Substituting bikeshare for taxis: Evidence from the launch of the Citibike program in New York City

Alejandro Molnar Vanderbilt University Francis Ratsimbazafy Vanderbilt University

Note: uses data as of Jan 2016 (draft available on request: <u>alejandro.i.molnar@vanderbilt.edu</u>)

### Bikeshare is a rapidly expanding form of transportation

- Cities worldwide are rolling out "3<sup>rd</sup> generation" bikeshare.
  - 980 cities in 2016, up from 450 in 2013 (MetroBike; O'Brien, 2014). 1.2M bikes.
  - In the US: 5 systems launched in 2010, about 40 currently and 15 in deployment.

- New York City launched Citibike on May 27, 2013.
  - Designed for high density, high availability within zone.
  - Designed for, and used by commuters:
    - 89.0% rides are from annual members.
    - Annual members exhibit commuter patterns.
  - Footprint did not change from launch to summer 2015.
  - Second phase expansion ongoing, started summer 2015.

What has been the impact of Citibike on commuters and transportation alternatives, in particular taxis?

- Evaluations of bikeshare as transportation alternative, for planning and cost-benefit analysis.
  - Citibike is privately funded and operated.
  - Increasing demand for service to outer boroughs, calls for subsidies.
- Ex-ante evaluation of the economic impact of transportation infrastructure relies on credible estimates of substitution effects. (Redding, Ahlfeld and Sturm, 2015)
- Fact: declining market value of taxi medallions
  - Long run impact on "asset" price:
    - Uncertainty over long run, discounted licensing rents
  - Short run impact on driver revenue:
    - Citibike.
    - Boro taxis. Ongoing project w/ Daniel Mangrum @ Vanderbilt

### Results

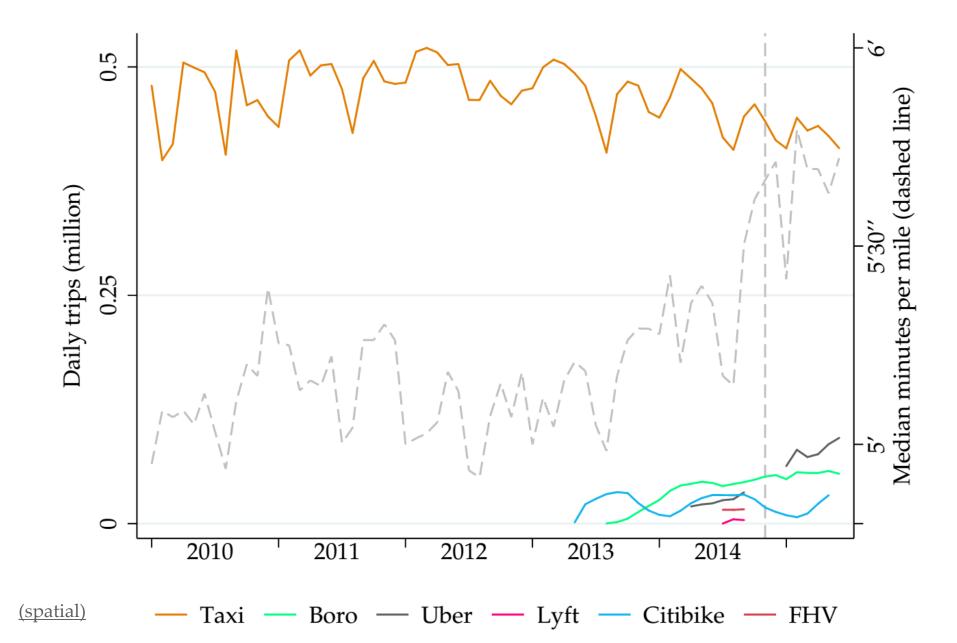
- Passenger substitution of taxi service.
  - Long-run substitution: Estimates in range of 3-4% decline in taxi trips due to roll-out of Citibike
    - Back-of-the envelope: 288-385 taxis.
  - Evidence of a short-run elasticity from outages of Citibike stations.

- Distributional impact on drivers.
  - Revenues for most impacted drivers decline by 5-6%
  - Somewhere between 50%-100% due to differential attrition.
- Driver responses (in progress)

### Recent changes to urban transportation in NYC

- Citibike: launched on May 27, 2013.
  - Second phase expansion ongoing, started summer 2015
- Boro medallion taxi (green taxi): launched on August 9, 2013.
  - Medallion taxis regulated by NYC TLC, banned from pickups below W 110 St and E 96 St in Manhattan + airports.
  - Medallion restriction is a "second best" approach to provision of street-hail service in outer boros.
- Growth of TNCs (e.g. Uber, Lyft) beyond black-car/livery.
- Subway: first additional station in >25 years
  No. 7 line Hudson Yards station opened Sep 13, 2015.
- Infrastructure: 24.2 miles of protected bike lanes since 2013.

### Trips by NYC medallion taxis and the competition



### Data

#### Taxi data

- Taxi and Limousine Commission TPEP trip records, obtained by FOIL. From 2009 to June 2015.
  - GPS coordinates and exact time of pickup and drop-off.
    - About 170 million records per year.
    - Fares, tips, number of passengers, tolls, means of payment.
  - For (only) 2009 and 2013, we also have matched medallion and hack license (i.e. driver) IDs.

