New York City Taxis in an Uber World

Kristin Mammen
City University of New York - College of Staten Island

Hyoung Suk Shim
City University of New York - College of Staten Island

We empirically examine the effect of Uber's presence on the demand for medallion taxi trips in New York City. We estimate the percent change in number of Yellow and Green cab trips given a one percent change in number of Uber rides – the elasticity - using rainfall as an instrumental variable. City-wide, Uber rides supplement, rather than replace, Yellow and Green cab rides. For Yellow cabs, this result is powered by the area of Manhattan below 110th street; however during the morning rush only, Uber rides replace Yellow cab rides there. These results suggests Uber competition will have quite different effects in markets depending upon the thickness and vigor of the existing taxi market, and site-specific commuting patterns.

INTRODUCTION

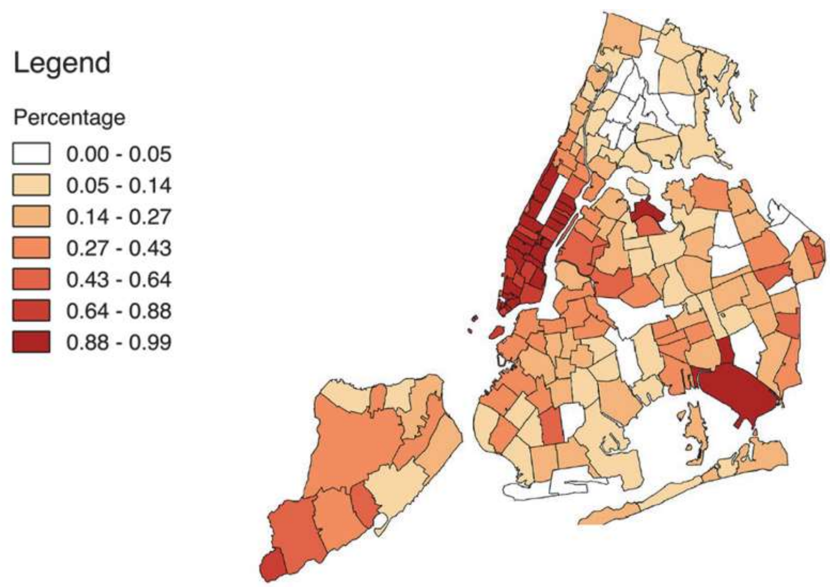
Uber, the leading smartphone app based ride-hailing company, has been touted for the efficiency of its service. The benefits to both passengers and drivers have recently been examined by Cohen et al. (2016), Cramer and Krueger (2016), Hall and Krueger (2016), and Chen et al. (2017). In the meantime, Uber has experienced temporary bans from major cities such as London, Delhi, India, and Austin, Texas. Critics have assailed Uber for opaque p assenger safety requirements, increased traffic congestion, labor practices regarding its "driver-partners," and its surge pricing policies. Another critical issue, that has not been well-studied, is Uber's impact on existing taxi cab services. Evaluating whether Uber competition is a threat to the traditional taxi industry will inform the strategies of Uber and cities with disputes.

New York City (NYC) has a large and well established taxi market, appropriate as an experimental field to conduct research on Uber's impact. The NYC Taxi & Limousine Commission (TLC) has regulated the medallion cab service (Yellow cabs) for almost fifty years. Recently, TLC launched a new medallion cab service, called street hail livery (SHL, or Green cabs).[1] In 2015, the medallion taxi fleet comprised 7,676 Green cabs and 13,587 Yellow cabs (TLC 2016 page 1).

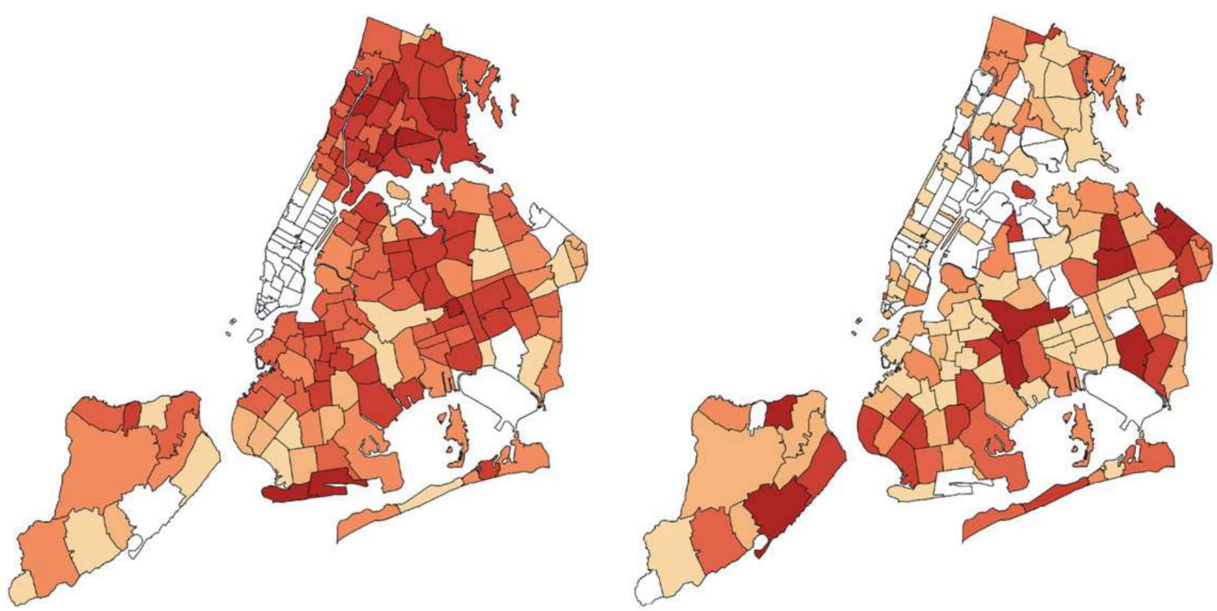
Specifically, we estimate the percent change in Yellow and Green cab trips given a one percent change in Uber rides - the elasticity - using NYC medallion taxi trip records and Uber pick-up records from April to September 2014, and January to June 2015. We use rainfall as an instrumental variable to control for endogenous factors affecting medallion taxi demand in a taxi trip demand model.

The Uber–rides elasticity of demand for cab rides for the entire New York City area is about 4.7% for Yellow cab rides and 9.1% for Green cab rides. These GMM estimates have strong statistical significance and sufficiently small overidentification test statistics.

