Reducing Paid Vehicle Service Congestion at LaGuardia Airport

DeWein Pelle[1], Zhenyi Chen[1], Xuan Di[1], Zhaobin Mo[1]
[1]Columbia University, New York, NY 10027, USA
(Dated: May 15, 2019)

ABSTRACT

Transportation in the United States has changed progressively over the past decade. Persons who have the luxury of ordering private vehicles have opted against the style of using long-established car services like taxis, limos, and town cars. At the start of the 2010s, people have been flocking to for hired vehicle (FHV) as their preferred mode of transportation. The competition between car services and FHV's can be translated as a fueled rivalry between NYC taxi drivers and FHV operators. At hubs like New York City's LaGuardia Airport (LGA), there is a significant issue of whether or not congestion at the port is affected by the increased volume of FHV's over recent years. Several laws and regulations have been put forth to dampen the growth of FHV companies to allow a fair and non-monopolized paid-private-vehicle transit system. Unfortunately, regulations on FHV services have hindered the FHV industry's operational growth in the city due to the belief that FHV cars are suffocating the taxicab industry. This study investigates the volume of travel from LGA to popularly traveled destinations in NYC. Specifically, we will be looking at all travel between LGA and the Columbia University community –a strong group of travelers to and from the port. To intensify the results, we added a traveler's attribute to each recorded ride between the two locations. This was done by affiliating every requested ride with an incoming flight at LGA. The study utilizes linear regression and Google API to look at the monthly behavior of FHV and Taxicab rides taken during several months in the year 2017. The study aims to provide a proposal that can ultimately reduce the flow of traffic and congestion rate entering and exiting the LaGuardia Airport area.

Keywords: for hire vehicles, congestion metrics, google maps api

Introduction

The advent of smartphones along with apps that provide a real-time supply and demand chain has ballooned in the transportation industry. Taxis were once the customary mode of travel through urban regions, and especially to and from hubs such as airports. However, for-hire vehicle (FHV) services such as Lyft, Uber, and Via-NYC have become staple modes of transportations. Subsequently, the value of taxi services has staggered. According to John Tamny, "a medallion which once fetched up to \$1.3 million has plummeted to \$175,000, and sometimes less[1]." These medallions, which are permits that allow parties to operate a taxi cab, come at a hefty price to middle-class citizens, unlike FHV.

In continuation, the issue that initially overshadowed the race between hail-rides versus taxicabs is the fact that there were few, if not any, requirements for a person to operate an FHV in New York City. Recently, legislation was put in place to put a cap on the total amount of FHV operating in the city at any given point. This effort was made to seize the upending of the taxicab industry by companies like Uber, and to figure out "how to regulate the company[2]," according to

Fitzsimmons. Outside of the perceived economic warfare, the warpath set out against FHV's is aided by the belief that these parties saturate the city's transportation systems. However, Josh Gold of Uber states that] the legislation limiting the number of hail-rides allowed in NYC "will threaten one of the few reliable transportation options while doing nothing to fix the subways or ease congestion[3]."

Using the NYC Taxi and Limo Commission (TLC) trip and record data, an investigation on how Uber has impacted NYC traffic may provide a better idea on how the rideshare has immersed itself into the already boisterous traffic dynamics of the area. If we telescope the study into one specific area, travel associated with local airports, we can investigate what impact ridesharing has contributed to the congestion of NYC traffic in these areas. Supplementary data could include the frequency of flight arrivals and departures associated with the requests for FHV and taxicab rides. Gaining a better understanding of what has caused congestions at hubs the major NYC airports may be a great indicator of how the city traffic behaves.

At this time, TLC has available data on FHV and taxicabs. Unfortunately, for flight-related API, the datasets are more privatized. However, the flight data obtained for this study reflect monthly data collected on the on-time arrival of commercial flights in the United States. Another way to track the interaction between the airline and vehicular infrastructures is to look at the normal travel rhythm of flights, as companies tend to run daily departures on a similar schedule. The issues with implement this project will likely come into play with the lack of flight data, a probable result of airline-related terrorism. However, a study like this can analyze the roles of the FHV and taxis can play in the stabilization of roadway travel and the corresponding market for it. As a result, faster road travel, along with decongestion could assist in ensuring a diverse travel network that meets the needs of the socially intricate city that is New York City.

