# *Uber Express*: A Case of Study

Mariana Ávalos Arce
Universidad Panamericana
0197495@up.edu.mx

Daniel Heráclito Pérez Díaz
Universidad Panamericana
0190575@up.edu.mx

**Abstract**

In the following document an analysis is presented about Uber's product, Express, which was launched to a polarizing market and management team, since it was similar to any Pool service, with only the setbacks being two: the service required the user's walking to a pick-up point and the waiting time for the algorithm to match as many users as possible. The analysis is a brief comparison between the experiment result of Express with two minutes of waiting time against five minutes, some suggestions and overview of the experiment in general. The data set and details about Uber's experiments where taken from the article cited [1].

## 1 Introduction

Uber was the first company to explore the ride-sharing industry. Matching idle drivers and passengers willing to pay a fee for a ride. Since then, several others competitors have arisen around the globe, offering basically the same on-demand service, with barely small differences. But an increasing competence, usually implies the need of changes and innovation within a company and Uber reacted by adding multiple services inside its platform, at the beginning, the main differences between them were merely about the comfort level, luxury and/or number of places, but back in 2014, they developed UberPOOL (POOL), a service that allows passengers to share the same driver to get cheaper rides.

The central idea of POOL was that passengers going from point A to point B could share the same ride for half of the price, increasing the seat occupation and also the company's revenue. But, the "greedy algorithm" they implemented for the matches and the door-to-door model that they were using caused trips to get many unnecessary loops and sometimes even to go back just to pick up another passenger, with left a very negative feeling about the service in both passengers and riders. By 2016, the POOL service remains unprofitable and its popularity continues to go down, so, in 2017, a new idea was developed: Uber Express [1].

In contrast to POOL, Express was planned to establish pick-up points in order to minimize the deviations from the main route and added a waiting time, wondering that the algorithm makes better matches. This two features were adopted from previous tests that the company had been running in different cities and after a merge from the shared-rides and marketplace teams, the Express Project had born. Years before than the creation of this project, the importance of data science had grown at Uber, so its no surprise the relevance it took on the Express' tests.

The three main kinds of experiments that Uber runs to test new features: **a) User level A/B:** in this type of experiments, the users are random and simultaneously allocated in the testing or control group. **b) Switchbacks:** in this case, the experiments are turned on for a certain time lapse, and then turned off for the same period, so within a day, all users are switched from testing or treatment group to control

each time lapse. **c) Synthetic control experiments:** this is the last kind of experiment and consists in comparing different cities using some cities as control and some cities as treatment. The tests whose info is available so far were Switchbacks, and those that were ran after in the 12 cities group were Synthetic control experiments.

The result of those tests, was the data set that was used for this analysis, whose columns are described below:

- **city_id (String):** Location where the experiment took place. In the data it is always equal to "Boston."
- **period_start (Date):** Start date and time for the 160-minute time period of the current observation.
- **wait_time (String):** This variable takes on two possible values: "2 mins" if the matching algorithm let riders wait up to 2 minutes during the current time period; "5 mins" if the matching algorithm let riders wait up to 5 minutes during the current time period.
- **treat (Boolean):** This variable takes on two possible values: "TRUE" if wait_time equals "5 mins"; "FALSE" if wait_time equals "2 mins".
- **commute (Boolean):** This variable takes on two possible values: "TRUE" if the time period happens during rush hours (7-9:40AM or 3-5:40PM), "FALSE" otherwise.
- **trips_pool (Numeric):** Total number of POOL trips completed in the current time period. Each matched ride request is a separate trip.
- **trips_express (Numeric):** Total number of Express POOL trips completed in the current time period. Each matched ride request is a separate trip.
- **rider_cancellations (Numeric):** Total number of requested trips that were cancelled by the rider in the current time period.
- **total_driver_payout (Numeric):** Total dollars paid to drivers for trips completed in the current time period. This is equal to Uber's total costs for matching trips in the current time period.
- **total_matches (Numeric):** Number of completed trips during the current time period that were paired with at least another rider for part of the trip. Each matched ride request is a separate trip, so two separate riders matched together would count as two matches.
- **total_double_matches (Numeric):** Number of completed trips during the current time period that were paired with at least two other riders for part of the trip. Each matched ride request is a separate trip, so three separate riders matched together would count as three double matches.

## 2 The effects of waiting time

Uber's main experiment involved switching between 2 minutes and 5 minutes of waiting time, to see its effects on different quantitative variables. If we wanted to know **what was the effect of extending the waiting time from 2 to 5 minutes on the total of completed trips, proportion of trips correctly matched and driver's earnings**, we would need to make a comparison between these three variables when the app's waiting time was 2 and 5 minutes.

