Journal of Applied Economics and Policy Studies (2024) Volume 5 EWA Publishing
Published online: 23 May 2024 DOI: 10.54254/2977-5701/5/2024028

The Analysis of Airbnb Blocking Behavior and Relative Position to Top Tourism

Longji Xia

College of Literature, Science and Arts, University of Michigan, 855 South University Ave, Ann Arbor, U.S.A

xlongji@umich.edu

Abstract. With the rapid development of Airbnb in recent years, it has taken the main position in the sharing economy and also brought a great number of challenges to hotel competitors and regulators for its volatile supply. The benefit of agglomeration relative to Airbnb listing position and blocking behavior of listings has been discussed by Xie et al. (2019) and Peng (2020), and how the locational factor affects the blocking behavior remained unanswered. This paper will select the New York Airbnb listing data from Oct. 2014 to Dec. 2016 to compile the distance binary variables in 3 groups to identify whether they are near the top 10 attraction and then perform the logistic regression and t-test, which delineates the relationship between blocking behaviors and distance to attractions. We find that almost all top attraction variables significantly affect the Airbnb listings' blocking behaviors, but no clear pattern of direction is shown. Considering the negative elasticity of income for NYC cab working hours, this paper performs t-tests on attractions' yearly revenue indicating that while the coefficient estimates are both positive, their yearly revenue consistently has no significant differences of income in the 200m radius distance variable and no consistent result can be obtained in comparing attractions' yearly revenue with opposite signs coefficient.

Keywords: sharing economy, blocking behavior, locational factor, Airbnb

1. Introduction

With the advent and surge of companies like Airbnb and Uber, the sharing economy remains a heated topic because its common mechanism and potential effects on existing industries are still not quite clear. The key assumption of the sharing economics is that the consumers are allowed to "share creation, production, distribution, trade, and consumption of goods and services"[1]that is brought by the technology with lower transaction cost and supply side flexibility (Zervas et al., 2016). Like Airbnb, it constructs an online platform that enables the household to rent out apartments or houses with low utilization to enlarge the overall welfare by improving the overall allocative efficiency constrained by the same amount of factor of production.

Recently, Airbnb has kept a strong growing trend and plays a more crucial role in the economy. Its revenue steps up to $8.4 billion in 2022 from $0.4 billion in 2014, while the second quarter of 2023's revenue reaches $2.5 billion. Not only the revenue, but the rise of annual listings and bookings also illustrate the boosted scale in both the demand and supply side of the Airbnb industry: while the bookings jumped from 52 million to 393 million, the listings even increased almost 21 times from 0.3 million to 6.6 million. Airbnb also turns into the third largest firm in the industry with 16 percent of the market share which is only 3 percent lower than the market leader, Expedia.[2]

However, the large scale of Airbnb affects other industries, especially the hotel industry. According to (Zervas et al., 2016), Airbnb has a strong detrimental effect on hotel revenue which leads to a revenue drop of 8 to 10 percent: the regression results indicate that every 10 percent increase in the size of the Airbnb market will lead to 0.39 percent decrease hotel revenue by using the DID model that treat the entry of Airbnb as the variable intervention in space and time. For the hotel industry, the supply side gains tremendous pricing power during the high demand situations like festival times including Thanksgiving and New Year's Eve: the price of the same room in the peak season would be several times of the price in the off-season for excess demand in the market. However, the entry of Airbnb provides the flexibility of scale related to the high demand which strongly limits the price power of the hotel industry. The Airbnb hosts will evaluate welfare change between the utility of the space and outside renting price to decide whether to rent or not. The combination of hotels and Airbnb will result in lower variations in prices that

[1] https://www.thebalancemoney.com/what-is-the-sharing-economy-5188892
[2] https://www.searchlogistics.com/learn/statistics/airbnb-statistics/

accommodate demand within a more efficient market.

New York, which was always in the top 10 cities of Airbnb in the last few years, has suffered a huge loss from the new registration enforcement in September 2023. This city has been the third city in the top lists both for Airbnb listings and nights stayed in 2020 and further stepped to the second position in the next year proving the success in the local market. The constrained land space of hotels and heated traveling demand lay a solid base for the rapid development of Airbnb. However, Local Law 18, the Short-term Rental Registration Law, came into force then. It is so strict that it almost bans entirely for lots of short-term guests and hosts. The law requires every short-term rental host to register with the Mayor's Office of Special Enforcement (OSE) and prohibit the platform from processing any orders for unregistered hosts. It also limits rental to two guests. In the short term, the number of Airbnb listings has plunged by more than 75% which offers a more favorable for the hotel industry. By the Trivago's Hotel Price Index, the average room price is $529 per night which has increased nearly 20% from November during Macy's Thanksgiving Day Parade and it also marks an 8% increase from last December which indirectly implies the effect of diminishing Airbnb.

In this way, it is shown that the variation in Airbnb supply will affect the hotel industry strongly. Besides the local policies as the uncontrolled factor, we are more interested in the blocking behaviors of Airbnb hosts that more obviously interfere with other traits in the markets. We postulate that location might be one of the determining factors and assume that Airbnb's location near the top 10 tourism tends to have a lower frequency of blocking behavior due to huge demand. With the aid of New York's Airbnb database from 2014-2016, we estimate the quantitative relationship between the blocking behavior and location binary variables by the logistic linear regression. Several groups of binary variables that whether Airbnb is around this tourism are constructed for different distances to the tourism (200m, 300m, 500m) to check the commonality of the pattern in the regression. Other numeric traits including the revenue, rating, and occupation rate are the controlled variables in the model. Although the majority of the binary variables are significant, the results do not follow our assumption. It is divided into two groups: one is significantly positive, and the other is significantly negative. Further assumptions are made: the blocking pattern might depend on the yearly revenue of Airbnb hosts around certain locations. Airbnb might be less willing to block when the price is relatively higher for the same room, which contributes to the negative intercept of the binary variable. Applying the T-test within and without the positive and negative groups is our method to assure our assumption. However, this situation only occurs in the 200m case but does not occur in 300m and 500m cases. We supposed several possible reasons for this result: a) we only collected the Airbnb database to determine the relationship while Airbnb only represents 16% of the market share. The selection bias could affect the final result. b) Here, we only pick New York City's Airbnb as the data due to the availability of the dataset. The more evident trend may occur in the overall Airbnb market combining majority cities. c) The suitable radius and more controlled variables should be considered in regression.

The rest of paper is organized as follows. In Section 2, we summarize sharing economics and Airbnb literature. In Section 3, we describe our data and variables with base statistics. In Section 4, we discuss our methodology and 3 hypotheses. Then in Section 5 we discuss our results, provide future research side in Section 6, and conclude in Section 7.

