2019

Would Uber Help to Reduce Traffic Congestion?

Qianxi Zheng
Skidmore College, qzheng2@skidmore.edu

Would Uber Help to Reduce Traffic Congestion?

Qianxi Zheng

Skidmore College

This thesis is submitted in partial fulfillment of the requirements for the course Senior Seminar (EC 375), during the Spring Semester of 2019

While writing this thesis, I have not witnessed any wrongdoing, nor have I personally violated any conditions of the Skidmore College Honor Code.

Name: Qianxi Zheng

Signature: Qianxi Zheng

Abstract:

This research explores the effects of Uber entry on New York City's traffic. The two major questions I am trying to answer that might be of vital importance to transportation authorities are 1) does Uber substitute public transits? 2) does an introduce of Uber slow down average travel speed? After Uber was first introduced in year 2009, there are continuous debates on distinguishing its impact on traffic (Rayle et al., 2014; Li et al., 2016; Schaller, 2018; Castiglione et al., 2018). Considering that Uber is relatively new, relevant traffic data such as congestion indices are in general unavailable, which appears as a common limitation in previous analysis. In this research, I use monthly number of public transit trips in NYC to estimate a substitution effect of Uber on public transit ridership. To measure its direct impact on road traffic, I use Average Travel Speed generated from NYC yellow cab trips as a proxy for the citywide Average Travel Speed. A further application of monthly number of vehicles crossing nine major bridges and tunnels is used to capture a trend of traffic volume in NYC. The final dataset comprises 133 observations range from January 2008 to January 2018. Perceiving that Uber was introduced to NYC on May 2011 and was suspended on issuance of new vehicle licenses starting from August 2018, I use a regression discontinuity (RD) design and set the two events as cutoff points in the model. Additional use of Google Trend helps to more precisely determine the cutoff point. The regression results suggest that after Uber was introduced to NYC, 1) number of public transit trips has increased by about 3%; 2) average travel speed has decreased by .127 mph; and 3) traffic volume was not affected.

1. Introduction:

Traffic is getting worse in cities. From year 2012 to 2013, 61 of America's 100 largest metropolitan areas experienced an increase of traffic congestion. From year 2013 to 2014, the number increased to 91 cities (Schrank et al., 2015). Congestion as one of the most rooted problems has long been addressed. Enhancing public services including public transits and roads are major approaches transportation authority suggests in hopes of easing congestions (FHWA, 2017).

In recent years, there emerge private companies that try to provide alternative transportation solutions with an implementation of modern technologies. One disruptive innovation is ride-hailing service, sometimes called ride-share or Transportation Network Companies (TNCs), which aims to improve current on-call transportation sector. In general, ride-hailing service refers to the activity of asking for a car or driver to come immediately and take passengers somewhere (Ride-hailing, n.d.). But it has been more frequently referred to Uber or Lyft trips, the two most popular companies that provide ride-hailing service. By decomposing the concept of ride-hailing service, we realize that the service is fundamentally no different from traditional taxi service, but more an improved version of the latter. Some features being introduced include a combined use of the internet and Global Positioning System (GPS) to match drivers and passengers, the big data technology implemented to automatically schedule best traveling routes and the shareability feature provided to improve transportation efficiency.

Though these features sound exciting, there are different voices on distinguishing an impact of ride-hailing service on traffic. In general, boosters claim that the service might help to mitigate traffic congestion by providing travelers with a more efficient transportation mode and by reducing car use as well as car ownership (Li et al., 2016; Rayle et al., 2014; Cramer & Krueger, 2016). Opponents on the other hand criticize that the service would worsen current traffic, introduce more vehicular trips, compete public transits, mislead consumers with varying price system and cater to only specific groups of people who are well-educated and young (Schaller 2018; Castiglione et al., 2018; NYCDOT, 2018).

The conflicting arguments raise an interest to conduct further empirical research by applying up-to-date real-time data in order to more closely examine the effects of ride-hailing service on traffic congestion. Considering that attitude towards ride-hailing service has been shifted in recent years since the idea was first introduced by Uber in year 2009, while some cities and countries such as London and New York City have even fully or partially banned Uber, it gives us an advantage to use the regression discontinuity (RD) design to estimate both effects of Uber entry and restriction on congestion. The two primary measures I use in this research to estimate such impact include average travel speed, which helps to estimate Uber's direct impact on road traffic, and number of public transit trips that helps to measure its substitution effect on public transit ridership. The major goal of this research is to provide a more comprehensive and up-to-date analysis of Uber's impacts on traffic. Based on that, I propose relevant policy suggestions. Specifically, the two major questions I try to answer in this research that might be of vital importance to transportation policymakers are, does ride-hailing service slow down average travel speed? Second, does ride-hailing service substitute public transits?

The rest of this paper is organized as follows. Section 2 includes literature reviews on discussion of related background information including congestion, measures of congestion, traditional approaches to ease congestion and concept of sharing economy, as well as previous researches that explore the impacts of ride-hailing service on traffic; In Section 3, I elaborate details of this research by introducing data, research framework and methods being applied to check for robustness; Section 4 includes results from regression analysis and a discussion on the result; Section 5 concludes the research. Related tables and graphs are reported in Section 6.

2. Literature Reviews:

2.1 Background:

Traffic congestion has been generally defined as an excess demand for road travel while the supply of road infrastructures is insufficient to meet such demand. Rao (2012) concludes two categories of causes of congestion: micro-level and macro-level factors. Micro-level factors relate to traffic on the road, such as many people and freight need to move at the same time or delay caused by irregular but frequent incidents like car accidents and poorly designed traffic systems, or even drivers' driving behaviors. Macro-level factors relate to overall demand for traffic, such as economic change like an improved income level, change in car ownership, employment and land-use patterns. Another explanation of congestion relates to the economic theory of “the tragedy of the commons”. While roads are in general free to use, people tend to overuse this public resource up to the point when traffic collapses into a jam. As traffic volumes continue to grow faster than the designed road capacity and economic impacts of congestion are usually not negligible, congestion has widely been viewed as a growing problem in many urban areas. The “2018 Global Traffic Scorecard” reported by INRIX (Reed and Kidd, 2019) estimates that U.S. citizens lost an average of 97 hours a year due to congestion, costing them nearly $87 billion in 2018, an average of $1,348 per driver. Negative impacts of congestion include not only social-economic loss such as a waste of time, but also irreversibly environmental loss such as deteriorated air quality caused by additional emission.

Two long-lasting approaches that transportation departments apply in hopes of easing traffic congestion are: 1) to widen and build more roads, and 2) to enhance public transits system. However, continuously keep building more roads and improving infrastructures does not fundamentally fix the problem but might be fairly impractical and expensive. Duranton and Turner (2011) in an empirical research find that vehicle kilometers of travel (VKT) increases in direct proportion to the available lane-kilometers of roadways, which implies that new roads only result in additional traffic that continues to rise until congestion returns to the previous level. This phenomenon is also known as Downs-Thomason paradox (Downs, 2000). The paradox suggests that any improvements on current road systems such as an expansion of roads will only result more traffic following Downs’s Principle of Triple Convergence. When commuters realize that new roads are faster, they who had traveled before or after peak-hours to avoid congestion would shift back into the peak period. Drivers who were choosing alternative routes to avoid congestion would also shift back to using the more convenient roads. Commuters who originally took public transits might also start to drive cars for a more comfortable ride (Downs, 2004). Perceiving this fact, U.S. Department of Transportation Federal Highway Administration (FHWA) proposed that it requires three coordinated approaches including construction, preservation and operation in order to reduce congestion in the new era. Among the three approaches, FHWA views a better operation as the newest and a more sustainable approach to confronting transportation challenges in the 21st century. In recent years, FHWA takes advantages of new technologies and is working on to include congestion data analysis, development of policy and legislation and promotion of an information infrastructure in their system (FHWA, 2017).

