“2019 is the year Williamsburg dies”?*

Taxi Ridership Amidst a New York Subway Slowdown

Team Number 3

April 15, 2017

Abstract

This paper uses New York City yellow cab taxi data to estimate changes in the New York City transit market in response to unexpected transit delays. We exploit exogenous changes in delays for public transit using difference estimation techniques to estimate changes in matching rates for taxis, bargaining power shifts, and potential welfare losses for New York city commuters. We find small changes in response to public transit delays, suggesting that people value commuting time at a relatively low value. We also find that bargaining does not increase when delays happen, corroborating the claim that commuters may not value time too highly.

*http://nypost.com/2016/07/25/the-l-train-will-shut-down-for-18-months-in-2019/

Contents

1 Introduction

New York City plans to close a line between Brooklyn and Manhattan for 15 months in 2019. The closure, on the "L" subway line, could slow down 300,000 commuters each day by closing one third of the cross-town lines.[1] The New York Post has gone as far as arguing it would cause the economic death of the Williamsburg neighborhood, where visitors and residents primarily use the L to travel to and from Manhattan (hence the title of our paper).[2]

Policy makers are more sanguine. They argue that "it is better to have shorter duration of pain than a longer more unstable process."[3] In this paper, we consider whether the short-term response to a slowdown suggest much pain.

To tackle this question and understand other interesting economic phenomena related to changes in public transit access, we look at a measure of the spillover effects from such slowdowns. We look at data for New York taxi pickups and dropoffs, comparing aggregate variables on a day of major delays to days with normal transit operations.

To do this, we employ a difference comparison empirical design between taxi usage, both near L stations and across New York City more generally, within March 2015. In particular, we analyze the response to significant delays on the L train on Wednesday March 11, 2015, finding that few consumers switch to taxi ridership in response. This may suggest that subway riders do not value their time very much; The mean Wednesday trip from Manhattan to Brooklyn – the affected bridge's crossing – only incurred a fare of between \$17 and \$19, regardless of day and afternoon rush hour status.

On March 11, commuters suffered "massive delays" on the L train.[4] Many commuters cited rides that took more than an hour over the usual time. It is important to note that the delays were bilateral, meaning that delays occurred both going into and out of Manhattan and Brooklyn, despite the initial problem only occurring on the side heading into Manhattan. The delays persisted between 5 p.m. and 6:30 p.m., aligning right during the evening rush hour.

We also explore the relationship between public transit delays and the propensity for taxi fares to be determined by bargaining. We again use a differencing strategy to address this phenomenon. We find evidence that, despite potentially increased traffic, there is no increase in reported bargaining, but there is an increase in cash payment. We also find little effect on the number of trips; If anything, our difference-in-difference analysis suggest that taxi rides went slightly down. However, some of this may be caused by a confounding basketball game (see Section 3.1). We note some potential other delays ripe for similar studies in our conclusion.

The rest of our paper is organized as follows. Section 2 summarizes the relevant literature and

[1] http://www.amny.com/transit/l-train-shutdown-explained-facts-figures-proposals-and-more-1.11761564

[2] http://nypost.com/2016/07/25/the-l-train-will-shut-down-for-18-months-in-2019/

[3] http://www.slate.com/blogs/moneybox/2016/07/25/the_l_train_shutdown_will_be_a_nightmare_but_also_a_blessing_in_the_long.html

[4] https://www.dnainfo.com/new-york/20150311/murray-hill/commuters-suffer-massive-delays-service-changes-c

places the contributions of our paper in the context of other related work. Section 3 describes and summarizes the data. Section 4 provides a basic theoretical framework that we work in and provides some intuition for what we would expect in our results. Section 5 gives our results. Section 6 concludes.

2 Literature Review

There is a growing literature understanding economic phenomena that may or may not be exclusive to the market for taxis using New York City taxi data. The taxi market is interesting to study in its own right for welfare and search friction reasons (Buchholz 2016). It also provides a useful data source for understanding more broad economic phenomena like learning by doing (Haggag et al. 2017).

Our estimates of shifts in taxi matching contribute to the expanding literature on taxi labor supply and demand. This literature started with classic papers by Camerer et al. (1997) and Farber (2005, 2008) but has recently been expanded on by Buchholz (2016).

Our finding that consumers did not, by and large, shift to taxis contributes to the literature on social welfare in transportation and taxi markets. In particular, Buchholz (2016) estimates the differences between a decentralized taxi search equilibrium and the social planner's optimal equilibrium. Our finding adds nuance to these results by considering optimal social planning of alternate modes of transportation: i.e. the location and subsidy amount of public transit stops.

Finally, our findings related to bargaining propensity contribute to the literature on matching and bargaining. This literature is strong in labor search (e.g. Shimer 2005) but also in general matching markets (e.g. Samuelson 1992).

3 Data

For these analysis we used several different datasets. For one, we used publicly available data on taxi services in New York City. We looked at the Yellow Cab data for the month of March 2015. We excluded Green Cab data, because these cabs are ”Boro taxis”, taxis that are only permitted to pick up in outer boroughs and parts of Manhattan above East 96th Street and West 110th Streets. We find that, in the case of the yellow taxis, the overwhelming majority of pickups happen in those non-Boro zones.