2. Literature Review

2.1. Sharing Economy's Definition and Recent Research Trend

Due to the success of Airbnb and Uber, the rapid growth of sharing economy (SE) has stimulated great public interest in the last one and a half decades for its mechanism and impact. However, no exact definition of SE can be approved by public researchers because SE practices are strongly varied and persistently changing (Hossain, 2020) while no specific criteria can be applied to every SE transaction. A reasonable and balanced definition is provided by Munoz and Cohen (2016: p21): "a socio-economic system enabling an intermediated set of exchanges of goods and services between individuals and organizations which aim to increase efficiency and optimization of sub-utilized resources in society" Munoz and Cohen (2016: p 21).

Two literature reviews by Cheng and Hossain trace the number of scholars' work on SE. Chen (2016) analyzed 66 papers on SE with ten publications about tourism and hospitality from 2010 to 2015 by applying co-citation network analysis and Leximancer content analysis. We can track a rapid growth of research interest in the SE field: published started at one or two in the first two years and then grew to 5 for 2012-2013. It reached ten in 2014 and had an explosive growth to 35 in 2015. 3 main areas in the sharing economy in 2010-2015 include SE business models and its impact, the nature of SE, and sustainability development. Concerning sustainability, even though Airbnb states that its business model results in sustainable business practices for reducing greenhouse emissions, energy, and water use, no evidence from empirical studies can support their ecological impact in SE (Juul, 2015). For tourism and hospitality, major papers focus on SE impacts on destinations and tourism and impacts on tourists.

After Chen's conclusion, Hossain (2020) then continued work on the literature review of SE from 2016 to 2018. 219 articles were collected and the number of publications reached to climax of 120 in 2017 and dropped to 40 in the next year. Two dominant sectors are still accommodation and transportation for the success of Uber and Airbnb. The scholars explore 3 main fields: definitional dilemma, sharing economy as a phenomenon, and key theories. More specifically, the SE phenomenon involves several aspects: motivation, stakeholders, challenges, impact, SE models, policy, and regulations.

2.2. Airbnb

Due to the quick development of Airbnb, lots of scholars pay great attention to heated topics of Airbnb and we will go through and focus on several of them: consumer segmentation, regulation, impact, superhost, blocking behavior, and agglomeration effect.

2.2.1. Consumer Segmentation

Researchers firstly try to figure out the major difference between the users and non-users of the sharing economy and picking Airbnb as the case study. Smith and Flash Eurobarometer (2016) have both shown a significant demographic divide between users and non-users of the sharing economy. Smith finds that income and education are the main distinctions between US users and non-users as one-fourth of American adults in the highest income and education had used home-sharing services before while only 4% of those in the lowest income and education had used so. Similar results were found in the large-scale investigation of 28 EU member countries (Flash Eurobarometer, 2016). The survey indicates that education and age are crucial determinants for whether to use the sharing economy services.

After that, scholars work on the consumer segmentation between shared rooms and entire rooms in the Airbnb service. For entire room and shared room customers, they have strong different demographic factors (Lutz et al., 2018). A huge distinction has been found while comparing the consumer segmentation and hosts' marketing targeting strategy. By conducting a quantitative survey of 659 Airbnb users in US and a qualitative content analysis of 500 listings on Airbnb.com, Lutz, and Newlands concluded such disconnection leads to lower efficiency in the Airbnb market.

2.2.2. Regulations

Although Airbnb brings new experiences, overcrowded resident areas situation caused by high demand is disruptive to the traditional industry and availability of housing (Nieuwland et al., 2018). More Airbnb users favor the local culture of residential than tourism enclave which further raises the problem of surging rent and inferior environment. The complaints range from visitors' noises, traffic & parking issues, waste management, and safety concerns (Gallagher, 2017). The main problem here is that regulation should also evolve with the online platform to avoid negative externalities.

In this way, Nieuwland conducts the Airbnb regulation in 11 European and American cities and zooms in Denver. There are three main approaches of regulation in the sample: prohibition, Laissez-faire, and allowing it with certain restrictions. Restrictions include quantitative restrictions (limit the number of days, visitors or time), locational restrictions, density restrictions, and qualitative restrictions. By zooming into Denver's case, how to introduce online regulation in an understandable way such as the online platform will be essential to sustainable development. Even though the listing number in Denver dropped a little from 4103 to 3540 for the first six months, such regulation still exploits the under-utilized capacity of apartments with low negative externalities. Koopman (2014) also indicates the essence of policy evolving when market circumstances change dramatically like Airbnb. Ranchordas mentions that innovation or technology like Airbnb should not be stifled by regulation but should not be unregulated which corresponds to Denver's case.

2.2.3. Impact

Airbnb strongly affects diverse markets by improving the underutilized capacity of rooms to reach higher efficiency in the economy. During the process, how the stakeholder's welfare changes in each market is a concerned and hotly debated topic among scholars.

From the macro aspect of Airbnb, the peer-to-peer rental market has altered the goods allocation dramatically, substituting rental for ownership and lowering the used goods price to raise the consumer surplus (Frailberger et al., 2015). Both consumption and supply shift heavily to below-median income users who gain a disproportionate part of welfare from the sharing economy by boarder inclusion, better quality consumption, and new ownership. Then, zooming into the accommodation market, Farronato and Fradkin (2018) propose a demand and supply framework to quantify the welfare effect of Airbnb's entry into the accommodation market by using data from top US cities. The results indicate that $41 of consumer surplus per room-night and $26 of landlord surplus are generated by Airbnb while hotel variable profits from accommodation decreased by up to 3.7 percent leading to a total welfare gain of $137 million in 2014 in these cities. Effects concentrate in New York and times when hotel capacity was constrained and had a higher pricing power. Moreover, by applying Texas data to quantify the Airbnb effects on the hotel industry, the base estimate shows that a 1% increase in Airbnb listings in Texas causes a 0.05% decrease in quarterly hotel revenue (Zervas et al., 2016). However, the impact is distributed unevenly in hotels not catering to business travelers and lower-end hotels.

Besides the hotel market, Airbnb also faces many criticisms or even lawsuits for stimulating the housing prices in the residential real estate market and the rental price as well (Chen et al, 2019; Horn et al., 2017). Theoretically, the landlord can switch from the long-term rental market to the short-term market which drives up the rental rate in the long run that will indirectly affect the housing price. With the data sources from Airbnb, Zillow, and the Census Bureau from all United States, a 1% increase in Airbnb listings number leads to a 0.018% increase in rent and a 0.026% increase in house price at the median owner-occupancy rate (Barron et al., 2018).