FIGURE 1
SPATIAL DISTRIBUTION: PROPORTION OF PICK-UPS BY ZIP CODE



(a) Yellow cab pick-ups



(b) Green cab pick-ups **(c) Uber pick-ups**

Data: NYC medallion taxi trip records and Uber pick-up records, April to September 2014 and January to June 2015. The unit of observation in the dataset is zipcode-hour-day-month-year.
Unit of observation in the figure: zipcode. The proportion of pick-ups by zip code is calculated by summing the pickups for each type of service over all hour-day-month-years of a given zip code, and dividing each by the sum of pick-ups over all hour-day-month-years for *all three* services in that zip code.

We find that the distribution of rides across the boroughs differs for Yellow cabs, Green cabs, and Uber cars. Figure 1(a) shows the proportion of Yellow cab rides relative to the sum of all three types of rides in each zip code, 1(b) the proportion for Green cab rides, and 1(c) for Uber rides. Yellow cabs predominate in the core Manhattan zone, below West 110th Street and East 96th Street[2], and at the airports, Green cabs in the Bronx and in patches of Brooklyn and Queens. Uber cars predominate in some of the outer areas of the Bronx and Queens, and parts of Staten Island. We therefore disaggregate the data into boroughs and divide Manhattan into the areas above and below 110th Street. The Uber-ride elasticity estimate of demand for Yellow cab rides is statistically significant only in Manhattan below 110th Street, about 4.1%. Looking at the median daily trip statistics in Table 1, 91% of City-wide Yellow cab trips and 70% of City-wide Uber trips occurred in Manhattan below 110th Street.

TABLE 1
DESCRIPTIVE STATISTICS: NUMBER OF TRIPS BY BOROUGH

	Yellow cabs			Green Cabs			Ubers		
	Median	Std. Dev	Total Sum	Median	Std. Dev	Total Sum	Median	Std. Dev	Total Sum
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Manhattan									
below 110th Street	386,145	44,875	139,228,820	3,578	1,025	1,237,478	28,928	18,213	12,489,445
above 110th Street	5,234	1,279	1,987,334	10,135	2,710	3,609,152	891	1,132	515,003
Bronx	297	117	121,565	3,601	1,144	1,308,222	369	637	254,801
Brooklyn	8,356	3,810	3,513,685	16,440	6,567	6,074,914	6,544	5,327	2,811,660
Queens	15,320	2,148	5,563,737	13,023	4,021	4,703,185	2,408	2,858	1,399,821
Staten Island	5	4.62	2,064	7	4.94	2,616	15	20	7,992
Airports	8,439	1,169	3,093,857	161	52.1	58,952	828	626	378,678

See data note for Figure 1. *Unit of observation in the table:* day-area. For each geographic area, the number of rides in each zipcode in the area and in the 24 hours of a given day are summed to find the number of rides for that day in that area. The statistics Median and Standard Deviation (Std. dev) for the area are calculated from these rides per day figures. Summing over the rides per day for the area gives Total Sum.

We also disaggregate the data by time of day and weekday/weekend. In Manhattan below 110th Street, we estimate a -2.1% Uber-ride elasticity during the morning rush hour between 6 am and 9 am on weekdays, and 3.9% during the weekend. The negative coefficient for the Uber elasticity in Manhattan below 110th Street during the morning rush hour implies that Uber rides replace, rather than supplement, medallion taxi trips during the rush in the central business district of Manhattan. In the boroughs outside Manhattan, we estimate 4.6% Uber elasticity during the morning rush hour, and 9.7% during the weekend, suggesting that Uber rides supplement, rather than replace, medallion taxi trips during the morning rush outside Manhattan, and on the weekends.

We observe an opposite pattern in the Uber elasticity of Green cab trip demand. In the outer boroughs, we find statistically significant Uber-ride elasticities of about 5.3% during the morning rush hour, and -6.5% during the weekend. This result can be interpreted as, for Green cab passengers, Uber rides supplement Green cab rides during the morning rush hour, but replace them during the weekend. However, the overidentification test statistics for Uber-ride elasticity estimates by different times of day and weekday/weekend are too large to accept them as supporting evidence.

The studies most closely related to our topic, on the supply side of the taxi market, are Farber (2015) and Brodeur and Nield (2016). They examine the NYC cabdrivers' labor supply, the well-known behavioral economics topic established by Camerer et al. (1997), Farber (2008), and Crawford and Meng (2011). Farber (2015) revisits the issue and shows that the wage elasticity of NYC cabdrivers' labor supply is positive, consistent with the prediction of the neoclassical labor supply model. He finds that when it rains, the number of taxi trips in NYC increases while the total fare income does not change; and shows that cabdrivers' heterogeneous preferences may yield negative wage elasticities. Brodeur and Nield (2016) use a similar research design, and find that the number of daily Uber rides increases on rainy days, suggesting that Uber drivers respond positively to increases in demand. The validity of the instrumental variables in the current investigation relies on the positive effect of precipitation on the number of Uber and medallion taxi rides.

The closely related studies on the demand side of the taxi market are Cohen et al. (2016) and Buchholz (2016). Cohen et al. (2016) utilize a large-scale dataset of individual Uber trip records for four U.S. cities: New York, Chicago, Los Angeles, and San Francisco. They use Uber's surge pricing algorithms to identify the price elasticity of demand for Uber rides at each price point, and then calculate the total associated consumer surplus. In the current investigation, we focus on estimating the elasticity of demand for medallion taxis relative to changes in the quantity of Uber rides. Buchholz (2016) investigates the consumer surplus of the taxi market in NYC with respect to search friction and regulated taxi fares with a large dataset of taxi ride characteristics. He shows that if search costs are removed (as they might be if medallion taxis adopted ride-rider matching technologies like Uber's), consumer surplus is doubled by substantially increasing number of daily trips (matching taxi supply to taxi demand).

Random utility maximization has been a predominant model in the travel demand literature, since the seminal work by Domencich and McFadden (1975) and McFadden (1974). We use an aggregate version of the travel demand model, proposed by Peters et al. (2011), to develop a demand model for the count of taxi rides with a single trip mode (taxi), which allows us to estimate the elasticity of demand for taxi trips relative to the quantity of Uber rides. A number of papers have studied the demand for taxi trips using different model specifications: Douglas (1972), De Vany (1975), Beesley and Glaister (1983), Cairns and Liston-Heyes (1996), Arnott (1996), and Flores-Guri (2003). These studies analyze the taxi trip market with the fare as the unit price of the trip, and discuss whether the regulated fare yields the second best in terms of efficiency, given the monopoly pricing in the market which arises due to the use of medallion licensing as an entry control.

Jackson and Schneider (2011) and Schneider (2010) examine New York City taxi drivers' moral hazard which motivates the drivers to engage in risky driving and criminal activities. The unit of observation in these studies, however, is the individual driver's legal record, not individual taxi trips.

