

Part 2 - Experiment and Metrics Design

Key Measure of Success

The key metric for evaluating the success of reimbursing toll costs should be the percentage increase in cross-city trips (i.e., trips where a driver starts in one city and ends in the other).

Why?

- If the toll reimbursement is effective, more drivers should be willing to cross between Gotham and Metropolis.
- Measuring the proportion of cross-city trips before and after the experiment will indicate whether drivers are now serving both cities instead of being exclusive to one.
- This metric is directly linked to the experiment's goal of encouraging driver flexibility across cities.

Additional supporting metrics:

- Number of unique drivers serving both cities before and after the experiment.
- Change in driver earnings from cross-city trips.
- Passenger wait times in each city (a decrease in wait times may indicate increased driver availability).

Experiment Design

To assess the impact of the toll reimbursement, we will conduct an **A/B test (Randomized Controlled Trial - RCT)**.

a) Implementation of the Experiment

1. Randomly split drivers into two groups:
 - **Control Group (A):** No toll reimbursement (business as usual).
 - **Treatment Group (B):** Toll costs reimbursed for cross-city trips.
2. **Monitor both groups for a fixed period** (e.g., 4-6 weeks) and collect key metrics:
 - Percentage of trips that cross between Gotham and Metropolis.
 - Unique drivers making cross-city trips.

- Total driver earnings.
- Rider wait times.

3. **Ensure randomization is balanced**, meaning:

- Each group has a similar number of active drivers.
- Groups should have similar prior behaviors (e.g., cross-city trip frequency before the experiment).
-

b) Statistical Tests for Significance

To determine whether the reimbursement has a meaningful impact, we will conduct:

1. **Two-Sample Proportion Test (Chi-Square or Z-Test)**

- **Metric:** % of cross-city trips before and after the experiment.
- **Null Hypothesis (H_0):** The proportion of cross-city trips is the same in both groups.
- **Alternative Hypothesis (H_1):** The treatment group has a significantly higher proportion of cross-city trips.

2. **T-Test for Driver Earnings**

- Comparing **average earnings of drivers in each group** to see if those receiving reimbursement earn significantly more.

3. **Mann-Whitney U-Test (if distributions are non-normal)**

- If the earnings or trip distribution is skewed, we use this **non-parametric test** instead of a t-test.

c) Interpreting Results & Recommendations

- If cross-city trips increase significantly in the treatment group, the policy is successful and should be expanded.
- If driver earnings increase without worsening wait times, this suggests that drivers benefit without negatively impacting rider experience.
- If the difference is not statistically significant, it may mean:

- Toll costs were not the main deterrent for cross-city trips.
- Other factors (e.g., rider demand imbalance, safety concerns) are limiting driver flexibility.

Caveats & Considerations

1. **External Factors:** Seasonality, special events, or economic conditions may influence results.
2. **Driver Behavior Adaptation:** Some drivers may take advantage of the reimbursement but still avoid cross-city trips.
3. **Long-Term Impact:** A short-term experiment may not capture the **long-term sustainability** of the incentive.

Final Recommendation

If the toll reimbursement **significantly increases cross-city trips**, it should be **fully implemented**, potentially adjusting other incentives (e.g., higher pay for cross-city trips) to further encourage multi-city driving.