Such comparison is shown in Figure 1, where the three plots correspond to trips, matched trips and driver's pay, each compared with 2 and 5 minutes of waiting time. The first plot is a stacked horizontal histogram of trips, which contain three categories: Express Trips, Pool Trips and Cancelled trips, where the last two were added for contrast purposes. We would read the plot, and thus the effects on the first variable, as follows: the number of Express trips decreased by 7.4% with 5 minutes of waiting time, com-

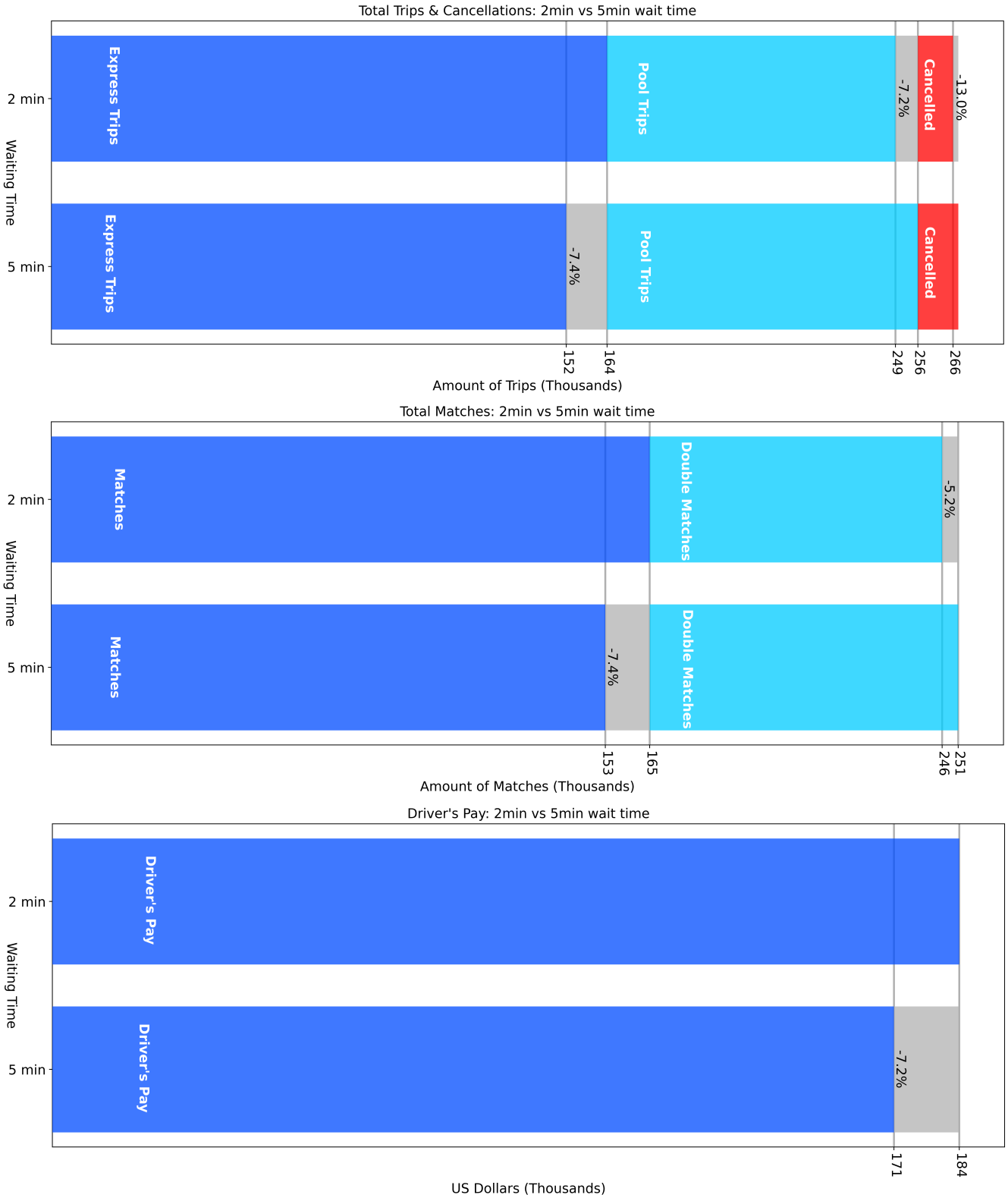


Figure 1: Completed trips, proportion of matched trips and driver's pay with 2 and 5 min of waiting.