We pre-processed and cleaned the data for analysis to include only points with positive fare amount, trip distance variables figures, and passenger counts. As other papers have done, for all points with pickup times that come after dropoff times, we switched the pickup and dropoff time values for that point. We also computed trip time for each ride and excluded all trips that had trip times exceeding 5 hours.

We further excluded data with pickup and dropoff longitude less than -20 degrees and pickup and dropoff latitude greater than 20 degrees. This loosely restricted our dataset to more or less the trips with pickup and drop-off in New York City (i.e. excluded the rare longer trip outside of New York City to, for example, Connecticut or Pennsylvania).

Though we did use the full March 2015 data for our conclusions, to avoid inference which assigns weekday affects to other dates we often focus on the Yellow Cab data for the four Wednesdays: March 4, 11, 18, and 25. We looked at these particular dates to identify changes in trip volume, distance, and times during significant delays. We compared taxi activity on these four dates during particular high-traffic times: the morning rush hour (6 A.M. to 9 A.M.) and the afternoon rush hour (4 P.M. to 7 P.M.). We did not find substantial differences in mean trip distances (see Table 1), so it is reasonable to think that the subway problems would drive any changes fully.

We also used data from New Yorks's MTA on the location of train stations throughout cities. A map of these stations can be found in Figure 1 below. For all of the taxi rides in our subsample we calculated how far the pick-ups and drop-off were from the closest L-train stop, since we focused on this train line in our analysis. Using latitude and longitude we also found which Census tracts of these starting and locations. We also use these tracts to infer borough.

Though in this paper we did not have time to merge in ACS tract demographic information, in future papers we are interested in examining key population factors that we suspect impact taxi ride behavior like average age and income.

3.1 Kernel Density Estimation: Heat Maps

We created heat maps of the taxi pick-ups and drop-offs for the three Wednesdays of March 2015 to have comparisons. We generated heat maps to visualize our density estimates of taxi volumes. To generate these estimates, we used the multivariate kernel density estimator:

$$\hat{f}(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{p}\frac{1}{h_j}K\left(\frac{\|\mathbf{x}-\mathbf{x_i}\|}{h_j}\right)$$

where $p = 2$, $(\mathbf{x_1}, \cdots, \mathbf{x_n})$ is the sample of longitude and latitude coordinate pairs, $K(\cdot)$ is the kernel, and $\mathbf{h} = (h_1, h_2)^T$ is the smoothing parameter, with different bandwidths for longitude and latitude. For our estimates, we selected the default Gaussian kernel. To determine optimal bandwidth, we performed cross-validation to find the optimal $\mathbf{h}$. Nonparametric density estimates (with normal kernal bandwidth unfortunately not estimated via cross-validation due to computation time constraints) are available below.

We include the heatmaps of Yellow Cab taxi pickups and dropoffs between the afternoon rush hour on March 4, March 11 and March 18. We limited the visualization to pickups within Manhattan and Brooklyn boroughs in Figure 3. Most trips during this rush hour period end in Manhattan, and though it may be difficult to discern, there is some taxi activity in other Brooklyn as well.

Figure 4 compares density of dropoffs only in Brooklyn on March 4 and March 11, respectively. What we observe is larger density of taxi traffic in upper Williamsburg near L train line, and in particular a greater concentration, though faint, of taxi activity near the L stops. During the evenings of March 4 and 11, there were large basketball games in Brooklyn, which perhaps attracted some of the crowds near the Barclays Center in SE Brooklyn. However, the March 11 game was only an Atlantic 10 college tournament play-in round featuring lowly Fordham, George Mason, Duquense, and Saint Louis. The March 4 game, on the other hand, featured the Nets and the Hornets, and 16,691 attendees who far outnumbered the 5,985 for the college games.[5], both large relative to the 61,888 March 11 afternoon rush hour trips overall. Therefore, we expect this factor likely underestimates the number of trips on March 11.

Figure 6 compares pickups in Manhattan and Brooklyn on March 4 and March 11. There is no discernable difference in the volume of taxi density. We do observe that a majority of the taxi volume occurs in central to upper Manhattan. The Green Cab taxis could not have picked up where the largest concentration of pickups occurred for the Yellow Cabs.

4 Theoretical Frameworks

4.1 Dynamic Taxi Market

We consider a simple model adapted from Buchholz (2016) and essentially ignore dynamics, comparing similar cross-sections of time across different days and events. We split New York into two locations for taxi drivers to search at, A and B. We assume that A experiences the public transit delays, and B is some external market for taxis.

For location i, we suppose that there is an arrival rate of customers for λ_i There is some number of customers u_i in a market for taxis that is distributed according to the λ_i parameter. There are a number of vacant cabs at any time v_i in an area. Matches are formed according to some matching function $m(u_i, v_i)$, which is itself a Poisson arrival rate of matches.

