"2019 is the year Williamsburg dies"? Taxi Ridership Admist a New York Subway Slowdown

Jacob Dorn, Sylvia Klosin, Cameron Taylor, Weian Wang

3rd Annual University of Chicago Econometrics Game

April 16, 2017

Introduction

- In 2019 New York City will close the L line train between Brookyln and Manhattan for 15 months
  - It is one of three cross-town lines, serves 300,000 commuters each day
  - Big deal in press: " 2019 is the year Williamsburg dies" (New York Post)
- Inspiration: Study economic phenomena related to public transit and taxi services, using public transit disruptions
- Empirical design:
  - Use taxi data on March 11, 2015 (delays on L during evening RH)
  - Compare to other days of March to see changes in taxi usage
- Results:
  - Little effect on number of trips
  - Less bargaining on fares, maybe more cash payment
  - Suggests small welfare value of time for commuters

NY Subway System

March 11th, 2015

- On March 11, commuters suffered "massive delays" on the L train.[1]
- Over hour long delays were bilateral, meaning that delays occurred both going into and out of Manhattan and Brooklyn
- The delays persisted between 5 p.m. and 6:30 p.m., aligning right during the evening rush hour.

[1] https://www.dnainfo.com/new-york/20150311/murray-hill/commuters-suffer-massive-delays-service-changes-on-14-lines-wed-mta

Literature Review

We relate our paper to two main strands of literature.

1. Taxi Supply and Demand: There is a strong literature starting with Camerer et al. (1997) and Farber (2005, 2008). Our paper mostly relates to Buchholz (2016): proxies for supply and demand changes in taxi markets and welfare implications.
2. Our small results on bargaining relate to the literature on matching and bargaining which includes labor (e.g. Shimer 2005) and general matching markets (e.g. Samuelson 1992).

Theoretical Framework I

We setup three theoretical frameworks to better frame our empirical approaches and understanding of the economic problems.

- To understand the theoretical effect of a public transportation slow down
  - Adapt static model of taxi labor supply and demand from Buchholz (2016)
  - Two locations A and B, identical
  - Poisson arrival rate of customers λ_i determines the number of customers u_i.
  - v_i number of taxis
  - Matching function $m(u_i, v_i)$
  - Subway slowdown in A: raising the cost of public transportation, and increasing λ_A to λ'_A.
  - Holding fixed the number of taxis, this will increase the number of matches, also shifts in taxi numbers in equilibrium.

Theoretical Framework II

- To understand bargaining issues and public transit delays, we think of the removal of public transit as an "outside option", which may cause market agents to bargain more.
- To understand welfare
  - Commuters have inelastic demand for going to their "destination".
  - Each mode of transportation has a time cost and a monetary cost.
  - Assume public transit delays only affect the time cost of public transit methods.
  - We observe aggregate deviations $\Rightarrow$ can bound time valuation for marginal consumers.

Data

- Dataset was Yellow Cab trips in March 2015.
- Excluded all trips with nonpositive fare amounts, passenger counts and distances and all trips exceed 5 hours.
- Mapped pickups and dropoffs longitude and latitude coordinates to census tracts and Boroughs.
- Compared rush hour taxi activity on March 4, 11, and 18.
  - Morning rush: 6-9 AM
  - Afternoon rush: 4-7 PM
- March 11, 2015 data set post-filtering had 13,250,205 total rides.

Added Data

- Downloaded locations of all train stops in New York
- Found distance from L stations using the lat and long points of pick up and drop off location
- "Near" L train: Within 0.5 mi of station