While improving operations is essential to make vehicular traffic more efficient, public transit plays an important role in transporting commuters especially during rush hours. Aftabuzzaman et al. (2010) have developed a framework to estimate the monetary values of congestion reduction effects led by efficient public transits. They find that congestion relief impacts are valued between 4.4 and 151.4 cents (Aus$, 2008) per marginal vehicle kilometers of public transits travel, with an average of 45.0 cents. Considering that the major focus of Aftabuzzaman's research is on Australia, the same conclusion might not apply to U.S. because the two countries might have distinctive characteristics of public transit ridership. In a statistical analysis of the 74 largest urbanized areas in the U.S. from year 1982 to 2007, Rubin and Mansour (2013) find that increasing public transit utilization does not lead to a reduction in traffic congestion as measured by the Travel Time Index (TTI), nor does decreasing transit utilization lead to an increase in traffic congestion in general. Specifically, the regression analysis does not reveal any significant statistical relationship between increased annual transit unlinked trips per capita (UTP) and reduced traffic congestion, or vice versa. However, they have found that the impacts of increased annual transit trips per capita on traffic congestion vary among cities with different public transits mode share. For cities with higher public transits mode share such as NYC, where transits use accounts for 12.2% of daily vehicles miles travelled (VMT), the impact is greater than those of cities with lower transit mode share like Los Angeles where transits only account for 2.2% of daily VMT. This finding is especially meaningful if we want to explore the effects of alternative transportation modes on public transits and its further impact on congestion by directing us to focus more on areas with higher public transits share. Rubin and Mansour also find that transit fare is a significant driven factor of public transit utilization in cities such as NYC and Los Angeles, which might suggest that commuters are highly price-driven, and a less costly transportation tool, like Uber, might be more attractive.

As mentioned before, micro-factors that induce congestion include not only inefficient traffic systems, but also characteristics of traffic participants. Improving road monitoring or control systems is only one approach that requires more of governments' intervention. Economists like Anthony Downs (2004) as well as Duranton and Turner (2011) suggest that rush hour congestion is inevitable and thus shifting the focus to the demand-side. They propose the use of congestion pricing and car-pooling to obliquely encourage commuters to also participate in a congestion reduction process. Insights for congestion pricing come from the observation that people tend to make socially efficient choices when they face all the social benefits and social costs of their action. Incorporating such the price mechanism makes demand for road more elastic and stimulates commuters to self-adjust the balance between the monetary and time costs. However, roads could no longer be considered as “perfect public goods” in this situation. Further alternative solutions such as ride-hailing service introduced by private sectors are brought into view.

The three main concepts behind ride-hailing service are sharing economy, urban mobility and sustainability. Attitudes toward consumption have gradually shifted in recent years. Consumers now enjoy not only utilization provided by products but further from their inherent values, which is a trend being especially popular among young generations. This shift is facilitated by a complex interaction between social and climate changes. For example, an increasing concern on sustainability. Sharing economy, the peer-to-peer-based activities of obtaining, giving, and sharing access to goods and services, aims to cater for such demand (Investopedia, 2017). Some examples include Airbnb, which allows house owners to provide short-term house renting service with their own living apartments; Upwork, a global freelancing platform where businesses and independent professionals connect and collaborate remotely; and Uber, which allows car owners using personal cars to provide private ride service. The Google Trend shows that interest on sharing economy gains its peak in year 2015 and then falls gradually (Google Trends, 2019). Hamari (2015) claims that sharing economy “has been expected to alleviate societal problems such as hyper-consumption, pollution, and poverty by lowering the cost of economic coordination within communities”. The study also shows that “participation in sharing economy activities is motivated by factors such as sustainability, enjoyment of the activity as well as economic gains”. Many of those sharing companies are advertising for their environmentally friendly attitude while consumers announce that one major reason why they find sharing economy appealing is because how it anchors to their awareness of an ecological harmony. One by-product of this economic model is how it encourages consumers to also engage in a self-driven societal activity with the economic mechanism being similar to that of congestion pricing. The difference is that the former is driven by a sense of value while the latter is more driven by monetary incentives.

Impacts of sharing are obvious. Given that one essential idea of sharing economy is to redistribute unused or underutilized assets in order to boost efficiency, a secondary market has the potential to reduce demand for new products, waste and extra consumption. We could take Uber as an example to illustrate this idea. When hailing an Uber cab is conveniently enough, it may discourage people to buy new cars but encourage them to take either Uber trips or public transits. This would possibly reduce congestion. Furthermore, as Paul Barter (2013) has found, cars are parked 95% of the time on average, which suggests that personal cars are generally underutilized. Uber's part-time drivers may help with such a low utilization rate by driving more of their private cars in spare time. However, this might incur more traffic on roads and makes road more congested. There are also claims saying that sharing economy has the potential to further reduce carbon emission as well as other pollutions, but there lacks comprehensive research to show its overall impacts. According to U.S. Environmental Protection Agency, the biggest contributor of greenhouse emission is transportation, which accounts for nearly 28.5% of 2016 greenhouse gas emission (EPA, 2018). Exploring the effects of Uber on traffic may in addition contribute to research that estimates the environment impacts of sharing economy model.

2.2 Ride-hailing Services:

Before the launch of ride-hailing services, people mostly consider either driving personal cars or taking public transits as major approaches to commute. Neither of those choices appears to be satisfactory for following reasons. Buying or renting cars confronts car owners with potential risks of car damages and costs of insurance and repairs. Taking public transits may not always fit to one's schedule while possible transits operating delay has to be expected. Taxi could be another choice, but the cost is usually expensive while hailing a cab is not that accessible when demand is overwhelming. To provide a more convenient commute solution, Uber, for example, implements internet-based technologies like Big Data and Artificial Intelligence (AI) to shorten waiting time that passengers hail an Uber cab and the lapse of time Uber drivers pick up next passenger. This feature allows people to reserve a cab by simply tapping on their smartphones. Recently, those ride-hailing companies further introduce their own car-pooling service that allows passengers with different departure or arrival locations but similar traveling routes to share one ride-hailing cab. We observe that ride-hailing services are similar to traditional cab service but with some new technologies being implemented. To better understand ride-hailing service, we can break it down into two essential parts, the rideshare part that being similar to car-pooling and an enhanced on-demand cab service.

Ridesharing has a long history. In the late 1990s, cities such as Los Angeles (Golob and Giuliano, 1996) and Seattle (Dailey et al., 1999) have implemented ride-matching services. Impacts of these traditional ride-sharing services on traffic have been extensively studied. Salomon and Mokhtarian (1997) discussed the effectiveness of various ridesharing policies to reduce traffic congestion. In a cost-benefit analysis, Fellows and Pitfield (2000) have found that ridesharing reduces personal commute costs in half and benefits the whole economy by reducing vehicle kilometers travelled (VKT) as well as increasing average travel speeds. Combining all those findings, it suggests that car-pooling service implemented by Uber should be expected to reduce congestion. But given that most of the above-mentioned researches lack a further investigation into how car-pooling eventually affects traffic volume in the long-run while an existence of Downs-Thomason Paradox might mitigate positive effects brought by car-pooling, additional check on congestion indices or even its substitution effects of public transits would consolidate the arguments.

On the improved on-demand hailing service side specifically, Li, Hong and Zhang (2017) apply a difference in difference (DiD) model to investigate the effects of Uber entry on an area's congestion level. The dataset being used comes from Texas A&M Transportation Institute (TTI). This dataset has been widely used in transportation analysis. Some controls include fuel cost, social economic characteristics such as population size and GDP, and characteristics of road systems such as miles and number of commuters on each of the roads. To measure traffic congestion, Travel Time Index (TTI), Commuter Stress Index (CSI) and Daily Vehicle Hours of delay have been used. TTI refers to the ratio of travel time in the peak period to travel time at free-flow conditions, and CSI is calculated for only the peak direction in each peak period. Eventually, they construct an urban area-year level panel dataset consisting of 957 observations over 87 urban areas starting from year 2003 to 2014. The regression result suggests that entry of Uber has improved traffic by reducing TTI by 0.19%, CSI by 0.3% and Daily Vehicle Hours of delay by about 1.2%. Though their results are statistically significant, we observe that 57 out of those 87 urban areas in the sample have Uber being introduced after year 2014. This finding leads me to suspect the validity of data they have been used considering that the introduce of Uber to an area does not necessarily indicate that local residents start to use Uber immediately. Hence, the given data by TTI might not fully capture the effects of Uber entry in either a short or even a longer time period if the effects are lagged. As more up-to-date data are released after Li et al.'s research, we could take this advantage and conduct further analysis by applying those updated data.