We consider a shift in public transit to the following: increasing λ_i exogenously. This increase, holding fixed the supply of taxis, should increase the matching rate.

We model the driver's decision simply based on a *directed matching framework*. Driver's decide to go to either location. We assume the number/supply of driver's is fixed and their supply decision is simply where to search for customers.

Assume that A and B are totally identical and that equilibrium is pinned down by the matching function. Suppose that location A has a public transit delay. This will increase the cost of public transit and will induce an increase in λ_A and subsequently the expected u_A. Holding fixed v_A and v_B this will increase the matching rate in A, which induces a strict increase in value for taxis in A. Thus, equilibrium will shift where taxis enter A until again, given the shift in B, the matching

[5] http://www.espn.com/mens-college-basketball/game?gameId=400766863

rates in the two sectors for taxis are equal (then by homogeneity of the two sectors, the values will be equal).

Formally the equilibrium condition will be

$$\frac{\mathbb{E}\Big[m(u_A, v_A)|\lambda_A\Big]}{v_A} = \frac{\mathbb{E}\Big[m(u_B, v_B)|\lambda_B\Big]}{v_B}$$

where taxis shift between A and B to balance out these taxi-specific matching rates.

In the taxi data, we will be interested in estimating the shift from $m_A = \mathbb{E}\Big[m(u_A, v_A)|\lambda_A\Big]$ to $m'_A = \mathbb{E}\Big[m(u_A, v_A)|\lambda'_A\Big]$, after the exogeneous shock of public transit delay $\lambda'_A > \lambda_A$. To do this, we will exploit the poisson properties of the matching function and the counts to estimate the poisson arrival rate of matches.

4.2 Bargaining and Competition

In this paper we also look at the reduction in the value of public transportation increases the propensity for customers and taxi drivers to bargain for their fares. The idea here is that with the removal of public transportation as an option, the endogenous "value" of a taxi goes up and the threat point of the commuter lowers. This leads to potential for shifting bargaining power in a simple bargaining game. One can refer to standard labor search matching models as a reference for how one might model this (see Shimer et al. 2005 for a survey of these methods).

4.3 Welfare

To consider inference of welfare effects via taxi demand, we suppose that commuters have inelastic demand for going to their "destination" (work, home or other) – i.e. they will get where they need to go somehow. Commuters have a few options, which include but are certainly not limited to public transportation and taxis.

Consider any single commuter. We simplify them into facing a binary decision problem between choosing to take a taxi τ or taking public transportation ρ. Their choice depends on a few critical values including the monetary costs and time components of each method.

In particular, commuters choose public transportation when $c_\rho \leq c_\tau$ where

$$c_\rho = v_T T_\rho + m_\rho$$

$$c_\tau = v_T T_\tau + m_\tau$$

where v_T is the value of time to the commuter, T_i is the time it requires to take some method and m_i is the monetary component. We note there that v_T should depend on individuals own wage among other things, T_i will depend on the distance the individual needs to travel, and m_i will depend on distance (although it will likely depend on distance more for taxi due to the differing pricing scheme).

We observe aggregate extensive margin deviations in the form of total people taking taxis and their deviation from the norm. Thus, we observe a marginal group of people who switch from public transport to taxis. During a subway delay, we assume that the time and monetary costs for taxi travel remains the same, and that the monetary costs of public transportation remain the same while there is a shift in time to $T'_\rho > T_\rho$. Then if $c'_\tau \leq c_\rho$ we can infer that

$$v_T \leq \frac{m_\rho - m_\tau}{T_\tau - T'_\rho}.$$

This framework will allow us to understand the welfare implications of our findings of the propensity for commuters to switch from public transit to taxis.

Since we find little effect on taxi demand, we infer that consumers have limited welfare loss – in pure monetary terms – from the train delays.

5 Results

In this section all models are estimated with the `lfe` package in R. The package uses the Method of Alternating Projections to estimate linear models with multiple group fixed effects using techniques outlined by Gaure (2013). This method is a generalization of the within estimator. Because of the large size of the dataset this new technique needed to be used in order to make estimation feasible. Including multiple levels of fixed effects allows us to control for unobserved heterogeneity specific to each group, which could otherwise preclude causal inference due to omitted variable biases (Gormley and Matsa 2014). Notably, throughout our paper we are able to control for census tract and day of the week. For this package the standard errors will be systematically overestimated, which means that we have more confidence when we find statistically significant results.

Our empirical methods center around using the public transit delays as a potential shock to demand for taxi use due to increased time-related costs. Inspired by the policy aspect of our paper, we look at delays specific to the L train. Our main method will be to compare taxi usage on the day of interest to other days in which delays were non-existent or less significant.

5.1 March 11th 2015 Slowdown

On March 11, commuter suffered "massive delays" on the L train.[6] Many commuters cited rides that took more than an hour over the usual time. It is important to note that the delays were bilateral, meaning that delays occurred both going into and out of Manhattan and Brooklyn.