The goal of this study is to produce a proposal fit to reduce traffic imprint in the New York City area. Congestion at ports such as the LaGuardia airport can indirectly contribute to hazards. These include, but are not limited to, more socially influenced hazards such as traffic accidents, impact on carbon footprint and gas emissions from vehicles, and even a less enjoyable travel experience for tourists. Nonetheless, a more pleasant travel network in the U.S. can make the city competitive in the transportation market as Canada. Currently Canada has the most diverse travel network in all of North America. Putting the scope of this project into respect of the United States' need to improve public infrastructure, ridership level increases can be promoted at many of the nation's strongest port to reduce traffic congestion. As a result of this, travel networks will have a reduced vehicular footprint and could potentially encourage stronger roadways.

Methods

The first step of the project was to scrape the raw data for the information we wanted to use for this study. Relevant columns in the flight data file included departure and destination airports, date and time of arrival, the amount of time the aircraft spent taxiing (arriving at the gate), as well as information on flight cancellations and diversions. We then narrowed down this data by filtering for flights that arrived at LGA during the three target months of July, November, and December. We chose these three months because we believed they would have ample data as they are all months with holidays associated; thus producing a significant amount of travel footprint. Both December and November have major holidays, whereas July is a summer month, suggesting a large volume of summer vacation travel. As for the vehicular data, we narrowed the file down to columns containing the location of pickup, date and time of pickup, and date and time of drop-off. Likewise, we narrowed down this file further to only contain trips originating from LGA (zone 138) then traveling to the vicinity of Columbia University (zones 24 and 166) within the three target months.

Our next step was to correlate the flight data and vehicle trip data. Due to the fact that the flight datasets did not indicate passenger counts, it was not plausible to associate the taxi/FHV trips to the flights as a 1:1 ratio. Resultantly, it was more challenging to estimate the number of people on each flight that were likely requesting a ride, in addition to not knowing the exact movements of each passenger on the plane. Resultantly, we dealt with this hurdle by grouping flights and rides in 15-minute intervals. We decided on a buffer of 15 minutes after using LGA's website, which indicated an estimated wait time in Terminal B for passengers that need to pick up their luggage from baggage claim. We based the correlation between flight arrivals and taxi/FHV pickups on the following standard: all flights within a 15 minute interval (0:00 to 0:14, 0:15 to 0:29, 0:30 to 0:44, and 0:45 to 0:59 minutes) are to be associated to all taxi/FHV pickups in the next 15 minute interval.

Table 1 is a screenshot of the data frame of merged files. On the left side is the flight data, and on the right side is the vehicular data. The two sides are linked by a common column "UNIT," which is a count of 15-minute intervals starting from the first day of the month up until the last 15-minute period of the month. We employed this variable to account for the time shift in our comparison, as the vehicular data had to count as one 15 minute interval after the flight data. Looking at the two "GROUP" columns in the merged data frame screenshot, the problem that prompted us to use the "UNIT" variable is apparent. The "GROUP" values originally span from 0 to 3 in one hour, (group 0 meaning the interval from 0:00 to 0:14), but it was impossible to account for the 15-minute shift when the flight arrived at the last period of the hour. The "UNIT" variable takes this problem into account, assigning a unique value for each 15-minute interval of the month.

	TAG_x	day_x	hour_x	GROUP_x	pass_tag_x	UNIT	TAG_y	day_y	hour_y	min	GROUP_y	pass_tag_y	DOLocationID	bus
0	2017-12-01 19:42:00	1	19	3	FLY	80	2017-12-01 19:54:11	1	19	54	4	TXY	166.0	1
1	2017-12-01 19:42:00	1	19	3	FLY	80	2017-12-01 19:49:03	1	19	49	4	TXY	166.0	1
2	2017-12-01 19:45:00	1	19	3	FLY	80	2017-12-01 19:54:11	1	19	54	4	TXY	166.0	1
3	2017-12-01 19:45:00	1	19	3	FLY	80	2017-12-01 19:49:03	1	19	49	4	TXY	166.0	1
4	2017-12-01 19:36:00	1	19	3	FLY	80	2017-12-01 19:54:11	1	19	54	4	TXY	166.0	1
5	2017-12-01 19:36:00	1	19	3	FLY	80	2017-12-01 19:49:03	1	19	49	4	TXY	166.0	1
6	2017-12-01 19:41:00	1	19	3	FLY	80	2017-12-01 19:54:11	1	19	54	4	TXY	166.0	1
7	2017-12-01 19:41:00	1	19	3	FLY	80	2017-12-01 19:49:03	1	19	49	4	TXY	166.0	1
8	2017-12-01 19:36:00	1	19	3	FLY	80	2017-12-01 19:54:11	1	19	54	4	TXY	166.0	1
9	2017-12-01 19:36:00	1	19	3	FLY	80	2017-12-01 19:49:03	1	19	49	4	TXY	166.0	1