Cramer and Krueger (2016) use capacity utilization rate to evaluate Uber's effects on congestion. Capacity utilization rate is usually measured by either the fraction of time or the fraction of miles that drivers have a fare-paying passenger in the cab. A higher capacity utilization rate suggests that the mode is more efficient. Interested in major metropolitan areas, they investigate traffic in Boston, Los Angeles, New York, San Francisco and Seattle. By also applying a DiD model, they have found that on average, the capacity utilization rate is 30% higher for Uber drivers than taxi drivers when measured by time, and 50% higher when measured by miles. In another research, Rayle et al. (2016) apply intercept surveys to explore the use of Uber in San Francisco. After analyzing 380 responses, they find that vehicle ownership does not change much after ride-hailing service has been introduced but some people drive less frequently as a result of using ride-hailing service. A further look at the occupancy level suggests that ride-hailing cab on average transports more passengers in a single trip than does traditional taxi cab, where average occupancy of ride-hailing cabs is estimated at 2.1 while only 1.1 of the matched taxi sample. Both findings of a higher utilization rate and the occupancy level for Uber cabs suggest that Uber might be more efficient. However, the result only implies that Uber might be a good substitute of traditional car-pooling service or taxi cabs. Considering the possibility that the majority users of ride-hailing service might be shifted from taking non-driving modes (Schaller, 2018), road being less congested might attract more drivers, which leads to an increase of traffic volume and eventually worsens overall traffic condition. Instead of using those micro-level data, a use of macro-level measures such as congestion indices, average travel speed or even its impact on public transit ridership would help us to more comprehensively analyze effects of Uber on traffic.

Another research by Schaller (2018) analyzes ride-hailing service's ridership, users and usages to examine how ride-hailing services affect traffic. Various datasets have been used, which include 2016-17 National Household Travel Survey (NHTS); ride-hailing trip volumes for Massachusetts municipalities released by the Massachusetts Department of Public Utilities; and data from industry sources presenting relative trip volumes for different size metro areas and urban and suburban/rural population densities. They find that ride-hailing services transported 2.61 billion passengers in 2017; 70 percent Uber and Lyft trips are located in nine large and densely-populated metropolitan areas (Boston, Chicago, Los Angeles, Miami, New York, Philadelphia, San Francisco, Seattle and Washington DC); and riders are relatively young and mostly affluent and well-educated. Those findings are helpful to provide a general understanding of ride-hailing services. Furthermore, the fact that most of ride-hailing trips are located in metropolitan areas suggests that it is better to focus on those large metropolitan areas in order to more comprehensively capture their effects on traffic. On traffic congestion side, the report shows that ride-hailing services put 2.8 new vehicle miles on the road for each mile of personal driving removed, which counts for an overall 180% increase in driving on city streets. Moreover, based on results of the NHTS survey, they conclude that about 60% commuters would go by transit, walking, biking (or not make the trip) while about 20 percent would have used their own car and 20 percent a taxi when had ride-hailing services not an option. Thus, Schaller argues that ride-hailing services add to traffic because the services neither appeal to private car owners nor will replace private cars but attracts those who were transferred from public transit, walked or biked. The results seem to be reasonable, but we observe that ride-hailing trips account for only less than 0.4% of the total recorded trips in NHTS survey. Meanwhile, according to the NHTS survey, ride-hailing trips account for just 1.7% of the total miles travelled, while travel by personal cars account for 86% (NHTS, 2017). This finding leads me to suspect that the value of 2.8 new ride-hailing vehicle miles being added might be exaggerated. Furthermore, according to NHTS survey, 4% responses report to commute by bus while 73% people by private vehicles. We then could similarly presume that the same fraction of trips taken in a private car could be substituted by transit, walking or biking. To address the discrepancy between different studies and to conclude a more comprehensive estimation of Uber's effects on traffic including both road traffic and public transits ridership, I apply some most updated real-time data to conduct further research.

2.3 Measures of Traffic:

There are several ways to measure congestion level, for example, by measuring average speed, flow/density, delay and travel time variability. In U.S., California Department of Transportation defines congestion as occurring on a freeway when the average speed drops below 35 mph for 15 minutes or more on a typical weekday (Varaiya, 2001); Denver Regional Council of Governments (DRCOG) examines traffic congestion by studying the regional vehicle miles travelled including vehicle miles and hours of travel, average travel speed (mph) and person hours of travel (DRCOG, 2011). Since these data are not always available and discrepancy exists in measuring vehicle trip information considering the difficulty to monitor cars continuously on a traffic jam, some researchers propose to use Cell Dwell Time (CDT) information available from cellular networks to measure congestion level (Hongsakham et al, 2008). Alexander and González (2015) develop a model of urban travel demand based on mobile phone data to measure the effects of ridesharing services on congestion. To evaluate the impact of ridesharing on cumulative vehicle travel time and distance, they simulated traffic for several rideshare adoption scenarios and found that under moderate to high adoption rate scenarios, ride-hailing services would likely have noticeable impacts on congested travel times.

On measure of public transits ridership, Taylor and Fink (2003) in a literature review paper conclude that there are three main categories of factors affecting public transit ridership. The first are variables that directly or indirectly measure automobile access and utility, such as auto ownership and parking availability. They find that those variables explain the most variation in public transits ridership. The second are economic factors, such as unemployment levels, CBD employment levels, and income levels, which explain substantial portions of transit use. The third are spatial factors, such as population and employment density, traffic congestion levels, and parking availability. Those findings and suggestions on measures of congestion and transit ridership direct me to collect relevant data for this research.

3. Data and Methods:

3.1 Data:

Since this paper is interested in how an introduce of Uber affects traffic in city areas, while previous (Li et al., 2017; Schaller, 2018; Hall et al., 2018) researches show an overall ambiguous impact. Some say Uber induces more traffic as it potentially attracts people originally taking public transits to a driving mode (Schaller, 2018). While some say it relieves road congestion because Uber provides a more efficient on-demand transportation solution (Rayle et al., 2014). To continue exploring the effects of Uber on traffic and to provide a more comprehensive analysis, it is sensible to apply up-to-date real time data to testify those claims from both the public transits side and the road traffic side. In this research, I analyze the public transits side by estimating Uber's substitution effects on public transits ridership and the road traffic side by analyzing its impact on congestion. Hence, I collect data on monthly passenger volume taking public transits and congestion level in an area.

Since regional characteristics significantly affect traffic in an area, it is important to first decide what types of urban areas to choose as the research sample. There are several factors we put into consideration. First, since Uber entered different areas in various time points and has been introduced to areas like Albany, NY in just recent couple years, available real-time data are expected to be scarce. This fact motivates me to first focus on areas where Uber has been introduced. Second, according to a report from U.S. Census Bureau (2013), workers living in principal cities within metro areas have a lower rate (78%) of private vehicle usage and a higher rate of public transits usage than their suburban or non-metropolitan counterparts (89% and 91% respectively), so focusing on those principal cities is expected to help us better capturing Uber's substitution effects on public transits. Also, according to a report by Schaller Consulting (2018), ride-hailing trips are mostly concentrated in large, densely-populated metro areas. For the interest of this paper, I then decide to only focus on metropolitan areas. The first city choice is New York City (NYC), one of the busiest metropolitan area in the United States. NYC might be a good candidate for this research for following reasons. First, according to INRIX, a global company that specializes in connected car service and transportation analytics, NYC is listed as one of the most congested areas in 2018 with 133 hours lost due to congestion (Reed & Kidd, 2019). Second, NYC has the highest level of public transits usages at about 56% according to American Community Survey (2015). Furthermore, checking the Uber website as well as other news sources, we know that Uber was introduced to NYC on May 2011 and was suspended on issuance of new vehicle licenses starting from August 2018. Separate stages of Uber in NYC provide us with a natural advantage to analyze traffic conditions before and after those events in order to conclude an overall impact of Uber on traffic.

Eventually, I gathered three major datasets from National Transit Database (NTD), NYC Taxi & Limousine Commission (TLC) and MTA Bridges and Tunnels (B&T). Data for controls are also collected from various sources. National Transit Database (NTD) provides public transits ridership. It includes monthly ridership for almost all transit agencies which receive federal funding and reports ridership from January 2002 to January 2019. In NYC specifically, 41 agencies operating in New York-Newark area are included. Because not all 41 agencies are normally operating continuously after Uber was introduced to NYC, I exclude those agencies with missing data to make sure that number and components of observations are consistent though time. To get a monthly passenger volume taking all selected NYC transits, I aggregate monthly number of passengers using each of those agencies' transit service. Finally, the sample is restricted to a range from January 2008 to January 2019 and contributes 133 samples in total. One concern of using this dataset is that only nine months of data are included after Uber was suspended with the new restriction on August 2018. Hence, available data might not fully reveal the restrictive effects considering that the effects might be lagged.

