To: NYCDOT and Motivate

Subject: Determining Eviction Correlates: CitiBike

Date: June 20, 2018

Introduction

There is a huge crisis of eviction in the US today. In 2016 1 million people were evicted from their homes[1]. Gentrification, the shifting in demographics of a neighborhood towards higher income residents, is often associated with the removal of previous low-income tenants, by the setting of untenably high rents or other means. One mean of removing renters from their housing is through eviction[2].

Bike sharing is accessed much less by lower income populations, and the presence of bike sharing stations is thought to be correlated with the gentrification of a neighborhood[3]. According to a study done by Portland State University, 2 percent of low-income people of color surveyed had a bike-share membership[4]. Prices and safety concerns, among other factors were hypothesized to contribute to this discrepancy.

Our research did not find a clear connection between eviction and CitiBike stations for census tracts in the borough of Brooklyn in 2016, but there was more evidence for a connection with poverty rates and median household income. More research has to be done, expanding geographies, years, and increasing the precision of the statistical methods.

Methodology

To answer the question of how the amount of evictions and eviction filings in census tracts connected to the presence of a citibike station, we gathered data on station locations from CitiBike's own feed, and data on evictions from The Eviction Lab. We narrowed our analysis to Brooklyn due to the fact that it was the only borough that had at least one census tract with a CitiBike station, in addition to accessible data on evictions. We matched each CitiBike station to the census tract it was located within. Then we linked this to the table, from the Eviction Lab, containing columns for median household income, population, eviction filing counts, eviction counts, eviction rate, and poverty rate. Using these columns we further calculated eviction filing counts per person, eviction counts per person, and 3 binary variables that indicated if the census tract in consideration had no stations, one or more stations, or two or more stations.

Results

We found a slight decrease in eviction filing counts, eviction filings per person, eviction counts, and eviction counts per person with the increase in CitiBike stations (Fig 1 and Fig 2). However, eviction rate showed no pattern (Fig 3). Clearer variations were seen in poverty rate and median household

---

[1] https://www.curbed.com/2018/4/9/17216054/eviction-records-data-matthew-desmond-princeton
[2] https://www.wsws.org/en/articles/2018/01/06/evic-j06.html
[3] https://www.citylab.com/equity/2017/07/what-keeps-bike-share-white/533412/
[4] https://trec.pdx.edu/research/project/884/Evaluating_Efforts_to_Improve_the_Equity_of_Bike_Share_Systems

income, with poverty rate decreasing and median household income increasing
with each progressively larger bin of CitiBike counts (Fig 4 and Fig 5).

Figures
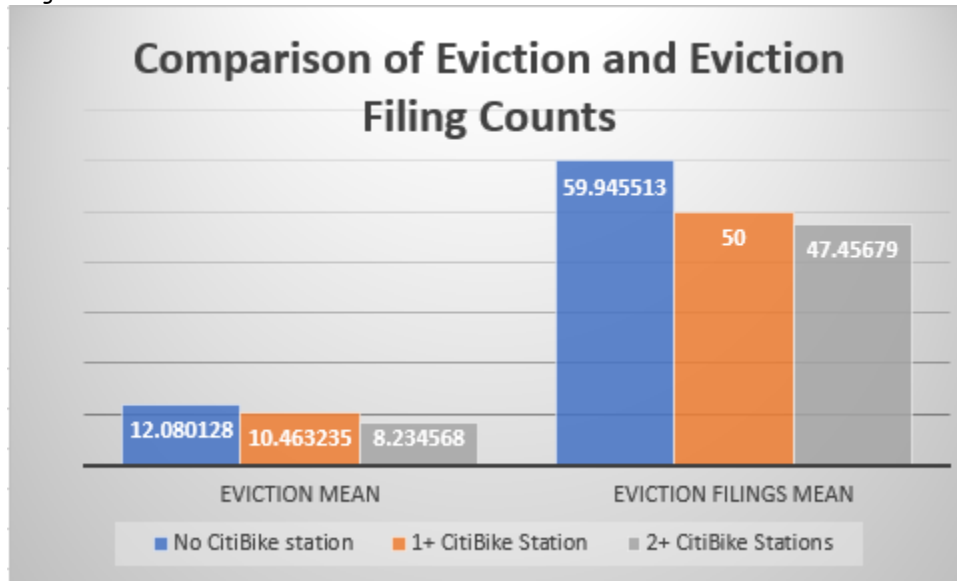(Sources: Eviction Lab, US Census Bureau, CitiBike)

