Bike Sharing Does Not Change Subway Usage: Evidence From New York City

Honors Thesis

Michel Gutmann *

April 28, 2022

Abstract

Public transit systems are challenged by high utilization and some communities remain largely underserved by these transit networks. Policymakers are increasingly interested in ways to alleviate these deficiencies without direct investment into transit systems, which are often extremely costly. I study the CitiBike bike sharing scheme in New York City to determine how alternative modes of transportation interact with the subway network. I implement a differences-in-differences design and find that the introduction of CitiBike had virtually no effect on subway usage in Manhattan and a small positive effect in the other boroughs, suggesting that policies relying on bike sharing to change subway usage may be ineffective. To investigate the role of pricing in bike sharing usage, I implement a spatial matching model to estimate the price-elasticity of demand for CitiBike trips, which I find to be close to zero. However due to the unrepresentative composition of the sample I use to estimate the model, I do not draw conclusions regarding the possible effect of bike sharing subsidies.

*Columbia University. Email: michel.gutmann@columbia.edu. I am grateful to my supervisor, Professor Bernard Salanié, whose support and advice has been invaluable throughout the project. I also thank Professor Michael Best for his advice and for organizing the Honors Thesis Workshop, and Bennett Smith-Worthington for his helpful comments.

1. Introduction

Although New York City's (NYC) subway system is one of the world's largest, both in terms of geographic spread and daily use, many communities within NYC remain largely underserved by the subway network. Some neighborhoods, like North Corona and Pomonok in Queens, do not contain any subway stations and many areas do not have a subway station within a 15 minute walk, especially outside of Manhattan. Extremely high subway expansion and operations costs can make subway expansion unfeasible in areas where usage would be too low to justify the costs of subway extension. The existing subway network itself is also under pressure from high usage. In NYC, subway delays, which are reported to be due to overcrowding in more than a third of cases[1], are estimated by Treffeisen (2017) to have an annual cost of $307 million, with most of the burden falling on users. Delays can make holding a job down more difficult or can imply lost wages, both of which particularly affect low-income users. In his survey, Forman (2017) finds that the per-user cost of subway delays is unevenly distributed among users, with users in low-income communities being much more affected than users in high-income communities.

One possible solution to both overcrowding and lack of subway access is to develop other transportation modes, that can serve either as alternatives to the subway or as a means to connect underserved areas to the subway. If commuters have access to other modes of transportation near the start or end of their trips in areas where the subway is not accessible, they can use these transportation options either to reach a subway station or to replace the subway entirely. This solves the "last-mile" problem (and the symmetric "first-mile" problem), where a user's trips are mostly covered by the subway, but the user does not have a means of reaching their destination from the subway station closest to the destination, i.e. the "last mile" of their trip is not covered by any transportation mode. Subway overuse could also be alleviated if users

1. https://www.nytimes.com/interactive/2017/06/28/nyregion/subway-delays-overcrowding.html

substitute the subway with other modes. Hence, evaluating whether these alternative modes of transportation act as subsitutes or complements to the subway is necessary for comprehensive urban transportation planning. In the context of subway usage, mobility-as-a-service platforms offering light transportation modes such as bikes or electric scooters are particularly interesting, given the relatively low infrastructure costs associated with their development, low expansion times and high flexibility for users. Naturally, these particular modes of transportation are not suited for use by all subway users, but understanding how they interact with existing public transit networks is likely to be useful to understand how more inclusive mobility services which do not require physical effort (such as Phoenix's autonomous Waymo One Cabs[2]) are developed.

This thesis aims to inform policy by documenting complementarity in bike sharing and subway usage in NYC and by estimating the demand elasticity for bike rides in the CitiBike sharing scheme. The latter can help inform possible subsidies to incentivize bike sharing usage.

1.1. Context and Data

The CitiBike scheme was first introduced in Manhattan and Brooklyn in 2013 and has continuously expanded in multiple waves of station openings since. As of 2021, CitiBike has slightly more than 1500 stations open and is the bike sharing scheme with the highest number of average daily users outside of China. Although stations still are concentrated in Manhattan, the system has progressively expanded to all boroughs, except Staten Island. Coverage in boroughs other than Manhattan is sparser and is mostly concentrated in areas closer to Manhattan, i.e. Northwestern Queens, South of the Bronx river and on the Western side of Brooklyn. There also are some stations in Jersey City and Hoboken. Pricing for CitiBike usage until the introduction of single rides in 2019 followed a two-part tariff structure. Customers would have to

2. See https://waymo.com/waymo-one/.

buy an annual subscription or a 1, 3 or 7 day pass before accessing bikes. Then, each trip would be free if its duration was under a certain threshold, and would be charged overtime fees if they went over this threshold. Threshold and overtime fees vary depending on the type of pass. Like most other bike sharing schemes, CitiBike "rebalances" stations over the day, moving bikes from stations where they accumulate to emptying stations.

CitiBike provides publicly accessible data on all trips made with the service, including start and end stations and their location, start and end time, type of subscription, unique bike identifier and trip duration[3]. CitiBike also provides real-time data on the number of available bikes at each station, which has been archived by The Open Bus[4]. In addition, I reconstruct a history of subscription types and corresponding prices, as well as of overtime pricing structures, from snapshots of the pricing page on CitiBike's website, archived by the Internet Archive. Since each trip's individual price is only determined by its duration and the subscription type (until the introduction of single rides in 2019), this pricing history allows me to compute trip specific prices for any trip before the introduction of single rides.

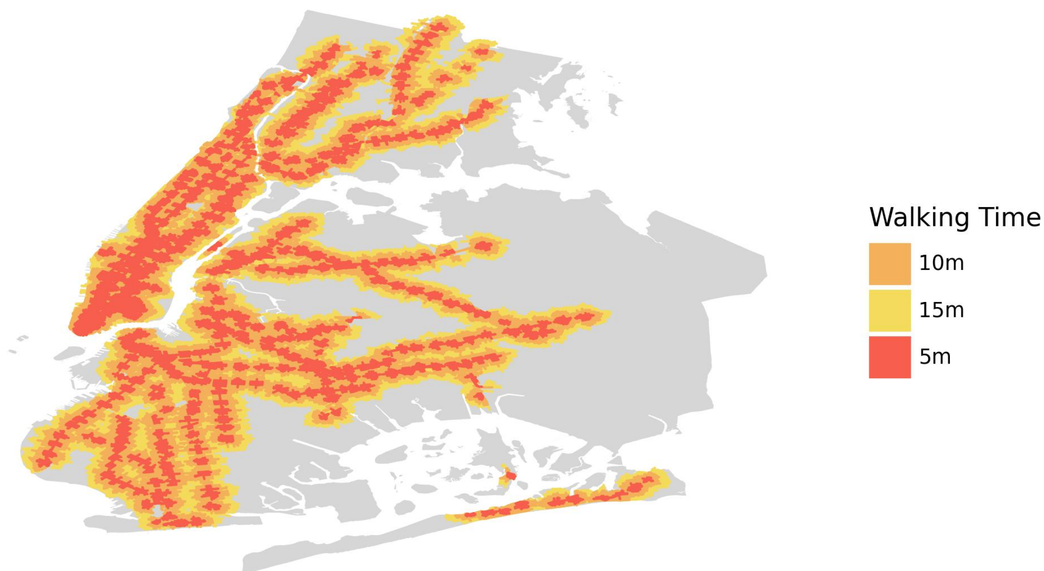
I study the interaction between CitiBike and the New York City subway network. The NYC subway covers all boroughs and all boroughs are connected by the subway, except Staten Island since the Staten Island Railway is not connected to the rest of the network. Station density is higher in Manhattan than in other boroughs. Figure 1 presents a summary of areas covered by the subway. The Metropolitan Transportation Authority (MTA) publishes turnstile counts which can be aggregated at the subway station - day level, with separated values for the number of entries and exits[5].

3. The data can be accessed at https://ride.citibikenyc.com/system-data

4. This data can be accessed at https://www.theopenbus.com/raw-data.html

5. The data can be accessed at http://web.mta.info/developers/turnstile.html. The turnstile counts are published at the turnstile level over 4 hour periods. These periods are not identical for all turnstiles, so aggregating at the daily level provides a consistent measure of entries and exits.

Figure 1 – Subway Coverage



1.2. Related Literature

Previous research has found that alternative modes of transportation can serve as a complement to the subway and provide a solution to the last mile problem. Chu et al. (2021) use a differences-in-differences strategy on data from multiple major Chinese cities to show that when a bike sharing service enters, the price premium for homes near the subway decreases to approximately zero. This can be explained by either substitution or complementarity between subway and bike usage. Home buyers may think that subway stations are more accessible once bikes are introduced or may simply rely on the subway less for their transportation. This result however indicates that bike sharing introduction should have some effect, whether positive or negative, on subway ridership. Fan and Zheng (2020) use Beijing data to show that bike sharing service increases the use of nearby subway lines and decreases road congestion, especially in areas with poor subway access. In particular they are able to differentiate between subway subsituting and subway complementing bike sharing trips. This allows them to show that there is strong complementarity and little subsitution between bike sharing and subway usage. Research has also been conducted on other similar modes of transportation. Hall et al. (2018) implement a differences-in-differences strategy across multiple American cities to show that the introduction of Uber ridesharing services in a city is associated with increased public transit usage in that city. Christensen et al. (2021) plan to run a randomized controlled trial with subsidies for ride hailing trips from and toward subway/train stations in Chicago to evaluate whether subway usage increases when access is facilitated. With respect to this literature, this thesis is most closely related to Hall et al. (2018), Chu et al. (2021) and Fan and Zheng (2020). This thesis implements a differences-in-differences design similar to Hall et al. (2018) at the scale of a city rather than over multiple cities to study the effect of a bike sharing station opening on nearby subway station usage. This study also differs from Chu et al. (2021) as it studies the direct effect on subway usage rather than the subway station proximity price premium on the real estate market. This allows me to determine whether the effects of bike sharing on subway usage are substitutive or complementary. Lastly, it differs from Fan and Zheng (2020) as it studies a significantly different urban context and considers a different definition of treatment.