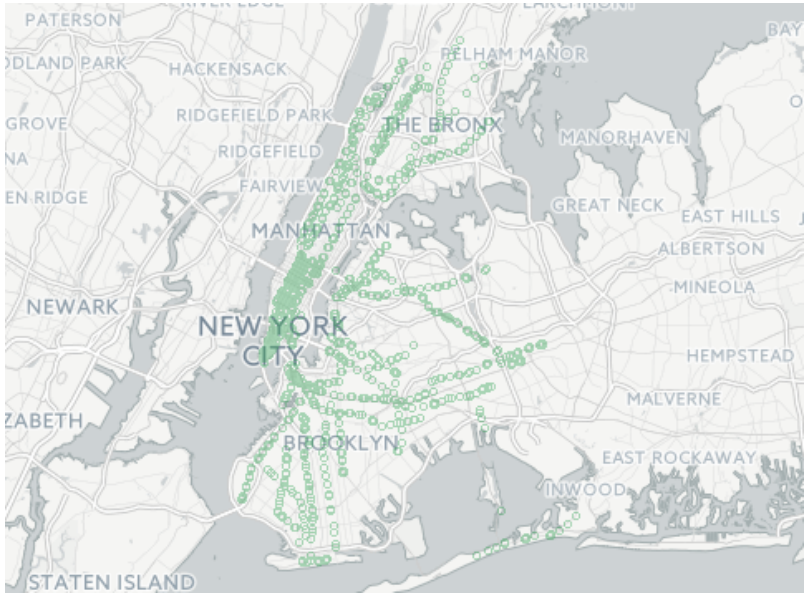


Figure 1: Location of NY subway stops

Data Visualization

- Kernel density estimation to visualize distribution of taxi volume across Manhattan and Brooklyn
- Compared March 4 and March 11 afternoon rush hour taxi activity
- Restricted visualization also to pickups and dropoffs in Brooklyn