Table 1. The table above shows the merging of data frames based on the association of on the ground vehicles and inbound flights at LaGuardia Airport.

After associating each taxi/FHV ride with a flight, we also took into consideration an alternative to taking a private vehicle back to campus, the M60 SBS. Moreover, this was our validation policy. We hypothesized that if we introduced a highly accessible travel option between the access points, then the congestion rates would be reduced. To accomplish this validation, we downloaded the weekday schedule for the M60 SBS and tagged each ride using the following criteria: if there was a bus pickup scheduled at LGA 5 minutes before or after a ride was taken, that passenger/group of passengers could have elected to take the bus back to campus instead of a private vehicle. As seen in Table 1, such rides were associated with a "1" value instead of the default "0" value.

To express our validation, we first plotted the behavior of vehicular footprint on the path per day. After realizing our initial results, we wanted to show the impact that taking the bus would have had on the congestion on the route from LGA to Columbia University. We decided to show this first using a frequency plot from python library matplotlib, which details the number of rides taken on the route by each party (taxis, FHV, and the M60 SBS). We then plotted the number of trips taken on the path by the taxis/FHV. Following this step, we integrated the bus schedule as an "encouragement" for individuals to take the bus upon arrival as opposed to ordering a car service. Using this algorithm, we provided a comparison between the amount of traffic on the route before and after we diverted vehicular congestion to the buses.

In continuation, we wanted to display the difference in congestion by using a Google Maps timelapse. Using the python library gmplot to access the Google Maps API, we plotted the daily congestion on the route from LGA to Columbia University by estimating the frequency of trips as the thickness of the access points' pathway. We showed a variance between the amount of congestion each day by changing the width of the route on the gmplot. Table 2 shows the process behind deciding on the width of the path. The "Days" column indicates which day of the month the row is referring to, while the "Taxi" and "FHV" columns are a count of the number of rides on the route each day. The columns "w_t" and "w_f" expressed the width of the line plot on the gmplot. We calculated these values by dividing the respective count values by 20. We chose the modifier of 20 because it provided the most appropriate overall visual width across all of our data.

	Days	Taxi	FHV	w_t	w_f	s_lat	s_lon	d_lat	d_lon
0	1	8.0	10.0	0.40	0.50	40.774391	-73.871469	40.807791	-73.963762
1	2	5.0	5.0	0.25	0.25	40.774391	-73.871469	40.807791	-73.963762
2	3	1.0	10.0	0.05	0.50	40.774391	-73.871469	40.807791	-73.963762
3	4	0.0	0.0	0.00	0.00	40.774391	-73.871469	40.807791	-73.963762
4	5	5.0	8.0	0.25	0.40	40.774391	-73.871469	40.807791	-73.963762
5	6	12.0	5.0	0.60	0.25	40.774391	-73.871469	40.807791	-73.963762
6	7	3.0	11.0	0.15	0.55	40.774391	-73.871469	40.807791	-73.963762
7	8	8.0	0.0	0.40	0.00	40.774391	-73.871469	40.807791	-73.963762
8	9	1.0	4.0	0.05	0.20	40.774391	-73.871469	40.807791	-73.963762
9	10	16.0	9.0	0.80	0.45	40.774391	-73.871469	40.807791	-73.963762

Table 2. The table above shows the calculated width of the lines used by FHV's and Taxis during their path between Columbia University and LaGuardia Airport after accounting for the diversion of traffic to the M60 SBS.