Within Airbnb, authors try to find out which party gains the most benefit from the sharing economy. To find that, Quattrone & Proserpiio (2016) match the listing information from the census and hotel in London for panel data and then split them into the ward unit area to do the regression with the control variables. After using different dependent variables such as the number of Airbnb listings, reviews & hotels, scholars find a similar pattern: Airbnb is inclined to be offered in areas with highly educated non-UK-born renters while homes tend to be offered in high-end areas. Considering the periodical element, central areas become less predominant year on year base while income and owned property numbers become significantly negatively related.

2.2.4. Superhost & Blocking Behavior

Superhosts are automatically rated by Airbnb if certain hosts are certified with four criteria Airbnb announced (reliable cancellation behavior, host responsiveness, and sufficient Airbnb demand). However, as the Superhost is given by the platform without explicit process information, it is not clear which one of the four criteria is more crucial to gain the Superhost status. In this way, Gunter (2018) applies logit, probit, and IV probit models to quantify the marginal contribution of the four criteria. The result reveals that an excellent rating is the most important criterion followed by reliable cancellation behavior, host responsiveness, and Airbnb demand. Moreover, Wang and Nicolau (2017) find that the Superhost will have a significantly higher price on the same listing while Liang et al. (2017) also show that Superhost can process more extensive host profiles than others.

As the Airbnb listing number grows extensively and even larger than the biggest hotel chain, how the Airbnb supply chain would be crucial to stakeholders. Peng (2020) investigated the blocking behavior of homeowners with an empirical analysis of Airbnb proprietary data to find factors that affect their product supply. 10 factors significantly affect the blockade of houses, which include prices, bedrooms, consecutive booking, property type, and listing types.

2.2.5. Agglomeration Effect

Agglomeration theory was first introduced by Marshall (1890) and gives two explanations to interpret the reasons why competitors decide to colocate in the nearby places including production enhancement and increased demand. This theory offers a strong theoretical foundation for why lodging companies always agglomerate in close geographic locations (Lee & Jang, 2015).

Step by step, location becomes a vital feature of lodging products and strongly affects the hotels' performance (Balaguer et al., 2013). Then Marco-Lajara et al. (2016) figured out the U-shaped relationship between the degree of agglomeration and the company's growth in the lodging industry. The profits will firstly decrease with high competition at the beginning and then rise for higher agglomeration levels. Such an agglomeration of homogeneous suppliers provides positive externalities that are brought by strategic price positioning and spillover demands. The lodging products are all inclined to be colocated in popular locations like the tourist attractions that depend on nearby amenities. Tsang and Yip (2009) checked the agglomeration effects in Beijing and proved that the high star-ranking joint-venture hotels all contribute to improved demand and all hotels can gain such benefits.

Zooming into Airbnb, Airbnb listings are also found agglomerated in popular locations like tourist attractions (Blal et al., 2018; Heo & Blengini, 2019; Wegmann & Jiao, 2017), as they also heavily rely on amenities that the facility provides (Davidson & Infranca, 2016). Zhang and Chen sought to learn Airbnb's geographic dynamics by the data of Chicago, NYC, and Los Angeles: listings are centered in such population positions with center-peripheral patterns that are consistent with other literature. Xie et al. (2019) continue to bridge the research gap in quantifying the agglomeration effect. Results explain that it will positively affect revenue performance and it will be stimulated by operation experience and mitigated by longer tenure. As the tourist attractions are the places that Airbnb will often colocate, the listings can earn more due to the agglomeration effect in the above analysis. In this way, we want to explore whether Airbnb nearby these places that enjoy the agglomeration effect will have certain common blocking behavior patterns or the same as others.

3. Data

Two data frames we used are from the AirDNA website (http://www.airdna.com) which offers the listing information in New York. One includes 46 traits of each Airbnb listing while the other is the panel data that contains the status of the listing (available, blocked, or rented) from October 2014 to December 2016. As dealing with around 40 million observations is a truly heavy workload for further analysis, we only leave 1 dependent variable (blocking status) and 9 control variables from the traits dataset. For control variables, we pick the numerical ones that are more likely to be significant from Peng's (2020) regression (Price, occupancy rate, number of bookings, rating, number of reviews, bedroom and bathroom numbers, number of photos).

For identifying the tourist attractions, we choose the top 10 list of New York attractions on the Trip Advisor website for recent five years: Central Park, National 911 Museum, Museum of Art, Empire State Building, High Line, Top of Rock, Broadway, Brooklyn Bridge, Statue of Liberty and Manhattan Skyline. In order to distinguish whether the listing is nearby the attraction, we obtain the longitude and latitude of each attraction from their official website to at least 3 decimal places to limit the error to 50 meters. After that, by using the longitude and latitude of each listing in the traits dataset, the great circle distance from listing to ten attraction can be computed through the law of cosine by R. The next problem is which distance is a better choice. We first try a relatively long distance like 1 kilometer, but lots of noise variables will potentially pollute the regression result indeed while a too-short distance will also cause a too small sample size for further regression. Therefore, 3 groups of distance-independent variables (200, 300, 400 meters) are chosen to balance the noise and sample size problem. Within each group, 10 binary independent variables for 10 attraction works to convey the listing locational information.

Table 1. Descriptive Statistics of Independent Variables

Variable	Definition	Min	Max
	Independent Varibles: Distance		
Nearby Centpark	1=distance to central park is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby National911	1=distance to National 911 museum is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Museum of Art	1=distance to Museum of Art is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Empire State Building	1=distance to Empire State Building is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby High line	1=distance to National 911 museum is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Top of Rock	1=distance to Top of Rock is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Broadway	1=distance to Broadway is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Brooklyn Bridge	1=distance to Brooklyn Bridge is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Manhattan Skyline	1=distance to Manhattan Skyline is smaller than 200/300/400m 0=otherwise	0.00	1.00
Nearby Statue of Liberty	1=distance to Statue of Liberty is smaller than 200/300/400m 0=otherwise	0.00	1.00

Table 2. Descriptive Statistics of Independent Variables

Variable	Mean (200m)	SD (200m)	Mean (300m)	SD (300m)	Mean (400m)	SD (400m)
Nearby Centpark	0.0003	0.02	0.0011	0.035	0.0023	0.049
Nearby National911	0.0004	0.018	0.0017	0.037	0.0038	0.057
Nearby Museum of Art	0.00004	0.005	0.0002	0.011	0.0004	0.017
Nearby Empire State Building	0.0019	0.044	0.0044	0.064	0.0067	0.078
Nearby High line	0.0014	0.04	0.0028	0.053	0.0041	0.063
Nearby Top of Rock	0.0002	0.013	0.0008	0.025	0.0019	0.037
Nearby Broadway	0.002	0.04	0.0029	0.054	0.0050	0.070
Nearby Brooklyn Bridge	0	0	0	0	0.0001	0.011
Nearby Manhattan Skyline	0.001	0.032	0.0031	0.048	0.0046	0.056
Nearby Statue of Liberty	0	0	0	0	0	0