Data Visualization

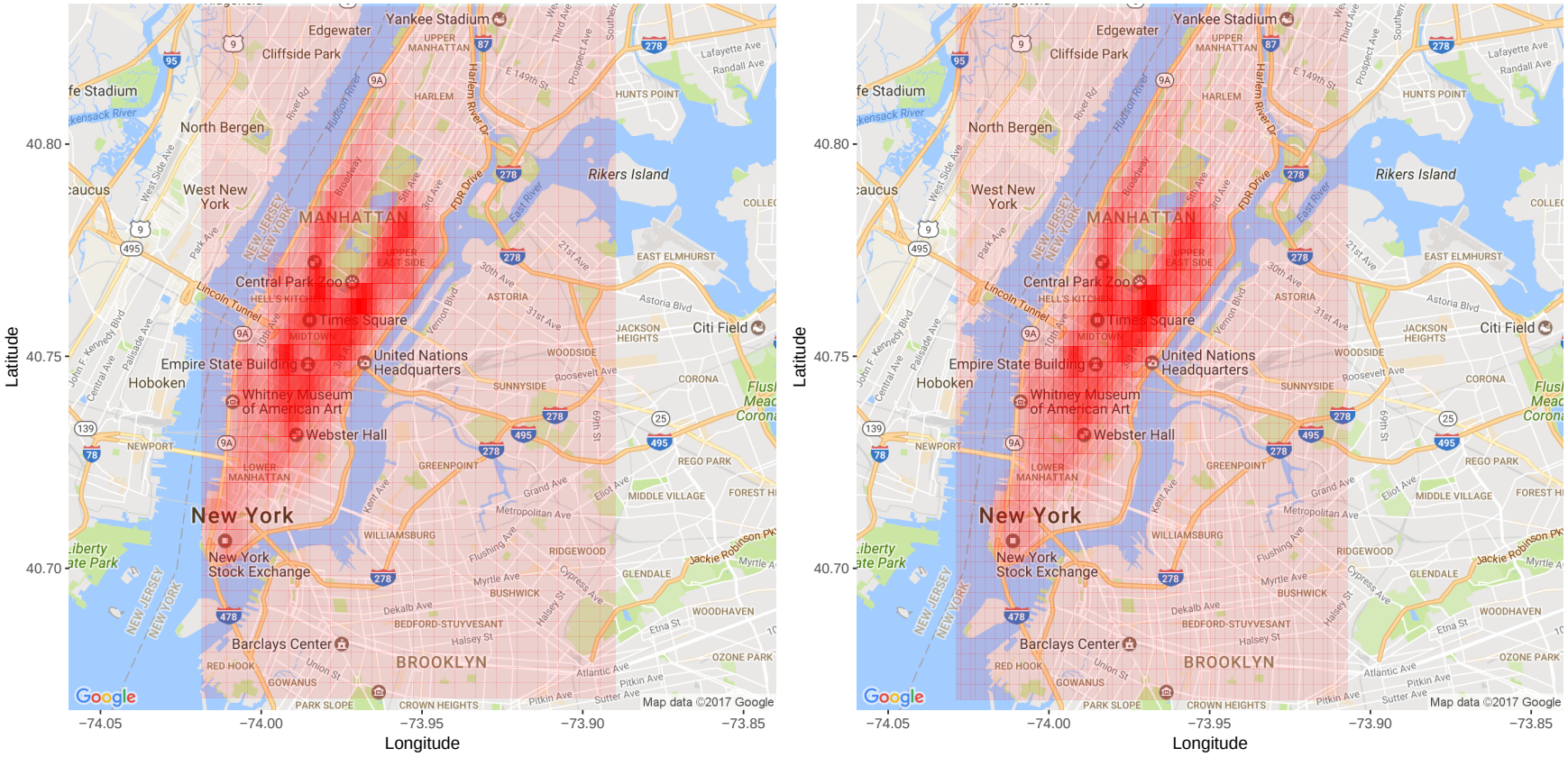


Figure 2: Nonparametric density estimate equivalent to Figure 3, but for pick-ups. The majority of the rides are concentrated in central to lower Manhattan, with little discernable difference in taxi activity concentration near the L train lines.

Data Visualization

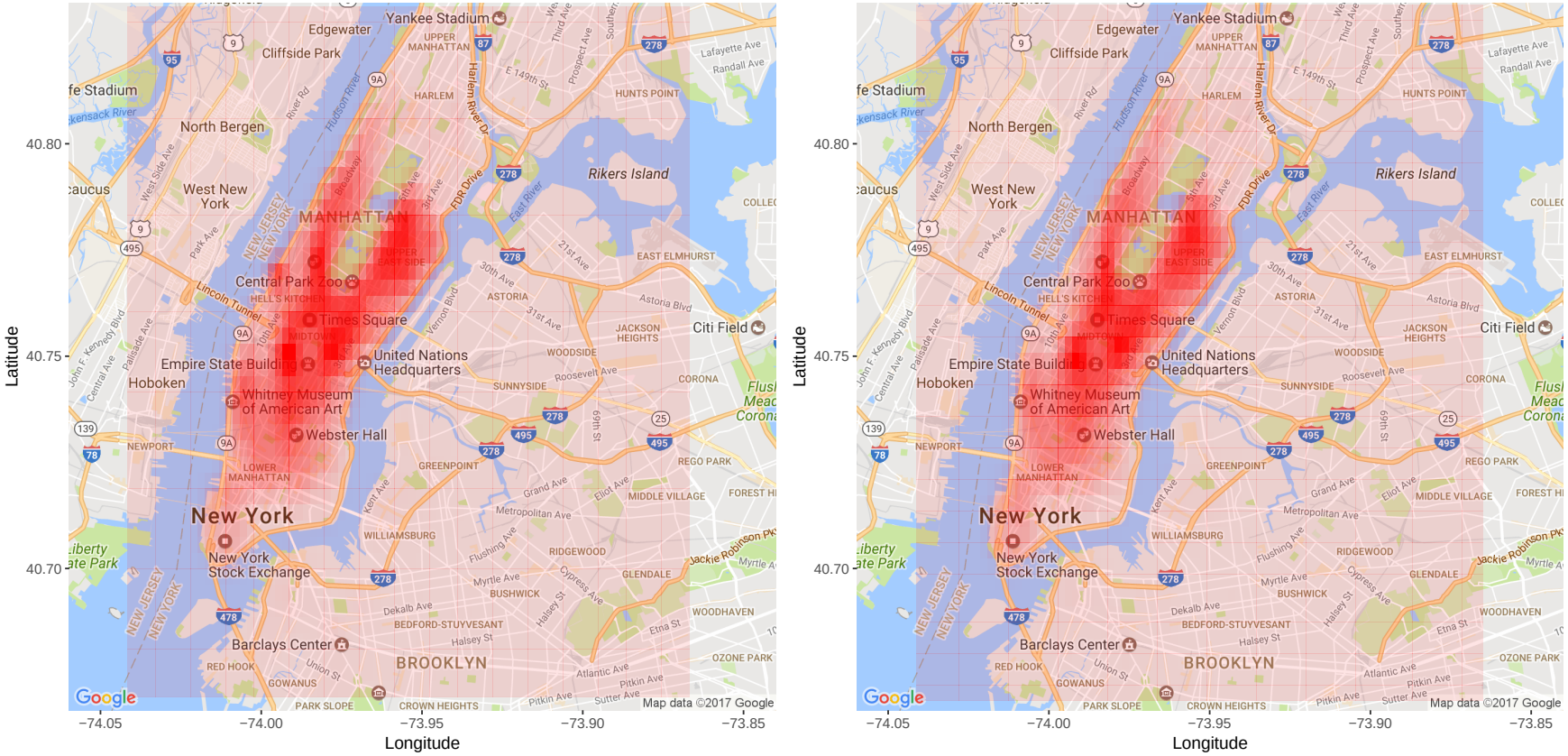


Figure 3: Nonparametric density (normal kernel, binwidth via normal score rule) plots dropoffs during the evening rush hours of March 4 (left) and March 11 (right).