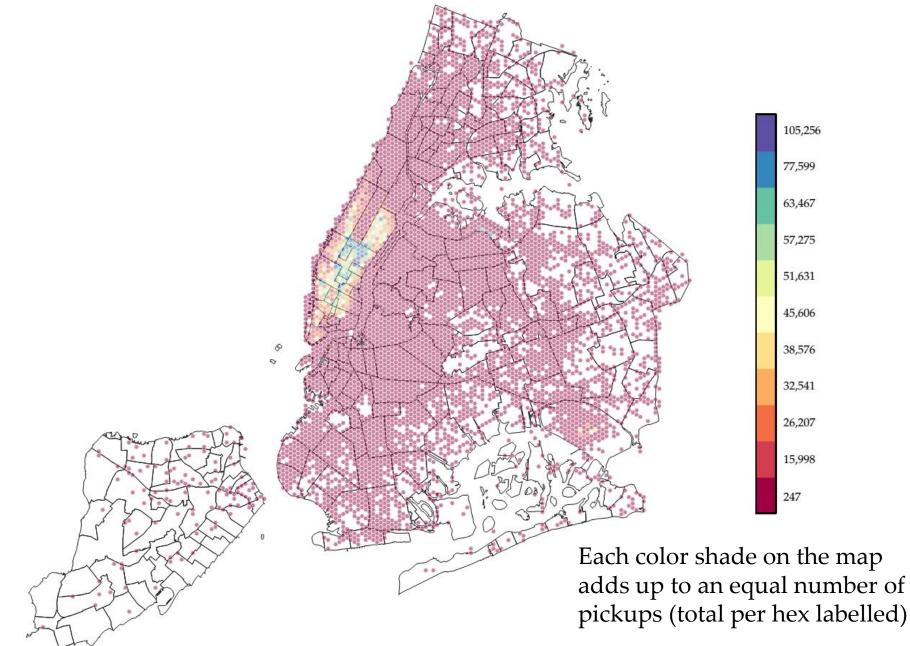
#### Citibike data

Station locations.

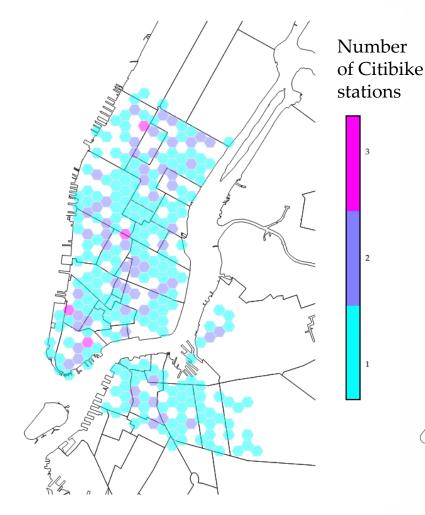
Bike trip records since July 2013. (used to reconstruct station inactivity). Since March 2015: station status in 5 min intervals (online/offline status, bikes docked).

**Other** related data: Travel times by mode from Google Distance API.

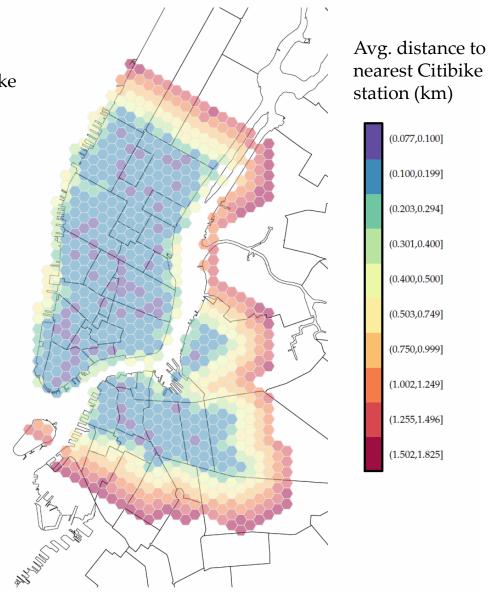
### Spatial distribution of taxi pickups. October 2013.



### Citibike roll-out. Station counts and distance to station May 27, 2013



Area of each hexagonal cell: 0.081 sq. km (19.9 acres, 0.03 sq. miles) Width is ~3 short or 1 long Manhattan blocks For 5 boros of NYC: 12,205 cells in total.



### Descriptive stats: a typical driver/day in May 2013

	mean	sd	min	max
Number of trips	20	9	1	105
Trip duration	265	112	0	1,384
Total trip distance (miles)	59	26	0	425
Total revenue	262	104	3	1,226

Table 1: Descriptive statistics by driver-day in May 2013

Table 2: Descriptive statistics by day, May 2013

	mean	sd	min	max
Medallions on road	12,935	310	12,075	13,174
Licenses on road	24,177	1,103	20,583	25,104
Total revenue	6,333,694	604,745	4,446,391	7,109,536
Number of trips	491,522	46,102	334,906	552 <i>,</i> 914
Observations	31			

### Estimation strategy: substitution of taxi trips

- Subdivide NYC into 12K hexagonal cells.
  - Count trip flows between origin and destination cells
  - Pairs of cells can be treated (Citibike trip possible) or not.
  - Launch occurred sharply on Monday, May 27, 2013.
- Differences in differences:

 $\log trips_{odm} = \beta treat \times post + \delta_m + \delta_{od} + \varepsilon_{odm}$ 

For  $m = \{Jan 2010, \dots, Jun 2015\}$  and a subset of od-pairs.