After generating our google map HTML files, we used the python packages selenium and virtual machine "chromedriver" to produce image files of the maps. We then merged our files together into a gif timelapse. We performed a timelapse for the congestion created by the two private vehicle car services, as well as a separate comparative timelapse for the effect of our validation policy on the vehicular footprint on the pathway. Resultantly, we returned a total of 4 timelapse gifs for each month, and 12 for the three target months. The timelapse gifs allow for a more direct comparison of the congestion with regards to each day.

Results

After evaluating the vehicular association, we produced frequency plots for July, November, and December 2017. During July, our graphs expressed a more aggressive usage of Taxis by Columbia University affiliates or tourists. The usage of this vehicle service increased over time, only taking small dips for a brief time. The overall frequency of the Taxi ridership seemed to have actively beat FHV ridership. As shared in our methods, the months used in the analysis were selected based on significant dates observed in the United States. For July, that includes the celebration of Independence day on the fourth of the month. On this day, we recorded an estimated 250 taxis traveled, while on 150 FHV's operated between the two access points. Interestingly enough, the highest peak for taxicabs occurred around July 30th, with roughly 350 taxis traveling in our study zone. Surprisingly, the lowest travel rate for FHV's during this period. Additionally, the FHV presence was not as high compared to Taxis for this month. The only time there was a higher frequency for FHV's was coincidentally when there was both a dip in ridership for both car services. With the M60 SBS ride encouragement, both car services plummeted in ridership. This could strongly indicate that stronger encouragements can reduce vehicular footprint; however, there are more implications associated with this argument.

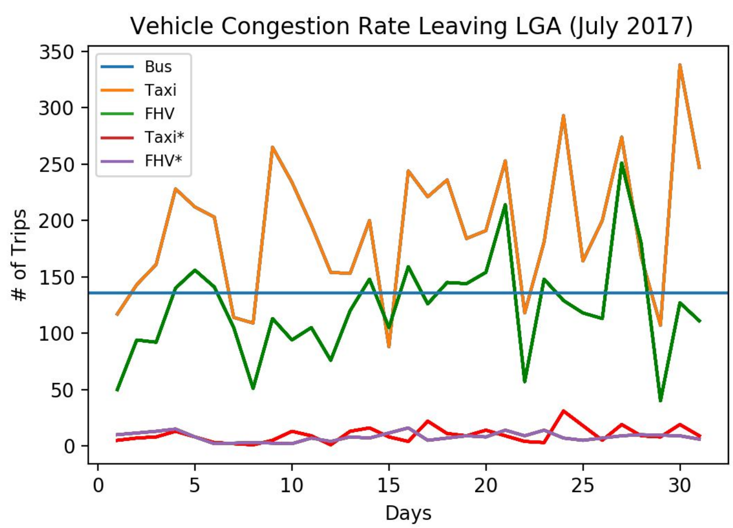

Figure 1. This figure expresses the daily volume of cars that are traveling between LaGuardia Airport and Columbia University during July 2017. The lines that are labeled with an (*) in the legend indicate the volume of vehicles on the path after travel is encouraged via the M60 SBS.

For December 2017, the competition between taxicabs and for-hire vehicles between our access points was the most robust that we've seen.

Early on, Taxi ridership was significantly comparative to FHV congestion on our travel pathway. At different points in the month, FHV usage directly competed with Taxicab usage. The highest peak increased by roughly 342.86% from the previous month of study. Interesting, taxicab did not hold the highest peak for this month. FHV usage grew larger than taxicab ridership for the latter part of November. Ridership noticeably surged around the Thanksgiving holiday for both services. This can have implications that include but are not limited to the fact that many people travel to the campus for the holidays. So, looking at the vehicular imprint in the Columbia University to LaGuardia AIrport direction would prove to be meaningful, and could validate this hypothesis. Nonetheless, the M60 SBS ride encouragement also seemed to have significantly diminished vehicular congestion. This is so far a great indicator that frequent bus employment can be an effective solution to reduce the vehicular footprint between Columbia University and LaGuardia Airport. However, the dampening of the data based on this metric was not as low as the dampening seen with the July data. This has a lot more to do with the general volume of ride requests that occurred in November 2017, as compared to July 2017.