To estimate the price-elasticity of demand for bikes, I build on a growing literature that uses matching function methods to recover market parameters in spatial contexts. Brancaccio et al. (2020a) provide a general introduction to these techniques. In particular, there have been multiple applications to transportation markets, both of goods (Brancaccio et al. (2020b)) and of people (Buchholz (2021), Rosaia (2020), Fréchette et al. (2019)). These methods enable me to recover the number of searching consumers and demand parameters.

2. Documenting CitiBike - Subway Multimodality

In this section, I implement a differences-in-differences (DiD) approach to explore how the CitiBike scheme has been integrated with the subway network, i.e. whether CitiBike riders use the bike sharing scheme as a means to connect to the network or as a subsitute for the subway. The outcome variable of interest is the monthly average of daily entries in the subway station. To estimate the effect of connection to the CitiBike network on subway entries, I make use of the staggered station deployment to estimate a CitiBike station effect on subway trips. I use the CitiBike trip data to identify which bike stations were active at different points in time. I track the opening of new stations by recording their first appearance in the trip data. I combine this information with New York City Metropolitan Transportation Authority data on subway entries and exits to construct a panel at the subway station - month level[6] for all subway stations from January 2011 to December 2018. Summary statistics for this panel are provided in Appendix A.

2.1. Strategy

I define treatment as a binary indicator of the presence of an active CitiBike station within walking time[7] of each subway station. This definition allows me to capture the effect of the connection of a subway station to the CitiBike network, rather than the effect of a bike station itself, which would be obtained by defining a continuous treatment that would be equal to the number of nearby bike stations. To determine whether a subway station is treated, I use OpenStreetMap street network data through the OSMnx Python library[8] to reconstruct the area that is within a 5 minute walk of each subway station. Then, for each subway station, I identify whether the

6. Data is available at the daily level but I aggregate it at the month level due to computational constraints in estimation.

7. Walking time is taken to be 5 minutes, assuming a walking speed of 4.5 kilometers per hour.

8. Developed by Boeing (2017) and available at https://github.com/gboeing/osmnx.

corresponding 5 minute walk area contains bike stations. Since I observe each bike station's first and last appearance in the trip data, I am also able to determine if treatment for any given subway station stops, i.e. the bike stations around it are shut down and the station loses its connection to the CitiBike network[9]. I therefore consider a binary treatment variable $T_{i,t}$ defined as follows:

$$T_{i,t} = \begin{cases} 1 \text{ if there is a bike dock within 5 minutes of } i \text{ at period } t \\ 0 \text{ otherwise} \end{cases} \tag{1}$$

Then, the canonical formulation of the DiD strategy is expressed by the following regression specification:

$$Y_{it} = \beta_0 T_{it} + \alpha_t + \eta_i + \epsilon_{it} \tag{2}$$

where Y_{it} is the average daily number of entries at subway station i in month t, T_{it} is the treatment binary variable, β_0 the effect of treatment, α_t is a month fixed effect and η_i is a subway station fixed effect. However, if the effects take time to appear because users need time either to notice that a bike station has opened near the subway station or need time to realize they can use bikes as either a substitute or complement to the subway, then the specification described in Equation 2 may not capture the desired effect. To adress this, we can implement an event-study specification, as presented in the below equation

$$Y_{it} = \sum_{k=0}^{l} \beta_k T_{it} + \alpha_t + \eta_i + \epsilon_{it} \tag{3}$$

where l is the number of months before and after treatment for which we want to compute effects. Such a specification captures dynamic effects of treatment. It estimates effect of treatment β_k after $k \in \{0, \ldots, l-1\}$ periods, which is of interest if

9. Although this data does not allow me to distinguish an inactive station and a station receiving no traffic, it is unlikely that any given station would not receive any traffic over a long period. The sample of trip data I consider to determine first and last appearances extends into 2022, which allows me to ensure that bike stations are not suddenly marked as inactive in the constructed panel, which only considers 2011-2018. Regardless, the first and last appearance dates likely constitute good proxies of the station opening/closing and otherwise indicate when stations are effectively used.

it takes time for the treatment to have an effect.

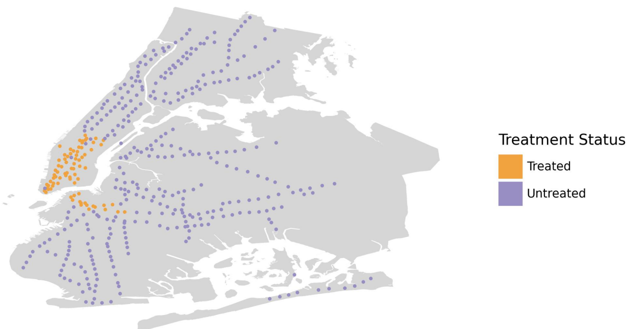
This approach yields interpretable results. If the addition of a bike dock in walking distance of a subway station has a positive effect on the number of subway departures at a given station, it seems likely that the effect of bike sharing is to bring more users to subway stations. This would imply that subway and CitiBike usage are mostly complementary. On the contrary, if the effect is negative, then the effect of bike sharing is to offer an alternative to commuters who would otherwise have taken the subway. This would in turn imply that bike sharing serves as a substitute for the subway. Lastly, finding no effect could also indicate that bike sharing and subway usage are unrelated, suggesting that the two modes of transportation have distinct uses. For instance, the subway could be mostly used for commuting while bike sharing could be mostly used for recreational or occasional trips.

The first CitiBike stations opened on May 27th 2013 in Manhattan and Brooklyn and most of these stations remain active today. Additional major station opening waves occurred in August-September 2015 and August 2016. In 2017-2018, additional stations open throughout the year, with no clearly identifiable opening waves. Stations are mostly concentrated in Manhattan at first and expand to the outer boroughs starting from the second expansion wave. Maps of treatment status at subway stations over time are presented in Figure 2.

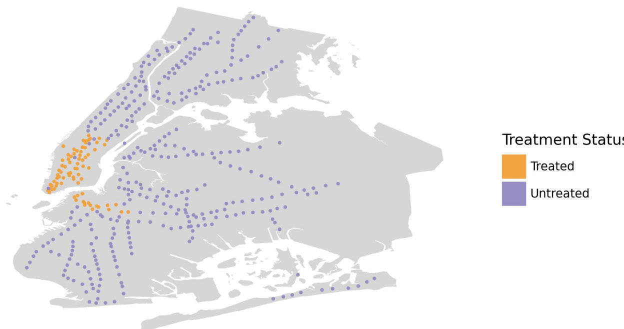
Goodman-Bacon (2021) shows that the simple two-way fixed effects (TWFE) DiD estimator as presented in Equation 2 requires treatment effects to be constant over time in a staggered treatment context. In particular, he shows that the TWFE estimator is a weighted average of the group-specific treatment effect estimates, which can vary over time. In the case where some weights are negative, it is possible that the TWFE estimate has a sign that is different from the sign of all the treatment effect estimates. de Chaisemartin and D'Haultfoeuille (2021) and Roth et al. (2022) provide an introduction to this issue and survey approaches to recover unbiased estimates. There are reasons to believe that treatment effects are heterogeneous in this context.