To measure the effects of Uber entry on road congestion, I use data from NYC Taxi & Limousine Commission (TLC) to achieve average travel speed and yellow cab ridership in NYC. TLC includes monthly NYC yellow cab trip records consisting of pick-up and drop-off dates/times, pick-up and drop-off locations, trip distances, itemized fares, rate types, payment types, and driver-reported passenger counts of each taxi trip. The data are available from January 2009 to June 2018 and provide millions of yellow cab trip records for each month. Though TLC dataset provides number of taxi trips in each month, which helps to determine the effects of Uber entry on people's use of traditional taxi, it does not help to capture an overall trend of traffic volume in NYC.

To further estimate the effects of Uber entry on the total number of operating automobiles in NYC in order to analyze how does the introduce of Uber affect people's choice on transportation modes, I use number of vehicles crossing bridges and tunnels in NYC provided by MTA Bridges and Tunnels (B&T). Previous researchers including Schaller (2018) and Hall et al. (2018) use a dataset from American Community Survey (ACS) to achieve a measure on traffic volume. However, this dataset might be inapplicable to the current research not only because it is updated annually, but also as it is based on survey reports that might include distorted information. In hopes of matching levels of time measurements among major datasets and improving the validity of observations, I decide to use a dataset from MTA Bridges and Tunnels (B&T). B&T reports monthly number of vehicles crossing nine major bridges and tunnels including Robert F. Kennedy, Throgs Neck, Verrazzano-Narrows, Bronx-Whitestone, Henry Hudson, Marine Parkway-Gil Hodges Memorial, and Cross Bay Veterans Memorial bridges, as well as Hugh L. Carey and Queens Midtown tunnels in NYC. Eventually, it provides 112 observations for all consecutive months from January 2010 to January 2019. The major goal of including those data is to use it as a proxy to represent the total number of operating vehicles on NYC roads, which is similar to the use of taxi average travel speed to represent citywide average travel speed. Though the dataset provides finer-grained measurements, one apparent limitation of using this dataset concerns with how well it could capture the overall traffic volume trend in NYC. Meanwhile, when skimming through the original dataset to check for data consistency, I realize that there is a sudden decline in the recorded number of trips for each observation in year 2017. A further comparison between those observations in year 2017 with them in other years suggests that the number is abnormal. By checking the MTA website and maintenance schedule of the nine bridges and tunnels, I did not find anything special to explain the sudden decline, except that some tolls have updated from magnetic token pass to the E-ZPass. Adjustments of those data will be discussed in Section 3.5.

3.2 Dependent Variables:

To determine the effects of Uber entry on public transits ridership, I use monthly total number of transit trips in NYC from NTD, which follows Hall et al. (2018)'s research that explores whether Uber substitutes public transits. To determine its impact on road congestion, I use a dataset from TLC. Though the TLC dataset does not provide us with any direct measure on congestion, it provides us with detailed information on each yellow cab trip in NYC. Hence, I use the monthly average travel speed as a proxy for citywide average travel speed to estimate the effects of Uber entry on road congestion. To achieve a monthly average travel speed, I simply divide each recorded taxi trip's distance by the trip duration, then compute an average value for all trip records of a given month. A simple mathematical formula is given below:

$$\text{AvgSpeed}_\text{t} = \frac{1}{n} \sum_{k=0}^{n} \frac{\text{tripDistance}_\text{kt}}{\text{tripDuration}_\text{kt}}\text{, for n observations in month t.}$$

Further data cleansing is included to discard trips that are too short either of trip duration or distance to reduce disturbing factors. As a result, it provides us with 114 observations where each observation is composed of millions of taxi trips in a given month. One apparent limitation of using the TLC dataset concerns with its applicability to this research since the dataset is not specifically designed for a congestion analysis. Applying average travel speed as a measure of congestion also has the limitation to fully capture the overall traffic condition. To measure the effects of Uber entry on traffic volume, we use monthly number of vehicles crossing nine major bridges and tunnels in NYC as another proxy variable. Table 1 presents descriptive statistics of our measures on congestion. "num_transit" is total number of public transit trips in NYC for a given month; "taxi_speed" represents the average travel speed in NYC generated from yellow cab trips for a given month with the assumption that a decrease in average travel speed indicates a more congested traffic; "num_taxi" accounts for total number of yellow cabs in NYC for a given month; and "num_bridge_total" is total number of vehicles cross the nine major bridges and tunnels in NYC for a given month, where it serves as a proxy for traffic volume in NYC as mentioned before.

3.3 Control Variables:

To control for externalities, I follow previous researches' suggestion on independent variables that are expected to affect the congestion level and public transit ridership. Following Hall et al. (2018) paper, I include non-farm employment data from the Bureau of Labor Statistics (BLS) and monthly gas prices from the U.S. Energy Information Administration (EIA). The expectation is that an increase of non-farm employments will introduce more traffic. And an increase of gas prices will shift more people to use public transits and hence less traffic on roads. Additional control of weather condition including precipitation level in NYC is also included. Table 2 shows descriptive statistics of controls. Non-farm employment ("nonfarm_emp") is counted in thousands, precipitation level ("Precipitation") is counted in millimeters (mm) and gas price ("gas_price") is counted in U.S. dollar cents. I was also considering other controls such as median income and population size, as well as characteristics of the road transportation system including number of parking lots in NYC that have been suggested in Li et al. (2016)'s research. Since datasets report those measures in general update only annually, I choose to discard those controls in order to maintain the consistency among major datasets. Meanwhile, because an increase of non-farm employments potentially indicates a growing economy and population size, the use of non-farm employments might already comprise effects of both change on median income and population growth rate.

3.4 Empirical Model and Specification:

To conduct the research on estimating the effects of Uber entry on public transits ridership, average travel speed and number of operating automobiles, this research applies a regression discontinuity (RD) design model. RD design was first introduced by Thistlewaite and Campbell (1960) in a research to evaluate scholarship programs and has been used vastly in recent economic studies, including Lee (2007)'s research on U.S. house election and Cellini (2010) on exploring value of school facility investments. As a quasi-experimental pretest-posttest design, RD design reveals a causal effect of interventions by assigning cutoff points or a threshold that is used to determine whether groups of observations receive treatments (Calonico et al., 2014). The design is especially useful to estimate the average treatment effect when randomization is unfeasible. Following the nature that Uber's development in NYC includes several stages, it gives us an advantage to use the RD design and set the date Uber entered or being restricted in NYC as cutoff points. Comparing observations lying beside the cutoff points will help to estimate the effects of Uber entry and restriction and further conclude an overall effect of Uber on traffic. According to Uber's website and other news sources, I determine that Uber entered NYC on May 2011 and was restricted on August 2018. To design the model for this research, I use the date as a running variable. The dummy variable for treatment group equals "1" if the date is before the selected cutoff point, and "0" otherwise. Our basic model is then given by:

$$Outcome = \alpha + \beta_0 runningVar + \beta_1 treatment + \beta_2 Controls + \epsilon$$

, "Outcome" are measures on traffic, including number of transit trips, average travel speed, number of operating automobiles and yellow taxi cabs for a given month; "α" is a constant term; "$runningVar$" is the running variable that equals "0" on date Uber entered or being restricted in

NYC and increments or decrements before or after selected cutoff points; "$treatment$" is a dummy variable that indicates whether Uber has entered or being restricted in NYC for a given date; "$Controls$" is a collection of control variables including non-farm employment, gas prices, and precipitation level in NYC; "ϵ" is the error term; "β_{0-2}" are corresponding estimated coefficients.

Based on Hall et al. (2018)'s research, I expect that after Uber entered NYC, there would be an increase of public transits trips (positive sign for β_1). According to Li et al. (2016)'s finding that after Uber entered an area, two measures on congestion including Travel Time Index (TTI) and Commuter Stress Index (CSI) have been improved, I expect that the entry of Uber would increase average travel speed (positive sign for β_1). For the remaining controls, I expect that an increase of non-farm employments and precipitation level would cause more traffic, hence more use on public transits, a lower average travel speed and an increased number of vehicular trips.

3.5 Robustness Check:

First, considering Uber's penetration into NYC might be lagged as suggested by Li et al. (2016) and Hall et al. (2018), I consider an alternative approach to determine cutoff points for the RD design. Though Uber was officially introduced to NYC on May 2011, it does not necessarily indicate that people began to use Uber service immediately since then. To better capture the penetration level and trend of Uber use in an area, researchers such as Cramer (2016), Hall et al. (2018) suggest the use of Google Trend. Hall and other researchers have found that Google search for "Uber" are strongly correlated with the number of active drivers who drive 4 more trips per month per capita in each area. The correlation is estimated at 0.948. To more precisely determine cutoff points for RD model of this research, I replicate this approach. First, I pull original Google Trend data from Google by inserting searching keyword "Uber NYC". Then, I calculate the growth rate of Google Trend on Uber for each subsequent month in NYC. The result is shown in Figure 1. Specifically, the blue shades in Figure 1 represent searching popularity of Uber in NYC with its value ranges from 0 to 100. The growth trend is represented in orange bar in Figure 1. It suggests that starting at May 2013, the growth rate of Google Trend on Uber for the first time kept being positive for subsequent months. This finding encourages me to take May 2013 as an alternative cutoff point in addition to May 2011, the officially announced date Uber entered NYC.

