Airbnb and Its Effects on the Housing Market of New York City

Gaurav Bhardwaj, Baiyue Cao, Unisse Chua, Te Du, Nina Nurrahmawati

Introduction

I. Overview of Airbnb

Airbnb is a company that helps people rent out their spare space in their homes. The company is currently valued at $31 billion and has transformed the hospitality industry by offering cheap alternatives to hotels (Lunden, 2017). Recently, there has been a rising controversy about whether Airbnb is serving our cities for good or worse. In 2016, three US governors asked the Federal Trade Commission to investigate the commercial listings on Airbnb for its possible effects on the rental prices (Stulberg, 2016). Airbnb also gained the reputation on social media as the major contributor to gentrification and rising housing crisis (Rosenburg, 2017; Monroe, 2014).

II. The Problem of Airbnb in New York City

Commercial listings on Airbnb specifically refer to unoccupied units that are listed year-round for short-period rentals. These commercial listings are accused of exacerbating the existing housing shortage in many cities by causing a decrease in the long-term rental units listings on the housing market (Edward, 2016). In October 2016, New York State governor Andrew Cuomo signed a bill saying it is "illegal to advertise entire unoccupied apartments for less than 30 days on Airbnb" (Hawkins, 2016). With tighter regulations in place, people continued to post their apartments on Airbnb and have also been fined (Carman, 2017). Therefore, evaluating the effect of airbnb on local housing prices is very important. In this research, we plan to study the effects of the presence of Airbnb on gentrification in major metropolitan areas through the lens of housing prices in New York City. This research aims to find the correlation between airbnb listing numbers and housing sales prices in different area of the New York City.

Data Acquisition and Cleaning

This project utilizes three main datasets, the DOF Annualized Rolling Sales, Inside Airbnb listings and PLUTO data modelling and visualizations.

I. *DOF Annualized Rolling Sales Data*

The Department of Finance provides the data of all property sales in New York City annually and this data can be accessed through the NYC Open Data Portal as well as the official webpage of DOF. This data set includes information on sales including neighborhood, building type, square footage, and many more.

For this study, we considered only residential sales from 2015 and 2016. We considered the properties under tax class 1 and 2 as residential and we also did not include sales with zero sales price since we needed the price as a point of comparison. We then aggregated the data by BBL and got the median price per BBL. After looking at the distribution of the median sales prices, due to the large standard deviation, we decided to take the log transformation of the sales prices. To remove outliers that could potentially skew the results of our regression, we followed a statistical approach by removing the tail of the distribution 2 standard deviations away from the mean of the log transformed sales prices as it looked to follow a normal distribution. After the data cleaning, we ended up having 42,278 data points for 2015 and 41,683 data points for 2016. Fig. 2 and Fig. show the distributions before and after cleaning.

II. *Inside Airbnb Data*

In order to find the correlation between the number of Airbnb listings and property sales prices, we used data from Inside Airbnb, a crowd sourced listing database that shows monthly Airbnb listings in New York City along with their geographical coordinates.

Since listings were scraped monthly, we first needed to remove the duplicate listings and considered only all unique listings for the year. After dropping duplicates, we had 147,421 listings in the span of two years. Fig. 4 in the appendix shows the Airbnb listing locations with the heat map of the sales prices from DOF.

To attain the counts of Airbnb listings surrounding a sale, we first drew a 500-meter radius around the centroid of each BBL and conducted a spatial intersect join for the geographical area and the Airbnb listings. All these were merged together to have a data set of aggregated residential sales at the BBL level with the ZIP code, neighborhood, borough, the year of sale and the count of Airbnb listings within a 500-meter radius of the property.

Methodology

To evaluate the price effect of Airbnb presence for housing prices, we constructed a multivariate linear regression with the log-transformed housing sales prices as the endogenous variable. We then regressed log-transformed sales prices on the Airbnb counts within 500-meter radius circle for each BBL, along with locational control variables and time control variables to isolate the spatial-temporal effects such price movements related with locational difference or yearly inflations. The coefficient we observe in front of the Airbnb count can then be interpreted as the percentage change of property sales prices with one additional unit of Airbnb listing within the 500 meter radius circle.

Although we aggregated the sales prices by taking the median of all sales prices in that year of specific BBL, we were still able to get meaningful insights from this abstraction, considering how small each BBL is and the property types we are working with are only tax class 1 and 2, which are residential properties that are either 1-3 unit houses, cooperatives or condominiums. By assumption different units in the same BBL should not have dramatically varied sales prices, or dramatically varied attributes such as sizes or amenities. Since we are taking the yearly median, we do expect monthly price fluctuations, however, because of the infrequent nature of property sales, we assume that the valuation of property should be relatively stable before and after the point of sale. Therefore taking the BBL median sales prices as indicators for all properties sold in each BBL can reasonably represent the market price condition of each BBL.