Fig 1

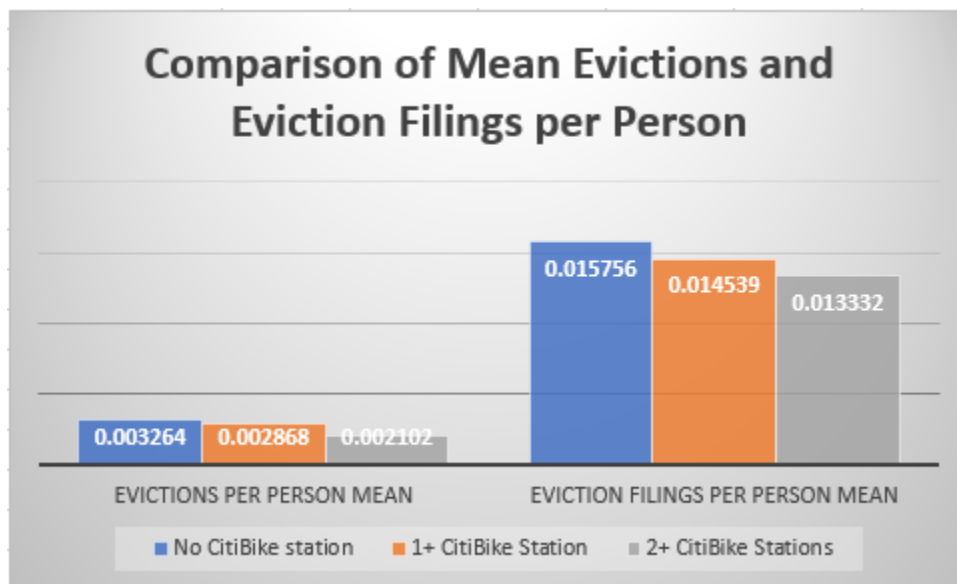Comparison of Eviction and Eviction Filing Counts

Fig 2

Comparison of Mean Evictions and Eviction Filings per Person

Fig 3

Comparison of Mean Eviction Rate

- No CitiBike station: 1.412051
- 1+ CitiBike Station: 2.157901
- 2+ CitiBike Stations: 1.755

MEAN EVICTION RATE

Fig 4

Comparison of Mean Poverty Rate

- No CitiBike station: 18.744984
- 1+ CitiBike Station: 14.377426
- 2+ CitiBike Stations: 12.053086

POVERTY RATE MEAN

Fig 5

Comparison of Mean Median Household Income ($)

No CitiBike station: 50148.71173
1+ CitiBike Station: 73125.71642
2+ CitiBike Stations: 80840.29114