Secondly, since traffic are usually affected by seasons and weathers, I integrate fixed month effect into the RD model and expect that demand for traffic varies among different months. For example, summer is considered as a traveling season in NYC. Tourists swarm into the city during summer may cause a significant increase of demand for transportation. And a worse weather condition during the winter months on the other hand is expected to deteriorate traffic. Meanwhile, including fixed effects not only helps to capture time-invariant factors but also allows other factors to be arbitrarily correlated with controls, which eventually helps to make regression results more robust (Angrist and Pischke, 2008).

Thirdly, I also generate two-way scatter points graphs before proceeding to regression analysis. An examination of those graphs provides us with graphical intuitions on the overall trend of the included measures and helps to predict certain adjustments for the model. The graphs are shown in Figure 2 with each dot represents an observation. The red line is the arbitrarily determined cutoff line of May 2013. Figure 2(a) shows a change of monthly number of transit trips in NYC from January 2008 to January 2019; Figure 2(b) shows observation of average travel speed in NYC from January 2009 o June 2018; Figure 2(c) shows a change of monthly number of vehicles crossing nine major bridges and tunnels in NYC from January 2011 to January 2019; and Figure 2(d) shows a change of monthly number of yellow cab trips in NYC from January 2009 to June 2018. Those graphs provide several interesting observations. First, I find that observations for monthly number of transit trips and vehicles crossing the nine bridges and tunnels are more scattered. Also, it seems that there are certain up and down patterns in observations for a given range of time. I hypothesize that the pattern corresponds to a change on seasons, which further encourage me to include fixed months effect in the RD design. Secondly, it shows an obvious downward trend of the yellow cab average travel speed and numbers after Uber entered NYC in Figure 2(b) and Figure 2(d). This finding motivates me to further implement polynomial regressions into the original RD model. Trying different orders of polynomials might be especially helpful for analysis of the average travel speed. To do so, I create a slope dummy variable by multiplying running variable and dummy treatment variable in the RD design. Thirdly, Figure 2(c) agrees with previous finding in data description section that there is a sudden decline of number of bridges and tunnels trips for observations in year 2017.

Since we do not know whether the decline was caused by internal, for example, systematic change, or external factors such as construction of new alternative roads, we tend to leave those observation unchanged. But because the overall trend of the observations is clear, I adjust each observation in year 2007 by using the average value of the same month in other years. The arbitrary adjustments might or might not bias the result based on future findings of what happened to those nine MTA bridges and tunnels in year 2007.

Finally, since observations for monthly number of transit trips seem much more scattered as suggested in Figure 2, I compare variance and mean of the measure of public transits ridership in order to determine severity of data dispersion. Result suggests that variances of the measure is much greater than conditional mean by each month. Hence, we could conclude that over-dispersion exists. To address this issue, I apply a negative binomial regression model. Negative binomial regression can be regarded as a generalization of Poisson regression model but with less restrictions of Poisson regression, for example, negative binomial regression loosens assumption that variance should be equal to mean (Hilbe, 2011), which is desired for the current analysis. Additional robustness check includes the use of variance inflator factor and correlational coefficients to test for multicollinearity among selected controls.

4. Results:

In this section, I start with a discussion on basic regression results derived from an unadjusted RD design. To consolidate the regression estimation, I apply the methods discussed in above robustness check section and proceed with additional regression analysis based on an adjusted RD model.

4.1 Basic Results:

Table 3 shows result of the baseline model with cutoff point set on May 2011. Each column specifies one measure of NYC traffic including monthly number of transit trips, yellow cabs, and vehicles crossing nine major bridges and tunnels in NYC, as well as a measure on the average travel speed. Overall, the result corresponds to our expectation. Specifically, it shows that after Uber entered NYC on May 2011, number of public transit trips has increased by about 1.52%; the average travel speed has increased by about 0.437 mph; and number of yellow cab trips has increased by about 17.2%. However, the results are only tested to be statistically significant for measures of the average travel speed and yellow cab ridership. The increased number of transit trips agrees with Hall et al. (2018)'s research. And an increase of the average travel speed agrees with Li et al. (2016)'s finding based on the assumption that an increased the average travel speed implies an improved traffic. On controls, the regression result suggests that non-farm employments and gasoline prices both have statistically significant impacts on traffic. It shows that change on non-farm employment increases number of public transit trips and yellow cab trips while slows down the average travel speed. Specifically, it has been estimated that a one percent increase of non-farm employment doubles (coefficient = 2.172) number of public transit trips, decreases the average travel speed by about 5.93 miles per hour and leads to seven times more taxi trips.

Directions of influence are expected following the prediction that an increase of non-farm employments would incur a greater demand for traffic. However, the estimated value is striking and far exceeds expectation. This incongruity encourages me to further check for robustness of the baseline model. First, I check severity of multicollinearity among selected controls by using variance inflation factor (VIF) and correlation coefficients. The result is shown in Table 4 & 5. Most of the controls are in good shape except that for non-farm employment, where I observe that it is almost perfect correlated with the running variable. And the VIF is estimated at 385.34. I realize that the strong correlation exists because non-farm employment continuously increases over time, which follows the same trend of running variable in the model. This finding leads me to use growth rate of non-farm employment instead of the direct measure on non-farm employment. A further test on multicollinearity after the adjustment shows that the multicollinearity issue is under control. Since the model is based on a time-series analysis, I further check and adjust for autocorrelation. Results from an adjusted is presented in the following section.

4.2 Results from an Adjusted Model:

Considering a possible lag on Uber's penetration into NYC as suggested by previous analysis on Google Trend of Uber, it provides me with more confidence to use May 2013 as the cutoff point instead of May 2011 in the model. Table 6 reports regression results of an adjusted model. In this adjusted model, I replace original non-farm employment measure with non-farm employment growth rate, add a slope dummy and shift the cutoff point to May 2013. To check and adjust for autocorrelation, I also include Durbin-Watson statistics. Results suggest that autocorrelation exists in all specifications except for measure of transit trips. A generalized least-squares model that estimates parameters where the errors are serially correlated was adopted to address this issue. Specification 2 – 4 in Table 6 report the updated estimation. Since observations for number of transit trips is over-dispersed as mentioned before, a negative binomial regression model has been further applied in hopes of achieving a more rigorous estimation of the effects of Uber entry on public transit ridership. The estimation is shown in Specification 5 in Table 6.

We first observe that R-squared value has been improved in each specification, which suggests that the adjusted model better fits to the observations. Regression result indicates that an introduce of Uber increases number of transit trips by about 3.28% and slows down average travel speed by about 0.122 mph holding other variables constant. Meanwhile, the measure of public transit ridership is also tested to be statistically significant at 95% level. The increase of transit trips agrees with previous finding from the baseline model as well as Hall et al. (2018)'s research. However, the decrease on average travel speed counters to Li et al. (2017)'s result in the view that entry of Uber would ease traffic congestion. This disagreement is based on the assumption that a drop on average travel speed indicates a worse traffic condition. In hopes of consolidating my finding to address this discrepancy, I examine the effects of Uber entry on number of vehicles crossing nine major bridges and tunnels in NYC. As mentioned in previous section, this research regards number of vehicles crossing bridges and tunnels as a proxy variable to represent total number of operating vehicles in NYC. If number of vehicles crossing bridges and tunnels increases after Uber entered NYC, it would possibly indicate that Uber introduces more vehicular traffic, or vice versa. However, regression result does not show that entry of Uber has any noticeable impact on this measure considering either the estimated coefficient or the statistically significant level. It also shows that entry of Uber does not seem to affect number of taxi trips. This finding disagrees with Rayle et al. (2016)'s survey result that shows ride-hailing service seems to be considered as a replacement for taxis. The discrepancy might exist due to a lack of detailed measure of Uber ridership in this research, which is also one limitation of the current analysis. Future research that includes more controls of Uber trips' characteristics are needed to conclude its overall impact on yellow cabs ridership.