THE EMPIRICAL FRAMEWORK

Our primary goal is to estimate the elasticity of NYC medallion cab demand with respect to quantity of Uber rides. In order to consider spatiotemporal variation, we estimate a panel data model for taxi trip demand

$$y_{it} = \delta \cdot u_{it} + \mathbf{x}_{it}\boldsymbol{\beta} + \gamma_i + \theta_t + c_{it}, \tag{1}$$

where y_{it} is the number of NYC medallion taxi trips, u_{it} is the number of Uber trips, $\mathbf{x}_{it}$ is a vector of medallion taxi trip attributes, γ_i and θ_t are location and time specific effects respectively; and c_{it} is the location-time specific error term. The unit of location, NYC zip code, is represented by i, while t represents the time period, hour of day-month-year. To estimate the coefficient of interest, δ, we must control for the endogeneity of the demand for Uber rides and of the medallion taxi trip attributes which stem from the cab drivers' labor supply behavior. We also must account for the non-uniform and nonstationary spatiotemporal variation in the data series of the demand for taxi trips. For y_{it} and u_{it}, we take the log of number of taxi trips and the log of number of Uber trips respectively; therefore the estimate of δ is interpreted as the elasticity of demand for taxi trip rides with respect to the number of Uber rides.

Data

We use NYC medallion taxi trip records and Uber pick-up records from April to September 2014, and January to June 2015. Medallion cabs' individual trip records are available to the public from the TLC's website.[3] The records have detailed information about individual taxi trips such as pick-up and drop-off date and time, pick-up and drop-off location in GPS coordinates (latitude and longitude), trip distance, itemized fares, number of passengers, etc. Uber does not make its trip records public, so we use data provided by FiveThirtyEight (2015) that have pick-up time and location only.

TABLE 2
DESCRIPTIVE STATISTICS

	# of pick-ups			Total fare		Total trip distance	
	Yellow cab	Green cab	Uber	Yellow cab	Green cab	Yellow cab	Green cab
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Median	440,246	47,700	40,520	$6,946,085	$703,083	1,378,623	136,663
Std. dev	53,393	14,838	29,783	$859,212	$229,988	8,073,866	45,088
Max	544,519	81,574	136,193	$10,000,912	$1,569,859	60,720,968	244,962
Min	0	0	0	$0	$0	0	0
Total sum	159,481,189	17,166,393	18,804,806	$2,496,244,821	$254,233,258	1,516,589,012	50,496,604

See notes for Table 1.

Descriptive statistics are reported in Table 2 for the daily number of pick-ups, total fares collected and total trip distances. In the first row we see the median number of daily pick-ups is about 440,000 for Yellow cabs, 47,000 for Green cabs, and 40,000 for Uber cars (columns 1, 2, and 3). Columns 3 and 4 show the median total taxi fare of daily trips is about 7 million dollars for Yellow cabs, and 700,000 dollars for Green cabs. The median total distance of daily trips is about 1.4 million miles for Yellow cabs, and 136,000 miles for Green cabs (columns 6 and 7). The total number of pick-ups over the sample period, in the bottom row, columns 1, 2, and 3, is about 159 million for Yellow cabs, 17 million for Green cabs, and 19 million for Uber cars.

We aggregate the individual trip records of Yellow cabs, Green cabs, and Uber cars separately by pick-up zip code (location identifier i) and hour-day-month-year (time period identifier t). We then match and merge the records for the three taxi trip services, with unit of observation pick-up zip code and hour-month-year. For Yellow Cabs, Green cabs, and Uber cars from 2014, we assign the 248 unique NYC zip code areas to each individual trip record according to the trip's pick-up geographic coordinates, longitude and latitude. The zip code assignment for Uber pick-up is the same for the 2014 records. Instead of the single point pick-up coordinates, the 2015 Uber records have "taxi *zone* identifiers." We therefore assign zip codes to the 2015 Uber trip records using the zip code area that overlaps most with the taxi zone. The sample period comprises 364 days and 12 months. The total number of time points, hour-day-month-year, is 8,736. With 248 zip code areas assigned to each time point, the total number of observations is 2,166,528.

The rain data that we use for an instrumental variable, produced by the National Centers for Environmental Prediction (NCEP), have 1121 × 881 four square kilometer boundary grids covering the entire U.S. territories on the North American continent. The stage IV weather radar measures three meteorological quantities in each grid, reflectivity, radial velocity, and spectrum-width base. Hourly precipitation accumulation is then calculated based on the three quantities. We use the hourly precipitation data for the 189 grids covering New York City in our empirical analysis. The grid is assigned to the zipcode polygon which encompasses the greatest proportion of the grid.[4]

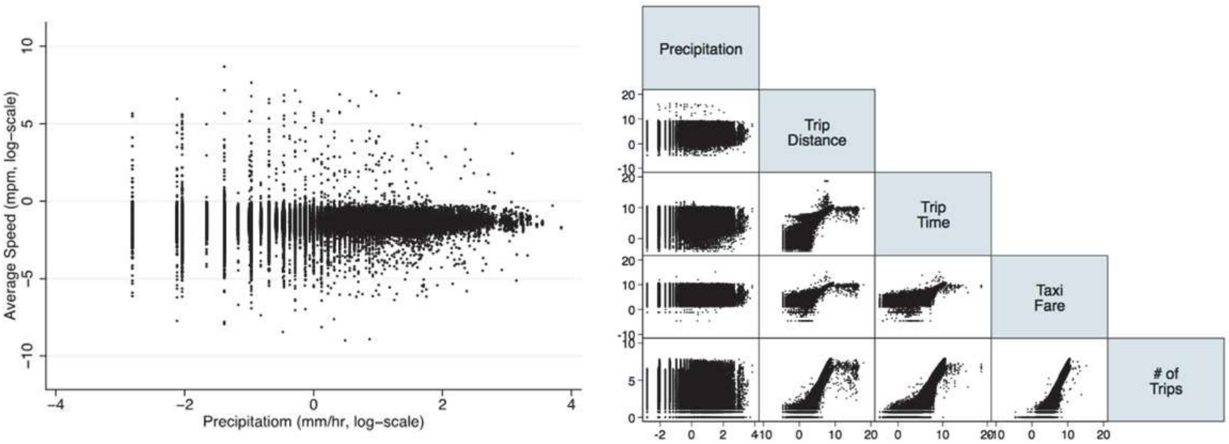
Identification

The regression model (1) is a demand model, and therefore controlling for endogeneity due to unobservable supply factors is crucial to identify the parameters in the model. Along with hourly precipitation, we use indicator variables for pick-up zip codes as instrumental variables for the number of Uber trips and the endogenous medallion taxi trip attributes such as trip distance, trip duration length, and number of passengers.