Figure 2 – Maps of Subway Stations with Treatment Status by Date

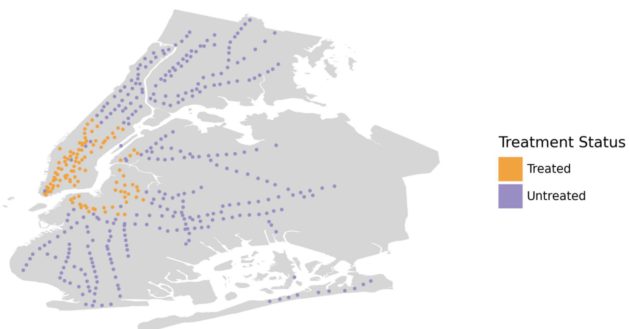
(a) Dec. 31st, 2013



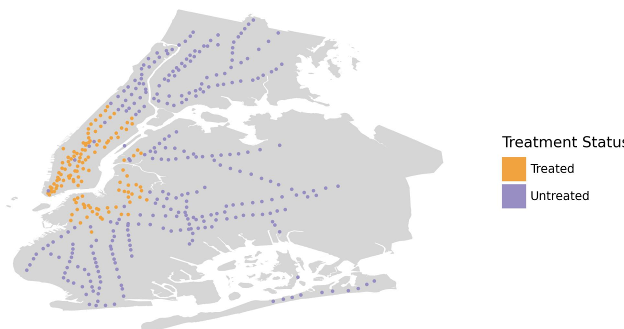
(b) Dec. 31st, 2014



(c) Dec. 31st, 2015

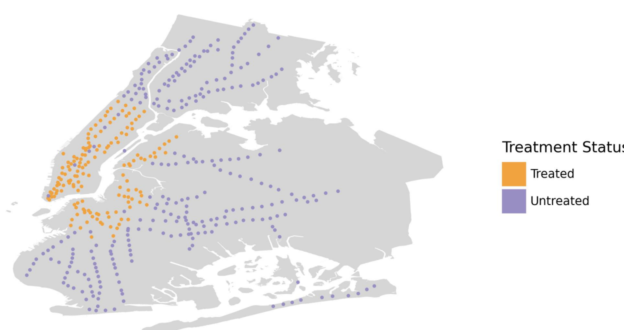


(d) Dec. 31st, 2016



(e) Dec. 31st, 2017

(f) Dec. 31st, 2018



In particular, since the number of CitiBike users has increased over time, assuming that new users are similar to the already existing users, one could expect treatment effects to grow in magnitude over time. I use the weight decomposition proposed by de Chaisemartin and D'Haultfœuille (2020b) and find that it includes negative-weighted group treatment effect estimates. To provide an unbiased estimate of the Average Treatment Effect on the Treated (ATT), I use the DiD_M instantaneous effect estimator proposed by de Chaisemartin and D'Haultfœuille (2020b). This estimate corresponds to the average effect of connecting a subway station to the CitiBike network at time t on the number of station entries at t. The analogous event-study specification is estimated using the DiD_l estimators proposed by de Chaisemartin and D'Haultfoeuille (2020a) with dynamic effects over 10 periods. Each DiD_l estimate corresponds to the average effect in $t+l$ of treatment in t, using period $t-l-1$ as a baseline. These dynamic effect estimates can be collapsed into a weighted average δ_+ that measures the average cumulative effect of treatment. In addition, I also provide results restricting the data to subway stations located in Manhattan. Results are provided in Table 1.

These estimators rely on multiple identification assumptions. The first assumption is that there is no anticipation, which in our case means that commuters do not change their use of a subway station before is is treated. This seems plausible given it is unlikely that users of a subway station use it more or less because they expect a bike station to open nearby at a later day. The second identification assumption requires that:

1. The $(Y_{i,t}(0)_{1\leq t\leq T}, T_i$ vectors are independent
2. $\mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0)|T_i] = \mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0)]$
3. $\forall t \geq 2, \forall i \neq j, \mathbb{E}[Y_{i,t}(0) - Y_{i,t-1}(0)] = \mathbb{E}[Y_{j,t}(0) - Y_{j,t-1}(0)]$

The first statement specifies that the potential outcomes and treatments of different entities are independent. This is a common assumption for DiD strategies where standard errors are clustered at the entity level. The second statement corresponds to a strict exogeneity assumption, i.e. treatment for entity i must be independent from shocks to never-treated outcomes for entity i. Lastly, the third statement requires that the expected never-treated outcome for entity i follows the same evolution over time. This corresponds to the usual parallel trends in DiD strategies.

Firstly, to account for the possibility that stations in different boroughs have different trends in expected never treated outcomes, I restrict comparisons between switchers and not yet switchers to stations within the same borough. Secondly, to validate these assumptions I compute long-difference placebo estimators as proposed in de Chaisemartin and D'Haultfoeuille (2020a), who show that if these identification assumptions hold, the expectation of each placebo is zero. Testing these placebo estimators is analogous to testing pre-period coefficients in the TWFE context and is a necessary but not sufficient condition for identification. A coefficient plot of these placebo estimates and their 95% confidence intervals (constructed with a normal approximation) and analogous dynamic effect coefficients is presented in Figure 3. In addition, when these placebos are not zero, the sign of the bias on the corresponding dynamic effect estimate is opposite to the placebo's sign.

2.2. Results and Discussion

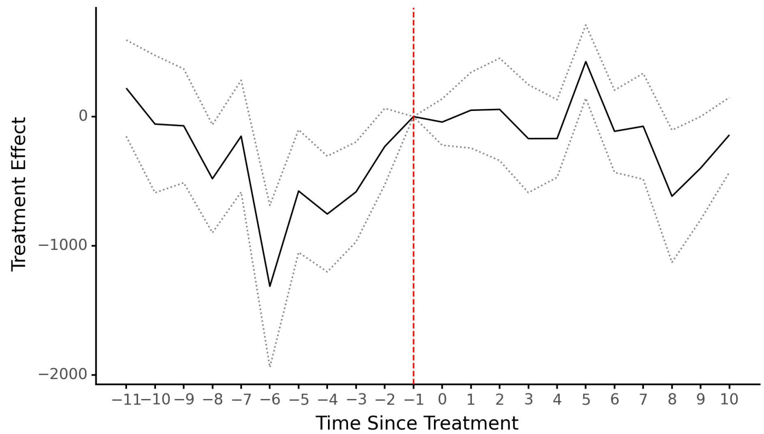
I present estimates of the effect of connecting a subway station to the CitiBike network in Table 1. This table presents results for multiple specifications, with differing panel lengths and restricting the estimation to different geographic areas. I provide estimates of instantaneous treatment effects and estimates computed with dynamic effects, to account for user learning time.

I present coefficient plots of treatment effects and long-difference placebos in Figure 3. In particular, I present corresponding event study plots for specifications 3-6. I provide analogous event study plots where the outcome variable is the number of subway exits rather than entries in Appendix B and find similar results.

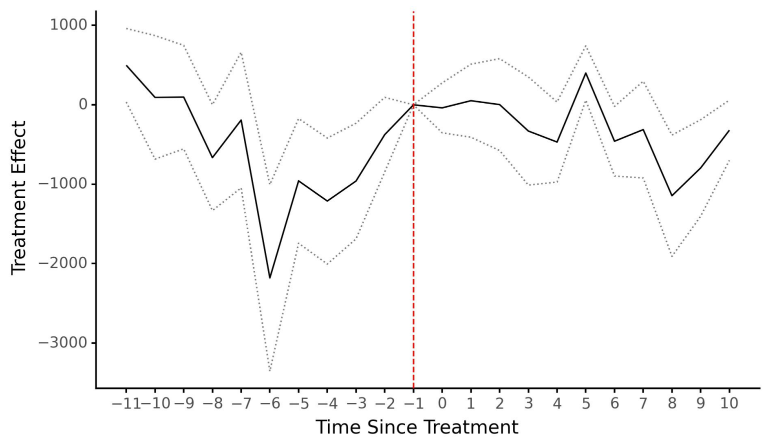
The event study plots presented in Figure 3 clearly show that when Manhattan

Figure 3 – DiD Event Study Plots

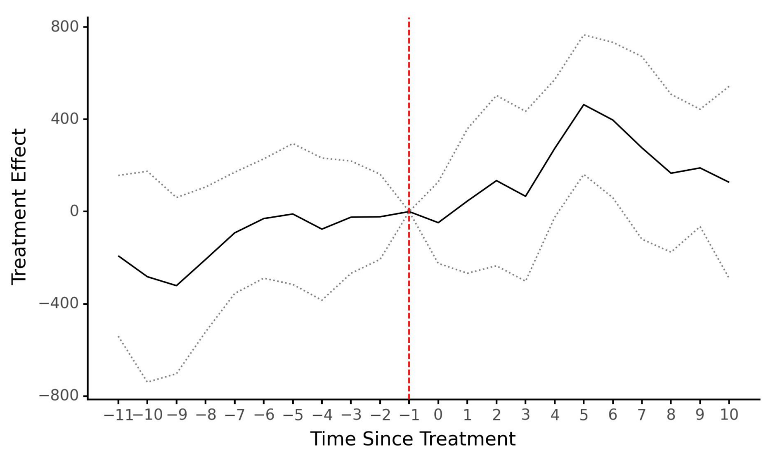
(a) All Boroughs



(b) Manhattan Only



(c) Bronx, Brooklyn, Queens



Notes. Standard errors are estimated with 50 bootstrap replications and are clustered at the subway station level. When multiple boroughs are included, borough specific trends are accounted for non-parametrically.

Table 1 – Estimates of Effect of CitiBike Network Connection on Subway Entries

	(1)	(2)	(3)	(4)	(5)	(6)
Estimate	-40	-109	-310	190	300	754
Standard Error	91	140	217	117	220	192
Dynamic Effects	0	10	10	10	10	10
Boroughs	All	All	Manhattan	Outer	Manhattan	Outer
Period	Full	Full	Full	Full	2016/18	2016/18
Estimator	DiD_M	DiD_l	DiD_l	DiD_l	DiD_l	DiD_l

Notes. The reported effect is the instantaneous effect for specifications with no dynamic effects. For specifications with dynamic effects, the weighted average of instantaneous and dynamic effects δ_+ is reported. Standard errors are estimated with 50 bootstrap replications and are clustered at the postal code level. Borough specific trends are accounted for non-parametrically when multiple boroughs are included in the estimation. Staten Island is excluded from all estimations.

is included in the estimation process, the parallel trends assumption does not hold. However, the DiD_l placebo estimators's signs are of the opposite sign of the bias on the corresponding dynamic effect estimate. If $\text{DiD}_{-l-1} < 0$, as is the case in these plots, the bias on DiD_l is positive. As a result, the estimates in Table 1 for estimations over the whole panel that include Manhattan constitute an upper bound for the true value of the effect. When restricting estimation to the outer boroughs, placebos are consistently close to zero and negative when not approximately zero, indicating that any violation of parallel trends is likely small in magnitude and the corresponding estimate in Table 1 is a tight upper bound of the ATT.