For the remaining controls, one interesting observation is that change on number of vehicles crossing bridges and tunnels, a proxy variable to represent traffic volumes in NYC, has statistically significant impact on all traffic measures except public transit ridership. Specifically, one percent increase of bridge and tunnel trips boosts average travel speed by more than 0.7 miles per hour and has increased number of taxi trips by about 18.3% while does not significantly affect public transits ridership. Impacts on number of taxi trips are within expectation because an increase of total number of vehicular trips in NYC might comprise an increase of taxi ridership. But its impact on increasing average travel speed is counterintuitive. A check of correlation among those three sets of data does not show significant correlational relationship except that for taxi trips and average travel speed, which is estimated at 0.8247. This finding suggests that number of taxi trips are strongly correlated with average travel speed. From Figure 2, we observe that number of yellow cab trips continuously decreases after Uber entered NYC, which follows a similar trend of that for average travel speed. One possible explanation could be that new transportation tools like Uber substitute traditional taxi as suggested by Rayle et al. (2016). But as mentioned above, regression result from this research does not show that Uber entry has noticeable impact on yellow cab ridership.

Effects of change on precipitation level is as expected. Regression results show that one percent increase of precipitation level decreases number of transit trips by 0.024% and slows down average travel speed by about 0.226%. Gasoline price in general has a trivial impact on all measures and is statistically insignificant except for measure on average travel speed, where it shows that one cent increase of gasoline price increased the speed by about 0.2%. Following previous expectation, it suggests that the fixed month factors in general have statistically significant impact on all four measures of traffic.

4.3 Effects of Uber Restriction:

Given that Uber was suspended with issuance of new vehicle license in NYC on August 2018, a further look at how the restriction of Uber affects traffic might help to consolidate previous findings. Since the suspension was enacted just recently, most traffic measures are unavailable except for public transits ridership as well as bridges and tunnels trips. Applying a similar RD model, I examine the restrictive effects of Uber on number of transit trips as well as bridges and tunnels trips in NYC. Some controls such as precipitation level are excluded because of data unavailability. Regression results are reported in Table 7. Though not statistically significant, sign of the coefficients for treatment dummy is within expectation and agrees with previous analysis on exploring the effects of Uber entry. It suggests that after Uber was restricted, number of public transit trips has decreased by more than 2%. Combining the previous results, it indicates that Uber complements public transit ridership instead of substituting it. This finding counters to Schaller (2018)'s claim that Uber introduces more traffic as it diverts people who originally took public transits into a driving mode. However, given that the regression results are not statistically significant, and number of observations is restricted while the restriction effects might be lagged, future studies are needed to consolidate the arguments. Estimations of remaining controls agree with previous analysis. Specifically, it suggests that an increase of non-farm employments incurs more public transit trips as well as bridge and tunnel trips. Also, an increase of bridge and tunnel trips leads to an increase of public transit trips.

4.4 Discussion:

Major findings from the regression analysis suggest that an introduce of Uber in NYC 1) increases number of transit trips by about 3.3%, 2) slows down the average travel speed by 0.122 mph, 3) has trivial effects on number of vehicles crossing nine major bridges and tunnels in NYC. A further investigation of effects of Uber restriction on traffic consolidates the first finding and suggests that Uber complements public transits.

To further conclude all those findings, I propose the following hypotheses. Considering the fact that an introduce of Uber in NYC brings additional vehicles on roads and drivers who might be similar to taxi drivers in regard to driving behaviors, overall road traffic is expected to be worse with a decreased average travel speed. The deteriorated traffic might further discourage commuters who originally drove personal vehicles for commute and divert them into alternative transportation modes such as public transits. This explains why there is an increase of public transit trips and why the total traffic volume (represented by bridge and tunnel trips in this research) was not affected. To further explain why average travel speed has decreased, I start from generalizing characteristics of Uber trips based on previous researches (Schaller, 2018; Hall and Krueger, 2018). It has been suggested that Uber drivers are in general much younger than taxi drivers who might also be more familiar with local traffic, for example, having knowledge of shortcuts or a provision on congestion. Also, considering that Uber drivers need to frequently check phone screens for navigation and prompt information about next passenger, it might suggest that Uber drivers are driving with more distractions and hence more slowly, which eventually interferes the whole traffic. Furthermore, unlike taxi cabs that could pick up a passenger immediately after the person is targeted, Uber might share a greater fraction of time on picking up the passenger because Uber drivers sometimes also need to search and match for the reserved passenger around the given location. The slower driving speed during the searching process may further impede the traffic. However, those hypotheses heavily rely on the manipulation of Uber drivers' characteristics that I could not testify in the current analysis. Future researches that include such measures on characteristic of Uber trips are needed.

Though the regression results are significant and follow expectation, there are several limitations of this research. Firstly, instead of using a DiD model that has been applied by most of the previous researches (Li et al., 2017; Hall et al., 2018), this research uses an RD model and focuses on only one metropolitan area. This restriction on sample selection might bias the overall regression analysis, thus deterring me to generalize its implication to other urban areas. Including more diversified observations from various areas may help to mitigate this issue. Secondly, due to the fact that most congestion indices are outdated, I in this research use average travel speed as an alternative measure of congestion. There are two concerns of using this measure. First, because of the general data unavailability, I generate average travel speed by using yellow cab trips in NYC from the NYC Taxi & Limousine Commission dataset. The use of taxi average travel speed as a proxy for citywide average travel speed might bring bias to the analysis considering that taxi trips constitute only part of the whole traffic. Meanwhile, the use of average travel speed as a measure of congestion might be problematic since an increase of average travel speed does not necessarily mean traffic has been improved considering that there are many factors affecting traveling speed. Applying more rigorous measures on congestion such as TTI, CSI and travel delay time in future researches may help to mitigate this issue. Thirdly, similar to the use of taxi average travel speed to measure congestion, the use of bridge trips as a proxy to represent the total number of commuters or number of operating vehicles in NYC might be biased as bridge trips only account for certain road traffics. Including more comprehensive measures of traffic volume and number of commuters when relevant data become available will solve this issue. Fourthly, to provide a finer-grained analysis, I use monthly-level data in this research. Hence, controls such as median income, parking lot availability and line-miles that might affect traffic (Li et al., 2016; Hall et al., 2018) were discarded due to data unavailability. The exclusion of those potentially prominent controls might also bias regression results.

5. Conclusion:

Mass production leads to a glut of unnecessary and similar products, one example is automobiles. Though competition is welcomed on the production side for an industry's development, the industrial and over-consumption waste pose many societal as well as environmental concerns. The concept of sharing economy seems to be a panacea intuitively when it has been first introduced. In this paper, I evaluated one of the most successful products applying sharing economy model, Uber, and its impact on traffic. Specifically, this research explores the effects of Uber entry and restriction on traffic in New York City (NYC) by applying up-to-date real time data. Major regression results from a use of RD design suggest that after Uber was introduced to NYC, number of public transit trips has increased by about 3.3%, average travel speed slows down by 0.122 miles-per-hour and overall traffic volume was not affected. A further investigation of the effects of Uber restriction in NYC on Aug 2018 consolidates the major findings. The result counters to previous research's claim (Schaller, 2018) that Uber as well as other ride-hailing services are likely to attract people who originally took public transits and divert them into a driving mode. Instead, it suggests that traffic volume in NYC was not affected by the presence of Uber, but Uber might complement public transits, which are in-line with Hall et al. (2018)'s finding.

To further conclude the major findings in this research, I propose the following hypothesis. After Uber entered NYC, more Uber cabs and drivers are introduced, which potentially worsens traffic. The deteriorated traffic further discourages people to drive in the city and divert them into using public transits. The hypothesis is tentative considering several limitations embedded in this research. First, the use of proxy variables generated from certain sectors of the entire traffic to conclude the estimation, for example, a use of taxi average speed to represent citywide average speed in NYC, might bias the regression result. Secondly, the hypothesis is based on two assumptions and some manipulation on general characteristics of Uber driver: 1) Uber drivers are similar to taxi drivers in regard to driving behaviors, 2) and taxi drivers have distinctive driving behaviors, which might not hold without further empirical evidence. Thirdly, given that this research focuses only on New York City, the same implication might not apply to other areas with distinctive city characteristics. Hence, the recommended policy suggestions in this research might be more applicable to metropolitan areas that are similar to NYC, such as San Francisco and Boston.

