Uber Project

Team 13

March 25, 2017

1 Topic Question- Introduction

The emergence of Uber and other non-traditional for hire vehicle companies into New York City (NYC) has been rapid. First launched in 2011, by the end of 2016 Uber reported providing an average of approximately 250,000 rides per day in NYC[1] . This expansion has had drastic effects on traditional taxis and on other drivers. In this project, we use data on Uber and traditional taxi's to identify how the emergence of Uber has affected travel time in NYC, fares charged by traditional taxis, and number of traditional taxi fares.

The increase in competition, along with the very low fares have had a large impact on traditional taxi drivers. In some instances, traditional taxi drivers have identified Uber as a new opportunity and decided to drive for the new ride sharing app. Others who stuck with the traditional methods, have had to lower fares or change behaviors in order to compete with the new challenger.

The impact of Uber is not limited to taxi cabs. The large increase in Uber drivers may increase overall road congestion. This increase would manifest itself in an increase to average trip time.

We aim to find the impact of Uber to explore how it impacts the city. With this knowledge in hand, the public will be able to better understand how to structure legislation involving Uber and other non-tradition for hire vehicle companies.

2 Topic Question- Data

2.1 Data Sets

We use information on traditional green boro taxis. For each of these groups we utilize information on pickup time, total trip time, pickup location, drop off location, trip distance, and total fare.

Location is determined by mapping the pick up and drop off coordinates to NYC zipcodes.

[1]See: New York Times Winnie Hu "Yellow Cab, Long a Fixture of City Life, Is for Many a Thing of the Past"

Data on Uber pickups by date time is also available. This data is used to determine the frequency of uber usage by time.

We limit our analysis to the second quarter of 2014 and the borough Queens.

2.2 Aggregation

Data on the traditional taxi drivers are aggregated together based on the day of pickup, the drop off zipcode, and the pickup zipcode. For each unit of aggregation we calculate the number of fares, the average trip time, average trip distance, and average fare.

2.3 Outcome Variables

For each day, drop off, and pickup location we consider the impact of increased Uber activity on average trip time, average fare amount, and number of traditional cab fares.

Average trip time is measured in seconds using traditional taxi pick up and drop off times. Average fare amount is normalized for trip distance. Specifically, we take the average of amount of dollars spent on the ride divided by the distance traveled in miles. Number of traditional fares is the count of traditional fares occurring each day for each pickup and drop off pair.

3 Non-Technical Summary

We find that for every uber ride, the average trip time decreases by .85 seconds (p=.249) and the fares charged by traditional taxis are not affected.

This shows that uber is a substitute for traditional taxis and greatly reduced the profit potential available to taxis.

4 Technical Summary

A regression to estimate the effect of Uber's emergence on the average trip time, average fare amount, and number of traditional cab fares takes the following form

$$Y_{h,p,d} = \alpha + \beta_1 U_{h,p} + \beta_2 U_{h,d} + \delta_h + \delta_p + \delta_d + \epsilon_{h,p,d} \tag{1}$$

Where h is corresponds to the day of pickup, p corresponds to the pickup location, and d the drop off location.

The fixed effect δ_h controls for time level variation that may affect all trips. This variation may be due to traffic cyclicity, inclement weather, or holidays. Fixed effects for pick up and drop off location δ_d and δ_p account for all time invariant location specific factors that may affect outcomes. For example, this variable would control for variation due to infrastructure or population density.

The regression estimates how increase in Uber activity in either the pickup location ($U_{h,p}$) or the drop off location ($U_{h,d}$) separately affects each of the outcome variables $Y_{h,p,d}$. The coefficient of interest are β_1 and β_2. These coefficients reveal how increased Uber activity has affected outcomes.