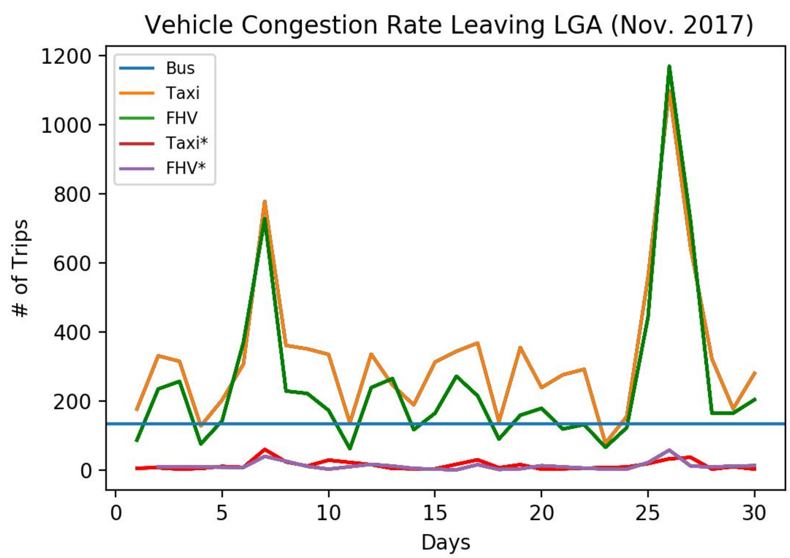

Figure 2. This figure expresses the daily volume of cars that are traveling between LaGuardia Airport and Columbia University during November 2017. The lines that are labeled with an (*) in the legend indicate the volume of vehicles on the path after travel is encouraged via the M60 SBS.

For December 2017, there was less robust competition between the car services. Early on, Taxi ridership overshadowed FHV congestion on our travel pathway. At different points in the month, FHV usage grew larger than Taxicab usage. Ultimately, taxicab ridership grew larger than FHV for the latter part of the year. Ridership noticeably plummeted for both services around the holiday. This can have implications that include but are not limited to the fact that most people travel away from the campus rather than towards it for the holidays such as Christmas. So, looking at the vehicular imprint in the Columbia University to LaGuardia AIrport direction would prove to be meaningful, and could validate this hypothesis. Nonetheless, the M60 SBS ride encouragement also seemed to have significantly diminished vehicular congestion. This is so far a great indicator that frequent bus employment can be an effective solution to reduce the vehicular footprint between Columbia University and LaGuardia Airport.

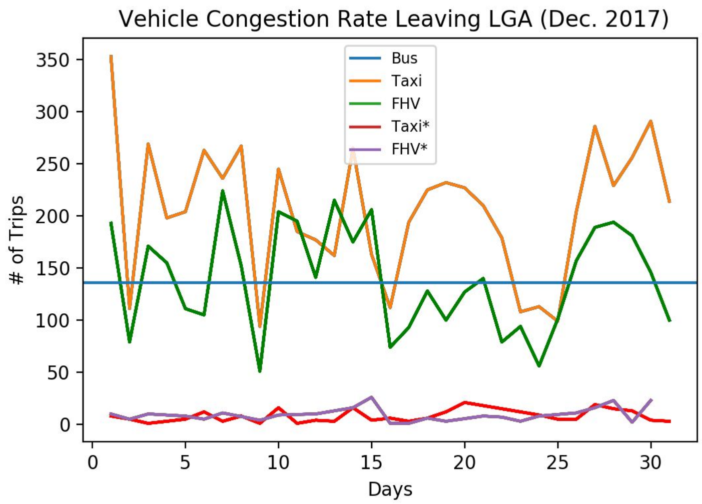

Figure 3. This figure expresses the daily volume of cars that are traveling between LaGuardia Airport and Columbia University during December 2017. The lines that are labeled with an (*) in the legend indicate the volume of vehicles on the path after travel is encouraged via the M60 SBS.

Now we will look to analyze the Google maps that we created in our methods.

Figure 4. The image above is a Google Maps plot of our data for July 4, 2017. The path shown is the optimal driving directions for vehicles between Columbia University and LGA. The route is also the same route assigned to the MTA M60 SBS. L1 expresses the congestion volume of FHV's on the pathway and R1 for Taxis. L2 and R2 describe the congestion volume after bus transportation competes with FHV's and Taxi's, respectively.