We argue that rainfall and the pick-up location of taxi trips are valid instruments because i) taxi trip demand is highly correlated with rainfall; but ii) cab drivers' labor supply is uncorrelated with rain because of the compliance rule for any passengers' trip request. Farber (2015) and Brodeur and Nield (2016) are the first studies of the effect of rain on NYC taxi cab and Uber drivers' labor supply respectively. Farber (2015) finds that taxi demand substantially increases when it rains, but drivers' income does not change. This is due to a decrease in the supply of taxi trips because i) traffic congestion gets worse when it rains; and ii) drivers prefer not to drive in the rain so they tend to stop their shifts early. Brodeur and Nield (2016) document evidence that Uber drivers positively respond to increasing demand when it rains. We therefore infer that the magnitude of the Uber drivers' response is substantially greater than the medallion cab drivers'.

We argue further that taxi trip supply is uncorrelated with rainfall due to the compliance rule. The TLC mandates drivers to accept any trip requests, unless the vehicle is occupied and the passengers do not want to pick-up additional passengers, or the prospective passenger is in possession of an article that would damage the vehicle or leave a stain or foul smell (TLC 2010 Section 2-50(e)(3)). According to the TLC rulebook (Section 2-50(a)), "a driver shall not seek to ascertain the destination of a passenger before such passenger is seated in the taxicab."

FIGURE 2
SCATTERPLOT: RAINFALL AND TAXI TRIPS (LOG-SCALE)



(a) Rainfall and Taxi Speed **(b) Rainfall and Taxi Trip Attributes**

Data: NYC medallion taxi trip records and hourly precipitation measurements, April to September 2014 and January to June 2015. The unit of observation in the dataset and in the figure is zipcode-hour-day-month-year. The taxi trip speed average is over all the trips in a given zipcode-hour-day-month-year.

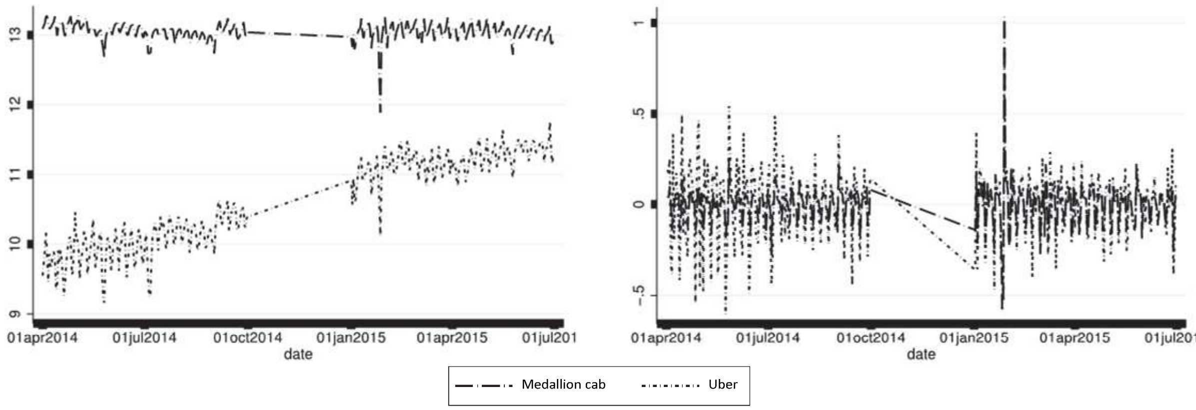
We find some support for this argument in our data. Precipitation in a given hour-location is almost uncorrelated with the average speed of taxi trips occurring in that hour-location. As shown in Figure 2(a) in log scale, the scatterplot of average taxi trip speed is almost flat with respect to precipitation. We further find that precipitation is negatively correlated with trip distance (-0.0018), trip duration (-0.0040), and total fare (-0.0013), and positively correlated with the number of taxi trips (0.0081), (all in log scale). These signs of the correlation coefficients are suggestive that taxi trips decrease in length and increase in frequency when it rains; however the magnitudes of the coefficients are very small, as illustrated in the first column of the scatterplot matrix in Figure 2(b).

Spatiotemporal Distribution

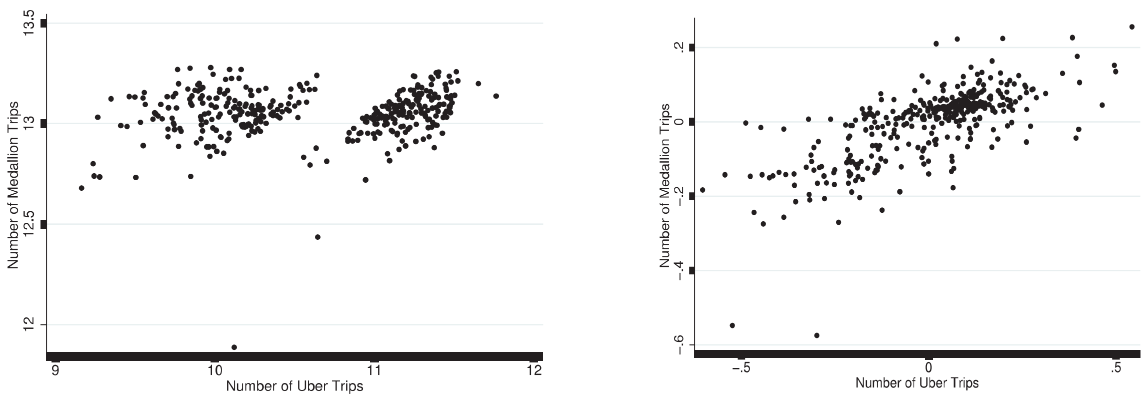
It is well-known that econometric estimation with nonstationary data may cause either inconsistent estimation of the target parameter due to serial correlation in the error term, or inefficient standard error estimation due to heteroskedasticity. To control for nonstationarity issues, we apply the first-differencing transformation by day-month- year for all variables in (1). Prima facie, the data series of the number of NYC medallion taxi trips and Uber trips are nonstationary over time, due to factors such as whether a driver's shift is a day shift or a night shifts and whether it is rush hour or not. Income targeting on the part of cab drivers, addressed by a number of behavioral economics papers, could also produce nonstationarity in taxi trips. Farber (2015), in particular, demonstrates the time variation in NYC medallion cab trips. He shows that day shift cab drivers have more rigid start times, whereas end times are more rigid for night shift drivers.[4]

To illustrate the daily variation in our dataset that may cause nonstationarity, Figure 3(a) shows time series plots for the number of medallion cab trips and Uber trips by day. Uber trips have a steady growth trend while the plot for medallion cabs is stable. These different long-run trends may prevent the estimation of the causal relationship between the number of Uber trips and the number of medallion cab trips. Unlike the log-scale scatterplot in Figure 3(c), the first differenced variables in Figure 3(d) show a clear positive linear relationship.