In terms of picking the geographical boundary of price effect, we make the assumption that Airbnb activities within 500 meters of specific property will have effects on its sales prices. Since a New York City block is usually 80m X 274m, a 500 meter radius circle covers the adjacent blocks of the sold property, which we assume will have the strongest price effect on sales prices. Also a 500 meter radius circle is what we assume how people perceive as their immediate neighborhood.

Results and Interpretation

After running the regressions we observed that Airbnb presence within 500 meter radius circle of particular property sold has minimal positive effects on property sales prices, and the result is statistically significant, controlling by Borough and ZIP code. Although not statistically significant, we observed the same result for the model that controls by neighborhood.

Our result tells us that Airbnb activities within 500 radius circle distance does not have significant impact on housing prices for different location across New York City. Despite how the presence of Airbnb seemingly correlate with neighborhood gentrification, its observed effect, through the lens of property prices, was not captured. Our interpretation is that people who are buying family houses or apartments must have been planning on the property transaction for a very long time, and property pricing often takes neighborhood infrastructure or property specific features into consideration, rather than unobserved indirect features such as Airbnb presence. Also, the effects of Airbnb is more likely to be captured through the rental market, because people who are paying high rent are more likely to rent out their rooms to relief rent burden. Therefore our hypothesis that in a short time period, a rise of Airbnb presence causing neighborhood gentrification and rising residential property prices is not supported by our current data and model.

Limitations and Next Steps

Because of the availability of the Airbnb listing data, we can only build models over two years of time frame, yet the lack of sufficient repeated sales data within the two year time frame makes it hard to establish causality model that compares before and after prices. Therefore, in future studies, we will collect Airbnb data before 2015 and use similar properties that have different levels of Airbnb activities as control and treatment, capturing price effect on a more direct and granular level.

Even though Airbnb's effect on property price is shown to be minimal based on our 500 meter radius model, its actual impacts on the neighborhood could still be significant. Because a small change in real estate price could lead to increase in rent and increase in general living expense in the area. As our literature review suggests that for some neighborhood even an increase in rent as small as fifty dollars is enough to distress the economic situations of low-income residents and therefore leads to gentrification. So to quantify this stress, a similar methodology could be applied to investigate the Airbnb effects on rent cost in the neighborhood. It is also interesting to build a living-expense index for other general spending including transportation, food, and others to show if Airbnb can have a significant impacts on those as well.

Furthermore, it is all but more important to answer the question of what are the underlying causes for an increase in property price. As a case study in Amsterdam suggested, the use of Airbnb as an extra source of income, the home buyers will be able to afford higher mortgage rates. This could potentially be an underlying multiplicative factors in further pushing the real estate upward. These effects, however, are more difficult to measure. But it could potentially be scoped out if the mortgage records for home buying in the area could be obtained. Another possible explanation is that Airbnb could raise property values by bringing in commercial prosperity and more mobility in population. To investigate such effects would require data on commercial activities of the area through Yelp or other census data.

References

1. Hawkins, A. (2016, October 21). Airbnb is now banned from listing short-term rentals in New York. The Verge. Retrieved from https://www.theverge.com/
2. Carman, A. (2017, February 7). New York City issues first illegal Airbnb fines. The Verge. Retrieved from https://www.theverge.com/
3. Lunden, I. (2017, March 9). Airbnb closes $1B round at $31B valuation, profitable as of 2H 2016, no plans for IPO. Tech Crunch. Retreived from https://techcrunch.com/
4. Stulberg, A. (2016, August 24). Airbnb Probably Isn't Driving Rents Up Much, At Least Not Yet. FiveThirtyEight. Retrieved from https://fivethirtyeight.com/
5. Edwards, J. (2016, October 20). Here's exactly what Airbnb does to rent in popular cities. Business Insider. Retrieved from http://www.businessinsider.com/
6. Cox, M. (2017, March 1). The Face Of Airbnb, New York City. Retrieved from http://insideairbnb.com/

Appendix

Table 1: Regression Result:

	log_price I	log_price II	log_price III
Locational Control	Borough	Zipcode	Neighborhood
Intercept	5.88***	5.94***	5.79***
	(0.00)	(0.26)	(0.18)
Airbnb count	0.00***	0.00***	0.00
	(0.00)	(0.00)	(0.00)
Time Control(2016)	0.01**	0.02***	0.03***
	(0.00)	(0.00)	(0.00)
N	32784	32784	32784
R2	0.16	0.28	0.29

