BUAN 6392.0W1 Assignment 2

**Group 10**

Kushal Ayyagari Satyanarayana

Melissa Cunningham

Shubham Hota

Rohith Thotakura

Faraz Mohammed Mumtaz

**1.**

Uber manages a complicated system with variables such as Interaction with customer through user interface, cancellation rates, rider’s satisfaction and pricing that might affect user behavior and business results, thus it has many options to conduct experiments. Depending on the type of improvements to be tested, it employs several types of experiments to improve its products. Every experiment type has distinct advantages and disadvantages and is used for a certain purpose.

Here are some common types of experiments Uber uses, and their pros and cons:

**A/B Tests**: Uber uses A/B testing to continually optimize its products and services to better meet the needs of its users. This is the most common type of experiment, where two different versions of a feature or product are tested on different groups of users, with one group acting as a control group. The pros of A/B testing include its simplicity, ease of implementation, and the ability to measure direct impact. However, A/B testing requires a large sample size and time-consuming as well, to produce statistically significant results and may not work for more complex experiments. A/B testing is often used to test small changes to a product or service, such as changes to the user interface, pricing, or messaging.

**Switchbacks**: In this study design, the changes are first tested on a small subset of users for a short period of time. If the results of the initial test are positive, the changes are then rolled back for a short period of time, and then reintroduced to a larger group of users. This rolling back and reintroducing process is called a "switchback." This approach allows Uber to compare the performance of the system with and without the changes, to better understand the impact of the changes on user behavior, experience, and system performance. The Switchbacks study design can be effective for testing changes that may have a significant impact on the product or service, and for minimizing the risks associated with large-scale changes. However, the approach may require a longer testing period and may be more resource-intensive compared to other study designs and fails to identify any interactions as we can really run one at a time in a city.

**Synthetic Control Tests**: To conduct a synthetic control test, Uber uses statistical modeling to create a "control group" of users who are like the group of users who were exposed to the change, but who did not receive the change. This is called a "synthetic" control group. Uber uses it to estimate what would have happened if they hadn't made the change. This allows Uber to compare the actual results they observed after making the change to the estimated results they would have seen if they hadn't made the change. The pros of synthetic control tests include the ability to test changes that cannot be easily rolled back and to generate more accurate results when a natural control group is not available. However, changes with low significance (below 5%) are less likely to be detected.

**2.**

In the dataset provided, we know the number of matched rides and double matched rides, but we don’t know what product they were from (POOL vs Express).

The proportion of total trips is split evenly between the 2 min and 5 min wait group. The proportion of total trips from the 2 min wait group was 50.6% and the 5 min wait group was 49.4%. The number of total trips taken by the 2 min wait group was 249,920 compared to the 5 min wait group with total trips at 244,445. The density plot below shows the distribution of total trips by wait time group and both groups follow a similar density.

Chart, histogram

Description automatically generated

The proportion of trips that were matched between the two wait groups was also similar. The proportion of matched trips for the 2 min wait group was 51.4% and the proportion of matched trips for the 5 min wait group was 48.6%. Changing the wait times did not drastically affect the match rate. The major difference between the groups is the sharp decline in match rate in the 5 min wait time after about 63%, where the 2 min wait group has a more gradual decline.

Chart, histogram

Description automatically generated

The driver payout per trip is higher for the 2 min wait group compared to the 5 min wait group. The average driver payout for the 2 min wait group was $7.36 and the average driver payout for the 5 min wait group was $6.99. This could be important for Uber as it affects their total cost. It is worth exploring further to see if there is a significant difference.

Chart, histogram

Description automatically generated

**3.**

Yes, I would agree with Rahematputra that Uber stood to lose $1.6m roughly if wait times were kept at 2 min in the 6 treatment cities of the launch experiment.

Based on the data and additional considerations, on the surface level I would not recommend Uber to increase wait times from 2 to 5 minutes in 6 treatment cities in the experimental launch. The actual decision requires deeper research.

Here are few considerations and reasons to support:

**2 min Wait Advantages/Disadvantages**-This approach would have a stronger sense of customer satisfaction, less prone to loss of user in long-term. The disadvantage is seat utilization wouldn’t be optimized to its full extent. Although this would lag in matching efficiency for optimized pick-up locations. JIT algorithm must be optimized further to respond efficiently.

**5 min Wait Advantages/Disadvantages**- Seat utilization could be utilized to its full extent. This would ace in matching efficiency for optimized pick-up locations. The downside is this approach would have a weaker sense of customer satisfaction, more prone to loss of potential users in the long-term.

The experimentation explicitly set up for this problem is Synthetic control Experimentation. Although this experimentation allows flexibility in alternating the wait times for rider experimentation, the major setback is that it generates significant results over a span which had been contradicting its results.

As mentioned in the article , in 3 months of experimentation, if they find that Express is performing poorly in Denver, they won’t be able to say with certainty whether it is due to defects with Express or because people in those markets are performing poorly to the increased wait times.

Additional considerations, such as network conditions could have had an impact on the response. Chances could be that network conditions in the 2-minute wait zone have been better compared to the 5-minute wait zone which could have a negative impact on the customer’s view.

They also need to consider the lifetime value of the customer. Increased wait times could lead to lower long-term earnings from dissatisfied customers.

**Rush and Non-rush hours: -**Timing could affect experiment results. At rush hour times, customers are less likely to want to wait for a ride if they are in a hurry. Uber should consider keeping wait times at 2 minutes for rush hour and possibly experiment with 5-minute wait times during non-rush hours.

**Local market conditions and user preferences: -** It might vary significantly when considered in the treatment group for experimentation. Hence Uber must carefully review both views and decide based on long-term revenue along with customer satisfaction.