[6]https://www.dnainfo.com/new-york/20150311/murray-hill/commuters-suffer-massive-delays-service-changes-c

5.2 Overall Comparisons

Comparing all taxi trips on the afternoons of March 11 and March 4, via naive t test, we find no statistically-significant difference of trip distance (p = 0.32) nor reported bargaining (p = 0.32). We do find a statistically significant difference in the rate of cash payment (p = 1.7e-09, difference of 0.016 percentage points). There was no statistically-significant difference in whether pickups happened near an L stop (p = 0.9415), but perhaps a reduction in dropoffs near an L stop (p = 0.1357, effect of 0.0028).

The non significant difference in reported bargaining further corroborates the suggested conclusion that commuters may not value time as highly as anticipated. Theory would suggest that riders in taxis with high time values may try to extract some of the surplus that taxis get from providing a new service, but because this is not the case, the value of high time must not be high enough for costs of bargaining to be rationalized.

We also consider a Poisson model of pickups and dropoffs. We find that while afternoon rush hour trips were lower on March 11 relative to March 4 – with a p-value of 0.0009 – the best estimate of a 1.9% reduction in trip rate (0.9% reduction in trips overall) is less than the 4.6% reduction in morning rush hour trip rate (2.3% overall). This suggests that perhaps some individuals responded to the news of the afternoon L slowdown by switching to a cab, but the margins are very small.

5.3 Panel Analysis

One model considers whether taxi drivers were more likely to drive to or from locations near an L stop – located as a half-mile from an entrance – during the afternoon rush hour on March 11.

The results seem to depend on the method of control. All of our results control for fixed effects for general morning and afternoon rush hour as well as second-order polynomials of trip distance. However, they differ based on the method of controlling for date within March.

Tables 2 and 3 control for each day in March. This is to say that evening effects are measured relative to the rest of the day on March 11. These suggest that taxis were perhaps less likely to pick up near L stops on that date (Table 3), but, while they were less likely to drop off near an L stop on that evening, the effect was reversed when picking up near the subway (Table 2). They all agree that those picked up or dropped off near an L stop are less likely to bargain or pay cash – we are unsure if cash payments can serve as a taxless alternative to on-the-books bargaining – but in the case of Table 3, that those paying that evening perhaps were more likely to pay with cash.

Once we control for week and day-of-week (i.e. Tuesday fixed effect and second-week-of-March fixed effect), many of the effects flip. This specifications, summarized in Tables 4 and 5) aimed to reduce overfitting of the implied alternative but perhaps at the cost of too coarse a construction, finds that taxis were, if anything, more likely to pick up near an L stop. The patterns for dropoffs near L stops are similar. However, there is now perhaps evidence that individuals are *less* likely to pay in cash during that evening

5.4 Difference estimation: Census Tract Fixed Effects

We also created a condensed version of the March 2015 dataset in order to calculate Census tract level "effect" fixed effects. Our condensed dataset calculated census tract-day-time level measures of total rides, whether the tract was within .5 miles of a L train stop, average fees, etc. Using this data we estimate equation (1) below. Note that to save space we did not write the baselines of the three-way interaction (i) below, but the baseline effects are included in the actual regression.

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

Our outcome variable of interest is an inverse hyperbolic sine transformed version of the number of trips. We chose this transformation over the traditional log transform because some cells received 0 trips, and log is not defined at 0 while inverse hyperbolic sine is. Even adding ϵ to the variable inflated coefficients due to the difference between $\log(0+\epsilon)$ and $\log(1+\epsilon)$. Inverse hyperbolic sine is a convenient transform because it is approximately equal to $\log(2y_i)$ or $\log(2)+\log(y_i)$, and so it can be interpreted in exactly the same way as a standard logarithmic dependent variable (Burbidge, 1988). Our results can be interpreted as percent changes.

We also ran versions with track dropoff instead of tract pickup on the left hand side of the model. The differences we are comparing are difference across days: March 11th vs other days in the month, comparing across whether is the afternoon or morning, and whether the tract is close to the L-train or not. By calculating differences across these groups we are trying to estimate the "treatment" effect of facing a broken public train on taxi rides. By comparing averages across these groups we hope to lessen the effects of extraneous factors and selection bias. In the future we would like to see how good of controls our groups are for our treatment group, by looking at census characteristics is more detail for example, so we could further investigate concerns with the direction of causality and omitted variable bias.

The results for this section are presented in Table 6. Model 1 has results using pickups, Model 2 is results using drop offs, and Model 3 uses both pick ups and drop offs. We find that the coefficients on the triple interaction are very small in magnitude. The change in the number of taxi rides is only about 3% in areas near the L-train during the afternoon of March 11. In future analysis we will look at more events and see if the treatment coefficient is always small, or only a relic of our particular example.

We also created a map of the census tract fixed effects, which is presented in Figure 2. We find that tracts with large fixed effects are also the areas that stand out in the heat-maps as high intensity areas, discussed in the next section. This makes sense as both methods are determining which areas reactive a lot of taxi activity. Later on we also ran a regression where we interacted our treatment variable with an indicator for tract to see if there were location treatment effects. The resulting map presents a similar picture, which may be due to the small treatment effects we observe.