- Alternative definitions of treatment: count vs. distance.
- Alternative control pairs
  - All od-pairs
  - od-pairs originating in CB zone (controls for driver response)
- Interactions with shifters of utility for bike vs. car

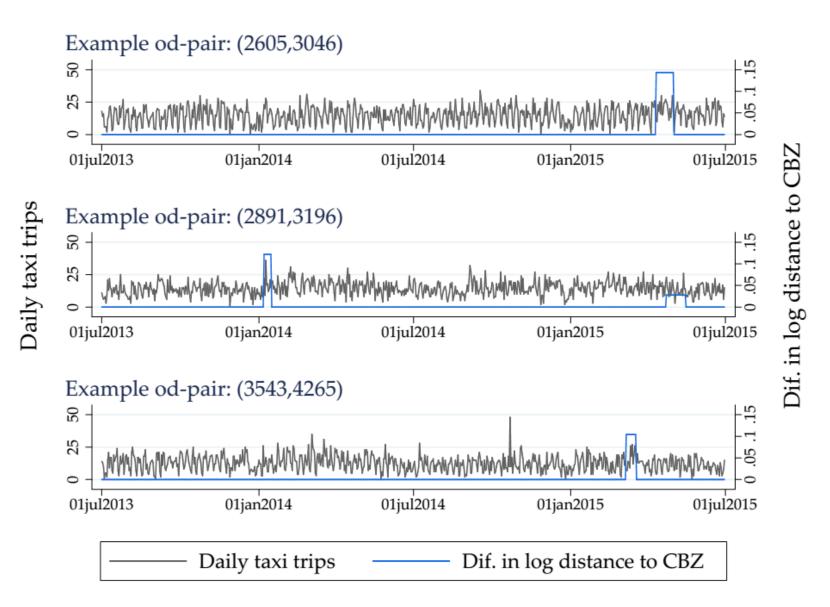
### Results: trip flows

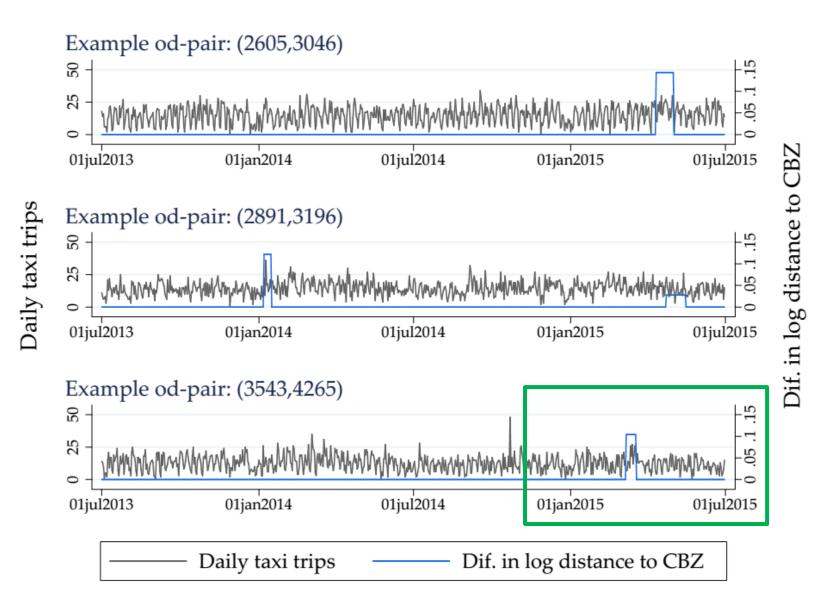
Dep. var: $log(trips_{odym})$	(1)	(2)	(3)	(4)	(5)
O&D cells contain CB station	-0.049***				
CB with avg. walk $\leq 2 \rm km$	[0.001]	$-0.043^{***}$ [0.000]	-0.069*** [0.000]		
CB with avg. walk $[0, 0.5]~\mathrm{km}$				-0.058*** [0.000]	
CB with avg. walk (0.5,1] km				-0.025*** [0.001]	
CB with avg. walk $(1,1.5]$ km				-0.025*** [0.001]	
CB with avg. walk (1.5,2] km $$				-0.018*** [0.001]	
CB with avg. walk $\leq$ 2km				[0.001]	-0.043*** [0.000]
$\times$ 1 Qtl bike/car travel time					-0.065*** [0.002]
$\times$ 2 Qtl bike/car travel time					-0.038*** [0.002]
$\times$ 3 Qtl bike/car travel time					-0.030***
$\times$ 4 Qtl bike/car travel time					[0.002] -0.032*** [0.002]
Origin-Destination f.e.	Y	Y	Y	Y	[0.002] Y
Month-year f.e.	Y	Y	Y	Y	Υ
$R^2$	0.92	0.91	0.92	0.91	0.91
N N of OD poirs	17,548,369	30,589,606 762,172	25,045,801	30,589,606 762,172	30,589,606 762,172
N of OD-pairs	423,606	763,172	604,889	763,172	763,172