Table 3. Dependent and Controlled Variables

Variable	Definition	Mean	SD	Min	Max
	Dependent Variable: Blocking Status				
Status_b	1 = blocked 0 = available or rented	0.389	0.474	0	1
Control variables					
Average Daily Rate	Average daily price rate of listing	161	138.06	10	8000
Occupancy Rate LTM	Rate of rented time to total time in preceding 12 months	0.604	0.2519	0	1
Number of Booking LTM	Number of bookings in preceding 12 months	10.13	19.647	0	184
Overall Rating	Online rating of Airbnb listing	4.584	0.448	1	5
Number of Reviews	Number of online reviews	14.3	30.522	0	367
Number of Photos	Number of listing photos online	11.9	10.355	0	240
Minimum Stays	Minimum days to book the listing	2.99	4.663	0	1000
Bedrooms	Number of bedrooms	1.13	0.694	0	10
Bathrooms	Number of bathrooms	1.13	0.370	0	8

4. Methodology

4.1. Binomial logistic model

In order to study the dynamic change of the binary dependent variable, the binomial logistic model is strongly helpful in distinguishing multiple variables' influence. When leaving the control variables unchanged, we can figure out how the same listing will be affected when located nearby the travel attraction in New York. The dependent variable is the blocking status where the listing is blocked or not. The blocking status will be recorded as 1 while other unblocking statuses (rented and available) will be recorded as 0. In this way, we suppose p_i is the probability that the individual host decides to block the listing and then the expression of the binomial model is:

$$p_i = \frac{e^{(\beta_0+\beta_1*x_1+...+\beta_n*x_n)}}{1+e^{(\beta_0+\beta_1*x_1+...+\beta_n*x_n)}} \tag{1}$$

In this equation, β_0 is the constant term that discuss the intercept of the regression as the start point for variable measurement. $\beta_1,\beta_2...\beta_n$ are the regression coefficients for variables x1,x2..xn that illustrate the marginal effect of this variable on the dependent variables which is the blocking probability here p_i here. Through the logit transformation of the first equation, it turns into the linear regression function with the same parameters $\beta_1,\beta_2...\beta_n$. This transformation can be considered as converting the binomial model into the fitted linear regression model.

$$Y = ln\frac{p_i}{1-p_i} = \beta_0 + \beta_1 * x_1 + ... + \beta_n * x_n \tag{2}$$

In the new linear regression form, β_i ($i \neq 0$) represents that 1 unit change of the variable x_i leads to the logarithmic change the ratio of blocking probability by β_i amount. Among the 10 independent variables this paper is going to investigate, all of them are binary variables that indicate the locational information of whether it is near any top travel attractions. For these binary variables, the coefficient estimates identify the marginal effect when they are close to certain attraction when other conditions stay unchanged instead of the one unit change for continuous variables. 9 more numerical controlled variables are also in regression form to avoid the potential bias and noise variable. In total, there are 19 variables in the final regression.

$$\begin{gathered} ln\frac{p_i}{1-p_i} = \beta_0 + \beta_1 * Nearyby\_Cent\_park \\ +\beta_2 * Nearby\_National\_911 \\ +\beta_3 * Nearby\_Museum\_of\_Art \\ +\beta_4 * Nearby\ Empire\ State\ Building \\ +\beta_5 * Nearby\ High\ line \\ +\beta_6 * Nearby\ Top\ of\ Rock \\ +\beta_7 * Nearby\ Broadway \\ +\beta_8 * Nearby\ Brooklyn\ Bridge \\ +\beta_9 * Nearby\ Manhattan\ Skyline \\ +\beta_{10} * Nearby\ Statue\ of\ Liberty \\ +\beta_{11} * Average\ Daily\ Rate \\ +\beta_{12} * Occupancy\ Rate\ LTM \\ +\beta_{13} * Number\ of\ Booking\ LTM \\ +\beta_{14} * Overall\ Rating \\ +\beta_{15} * Number\ of\ Reviews \\ +\beta_{16} * Number\ of\ Photos \\ +\beta_{17} * Minimum\ Stays \\ +\beta_{18} * Bedrooms \\ +\beta_{19} * Bathrooms \end{gathered}$$

4.2. Hypothesis

According to the Agglomeration Effect research by Xie et al.(2019), when the Airbnb listings are close to the travel attraction, it means that the agglomeration effect will have a positive effect on revenue growth. Peng (2020) also works on the potential significant factors that will affect the blocking behavior. While the top tourist attractions are the places that maximize the agglomeration effect, this paper tries to investigate whether there is a potential pattern under their blocking behavior. Then we propose the following hypothesis:

Hypothesis 1: The binomial regression coefficient for independent binary variables will be all positive or negative.

When the results are in accordance with the H1, a direct correlation between the agglomeration effect and blocking behavior can be revealed. However, if the regression results are not the same with H1, we will continue to investigate other situations.

4.3. T-test analysis

While Faber (2005) and Crawford et al. (2011) have found the negative wage elasticity of working hours of labor supply in the New York cab market, a similar pattern might also occur in the Airbnb accommodation market. We first separate the top attraction binary variables with values into two groups according to the sign of the coefficient estimates: one with the positive coefficient and the other with the negative coefficients (The following paper will use this separation). To explore this effect, the t-test will be a useful statistical tool to determine whether there is a significant difference between means of positive and negative groups or means in the same group. All listings that are nearby one certain attraction are gathered and form a new data frame. This step will repeat for all attractions. The goal is to look for whether there is any significant difference in earnings between two groups of listings or within the group. In this way, with the yearly revenue and maximum number of guests from the merged listing data frame, yearly revenue by size can be obtained by revenue divided by the maximum guest number. Wald T-test will then be done between and within two groups for both yearly revenue and yearly revenue by size. Two following hypotheses are proposed in accord with the cab patterns:

Hypothesis 2: Significant differences will consistently be found between two attraction data frames in different groups.

Hypothesis 3: Insignificant differences will be consistently found between two attraction data frames in same groups

Two above hypotheses try to prove that listings that nearby different attractions will have a similar blocking pattern when listings' yearly incomes are at a similar level while different behavior patterns could be caused by income differences like the cab market.