6 Conclusion

Though the results presented in this paper start our analysis, given more time there are several other factors we would like to examine. For one, we focused on only on one train delay in this paper. Given that there are many other major train failures in New York over the time period provided in the raw dataset, we would like to broaden our scope. In particular, there are other, perhaps more impactful, L train disruptions. On July 25, 2016, L train service was entirely shut down due to NYPD activity. On October 6, 2016, L train service between the boroughs was disrupted for 8 hours after a fire. Furthermore, we would like to do an extension looking at the train fare increase that occurred on March 23, 2015. On this day fare increase from $2.50 to $2.75, and we would like to see if there were changes in people's taxi behavior.

In this paper we specifically find that there is not a large change in the number of taxi trips around the L-line on March 11th 2015, a day of major L train delays. We ueconclude this do to a series of empirical work. By just looking at the heatmaps, we find it is difficult to find a difference in certain taxi behaviors. In the Panel Analysis we find that in afternoon of March 11th people were less likely to move to areas around L stops. In the difference estimation we find small effects of the L-train problems on the number of trips, though the effects are significant. We look forward to estimating the effects of other train problems, as we think it will enable us to obtain more general and useful results.

7 References

1. Buchholz, Nicholas. Spatial Equilibrium, Search Frictions and Efficient Regulation in the Taxi Industry. Working Paper. August 2016.

2. Burbidge, John, and Lonnie Magee and Leslie Robb. "Alternative Transformations to Handle Extreme Values of the Dependent Variable." *Journal of the American Statistical Association* 83.401(1988):123-127

3. Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. Labor supply of New York City cabdrivers: One day at a time. The Quarterly Journal of Economics. 1997.

4. Farber, Henry. Is Tomorrow Another Day? The Labor Supply of New York City Cabdrivers. The Journal of Political Economy. 2005.

5. Farber, Henry.s Reference-dependent preferences and labor supply: The case of New York City taxi drivers. The American Economic Review. 2008.

6. Gaure, Simen. "OLS with multiple high dimensional category variables." *Computational Statistics & Data Analysis* 66(2013):8–18

7. Haggag, Kareem, Brian McManus and Giovanni Paci. Learning by Driving: Productivity Improvements by New York City Taxi Drivers. American Economic Journal: Applied Economics. Vol 9, No 1. 2017.

8. Samuelson, Larry. Disagreement in Markets with Matching and Bargaining. The Review of Economic Studies. Vol. 59, No. 1. January 1992.

9. Shimer, Robert and Randall Wright. Search-Theoretic Models of the Labor Market: A Survey. Journal of Economic Literature. 2005.

8 Appendix

Table 1: Mean fare of trips from and to within a half-mile of an L train stop from Manhattan to Brooklyn on three days in April, based on whether or not they were during afternoon rush hour. While the mean fare did increase during the afternoon rush hour on April 11, it was increased by less than $1.50, and the mean trip time was increased by less than four minutes, relative to the rest of the day, so travelers did not have a good reason to adjust their expectations of taxi rides substantially.

afternoon_rh	day_of_month	Mean Fare	Mean Trip Time (Hours)
FALSE	4	17.15	0.31
TRUE	4	17.47	0.33
FALSE	11	17.80	0.33
TRUE	11	19.25	0.39
FALSE	18	18.36	0.35
TRUE	18	18.38	0.38
FALSE	25	18.07	0.34
TRUE	25	19.13	0.40

Table 2: Panel regressions for all taxi rides, with evening_rh_11 being measured relative to March 11 overall. Estimates reflect OLS (linear probability model) controlling for day, both morning and afternoon rush hour Fixed Effects, and a quadratic polynomial of trip distance. We do not find statistically significant effects on cash payments during the March 11 evening rush hour. While we do find a reduction in dropoffs near L stations during the evening rush hour, those departing from or near an L station were more likely to come close to tracing the subway's path, but only by one or two percentage points.

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
	(1)	(2)	(3)	(4)
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

Table 3: Panel regressions for all taxi rides, with evening_rh_11 being measured relative to March 11 overall. Estimates reflect OLS (linear probability model) controlling for day, morning and afternoon rush hour Fixed Effects, and a quadratic polynomial of trip distance. There is perhaps some weak evidence that riders during that evening were more likely to pay cash, and more likely to bargain, but all of these effects are small. There is better evidence that fewer rides began near L stations.

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
	(1)	(2)	(3)	(4)
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

Table 4: Relative to general patterns (also controlling for second-order polynomial of trip distance). Similarly to Table 2, we find that while March 11 evening taxi passengers were less likely to go to an L station, those departing from near one were more likely to do so. However, we no longer find any statistically significant effect on cash payment or bargaining.

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
	(1)	(2)	(3)	(4)
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

Table 5: Relative to general patterns (also controlling for second-order polynomial of trip distance). We now find evidence *against* cash payments and no effect on pickups near L stations.