As shown in Figure 4, there is not a dramatic congestion rate, or rather line thicknesses of the routes taken by both car services. L represents for-hire vehicle services, while R represents taxicabs services. L1 and R1 are representative of the uninterrupted competition between the respective car services, while L2 and R2 represent the respective car services after M60 SBS bus encouragement is integrated into the evaluation process. L1 and L2, as well as between R1 and R2. In both L and R images, the thickness of both plots go from a solid line to a barely visible line, displaying how much congestion the M60 is theoretically able to take. Additionally, when we compare Figure 4 to Figure 5, which is the plot for November 7, 2017, we immediately notice a large jump in traffic volume for L1 and R1, but the corresponding change in traffic volume for L2 and R2 is not much different from the plots on July. The large discrepancy in volume in L1 and R1 can be attributed to July 4th being a national holiday as well as the hotter weather encouraging the population not to go outdoors as much. However, the similar thickness in both L2 and R2 plots for the days from both months hints at the sheer amount of congestion the bus system is able to take away from the road.

Figure 5. The image above is a Google Maps plot of our data for November 7, 2017. The path shown is the optimal driving directions for vehicles between Columbia University and LGA. The route is also the same route assigned to the MTA M60 SBS. L1 expresses the congestion volume of FHV's on the pathway and R1 for Taxis. L2 and R2 describe the congestion volume after bus transportation competes with FHV's and Taxi's, respectively.

Discussion

Looking at the results generated from our study, it is apparent that the policy validation allows us to absorb large amounts of traffic and congestion on the pathway. There are advantages to taking a taxi or an FHV such as privacy, comfort, and timeliness; however, if faculty and students traveling from LGA to Columbia University made an effort to take the M60 SBS more often, it would make an apparent impact on the roads between the two locations. Ridesharing and carpooling are already incentivized by factors such as carpool lanes and reduced fares, and the bus system can be perceived as an already existing, cheap, consistent, and large scale ride sharing mechanism.

In continuation, we made assumptions in this study, and there is ample space for an improved version. Our results would be more accurate if we had passenger data for the plane flights, allowing us to possibly trace taxi/FHV customers from the plane to the vehicle. Considering a traveler count would allow us to make observations regarding the proportion of people taking specific methods of transportation out of the airport as well. Additionally, we restricted the route between LGA and Columbia University, but it would be more insightful if we expanded the study to the broader NYC area. Originally, we planned to observe congestion patterns going from LGA to all of the taxi zones in NYC, but to accomplish this, we would need real-time data on the multitude of routes taken by the car services. Also, we diverted all the traffic to the M60 SBS without considering the capacity of the bus. If we take into account the capacity of each bus, the current bus schedule may not be able to absorb congestion at the rate discussed.

We believe that this study provides significant insight into the role that paid vehicle services play in the congestion of traffic outside of major airport roadways. All of our considerations are put into play but do not account for private and commercial vehicles that share this pathway at any point to access either Columbia University, LaGuardia Airport or the broader NYC and NY state areas. We strongly believe that implementing our policy to promote bus rides is a great metric to approach reducing the amount of vehicular footprint made on our pathway of study. We can broaden this study to look at the envelope of other travel issues that occur along the path, and even the demographics of people traveling from the origination point to the destination zone. Implementing a policy like this can be difficult; however, using monetary and environmental compensations to promote the campaign can be a forceful approach. Resultantly, we would hope that the campaign would spur other studies across the United States, and even provide a study for the MTA on why they should increase the number of buses in their M60 SBS fleet. There is, of course, other implications that can hinder this outcome, but based on our results, we believe that it is a robust solution to the congest issues occurring at the port. One last interesting aspect was that taxicab seemed to rain supreme over FHV in this neighborhood; so, taking a look into the demographic correlation between ridership from LaGuardia to the Columbia University would be a worthwhile case to telescope on in the future.