5. Result

5.1. Regression Results

Table 4. Logistic Regression Model Results (200m radius)

Variable	Coefficient Estimates	Standard error	Z-value	P-value
Intercept	-2.713***	0.0049	-550.271	0.000
Nearby Centpark	0.423***	0.02	21.532	0.000
Nearby National911	0.253***	0.02	11.614	0.000
Nearby Museum of Art	0.969***	0.095	10.235	0.000
Nearby Empire State Building	-0.057***	0.01	-5.705	0.000
Nearby High line	-0.042***	0.011	-3.888	0.000
Nearby Top of Rock	-0.122***	0.036	-3.434	0.001
Nearby Broadway	-0.055***	0.01	5.366	0.000
Nearby Brooklyn Bridge	n.a.	n.a.	n.a.	n.a.
Nearby Statue of Liberty	n.a.	n.a.	n.a.	n.a.
Nearby Manhattan Skyline	-0.015***	0.014	-11.114	0.000
Average Daily Rate	0.00009***	0.000004	2.551	0.010
Occupancy Rate LTM	1.718***	0.002	957.436	0.000
Number of Booking LTM	-0.038***	0.00004	-882.667	0.000
Overall Rating	0.376***	0.001	374.887	0.000
Number of Reviews	-0.006***	0.00003	-213.369	0.000
Bedrooms	0.033***	0.0007	44.426	0.000
Bathrooms	-0.037***	0.001	-27.730	0.000
Minimum Stays	0.001***	0.00009	15.032	0.000
Number of Photos	-0.005***	0.00005	-112.596	0.000

Note: *=10% significant, **=5% significant, and ***=1% significant

Table 5. Logistic Regression Model Results (300m radius)

Variable	Coefficient Estimates	Standard error	Z-value	P-value
Intercept	-2.711***	0.005	-549.748	0.000
Nearby Centpark	0.357***	0.012	31.023	0.000
Nearby National911	-0.123***	0.012	-10.667	0.000
Nearby Museum of Art	-0.164***	0.037	-4.475	0.000
Nearby Empire State Building	-0.121***	0.007	-16.988	0.000
Nearby High line	-0.135***	0.008	-16.705	0.000
Nearby Top of Rock	-0.057***	0.019	-3.004	0.002
Nearby Broadway	0.176***	0.008	23.139	0.000

Table 5. Continued

Variable	Coefficient Estimates	Standard error	Z-value	P-value
Nearby Brooklyn Bridge	n.a.	n.a.	n.a.	n.a.
Nearby Statue of Liberty	n.a.	n.a.	n.a.	n.a.
Nearby Manhattan Skyline	-0.148***	0.009	-16.056	0.000
Average Daily Rate	0.00001***	0.000004	3.995	0.000
Occupancy Rate LTM	1.718***	0.002	957.372	0.000
Number of Booking LTM	-0.037***	0.00004	-882.233	0.000
Overall Rating	0.376***	0.001	374.541	0.000
Number of Reviews	-0.005***	0.00003	-213.836	0.000
Bedrooms	0.032***	0.0007	43.899	0.000
Bathrooms	-0.038***	0.001	-28.072	0.000
Minimum Stays	0.001***	0.00009	14.651	0.000
Number of Photos	-0.005***	0.00005	-112.201	0.000

Note: *=10% significant, **=5% significant, and ***=1% significant

Table 6. Logistic Regression Model Results (400m radius)

Variable	Coefficient Estimates	Standard error	Z-value	P-value
Intercept	-2.711***	0.005	-549.524	0.000
Nearby Centpark	0.331***	0.008	40.780	0.000
Nearby National911	-0.074***	0.007	-9.996	0.000
Nearby Museum of Art	0.179***	0.025	7.205	0.000
Nearby Empire State Building	-0.094***	0.006	-16.259	0.000
Nearby High line	0.050***	0.007	-7.358	0.000
Nearby Top of Rock	0.050***	0.012	4.085	0.000
Nearby Broadway	0.223***	0.006	38.419	0.000
Nearby Brooklyn Bridge	-0.566***	0.040	-14.084	0.000
Nearby Statue of Liberty	n.a.	n.a.	n.a.	n.a.
Nearby Manhattan Skyline	-0.005	0.008	-0.641	0.522
Average Daily Rate	0.000**	0.000	2.239	0.025
Occupancy Rate LTM	1.719***	0.002	957.668	0.000
Number of Booking LTM	-0.037***	0.000	-882.026	0.000
Overall Rating	0.376***	0.001	373.966	0.000
Number of Reviews	-0.006***	0.000	-213.990	0.000
Bedrooms	0.034***	0.001	45.365	0.000
Bathrooms	-0.038***	0.001	-28.350	0.000
Minimum Stays	0.001***	0.000	14.084	0.000
Number of Photos	-0.005***	0.000	-111.710	0.000

Note: *=10% significant, **=5% significant, and ***=1% significant

We estimate the equation (1) with the logistic regression 3 times for each distance group by R studio and report the software's outputs in Table 4,5&6. One main problem is the n.a phenomenon of variable Nearby_Statue_of_Liberty and Nearby_Brooklyn bridge. Their results are intentionally done by the software for multicollinearity. It is caused by the the special location of these two attractions where it is not realistic to open the Airbnb listing so close to these two attractions. Both Statue of Liberty and Brooklyn Bridge are far away from the residential area and the Statue of Liberty is even on the island. In this way, no observations are recorded as 1 in 200 and 300m for both attraction's dummy variables, while Nearby_Statue_of_Liberty's result is still n.a for the 400m case.

For the significance level and coefficient estimates, the tables indicate the following findings: (a) all of the independent variables are significant at 1% level except the Manhattan Skyline in the 400m case, reflecting that nearby the travel attraction should be the crucial factor affecting the listing supply. (b) Besides the pricing factor only passing the 5% significance test in 400m, all other controlled variables are also significant at the 1% level which corresponds to Peng's (2019) research results. (c) However, the coefficient estimates of the independent distance variables contradict to the hypothesis 1 we propose in the above section. In each distance group, no general explicit pattern we expected can be found and they are all blended with positive and negative coefficient estimates.