MEDIAN HOUSEHOLD INCOME MEAN

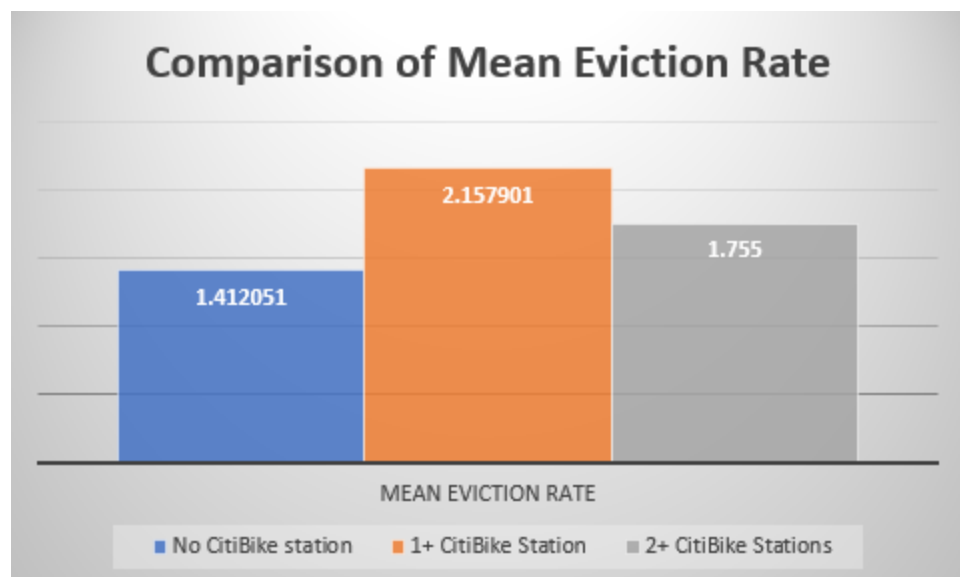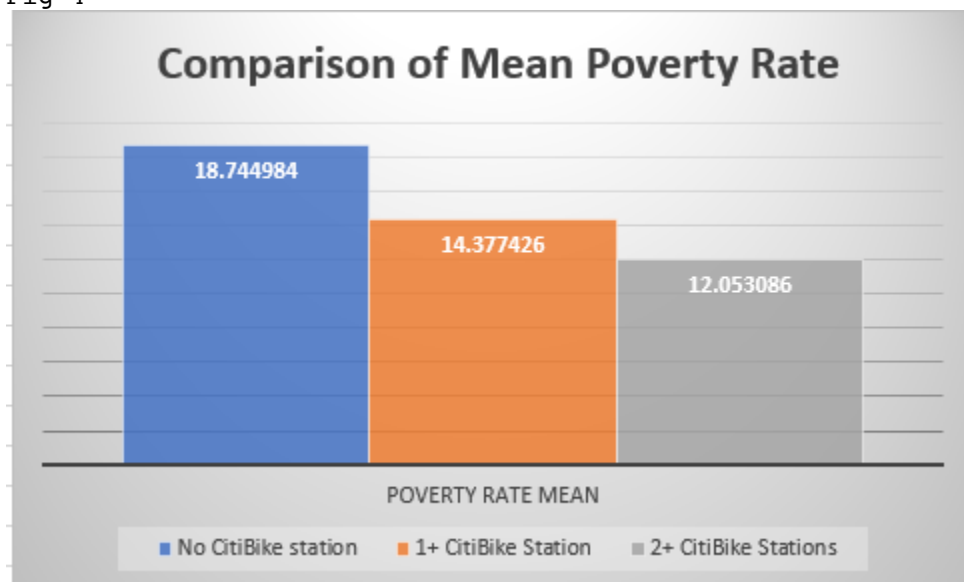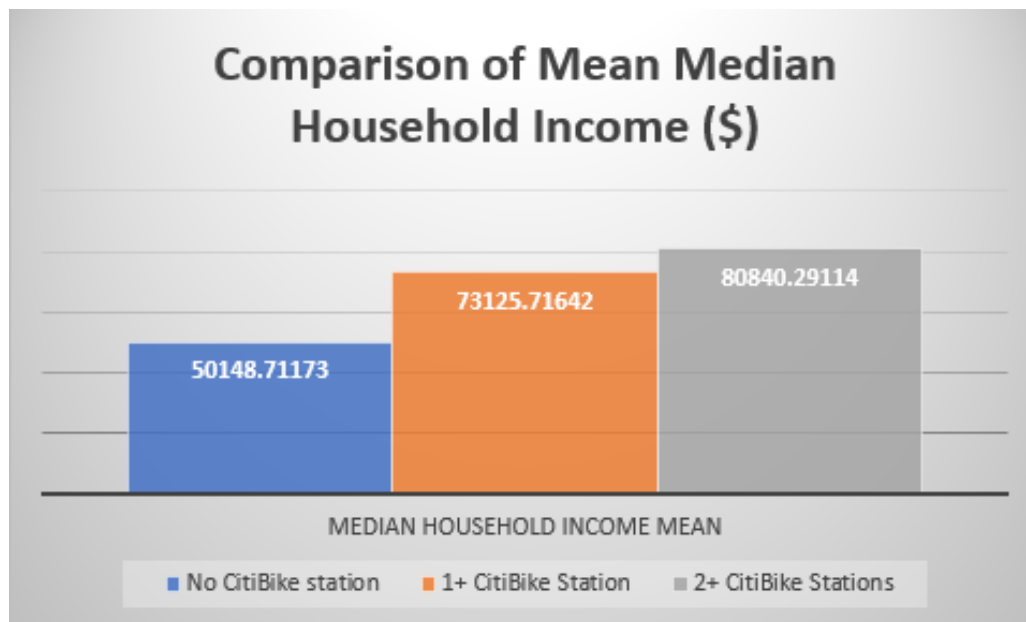■ No CitiBike station   ■ 1+ CitiBike Station   ▨ 2+ CitiBike Stations

Assumptions

Our main assumption is that the census tract a CitiBike station is located within is the one whose characteristics residents are most associated with it. This is likely often not the case as there are no physical boundaries between census tracts and the area included in a circle with the radius of a certain walking or biking distance is probably a better predictor of who the station serves. However, garnering statistics on this spatial unit presents difficulties, but could be managed using point data by household aggregated to our geography. For example, if we had location data on every eviction we could match it to our circle zones. Poverty Rate and Household Income data may be more difficult as there may not be household-by-household data on these categories.

Median household income, population and poverty rate, although attached to the Eviction Lab shapefile, all come from the 2015 5-year American Community Survey data. The ACS uses a sample to extrapolate estimates about the total population, which often include a large margin of error, so the numbers we used for these variables may be greatly inaccurate[5].
We want to expand analysis to include more areas and run a regression to get a correlation coefficient between the amount of CitiBike stations in a certain census tract and evictions/eviction filings in that same census tract before drawing any conclusions.