Though restricted, these findings nevertheless indicate that Uber as well as other ride-hailing services expand urban mobility. As traditional remedies that have been used to ease congestion, for example, widening and building more roads, become less practical and more expensive (Downs, 2004), how to more efficiently utilize the existed infrastructures turns to be transportation authorities' major concern (FHWA, 2017). Improving and encouraging the use of public transits have always been supposed as one of the most efficient way to ease congestion. But researchers (Rubin and Mansour, 2013; Hall et al., 2018) have found that public transits utilization rate in U.S. is relatively low, and the reason is that residents in U.S. on average are more scattered while the general design of public transits is more fixed, route-wise and schedule-wise. Considering that the introduce of Uber to NYC leads to an increased use of public transits and given that it is possible that commuters are taking Uber cabs to a transportation hub, local transportation agencies could collaborate with Uber as well as other ride-hailing service to encourage the use of public transits. But since we do not yet know the fundamental reason why such the increase happened while I assume it was because of commuters' combined use of the two modes, future researched that incorporate an analysis on pick-up and drop-off locations of Uber trips will help to determine the Uber's use case and to further distinguish impacts of Uber on other transportation modes.

Overall, this research provides a further up-to-date analysis of Uber's effects on traffic and suggests that Uber complements public transits in NYC but worsens road traffic by slowing down average travel speed. To conclude an overall effect of Uber on traffic, however, it requires future researches to also look at characteristics of Uber trips as well as to include more areas as sample. Since one major concept of Uber is based on a peer-to-peer sharing economy model, this research also provides an analytical framework for researchers who are interested in evaluating similar sharing economy products like Airbnb and eBay.

VI. Tables and Graphs

Table 1 Descriptive statistics of measures on congestion

		(1)	(2)	(3)	(4)	(5)
VARIABLES	units	N	mean	sd	min	max
num_transit	(person)	133	3.397e+08	1.904e+07	2.818e+08	3.856e+08
taxi_speed	(mph)	114	12.48	0.764	10.91	13.91
num_taxi	(#vehicle)	114	1.283e+07	2.111e+06	8.340e+06	1.600e+07
num_bridge_total	(#vehicle)	109	2.17e+07	2117800	1.43e+07	2.70e+07

Table 2 Descriptive statistics of controls

VARIABLES	units	N	mean	sd	min	max
nonfarm_emp	(person)	133	4,090	288.5	3,706	4,603
Precipitation	(mm)	126	4.141	2.347	0.360	18.95
gas_price	(U.S. cents)	121	292.4	61.49	173.3	402.1

Table 3 Regression result from baseline model

VARIABLES	(1) log(Transit Trips)	(2) Avg. Travel Speed	(3) log(B&T Trips)	(4) log(Taxi Trips)
Running variable,=0 on May 2011	-0.00418**	-0.0110*	-0.000611	-0.0209***
	(0.00170)	(0.00574)	(0.00258)	(0.00239)
Uber Entry (2011-05)	0.0152	0.437***	-0.0274	0.172***
	(0.0111)	(0.119)	(0.0191)	(0.0189)
log(Bridge & Tunnel Trips)	0.0914	0.689*		0.0904
	(0.0606)	(0.338)		(0.0920)
log(Non-farm Employment)	2.172**	-5.930*	0.552	7.048***
	(0.838)	(2.727)	(1.230)	(1.251)
log(Precipitation Level)	-0.00978	-0.102**	-0.0130	-0.0295***
	(0.00575)	(0.0376)	(0.00944)	(0.00767)
Gasoline Price	0.000169**	0.00251***	-0.000183*	0.000637***
	(7.12e-05)	(0.000602)	(9.29e-05)	(0.000160)
Constant	0.117	49.52**	12.41	-43.43***
	(6.907)	(21.57)	(10.16)	(10.72)
Observations	102	102	102	102
R-squared	0.293	0.925	0.177	0.913
Number of month	12	12	12	12
Month FE	YES	YES	YES	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4 Correlation Matrix

VARIABLES	running_variable	dummy_treatment	num_bridge_total	nonfarm_emp	precipitation	gas_price
running_variable	1					
dummy_treatment	0.630***	1				
num_bridge_total	0.215*	0.0951	1			
nonfarm_emp	0.997***	0.608***	0.223*	1		
precipitation	-0.136	-0.0485	-0.0269	-0.135	1	
gas_price	-0.581***	0.0229	-0.130	-0.619***	0.140	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5 VIF

Variable	VIF	VIF(Adjusted)
running_variable	351.82	4.06
dummy_treatment	2.84	3.74
num_bridge_total	4.3	4.42
nonfarm_emp	385.34	1.18
precipitation	1.09	1.09
gas_price	4.4	1.87
month		
2	1.96	1.98
3	2.19	2.19
4	2.47	2.45
5	3.62	3.75
6	3.39	3.5
7	3.07	3.18
8	3.22	3.26
9	2.34	2.34
10	2.57	2.59
11	2	2.04
12	2.02	2.02
Mean VIF	45.8	2.68

Table 6 Regression result from adjusted model

VARIABLES	(1) log(Transit Trips)	(2) Avg. Travel Speed	(3) log(B&T Trips)	(4) log(Taxi Trips)	(5) Transit Trips (nbreg)
Running variable,=0 on May 2013	0.000633	0.000980	-0.000257	0.00228**	0.000621
	(0.000377)	(0.00555)	(0.00192)	(0.000969)	(0.000421)
Uber Entry (2013-05)	0.0319**	-0.122	-0.00181	0.00576	0.0328***
	(0.0137)	(0.135)	(0.0419)	(0.0256)	(0.0118)
Non-farm Employment Growth Rate	0.0229	-0.0767	0.0382*	0.00507	0.0226
	(0.0155)	(0.0784)	(0.0196)	(0.0217)	(0.0147)
log(Bridge & Tunnel Trips)	0.0977	0.723*		0.183*	5.11e-09**
	(0.0601)	(0.404)		(0.0936)	(2.52e-09)
log(Precipitation Level)	-0.00927*	-0.0226	-0.00938	-0.0175**	-0.00242**
	(0.00448)	(0.0284)	(0.00722)	(0.00768)	(0.00114)
Gasoline Price	2.61e-05	0.00203**	-0.000197	-3.29e-05	3.37e-05
	(5.23e-05)	(0.000821)	(0.000259)	(0.000154)	(7.23e-05)
Slop Dummy	-0.00110**	-0.0297***	0.00115	-0.0119***	-0.00113**
	(0.000383)	(0.00702)	(0.00243)	(0.00123)	(0.000545)
Constant	17.99***	0.574	16.80***	13.47***	19.48***
	(1.016)	(6.800)	(0.103)	(1.575)	(0.0591)
Observations	102	101	101	101	102
R-squared	0.371	0.875	0.760	0.953	N/A
Number of month	12	12	12	12	12
Month FE	YES	YES	YES	YES	YES
Autocorrelation	N/A	YES	YES	YES	N/A

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 7 Restriction Effects

VARIABLES	(1) log(Transit Trips)	(3) log(Bridge Trips)
Running variable,=0 on Aug 2018	0.000435***	0.000215
	(8.91e-05)	(0.000419)
Uber Restriction (2018-08)	-0.0222	-0.165***
	(0.0232)	(0.0508)
Non-farm Employment Growth Rate	0.0376*	0.0349
	(0.0193)	(0.0235)
Gasoline Price	6.60e-05	-0.000344*
	(6.58e-05)	(0.000201)
Slop Dummy	-0.00681	0.0617***
	(0.00592)	(0.0159)
log(Bridge & Tunnel Trips)	0.0406	
	(0.0499)	
Constant	18.96***	16.86***
	(0.855)	(0.0525)
Observations	109	109
R-squared	0.188	0.995
Number of month	12	12
Month FE	YES	YES
Autocorrelation adjusted	N/A	YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure 1 Google Trend of Uber in NYC