After separating the independent variable into positive and negative groups, another problem is revealed when checking the pattern: when altering the distance to the larger radius in the regression, not all coefficients will stay positive or negative: 4 of all attractions' coefficient estimates sign stay unchanged and significant in 3 regressions while Central Park and Broadway's estimates are positive and Empire State Building and High Line's estimates are negative. Estimates of Manhattan Skyline also remain negative for all regression, but it is not significant in the 400m case. The others all change when the radius alters in the process: Top of Rock's coefficient retains negative then turns to positive in the 400m case; Museum of Art's coefficient is positive in the 200m and 400m case and changes to negative in 300m case; National 911 museum's estimate is only positive in 200m case and become negative in 300m and 400m case. The reason we suspect for the swap phenomenon is the sample size of the listing that is recorded as 1. Especially for the downtown, the density of airbnb listing will be extremely high. This will cause the number of listings in a 400m radius might be several times of listing number in a 200m radius. While only 35 observations are recorded as 1 for Nearby_Top_of_rock in a 200m radius, this number climbs to 313 observations in 400m cases. For the other 2 swapping cases, Museum of Art observations jump from 7 to 59 and the National 911 museum's cases also surge from 60 to 491. By calculating each attraction observation according to the appendix, the result shows that the 400m observation number normally is 5.91 times of the 200m observation number which could be a potential problem. In these expansions, added neighborhoods with diverse housing types will cause quite a different result for the listings blocking behavior in the market that greatly affects the estimates of independent variables.

5.2. T-test Results

200m t-test		Positive				Negative			
		Cent	911_mus	MOA	Broad	Empire	High	Top	Manhattan
Positive	Cent		×	×	×	×	×	×	×
	911_mus			×	×	√	√	×	√
	MOA				×	√	√	×	√
	Broad					√	×	×	×
Negative	Empire						×	×	×
	High							×	×
	Top								×
	Manhattan								

Note: √= 1% significant different, ×= 1 % insignificant

200m t-test by size		Positive				Negative			
		Cent	911_mus	MOA	Broad	Empire	High	Top	Manhattan
Positive	Cent		×	×	×	×	×	×	×
	911_mus			×	×	√	√	×	√
	MOA				×	×	×	×	×
	Broad					√	×	×	×
Negative	Empire						×	×	×
	High							×	×
	Top								×
	Manhattan								

Note: √= 1% significant different, ×= 1 % insignificant

300m t-test		Positive		Negative					
		Cent	Broad	Empire	High	Top	Manhattan	911_mus	MOA
Positive	Cent		×	×	×	×	×	√	√
	Broad			√	×	×	×	×	√
Negative	Empire				×	×	×	√	√
	High					×	×	√	√
	Top						×	√	√
	Manhattan							√	√
	911_mus								√
	MOA								

Note: √= 1% significant different, ×= 1 % insignificant

300m t-test by size		Positive		Negative					
		Cent	Broad	Empire	High	Top	Manhattan	911_mus	MOA
Positive	Cent		×	×	×	×	×	√	√
	Broad			×	×	×	×	×	√
Negative	Empire				×	×	×	√	√
	High					×	×	√	√
	Top						×	×	√
	Manhattan							√	√
	911_mus								×
	MOA								

Note: √= 1% significant different, ×= 1 % insignificant

400m t-test		Positive				Negative			
		Cent	Top	MOA	Broad	Empire	High	Brook	911_mus
Positive	Cent		×	×	√	×	×	×	×
	Top			√	×	×	×	×	×
	MOA				×	√	√	×	×
	Broad					√	×	×	×
Negative	Empire						×	×	√
	High							×	×
	Brook								×
	911_mus								

Note: √= 1% significant different, ×= 1 % insignificant

400m t-test by size		Positive				Negative			
		Cent	Top	MOA	Broad	Empire	High	Brook	911_mus
Positive	Cent		×	×	√	×	×	×	×
	Top			√	×	×	×	×	×
	MOA				×	√	√	×	×
	Broad					×	×	×	×
Negative	Empire						×	×	√
	High							×	√
	Brook								×
	911_mus								

Note: √= 1% significant different, ×= 1 % insignificant

The purpose of the t-test is to verify hypotheses 2 and 3 that significant should occur between positive and negative groups' revenue means but do not occur between means in the same group. After we form the data frames for each attraction and do the tests one by one, the above t-test results are shown in each distance regression for both yearly revenue and yearly revenue by size. Although the majority of Airbnb listings have a relatively small size on average, yearly revenue by size will bring a more accurate result as larger listings should have higher income requirements and different elasticity of income. While comparing the tables, no huge differences are found for total revenue or revenue by size.

For hypothesis 2, it is perfectly proved in the 200m case but becomes fragile in 300m and 400m cases. No significant differences are found between any means within the same group in the 200m case. While the hypothesis is true for the positive group member and 3 members in the negative group in the 300m case, 911 National Museum and Museum of Art's means are significantly different from almost all other negative group members. In the 400m t-test, still 4 pairs of means in the same group have significant differences. By collecting more samples in the data frames, it seems the income differences within the same group will be more significant. However, hypothesis 3 could not be supported by any radius cases: in all cases, around half or more t-tests indicate that no significant differences are detected between the means in different groups. To sum up, the potential pattern of hypothesis 2 might occur for the same group of coefficients estimate but needed more attraction and samples to support that while significant differences could not be consistently found among means in different groups that contradicts hypothesis 3. For the negative coefficient group, the income negative elasticity should not be the only motivation that affects the blocking behavior and more correlated factors are required to be considered to obtain the final pattern.

6. Future Research

6.1. Data limitation

For this research, we selected the New York City Airbnb data for around two years to do further analysis. However, the pattern this paper analyzes here might not be the general pattern that applies to every city due to differences in economic, cultural, and national environment. Therefore, a method like conducting the data on a larger scale such as the whole US will be needed for a more solid result. Besides that, even though Airbnb is the market leader in the accommodation industry of the sharing economy, its market share percentage still could not symbolize the whole market. Simply digging into Airbnb data instead of the whole market to do a comparison might misunderstand the comprehensive effect of tourist attractions’ agglomeration influence on blocking behavior. Moreover, in this numerical analysis, we only pick 10 top attractions for limited calculation capacity for running the regression. With the existence of strange cases like Brooklyn Bridge and the Statue of Liberty, more top attractions are required to be included for a more general pattern of blocking behavior. Later researchers can collect the US data of the whole industry and add more attractions to test the significance of the estimates.