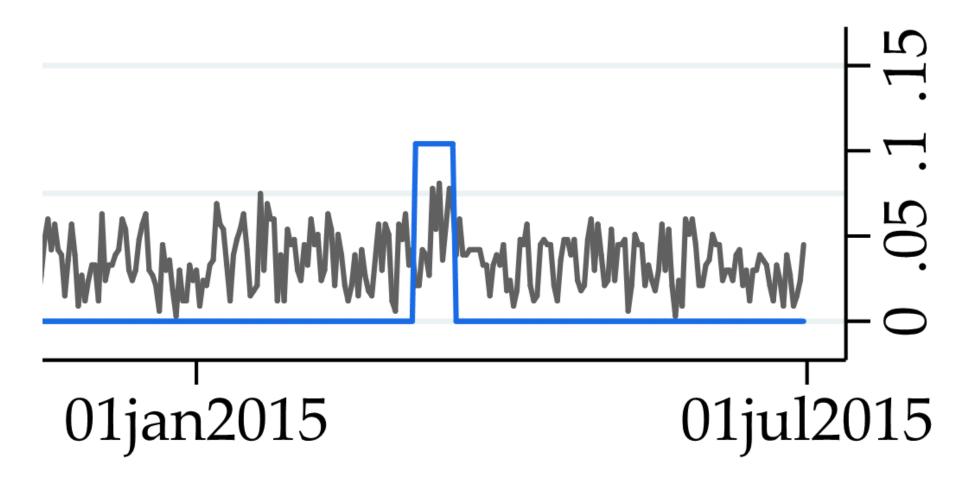
12

### Proximity to Citibike varies with station outages and relocations

- Citibike system footprint remained constant from launch through summer 2015.
- Down to the city block level, however, there has been variation in the proximity/convenience of Citibike:
  - Small number of station moves
  - Offline status (station maintenance, street repaving).
- In the data: 135 alterative "spatial configurations" of the Citibike station system from July 2013 to June 2015.
- We compute a cell's "average distance to closest Citibike station" under each alternative configuration.



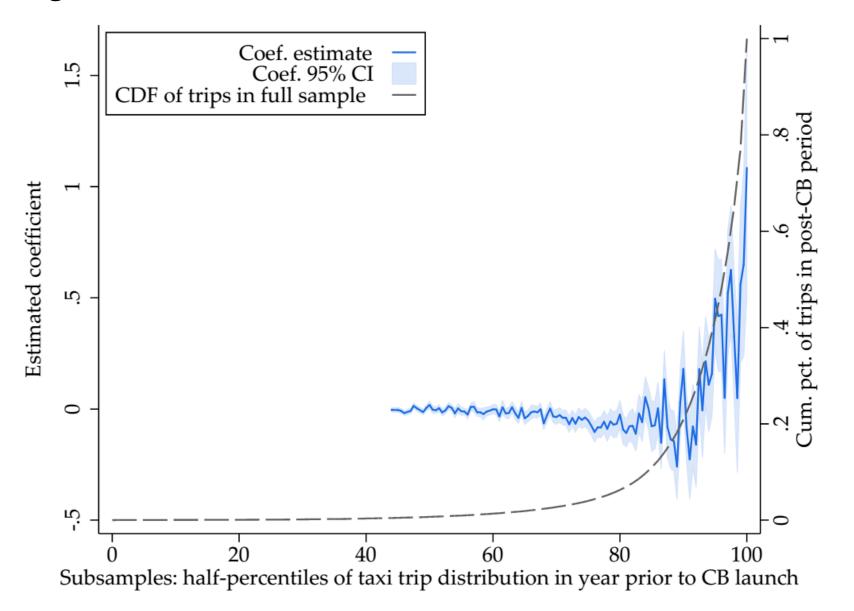




	(1)	(0)	(2)	(4)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta$ log avg. distance to CBZ	-0.005	0.008	$0.058^{*}$	0.044	0.078	0.496***	0.486***
	[0.027]	[0.014]	[0.027]	[0.111]	[0.060]	[0.109]	[0.109]
Date f.e.	Y	Y	Y	Y	Y	Y	Y
(od)-pair f.e.	Y		Y	Y		Y	Y
(od)-pair and day-of-week f.e.		Y			Y		
o-trends and d-trends			Y			Y	
(od)-pair trends							Y
$R^2$	0.81	0.86	0.81	0.72	0.81	0.72	0.72
Ν	30,518,380	30,518,380	30,518,380	6,505,030	6,505,030	6,505,030	6,505,030
Percentile cutoff for sample (%)	25	25	25	5	5	5	5
Sample of trips pct.	96.0%	96.0%	96.0%	58.3%	58.3%	58.3%	58.3%
N of od-pairs	41,806	41,806	41,806	8,911	8,911	8,911	8,911

Table 3: Effect of Citibike station outage shocks. Dependent variable: daily trips per (od)-pair

Note: Standard errors clustered at the level of the f.e. unit. Sample restrictions are described in bottom three rows: estimates in the first three columns correspond to the top quartile of (od) pairs in the full sample (ranked by trips in the year prior to Citibike launch). These pairs account for 96.0% of all taxi trips within the CBZ in the post period. Estimates in the last four columns correspond to the top 5% of (od) pairs in the full sample, and account for 58.3% of all taxi trips. Results by half-percentile subsamples with fixed effect and a linear trend by (od)-pair and day-of-week are plotted in Figure 6.



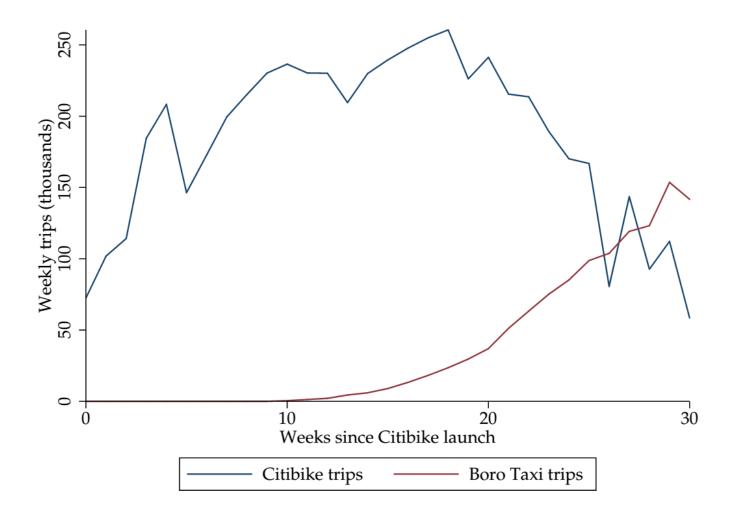
### Impact on driver revenue

- Exploit pre-existing heterogeneity in drivers.
  - Driver A lives in Jersey City, across tunnel from downtown Manhattan. Starts day with hails in CB zone.
  - Driver B lives in the Bronx. Starts day with hails in the Bronx.
  - Driver C lives in Long Island. Starts day at JFK.
- Define a shift start as first trip that follows at least 6 hours of inactivity by same driver (as in Farber, QJE 2015).
- For each driver, calculate median distance of first trip in each shift to closest Citibike station.
  - For  $1^{st}$  decile, it's < 200 m
  - For 10<sup>th</sup> decile, it's > 10 km (distribution)

