.

My recommendations to the operations team would depend largely on some of the considerations already mentioned above. In other words, does demand remain strong enough in the city to which drivers are going and do drivers perform enough

trips in the ‘new’ city to justify the cost of having Ultimate pay the toll? Ideally there should be some analysis performed on this question before the change is implemented to know if the change should be made in the first place. But some questions can only be answered with real data, so from my perspective, to maximize Ultimate’s revenue, I would want to know the net effect on revenue from paying the tolls before recommending to city managers whether to keep the change or not.

Of course, if I was also concerned about competitors, then I might be willing to tolerate losing revenue from paying the tolls in order to harm the bottom line of my competitors and therefore increase Ultimate’s overall revenue. This would need to be incorporated into the analysis, but it would lead to a greater likelihood of me recommending to keep the change.