6.2. Other Controlled Variables Ignored

Although Airbnb listings are always located at popular positions like tourist attractions (Blal et al., 2018; Heo & Blengini, 2019; Wegmann & Jiao, 2017), the locational factors of each attraction might be extremely different such as the surrounding amenities and demographic elements. It has become a common phenomenon that majority of the Airbnb listings are operated for places with a higher standard of amenities, though the listings also diverged to the suburbs in recent years (Quattrone et al., 2016). The target customer profile will also be diverse in two cases. While the downtown listings might focus more on wealth and highly educated customers, the suburb listing loves customers on the other side. Future researchers can look for the demographic data frame that consists of the demographic factors in each zipcode area and then merge the information according to the zipcode. With the help of the demographic factors, the regression estimate will give a more accurate result

6.3. Suitable Radius

Considering the radius of the distance variables, we are constrained by several factors. Getting a higher radius will truly take in more observations into the regression, especially for the attraction including the Statue of Liberty and Brooklyn Bridge which helps a more comprehensive view of the attraction effects. In real-life situations, “nearby the attraction” could consist of a distance from 100m to 5km depending on the transportation. However, for attractions in downtown like New York, the extremely crowdedness of buildings and population will exaggerate the noise variables in large radius situations. Even only for 2 kilometers, various factors like the substitution effect of the hotels or the amenities will be the vital elements related to blocking behavior. How to balance the noise variables while obtaining more observations is one problem.

7. Conclusion

From the previous sections, through 4 million observations of 40,000 listings in New York City, we used binomial regression and t-test to distinguish whether locating nearby the travel attraction in New York will affect the Airbnb listing’s blocking behavior. Regressions indicate that among the 10 top attractions, almost all coefficient estimates are significantly affecting the listings blocking behavior. However, instead of having a pattern like hypothesis 1 that all coefficient estimates should have the same sign, the sign of coefficient estimates are blended with both positive and negative.

Mirroring Faber (2005) and Crawford et al. (2011) study on the negative elasticity of income on hours in NYC’s cab market, we then focus on using a t-test to check whether attractions with positive and negative coefficients will have significantly different means of yearly revenue according to hypothesis 2&3. Results identify that attraction with the same sign of coefficient will tend to have a similar coefficient in the 200m case but failed in other cases. While comparing the income of different signs of coefficients, the yearly income does not always have a significant difference in any cases that refute hypothesis 3 which means more factors such as potential motivation are possibly needed for more explicit results.

Our attempt contributes to the general idea that locating nearby tourist attractions will have a significant influence on Airbnb listings’ blocking behavior in New York. As Airbnb plays an increasingly crucial role in the whole accommodation market as the supply side, the stakeholders including the government and hotel competitors can be more responsive to the blocking behavior of Airbnb according to the position. Although there are no hard-and-fast rules, the implication enables the stakeholders to consider more on how to deal with the flexible supply of Airbnb. Future work might propose more ideas on how to solve data limitations, more controlled variables and a suitable radius to dig into the motivation base of behavior patterns.

We hope you find the information in this template useful in the preparation of your submission.

Acknowledgements

I would like to thank Dr Zhu in IESE and Cora for providing research assistance throughout this paper.

References

[1] Balaguer, J., & Pernías, J. C. (2013). Relationship between spatial agglomeration and hotel prices: Evidence from business and tourism consumers. *Tourism Management*, 36, 391–400Bootorabi, F., Haapasalo, J., Smith, E., Haapasalo, H. and Parkkila, S. (2011) Carbonic Anhydrase VII—A Potential Prognostic Marker in Gliomas. Health, 3, 6-12.
[2] Blal, I., Singal, M., & Templin, J. (2018). Airbnb's effect on hotel sales growth. International Journal of Hospitality Management, 73, 85–92.
[3] Barron, K., Kung, E., & Proserpio, D. (2018). The Sharing Economy and Housing Affordability: Evidence from airbnb. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3006832
[4] Crawford, V. P., & Meng, J. (2011). New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. American Economic Review, 101(5), 1912–1932. https://doi.org/10.1257/aer.101.5.1912
[5] C .Koopman, M. D. Mitchell, and A. D. Thierer. The sharing economy and consumer protection regulation: The case for policy change. Available at SSRN 2535345, 2014.
[6] Chen, W., Wei, Z., & Xie, K. (2019). The battle for homes: How does home sharing disrupt local residential markets? https://ssrn.com/abstract=3257521
[7] Cheng, M. (2016). Sharing economy: A review and agenda for future research. International Journal of Hospitality Management, 57, 60–70. https://doi.org/10.1016/j.ijhm.2016.06.003
[8] Cohen, B., & Muñoz, P. (2017, April 6). Mapping out the sharing economy: A configurational approach to sharing business modeling. Technological Forecasting and Social Change. https://www.sciencedirect.com/science/article/abs/pii/S0040162517304365
[9] Coles, P. A., Egesdal, M., Ellen, I. G., Li, X., & Sundararajan, A. (2017). Airbnb usage across New York City neighborhoods: Geographic patterns and regulatory implications. In N. Davidson, M. Finck, & J. Infranca (Eds.), Cambridge handbook on the law of the sharing economy (pp. 108–128). Cambridge University Press.
[10] Davidson, N. M., & Infranca, J. (2016). The sharing economy as an urban phenomenon. Yale Law & Policy Review, 34(2), 16–10.
[11] Edelman, B. G., & Luca, M. (2014). Digital Discrimination: The case of airbnb.com. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2377353
[12] Farber, H. S. (2005). Is tomorrow another day? the labor supply of New York City cabdrivers. Journal of Political Economy, 113(1), 46–82. https://doi.org/10.1086/426040
[13] Farronato, C., & Fradkin, A. (2018). The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb. https://doi.org/10.3386/w24361
[14] Flash Eurobarometer 438 (2016). The use of collaborative platforms. Retrieved from file:///C:/Users/a1510623/Desktop/fl_438_en.pdf (last accessed 3 June 2017).
[15] Fraiberger, S. P., & Sundararajan, A. (2015). Peer-to-peer rental markets in the sharing economy. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2574337
[16] Gallagher, L. (2017). The Airbnb story: How three ordinary guys disrupted an industry, made billions … and created plenty of controversy. Boston: Houghton Mifflin Harcourt.
[17] Gunter, U. (2018). What makes an Airbnb host a superhost? empirical evidence from San Francisco and the Bay Area. Tourism Management, 66, 26–37. https://doi.org/10.1016/j.tourman.2017.11.003
[18] Heo, C. Y., & Blengini, I. (2019). A macroeconomic perspective on Airbnb's global presence. International Journal of Hospitality Management, 78, 47–49.
[19] Hossain, M. (2020). Sharing economy: A comprehensive literature review. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3536188
[20] Juul,M.,2015.The Sharing Economy and Tourism, from http://www.europarl.europa.eu/RegData/etudes/BRIE/2015/568345/EPRS BRI(2015)568345 EN.pdf.
[21] Karlsson, L., & Dolnicar, S. (2016). Someone's been sleeping in my bed. Annals of Tourism Research, 58, 159e162.
[22] Lee, S. K., & Jang, S. (2015). Conditional agglomeration externalities in lodging markets. Journal of Hospitality & Tourism Research, 39(4), 540–559.
[23] Liang, S., Schuckert, M., Law, R., & Chen, C. C. (2017). Be a "Superhost": The importance of badge systems for peer-to-peer rental accommodations. Tourism Management, 60, 454e465.
[24] Lutz, C., & Newlands, G. (2018). Consumer segmentation within the Sharing Economy: The Case of Airbnb. Journal of Business Research, 88, 187–196. https://doi.org/10.1016/j.jbusres.2018.03.019
[25] Marco-Lajara, B., Claver-Cortés, E., Úbeda-García, M., & del Carmen Zaragoza-Sáez, P. (2016). A dynamic analysis of the agglomeration and performance relationship. Journal of Business Research, 69(5), 1874–1879.
[26] Marshall, M. (1890). Principles of economics. Macmillan
[27] Nieuwland, S., & van Melik, R. (2018). Regulating airbnb: How cities deal with perceived negative externalities of short-term rentals. Current Issues in Tourism, 23(7), 811–825. https://doi.org/10.1080/13683500.2018.1504899
[28] Peng, Y. (2020). The analysis of blocking behavior and pricing: The case of airbnb. Business & Entrepreneurship Journal, 1–28. https://doi.org/10.47260/bej/911
[29] Quattrone, G., Proserpio, D., Quercia, D., Capra, L., & Musolesi, M. (2016). Who benefits from the "Sharing" economy of airbnb? Proceedings of the 25th International Conference on World Wide Web. https://doi.org/10.1145/2872427.2874815
[30] Tsang, E. W., & Yip, P. S. (2009). Competition, agglomeration, and performance of Beijing hotels. The Service Industries Journal, 29(2), 155–171