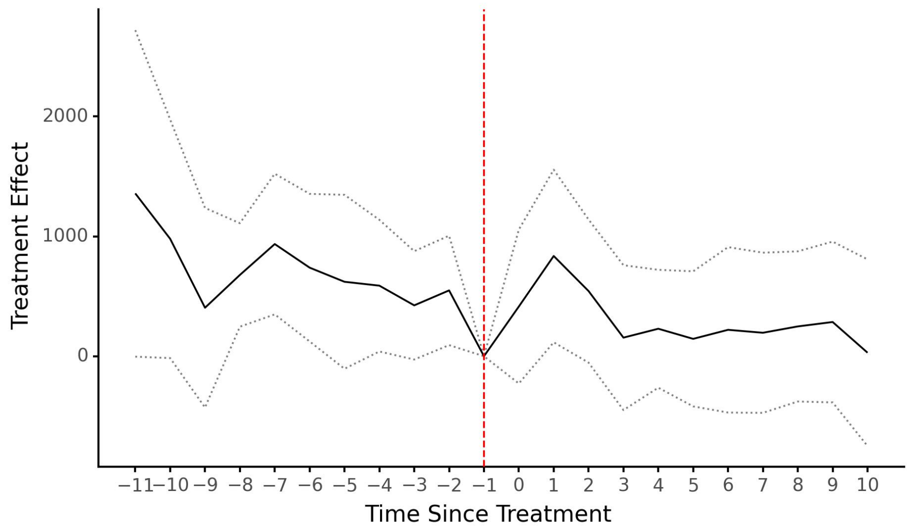
The values presented in Table 1 suggest that connecting a subway station to the CitiBike network has little to no effect on subway station entries. When Manhattan is included the effect estimates are negative and very small relative to daily station entries. CitiBikes may be a weak substitute for the subway. However, the relatively large standard error also makes it difficult to assert this with confidence. When restricting estimation to the outer boroughs, the estimate is positive but still very small relative to daily station entries, which would imply weak complementarity.

These results are consistent with economic intuition. Manhattan is very dense, and most areas in the borough are directly covered by the subway. This makes it unlikely that users take a bike to access the subway, i.e complementarity is unlikely. On the other hand, because most of Manhattan is covered by CitiBike stations, substitute trips are possible. In the outer boroughs, which are not entirely covered by the subway, users may find bikes useful to access the subway network. Still, until 2015-2016 the CitiBike network of stations in these boroughs was much less extensive than it is in Manhattan so few complementary trips could be made. To check whether such an explanation is plausible, I compute estimates of the δ_+ parameter when the panel is restricted to 2016-2018. In this period, the CitiBike dock network in the outer boroughs was extended and covered areas the subway does not reach. These estimates are presented in Table 1 and corresponding event study plots are presented in Figure 4.

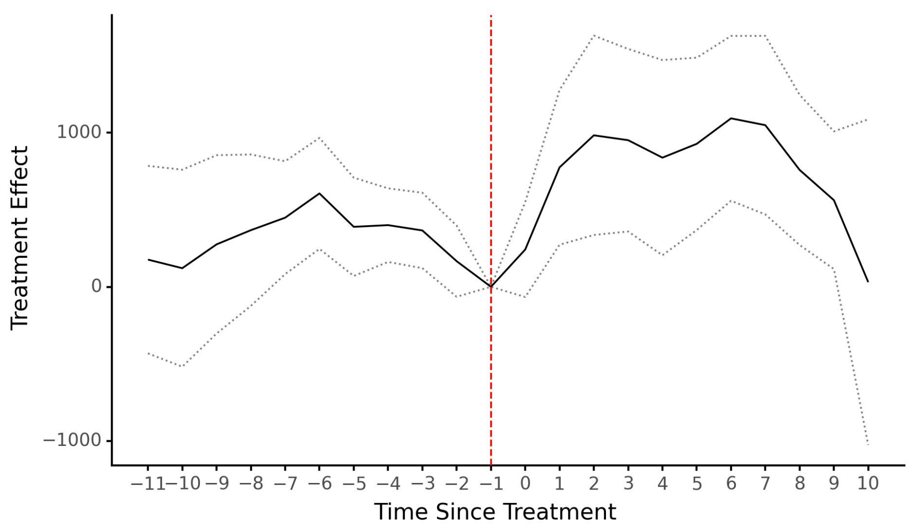
In Manhattan, the placebos are positive, which suggests the estimate is a lower bound. Placebos in the outer boroughs are consistently close to zero, small and positive, indicating the estimate is a tight lower bound. In both cases, the estimates are positive, which implies that the ATT is positive as well. These results are consistent with the previous explanation in the outer boroughs, where there appears to be complementarity between CitiBike and the subway. In Manhattan, the effect is likely smaller then in the outer Boroughs but seems to be weakly positive. These results suggest that the effect is quite small relative to daily entries for stations in Manhattan, which is partially consistent with the explanation described above. If the effect is approximately zero in Manhattan, then there is no complementarity, which is what intuition predicts, but there also is no substitution, which is a more unexpected result. One possible explanation for the absence of subsitution is that users perceive CitiBike to serve a very different purpose to the subway. For instance, they could view the subway as a means of transportation and bike sharing as leisure. In this case, adding a bike station near a subway station does not take away users from the subway, since each service attracts different users.

Figure 4 – DiD Event Study Plots 2016-2018

(a) Manhattan Only



(b) Bronx, Brooklyn, Queens



Notes. Standard errors are estimated with 50 bootstrap replications and are clustered at the subway station level. When multiple boroughs are included, borough specific trends are accounted for non-parametrically.

These results are partially consistent with those of Chu et al. (2021) and Fan and Zheng (2020). Fan and Zheng (2020) find that, in Beijing, dockless bike sharing serves as a complement for the subway and that substitution is negligible. The context in which their studies are conducted, 10 major Chinese cities and Beijing respectively, is significantly different from the context of NYC. Chinese cities tend to be spatially much larger and have a more connected subway network. In such a context, subway access is very valuable since accessing one line gives easy access to other lines, and the subway can reach areas that would be too far to reach on a bike. My finding on complementarity in the outer boroughs is consistent with these results. The outer boroughs are the areas within NYC which resemble Beijing most in terms of spatial spread, and I also find bike sharing to be complementary to subway ridership. My results in Manhattan differ from the results in Beijing, but it is likely that this difference is due to Manhattan being significantly less spread out than Beijing. In addition, there may be significant differences in cycling infrastructure and attitudes towards cycling that explain the differences in results.

In the specific context of NYC, these results indicate that regardless of the exact value of the effect, it is likely quite small, and bike-sharing does not strongly influence subway usage, except in the outer boroughs where there is some weak complementarity. These results imply that, in Manhattan, it is likely that bike-sharing and the subway serve different purposes and have different user types, i.e. if the subway is mostly used for commuting, then bike sharing is mostly used for recreational purposes or occasional trips. In addition, this suggests that connecting subway stations on overcrowded transit network sections is not an effective means of reducing subway usage. Conversely, in the outer boroughs, the weak complementarity implies that bike-sharing can partially connect communities to the subway, but this effect is small in relative terms.

3. Estimating Demand Elasticities For CitiBikes

This section describes and estimates a model of the CitiBike market, which replicates the NYC taxi cab market model constructed by Buchholz (2021)[10]. It also draws on similar estimation techniques to recover the demand elasticity for bikes in the CitiBike scheme.

3.1. CitiBike Market Model

Each day is divided in T periods, indexed by t. The market contains locations, or bike stations, indexed by $i \in I$. Bikes can move between locations in two ways. Customers can pick a bike up at a location and ride it to another location. Bikes can also be "rebalanced" by the system operator, i.e. moved by the operator from a station to another. There are two types of customer, indexed by c, annual subscription holders and short pass holders. The market is characterized at each period by a state

$$\mathcal{S}^t = \left\{ \{v_i^t\}_i, \{e_{ij}^k\}_{i,j,k<t}, \{r_{ij}^k\}_{i,j,k<t} \right\} \tag{4}$$

where:

- v_i^t is the number of unused bikes in i at t
- e_{ij}^k is the number of bikes en route from i to j that were matched with a customer in period k (before t)
- r_{ij}^k is the number of rebalanced bikes en route from i to j whose rebalancing process started in period k (before t)

10. Some changes are made to the taxi cab market model to better represent the CitiBike market.

3.1.1 Trip Prices

The price of a trip is uniquely determined by its duration and the customer type at any given time. We can define a pricing structure $PS_{c,t}$ as the following two term sequence:

$$(a_n, b_n)_{n\in\mathbb{N}}^{c,t} \tag{5}$$

where:

- a_n is the amount of time before the n-th overtime period starts
- b_n is the overtime fee for overtime period n

The price for a trip is given by:

$$\pi_{i,j,c}^t = \sum_{n\in\mathbb{N}} b_n^{c,t} \mathbb{1}_{\{\tau_{i,j} > a_n^{c,t}\}} \tag{6}$$

where $\tau_{i,j}$ is the trip duration.

3.1.2 Demand

The number of customers of type c wishing to move to a new location j in a location i at time t is $u_{i,j,c}^t \sim \text{Poisson}(\lambda_{i,j,c}^t)$. Each term is a function of the trip price $\pi_{i,j,c}^t$. I also denote the total number of customers looking for a bike in station i at period t by $u_{i,c}^t \sim \text{Poisson}(\lambda_{i,c}^t)$. Note that $\lambda_{i,c}^t = \sum_{j\in I} \lambda_{i,j,c}^t$.