Data Visualization

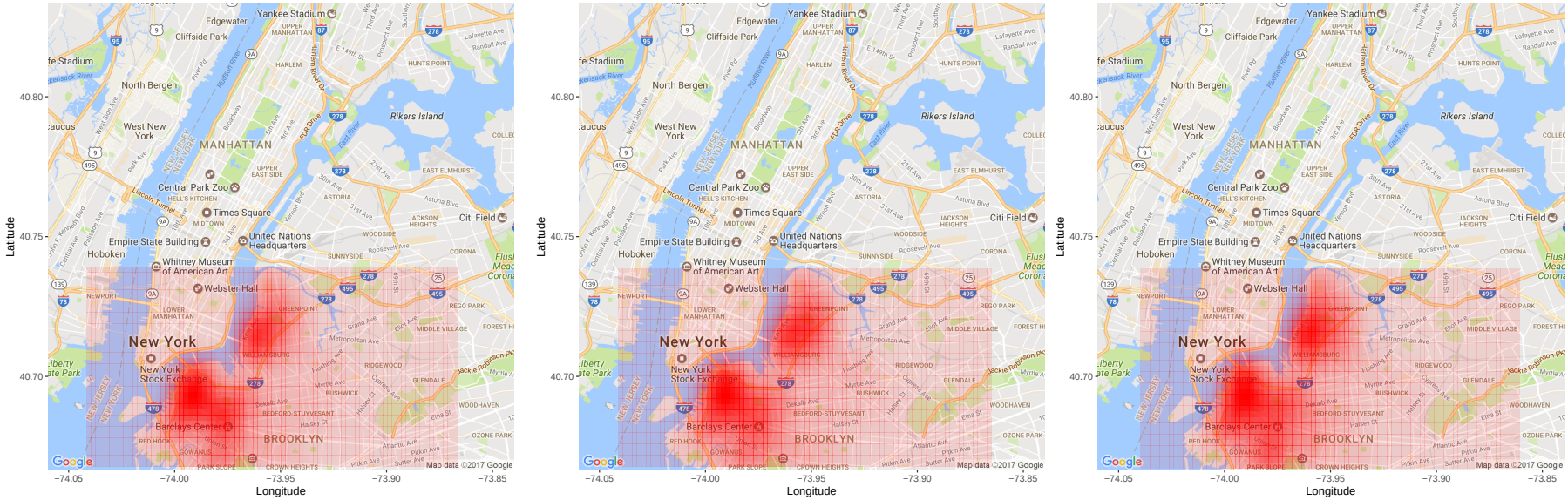


Figure 4: Nonparametric density estimation(normal kernel, binwidth via normal score rule) plots dropoffs only in Brooklyn during the evening rush hours of March 4, March 11, and March 18.

NYC-Level Analysis: March 11 vs 4

- Analysis without considering confounding variables (esp. general calendar)
- Comparing day statistics overall

Variable	Nominal Diff (11 vs 4)	Test	P-val
Trip distance (Mi)	+0.01	T test	≈ 0
Reported bargaining	-0.0001	T test	0.32
Cash payment	+ 0.01	T test	1.7e-09
Pickups near L	+0.001	T test	0.94
Dropoffs near L	-0.003	T test	0.14
Pickups (afternoon RH)	0.981	Poisson	0.001
Pickups (morning RH)	0.956	Poisson	≈ 0

Panel: Motivation

- Individual unit of analysis: A taxi ride
- Question: Are riders behaving like there are many more people on the streets?
  - Can't see Uber
- Another question: Are those picked up up or dropped off near an L stop likely to trace the path?
- How to control for time?
  - Day of month FEs
  - Week (days 1-7) and day of week (Friday) FEs