FIGURE 3
VARIATION OVER TIME AND SCATTERPLOTS FOR NUMBER OF TAXI AND UBER PICK-UPS



(a) Time Series Plot (Log-scale) **(b) Time Series Plot (Log-differenced)**



(c) Scatterplot (Log-scale) **(d) Scatterplot (Log-differenced)**

See data note for Figure 1. *Unit of observation in the figure:* day. The number of taxi and Uber pickups (separately) in the 24 hours of each given day in New York City are summed to find the number of rides for that day.

As we saw in Figure 1, the spatial distribution of rides shows that Yellow cabs, Green cabs, and Uber cars serve different areas of New York City. Yellow cab pick-ups occur mostly in the core Manhattan zone and the two airports, whereas Green Cabs and Uber cars have broader pick-up distributions.[6]

EMPIRICAL RESULTS

TABLE 3
ESTIMATES FOR NUMBER OF PICK-UPS WITH LOG-DIFFERENCED VARIABLES

	Yellow cabs			Green cabs		
	OLS	TSLS	GMM	OLS	TSLS	GMM
	(1)	(2)	(3)	(4)	(5)	(6)
Uber	0.0242***	0.0706***	0.0466**	0.0154***	0.1127***	0.0912***
(Log-differenced)	[0.000]	(0.017)	(0.018)	[0.001]	(0.027)	(0.025)
Trip distance	-0.0895***	-0.2494***	-0.2407***	-0.0906***	0.0543*	0.0212
(Log-differenced)	[0.004]	(0.053)	(0.063)	[0.005]	(0.029)	(0.027)
Trip duration	0.1017***	0.1044**	0.0674**	0.0461***	0.0209	0.0170***
(Log-differenced)	[0.002]	(0.047)	(0.032)	[0.001]	(0.013)	(0.004)
Passengers	0.4463***	0.5316***	0.6236***	0.4603***	0.4810***	0.4932***
(Log-differenced)	[0.002]	(0.027)	(0.045)	[0.002]	(0.035)	(0.043)
Meter fare	0.4514***	0.5269***	0.4800***	0.4841***	0.3177***	0.3513***
(Log-differenced)	[0.006]	(0.052)	(0.096)	[0.007]	(0.053)	(0.063)
Tip	-0.0420***	-0.0469***	0.0388***	-0.0401***	-0.0359***	-0.0363***
(Log-differenced)	[0.001]	(0.004)	(0.007)	[0.001]	(0.004)	(0.005)
Constant	0.0238***	-0.0297***	-0.0177**	-0.0243***	0.0065	0.0010
	[0.002]	(0.006)	(0.008)	[0.002]	(0.007)	(0.007)
# of obs	391,181	391,181	391,181	208,385	208,385	208,385
R^2	0.9008	0.8867	0.8787	0.8944	0.8774	0.8837
χ^2 Test statistic (df)		375.92 (150)	102.70 (150)		276.89 (126)	74.91 (126)
(p-value)		(0.000)	(0.999)		(0.000)	(0.999)

Data: NYC medallion taxi trip records, Uber pick-up records, and hourly precipitation measures, April to September 2014 and January to June 2015. The unit of observation in the dataset is zipcode-hour-day-month-year. These observations are log-differenced in the table.
Standard errors are reported in parentheses, heteroskedasticity robust standard errors in square brackets. The symbols *, **, and *** indicate respectively that the estimated coefficient is statistically significant at the 10%, 5%, and 1% levels. The Two-Stage Least Squares (TSLS) and Generalized Method of Moments (GMM) estimates treat "# of Uber pickups", "trip distance", "trip duration", and "# of passengers," as endogenous covariates. The instrumental variables are precipitation and the indicator variables for trip origin ZIP Code. The row for $\chi 2$ test, second from the bottom, reports the overidentification test statistics with degrees of freedom in parentheses. The associated p-values are reported in the bottom row. Note that all model estimates contains fixed effect indicator variables for i) month, ii) year, and iii) weekday.

Table 3 reports estimation results of the model (1), with Yellow cab trips in columns 1, 2, and 3, and Green cab trips in 4, 5, and 6. All variables, excluding indicators, are log-differenced from the same zip code-hour-day-month-year hour of 24 hours previous. Since all the variables are in log scale, each coefficient represents the elasticity of the designated cab rides (Yellow or Green) with respect to the corresponding variable.