It is assumed that the demand curve has constant elasticity. Demand depends on customer type, time of day, start location, end location and trip price. As a result, bike demand is given by:

$$\ln(\lambda_{i,j,c}^t) = \beta_1 \ln(\pi_{i,j,c}^t) + \gamma_{i,j} + \nu_t + \alpha_c + \epsilon_{i,j,c}^t \tag{7}$$

where α_c is a customer type specific term, $\gamma_{i,j}$ is an origin-destination specific term and ν_t is a period specific term. Demand elasticity β_1 is assumed to be identical across locations, time and customer types.

3.1.3 Search Process

Matching The following assumptions are made regarding matching:

1. Matches only occur among bikes and customers in the same location

2. Each customer is matched to a bike randomly

The expected number of matches with customers of type c in i wishing to move to j at t is given by a matching function $m(\lambda^t_{i,j,c}, v^t_i)$. The aggregate matching function is given by:

$$m(\lambda^t_{i,j,c}, v^t_i) = v^t_i \cdot \left(1 - \exp\left(-\frac{\lambda^t_{i,j,c}}{v^t_i}\right)\right) \tag{8}$$

This aggregate matching function is a spatially homogeneous frictions version[11] of the matching function used by Buchholz (2021)[12]. Recovering friction parameters is out of the scope of this project, which is why frictions are assumed to be identical across locations. Any possible region-specific differences in matching will not influence the elasticity estimation process, which uses origin-destination fixed effects.

Rebalancing Process The rebalancing of bikes toward a location i in period t is chosen at the system level and is not explicitly modeled since no rebalancing cost information is available. In addition, since supply is observed in the CitiBike context, observing or reconstructing rebalancing flows is not necessary since it is not necessary to reconstruct supply at each period.

11. It takes the friction parameter α_r to be 1 in all regions, which is close to the average of the empirical estimates in the NYC taxi cab market. See table 5 in Buchholz (2021).

12. This form of matching function was first used by Butters (1977) and Hall (1979). See Petrongolo and Pissarides (2001) and Appendix A.9 of Buchholz (2021) for a derivation.

3.1.4 Timing

Each day starts with an initial distribution of bikes in stations $\mathcal{S}^1$, which is constant across weekdays. At the end of each period bikes that have found a customer during the period are removed from the stock of bikes in stations and customers are removed from the searching customers. In the next period, the stock of vacant bikes in each location is updated with the number of bike trips that end in a location and the net rebalancing in a location. In summary, the day proceeds as follows:

1. Bikes are distributed according to $\mathcal{S}^1$
2. $\sum_c m(\lambda^1_{i,c}, v^1_i)$ bikes are matched to customers and leave their location
3. The remaining $\sum_c \lambda^1_{i,c} - m(\lambda^1_{i,c}, v^1_i)$ unmatched customers leave the market and choose another mode of transportation
4. The remaining $v^1_i - \sum m(\lambda^1_{i,c}, v^1_i)$ unmatched bikes are given a new location (either they are rebalanced or they do not move)
5. Rebalanced bikes and bike trip arrivals form a new distribution of bikes $\mathcal{S}^2$
6. The process is repeated to form $\mathcal{S}^3, ..., \mathcal{S}^T$

The distribution of bikes $\mathcal{S}^t$ at each period t is directly observed.

3.2. Estimation Process

In this section, I explain how I estimate model parameters using data on supply and matches. Match data is obtained from the CitiBike trip data, while supply data is obtained from the CitiBike real time data archived by The Open Bus.

3.2.1 Estimating Model Parameters

Estimating $\lambda^t_{i,j,c}$ Assuming that $\forall i, t\ v^t_i \neq 0$, I rearrange Equation 8 to recover $\lambda^t_{i,j,c}$, the demand for bikes in i for trips with destination j, by the following formula:

$$\lambda^t_{i,j,c} = -v^t_i \cdot \ln\left(1 - \frac{m^t_{i,j,c}}{v^t_i}\right) \tag{9}$$

where $m^t_{i,j,c}$ is the average over weekdays of the number of trips from i to j that start in period t. Both supply v^t_i and matches $m^t_{i,j,c}$ are observed. I present maps of average demand by bike dock in each postal code area for May and June 2017 in Figure 5.

Estimating Demand Elasticity To estimate the demand elasticity in Equation 7, I use a change in overtime pricing structure that occurred on June 1st, 2017 and affected the overtime prices of trips made by annual subscribers. The pricing structure for annual subscribers before and after the change is presented in Table 2.

Table 2 – Pricing Structure Change Summary

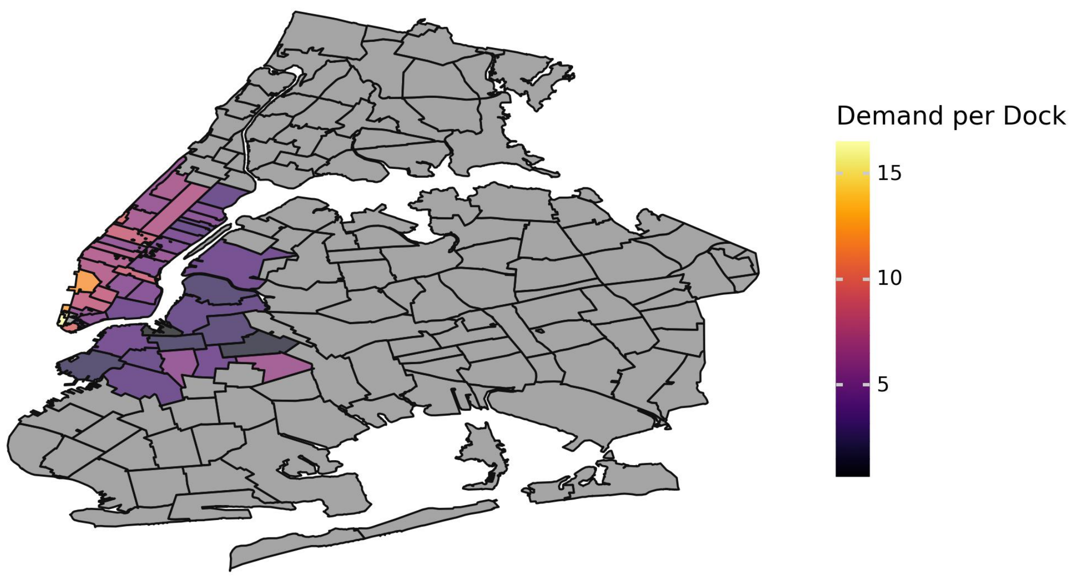
Trip Duration τ (Minutes)	Initial Price ($)	New Price ($)	Matches
$\tau < 45$	Free	Free	2722659
$45 \leq \tau < 75$	2.5	5	36950
$75 \leq \tau < 105$	6.5	10	8378
$105 \leq \tau$	$6.5 + 9 \cdot \lceil \frac{\tau - 105}{30} \rceil$	$10 + 2.5 \cdot \lceil \frac{\tau - 105}{15} \rceil$	13046

Notes. $\lceil . \rceil$ is the ceiling function. Information on prices was reconstructed from archived snapshots of the CitiBike pricing page.

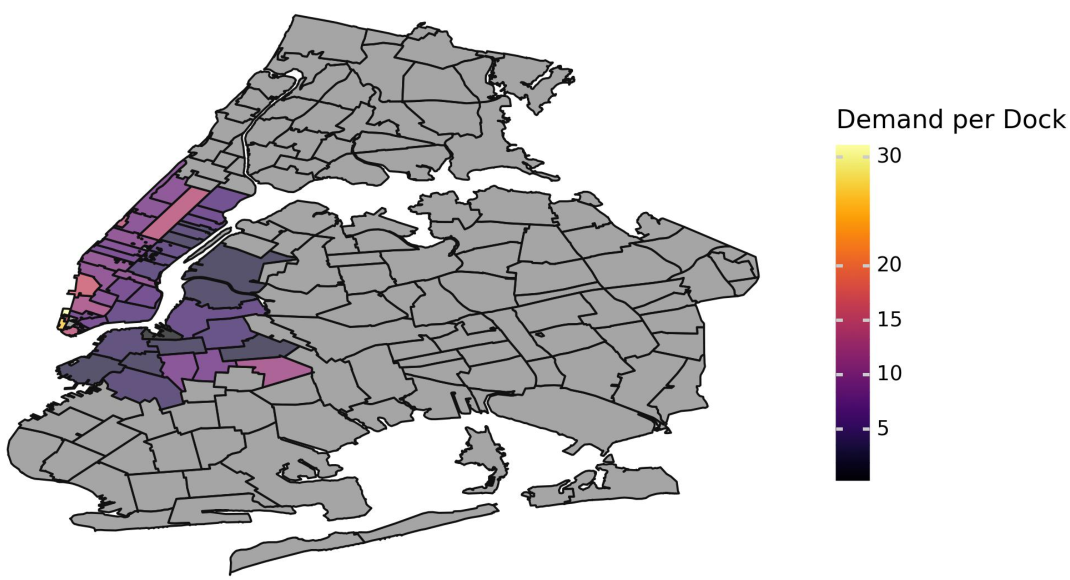
Note that the annual subscription price was also raised from 155$ to 163$ on June 1st 2017, potentially reducing the number of active subscribers and the resulting demand. I assume that this effect is negligible for two reasons. Firstly, the relatively small magnitude of this change makes it unlikely that the total number of subscribers would sharply drop. Second, it is unlikely a very large portion of annual subscribers was due for a subscription renewal exactly at this time. This implies that even if

Figure 5 – Average Monthly Demand for Trips Above 45 Minutes

(a) May 2017



(b) June 2017



Notes. These plots provide a visualization of average monthly demand by bike dock in each postal code area for the months of May and June 2017. The demand values are reconstructed using Equation 9. Postal code areas in gray do not contain any CitiBike stations.

some subscribers chose not to renew, the proportion of them who did not renew in the considered period is likely quite small. Also note that single rides and electric bikes are not offered throughout May and June 2017, so the only trip specific costs for annual subscribers are overtime fees.