	Dependent variable:			
	near_L_pickup	near_L_dropoff	bargain_pay	cash_pay
	(1)	(2)	(3)	(4)
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

Table 6: Difference estimation: Census Tract Fixed Effects

	Dependent variable:		
	ihs(trips)		
	(1)	(2)	(3)
near_L_dropoff	0.191^{***}		0.193^{***}
	(0.002)		(0.002)
near_L_pickup		0.068^{***}	0.046^{***}
		(0.003)	(0.002)
afternoon_hour	-0.256^{***}	-0.207^{***}	-0.299^{***}
	(0.001)	(0.001)	(0.001)
March11	0.017^{***}	0.012^{***}	0.003
	(0.003)	(0.003)	(0.003)
near_L_dropoffTRUE:afternoon_hour	-0.123^{***}		
	(0.003)		
near_L_dropoffTRUE:March11	0.100^{***}		
	(0.008)		
near_L_pickupTRUE:afternoon_hour		-0.083^{***}	-0.098^{***}
		(0.003)	(0.003)
near_L_pickupTRUE:March11		0.063^{***}	0.078^{***}
		(0.008)	(0.007)
afternoon_hourTRUE:March11	-0.011^{*}	0.009	−0.009
	(0.006)	(0.006)	(0.006)
near_L_dropoffTRUE:afternoon_hourTRUE:March11	-0.033^{**}		
	(0.016)		
afternoon_hourTRUE:near_L_dropoff			-0.117^{***}
			(0.003)
March11:near_L_dropoff			0.108^{***}
			(0.008)
near_L_pickupTRUE:afternoon_hourTRUE:March11		0.020	0.007
		(0.016)	(0.014)
afternoon_hourTRUE:March11:near_L_dropoff			-0.041^{***}
			(0.015)
Observations	3,230,594	3,229,279	3,218,195
R^2	0.184	0.101	0.269
Adjusted R^2	0.184	0.100	0.268
Residual Std. Error	0.788 (df = 3229248)	0.828 (df = 3227995)	0.747 (df = 3215580)
Note:			*p<0.1; **p<0.05; ***p<0.01

We control for day of the week and the particular census tract, but do not show these coefficients due to the large number of fixed effects.

Location of All Train Stations in New York

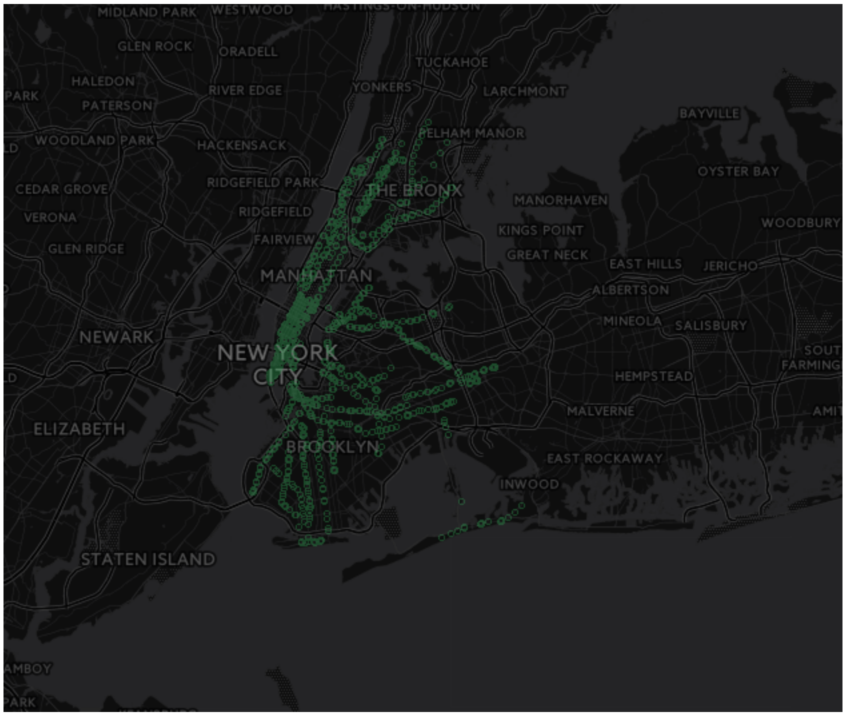

Figure 1: Location of NY subway stops.

Census Tract Fixed Effects

Figure 2: Fixed effects from transit regression.