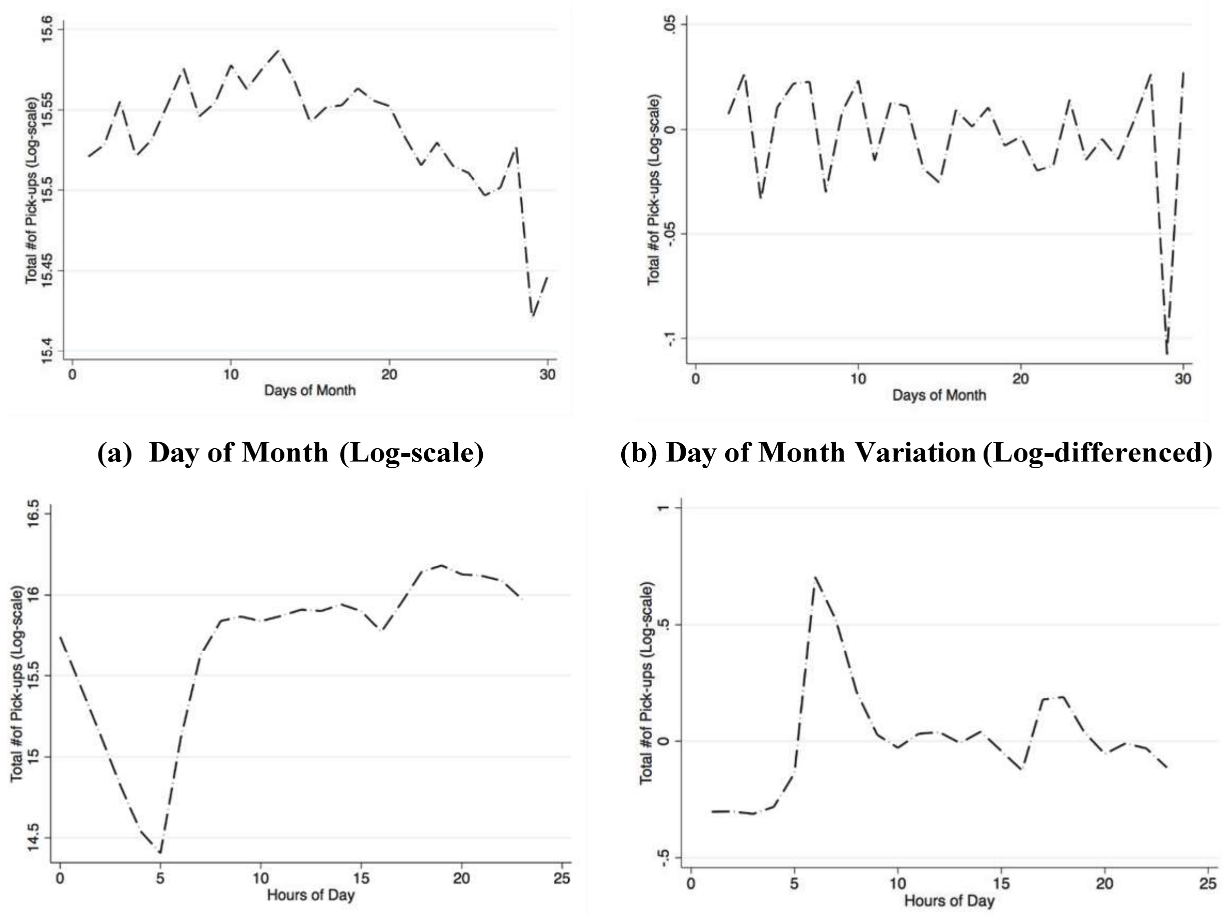
We find a City-wide Uber-ride elasticity of 4.7% for Yellow cab trips with respect to Uber rides, and 9.1% for Uber-ride elasticity for Green cab trips in the generalized method of moments (GMM) estimations in columns 3 and 6. These positive and statistically significant coefficients imply that Uber rides supplement both Yellow and Green cab trips. In particular, a 1% increase in Uber trips causes a 4.7% increase in Yellow cab trips and a 9.1% increase in Green cab trips. Although the magnitudes differ, the sign and statistical significance of the GMM estimates are the same as those from the OLS specification (columns 1 and 4) and the two-stage least squares (TSLS) specification (columns 2 and 5).

The Uber-ride elasticity of Green cab trips found in Table 3 is twice the size of Yellow cab trips. This difference results from the large disparity in market share held by Yellow and Green cabs in the five New York City boroughs. As shown in Table 1 about 91% of daily Yellow cab trips and 70% of daily Uber trips occurred in Manhattan below 110th Street, whereas only 7% of Green cab trips occurred in that area.[7] The 4.7% Uber-ride elasticity of Yellow cab trips is therefore powered by rides in Manhattan below 110th Street, and the 9.1% Uber elasticity of Green cab trips is powered by rides in Brooklyn and Queens, where 63% of Green cab trips occurred.

The GMM estimates are our preferred results, because this specification controls for the endogeneity of cab drivers' labor supply and the nonstationarity of taxi rides, providing statistically consistent Uber elasticity estimates. In addition, the overidentification (overid) test statistics show that the GMM results do not reject the null hypothesis that the instrumental variables are exogenous. Although the TSLS estimates are qualitatively similar to the GMM estimates, the TSLS overid test statistics strongly reject the null hypothesis of exogeneity. We do not believe these statistics invalidate the instrumental variables. Rather, we suspect that the heteroskedasticity resulting from the nonstationary data causes the rejection of the overidentifying restriction in TSLS.

Examining our data, we find suggestive evidence of heteroskedasticity in the hourly data series for number of taxi trips, which is nonstationary, and is successfully controlled for in the Table 3 GMM estimation. The data series for the daily number of taxi trips appears (relatively) stationary in log-differenced form, but the hourly data series does not. Figure 4 shows daily and hourly variations in the number of pick-ups before and after log-differencing. Comparing panels 4(a) and 4(b), the log-differencing appears to make the daily data series stationary, that is, the series randomly fluctuates around zero. The log-differencing for the hourly data series, however, seems to amplify the morning and evening rush hours, causing nonstationarity. In 4(c), the log-scale series declines substantially in the middle of the night. In 4(d) the log-differenced series remains nonstationary, with two peaks, one in the morning rush hour between 6 am and 9 am, and the other at the evening rush hour between 5 pm and 7 pm.

FIGURE 4
VARIATION OVER TIME BY DAY AND HOUR: MEDALLION TAXI PICK-UPS



(a) Day of Month (Log-scale) **(b) Day of Month Variation (Log-differenced)**

(c) Hour of Day (Log-scale) **(d) Hour of Day (Log-differenced)**

See data note for Figure 1. The unit of observation in figures a and b is day of month; for a given day of the month (from 1 to 31) the number of taxi pickups in New York City are summed over the 24 hours of that day over all month-years. The unit of observation in figures c and d is hour of day; for a given hour in a day (from 1 to 24) the number of taxi pickups in New York City are summed over that hour in all day-month-years.

The Effect of Uber Trips on Yellow Cab Trips

Table 4, column 1, reproduces the GMM estimate of the Uber-ride elasticity of the demand for Yellow cab trips for all of New York City from Table 3, column 3, and then reports estimates by borough and Manhattan below and above 110th Street. The overid test statistics do not reject the overidentifying restriction, and therefore the instrumental variables are valid for the estimates in Table 4.

TABLE 4
GMM ESTIMATES OF YELLOW CAB DEMAND (# OF PICK-UPS)

	Entire sample	Manhattan			Brooklyn	Queens	Bronx
		All	Below 110th	Above 110th			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Uber	0.0466**	0.0330	0.0407*	0.2278	0.1082	0.1236	0.0344
(Log-differenced)	(0.018)	(0.022)	(0.021)	(0.271)	(0.097)	(0.096)	(0.077)
# of obs	391,181	259,791	228,679	31,112	75,543	46,014	1,785
R^2	0.8787	0.8889	0.9208	0.6974	0.8758	0.8173	0.7541
χ^2 Test (df)	102.70(150)	14.75(63)	9.90(51)	0.81(8)	19.74(32)	21.40(30)	8.49(11)
(p-value)	(0.9988)	(1.0000)	(1.0000)	(0.9992)	(0.9556)	(0.8751)	(0.6690)

See notes to Table 3.