The supply archive contains data for May 1st - May 26th and June 8th - June 30th. Although this does not include all days in each month before and after the pricing structure change, the relatively low number of missing days is unlikely to significantly change results. In addition, the absence of the immediate post-pricing change days is likely to attenuate any effect caused by knee-jerk reactions to the change. Periods are taken to be 30 minutes long, and I take the first period to be 6-6:30am and the last period to be 8:30-9pm. Lastly, to target the users that were affected by this pricing change, I restrict matches to trips made by annual subscribers whose duration is greater than 45 minutes to compute the $m_{i,j}^t$ and $\pi_{i,j}^t$ values[13].

Assuming a negative elasticity of demand, one could expect two possible reactions to such a price change. Firstly, we could observe lower demand over all origin-destination pairs where the travel time is higher than 45 minutes, in which case we would expect the number of daily matches to decrease. Secondly, we could observe demand to shift towards origin-destination pairs with lower travel times, even among those where travel time is above 45 minutes. In this second case, we would expect the daily average trip duration to decrease. As I show in Figure 6, even though the series are noisy, it appears that neither average trip duration nor the number of matches decrease after the price change. This suggests that demand is fairly inelastic.

I estimate demand elasticity with the following specification:

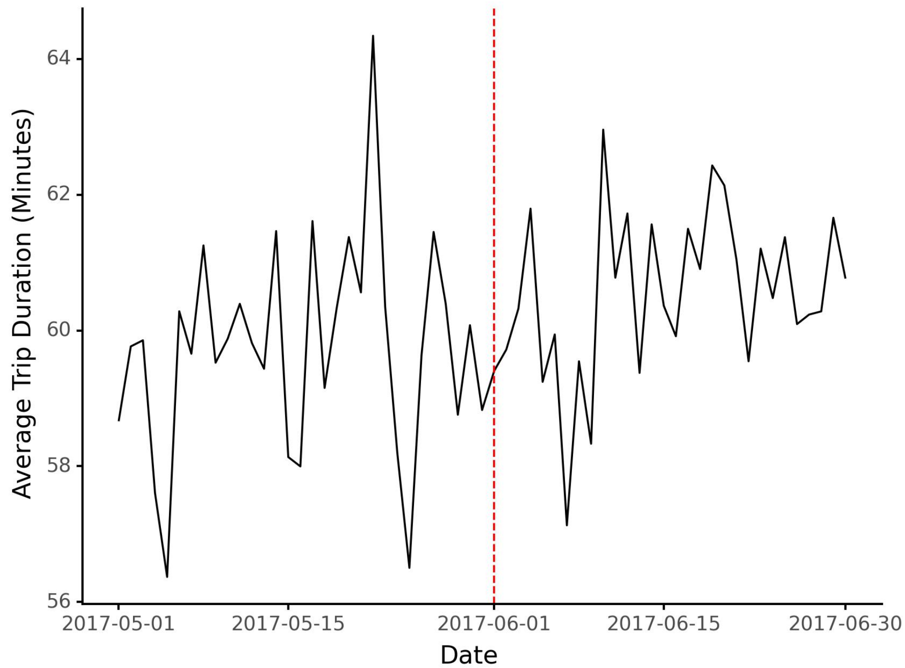
$$\ln(\lambda_{i,j}^t) = \beta_1 \ln(\pi_{i,j}^t) + \gamma_{i,j} + \nu_t + \epsilon_{i,j}^t \tag{10}$$

where $\gamma_{i,j}$ is a origin-destination fixed effect and ν_t is a period fixed effect. Since we

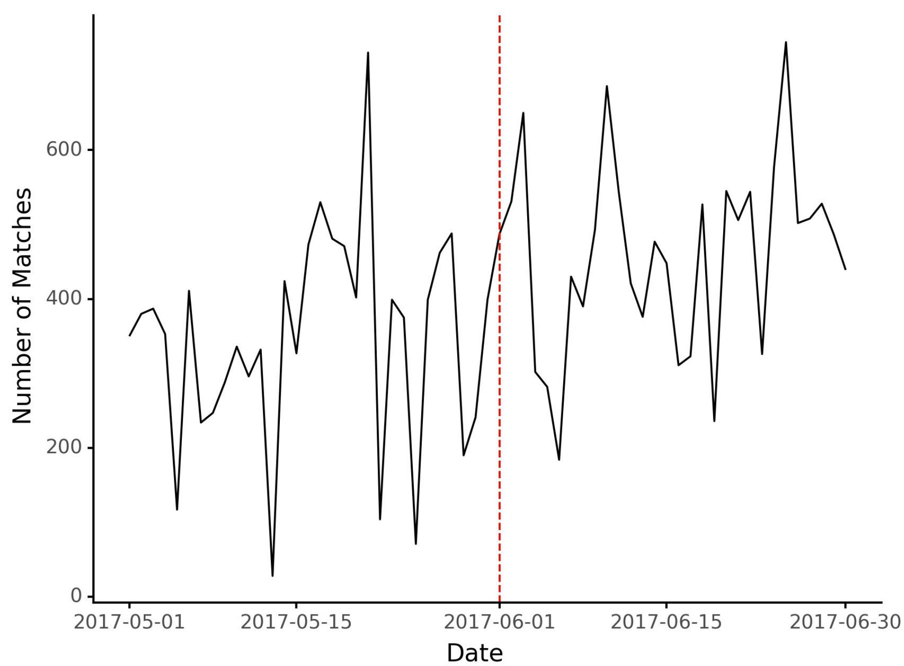
13. Given the sample is restricted to annual subscribers, I remove the c index from these terms.

Figure 6 – Summary of Trips Above 45 Minutes Before/After Price Change

(a) Daily Average Trip Duration



(b) Daily Number of Matches



Notes. These plots provide visualizations of the time series for daily average trip duration and daily number of matches. The sample is restricted to trips longer than 45 minutes (i.e. trips with an overtime fee) made by annual subcribers. The dashed vertical red line marks the price change.

restrict the sample to a single type of consumer, prices are uniquely determined by trip duration. In addition, the inclusion of origin-destination fixed effects implies that most credible confounders, such as hills between i and j, which may reduce demand, are controlled for. Similarly, period fixed effects allow us to control for any peaks in demand at rush hours. This makes it plausible that price is exogenous in the above specification. I compute a standard error using 200 bootstrap replications to estimate $m_{i,j,c}^t$, $\pi_{i,j,c}^t$ and v_i^t and then recovering the corresponding β_1 elasticity parameter. Results from Equation 10 are presented in Table 3.

Table 3 – Elasticity Estimation Results

	$\hat{\beta}_1$
Estimate	0.007
Standard Error	0.005
Bootstrap Replications	200
Fixed Effects	Period, Origin-Destination

Notes. Estimation result for specification described in Equation 10. Standard error computed by bootstrap on estimated matches, supply and prices.

3.3. Discussion

The results presented in Table 3 would suggest that demand for CitiBike is strongly inelastic. Note that the estimation process is conducted on a restricted sample of users, which may not entirely correspond to the population of commuters. In particular, since trip-specific prices become positive only after 45 minutes, the process estimates the elasticity for users that have particularly long travel times. This could be an issue if annual subscribers with long travel times use CitiBikes for purposes other than commuting.

The particularly high estimate is also likely to reflect a selection issue. $\lambda_{i,j}^t$ is a function of $m^t_{i,j}$, which is often zero [14]. Note that the assumed form of the demand function as implied by Equation 7 is inconsistent with the possibility that no matches occur, since the form of the matching function (see Equation 9) implies that when no matches occur, there is no consumer demand. The assumed form of the demand function is an approximation, and it is likely that for an origin-destination pair there does exist a price where demand would become zero. However, when $m^t_{i,j} = 0$, I cannot observe $\pi^t_{i,j}$, since this price is taken to be the average of price trips from i to j at t. As a result, any observations where $m^t_{i,j} = 0$ cannot be included in the estimation process. If one were to assume that the price was not period specific, one could impute the value of $\pi_{i,j}$ for all periods. This would allow observations where $m^t_{i,j} = 0$ to be included in the estimation process. This assumption would be fairly reasonable: the duration of trips between stations entirely determines price, and considering the day is restricted to 6am-9pm, it is unlikely these trip durations change significantly over the day. However, this assumption would not be sufficient, as we would have to additionally assume that prices are similar when $m^t_{i,j} = 0$ and $m^t_{i,j} \neq 0$, the latter being the observed price values. This assumption is unlikely to hold (unless the demand elasticity parameter were positive), and it is much more likely that prices when there are no matches are greater than when there are matches. If the latter occurs, then the $\pi^t_{i,j}$ value used in the estimation process are lower than the true values, which could bias the estimate of β_1 presented in Table 3 upward.