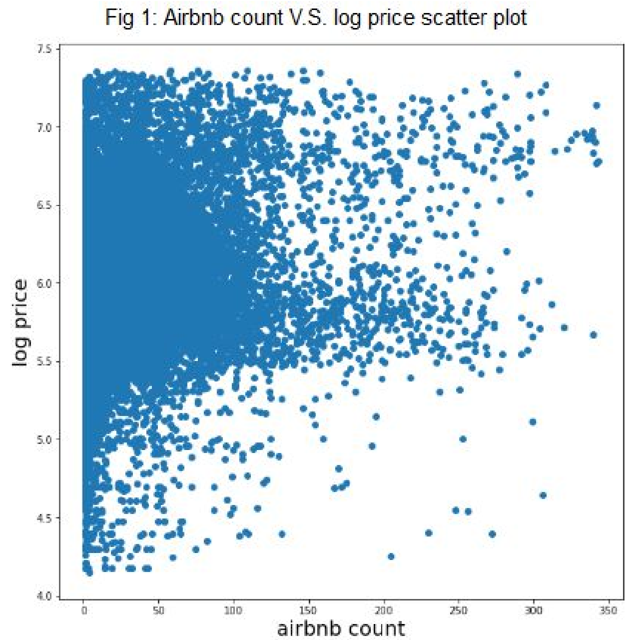
Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01

Table 2: Sample data frame

	BBL	ZIP_CODE	NEIGHBORHOOD	SALE_PRICE	YEAR	log_price	Borough	airbnb_count
0	1000280001	10004	FINANCIAL	713500.0	2015	5.853394	MN	186
1	1000290026	10004	FINANCIAL	1325000.0	2015	6.122216	MN	110
2	1000300001	10004	FINANCIAL	10000000.0	2015	7.000000	MN	56
3	1000390040	10005	FINANCIAL	20125000.0	2015	7.303736	MN	148
4	1000640008	10005	FINANCIAL	1937500.0	2015	6.287242	MN	52

Table 3: Summary statistics

	SALE_PRICE	log_price	airbnb_count
count	32784.00	32784.00	32784.00
mean	1334804.11	5.87	23.05
std	2263861.61	0.43	38.96
min	14000.00	4.15	1.00
25%	420000.00	5.62	1.00
50%	700000.00	5.85	5.00
75%	1200000.00	6.08	30.00
max	23009004.00	7.36	343.00



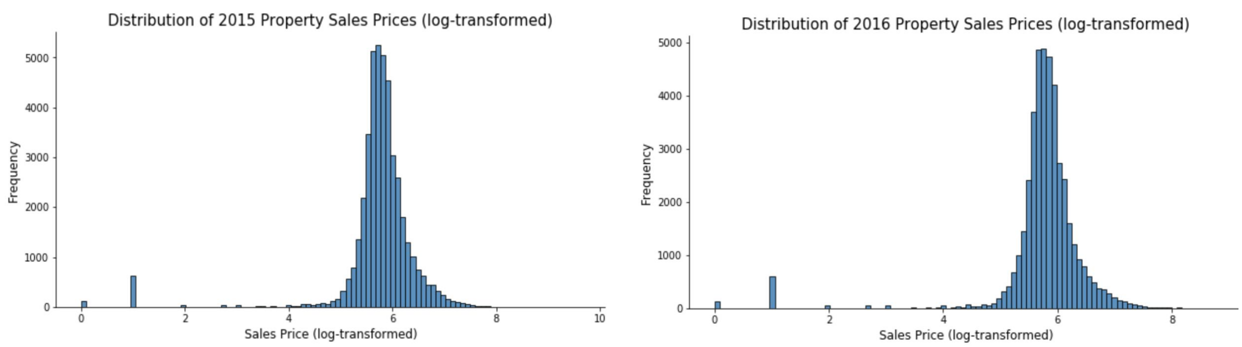


Fig. 2. Distribution of 2015 and 2016 Property Sales Prices (log-transformed) prior to cleaning