Panel: Day of Month Control, Riders by L Route Likeness

	Dependent variable:			
	near_L_pick	near_L_drop	bargain_pay	cash_pay
evening_rh_11	0.001	−0.003**	−0.0001	0.001
	(0.002)	(0.002)	(0.0002)	(0.002)
near_L_dropoff	0.124***		−0.001***	−0.060***
	(0.0003)		(0.00003)	(0.0004)
near_L_pickup		0.112***	−0.001***	−0.059***
		(0.0003)	(0.00003)	(0.0004)
evening_rh_11:near_L_dropoff	0.003			
	(0.004)			
evening_rh_11:near_L_pickup		0.016***		
		(0.004)		
Observations	13,250,205	13,250,205	13,250,205	13,250,205
R^2	0.017	0.017	0.0003	0.005
Adjusted R^2	0.017	0.017	0.0003	0.005
Note:				*p<0.1; **p<0.05; ***p<0.01

Panel: Day of Month Control, Regardless of Similarity to L Route

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
evening_rh_11	-0.032^{***}	-0.0001	-0.001^{***}	0.004^{*}
	(0.001)	(0.001)	(0.0002)	(0.002)
near_L_pickup			-0.001^{***}	-0.059^{***}
			(0.00003)	(0.0004)
near_L_dropoff			-0.001^{***}	-0.062^{***}
			(0.00003)	(0.0004)
Observations	12,654,756	12,654,756	12,654,756	12,654,756
R^2	0.004	0.004	0.0003	0.011
Adjusted R^2	0.004	0.004	0.0003	0.011
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Panel: Week-Day_of_Week Ctrl, Riders by L Route Likeness

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
evening_rh_11	0.002	−0.005***	−0.0001	−0.012***
	(0.002)	(0.001)	(0.0002)	(0.002)
near_L_dropoff	0.124***		−0.001***	−0.060***
	(0.0003)		(0.00003)	(0.0004)
near_L_pickup		0.112***	−0.001***	−0.059***
		(0.0003)	(0.00003)	(0.0004)
evening_rh_11:near_L_dropoff	0.003			
	(0.004)			
evening_rh_11:near_L_pickup		0.016***		
		(0.004)		
Observations	13,250,205	13,250,205	13,250,205	13,250,205
R^2	0.017	0.017	0.0003	0.005
Adjusted R^2	0.017	0.017	0.0003	0.005
Note:				*p<0.1; **p<0.05; ***p<0.01

Panel: Week-Day_of_Week Ctrl, Regardless of L Route

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
evening_rh_11	0.002	-0.003^{**}	-0.0001	-0.012^{***}
	(0.001)	(0.001)	(0.0002)	(0.002)
near_L_pickup			-0.001^{***}	-0.059^{***}
			(0.00003)	(0.0004)
near_L_dropoff			-0.001^{***}	-0.060^{***}
			(0.00003)	(0.0004)
Observations	13,250,205	13,250,205	13,250,205	13,250,205
R^2	0.003	0.003	0.0003	0.005
Adjusted R^2	0.003	0.003	0.0003	0.005
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Difference Estimation Approach