It is also important to recall that the sample on which the estimation process is conducted is a very particular sample of the population of matches. Since I only consider matches for annual subscribers whose trip has a positive price, i.e. whose trip lasts for more than 45 minutes, it is likely that the sample is not representative of the population. In particular, it appears that the bike usage for these consumers is quite different from the original target population of commuters, who have multiple possible substitutes for bikes in their use case. This is suggested by the relatively high

14. There often are no trips between two locations between i and j at period t

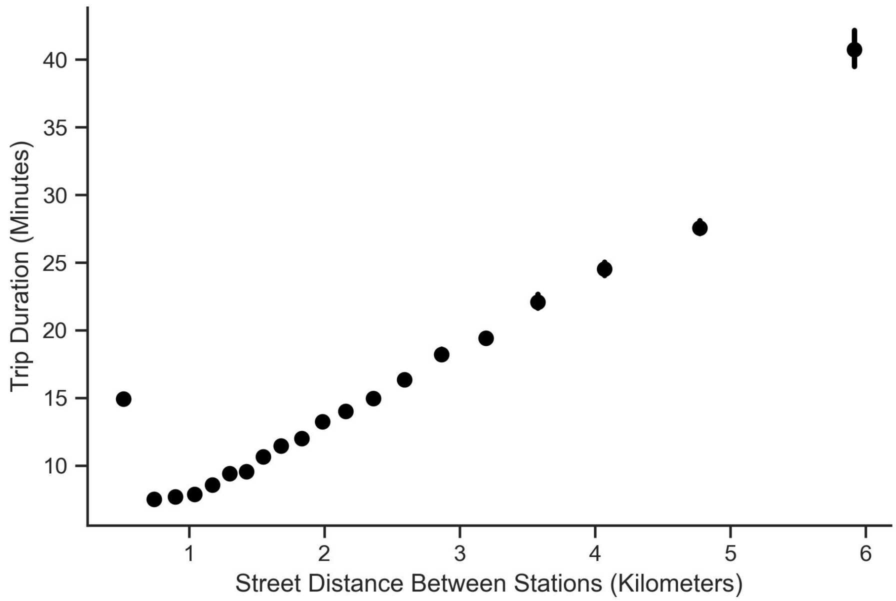
proportion of trips that start and end in the same location. When we consider the sample of trips over 45 minutes, such trips account for approximately 12% of trips, while they only account for 2% of trips in the full population of trips.

I generalize this observation in Figure 7, which provides plots of trip duration against the distance between start and end docks for both the entire population and the sample of interest. The distance between docks is taken to be the distance of the shortest route rideable by a bike between the two docks. The graphs suggest that besides the mechanical difference in trip duration levels between the sample and population, the relationship between trip distance and trip duration is also different. In particular, in the population, the relationship is increasing, as one would expect. In the restricted sample, the relationship is decreasing, which suggests that many trips between nearby stations do not employ the shortest path between docks. This may in turn suggest that these trips may have purposes other than direct transportation between points. In particular, it could be that these users' usage of CitiBikes is a more recreational one, where there are fewer direct substitutes to CitiBike.

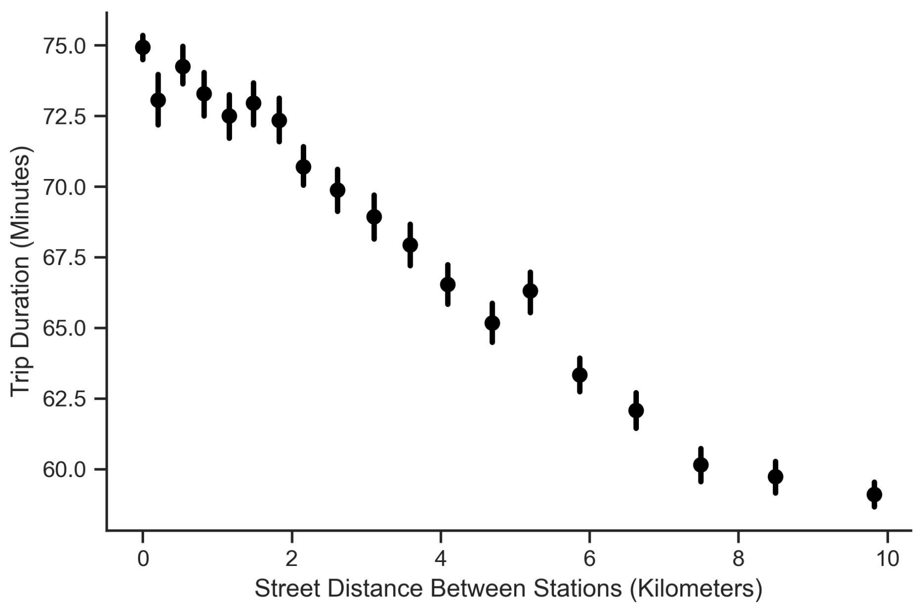
As a result, the estimate presented in Table 3 should be taken with a grain of salt. Given the sample composition and the possible selection problems, it is likely that the estimate constitutes an upper bound for the elasticity of demand, rather than an accurate estimate of the elasticity in the population of CitiBike users. In light of this, it is difficult to formulate any policy recommendations regarding possible subsidies to incentivize certain types of bike sharing usage. Given the pricing structure in bike-sharing, where most trips are free, it is likely that accurately estimating elasticities would require an experiment similar to that planned by Christensen et al. (2021) on ride-sharing markets.

Figure 7 – Trip Duration Against Distance

(a) Population



(b) Trips Above 45 Minutes



Notes. These graphs plot trip duration against the street distance between the start and end stations of each trip. The street distance is computed by taking the shortest bike-rideable route between the start station and the end station and taking the distance of this route. The plots are binned into 20 equal size groups based on street distance.

4. Conclusion

This paper implements a differences-in-differences design to estimate the relationship of bike sharing and subway usage in NYC and models the CitiBike bike-sharing market as a spatial matching market to recover demand values from supply and match information and estimate demand elasticity for bike sharing trips. Firstly, it finds that connecting a subway station to the bike sharing network has little effect on subway ridership in Manhattan and a small positive effect in other boroughs. Secondly, it finds a price elasticity of demand that is not significantly different from zero on a sample of CitiBike users. However, the sample may not be representative and this estimate likely constitutes an upper bound for the population elasticity. These results suggest that bike sharing policies are unlikely to be effective to influence subway usage but do not make it possible to accurately predict the effect of subsidies on bike usage.

References

Boeing, Geoff. 2017. "OSMnx: New Methods for Acquiring, Constructing, Analyzing, and Visualizing Complex Street Networks." *Computers, Environment and Urban Systems* 65 126–139. 10.1016/j.compenvurbsys.2017.05.004.

Brancaccio, Giulia, Myrto Kalouptsidi, and Theodore Papageorgiou. 2020a. "A Guide to Estimating Matching Functions in Spatial Models." *International Journal of Industrial Organization* 70 102533. 10.1016/j.ijindorg.2019.102533.

Brancaccio, Giulia, Myrto Kalouptsidi, Theodore Papageorgiou, and Nicola Rosaia. 2020b. "Search Frictions and Efficiency in Decentralized Transportation Markets."Technical Report w27300, National Bureau of Economic Research, Cambridge, MA. 10.3386/w27300.

Buchholz, Nicholas. 2021. "Spatial Equilibrium, Search Frictions, and Dynamic Efficiency in the Taxi Industry." *The Review of Economic Studies* rdab050. 10.1093/restud/rdab050.

Butters, Gerard R. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *Review of Economic Studies* 44 (3): 465–491.

Christensen, Peter, Lewis Lehe, and Adam Osman. 2021. "The Last Mile Problem – A Grand Transportation Challenge." conference Presentation, NBER.

Chu, Junhong, Yige Duan, Xianling Yang, and Li Wang. 2021. "The Last Mile Matters: Impact of Dockless Bike Sharing on Subway Housing Price Premium." *Management Science* 67 (1): 297–316. 10.1287/mnsc.2019.3550.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2020a. "Difference-in-Differences Estimators of Intertemporal Treatment Effects."

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020b. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–2996. 10.1257/aer.20181169.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2021. "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey."

Fan, Yichun, and Siqi Zheng. 2020. "Dockless Bike Sharing Alleviates Road Congestion by Complementing Subway Travel: Evidence from Beijing." *Cities* 107 102895. 10.1016/j.cities.2020.102895.

Forman, Adam. 2017. "The Human Cost of Subway Delays: A Survey of New York City Riders."Technical report, New York City Comptroller.

Fréchette, Guillaume R., Alessandro Lizzeri, and Tobias Salz. 2019. "Frictions in a Competitive, Regulated Market: Evidence from Taxis." *American Economic Review* 109 (8): 2954–2992. 10.1257/aer.20161720.

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–277. 10.1016/j.jeconom.2021.03.014.

Hall, Jonathan D., Craig Palsson, and Joseph Price. 2018. "Is Uber a Substitute or Complement for Public Transit?" *Journal of Urban Economics* 108 36–50. 10.1016/j.jue.2018.09.003.

Hall, Robert E. 1979. "A Theory of the Natural Unemployment Rate and the Duration of Employment." *Journal of Monetary Economics* 5 (2): 153–169. 10.1016/0304-3932(79)90001-1.