Conclusion

Bike sharing is an environmentally friendly method of transportation. The results of our spatial analysis indicate that census tracts with lower household incomes and higher poverty rates may be less likely to have access to this resource. We want to increase awareness of these trends to encourage bike sharing companies to cater better to a more diverse population.

---

[5] https://www.census.gov/content/dam/Census/programs-surveys/acs/guidance/training-presentations/20180418_MOE.pdf

As for the slight negative correlation between eviction rates and CitiBike presence, it may be that the census tracts, 217 out of Brooklyn's total 841, have already undergone the population shift associated with gentrification. Additionally to honing our 2016 analysis, we would like to look at historical eviction data, and compare that to the current presence of CitiBike stations. This is to attempt a conclusion about evictions leading to a population shift, which we can then correlate with CitiBike station presence.

<div align="center">Technical Appendix</div>

Data Sources

The datasets used in this analysis include:

- Cartographic Boundary File by Census Tract from the US Census Bureau for New York State[6]
- New York City Borough Boundaries, clipped to shoreline[7]
- Eviction Data for New York State, by census tract, 2016 from Eviction Lab[8]

    This data includes additional tabular data about each census tract including Median Income, Poverty Rate and Population, coming from the 2015 5-year ACS.

    The Data Dictionary gives detailed source information and a full list of tabular data[9]

    "This research uses data from The Eviction Lab at Princeton University, a project directed by Matthew Desmond and designed by Ashley Gromis, Lavar Edmonds, James Hendrickson, Katie Krywokulski, Lillian Leung, and Adam Porton. The Eviction Lab is funded by the JPB, Gates, and Ford Foundations as well as the Chan Zuckerberg Initiative. More information is found at evictionlab.org."

- CitiBike Station Location Data, from a live json feed, gathered June 18 2018[10]

Process Outline:

(all names are suggestions, given for clarity of continuity)

Download all of the above datasets

Convert CitiBike Station Location Data json to csv

Clean Eviction data, remove all columns except the variables of interest and format column names to be ArcGIS compliant

Spatial Join Citibike Station points with NY State Census Tract polygons, choosing to sum the amount of stations in each census tract, which creates a new column for count (SpJoin1)

---

[6] https://www.census.gov/geo/maps-data/data/cbf/cbf_tracts.html
[7] https://www1.nyc.gov/site/planning/data-maps/open-data/districts-download-metadata.page
[8] https://data-downloads.evictionlab.org/. We downloaded tracts.csv
[9] https://eviction-lab-data-downloads.s3.amazonaws.com/DATA_DICTIONARY.txt
[10] https://feeds.citibikenyc.com/stations/stations.json

Run a definition query on New York City Borough Boundaries shapefile to only keep Brooklyn. We will name this BkBoro

Clip SpJoin1 to the shape of BkBoro. We will name this SpJoinBk

Create new column in SpJoinBk that registers in GEOID as a number rather than a string

Table join eviction data with SpJoinBk on GEOID. We will name this BkBikes_Eviction.

Export BkBikes_Eviction as a txt file and open in Excel

Create the following columns:

> *Eviction_Filings_by_Pon*, which equals Eviction Filings divided by Population
>
> *Evictions_by_Pop*, which equals Evictions divided by Population
>
> *No_Citibike_Stations*, which can be gathered from an IF statement. If the count of citibike stations is 0, this column will have a 1, if not, it will have a 0
>
> *1plus_Citibike_Stations*, which can be gathered from an IF statement. If the count of citibike stations is equal to or greater than 1, this column will have a 1, if not, it will have a 0
>
> *2plus_Citibike_Stations*, which can be gathered from an IF statement. If the count of citibike stations is equal to or greater than 2, this column will have a 1, if not, it will have a 0

Save this file as a CSV and import it into ArcGIS. We will name this Evc_Math

Table join Evc_Math with SpJoinBk. We will name this Math_Join.

Make 2 copies of Math_Join

Run a definition query on each of the copies and rename the shapefiles as follows:

> Name: NoBikeStations. Definition query parameters will be: *No_Citibike_Stations = 1*
>
> Name: OneplusBikeStations. Definition query parameters will be: *1plus_Citibike_Stations = 1*
>
> Name: TwoplusBikeStations. Definition query parameters will be: *2plus_Citibike_Stations = 1*

Gather Summary Statistics from each of these shapefiles for variables of interest and copy the mean and standard deviation into an Excel file, where it can be analyzed and graphed.