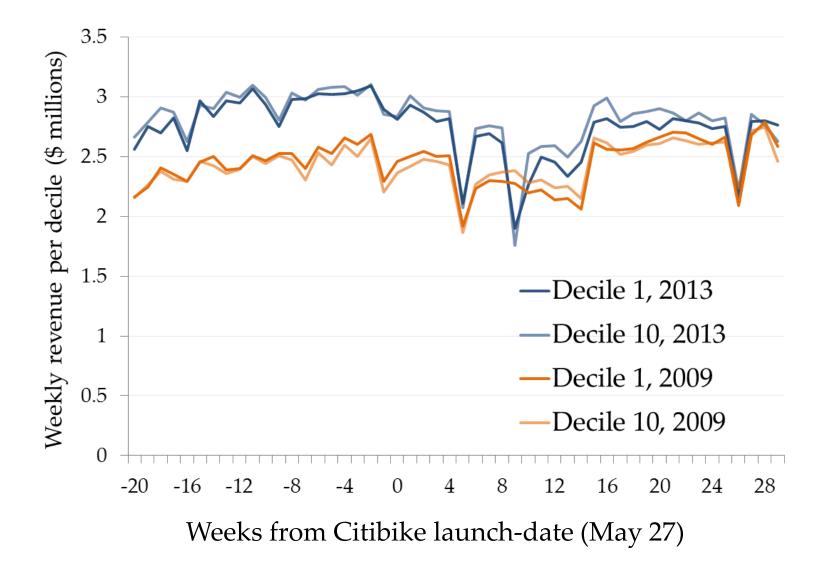
Data on driver IDs is available only for 2009 and 2013.

### Specification

- Decile x normalized Citibike trips, by week from Citibike launch
- Decile x normalized Boro-taxi trips, by week from Citibike launch



## Revenue trends for drivers in 1<sup>st</sup> and 10<sup>th</sup> decile of distance to Citibike



### Drivers with location advantage in CB zone see revenue decline with roll-out, relative to baseline drivers with location furthest from CB zone

		(1)	(2)	(3)	(4)	(5)
	Decile $\times$ Norm. Citibike trips					
Observation:	1	$-61.076^{***}$	-53.738*	$-71.230^{***}$	-30.553**	-12.507
Driver-week in 2009 or 2013		[9.243]	[25.444]	[21.088]	[9.304]	[6.452]
	2	-51.659***	-26.154	-80.857***	-51.561***	-44.431***
		[9.362]	[27.836]	[20.660]	[8.838]	[6.111]
	3	$-40.351^{***}$	-48.944	-20.789	-14.096	-0.443
Dependent variable:		[8.621]	[28.223]	[19.111]	[8.205]	[6.223]
Weekly revenue (fare+surcharge)	4	$-51.076^{***}$	-58.598*	-28.096	11.097	7.738
•		[8.248]	[26.802]	[18.170]	[8.627]	[5.450]
	5	-43.842***	-47.527	-20.674	-1.912	0.044
		[9.491]	[26.006]	[19.495]	[8.709]	[5.327]
Drivers binned into deciles by	6	-46.949***	$-59.531^{*}$	-10.703	-0.929	22.794***
median distance to closest		[9.174]	[25.936]	[20.797]	[8.880]	[5.244]
Citibike station of shift first	7	$-41.455^{***}$	-41.517	$-65.646^{***}$	-22.287**	31.183***
		[8.388]	[28.374]	[18.068]	[8.100]	[6.138]
trips, Jan-May, 2013 (specs 1-4),	8	-14.280	-27.054	-36.576	-15.827	-12.178*
or Jan1-May, 2009 (spec 5).		[10.364]	[29.819]	[20.670]	[9.641]	[5.476]
<i>y</i> , (1)	9	-35.008***	-44.867	-18.575	-36.189***	0.244
		[9.021]	[29.211]	[19.364]	[9.845]	[5.980]
	Decile $\times$ Norm. Boro taxi trips					
Deciles are interacted with		(	omit	ted, see <u>he</u>	<u>ere</u> )	
weekly, system-wide Citibike					,	
	Driver f.e.	Y	Y	Y	Y	Y
(or Boro taxi) trips.	Week-year f.e.	Y		Y	Y	Y
	Decile-week f.e.		Y			
	Decile-year f.e. and trend			Y		0.40
S.e. clustered at week-decile	$R^2$	0.36	0.36	0.36	0.45	0.49
	N N of driver	2,342,912	2,342,912	2,342,912	1,798,732	1,666,288
(104 clusters)	N of drivers	22,528	22,528	22,528	34,591	32,044
						~~