Appendix

November 23, 2017

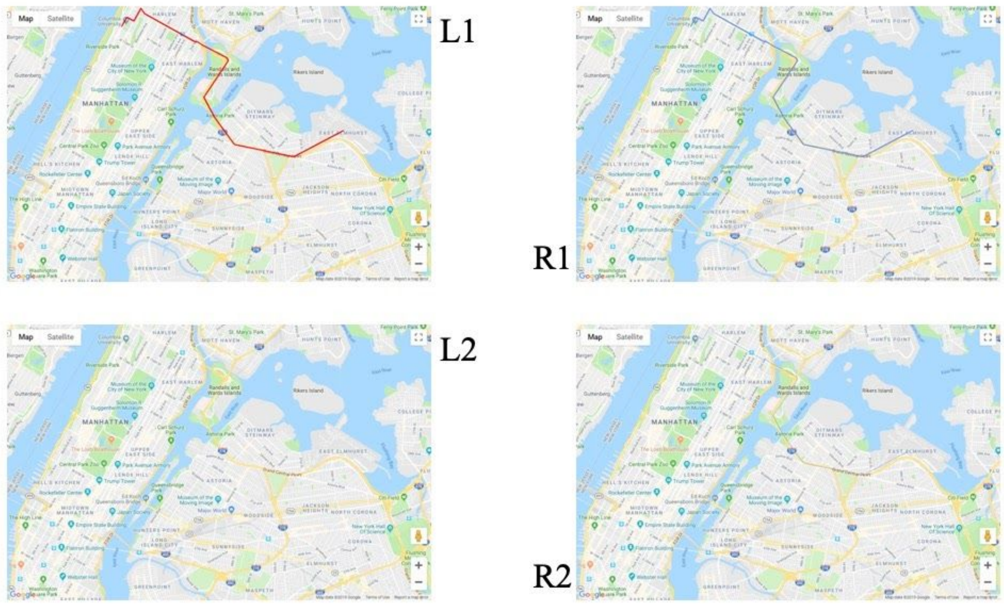

Figure A1. The image above is a Google Maps plot of our data for November 23, 2017. The path shown is the optimal driving directions for vehicles between Columbia University and LGA. The route is also the same route assigned to the MTA M60 SBS. L1 expresses the congestion volume of FHV's on the pathway and R1 for Taxis. L2 and R2 describe the congestion volume after bus transportation competes with FHV's and Taxi's, respectively.

December 25, 2017

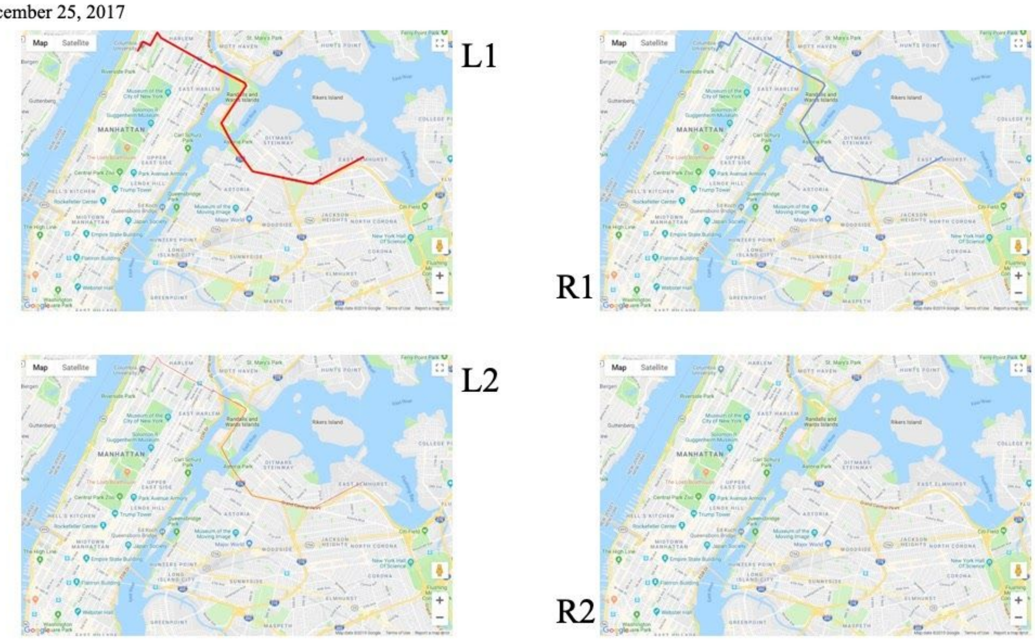

Figure A2. The image above is a Google Maps plot of our data for December 25, 2017. The path shown is the optimal driving directions for vehicles between Columbia University and LGA. The route is also the same route assigned to the MTA M60 SBS. L1 expresses the congestion volume of FHV's on the pathway and R1 for Taxis. L2 and R2 describe the congestion volume after bus transportation competes with FHV's and Taxi's, respectively.