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the

Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431.

Rosaia, Nicola. 2020. "Competing Platforms and Transport Equilibrium: Evidence from New York City." 77.

Roth, Jonathan, Pedro H C Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." 54.

Treffeisen, Alan. 2017. ""We Are Being Held Momentarily:" How Much Time and Money Are New York City Subway Riders Losing to Delays?"Technical report, New York City Independent Budget Office.

Appendices

A. Subway Summary Statistics

Table 4 – Mean Daily Subway Entries By Station

Month	Bronx	Brooklyn	Manhattan	Queens
1	5661	6263	20533	8148
2	5998	6623	21699	8561
3	6271	6889	22625	8937
4	6198	6790	22874	8940
5	6372	6906	22849	9047
6	6261	6875	23188	8981
7	5754	6325	21576	8375
8	5713	6268	21390	8439
9	6205	6764	22474	8946
10	6363	6949	23447	9068
11	6145	6646	22235	8756
12	5895	6439	21942	8507

Table 5 – Mean Daily Subway Exits By Station

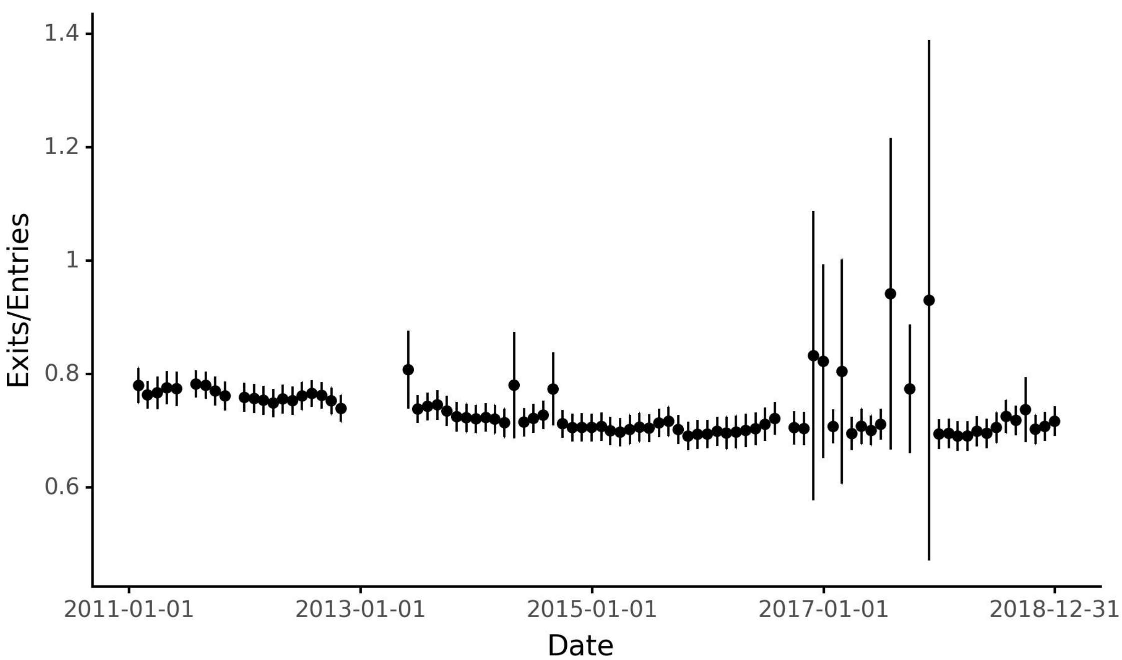
Month	Bronx	Brooklyn	Manhattan	Queens
1	3475	4383	17115	6145
2	3656	4592	17995	6435
3	3818	4769	18735	6678
4	3809	4733	19005	6738
5	3912	4811	18929	6810
6	3900	4860	19332	6821
7	3661	4533	18127	6437
8	3643	4497	17976	6453
9	3855	4746	18628	6756
10	3890	4799	19354	6793
11	3748	4576	18378	6527
12	3589	4449	18242	6335

B. Exits Instead of Entries

In the first part of the thesis, where I attempt to determine whether bike sharing and the subway are substitutes or complements, my empirical strategy estimates the effect of connecting a subway station to the subway on the number of entries into that subway station. This measures complementarity and substitution at the start of trips: either where the user takes the bike or where they drop their bike and take the subway. It would have been possible to focus on subway station exits rather than entries. Then, one would have looked at the end of trips, either where users drop off their bikes or where they exit the subway to take a bike. The interpretation would have been the same. If more users exit the subway in stations where there is a bike dock nearby, their trips are likely making use of both the subway and bike sharing. In a sense, this approach would have more directly targeted the last-mile problem instead of the first mile problem.

I focus on subway entries rather than exits for multiple reasons. Firstly, the data on entries is much more reliable than data on exits. Entering a subway station must be done through a turnstile, with very few exceptions (people with certain disabilities, people with a stroller or a bike, etc.). As a result, nearly every single subway user that pays for their trip is accounted for by the turnstile data. On the other hand, users may exit stations through turnstiles or through revolving doors, which do not record the number of users who go through them. This makes the exit data prone to systematic undercounting of users and more prone to noise. Secondly, the exit data closely mirrors the entry data. In essentially all stations, the trends in exit data closely follow that of the entry data, with differences remaining approximately constant over time. In Figure 8, I plot the average ratio between exits and entries over time, which is lower than 1 and approximately constant over time. Lastly, I provide event study plots analogous to those presented in Figure 3 where the outcome variable is exits rather than entries in Figure 9 and find similar results.

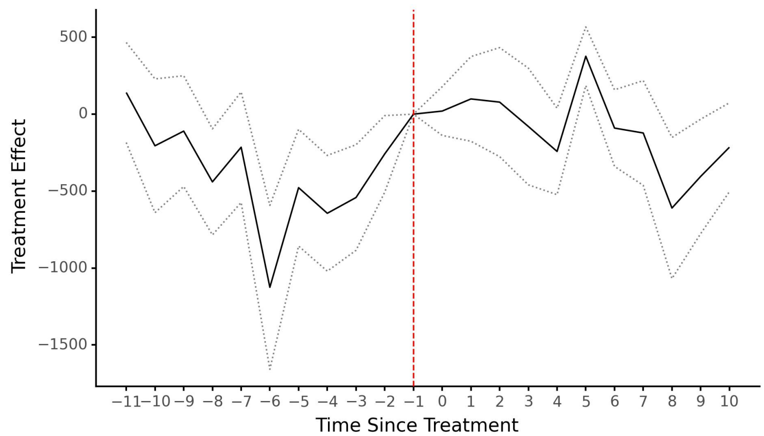
Figure 8 – Ratio of Entries to Exits Over Time



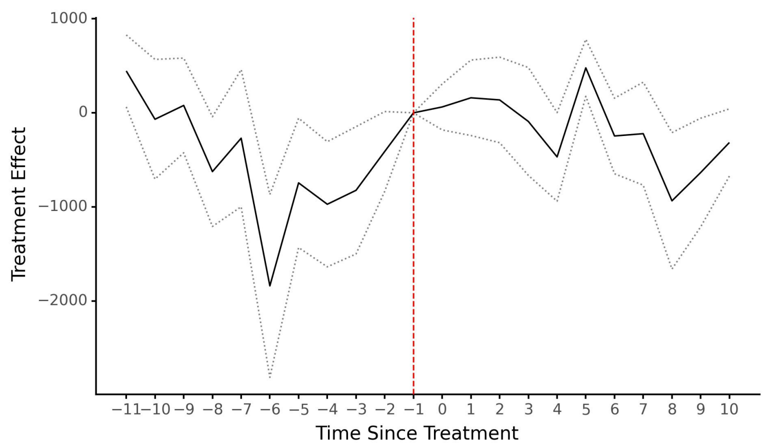
Notes. This plot provides a visualization of the mean across stations of the ratio of exits to entries each day. The corresponding 95% confidence intervals are constructed with a normal approximation.

Figure 9 – DiD Event Study Plots On Exits

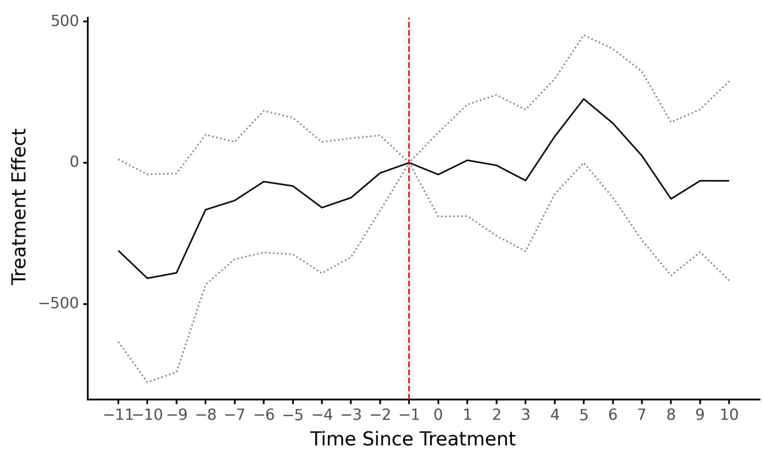
(a) All Boroughs



(b) Manhattan Only



(c) Bronx, Brooklyn, Queens



Notes. Standard errors are estimated with 50 bootstrap replications and are clustered at the subway station level. When multiple boroughs are included, borough specific trends are accounted for non-parametrically.