#### Revenue decline decomposed into margins: robust effect on the extensive margin (differential attrition in shifts)

	D ''	D	C1 : 6	Revenue	01:6	Hours	Trips	Revenue	Minutes	Miles	Revenue
Specification	Decile	Revenue	$_{ m hours}^{ m Shift}$	per hour	Shifts	per shift	per hour	per trip	per trip	per trip	per mile
	1	-0.049	-0.027	-0.022	-0.032	0.006	0.047	-0.069	-0.049	-0.132	0.062
	2	-0.041	-0.020	-0.022	-0.027	0.007	0.049	-0.071	-0.052	-0.125	0.054
	3	-0.033	-0.012	-0.021	-0.017	0.005	0.045	-0.065	-0.046	-0.121	0.056
	4	-0.041	-0.017	-0.024	-0.020	0.003	0.039	-0.063	-0.048	-0.113	0.050
Week-year f.e.	5	-0.034	-0.010	-0.024	-0.017	0.007	0.047	-0.071	-0.051	-0.126	0.056
(1)	6	-0.037	-0.007	-0.030	-0.006	-0.000	0.035	-0.065	-0.038	-0.119	0.054
	7	-0.035	-0.002	-0.032	-0.003	0.000	0.035	-0.068	-0.048	-0.118	0.050
	8	-0.014	0.017	-0.030	0.019	-0.002	0.029	-0.059	-0.042	-0.108	0.049
	9	-0.030	-0.022	-0.008	-0.000	-0.022	0.027	-0.035	-0.038	-0.073	0.038
	1	-0.041	-0.011	-0.030	-0.016	0.006	0.058	-0.088	-0.064	-0.172	0.083
	2	-0.020	0.013	-0.032	0.007	0.006	0.059	-0.091	-0.069	-0.164	0.073
	3	-0.039	-0.006	-0.033	-0.009	0.003	0.055	-0.088	-0.064	-0.161	0.074
Decile-week	4	-0.045	-0.009	-0.036	-0.014	0.005	0.046	-0.082	-0.063	-0.151	0.069
and year f.e.	5	-0.035	-0.003	-0.031	-0.008	0.005	0.059	-0.090	-0.068	-0.165	0.075
(2)	6	-0.045	0.003	-0.048	0.003	-0.000	0.044	-0.092	-0.054	-0.164	0.072
	7	-0.035	0.008	-0.043	0.010	-0.002	0.052	-0.095	-0.067	-0.160	0.066
	8	-0.025	0.015	-0.041	0.022	-0.007	0.043	-0.084	-0.059	-0.149	0.065
	9	-0.039	-0.030	-0.009	0.005	-0.035	0.037	-0.046	-0.051	-0.098	0.052
	1	-0.061	-0.057	-0.003	-0.064	0.006	0.011	-0.015	0.005	-0.010	-0.005
	2	-0.063	-0.058	-0.004	-0.067	0.008	0.014	-0.018	0.001	-0.018	-0.000
Week-year f.e.,	3	-0.025	-0.020	-0.005	-0.035	0.015	0.012	-0.017	-0.001	-0.014	-0.004
decile-year f.e.	4	-0.034	-0.031	-0.002	-0.035	0.003	0.018	-0.020	-0.001	-0.010	-0.010
and	5	-0.028	-0.009	-0.018	-0.030	0.021	-0.004	-0.015	0.008	-0.006	-0.009
decile-year-trend	6	-0.022	-0.011	-0.010	-0.024	0.012	0.014	-0.024	-0.021	-0.006	-0.019
(3)	7	-0.055	-0.043	-0.013	-0.053	0.010	0.009	-0.021	-0.011	-0.013	-0.009
	8	-0.033	-0.013	-0.020	-0.019	0.006	0.005	-0.025	-0.014	-0.019	-0.006
	9	-0.016	-0.006	-0.010	-0.011	0.006	0.011	-0.021	-0.007	-0.016	-0.006

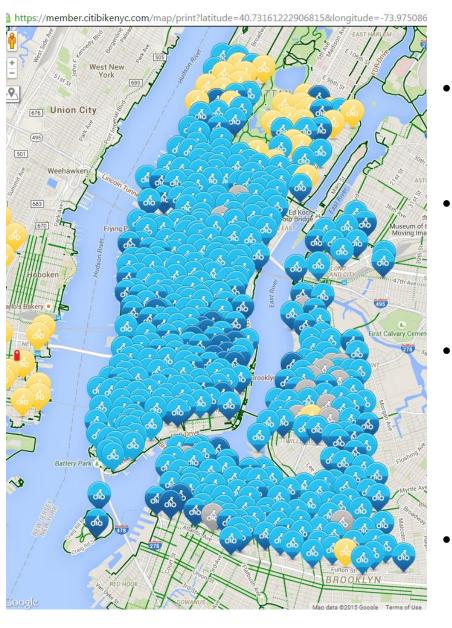
### Conclusions

- 17M Citibike trips from launch through June 2015.
  - If 55.1% of Citibike trips substitute taxis
    - 12,367 taxi trips substituted per day: 385 taxis
  - If 42.3% of Citibike trips substitute taxis
    - 9,513 taxi trips substituted per day: 288 taxis
- Alternative margins (walking/subway) not yet measured.
- Substitution is greater with:
  - Proximity to station
  - Availability of nearby station
  - Infrastructure that supports cycling (measured indirectly through travel times)
- Drivers operating in the Citibike zone have seen relative decline in revenue.

### Future directions

- Working on driver supply response
  - Changes to transition probability matrix when taxi is empty.
- Exploit station status at 5 minute frequency (since March 2015):
  - Finer measure of station inactivity periods
  - Station stock-out
- Substitution of walking and subway
  - NYC data on pedestrian counts
  - Subway station turnstile counts
- Quantify public transportation provision with a model of demand.