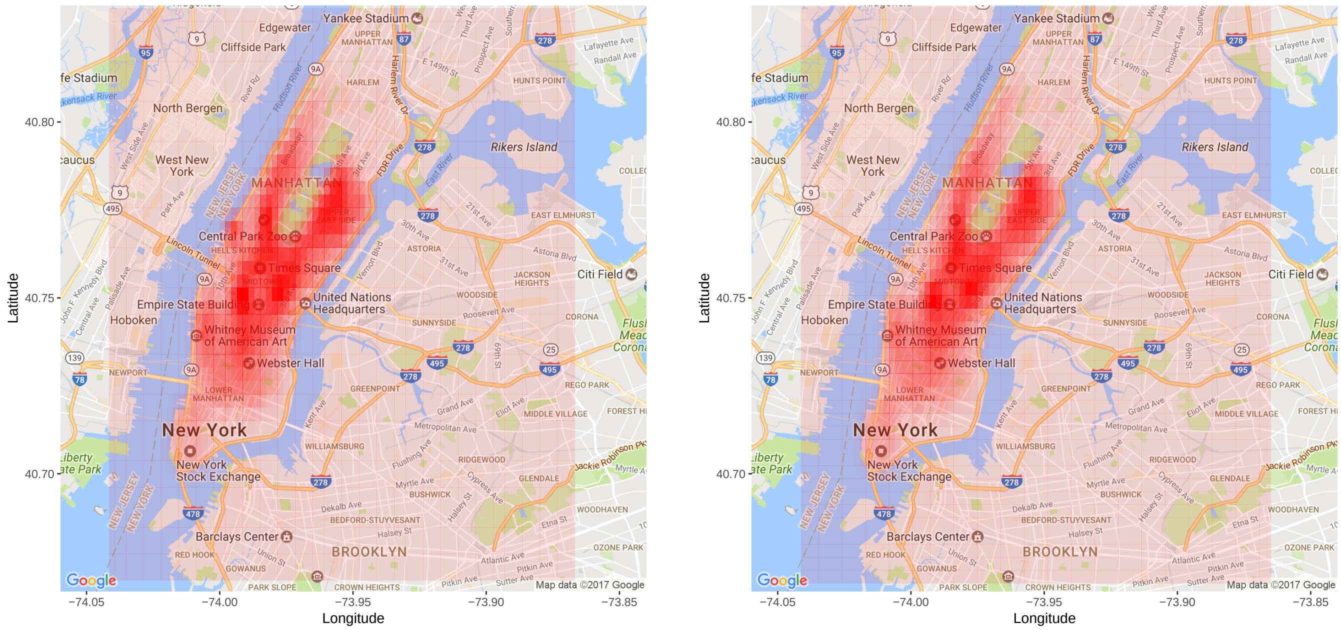

Figure 3: Nonparametric density (normal kernel, binwidth via normal score rule due to computational time limits) plots dropoffs during the evening rush hours of March 4 (left) and March 11 (right). Most trips do end in Manhattan, although, as is hard to see, there is some density elsewhere as well.

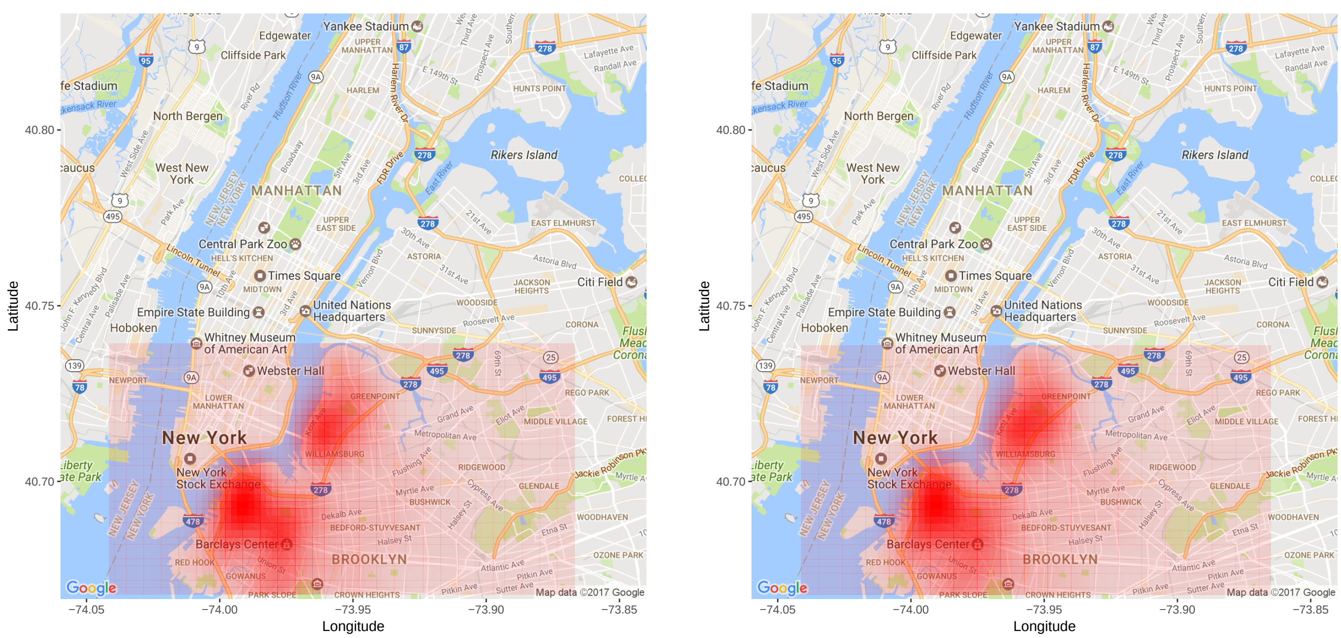

Figure 4: Nonparametric density estimation(normal kernel, binwidth via normal score rule) plots dropoffs only in Brooklyn during the evening rush hours of March 4 (Left) and March 11 (Right). The majority of taxi traffic is directed to two specific areas of Brooklyn, though the volume is more