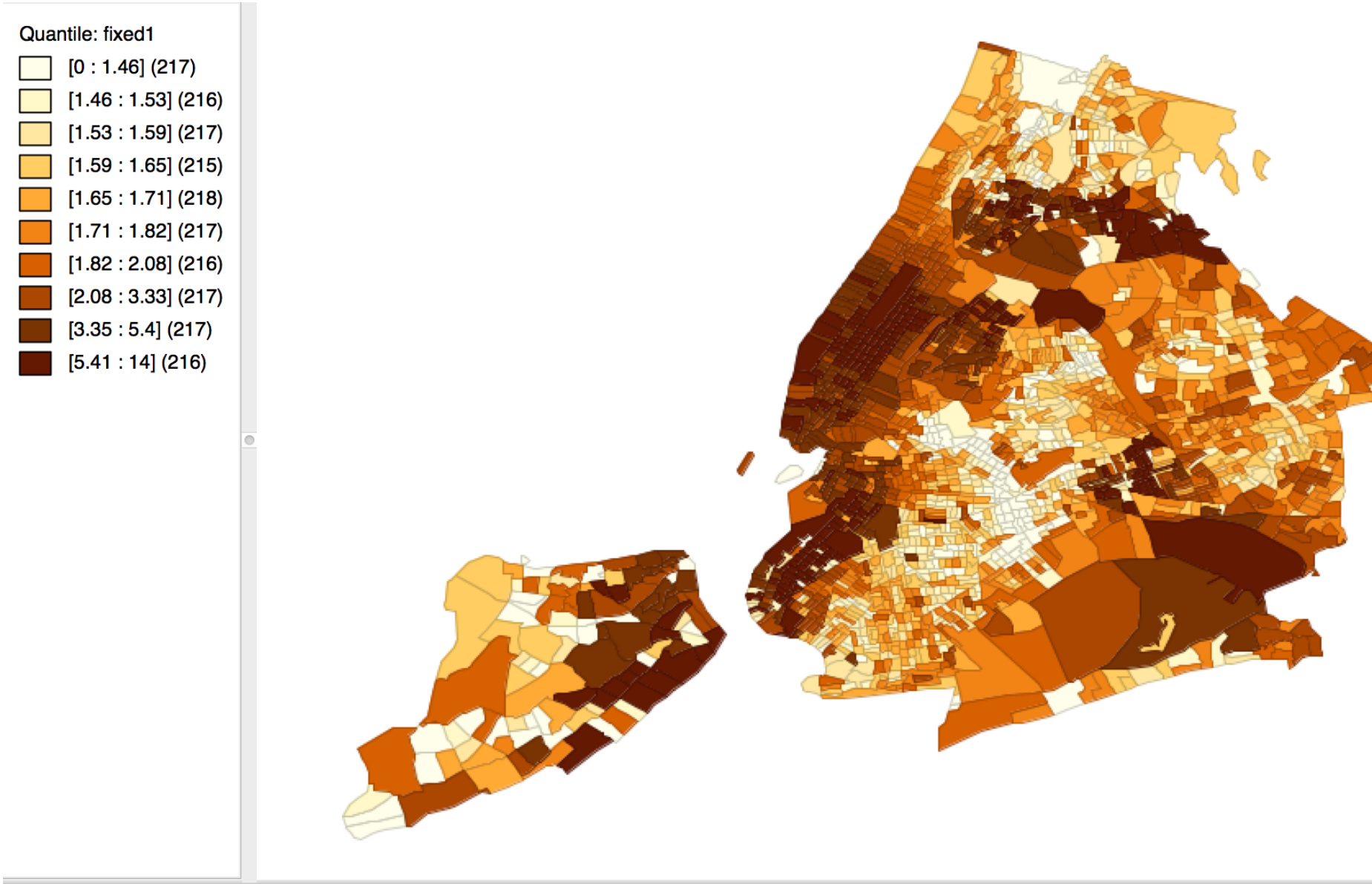
- Use a condensed version of the March 2015 data
- We wanted to estimate the effect the "treatment" of being around broken public transportation

$$\text{ihs(number trips)}_{ijkl} = \sum_i (\text{near L} \times \text{March 11th} \times \text{afternoon})_i + \sum_j \text{pickup tract}_j + \text{day of week}_k + \epsilon_{ijkl} \quad (1)$$

Difference Estimation Results

	Dependent variable:	
	ihs(trips)	
	(1)	(2)
near_L_dropoff	0.191***	
	(0.002)	
near_L_pickup		0.068***
		(0.003)
afternoon_hour	−0.256***	−0.207***
	(0.001)	(0.001)
March11	0.017***	0.012***
	(0.003)	(0.003)
near_L_dropoffTRUE:afternoon_hourTRUE:March11	−0.033**	
	(0.016)	
near_L_pickupTRUE:afternoon_hourTRUE:March11		0.020
		(0.016)
Observations	3,230,594	3,229,279
R^2	0.184	0.101
Adjusted R^2	0.184	0.100
Residual Std. Error	0.788 (df = 3229248)	0.828 (df = 3227995)
Note:		*p<0.1; **p<0.05; ***p<0.01

Difference Estimation Census Tract Events



Conclusion

- We used public transit delays to study shifts in behavior in the NYC taxi market.
- Value of time may not be too high among commuters, and no significant change in bargaining results.
- Our time constraint was definitely binding!
  - Demographic differences
  - There are many delays in New York public transit system, so we would liked to have increase our sample.
  - We would also like to extend to looking at the effects of a train fare increase.[2]

[2] There was a \$0.25 increase from \$2.50 to \$2.75 on March 23, 2015.