pared to the amount of Express trips done with 2 minutes; the number of Pool trips decreased by 7.4% when the waiting time changes from 5 to 2 minutes; and finally, the amount of cancelled trips decreased 13% when the waiting time changes from 5 to 2 minutes, or increased 13% when the waiting time changes

from 2 to 5 minutes. Overall, even though the percentage lost in Express trips (7.4%) is quite similar to the percentage gained of Pool trips (7.2%) when the waiting time is 5 minutes, the plot shows that the difference in numbers of trips is 12 thousand against 6 thousand trips, respectively. Pool trips are more expensive to the user, sometimes even **by half** [1], which is an exact recovery in this case, since the Pool trips increased by half the amount of the decrease of Express trips, and the cost is twice the Express trips, so the loss is recovered with Pool, economically speaking. What Uber lost with Express, is recovered by Pool.

The middle plot of Figure 1 presents the effects of the increase in waiting time as follows: Simple matches decreased by 7.4% when the waiting time changed from 2 to 5 minutes, but the amount of double matches (that means the app matched a rider with two or more people for the same Express ride) increased by 5.2% when the waiting time was 5 minutes, since the algorithm was able to make more efficient matches with more computation time.

The last effect shown on Figure 1 is the one with the driver's earnings, or *Uber's costs*, and it shows again similar numbers: when the waiting time increases to 5 minutes, the driver's earnings drop by 7.2%, but that also means that Uber's costs, since the drivers are their contractors, drop as well by 7.2%, which is a positive effect added on top of the recovery that Pool's services offer.

## 3 Recommendations: is 5 minutes of waiting advisable?

Given the previous analysis, we can grab a few conclusions, but **should Uber's data science team increase the waiting time of Uber Express?, or when is this increment applicable?** For such purposes, the amount of Express, Pool and Cancelled trips was counted *per hour of the day* during Uber's experiments and plotted as a time series in Figure 2. In addition, Cancelled trips during the day is plotted in the right side of the figure, since in the left side plot the numbers for Cancelled are so small in comparison that the plot looks almost horizontal.

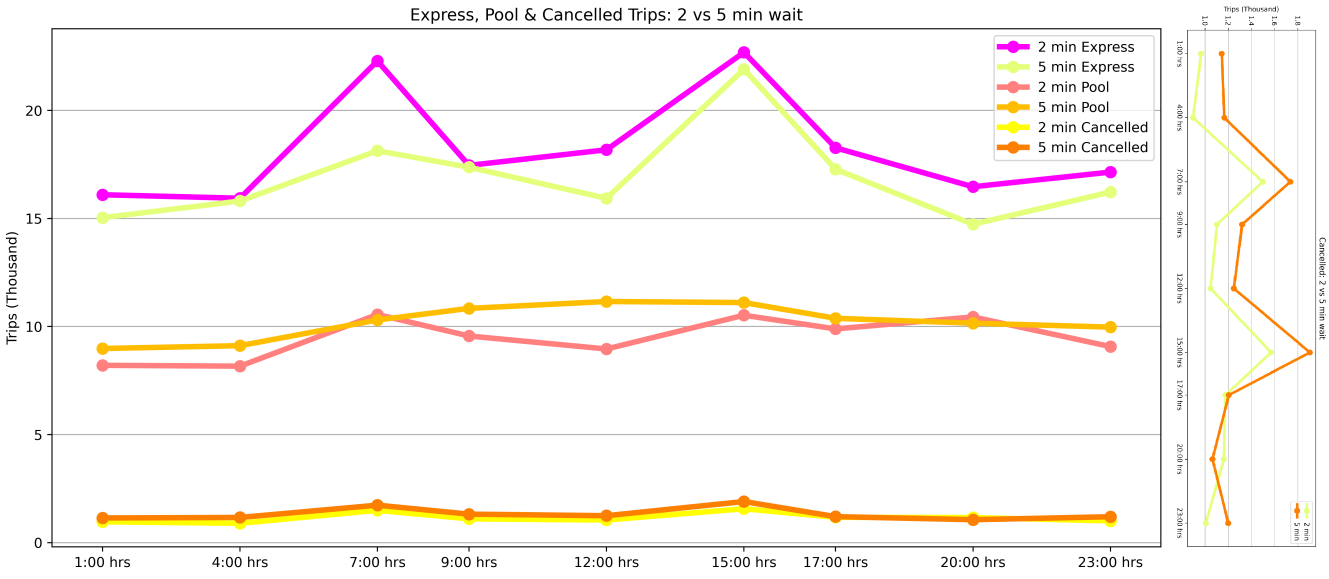


Figure 2: Express, Pool and Cancelled trips during the hours of a day.