We see in column 3 that the 4.7% City-wide Uber-ride elasticity for Yellow cab trips comes mostly from Manhattan below 110th Street, where the Uber-ride elasticity estimate is 4.1%. The elasticity estimate for all of Manhattan is about 3.3%, and in Manhattan above 110th Street, 23%, although neither is statistically significant. Uber trips appear to supplement Yellow cab trips in Manhattan below 110th Street.

The Uber elasticity estimates outside Manhattan are positive but statistically insignificant. The elasticity estimates in Brooklyn and Queens have, however, Z-statistics that exceed one. It is thus too early to conclude that Uber trips have no impact on Yellow cab trip outside of Manhattan below 110th Street. We are unable to estimate the elasticity in Staten Island due to the insufficient number of observations.

In Table 5 we report Uber elasticity estimates for weekday rush hours and on the weekend, in Manhattan below and above 110th Street, and for the other boroughs grouped together. Interestingly, we have a negative Uber elasticity estimate of -2.1% in Manhattan below 110th Street during the morning rush hour, statistically significant at the 5% level. Below 110th Street, the elasticity estimate for the weekend is about 4%, close to the City-wide elasticity of Yellow cab trips. The negative morning rush hour elasticity suggests that Uber trips replace Yellow cab trips at that hour. Note, however that that the overid test statistics for both of these specifications strongly reject the overidentifying restriction. Thus, these two Yellow cab trip samples need to be re-examined with more observations. The elasticity estimate for the evening rush is less than 1% but not statistically significant.

TABLE 5
GMM ESTIMATES FOR YELLOW CAB TRIPS BY WEEKDAY RUSH HOUR AND WEEKEND

	Manhattan Below 110th st			Manhattan Above 110th st			Other Boroughs		
	Morning	Evening	Weekend	Morning	Evening	Weekend	Morning	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Uber	-0.0207**	0.0041	0.0389***	0.0254	0.1278	-0.0488	0.0458*	0.0443	0.0969***
	(0.010)	(0.009)	(0.008)	(0.141)	(0.118)	(0.086)	(0.024)	(0.028)	(0.028)
# of obs	26,714	24,604	63,605	4,070	3,112	10,115	18,352	11,750	43,390
R^2	0.9309	0.8540	0.9139	0.7698	0.8252	0.7647	0.7194	0.8386	0.7587
χ^2 Test (df)	79.73(49)	41.67(51)	233.41(51)	1.81(8)	3.00(8)	8.79(8)	88.55(70)	74.09(60)	128.71(78)
(p-value)	(0.0036)	(0.8213)	(0.0000)	(0.9863)	(0.9346)	(0.3607)	(0.0665)	(0.1043)	(0.0003)

See notes to Table 3. Morning (evening) rush hour is between 6 am (5 pm) and 9 am (7 pm) on weekdays

The Effect of Uber on Green Cab Trips

Table 6 reports GMM estimates of the Uber-ride elasticity for Green cab trips during the rush hours on weekdays, on the weekend in Manhattan above 110th Street, and in the other boroughs grouped together. There are no statistically significant Uber-ride elasticity estimates for Green cab trips at any time in Manhattan above 110th Street. In the boroughs outside Manhattan, the Uber elasticity estimate is about 5% during the morning rush hour, and is statistically significant at the 10% level. During the weekend, however, the elasticity estimate is about -6.5% and statistically significant at the 1% level, in contrast to the positive Uber elasticity for Yellow cab trips in Table 5, column 9. The negative elasticity suggests that Uber rides replace Green cab trips in the outer boroughs on the weekend. But this elasticity estimate needs to be re-examined with more observations because the overid test statistic strongly rejects the overidentifying restriction.

TABLE 6
GMM ESTIMATES FOR GREEN CABS BY WEEKDAY RUSH HOUR AND WEEKEND

	Manhattan Above 110th St			Other Boroughs		
	Morning	Evening	Weekend	Morning	Evening	Weekend
	(1)	(2)	(3)	(4)	(5)	(6)
Uber	-0.0080	0.0107	-0.0704	0.0526*	-0.0000	-0.0648***
	(0.106)	(0.086)	(0.051)	(0.031)	(0.018)	(0.024)
# of obs	4,346	3,472	10,719	20,802	17,655	52,202
R^2	0.8649	0.6257	0.8953	0.8244	0.7760	0.7799
χ^2 Test (df)	4.23(8)	0.31(8)	5.65(8)	97.81(75)	91.36(78)	208.95(89)
(p-value)	(0.8358)	(1.0000)	(0.6868)	(0.0397)	(0.1430)	(0.0000)

See notes to Table 5.

CONCLUSION

We have empirically examined the effect of Uber's presence on the demand for medallion taxi trips in New York City. Specifically, we estimate the percent change in Yellow and Green cab trips given a one percent change in Uber rides – the elasticity - using NYC medallion taxi trip records and Uber pick-up records from April to September 2014, and January to June 2015. We use rainfall as an instrumental variable in a taxi trip demand model to control for endogenous factors affecting medallion taxi demand.

We find that, City-wide, Uber rides supplement, rather than replace, Yellow cab and Green cab rides. For Yellow cabs, the City-wide positive and significant Uber-ride elasticity of the demand for Yellow cab trips is powered by the positive and significant Uber-ride elasticity in Manhattan below 110th Street, where 91% (70%) of daily Yellow cab trips (Uber trips) are initiated.

However, results indicate that whether Uber's presence supplants medallion taxi rides or increases demand for them depends on location and traffic conditions influenced by time of day and weekday/ weekend status. Our statistically significant results most often show a positive effect of Uber rides on taxi demand. However, Uber pick-ups decrease the number of Yellow taxi rides in Manhattan below 110th Street during the morning rush hour. They also decrease the number of Green cab rides on the weekend in the outer boroughs grouped together.

We view these two results with caution because in both specifications the overidentifying restriction is strongly rejected. But they suggest that Uber competition will have quite different effects in markets depending upon factors such as the thickness and vigor of the existing taxi market and site-specific commuting patterns. Documenting the market characteristics which make the presence of Uber a positive or negative force on the demand for traditional taxis is an important area for future research.