### Available now: data on additional roll-outs



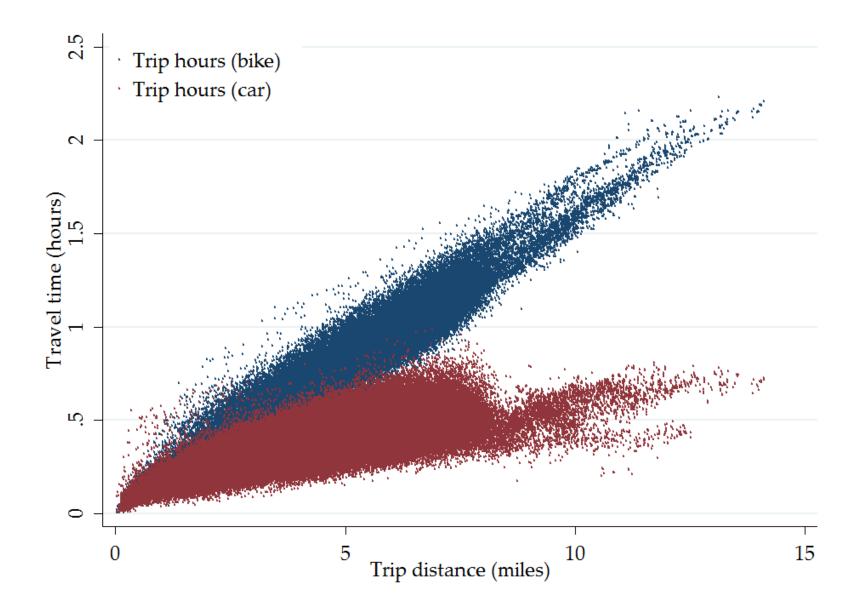
Brooklyn rolled out in summer 2015.

Upper East Side and Upper West Side rolled-out fall 2015.

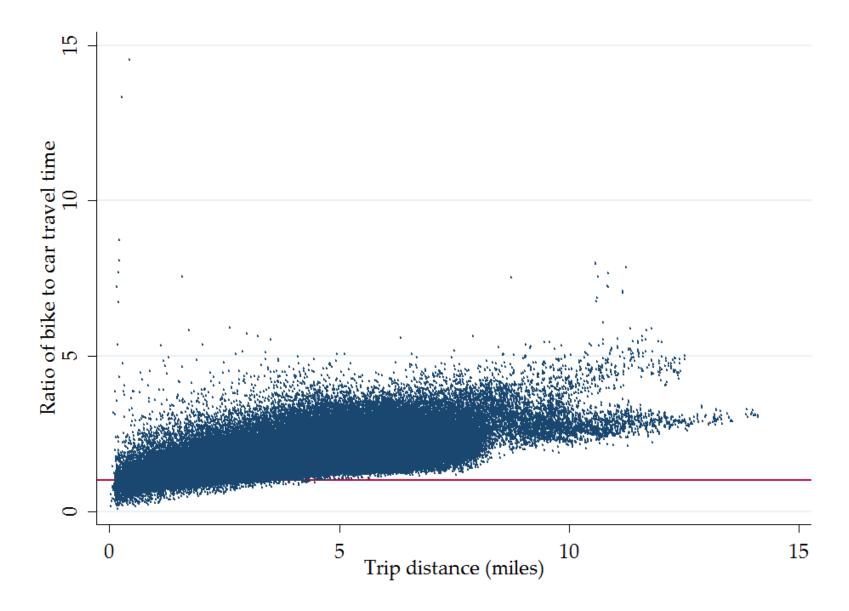
- Jersey City rolled-out fall 2015.Compatible membership, but travel between states not possible.Funded by Jersey City, not Citibike.
- Since: 3<sup>rd</sup> phase roll-out

#### Extra stuff

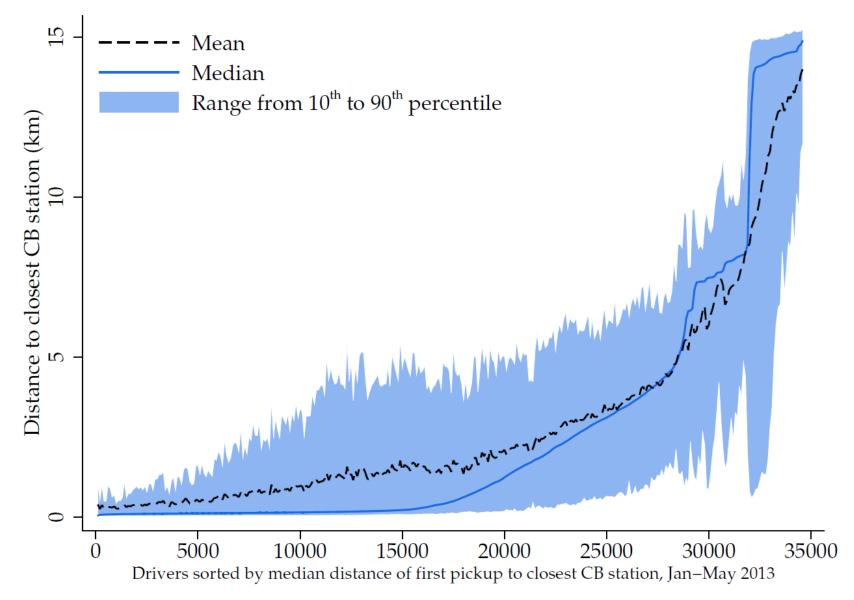
#### Google Distance Matrix API data Travel times collected since June 2015 at 8.30 am EST