From Figure 2, the data shows that Express trips are much more requested than Pool trips, and thus when the waiting time increases to 5 minutes, the decrease in Express trips is quite notorious, but more

prominent in peak hours such as the morning hours from 7 am until around noon. Interestingly, an increase in Pool trips happens at these exact hours as well, confirming that Pool is a very strong substitute for Express. But in more detail, the numbers that Pool trips reach after this increase are still very low compared to the level of numbers that Express trips have. Pool increases when Express decreases during 5 minutes of waiting, but can Uber profit from this? Let's look at Figure 3.

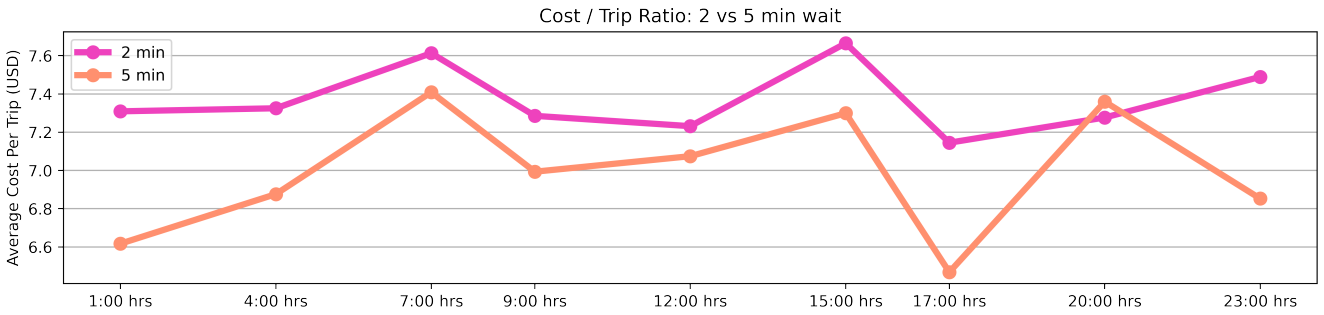


Figure 3: Average Cost per trip during the hours of a day.

In Figure 3, a time series of the cost per trip for Uber during the day is shown. The trend is noticeable: when **the waiting time increases to 5 minutes, the cost per trip for Uber is almost always lower than when the waiting time is 2 minutes**. This cost decrease during almost all day represents quite an opportunity for Uber, since Express proves to be cheaper when the algorithm is given 5 minutes to match the best possible riders, resulting in more matched users and thus less resources are used and Uber's expenses decrease. The fact that the costs are lower during almost all day is the base for our main advice: since the costs are almost always lower when waiting 5 minutes for a match in Express, the increase of waiting time to 5 minutes is suggested and even convenient for Uber **during the hours from 3 pm until 4 am** where the numbers for Express trips with 5 minutes are fairly close to those with 2 minutes; and **from 5 am to 2 pm, the waiting time can be left at 2 minutes**, given that peak hours marked a decrease in Express trips and a Pool increase that shows a smaller growth ratio than the enormous decrease ratio in Express trips, thus suggesting that a waiting time of 5 minutes might not be convenient for Uber at peak hours. Therefore, what could benefit Uber would be a **dynamic waiting time rather than a fixed waiting time**: 2 minutes from 5 am to 2 pm, and 5 minutes from 3 pm until around 4 am. The adoption of a 5 minute waiting time during these hours benefits Uber since the costs per trip are constantly lower for the company when Express is chosen.

## 4 Future Improvements

According to the graphs and information presented along this document, and the project's information on the paper [1], we thought that it was very accurate from Uber to gather the comparative information from both POOL and Express rides and from the 2 and 5 minutes, however, in order to make the economic comparison easier and more feasible, it would be better to get the separate information of both services and, also, the gross revenue obtained from the trips, to be able to compare the real profit obtained from each one at the given time. Another remarkable point to mention, is the 5 weeks freeze time that the data science team considered, to avoid any "contamination" from the market changes and letting the market to reach equilibrium the first week. In conclusion, we consider that more specific information about each one of the services, would lead to a more accurate comparison and will allow better recommendations.

## References

[1] Ch. Farronato, A. MacCormack, and S. Mehta. “Innovation at Uber: The Launch of Express POOL”. In: *Harvard Business School* 619.003 (2020), pp. 1–21.