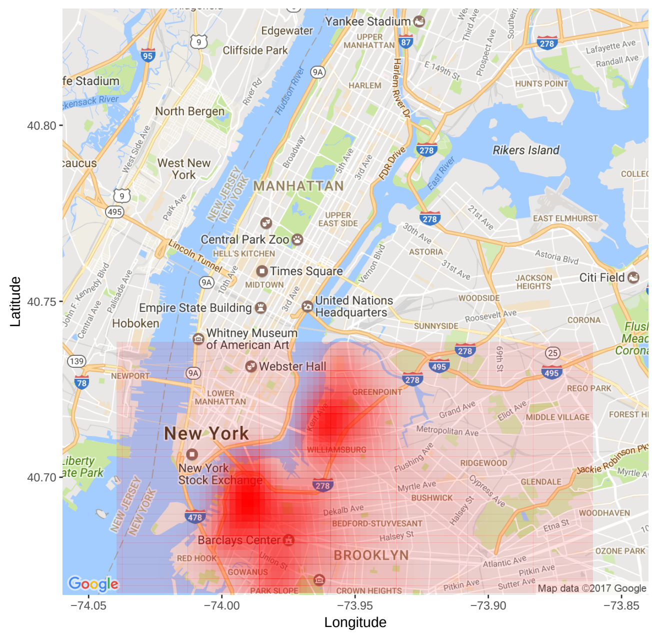

Figure 5: Nonparametric density estimation (normal kernel, bindwidth via normal score rule plot) plots dropoffs only in Brooklyn during the evening rush hours of March 18. We observe less centralized taxi volume near the Barclay's center.

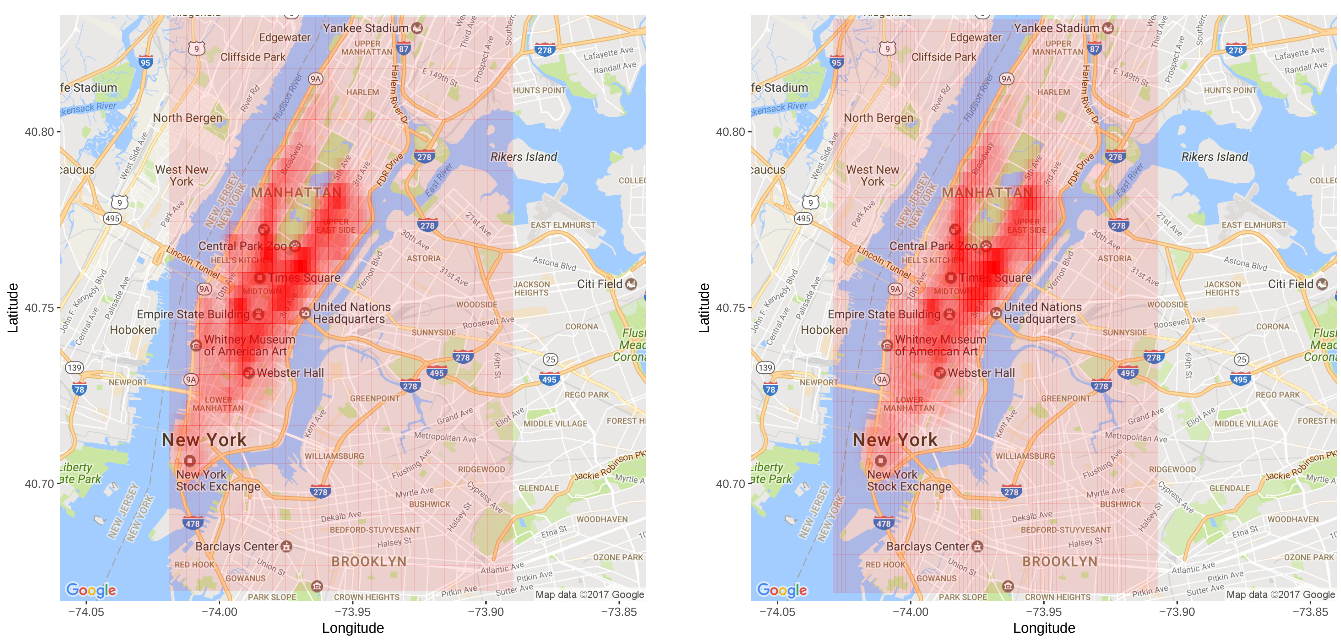

Figure 6: Nonparametric density estimate equivalent to Figure 3, but for pick-ups. The majority of the rides are concentrated in central to lower Manhattan, with little discernable difference in taxi activity concentration near the L train lines.