ENDNOTES

1. Green cabs are restricted from picking up passengers in the core Manhattan zone and at the two NYC airports (TLC 2013). Other than this restriction, TLC regulations are the same for both Yellow and Green cabs.
2. NYC documents define the core Manhattan zone as Manhattan below West 110th Street and East 96th Street (TLC 2013). Technically, this zone contain two major business districts, Midtown Manhattan, the central business district (CBD)), and Lower Manhattan/Wall Street, the financial district. In practice, observers often use CBD refer to all of Manhattan below 59th Street as the CBD (e.g., Merrill and Coote 2015).
3. TLC Raw Data landing page: *http://www.nyc.gov/html/tlc/html/technology/raw_data.shtml*
   TLC Trip Record Data landing page: *http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml*
   Dataset webpages:
   https://data.cityofnewyork.us/Transportation/2014-Green-Taxi-Trip-Data/2np7-5jsg
   https://data.cityofnewyork.us/Transportation/2015-Green-Taxi-Trip-Data/gi8d-wdg5
   https://data.cityofnewyork.us/Transportation/2014-Yellow-Taxi-Trip-Data/gn7m-em8n
   https://data.cityofnewyork.us/Transportation/2015-Yellow-Taxi-Trip-Data/ba8s-jw6u
4. See Hamidi et al. (2017) for details about the stage IV radar data. Many thanks to Ali Hamidi and Naresh Devineni of the National Oceanic and Atmospheric Administration/Cooperative Remote Sensing Science and Technology Center at the City College of the City University of New York for sharing the data.
5. Farber (2015) finds that there are two peaks in the hourly distribution of shift start times, additional evidence that the time variation of taxi trips is nonstationary.
6. Recall Green cabs are restricted from picking up passengers in the core Manhattan zone (and at the two NYC airports TLC (2013). This is the reason that Figure 1(b) shows Green cabs with almost no pick-ups in those areas.

7. For each type of service, find the percent of rides in Manhattan below 110th Street by dividing the total number of rides below 110th Street (in the top row of columns 3, 6, and 9 of Table 1) by the sum of the numbers in the "Total" column.

REFERENCES

Arnott, R. (1996). Taxi travel should be subsidized. *Journal of Urban Economics, 40*(3), 316-333. doi:10.1006/juec.1996.0035

Beesley, M. E., & Glaister, S. (1983). Information for regulating: The case of taxis. *The Economic Journal, 93*(371), 594-615. doi:10.2307/2232397

Brodeur, A., & Nield, K. (2016). *Has uber made it easier to get a ride in the rain?* Unpublished manuscript. Retrieved from https://www.iza.org/publications/dp/9986/has-uber-made-it-easier-to-get-a-ride-in-the-rain

Buchholz, N. (2016). *Spatial equilibrium, search frictions and efficient regulation in the taxi industry.* Unpublished manuscript. Retrieved from https://scholar.princeton.edu/sites/default/files/nbuchholz/files/taxi_draft.pdf

Cairns, R. D., & Liston-Heyes, C. (1996). Competition and regulation in the taxi industry. *Journal of Public Economics, 59*(1), 1-15. doi:10.1016/0047-2727(94)01495-7

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics, 112*(2), 407-441. doi:10.1162/003355397555244

Chen, M. K., Chevalier, J. A., Rossi, P. E., & Oehlsen, E. (2017). *The value of flexible work: Evidence from uber drivers.* NBER Working Paper 23296.

Cohen, P., Hahn, R., Hall, J., Levitt, S., & Metcalfe, R. (2016). *Using big data to estimate consumer surplus: The case of uber.* NBER Working Paper 22627.

Cramer, J., & Krueger, A. B. (2016). Disruptive change in the taxi business: The case of uber. *The American Economic Review, 106*(5), 177-182. doi:10.1257/aer.p20161002

Crawford, V. P., & Meng, J. (2011). New york city cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *The American Economic Review, 101*(5), 1912-1932. doi:10.1257/aer.101.5.1912

De Vany, A. S. (1975). Capacity utilization under alternative regulatory restraints: An analysis of taxi markets. *Journal of Political Economy, 83*(1), 83-94. doi:10.1086/260307

Domencich, T. A., & McFadden, D. L. (1975). *Urban travel demand: A behavioral analysis*. Amsterdam: North-Holland Publishing Company.

Douglas, G. W. (1972). Price regulation and optimal service standards: The taxicab industry. *Journal of Transport Economics and Policy, 6*(2), 116-127.

Farber, H. S. (2008). Reference-dependent preferences and labor supply: The case of new york city taxi drivers. *The American Economic Review, 98*(3), 1069-1082. doi:10.1257/aer.98.3.1069

Farber, H. S. (2015). Why you Can't find a taxi in the rain and other labor supply lessons from cab drivers. *The Quarterly Journal of Economics, 130*(4), 1975-2026. doi:10.1093/qje/qjv026

FiveThirtyEight. (2015). Uber trip data from a freedom of information request to NYC's taxi & limousine commission. Retrieved from https://github.com/fivethirtyeight/uber-tlc-foil-response

Flores-Guri, D. (2003). An economic analysis of regulated taxicab markets. *Review of Industrial Organization, 23*(3-4), 255-266. doi:10.1023/b:reio.0000031368.93775.0a

Hall, J. V., & Krueger, A. B. (2016). *An analysis of the labor market for uber's driver-partners in the united states.* NBER Working Paper 22843.

Hamidi, A., Devineni, N., Booth, J. F., Hosten, A., Ferraro, R. R., & Khanbilvardi, R. (2017). Classifying urban rainfall extremes using weather radar data: An application to the greater new york area. *Journal of Hydrometeorology, 18*(3), 611-623. doi:10.1175/JHM-D-16-0193.1

Jackson, C. K., & Schneider, H. S. (2011). Do social connections reduce moral hazard? evidence from the new york city taxi industry. *American Economic Journal: Applied Economics, 3*(3), 244-267. doi:10.1257/app.3.3.244

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics, 3*(4), 303-328. doi:10.1016/0047-2727(74)90003-6

Merrill, J. B., & Coote, A. (2017, 27 July 2017). Blame uber for congestion in manhattan? not so fast. *New York Times* Retrieved from https://nyti.ms/1SJY3Ae

Peters, J. R., Shim, H. S., & Kress, M. E. (2011). Disaggregate multimodal travel demand modeling based on road pricing and access to transit. *Transportation Research Record: Journal of the Transportation Research Board, 2263*, 57-65. doi:10.3141/2263-07

Schneider, H. (2010). Moral hazard in leasing contracts: Evidence from the new york city taxi industry. *Journal of Law and Economics, 53*(4), 783-805. doi:10.1086/652423

TLC. (2010). *Old rule book.* ().The New York City Taxi & Limousine Commission. Retrieved from http://www.nyc.gov/html/tlc/html/about/tlc_old_rules.shtml

TLC. (2013). Background on the boro taxi program. Retrieved from http://www.nyc.gov/html/tlc/html/passenger/shl_passenger_background.shtml

TLC. (2016). *2016 TLC factbook.* ().The New York City Taxi & Limousine Commission. Retrieved from http://www.nyc.gov/html/tlc/downloads/pdf/2016_tlc_factbook.pdf