[31] Smith, A. (2016). Shared, collaborative and on demand: The new digital economy.

[32] Washington, DC: Pew Internet & American Life Project. Retrieved from http://www.pewinternet.org/2016/05/19/the-new-digital-economy/, Accessed date: 3 June2017

[33] S .Ranchordás. Does sharing mean caring: Regulating innovation in the sharing economy. Minn. JL Sci. & Tech., 16:413, 2015.

[34] Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on. International Journal of Hospitality Management, 62, 120e131. Airbnb.com.

[35] Wegmann, J., & Jiao, J. (2017). Taming Airbnb: Toward guiding principles for local regulation of urban vacation rentals based on empirical results from five US cities. Land Use Policy, 69, 494–501.

[36] Xie, K. L., Kwok, L., & Heo, C. Y. (2019). Are neighbors friends or foes? assessing airbnb listings agglomeration effect in New York City. Cornell Hospitality Quarterly, 61(2),128141. https://doi.org/10.1177/1938965519890578

[37] Zervas, G, Proserpio, D., &. Byers, J.(2013). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2366898

[38] Zhang, Z., & Chen, R. J. (2019). Assessing Airbnb logistics in cities: Geographic information system and convenience theory. Sustainability, 11(9), Article 2462

Appendix

T-test P-value data

200m t-test p-value		Positive				Negative			
		Cent	911_mus	MOA	Broad	Empire	High	Top	Manhattan
Positive	Cent		0.054	0.038	0.181	0.846	0.714	0.658	0.760
	911_mus			0.507	0.037	0.000	0.000	0.080	0.000
	MOA				0.106	0.001	0.007	0.056	0.006
	Broad					0.000	0.031	0.169	0.014
Negative	Empire						0.263	0.702	0.300
	High							0.454	0.899
	Top								0.478
	Manhattan								

200m t-test by size p-value		Positive				Negative			
		Cent	911_mus	MOA	Broad	Empire	High	Top	Manhattan
Positive	Cent		0.028	0.038	0.145	0.331	0.319	0.358	0.247
	911_mus			0.667	0.031	0.000	0.002	0.218	0.006
	MOA				0.181	0.067	0.076	0.206	0.104
	Broad					0.007	0.254	0.736	0.452
Negative	Empire						0.904	0.810	0.575
	High							0.855	0.701
	Top								1.000
	Manhattan								

300m t-test		Positive		Negative					
		Cent	Broad	Empire	High	Top	Manhattan	911_mus	MOA
Positive	Cent		0.015	0.819	0.353	0.530	0.475	0.002	0.000
	Broad			0.000	0.027	0.017	0.011	0.164	0.000
Negative	Empire				0.268	0.381	0.441	0.000	0.000
	High					0.175	0.766	0.001	0.000
	Top						0.228	0.005	0.000
	Manhattan							0.000	0.000
	911_mus								0.009
	MOA								

300m t-test by size		Positive				Negative			
		Cent	Broad	Empire	High	Top	Manhattan	911_mus	MOA
Positive	Cent		0.015	0.147	0.167	0.372	0.093	0.001	0.000
	Broad			0.053	0.101	0.240	0.232	0.102	0.001
	Empire				0.991	0.835	0.596	0.000	0.000
	High					0.838	0.653	0.002	0.000
Negative	Top						0.633	0.039	0.002
	Manhattan							0.006	0.000
	911_mus								0.047
	MOA								

400m t-test		Positive				Negative			
		Cent	Top	MOA	Broad	Empire	High	Brook	911_mus
	Cent		0.921	0.845	0.005	0.871	0.403	0.317	0.432
Positive	Top			0.000	0.020	0.988	0.432	0.309	0.038
	MOA				0.039	0.000	0.000	0.533	0.039
	Broad					0.000	0.015	0.954	0.814
	Empire						0.192	0.279	0.000
Negative	High							0.477	0.052
	Brook								0.906
	911_mus								

400m t-test by size		Positive				Negative			
		Cent	Top	MOA	Broad	Empire	High	Brook	911_mus
	Cent		0.078	0.760	0.007	0.099	0.347	0.055	0.274
Positive	Top			0.000	0.495	0.572	0.274	0.301	0.156
	MOA				0.040	0.002	0.001	0.744	0.137
	Broad					0.063	0.024	0.435	0.302
	Empire						0.410	0.187	0.002
Negative	High							0.120	0.001
	Brook								0.647
	911_mus								

Number of observation in 200m and 400m radius cases

Listing number in radius	200m	400m	400m/200m
Broadway	183	558	3.05
Brooklyn_bridge	0	14	
Cent_park	30	253	8.43
Empire_state_bulid	259	921	3.56
High_line	168	531	3.16
Manhattan_skyline	183	640	3.50
Museum_of_art	7	59	8.43
National_911	60	491	8.18
Statue_of_liberty	0	0	
Top_of_rock	35	313	8.94
Average	92.5	378	5.91