December 31, 2017

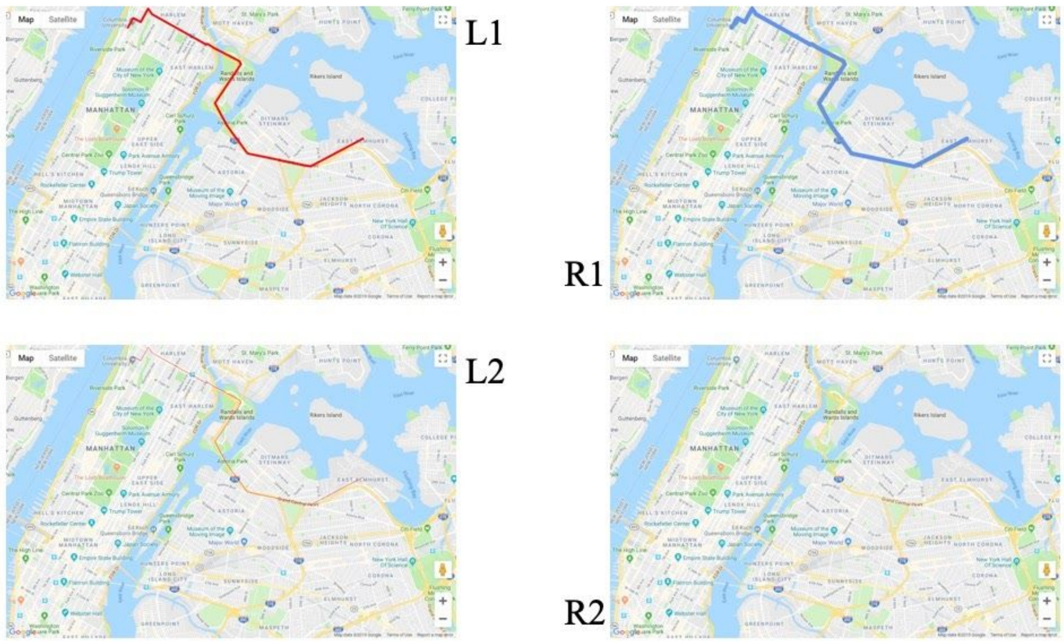

Figure A3. The image above is a Google Maps plot of our data for December 31, 2017. The path shown is the optimal driving directions for vehicles between Columbia University and LGA. The route is also the same route assigned to the MTA M60 SBS. L1 expresses the congestion volume of FHV's on the pathway and R1 for Taxis. L2 and R2 describe the congestion volume after bus transportation competes with FHV's and Taxi's, respectively.

References

[1]. Tamny, John. "Uber and Free Markets Are Not to Blame for Taxi-Driver Suicides In New York." *Forbes*, Forbes Magazine, 2 June 2018, www.forbes.com/sites/johntamny/2018/06/02/uber-and-free-markets-are-not-to-blame-for-taxi-driver-suicides-in-new-york/#1bdb3dc7a2fe.

[2]. Fitzsimmons, Emma G. "Uber Hit with Cap as New York City Takes Lead in Crackdown." *The New York Times*, The New York Times, 8 Aug. 2018, www.nytimes.com/2018/08/08/nyregion/uber-vote-city-council-cap.html.

[3]. Katz, Miranda. "Why Are New York Taxi Drivers Killing Themselves?" *Wired*, Conde Nast, 29 Mar. 2018, www.wired.com/story/why-are-new-york-taxi-drivers-committing-suicide/.

[4]. Stromberg, Joseph. "The Real Reason American Public Transportation Is Such a Disaster." *Vox*, Vox, 10 Aug. 2015, www.vox.com/2015/8/10/9118199/public-transportation-subway-buses.