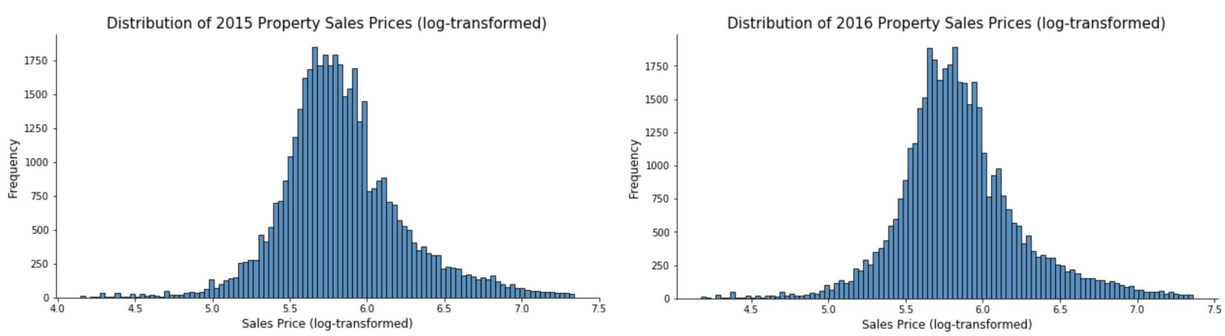


Fig. 3. Distribution of 2015 and 2016 Property Sales Prices (log-transformed) after removing outliers

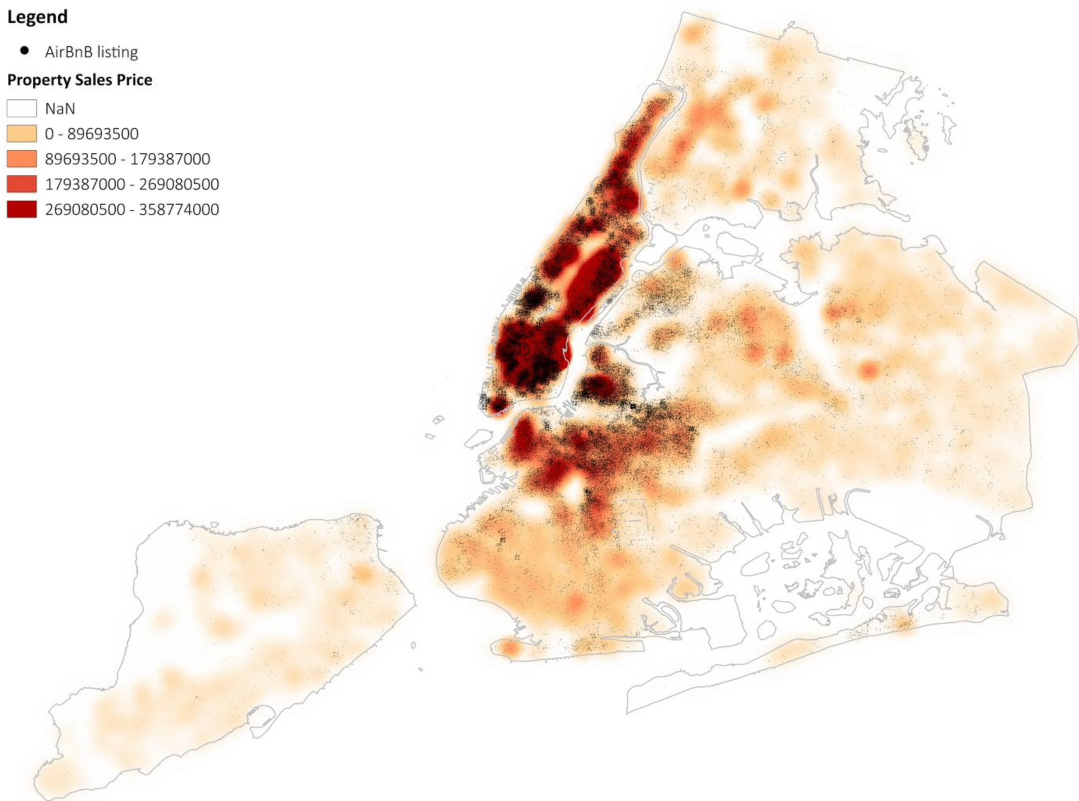


Fig. 4. Property Sales Price Heatmap and AirBnb listing location

Roles and Contributions of each member

Everyone in the team worked collaboratively

- **Gaurav Bhardwaj** worked on the data section of the project proposal and the spatial join for the data processing part.
- **Baiyue Cao** suggested the idea of doing a topic on Airbnb and came up with the methodology and approach on how to tackle the issue. He handled running the models and the discussion of the results.
- **Unisse Chua** conducted the background research on Airbnb and the effects reported on New York City and worked on part of the introduction. She also worked on cleaning the DOF and Airbnb data and discussed this in the paper.
- **Nina Nurrahmawati** worked on the introduction, found a research paper with a similar study done by the Department of Economics in Williams College and provided the shapefile needed for the spatial join. She also provided the map plots in the appendix.
- **Te Du** did background research on Airbnb and worked on part of the introduction. He also worked the spatial join for the data processing and writing the next steps of the research.