THE PROFITABILITY OF AIRBNBS RELATIVE TO LONG-TERM RENTALS IN LOS ANGELES

By

RICHARD BURDETTE HOUGHTON III

A thesis submitted in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE IN APPLIED ECONOMICS

WASHINGTON STATE UNIVERSITY
School of Economic Sciences

AUGUST 2020

To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of RICHARD BURDETTE HOUGHTON III find it satisfactory and recommend that it be accepted.

____

Joseph Cook, Ph.D., Chair

____

Thomas Marsh, Ph.D.

____

Liang Lu, Ph.D.

THE PROFITABILITY OF AIRBNBS RELATIVE TO LONG-TERM RENTALS IN LOS ANGELES

Abstract

by Richard Burdette Houghton III, M.S.
Washington State University
August 2020

Chair: Joseph Cook

With the introduction of Airbnb, the short term rental market has been greatly expanding. It is assumed that profit-maximizing property owners would only choose to use a short-term platform if expected net revenues from doing so exceeded the expected net revenues of renting to a long-term tenant. Information asymmetries in occupancy rates, net nightly rental fees, and other short-term rental factors, however, might lead property owners to overestimate the profitability of short-term rentals. This study estimates the yearly revenues earned from 19,970 short term rentals in metro Los Angeles from March 2018 to February 2020 using data scraped from the Airbnb website and compares it to yearly revenues over the same time period for 4,443 comparable properties in the long-term (annual lease) rental market. It was found that a property listed on Airbnb is expected to generate $17,027 less per year than a comparable long-term rentals, on average. This result is sensitive to assumptions about Airbnb occupancy rates, but the gap remains even under the most optimistic occupancy assumptions. This difference in revenues indicates that property owners may be over-estimating the profitability of Airbnb rentals or that the annual costs of operating an Airbnb is lower than a long-term rental.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Airbnb was started in 2008 as an alternative to other short-term housing such as hotels and bed and breakfasts. Over the past decade, Airbnb has grown immensely, becoming one of the most successful start-ups of the decade. However, their success has come with some externalities to other industries, particularly in terms to the housing market. When homeowners rent their houses full-time on Airbnb, they effectively remove that property from the long-term rental market and move it into a market that stands in between the hotel market and the long-term rental market. With the STR (short term rental) market being so new, its impacts were unknown and unsure up until the past few years. One study has shown that Airbnb listing density is directly correlated with an increase in rental prices (Lee 2016). While these studies focus more on the impact of rental prices within cities and neighborhood, there is very little to no focus on the revenues being earned through Airbnb. These studies have an underlying assumption that property owners maximize profits when choosing to rent through Airbnb rather than a long-term rental. Factors such as occupancy rate, price of the property being offered, whether that be nightly or longer, and the costs of placing a house on Airbnb such as cleaning fees are usually sparsely touched on, if at all.

Table 1.1: Costs between STRs vs. LTRs

	AIRBNB (STR)	LTR
Transaction	Easy and low-cost to signup but with ongoing costs per transaction (commission).	Upfront costs – advertising, negotiate lease, security deposit into escrow.
Legal risk	In theory, liability covered under Airbnb's umbrella	Full liability
Payment risk	Relatively low? Payment via Airbnb before check-in	Higher – tenant can default; require eviction proceedings.
Maintenance	Requires cleaning after each visit. Professional cleaning services.	Requires more "deep" but infrequent cleaning when apartment turns over. Professional cleaning services.
Option value	Retain option to remove it from Airbnb and rent long-term or use personally	Commit to year-long lease with tenant.

This paper will focus on estimated earnings differentials. Property owners will, of course, seek to maximize profits, so any differences in the costs of operating LTRs vs STRs is relevant. Unfortunately I have no cost data to draw upon. Instead we summarize the types of costs associated with operating an Airbnb/STR versus a long-term rental in Table 1.1. While there are distinct differences, especially in categories such as revenue certainty and legal risk, I feel that many of the costs seem equivalent at a first approximation. One key difference might the cost of evictions. In my study site of Los Angeles, the eviction rate in 2015 was 3% (Badger et al 2015), suggesting that this distinction will be too low to create any large difference between STR revenue and LTR revenue.

There may, however, be a significant market failure present within the short-term rental market: asymmetric information. In the long-term rental market, it is possible to search secondary databases (e.g. newspaper listings, Zillow, etc) and discover the general price range for a particular type of property in order to match one's asking price as close as possible to market value. However, in the short-term rental market, specifically Airbnb, information that would be needed to accurately estimate pricing, such as average nights stayed, is not made public, forcing owners to price their properties as best as possible. This lack of data from Airbnb's side also greatly impacts a property owner's ability to estimate their possible occupancy rates nearly impossible. Although an Airbnb operator can gather information on market prices, there is much less certainty in their expected total revenue because there is no easy way to project occupancy rates. An owner of an LTR has high certainty in annual revenues since they lock the tenant into a negotiated per-month rent for the duration of the contract. Although renting short-term allows the owner to price dynamically through the year (e.g. higher during tourist seasons, or lower during economic contractions), they must assess the probability of her occupancy expectations being incorrect. Under these circumstances, it is therefore possible that this information asymmetry drives owners to make sub-optimal choices, earning lower profits under STR than LTR, again, assuming no major differences in costs.

The goal of this paper is to test whether yearly revenues gained through renting a property on Airbnb are greater or less than if they were to rent their property through a long-term rental. In this study, I will analyze annual earnings of Airbnbs and long-term rentals in the city of Los Angeles over the time period starting in March 2018 and ending in March 2020. Los Angeles was chosen due to the high density of Airbnbs and availability of short-term rental data . The contribution to the existing literature examining rents with respect to Airbnb will be a deeper analysis into the revenues earned rather than the impact on the price increases due to Airbnb. Controlling for property characteristics, I estimate that owners renting their properties full-time on Airbnb earned less money than if they were to rent their property through an annual 12-month lease. This was robust to using a more conservative occupancy rate.

The organization of the remainder of the paper is as follows. First, I will discuss the relevant literature for this topic and how this study differentiates itself from other similar studies. Second, I will explain the data used in this study, limitations of that data, and the methods used to work with that data, including summary statistics. After briefly describing the model being used in this study, I present results and conclude with next steps.

CHAPTER TWO: LITERATURE REVIEW

Since this paper is looking at the housing market and the characteristics of a property, I apply a hedonic pricing model. A hedonic pricing model allows a measure of individual characteristics of a specific good, such as the numbers of bedrooms or bathrooms of a property (Sirmans et al (2005)). Since hedonic pricing models work most effectively when looking at one particular market, this model fits very well for the type of data being used in this study, as a look at the housing market in one city. Sirmans et al (2005) also recommends the use of a semilog model to measure the impacts of individual characteristics when housing values can vary substantially. For example, the marginal impact of an additional bathroom in a house worth \$300,000 may be different than in a house worth \$3,000,000. The semilog model allows a measure of percentage change on the dependent variable with a unit of change rather than a strict value shift. Given the particular sets of data being used for this study, with being over two years and in a large city where such a disparity can occur, a semilog model will allow a more even estimation across the entire city of Los Angeles.

A similar study has been done in Los Angeles, but focused more on the impacts on rents and policies to regulate Airbnb's growth within Los Angeles. Lee (2016) looked at the seven neighborhoods with the largest density of Airbnbs, which at the time encompassed nearly half of all Airbnb listings in Los Angeles for 2014. By properties originally for long-term leases being converted into Airbnbs, Los Angeles was hit with a reduction in supply of long-term rentals. Lee found that within these seven neighborhoods, a drop of 3% to 12.5% of the rental market supply lead to a supply shock due to the market being unable to effectively adjust. From this observation, Lee states that under a simple economic model with a constant demand curve and a relatively flat supply curve with a low price elasticity coefficient would lead to an increase in rents by 0.2% for each percent decrease in supply. Furthermore, with Airbnb effectively removing about 7,400 houses from the market with it expected to increase in years to come, affordable housing as a whole will begin to be impacted by this, possibly increasing rental prices even further.

The paper most closely related to this one is Coles et al (2017), who examined differences in short term rental earnings versus long term earnings. Coles et al (2017) looks at New York over the span of 2011 to 2016 using Airbnb data aggregated to the census-tract level (a co-author was an Airbnb employee) and the Zillow Rent Index data for census-tract level data on long-term rental prices. With this census-level data, they were able to determine qualities such as the median household income for the areas with a high density of Airbnbs and the demographics of these areas. They found that centrality to major landmarks such as the Empire State Building tends to create a higher density of Airbnbs in general along with the percentage of highly booked Airbnbs, measured by Airbnbs booked for 180+ days. They ran a regression looking at median income, median income squared, distance to the Empire State Buildings, and dummy variable controlling for location. Through this, they found that short-term rentals from the period of 2011 to 2016 were not as profitable as many assume, suggesting that in order for a short-term rental to break even with a median estimate, they would need to rent out their property for 216 days a year; the median during June 2017 was a meager 46. They note that revenues have drastic difference between neighborhoods, which changes where the break-even point lies and that lower income and middle-upper income neighborhoods have a better profit margin for short-term rentals.

Finally, Horn and Merante (2017) utilizes a hedonic pricing model to show that an increase in Airbnb listings increase asking rents within Boston. Like this study, they used scraped insideairbnb data from July 2015 to January 2016 and rental data from Rainmaker Insights Inc. from September 2015 to January 2016 and created a hedonic pricing model with a semilog specification using number of bedrooms, number of bathrooms, square footage, and the number of new rental units in each census tract in Boston, along with information on crime rate and building permit and restaurant permit issuances. Their model also includes fixed effects for month and tract. From their results, they found that the number of bedrooms, number of bathrooms, and square footage all had a positive impact on rental prices. Horn and Merante found that a one standard deviation increase in Airbnb density lead to a 0.4% increase in Boston rents.

Our study builds upon the previously mentioned studies in a number of ways. One of those ways is through the focus of the study. While there have been studies that have looked at a specific month of Airbnb revenues in dollar values, such as the study done by San Francisco's Budget and Legislative's Office, no study looks at the yearly revenue of Airbnbs, therefore capturing monthly variations in revenue. Second, our study uses more recent data (in case property owners have updated their information on expected profitability of short-term rentals) and studies the issue in a second metropolitan area. Another method is the model specification. In the studies that do utilize hedonic pricing models, none of the models control for the specific kind of residence being looked at, such as an apartment or single-family residence. By including dummy variable that control for this, we can determine which types of properties have the most benefit being rented through Airbnb.

CHAPTER III: DATA

The data on Airbnb listings was gathered from insideairbnb.com, a website where Airbnb data has been regularly scraped and uploaded since 2015. The datasets from Insideairbnb span from March 2018 to February 2020 in monthly scrapes and contain information such as the room type, property type, and payment information such as security deposits and cleaning fees. However, there are several limitations with this dataset. One such is that the scrapes uploaded to Insideairbnb are snapshots and only representative of the time of the scraping rather than an entire period. Due to this, there is the possibility that within any month properties both enter and exit the market that are not captured within each monthly scrape. An assumption will be made that each scrape is representative of the specific month. Another assumption that will have to be made is that each host within each scrape has not taken their property off of the market for extended periods of time. Through the nature of the scrape, we are unable to determine whether each host has a consistent schedule for days where their property can be rented.

The long-term lease data was gathered from the California Regional Multiple Listing Service (CRMLS) database. From the period of March 2018 to December 2019, 4,433 properties consisting of apartments, townhouses, and single family residences were rented within the city of Los Angeles. There are some limitations to this data, particularly because of the timeframe. There was no data available for January 2020 or February 2020, so the months of January 2019 and February 2019 were used as proxies by scaling the 2019 rent by the average city-wide inflation rate from 2019-20 of 4 percent (HCIDLA). The other issue is how bathrooms are categorized within the dataset. There are values for full bathrooms, three-quarters bathrooms, half bathrooms, and one-quarter bathrooms, each representing different bathroom sizes and amenities. An assumption will be made in that each of these different bathrooms impacts the rental price of the property the same, so the number of bathrooms, three-quarters bathrooms, half bathrooms, and one-quarter bathrooms will be summed to determine the actual number of bathrooms of the property.

There are some limiting factors of this dataset that also need to be addressed. The majority of the properties from this dataset were pulled from the CLAW MLS, or the Combined Los Angeles/Westside Multiple Listing Service. These listings are not associated with any particular firm due to the open nature of the CRMLS, but there is one shared factor between all of these listings: all of the listings only include properties that were handled by agents rather than handled directly by the owners. Due to the cost of hiring an agent, which is about 6% of a year's lease on average, this could be seen as a barrier of entry into this particular dataset. This could create a possible selection issue, in which only properties where owners are willing to pay an agent are captured. Still, I believe the dataset is representative of market rental prices for long-term leases across a broad range of apartments, homes and townhouses.

One other assumption that will need to be made is the costs needed to upkeep both short-term rentals and long-term rentals. There are numerous costs such as paying a rental company if the owner is having an outside company manage their property and maintenance costs. As discussed above, other less apparent costs come from renters defaulting on their payments, effectively removing any possible revenue earnings for at least one month, if not more. These costs will be assumed equal within this study and the predicted impacts, if they were to be present, will be talked about in the discussion section.

Although calculating annual revenues from a long-term lease is straightforward (12 x reported monthly rent), the calculations for short-term rentals require more assumptions. This is particularly true since the full set of data on the Airbnb platform, such as actual occupancy rates, is not publicly available. First, I narrow down the dataset to more closely match properties in the data obtained from the CRMLS database. This involved limiting the Airbnb data to properties that rent out the entire property rather than just a spare room in townhouses, apartments, or single family residences. I kept records listed in the Airbnb data as “Entire home/apt” in the room_type column and “Apartment”, “House”, or “Townhouse” in the propery_type column. I include, however, properties that may not have been rented for a full twelve months. Overall, there are 111,280 unique properties that are likely to be full-time, “entire home/apt” rentals that fall under the specification “Apartment”, “House”, or “Townhouse” through Airbnb in the time period from March 2018 to February 2020, with 70,290 observations being dropped.

Second, to calculate the occupancy rate and utilization model, I use a combination of the methods used by San Francisco Planning Department and the City and County of Sacramento’s Budget and Legislative Analyst’s 2015 policy analyst report on short term rentals as a base for my model. In these studies, these departments calculated occupancy and utilization rates based on this model:

[([(# Reviews per listing / (Review Rate) * Average # of nights a listing is booked] / # Days that have passed since the first review = occupancy rate of a property

In this model, they utilize the current number of reviews along with the time where the Airbnb was first utilized, which was either the day the host joined the platform or when the first review was posted. By using a measure of the property’s time on the market as a measure for occupancy, there are a couple assumptions made for this to hold. First of all, any properties offered during this time period should have been consistently offered on the market since the first review date or the date the host joined. This in turn implies the owner of the Airbnb is not necessarily profit maximizing, and even if their profits are poor, have kept their property on the

Airbnb market since its debut. The other assumption being made is that as time passes, a property's occupancy rate is inherently tied to the amount of reviews it is getting. For example, a property with a high amount of reviews in March 2018 but a small amount of new reviews will most likely have a higher occupancy rate than a relatively new Airbnb that just entered the market in March 2018 that is getting many reviews per month. With these assumptions along with data that Insideairbnb has scraped, I have adjusted this model to more closely measure the amount of revenue generated per month.

One portion of this model that is somewhat unclear is the review rate, or the percent of Airbnb customers who choose to leave a review for the property's owner on the Airbnb platform. In order to estimate the actual amount of booking for each Airbnb property, I inflate the number of reported of reviews by a specific review rate. For this particular study, review rates of 50% and 30.5% will be used. 30.5% is the review rate used by the Budget and Analyst Office. The 50% review rate used by Insideairbnb is based on the average between the 30.5% review rate used by the Budget and Analyst Office in their study on San Francisco and the 72% review rate reported by Airbnb's CEO Brian Chesky. The 50% review rate will be used as a baseline, while the 30.5% review rate will be used as a less conservative estimate.

My revised revenue model is:

Nights * Price = estimated monthly revenue of a property

where Nights = (# New reviews / (Review Rate)) * Average # of nights a listing is booked)

Due to the panel nature of this data, I am able to look more closely at the reviews being left per month and use that as the framework for how many nights an Airbnb is occupied for. An assumption this model has to make is that reviews are left within the same month that the property is rented. The variable "New reviews" is calculated by taking the difference between the number of reviews of the current month and the number of reviews in the previous month. In months where one of these two numbers are missing, the average number of reviews over the two year span starting in March 2018 is used. Another issue with this measurement is that it allows for properties to be rented out for a negative number of days if a review is deleted or for more days than there are in a month. This possibility is controlled for by hard capping the number of nights by the number of days during that month and setting a floor of zero days for any property.

Next, we need assumptions about the usage of average nights for a booked listing in the model. In the San Francisco Planning Department's study, they used both the minimum nights per stay and the average nights per stay, but noted that using minimum nights per stay creates a less conservative long-term estimate. Therefore, average nights per stay will be used instead. This value will be one reported by the Southern California policy manager for Airbnb, which is 4 (Madans 2019).

With monthly revenues determined, I then add revenue earned through the cleaning fee. A quarter of the cleaning fees will be added in as revenue to account for the possibility that cleaning fees are a hidden source of Airbnb profitability for the property owner (i.e. the cleaning fees overestimate actual cleaning costs). Since the cleaning fee is applied to each separate visit, the number of nights calculated above will be divided by the average nights per stay of 4 to determine visits, and then multiplied by a quarter of the cleaning fee as such:

$$.25 * (\text{Nights}/4) * \text{Cleaning Fee} = \text{Cleaning revenue}$$

As an example of how this calculation will work, a calculation for the 50% review rate and 30.5% review rate will be presented here. The listing chosen for this example has 3 new reviews during the June 2018 scrape. The daily price for this property is 303 dollars and the cleaning fee is 100 dollars.

50% Review Rate:

(3/0.5)*4 = 24 (nights booked during the month of June)

24 * 303 = 7272 (monthly revenue of the property)

24 / 4 = 6 (visits for the month)

.25 * 100 * 6 = 150 (portion of cleaning fee taken as revenue)

7272 + 150 = 7422 (Monthly revenue assuming cleaning fee profit)

30.5% Review Rate:

(3/.305)*4 = 39 (nights booked during the month of June) (hard capped at 30)

30 * 303 = 9090 (monthly revenue of the property)

30 / 4 = 7.75 (visits for the month)

.25 * 100 * 7.75 = 193.75 (portion of cleaning fee taken as revenue)

9090 + 193.75 = 9283.75 (Monthly revenue assuming cleaning fee profit)

From these calculations, it is easier to see how assuming a smaller review rate creates a less conservative estimate for Airbnbs.

Next, I clean the Airbnb data to further match the CRMLS dataset. This involved setting dummy variables to control for if the property is an apartment or townhouse and dropping all zip codes that are not present in the CRMLS dataset. This lead to another 71,002 observations being dropped. Once that is done, I clean the data to remove any outliers (such as properties with extremely high amenity levels) based on the daily price of the property. This is done through three standard deviations away from the mean and leads to 559 observations being dropped. Finally, any properties that were not on the market for at least 12 months are dropped to attempt to capture renters that are constantly in the market and saw reasons to stay, which lead to 19,479 observations being dropped. This cleaning created a dataset with 19,970 unique properties.

For the long term-rental properties in the CRMLS dataset, we used the Google API through R package ggmap to geocode zip codes. During this step, the geocode was unable to place a zip code to 13 properties with 1 property being identified as out of the state, so those properties were dropped. We assume that for each long-term rental the renter has signed a year-long lease.

CHAPTER FOUR: SUMMARY STATISTICS

Table 4.1: Descriptive Statistics (Airbnb)

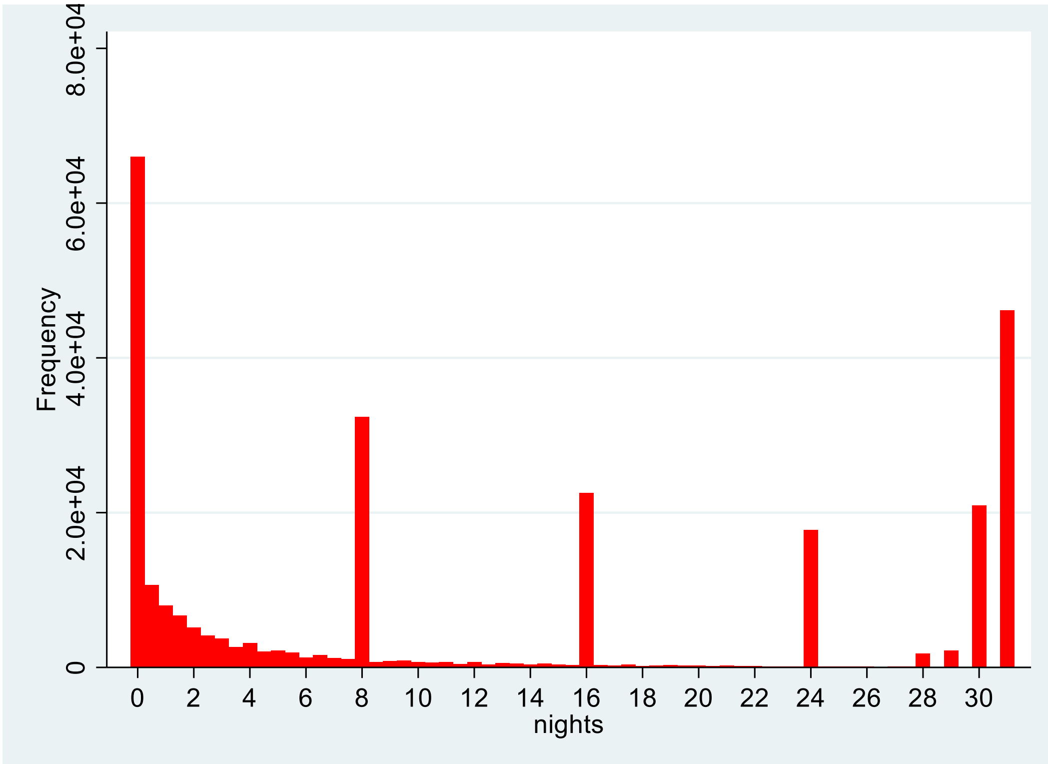
Variable	Obs	Mean	Std.Dev.	Min	Max
annualrev	19970	22318.85	28203.25	0	464000
annualrev_~h	19970	27796.81	35202	0	510000
bedrooms	19955	1.606	1.204	0	15
bathrooms	19960	1.519	.893	0	13.5

Table 4.2: Descriptive Statistics (Long Term Rentals)

Variable	Obs	Mean	Std.Dev.	Min	Max
annualrev	4433	77233.55	52424.71	4620	349000
annualrev_~h	4433	77233.55	52424.71	4620	349000
bedrooms	4433	3.015	1.137	0	22
bathrooms	4433	2.709	1.369	0	21

Before running the model, summary statistics are run to see how our variables are behaving as expected. Based on these results, yearly revenue for the Airbnb rentals averages at about $22,318 for the 50% review rate, which leads to an average monthly revenue of $1860. For the 30.5% review rate, yearly revenue averages at $27,796, or $2316 per month. Within the LTR data, average monthly yearly revenue comes out to $77,233, or $6436 per month. If we look at bedrooms and bathrooms, long term rentals tend to have more of both amenities, which could explain the raw difference between average revenues for short-term rentals and long-term rentals. We control for these characteristics, along with zip code fixed effects, in the regression model.

Figure 4.1: Nights Stayed Per Property Per Month (50% Review Rate)



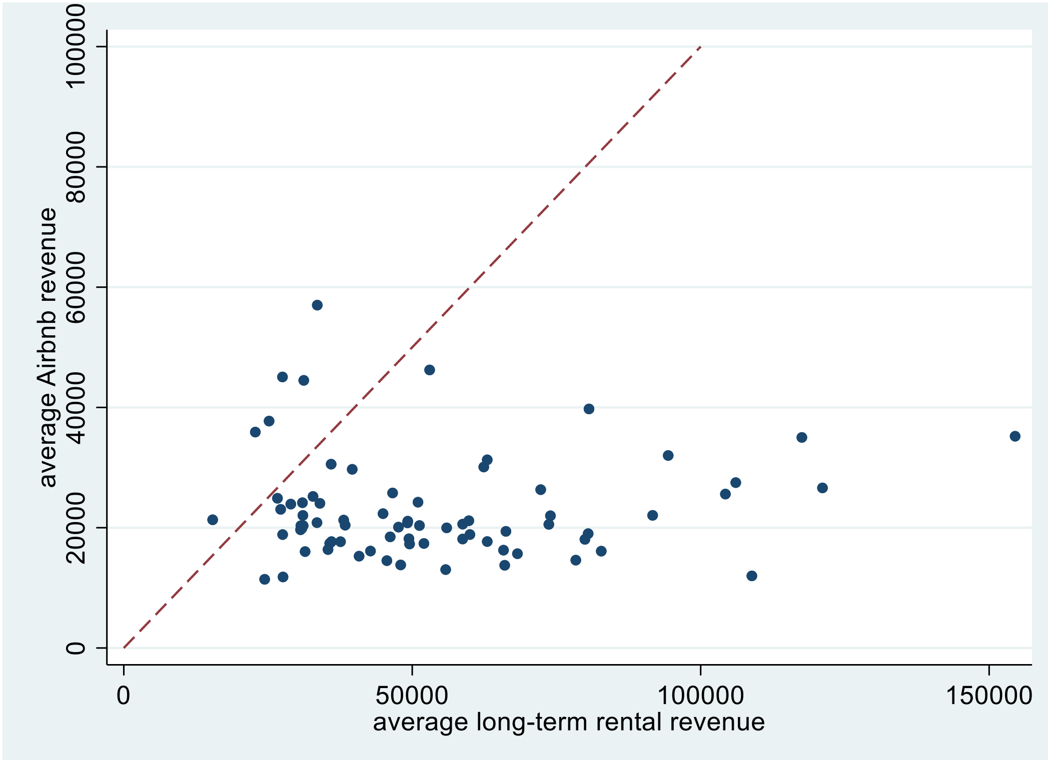
This figure visualizes the nights stayed during each month at the 50% review rate, capturing 277,050 different observations. There are a significant amount of observations focused at zero, eight, sixteen, twenty-four, thirty, and thirty-one days of usage. Given that this data is truncated at zero and capped at the number of days within each month, spikes at zero, thirty, and thirty-one makes sense. The other notable points each correlate to a specific amount of visits, being two visits for eight nights, four visits for sixteen nights, and six visits for twenty-four nights.

Table 4.3: Frequency of Months of Data in All Observations

monthsdata	Freq.	Percent	Cum.
12	916	4.59	4.59
13	1156	5.79	10.38
14	1139	5.70	16.08
15	1149	5.75	21.83
16	1121	5.61	27.45
17	986	4.94	32.38
18	1068	5.35	37.73
19	1376	6.89	44.62
20	1174	5.88	50.50
21	2066	10.35	60.85
22	930	4.66	65.50
23	998	5.00	70.50
24	5891	29.50	100.00

If we look at the frequency of how many months appear in each observation, we can see that it is very well spread out outside of properties that were on the market for 21 months with nearly 30 percent of properties in the dataset having a full 24 months of data available.

Figure 4.2: Revenues of Airbnbs vs. Rental Properties, averaged by Zip Code and assuming a 50% review rate



The disparity between yearly revenues between Airbnbs and rental properties is easier to visualize. If we compare the average revenues of all the properties in the dataset by zip code, only 6 of the 76 zip codes in the dataset show higher averages in the Airbnb data then the rental data. On top of that, no zip code has an average revenue above 60k in the Airbnb dataset, while there are a couple dozen zip codes that break above that range in the rental data.

CHAPTER FIVE: MODEL FRAMEWORK

From Sirmans et al, the strengths of a hedonic pricing model within the housing market are greatly detailed, and the same model is even used within some of the literature above, most notably in Barron. Barron uses a semilog model with the dependent variable logged and this model is also mentioned in Sirmans' work. Other models that work within hedonic pricing include linear and log-log models. Due to the nature of this study and data, a linear and semilog model are most applicable for this study. The log-log model was considered, but dropped due to the specifications of the data and the way log-log models are interpreted. In a log-log model, it looks to measure the percentage change of the dependent variable from a one percentage shift of an independent variable. Since this study is not looking at the square footage of the property's bedrooms and bathrooms but rather the amount of bedrooms and bathrooms located in the property, using percentage increases in relatively small units of measurements may poorly specify the model. Both linear and semilog are being used to account for the possibility that the rent disparity between Airbnb being rented long term and actual long term rentals are not large. If the disparity is large, the semilog model will more accurately predict impacts on the rent.

The yearly revenue generated through a property will be regressed on the number of bedrooms, the number of bathrooms, a dummy variable representing if the property was rented through Airbnb or not, a dummy variable representing which year of the two years of data the house is a part of, two dummy variables representing if the property was an apartment, townhouse, or single family residence, and a variable capturing the zip codes of the properties.

The model is:

$$YR_{ijt} = \alpha + \beta_1 BED_{ijt} + \beta_2 BATH_{ijt} + \beta_3 \, AIRBNB_{ijt} + \beta_4 \, Y2_t + \beta_5 \, APT_i + \beta_6 \, TWNHS_t + \beta_7 ZIP_j + \varepsilon_{ijt}$$

This model represents the linear model with yearly revenue as our dependent variable with all right-hand variables staying in linear form. YR represents the yearly revenue of house i in zip code j during time period t. BED and BATH both stand for the number of bedrooms and bathrooms found in house i and zip code j during time period t. AIRBNB represents the dummy variable indicating whether the property is an Airbnb or not and is the main variable we want to focus on in the results. To capture any yearly time trend, Y2 indicates whether the house was rented out during the time period of March 2019 – February 2020, indicated by a 1, or during the time period of March 2018 – February 2019, indicated by a zero. Finally, ZIP captures the zip codes present in the dataset and works as a spatial control to compare earnings from short-term and long-term properties with similar locational amenities. The semi-log model simply substitutes the natural log of estimated yearly earnings for the left-hand side. In this model, the percentage change of the dependent variable is measured for each unit of change in the right-hand variables.

CHATPER SIX: RESULTS

From the above explanation, four OLS regressions are run: two models using a 50% review rate as a linear and semilog model and two models with a 30.5% review rate under the same model specification. All models include fixed-effects for the 76 zip codes in the data.

Table 6.1: OLS Regression of Estimated Annual Earnings of Short-term and Long-term Rentals in Los Angeles (2018-2020)

	(1) annualrev	(2) lnannualrev	(3) annualrev_high	(4) lnannualrev_high
bathrooms	14880.7***	0.168***	15963.7***	0.163***
	(52.42)	(16.49)	(49.15)	(16.28)
bedrooms	3391.0***	0.220***	5062.5***	0.232***
	(13.71)	(24.68)	(17.89)	(26.43)
y2	-3067.5***	-0.0511***	-2743.8***	0.0666***
	(-8.70)	(-3.96)	(-6.80)	(5.24)
airbnb	-33368.7***	-0.877***	-23219.4***	-0.639***
	(-58.24)	(-42.90)	(-35.42)	(-31.72)
apt	-684.1	-0.268***	-2606.6***	-0.265***
	(-1.28)	(-13.66)	(-4.25)	(-13.73)
twnhs	-13214.4***	-0.0880**	-13490.3***	-0.0838*
	(-14.15)	(-2.66)	(-12.63)	(-2.57)
Zip Code FE	Included	Included	Included	Included
R^2	0.5286	0.4026	0.4742	0.3714
Adjusted R2	0.5270	0.4000	0.4725	0.3691
Constant	5156.4	9.414***	-1112.2	9.381***
	(0.50)	(26.07)	(-0.09)	(26.38)
N	24378	22169	24378	22169

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the 50% review rate linear model, everything but the constant is significant at the 0.01 level, with the constant being not statistically significant. The adjusted R-squared indicates this model captures about 52.70% of the variation in the data. As expected, properties with more bedrooms and bathrooms rent for higher prices, and townhouses and apartments rent for less than single-family homes (the omitted category). Most importantly, a property being rented through Airbnb generates about $33,368 less, or about $2780 per month, than a property being rented through a traditional long-term rental, all else held equal. In the semilog model, the results show that a property being rented through Airbnb makes about 88% less than a property being rented through a traditional long-term rental. The adjusted R-squared also shows a dip in its predictive power of about 12%. Both of these numbers seem high, but given the work done by the Budget and Legislative Office in 2015 where in certain zip codes the loss in revenue was about $1,700, these results are

Within the linear model for the 30.5% review rate, all variables besides the constant are significant at the 0.01 level, while the constant is not statistically significant. The adjusted R-squared indicated this model captures about 47.25% of the variation in the data. Most importantly, a property being rented through Airbnb generates about $23,219 less, or about $1,935 per month, than a property being rented through a traditional long-term rental, all else held equal. In the semilog model, the results show that a property being rented through Airbnb makes about 64% less than a property being rented through a traditional long-term rental. The adjusted R-squared shows a similar dip as the 50% review rate model, dropping by 10%. With this estimate of the model, the linear model shows number that are believable. In the study done by the Budget and Legislative Analyst's Office, they calculated monthly revenue for zip codes for casual hosts, defined as renting out a property for less than 5 days per month, which is held constant in our study to be 4 when calculating occupancy rates. What they find is that none of these hosts end up making any monthly profit than if they rented as a long term rental instead, reaching numbers in some zip codes of nearly $1,700 less per month. Given the time since this study, it could be feasible that in Los Angeles, when factoring in inflation, the number of $1,934.92 of tradeoff by renting your property through Airbnb instead of a long-term rental could hold.

Two factors of note within these regression results is the sign on y2 and the number of observations present in the linear and semi-log models. The sign on y2 indicates that during the second year of data, rental revenues decreased, which goes against the trend of rental prices in the current market. I believe that this is caused by the observation differences that occur within the linear and semi-log model. These differences mean that when the natural log of yearly revenue were taken, about 2,200 of those revenues were zero and therefore unable to have a natural logarithm. This indicates a particularly interesting trend within the short-term rental data. Properties were offered on the market for at least 12 months, but were never utilized once. This could be caused by a multitude of results. One possibility could be that the owner forgot to take their property off of Airbnb, so when someone did attempt to use the property, the owner informed the individual that the property was not available, but did not remove their listing. Another possibility was that in order to work around the commission when a property is booked through Airbnb, the owner offered a discount to the renter if they did not go through Airbnb and keep their property on Airbnb as an advertisement of their property. Finally, the owner may have just had poor luck in renting their property out during a 12 month period of time.

Table 6.2: OLS Regression of Estimated Annual Earnings of Short-term and Long-term Rentals in Los Angeles (2018-2020) (Robustness Check)

	(1) annualrev	(2) lnannualrev	(3) annualrev_high	(4) lnannualrev_high
bathrooms	15833.5***	0.168***	17012.4***	0.163***
	(54.32)	(16.49)	(51.46)	(16.28)
bedrooms	3361.2***	0.220***	5059.1***	0.232***
	(13.16)	(24.68)	(17.46)	(26.43)
y2	1354.0***	-0.0511***	2874.3***	0.0666***
	(3.65)	(-3.96)	(6.84)	(5.24)
airbnb	-28435.6***	-0.877***	-17027.9***	-0.639***
	(-48.46)	(-42.90)	(-25.58)	(-31.72)
apt	-2257.1	-0.268***	-4730.3***	-0.265***
	(-4.01)	(-13.66)	(-7.41)	(-13.73)
twnhs	-14494.2***	-0.0880**	-14975.4***	-0.0838*
	(-15.26)	(-2.66)	(-13.90)	(-2.57)
Zip Code FE	Included	Included	Included	Included
R^2	0.5456	0.4026	0.4984	0.3714
Adjusted R2	0.5439	0.4000	0.4966	0.3691
Constant	691.0	9.414***	-6630.0	9.381***
	(0.07)	(26.07)	(-0.56)	(26.38)
N	22169	22169	222169	22169

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

When the observations that showed zero yearly revenue were more closely examined, it turned out that over 2,200 of the entries were focused on the second year of data, with only a handful of these entries being in the first year of data. As a robustness check, these entries were dropped and the linear models were run again. The sign on y2 is now positive, indicating that the amount of zero revenue entries was forcing the coefficient on y2 down significantly enough to create a negative coefficient. With these new results, Model (3) estimates a monthly loss of $1,412, which given the 2015 San Francisco study, is much more feasible and would be the model of choice.

CHAPTER SEVEN: CONCLUSIONS

The results say that in both the two different specifications and occupancy rates that treating Airbnb as a method for long-term rentals does lead to less revenue than renting through a traditional long-term rental. In terms of data, the data available on insideairbnb, while the best available, still relies on several assumptions to calculate monthly and yearly revenue. Some listings did not have any values for monthly price or even daily or weekly prices and if one data point needed was never scraped over a whole year, this methodology dropped the value. If we were to redo this study, personally scraping Airbnb listing with our own tool or hand gathering the data would lead to less data being dropped and a hopefully more accurate result, albeit that method would take a significant amount of time. Another issue of the Airbnb dataset was a lack of more information on each property. I would have liked to add data such as the lot size of the property to get an idea of if the property had a yard or other extra amenities and square footage of the property itself to determine the actual size of each property. Nevertheless, these characteristics are likely correlated with information on bedrooms and bathrooms, which are observed. Along with that, the rental data had information for certain properties if they were attached to another property by a wall or not, which could also be an interesting avenue to explore given that this dataset had both apartments and townhouses.

In terms of costs such as maintenance and management, I would predict that these factors would create a smaller gap between long term rentals and Airbnbs, but not by a significant enough margin to change the overall results. Due to there being no monetary barrier to entry for Airbnb, there would be no real cost of management to the owner of an Airbnb unless they choose to put up multiple listings through a company. Along with this, there is no risk of renters defaulting through Airbnb due to the upfront payment needed to reserve an Airbnb. Despite this, it is still unlikely that Airbnb owners would be able to close the gap enough to break even or profit relative to long term rentals, even using the less conservative estimate of about $2,000 per month.

One point to note was the study done by Coles et al (2017) that not only used data directly from Airbnb, but also had employees from Airbnb working on that study. That study indicated that in New York, Airbnbs were incredibly unprofitable relative to long-term apartment and would need to be rented out for over quadruple the amount of days the median was in 2017 for an Airbnb to break even with a long-term rental. There are only two logical explanations behind these results: either the owners of these Airbnbs are not price maximizing or they are renting out multiple Airbnbs, therefore matching the revenue earned by one long-term rental, which has its own problems when examined. A study that focuses on Airbnb owners that rent two or more properties may show different results than the ones above and may be worth examining. For example, how many Airbnbs would you need to rent to break even with a long-term rental?

There are many different paths this research could go down, with the most notable expansion being determining when Airbnb owners begin to drop out of the STR market. For example, I would want to see if there any trend in how owners drop out of the STR market, such as if their predicted loss is amongst the highest of all owners within the past time period. If this trend was seen throughout the data, it would indicate that these owners are joining Airbnb in order to make a profit, stay in for a time period, and then realize that they aren't making as much as they would like to or could make more profit by renting their property out long-term and leave the market. Another direction this could go is determining a break-even point on how many nights would be needed to be booked for an Airbnb to match the revenue generated by a long-term rental. This would create a baseline that could be further specified to control for specific amenities of a property to determine whether it really would be worth it to put your property on Airbnb for an entire year.

REFERENCES

Badger, E., & Kang, I. (2018, April 10). California Today: How Do California Eviction Rates Compare to the Rest of the U.S.? Retrieved July 21, 2020, from https://www.nytimes.com/2018/04/10/us/california-today-housing-eviction.html

Barron, K. , Kung, E. , Proserpio, D. , 2017. The Sharing Economy and Housing Affordability: Evidence from Airbnb Working Paper

California Regional Multiple Listing Service. Retrieved 2020 from https://go.crmls.org/

City and County of Sacramento Board of Supervisors. *Policy Analyst Report*. Retrieved March 28, 2020 from https://sfbos.org/sites/default/files/FileCenter/Documents/52601-BLA.ShortTermRentals.051315.pdf

Coles, P. , Edesdal, M. , Ellen, I.G. , Li, X. , Sundararajan, A. , 2017. Airbnb usage across New York City neighborhoods: geographic patterns and regulatory implications. Cambridge Handbook of the Law of the Sharing Economy Working Paper

Cox, Murray. *Insideairbnb*. Retrieved from http://insideairbnb.com/

Horn, Keren & Merante, Mark, 2017. "Is home sharing driving up rents? Evidence from Airbnb in Boston," Journal of Housing Economics, Elsevier, vol. 38(C), pages 14-24.

Lee, D. , 2016. How Airbnb Short-Term Rentals Exacerbate Los Angeles's Affordable Housing Crisis: Analysis and Policy Recommendations. Harvard Law & Policy Re- view, pp. 229–254

Los Angeles Housing And Community Investment Department. Retrieved May 1, 2020 from https://hcidladev.lacity.org/

Madans, H. (2019, April 26). Short-Term Rentals Fill a Niche. Retrieved May 24, 2020 from https://labusinessjournal.com/news/2019/apr/26/short-term-rentals-fill-niche/

San Francisco Planning Department. *Executive Summary*. Retrieved May 28, 2020 from https://commissions.sfplanning.org/cpcpackets/2014-001033PCA.pdf

Sirmans, G., Macpherson, D., & Zietz, E. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature, 13*(1), 3-43. Retrieved March 28, 2020, from www.jstor.org/stable/44103506