#### Google Distance Matrix API data Travel times collected since June 2015 at 8.30 am EST



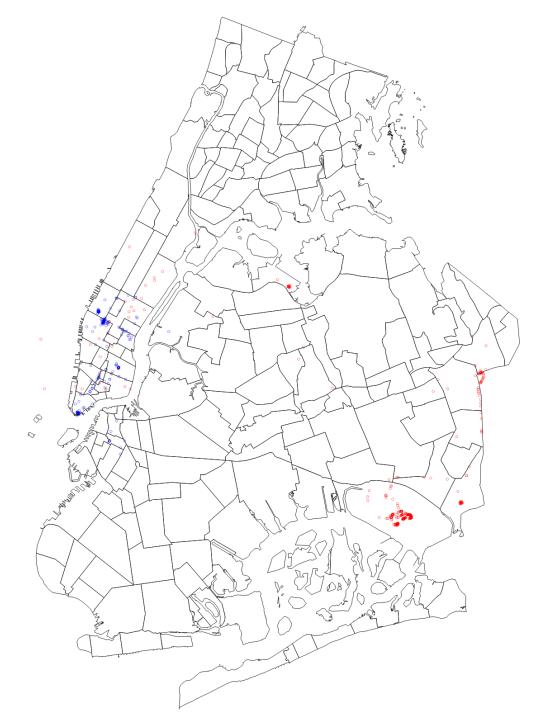
### Driver heterogeneity



(back)

# Driver starting locations

- Plotted: locations of first pickup in the shift for 10 drivers.
- In blue: top 5 closest drivers
- In red: bottom 5 furthest drivers (mostly airport pick-ups)



Drivers with pre-existing		(1)	(2)	(3)	(4)	(5)
Different pre existing	Decile $\times$ Norm. Citibike trips	01 050***	F0 700*	<b>71 000***</b>	00 550**	10 505
location advantage in	1	-61.076*** [9.243]	-53.738* [25.444]	-71.230*** [21.088]	-30.553** [9.304]	-12.507 [6.452]
e	2	-51.659***	-26.154	-80.857***	-51.561***	-44.431***
Citibike zone see		[9.362]	[27.836]	[20.660]	[8.838]	[6.111]
CHIDINC ZOIIC SEC	3	-40.351***	-48.944	-20.789	-14.096	-0.443
revenue decline with	4	[8.621] -51.076***	[28.223] -58.598*	[19.111] -28.096	[8.205] 11.097	[6.223] 7.738
	*	[8.248]	[26.802]	[18.170]	[8.627]	[5.450]
Citibike roll-out, relative	5	-43.842***	-47.527	-20.674	-1.912	0.044
CIUDINE IOII-OUL, IEIAUVE		[9.491]	[26.006]	[19.495]	[8.709]	[5.327]
to baseline drivers with	6	-46.949*** [9.174]	-59.531* [25.936]	-10.703 [20.797]	-0.929 [8.880]	22.794*** [5.244]
to baseline unvers with	7	-41.455***	[25.530] -41.517	-65.646***	-22.287**	[5.244] 31.183***
location advantage		[8.388]	[28.374]	[18.068]	[8.100]	[6.138]
location advantage	8	-14.280	-27.054	-36.576	-15.827	-12.178*
furthest from CBZ	0	[10.364]	[29.819]	[20.670]	[9.641]	[5.476]
Turmest from CDZ	9	-35.008*** [9.021]	-44.867 [29.211]	-18.575 [19.364]	-36.189*** [9.845]	0.244 [5.980]
	Decile $\times$ Norm. Boro taxi trips	[5:021]	[20:211]	[15.664]	[5:040]	[0.000]
	1	19.907	0.055	-19.094	31.824*	8.178
Observation:		[14.677]	[31.025]	[33.001]	[12.770]	[12.057]
Driver-week in 2009 or 2013	2	38.327** [13.638]	39.209 [22.847]	14.252	37.080**	-32.376**
Dirver week in 2009 of 2010	3	[13.030] 44.358***	[32.847] 61.097	[30.466] 48.756	[13.049] 51.409***	[11.318] -12.052
	C C	[11.524]	[32.241]	[27.987]	[10.440]	[10.052]
Dependent variable:	4	18.992	0.834	18.716	47.607***	18.269
Weekly revenue (fare+surcharge)	_	[10.910]	[30.501]	[27.225]	[11.949]	[9.568]
Weekly levellue (late-suicharge)	5	20.572 [13.273]	18.792 [30.523]	28.370 [27.457]	37.841** [11.727]	-8.127 [10.193]
	6	46.736***	46.880	72.848*	66.245***	7.454
Drivers binned into deciles by median		[11.872]	[31.085]	[30.083]	[10.945]	[10.139]
distance to closest Citibike station of	7	34.756**	15.975	-7.889	68.278***	6.304
	0	[11.601]	[31.142]	[27.806]	[10.873]	[9.525]
shift first trips, Jan1-May27, 2013 (specs	8	22.970 [14.855]	25.791 [30.917]	-16.329 [29.052]	18.195 [12.820]	3.714 [10.030]
1-4), or Jan1-May27 2009 (spec 5).	9	20.505	30.910	40.702	15.134	3.227
1 1), of juilt why 27 2007 (spec 0).		[12.051]	[33.386]	[28.521]	[11.945]	[9.810]
	Driver f.e. Week-year f.e.	Y Y	Y	Y Y	Y Y	Y Y
Interacted with weekly, system-wide	Decile-week f.e.	1	Y	1	1	1
Citibike (or Boro taxi) trips.	Decile-year f.e. and trend	<b>10 10 10</b>		Y		
Church (or boro taxi) trips.	$R^2$ N	0.36 2,342,912	0.36 2,342,912	0.36 2,342,912	0.45	0.49
	N of drivers	2,342,912 22,528	2,342,912 22,528	2,342,912 22,528	1,798,732 34,591	1,666,288 32,044
		,	,	,		

33