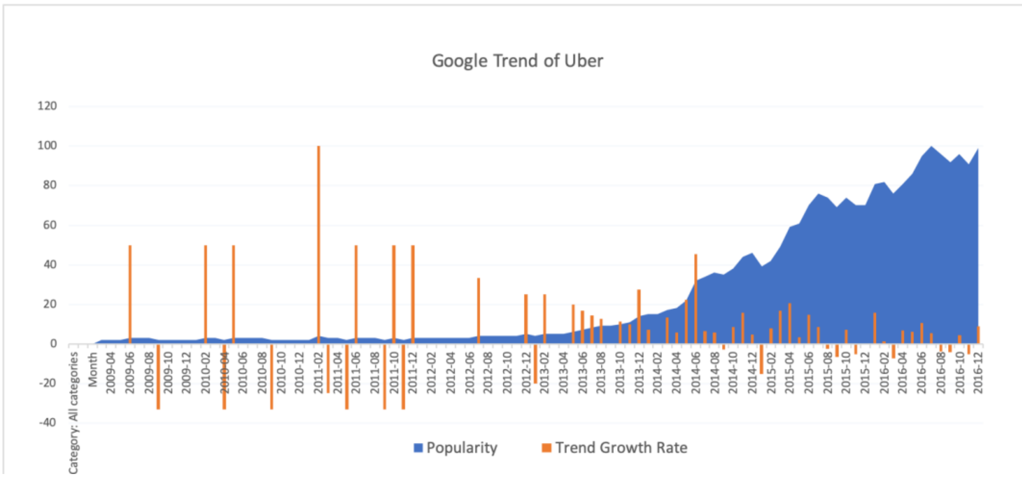

Figure 2 Two-way scatter plots showing each measure against months

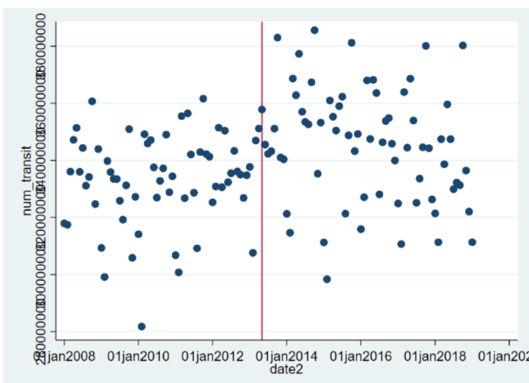

Figure 2 (a) monthly number of transit trips

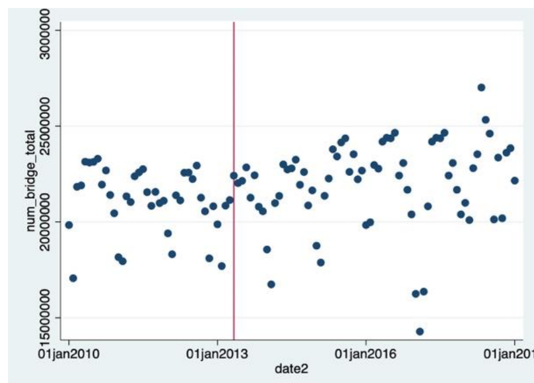

Figure 2 (c) bridge and tunnel trips

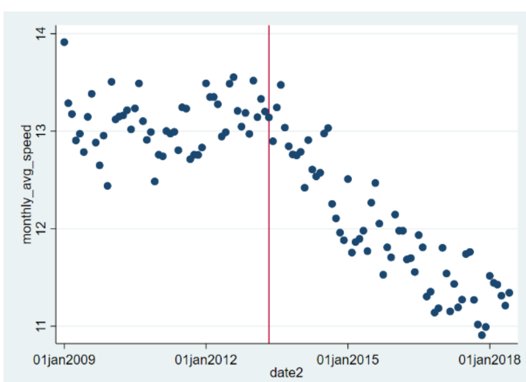

Figure 2 (b) average travel speed

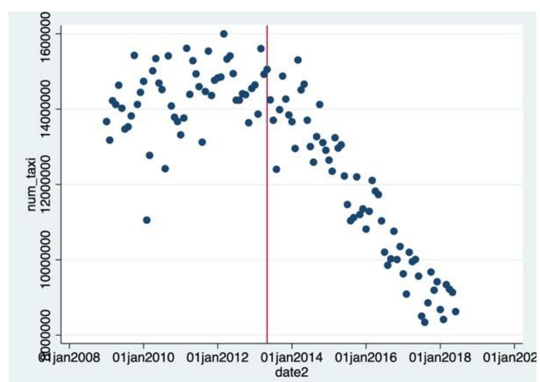

Figure 2 (d) yellow cab trips

Citation:

Aftabuzzaman, M., Currie, G., & Sarvi, M. (2010). Evaluating the congestion relief impacts of public transport in monetary terms. Journal of Public Transportation, 13(1), 1.

Barter, P. (2018, August 17). "Cars are parked 95% of the time". Let's check! Retrieved from https://www.reinventingparking.org/2013/02/cars-are-parked-95-of-time-lets-check.html

Caulfield, B. (2009). Estimating the environmental benefits of ride-sharing: A case study of Dublin. Transportation Research Part D: Transport and Environment, 14(7), 527-531.

Campbell, J. (2017, June 29). Uber, Lyft now available: What we know now. Retrieved April 2, 2019, from https://www.democratandchronicle.com/story/news/politics/albany/2017/06/27/uber-lyft-launch-thursday/103237788/

Castiglione, J., Cooper, D., Sana, B., Tischler, D., Chang, T., Erhardt, G. D., ... & Mucci, A. (2018). TNCs & Congestion.

Cellini, S. R., Ferreira, F., & Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. The Quarterly Journal of Economics, 125(1), 215-261.

Cramer, J., & Krueger, A. B. (2016). Disruptive change in the taxi business: The case of Uber. American Economic Review, 106(5), 177-82.

Dailey, D. J., Loseff, D., & Meyers, D. (1999). Seattle smart traveler: dynamic ridematching on the World Wide Web. Transportation Research Part C: Emerging Technologies, 7(1), 17-32.

Downs, A. (2000). Stuck in traffic: Coping with peak-hour traffic congestion. Brookings Institution Press.

Downs, A. (2004). Why traffic congestion is here to stay.... and will get worse.

Duranton, G., & Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. American Economic Review, 101(6), 2616-52.

Fellows, N. T., & Pitfield, D. E. (2000). An economic and operational evaluation of urban car-sharing. Transportation Research Part D: Transport and Environment, 5(1), 1-10.

FHWA. (2017). Operations Story. Retrieved February 23, 2019, from https://ops.fhwa.dot.gov/aboutus/opstory.htm

Golob, J., & Giuliano, G. (1996). Smart traveler automated ridematching service lessons learned for future ATIS initiatives. Transportation Research Record: Journal of the Transportation Research Board, (1537), 23-29.

Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit?. *Journal of Urban Economics*, *108*, 36-50.

Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. Journal of the association for information science and technology, 67(9), 2047-2059.

Hilbe, J. M. (2011). Negative binomial regression. Cambridge University Press.

Jacobson, S. H., & King, D. M. (2009). Fuel saving and ridesharing in the US: Motivations, limitations, and opportunities. Transportation Research Part D: Transport and Environment, 14(1), 14-21.

Kenton, W. (2018, December 13). Sharing Economy. Retrieved from https://www.investopedia.com/terms/s/sharing-economy.asp

Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. Journal of Econometrics, 142(2), 675-697.

Li, Z., Hong, Y., & Zhang, Z. (2016). Do ride-sharing services affect traffic congestion? An empirical study of uber entry. Soc. Sci. Res. Netw, 2002, 1-29.

NYC Department of Transportation. (2018). New York City Mobility Report.

Rao, A. M., & Rao, K. R. (2012). Measuring urban traffic congestion-a review. International Journal for Traffic & Transport Engineering, 2(4).

Rayle, L., Dai, D., Chan, N., Cervero, R., & Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transport Policy, 45, 168-178.

Rayle, L., Shaheen, S. A., Chan, N., Dai, D., & Cervero, R. (2014). App-based, on-demand ride services: Comparing taxi and ridesourcing trips and user characteristics in San Francisco (No. UCTC-FR-2014-08). Berkeley, CA: University of California Transportation Center.

Reed, T., & Kidd, J. (2019). INRIX Global Traffic Scorecard

Ride-hailing. (n.d.) In Cambridge Dictionary. Retrieved from https://dictionary.cambridge.org/us/dictionary/english/ride-hailing

Rubin, T. A., & Mansour, F. (2013). Transit Utilization and Traffic Congestion: Is There a Connection? (No. Policy Study 427).

Schaller, B. (2018). The New Automobility: Lyft, Uber and the Future of American Cities.

Schrank, D., Eisele, B., Lomax, T., & Bak, J. (2015). 2015 urban mobility scorecard.

Salomon, I., & Mokhtarian, P. L. (1997). Coping with congestion: Understanding the gap between policy assumptions and behavior. Transportation Research Part D: Transport and Environment, 2(2), 107-123.

Taylor, B. D., & Fink, C. N. (2003). The factors influencing transit ridership: A review and analysis of the ridership literature.

U.S. Environmental Protection Agency (EPA) (2018, October 09). Sources of Greenhouse Gas Emissions. Retrieved February 13, 2019, from https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions