Market Shifts in the Sharing Economy: The Impact of Airbnb on Housing Rentals

Hui Li, Yijin Kim, Kannan Srinivasan*

August 25, 2020

Abstract

This paper examines the impact of Airbnb on the local rental housing market. Airbnb provides landlords an alternative opportunity to rent to short-term tourists, potentially causing some landlords to switch from long-term rentals, thereby affecting rental housing supply and affordability. Despite recent government regulations to address this concern, it remains unclear whether and what types of properties are switching. Combining Airbnb and American Housing Survey data, we estimate a structural model of property owners' decisions and conduct counterfactual analyses to evaluate various regulations. We find that Airbnb mildly cannibalizes long-term rental supply. Cities where Airbnb is more popular experience a larger reduction in rental supply; however, these cities do not necessarily have a larger percentage of switchers. Interestingly, we find that affordable units are the major sources of both the negative and positive impacts of Airbnb, as they see a larger rental supply reduction and a larger market expansion effect. Although Airbnb harms local renters by reducing affordable rental supply, it also serves as a valuable income source for local hosts with affordable units. Policy makers need to trade off between local renters' affordable housing concerns and local economically disadvantaged hosts' income source needs. The counterfactual results suggest that imposing a linear tax is more desirable than limiting the number of days a property can be listed. We propose a new convex tax and show that it outperforms existing policies in terms of reducing cannibalization and alleviating social inequality. Finally, Airbnb and rent control can exacerbate each other's negative impacts.

*Hui Li, Carnegie Bosch Assistant Professor of Marketing, Tepper School of Business, Carnegie Mellon University. Email: huil1@andrew.cmu.edu. Yijin Kim, Ph.D. Student in Marketing, Tepper School of Business, Carnegie Mellon University. Email: yijink@andrew.cmu.edu. Kannan Srinivasan, H.J. Heinz II Professor of Management, Marketing and Information Systems, Tepper School of Business, Carnegie Mellon University. Email: kannans@andrew.cmu.edu.

1 Introduction

Sharing economy platforms have affected marketing mix decisions (e.g., product, pricing, and distribution channels) by providing an additional channel for individuals to market their products and services. For instance, peer-to-peer marketplaces for short-term accommodations such as Airbnb, HomeAway, and VRBO have emerged as an alternative channel for landlords to market their properties to short-term tourists in addition to the traditional long-term rental market for local residents. These home-sharing platforms have grown at an exponential rate in recent years. Airbnb, the most popular platform, had over six million listings around the world as of March 2019—more listings than the hotel rooms from the six largest hotel groups combined.[1] Given the opportunity to rent to short-term tourists, some property owners may switch from the traditional channel of long-term rental to the new channel of Airbnb, as the yields can be two to three times higher on Airbnb than on the long-term rental market.[2] Such switching behavior could impact rental housing supply and affordability.

A staggering rise in short-term rental platforms has prompted questions concerning whether they have contributed to rental housing shortages and the affordable housing crisis and whether they should be regulated to protect affordable housing. For example, the City of Los Angeles approved new rules for Airbnb-type rentals in December 2018, after more than 3.5 years of debate since the law was first proposed.[3] Similarly, in San Francisco, there has been a controversial debate and changes in the scope of the city's short-term rental regulation, which first went into effect in February 2015.[4] By 2019, many cities had imposed limits on the number of days that a property can be listed on short-term rental platforms (e.g., a maximum of 90 days in San Francisco and 180 days in Los Angeles). However, most of these policies were launched without empirical evidence on the actual impact of Airbnb on the rental housing market.

In this paper, we seek to answer two questions. First, how does Airbnb affect the supply and affordability of rental housing? In particular, we examine how many units are taken off the rental market (i.e., the impact on rental supply) and what types of properties are taken off (i.e., the impact on rental affordability). Second, what is the impact of various regulations on short-term rentals? Answering these questions requires an understanding of the underlying trade-offs, or benefits and costs, for property owners. The benefits of renting can be directly observed from the prices and occupancy rates in the long-term market and on Airbnb.

[1] See https://press.airbnb.com/airbnb-hosts-share-more-than-six-million-listings-around-the-world/

[2] See https://tranio.com/articles/airbnb_a_game-changer_for_the_commercial_property_market_4982/.

[3] See https://www.latimes.com/local/lanow/la-me-ln-airbnb-rental-ordinance-20181211-story.html

[4] In a ballot measure in San Francisco in 2015, 55% of voters rejected Proposition F, which would have reduced the number of days that owners can rent out their properties from 90 to 75. See https://www.theguardian.com/us-news/2015/nov/04/san-francisco-voters-reject-proposition-f-restrict-airbnb-rentals. Later, in 2016, San Francisco approved a new rule that requires short-term rental websites such as Airbnb to display each host's registration number next to their listings or email the information to the city's short-term rentals office. This rule supplements San Francisco's existing short-term rental regulations that require hosts to register with the city's short-term rentals office. See http://fortune.com/2016/06/07/sf-airbnb-new-rules/.

However, the costs of renting and how they differ by demographics, properties, and cities are unknown.

We estimate a structural model of property owners' decisions using Airbnb listings data and American Housing Survey data. We recover the underlying heterogeneous hosting costs, which allow us to simulate market outcomes in the absence of Airbnb to examine its impact and evaluate market outcomes under different policies. In the model, property owners first make a discrete choice based on their availability type. Owners who are available for the full year choose among Airbnb, long-term rental, and an outside option of keeping the properties vacant. Owners who are available for part of the year choose between Airbnb and the outside option. This decision is usually made yearly, as the length of rental leases is typically one year. Second, if owners choose Airbnb, they decide the number of days to list their properties on Airbnb, which can be a monthly decision. The two decisions are linked in that the ex ante expected profit from the second decision affects the first decision. The hosting costs and host availability are allowed to be heterogeneous by property characteristics, host demographics, various metro area characteristics (e.g., population, density, mortgage affordability, wage and employment in the accommodation industry, how long Airbnb has been present, and how favorable city regulations are to short-term rentals) and over time.

The results show that Airbnb mildly cannibalizes the long-term rental supply but creates a market expansion effect. The level of cannibalization varies significantly across metro areas. Interestingly, we find that although the reduction in the rental supply is larger in metro areas where Airbnb is popular, the percentage of switchers is not necessarily larger in these areas. For example, Miami and New York are among the cities with the highest Airbnb popularity and the largest rental supply reduction. However, their percentages of switchers are among the lowest, suggesting that most of the Airbnb listings in Miami and New York are from market expansion rather than cannibalizing the rental supply. Policymakers must take a holistic view when evaluating the impact of Airbnb.

Importantly, the results show that affordable units are the major sources of both the negative cannibalization impact and the positive market expansion impact of Airbnb. We find suggestive evidence that Airbnb does raise affordable housing concerns, as the rental supply reduction is larger among affordable units. However, the market expansion effect is also larger for affordable units, as the fraction of nonswitchers is larger among affordable units on Airbnb. Although Airbnb harms local renters by reducing affordable rental supply, it also serves as a valuable income source and benefits local hosts who own affordable units and are more likely to be economically disadvantaged. Therefore, policy makers need to trade off between local renters' affordable housing concerns and local hosts' income source needs.

We highlight the usefulness of the structural model in identifying the actual potential switchers. It is not appealing to take the observed data and assume, without modeling the hosts' decisions, that an observed "full-time" ("part-time") Airbnb listing always implies cannibalization (market expansion). First, if hosts list all year on Airbnb, this does not necessarily mean that they are switchers from the long-term rental market. They could have chosen to keep their properties vacant without Airbnb if their costs (revenues) of long-term rental are high (low). Second, if hosts list part of the year on Airbnb, it does not necessarily mean that the properties are not available for the rest of the year and are not switchers from long-term rentals. Hosts may choose to list for shorter periods if the Airbnb profit is large enough to allow them to list part time and still earn more than listing in the long-term rental market. Overall, one must systematically model the revenue-cost trade-offs of the hosts to identify switchers.

In the counterfactual analysis, we evaluate two sets of policies related to the supply and affordability of rental housing. The first set of counterfactuals are motivated by recent regulations on short-term rental. Policymakers are continuously searching for effective policies to prevent switching away from long-term rentals, especially in cities with tight housing markets such as San Francisco, New York, and Los Angeles. In addition to limiting the length of listings on Airbnb, local municipalities also require hosts to collect certain taxes from guests, which is similar to a hotel occupancy tax. We examine these two existing policies (day limit and a linear tax) and further propose a new convex tax that imposes a higher tax on expensive units and a lower tax on affordable units, which is motivated by our finding that the cannibalization rate or percentage of switchers is larger for expensive units.

A desirable policy should maintain the positive impact of Airbnb (nonswitchers or market expansion) and reduce the negative impact of Airbnb (switchers or cannibalization). Therefore, we assess the desirability of the three policies along three dimensions: (1) the ability to reduce the cannibalization rate or percentage of switchers; (2) the ability to reduce the number of luxury units among nonswitchers; and (3) the ability to reduce the number of economically advantaged hosts (e.g., high-income, older, or high-education hosts) among nonswitchers. The second and third measures relate to social inequality because they capture potential differential policy impacts on heterogeneous hosts. In particular, imposing regulations can induce a redistributive effect among Airbnb hosts and determine who can continue to benefit from Airbnb. A desirable policy should prevent economically advantaged hosts, who already have abundant resources, from using Airbnb as an additional income source, which can exacerbate social inequality. We find that the proposed convex tax outperforms the other two policies along all three dimensions. The linear tax is the second-best policy, and the day limit is the worst.

The second set of counterfactuals focus on rent control policy, which limits the rent in the long-term rental market. Economists are virtually unanimous in concluding that rent controls are destructive because they reduce the supply of available housing. When a rent control policy is imposed, property owners choose not to rent out their units for long-term rental. Despite the known adverse impacts, the states of California, Maryland, New Jersey, New York, Oregon, and the city of Washington D.C. still have some rent control or stabilization policies on the books (as of March 2019).[5] We show that this negative effect of rent control policy is aggravated when Airbnb is available, as Airbnb serves as an additional profitable option for property owners and can motivate them to further switch away from the long-term rental market.

The results have strong policy implications for short-term rentals and affordable housing. Airbnb has been debated and regulated in the cities it has entered. It is difficult for policymakers to assess the impact of Airbnb because it requires knowing whether the properties would have been in the rental market had Airbnb not been available. Our model can be used to assess the impact of Airbnb on rental supply and affordability. The results provide a detailed profile of potential switching hosts and properties, which can serve as a foundation for policy making. We also provide a thorough evaluation of the desirability of various short-term rental regulations and propose a new policy that can outperform existing policies. Finally, we show that rent regulation must be implemented with extra caution when Airbnb is available, as lower profits from long-term rentals can cause landlords to switch to Airbnb.

2 Literature Review

This paper contributes to the recent literature on the sharing economy (see Einav, Farronato, and Levin 2016 for a review of the sharing economy). In particular, our work relates to the literature on (1) the impact of Airbnb on the housing and rental market, (2) the impact of the sharing economy on traditional industries, (3) supply decisions in the sharing economy, and (4) how the sharing economy affects marketing mix decisions.

There have been papers in the fields of Economics and Marketing that study Airbnb's impact on the housing market. Lee (2016) and Gurran and Phibbs (2017) provide descriptive analyses of Airbnb and the rental housing market in Los Angeles and Sydney, Australia, respectively. Most of other studies focus on how home-sharing affects housing prices and rents in a particular city. Horn and Merante (2017) find that a one-standard-deviation increase in Airbnb listings is associated with an increase in asking rents of 0.4% in Boston. Sheppard and Udell (2018) find that doubling the total number of Airbnb listings within 300 meters of a house is associated with an increase in house prices of 6% to 9% in New York City. Koster et al. (2019) find that the adoption of home sharing ordinances reduced housing prices by 3% and rents by 3% in Los Angeles. Marketing researchers have also recently contributed to this topic. Barron, Kung, and Proserpio (2019) use a comprehensive dataset covering the U.S. and find that a 1% increase in Airbnb listings leads to a 0.018% increase in rents and a 0.026% increase in house prices.

Our research contributes to the literature and can be differentiated from the previous studies in important ways. First, while most studies focus on the impact of Airbnb on housing and rental prices, we focus on

[5] See https://www.curbed.com/2019/3/8/18245307/rent-control-oregon-housing-crisis

hosts' supply choices and assess how many and what types of properties are taken off the rental market, which are the causes of the change in housing and rental prices. To the best of our knowledge, we are the first to systematically and formally model hosts' decisions and to recover the underlying cost of hosting. Second, other studies mostly focus on a particular city. We leverage data on a wide variety of cities and show that there is significant heterogeneity across cities, which is helpful for localized policy making. Third, while other studies mostly provide descriptive analysis or conduct regression analysis, we use structural models to explicitly analyze the underlying trade-offs faced by individual hosts. The framework allows us to conduct counterfactual analysis. The results have strong policy implications regarding short-term rental regulations and rent control.

More broadly, our paper contributes to the literature on how the sharing economy affects traditional industries and incumbent firms. For instance, ride-sharing services are found to affect the earnings of taxi drivers (Berger, Chen, and Frey 2018), automobile ownership (Gong, Greenwood, and Song 2017), alcohol-related motor vehicle fatalities (Greenwood and Wattal 2017), and local entrepreneurial activity (Burtch, Carnahan, and Greenwood 2018). On the subject of Airbnb, in a pioneering work, Zervas, Proserpio and Byers (2017) study the impact of Airbnb's entry on hotels in Texas and find that Airbnb mildly cannibalizes hotels, with lower price hotels being the most affected. Li and Srinivasan (2019) study how Airbnb's flexible supply changes the way in which the industry accommodates seasonal demand and how incumbent hotels with fixed capacity should respond.

Our work also relates to the stream of literature on supply choices in the sharing economy. Zhang, Mehta, Singh, and Srinivasan (2018) model Airbnb hosts' decisions regarding whether to operate or block listings along with listing quality decisions (e.g., image quality in the listing description and host service effort). Li, Moreno and Zhang (2016) study the pricing decisions of Airbnb hosts and find that a substantial number of Airbnb hosts are unable to optimally set prices. We contribute to the literature by studying property owners' decisions regarding whether and how long to list on Airbnb.

Finally, our paper contributes to the literature on how the sharing economy affects marketing mix decisions (e.g., product choice, pricing, and distribution channels). Jiang and Tian (2018) study sharing-economy-enabled collaborative consumption and find that when a firm strategically chooses its retail price, consumers' sharing of products with high marginal costs is a win-win situation for both firm and consumers. Tian and Jiang (2017) study how consumer-to-consumer product sharing affects the distribution channel and find that the sharing market tends to increase the retailers' share of the gross profit margin in the channel. Dowling et al. (2019) study two common pricing strategies in car sharing services, pay-per-use and flat-rate pricing. They find a prevalent and time-persistent pay-per-use bias because of an underestimation of usage, a preference for flexibility, and the influence of physical context (e.g., weather). They suggest that the pay-per-use bias may be the prevalent tariff choice bias in the sharing economy.

3 Data

3.1 Data Description

The two main data sets used in this study are the 2015 and 2017 American Housing Survey (AHS) and Airbnb listings data for 9 representative metropolitan areas.[6] First, the AHS is the most comprehensive longitudinal national housing survey in the U.S. that gathers detailed property-level data on properties in metropolitan areas. Each observation includes a housing unit, its property characteristics (e.g., number of bedrooms and bathrooms, amenities, property type); occupant demographics (e.g., age, education, income, gender, marital status); tenure information (whether the unit is owner-occupied, renter-occupied, or vacant); and, if applicable, rent. As the survey is conducted biennially, we utilize the most recent two years' data at the time of this study. These two years also have a stronger Airbnb presence than the previous years as Airbnb continues to grow over time. We focus on the properties that are rented or vacant because they are potentially available for listing on either the long-term rental market or Airbnb and are thus relevant for our study.

The second data set contains information on every Airbnb property listed on Airbnb in 2015 and 2017 collected by AirDNA, a third-party company that specializes in data collection and analysis. Each property record contains monthly performance information such as the number of days available for booking, average daily rate, and occupancy rate. It also includes over 20 property characteristics such as location (zip code); property type (e.g., house, apartment); listing type (entire place or private/shared room); number of bedrooms and bathrooms; and amenities such as a kitchen, air conditioning, heating, washer, dryer, fireplace, and parking space. We also collect data on when a property is first listed on Airbnb to distinguish between no listing and a new listing.[7] In the 9 representative metro areas, there were 169,338 properties listed on Airbnb in 2015 and 252,459 properties in 2017.

Combining the two data sets provides a comprehensive data set on every property that is potentially available for listing in the selected area. A property is listed either on Airbnb (units in the Airbnb data set) or on the long-term rental market (rented units in the AHS data set) or on neither (vacant units in the

[6]The U.S. Office of Management and Budget (OMB) refers to a metropolitan area as a core based statistical area (CBSA), which corresponds to an urbanized core area containing a substantial population and its adjacent communities having a high degree of economic and social integration with that core. For convenience, we denote a metro area by its principal city in the CBSA (e.g., New York for New York-Newark-Jersey City, NY-NJ-PA).

[7]For instance, if a property was first listed in February 2015, we exclude January 2015 when estimating the host's second-stage decision of how many days to list in a month because zero days listed in January 2015 is due to not have yet having joined Airbnb instead of choosing not to list.

AHS data set). Hereafter, we refer to keeping the property vacant as the outside option.[8] We distinguish between two types of units that choose the outside option, "vacant full year" and "vacant part year", which correspond to units that are kept vacant for the full year and units that are kept vacant for part of the year due to occasional self-use in the AHS data set.

We focus on three sets of covariates in the empirical analysis: property characteristics, demographics, and market characteristics. The property characteristics are available at the property level in both the AHS data and the Airbnb data. Demographic information is available for each property in the AHS but not for the Airbnb listings. We collect zip-code-level demographics from the American Community Survey (ACS) and impute the host demographics for the Airbnb properties using the local zip-code-level demographics. The metro area characteristics include the metro-area-level population, density, employment and wage in the accommodation industry from the ACS data and mortgage affordability information from the Zillow Mortgage Affordability Index. We also collect data on an additional set of metro-area-level variables that serve as covariates in the estimation of hosts' choices and instruments in the hedonic regressions of revenues. These variables include rent-to-own ratio, unemployment rate, number of air passengers to the city, Airbnb regulation score, and Airbnb history. Rent-to-own ratio and unemployment rate are collected from the ACS. The number of air passengers to the city is from the T-100 Market (All Carriers) database published by the Bureau of Transportation Statistics.[9] Airbnb regulation score, which measures how friendly city policies are to short-term rentals, is published by the R Street Institute.[10] Lastly, Airbnb history, measured by the time since Airbnb reached 10% of the total rooms supplied by hotel and Airbnb in a city, is computed using our Airbnb data set and the hotel data from tourism-related reports and articles.[11]

3.2 Data Patterns

In this subsection, we describe the observed data patterns that motivate our empirical model specifications. In particular, we present the percentage of Airbnb, long-term rental, and outside option units, which relates to the first-stage decision of whether to list, and the listing patterns for the Airbnb units, which relate to the second-stage decision of how many days to list.

Table 1 presents the percentage of Airbnb, long-term rental, and outside option (vacant full year and vacant part year) units by year and the summary statistics of properties choosing each option. In 2015,

[8] Note that the properties in the Airbnb data set can overlap with the properties in the AHS data set. The AHS data set classifies housing units that are unoccupied, or occupied by anyone who is not the usual resident (such as an Airbnb guest), as vacant. However, the observed characteristics in the two data sets do not allow us to distinguish whether an AHS property is listed on Airbnb. Given that the number of Airbnb listings as a fraction of the total number of AHS properties is very small (1.75% in 2015 and 2.63% in 2017), we assume that there are no overlaps and combine the two data sets without removing overlapped properties.

[9] See https://www.transtats.bts.gov/DatabaseInfo.asp?DB_ID=111.

[10] See https://www.rstreet.org/wp-content/uploads/2016/03/RSTREET55.pdf.

[11] See, for example, https://washington.org/dc-information/washington-dc-facts.

Table 1: Summary Statistics by Airbnb, Long-Term Rental, and Outside Option

	Airbnb	Long-term Rental	Vacant Full Year	Vacant Part Year
2015: Number of Units	169,338	8,481,448	717,148	298,965
2015: Proportion (%)	1.75	87.74	7.42	3.09
2017: Number of Units	252,459	8,409,056	655,796	275,502
2017: Proportion (%)	2.63	87.66	6.84	2.87
Number of Bedrooms	1.38 (0.96)	1.97 (0.96)	2.38 (1.06)	2.26 (1.04)
Number of Bathrooms	1.31 (0.70)	1.68 (1.17)	2.46 (1.70)	2.82 (2.12)
Apartment (%)	71.88 (44.96)	72.15 (44.83)	43.42 (49.58)	45.87 (49.87)
Kitchen (%)	90.60 (29.18)	99.14 (9.22)	98.66 (11.48)	96.81 (17.60)
Air Conditioning (%)	80.28 (39.79)	87.74 (32.79)	70.75 (45.51)	84.18 (36.52)
Heating (%)	86.97 (33.67)	99.37 (7.93)	98.01 (13.99)	95.90 (19.85)
Washer (%)	62.36 (48.45)	46.50 (49.88)	48.68 (50.00)	60.71 (48.88)
Dryer (%)	60.44 (48.90)	43.83 (49.62)	47.97 (49.98)	59.69 (49.09)
Fireplace (%)	12.01 (32.50)	11.20 (31.54)	19.54 (39.66)	14.18 (34.91)
Parking Space (%)	41.79 (49.32)	32.14 (46.70)	46.22 (49.87)	48.03 (50.00)
Private or Shared Room (%)	43.89 (49.63)			
Airbnb Daily Price ($)	199.79 (1,566.05)			
Airbnb Occupancy Rate (%)	25.41 (35.93)			
Monthly Rent ($)		1,263.74 (593.36)		

Note: Standard deviations are shown in parentheses.

1.75% of the properties chose Airbnb, 87.74% chose long-term rental, 7.42% were vacant for the full year, and 3.09% were vacant for part of the year. The numbers changed to 2.63%, 87.66%, 6.84%, and 2.87%, respectively, in 2017 as the number of Airbnb properties increased by nearly 50% from 2015 to 2017. We find that the Airbnb units are comparable in property characteristics to the long-term rental units. For example, both have smaller numbers of bedrooms and a larger proportion of apartment units than the outside option properties. This suggests that properties on Airbnb and in the long-term rental market could come from the same pool. We also find that Airbnb could generate more rental income per month than a long-term rental, as implied by Airbnb's average daily price ($199.79) and occupancy rate (25.4%) and the average monthly rent for long-term rentals ($1,263.74). The payoff difference could motivate hosts to switch from long-term rentals to Airbnb.

We further examine the listing patterns of properties on Airbnb. Property owners choose the dates when the property is available for booking (i.e., listed) or blocked from accepting reservations (i.e., not listed). We find that the listing pattern is heterogeneous across hosts and also across months within a host. Figure 1 plots the monthly number of days listed for two representative Airbnb properties. We find that hosts often choose not to list at all for a particular month. If a property is listed, it is more likely to be listed for the full month than for some part of the month. This is also supported by the histogram of the number of days

Figure 1: Representative Airbnb Listing Patterns

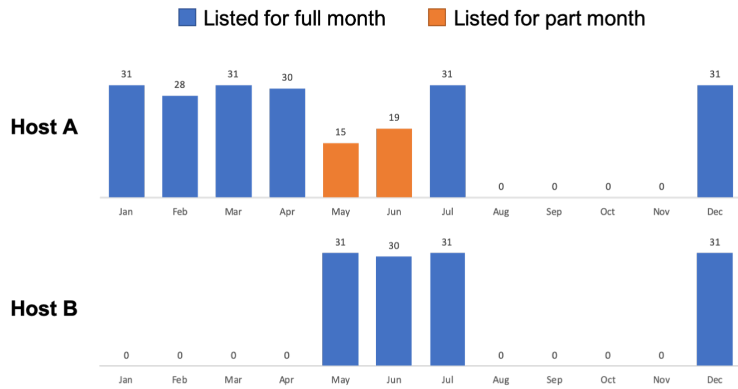
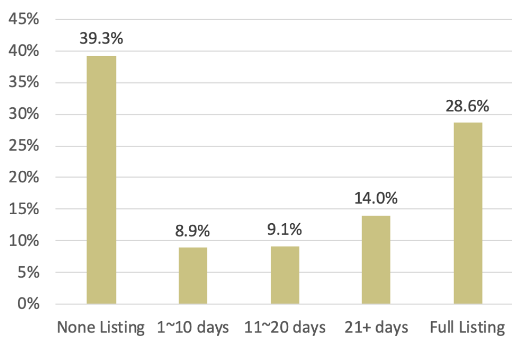


Figure 2: Histogram of Monthly Number of Days Listed

(a) 2015



(b) 2017

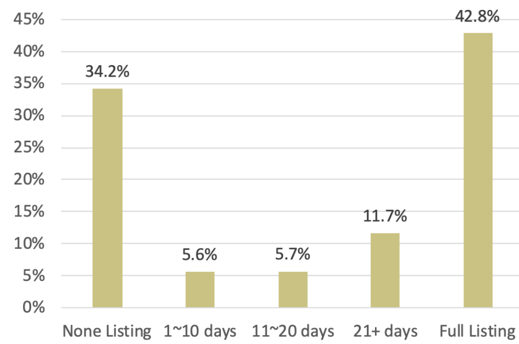
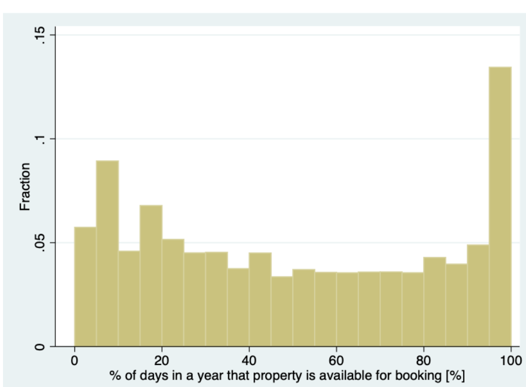
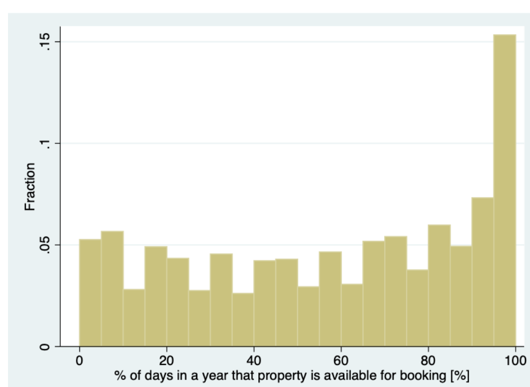


Figure 3: Histogram of Percentage of Days Listed in Each Year

(a) 2015



(b) 2017



listed in a month (Figure 2). The histograms for both 2015 and 2017 show a bimodal pattern with the two modes at "no listing" and "full listing". In addition, we find that hosts are more likely to list their properties longer in 2017 than in 2015.

We also explore the total number of days in a year that a property is listed on Airbnb, as the total revenue generated per year is more informative when compared with long-term rentals. Figure 3 shows the histogram of the percentage of days that a property is available for booking by year. In 2015, 51.9% of the properties are listed for less than half of the year. These "part-time" Airbnb hosts may list their properties only when they are not utilizing the property, for example, when they are away on vacation. In contrast, some properties are listed most of the time. The data show that 33.7% of the observations are listed for more than 70% of the year, 26.5% are listed for more than 80% of the year, and 18.3% are listed for more than 90% of the year. Some of these properties may have been in the long-term rental market had Airbnb not been available. The listing pattern in 2017 shows a very similar pattern with a slight shift to the right (i.e., longer listings).

3.2.1 Heterogeneity

The data patterns vary by metro area, property characteristics, and demographics. We present the data patterns averaged across years in this subsection, as they do not qualitatively change over time. Table 2 and Figure 4 show the observed percentages of Airbnb, long-term rental, and outside option units by metro area, number of bedrooms, and age group. First, these percentages vary significantly across metro areas. The percentage of Airbnb properties ranges from 0.34% in Detroit to 3.78% in San Francisco. The top three metro areas with the highest proportion of Airbnb properties are San Francisco, Miami, and New York. Second, the percentage of units choosing each option differs by property characteristics such as the number of bedrooms. As the number of bedrooms increases, the proportions of Airbnb properties and long-term rental properties both decrease in general, except the proportion of Airbnb increases for units with 4 or 5+ bedrooms. Third, the percentages of Airbnb units and long-term rental units first decrease with age and then significantly increase for seniors aged over 65, especially for Airbnb units. This is consistent with Airbnb's report that seniors are the fastest-growing demographic of Airbnb hosts.[12]

We also find that the Airbnb listing behavior varies by metro area, property characteristics, demographics, and season. Figure 5 shows a histogram of the number of days available for booking in a month by metro area, number of bedrooms, age, and season. Each observation represents a property-month combination. The overall bimodal pattern in Figure 2 holds for all subgroups, with variations across subgroups. For example, in terms of metro area, the percentage of no-listing months is 26.9% in Detroit and 39.8% in San Francisco.

[12]See https://www.airbnbcitizen.com/seniors-airbnbs-fastest-growing-most-loved-demographic/

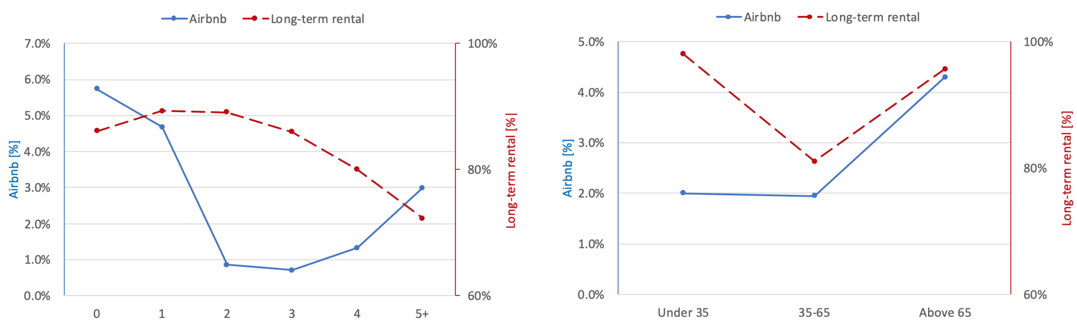
Table 2: Percentage of Units: By Metro Area

Metro Area	# Units	Percentage of Units [%]			
		Airbnb	Long-term Rental	Vacant Full Year	Vacant Part Year
Boston-Cambridge-Newton, MA-NH	1,232,717	2.47	91.70	3.97	1.87
Chicago-Naperville-Elgin, IL-IN-WI	2,370,656	1.19	89.41	8.38	1.02
Dallas-Fort Worth-Arlington, TX	1,938,368	0.58	94.98	3.55	0.89
Detroit-Warren-Dearborn, MI	1,010,764	0.34	84.06	14.10	1.50
Miami-Fort Lauderdale-West Palm Beach, FL	2,122,851	2.67	70.11	16.64	10.58
New York-Newark-Jersey City, NY-NJ-PA	6,425,847	2.96	89.43	5.43	2.18
Phoenix-Mesa-Scottsdale, AZ	1,263,736	1.19	85.69	6.17	6.95
San Francisco-Oakland-Hayward, CA	1,364,115	3.78	91.06	3.88	1.29
Washington-Arlington-Alexandria, DC-VA-MD-WV	1,530,661	2.26	90.81	5.29	1.64

Figure 4: Percentage of Units: By Property Characteristics and Demographics

(a) By Number of Bedrooms

(b) By Age Group

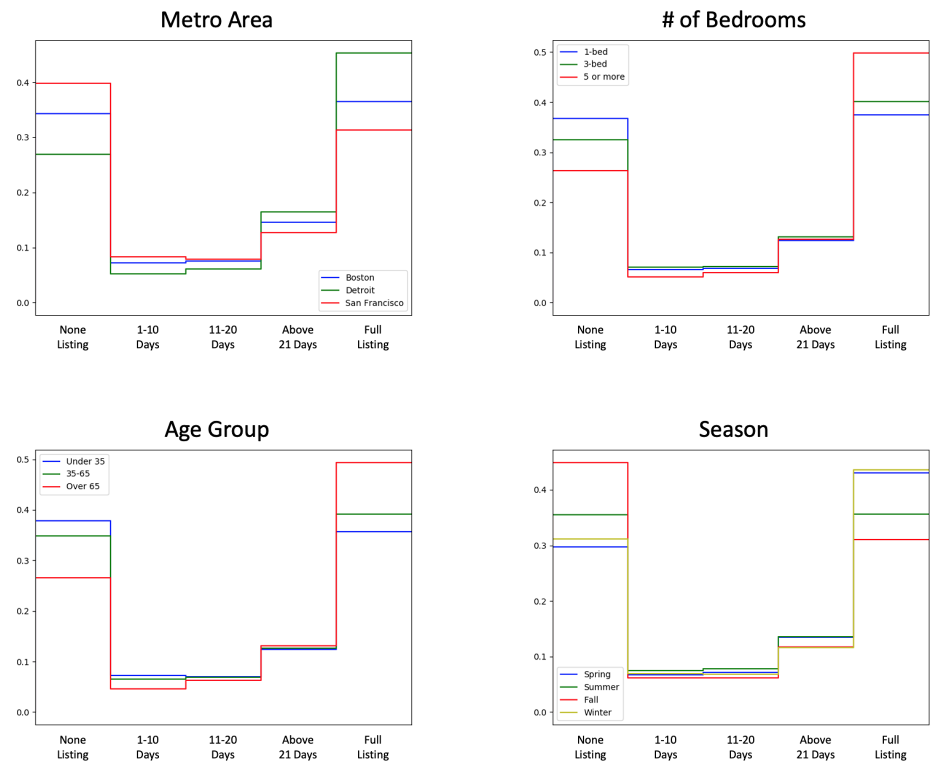


In terms of property characteristics, properties with more bedrooms are less likely to have no-listing months and more likely to have full-listing months. In terms of demographics, senior hosts are more likely to have full-listing months. Finally, properties are less likely to be listed in fall and more likely to be listed in spring and winter.

4 Model Setup

Property owners make endogenous decisions on whether and how long to list their properties based on cost-benefit trade-offs. Their decisions can be further affected by exogenously determined availability. Specifically, there are two types of properties. The first type is available for rent for the full year (hereafter, "full-available" type); the property owner does not need the property for self-use at all in a year. The second type is available for rent only for some part of the year (hereafter, "partial-available" type); the property owner may need to use the property for some part of the year. The availability type determines the choice set of the property owner. For full-available type, the property owners select the use of their properties among three options:

Figure 5: Monthly Number of Days Listed: Metro Area, Property Characteristics, Demographics, and Season



(1) Keep the property vacant without listing it on any market. (2) Rent on the long-term rental market during the next 12 months with some cost. (3) Rent on Airbnb with some cost. For the partial-available type, the property owners cannot choose the long-term rental option and only have options (1) and (3). The observed listing decisions are a result of both the endogenous decisions of the hosts and the exogenous availability types of the hosts.

In the model, property owners make decisions in two stages. In the first stage, at the beginning of each year, they select the use of their properties given their availability type and the corresponding options. In the second stage, if they choose Airbnb, they decide the number of days to list on Airbnb in each month. If they choose long-term rental or the outside option, there are no further decisions to make, as long-term rental hosts are bound by long-term leases during the lease period. This two-stage setup is a representation of the fact that property owners usually decide the use of their property at the year level and the number of days to list on Airbnb at the month level.

Property owners make decisions to maximize their profits given their expectations about the rent and occupancy rate they can obtain in the long-term rental market, and the price and occupancy rate they can obtain on Airbnb. We assume that the expectation is formed using a hedonic approach, which we detail in Section 4.3. We normalize the profit from the outside option to zero.

In addition to the revenues, property owners consider the costs of renting on the long-term rental market versus Airbnb. The costs can include both *tangible* costs (e.g., property maintenance) and *intangible* costs (e.g., hassle from dealing with renters, living with Airbnb travelers). Specifically, in the first stage, property owners may incur the cost of long-term rental and the fixed cost of Airbnb hosting. The cost of long-term rental may include fees, taxes, insurance, and maintenance costs. The fixed cost of Airbnb hosting may include the psychological cost of renting out property to transient guests, preparing property photos and descriptions, and preparing furnishings and amenities. In the second stage, Airbnb hosts may incur variable costs during the days they list their properties on Airbnb. These costs may include responding to guest inquiries and reservations, checking guests in and out, maintaining the property, and paying utility bills. We discuss how the cost functions are constructed and estimated in the following sections.

4.1 Second Stage: Continuous Decision of Listing on Airbnb

We first describe the model setup for the second-stage decision, as the profits from the second stage are nested into the first-stage decision and property owners must form expectations about the second stage before they make their first-stage decisions. In the second stage, conditional on choosing Airbnb in the first stage, the owner determines the number of days to list the property on Airbnb. Let s_{it} denote the number of days that property i is listed on Airbnb in month t. The owner chooses $s_{it} \in [0, \bar{s}]$, where the total number of days in each month $\bar{s}$ serves as the upper bound.[13] In the counterfactual analysis, we allow $\bar{s}$ to reflect the maximum listing length imposed by government regulations. In this section, we derive the model by allowing s_{it} to take any value between 0 and $\bar{s}$ for illustrative purposes. We account for the fact that s_{it} is an integer when we estimate the model and detail how we treat the integer issue in the appendix.

The optimal number of days to list is chosen to maximize the monthly profit from Airbnb for property i in month t:

$$\Pi_{it}^{A}(s_{it}) = p_{it}^{A}\phi_{it}^{A}s_{it} - c_{it}^{Av} \cdot \bar{s}\left(\exp\left(\frac{s_{it}}{\bar{s}}\right) - 1\right) \tag{1}$$

where p_{it}^{A} and ϕ_{it}^{A} are the expected average daily price and occupancy rate of property i. We discuss how the expectations are formed in Section 4.3. Here, c_{it}^{Av} is the heterogeneous *variable* cost of Airbnb hosting per day, to be parameterized later. The first term of the profit function represents the revenue, which is proportional to the number of days booked. The second term represents the cost, which increases with the

[13]In practice, there are government regulations that limit the maximum number of days in which a property can be listed on Airbnb. These regulations were imposed after our sample period ended; therefore, we do not account for them as $\bar{s}$ in our model estimation. The only exception is the Airbnb law in San Francisco, which went into effect on February 1, 2015, and restricts short-term rentals to a maximum of 90 days per year. However, the law was not strictly enforced, as the data show that 25% of the listings were listed for more than 90 days during the 9-month period from February 2015 (when the law went into effect) to October 2015. In fact, the lack of strict law enforcement was also reported during this time period. See https://www.sfchronicle.com/business/article/Airbnb-loses-thousands-of-hosts-in-SF-as-12496624.php.

number of days listed. Note that the profit is zero if the property is not listed, i.e., $\Pi_{it}^{A}(0) = 0$. Taking the derivative with respect to s_{it}, the optimal number of days to list is as follows:

$$s_{it}^{*} = \min\left\{\bar{s} \cdot \ln\left(\frac{p_{it}^{A}\phi_{it}^{A}}{c_{it}^{Av}}\right), \bar{s}\right\} \tag{2}$$

where the min operator accounts for the range of $s_{it} \in [0, \bar{s}]$. The solution suggests that the number of days to list on Airbnb is an endogenous function of the ex ante expected revenue $\left(p_{it}^{A}\phi_{it}^{A}\right)$ and heterogeneous cost of Airbnb hosting $\left(c_{it}^{Av}\right)$. It has the desirable property such that the larger the revenue-to-cost ratio is, the longer the property owner will choose to list on Airbnb.

The variable cost of Airbnb hosting for property i in month t is formulated as follows:

$$c_{it}^{Av} = \bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av} + \epsilon_{it}^{Av} \tag{3}$$

where $\bar{c}^{Av}$ is the baseline cost, X_{it}^{Av} are observed characteristics that affect the cost in a continuous way, ξ_{mT}^{Av} is a market-specific time fixed effect that captures any remaining time-varying unobservables, and ϵ_{it}^{Av} is an independently normally distributed idiosyncratic shock with mean zero and standard deviation σ_2.[14] Specifically, X_{it}^{Av} includes property characteristics (number of bedrooms/bathrooms/amenities, listing type), host demographics (age, education, income, marital status, gender), season fixed effects (spring, summer, fall, winter), and a set of metro-area-level characteristics that relate to the cost of hosting. The metro-area-level characteristics include population, density, mortgage affordability index, average wage and employment in the accommodation industry (measured as the percentage of the population who work in the accommodation industry), and Airbnb regulation score (measures how friendly city regulations are to short-term rental). Intuitively, a large property may be more costly to maintain and induce a larger variable cost of hosting. The hosting cost can vary by host demographics and across seasons even for the same host. Hosts in cities with more employees and lower wages in the accommodation industry may find it easier to obtain room maintenance services and thus have a lower variable cost of hosting. Hosts in cities with more favorable Airbnb regulations may face a lower variable cost of hosting. The level of mortgage pressure can further impact how long the hosts would like to list their property. Finally, the market-specific time fixed effect is specified as $\xi_{mT}^{Av} = \xi_{0}^{Av} \cdot 1\{T = 2017\} + \xi_{1}^{Av} \cdot \left(T - T_{m}^{0}\right)$, where T_{m}^{0} represents the year when Airbnb reached 10% of the total rooms supplied by hotels and Airbnb in a city. The first component captures any yearly unobservables that affect all metro areas (e.g., because of Airbnb's national marketing) relative to those in

[14] Note that this specification does not restrict the cost to be nonnegative. We keep the specification flexible to accommodate extreme cases; for instance, a host may derive positive utility (i.e., negative cost) from interacting with guests. In the case of a negative cost, the optimal number of days listed is bounded by $\bar{s}$; thus, the property owners choose to list for the full month ($s^{*} = \bar{s}$). We also bound the monthly profit Π_{it}^{A} by $p_{it}^{A}\bar{s}$, which is the maximum profit that a listing can possibly generate.

the baseline year 2015. The second component captures any market-specific time trend related to Airbnb history or how long Airbnb has been present in a city. For instance, Airbnb may be better received in markets where it has been present longer.

We summarize the covariates that enter the Airbnb variable cost (c_{it}^{Av}) in Column 3 of Table 3.

4.2 First Stage: Discrete Decision of Where to List

In the first stage, the property owners decide whether and where to list their properties given their availability types. The full-available property owners choose among long-term rental, Airbnb, and the outside option given the expected yearly profit from the second-stage decision for each option. Let d_{iT} denote the decision of property owner i in year T, and index the alternatives by superscripts A (Airbnb), R (long-term rental), and O (outside option). The full-available property owners solve the following problem:

$$\begin{aligned}
&\max_{d\in\{A,R,O\}} \Pi_{iT}^{d} \\
&\Pi_{iT}^{A} = \sum_{t\in T} \left(E\left[\Pi_{it}^{A}(s_{it}^{*})\right]\right) - c_{iT}^{Af} \\
&\Pi_{iT}^{R} = p_{iT}^{R}\phi_{iT}^{R} - c_{iT}^{R} \\
&\Pi_{iT}^{O} = 0
\end{aligned} \tag{4}$$

The partial-available owners solve a similar problem without the long-term rental option:

$$\begin{aligned}
&\max_{d\in\{A,O\}} \Pi_{iT}^{d} \\
&\Pi_{iT}^{A} = \sum_{t\in T} \left(E\left[\Pi_{it}^{A}(s_{it}^{*})\right]\right) - c_{iT}^{Af} \\
&\Pi_{iT}^{O} = 0
\end{aligned} \tag{5}$$

Here, Π_{iT}^{d} represents the yearly profit from each alternative $d \in \{A, R, O\}$. The profit of the outside option is normalized to zero. The profit of long-term rental comes from the ex ante expected yearly rent (p_{iT}^{R}) multiplied by the expected occupancy rate (ϕ_{iT}^{R}) minus the cost of long-term rental (c_{iT}^{R}). The profit of Airbnb comes from the sum of the ex ante monthly profit from Airbnb hosting $\left(\sum_{t\in T}\left(E\left[\Pi_{it}^{A}(s_{it}^{*})\right]\right)\right)$ minus the fixed cost of Airbnb hosting (c_{iT}^{Af}). The ex ante monthly profit from Airbnb hosting is obtained by substituting the optimal number of days to list in Equation 2 into Equation 1 and taking expectations over the error terms ϵ_{it}^{Av} in c_{it}^{Av}:

$$E\left[\Pi_{it}^{A}(s_{it}^{*})\right] = \left[\int_{-\infty}^{\infty} \Pi_{it}^{A}\left(s_{it}^{*}\right) f\left(\epsilon_{it}^{Av}\right) d\epsilon_{it}^{Av}\right] \tag{6}$$

Table 3: Summary of Covariates

	Owner type	Cost		Revenue Hedonic regression			
	γ_{iT}	c_{iT}^{Af}	c_{it}^{Av}	p_{it}^{A}	ϕ_{it}^{A}	p_{iT}^{R}	ϕ_{iT}^{R}
Property		yes	yes	yes	yes	yes	yes
Demographics	yes		yes	yes	yes	yes	yes
Metro area	yes		yes	yes (metro-year-month)	yes (metro-year-month)	yes (metro-year)	yes (metro-year)
Season fixed effect			yes				
Mortgage		yes	yes				
Wage and employment in accomm. industry		yes	yes				
Airbnb regulation score			yes				
Airbnb history		yes	yes	yes	yes		
Tourism (air passengers)				yes	yes		
Total Airbnb supply				yes	yes		
Airbnb price					yes		
Total rental supply						yes	yes
Rent							yes

Notes: Metro area characteristics include population and density. Property characteristics include number of bedrooms/bathrooms/amenities and property type: house (dummy). Demographics include age 35-65, age over 65, high school education, bachelor's education, 50-100k income, over 100k income, male, and marital status never. The baseline demographic group is age below 35, education below high school, income below 50k, female, and married.

We show in the appendix how we compute the ex ante monthly profit, accounting for s_{it} being integers.

Property owners are heterogeneous in their fixed cost of Airbnb hosting and the cost of long-term rental. The cost of long-term rental for property i in year T is assumed to be $c_{iT}^{R} = \epsilon_{iT}^{R}$, where ϵ_{iT}^{R} is independently normally distributed with mean 0 and variance σ_1^2 and is independent of the second-stage error terms $\{\epsilon_{it}^{Av}\}$.

The fixed cost of Airbnb hosting for property i in year T is:

$$c_{iT}^{Af} = \bar{c}_{\tau}^{Af} + \beta^{Af} X_{iT}^{Af} + \xi_{mT}^{Af} + \epsilon_{iT}^{Af} \tag{7}$$

where $\bar{c}_{\tau}^{Af}$ is the baseline cost that takes different values for full-available owners ($\tau = 1$) and partial-available owner ($\tau = 2$). Intuitively, the two types of owners can have different Airbnb fixed costs because their availability affects their psychological and tangible costs of adopting Airbnb. ξ_{mT}^{Af} is a market-specific time fixed effect that captures any time-varying unobservables; ϵ_{iT}^{Af} is an idiosyncratic shock that is independently normally distributed with mean 0 and variance σ_1^2 and is independent of the second-stage error terms $\{\epsilon_{it}^{Av}\}$. The linear component X_{iT}^{Af} includes metro area characteristics (mortgage affordability index, average wage and employment in the accommodation industry) and property characteristics (number of bedrooms/bathrooms/amenities, property type). The market-specific time fixed effect is specified as $\xi_{mT}^{Af} = \xi_0^{Af} \cdot 1\{T = 2017\} + \xi_1^{Af} \cdot \left(T - T_m^0\right)$, where T_m^0 represents the year when Airbnb reached 10% of the total rooms supplied by hotels and Airbnb in a city. The first component captures any yearly unobservables that affect all metro areas, and the second component captures any market-specific time trend that is related to Airbnb history in a city.

Owner availability types. In the data, we observe the owner availability types for properties that are on the long-term rental market and are kept vacant. Specifically, owners of units on the long-term rental market in the data are the full-available type because they are able to list the property for the full year. Owners of properties that are kept vacant for the full year are the full-available type, whereas owners of properties that are kept vacant for the some part of the year due to occasional self-use are the partial-available type. However, we do not observe the owner availability types for properties on Airbnb in the data.[15] There are two exceptions: 1) properties that are listed every month throughout the year in the data must be the full-available type, as availability is a necessary condition for listing; 2) properties that are listed as a “private room” (instead of an “entire place”) must be the partial-available type because the hosts live with the guests in this case. Apart from these two cases, we do not know whether an Airbnb host is full-available or partial-available.

[15] Note that one cannot conclude that Airbnb properties that are listed for some part of the year must be the partial-available type. This is because the observed listing pattern is a result of both the endogenous decision of the hosts and the exogenous host availability type. A full-available host may choose to list for only part of the year because the costs exceed the benefits for the rest of the year.

For model estimation purposes, we adopt a probabilistic view on the availability type of Airbnb hosts in the data. The probability that an Airbnb property i is the full-available type ($\tau = 1$) in year T is

$$\Pr(\tau = 1) = \gamma_{iT} = \frac{\exp(\beta X_{iT})}{1 + \exp(\beta X_{iT})} \tag{8}$$

The probability of being the partial-available type is $\Pr(\tau = 2) = 1 - \gamma_{iT}$. Here, X_{iT} includes host demographics (age, education, income, marital status, gender), metro area characteristics (population and density), and a dummy for being a single bedroom. Intuitively, a host's availability can be related to who the host is, where the host lives, and what type of property the host has.

We summarize the covariates that enter owner availability type (γ_{iT}) and Airbnb fixed cost (c_{iT}^{Af}) in Columns 1 and 2 of Table 3.

4.3 Expectation on Revenue

Property owners' decisions depend on revenue information, i.e., rent, rental occupancy rate, Airbnb price, and Airbnb occupancy rate. We assume that property owners form expectations over these variables using a typical hedonic approach when making their first-stage decisions. Hedonic regression is a widely used method to estimate property value by decomposing a property's value into its constituent attributes and obtaining contributory values for each attribute (see Sirmans, Macpherson, and Zietz (2005) for a review on using hedonic models to estimate house prices). We use the hedonic approach because it offers the following three advantages. First, the hedonic model incorporates property heterogeneity, which allows us to construct expected revenues for each property in the data. It also parsimoniously captures how hosts set prices and how occupancy rates are determined in practice. Second, this approach allows us to obtain rent, rental occupancy rate, Airbnb price, and Airbnb occupancy rate regardless of how the units are utilized. In the data, we observe rent and rental occupancy only for long-term rental properties and observe Airbnb price and occupancy rate only for Airbnb properties. However, the property attributes are observed for all properties. The hedonic model allows us to construct the expected rent and rental occupancy for properties listed on Airbnb and the expected Airbnb price and occupancy rate for properties listed on the long-term rental market. The underlying assumption is that properties with similar attributes will have similar revenues. Third, the hedonic approach allows us to generate counterfactual rent, rental occupancy, Airbnb price, and Airbnb occupancy under counterfactual scenarios, which we discuss in detail in Section 7.2. Li and Srinivasan (2019) adopt a similar approach by first estimating how prices and supply are determined in the data and using the estimates to generate new prices and supply in the counterfactual analysis.

The hedonic models of rent and rental occupancy for property i in year T are

$$p_{iT}^R = \rho_0 + \rho_1 x_i^P + \rho_2 x_i^D + \rho_3 S_{mT}^R + \psi_{mT}^{Rp} + \varepsilon_{iT}^{Rp} \tag{9}$$

$$\phi_{iT}^R = \eta_0 + \eta_1 x_i^P + \eta_2 x_i^D + \eta_3 S_{mT}^R + \psi_{mT}^{Ro} + \eta_4 p_{iT}^R + \varepsilon_{iT}^{Ro} \tag{10}$$

where we regress the rent of property i in year T, p_{iT}^R, on property characteristics x_i^P, household demographics x_i^D, rental supply of comparable units in the metro area S_{mT}^R (measured as the number of comparable units that choose to list on the long-term rental market), and metro-year fixed effects ψ_{mT}^{Rp}.[16] Here, m denotes the metro area to which property i belongs. The hedonic model of the rental occupancy ϕ_{iT}^R uses the same specification except it also includes rent as an additional regressor because the occupancy rate depends on the price. To run the rental occupancy regression, we supplement the original data set, which contains long-term rental properties (i.e., rented properties in the AHS data set), with data on properties that are for rent but not rented from the AHS data set.[17] We exclude outliers with rents below the 10th percentile when running the regressions.

The rental supply of comparable units S_{mT}^R and the rent p_{iT}^R are potentially endogenous variables. We address the endogeneity issue using BLP-type instruments. Specifically, we instrument the rental supply S_{mT}^R using the rental supply of noncomparable units in market m in year T (normalized by the total rental supply in market m in year T), and we instrument the rent p_{iT}^R using average property characteristics of noncomparable rental properties in market m in year T. The instruments pass the weak IV test.[18] We jointly estimate the rent and rental occupancy equations as a system of equations using three-stage least squares (3SLS) to allow for the correlation of the error terms in the two equations.

The hedonic models of Airbnb price and occupancy rate for property i in month t are

$$p_{it}^A = \delta_0 + \delta_1 x_i^P + \delta_2 x_i^D + \delta_3 S_{mt}^A + \delta_4 x_{mt}^A + \psi_{mt}^{Ap} + \varepsilon_{it}^{Ap} \tag{11}$$

$$\phi_{it}^A = \gamma_0 + \gamma_1 x_i^P + \gamma_2 x_i^D + \gamma_3 S_{mt}^A + \gamma_4 x_{mt}^A + \psi_{mt}^{Ao} + \gamma_5 p_{it}^A + \varepsilon_{it}^{Ao} \tag{12}$$

where we regress the monthly logged average daily price of property i in month t, p_{it}^A, on property characteristics x_i^P, household demographics x_i^D, Airbnb supply of comparable units in the metro area S_{mt}^A (measured as the number of days listed by all comparable units), Airbnb-related metro area variables x_{mt}^A, and metro-

[16]"Comparable" units are those with the same number of bedrooms. We conduct a robustness check by defining comparable units as those with the same numbers of bedrooms and bathrooms and obtain robust estimates. We keep the original definition because it produces a larger R-squared of the regressions.

[17]The average occupancy rate, or the fraction of rented properties among all for-rent (rented and for-rent but not rented) properties, is 91.9% in the data.

[18]The instruments pass the weak IV test with the first-stage regression F-statistics of 3976.76 for rental supply in the rent regression, 7825.27 for rental supply in the rental occupancy regression, and 88.48 for rent in the rental occupancy regression.

area-specific year and month fixed effects ψ_{mt}^{Ap}. The Airbnb-related metro area variables include tourism (measured as the number of air passengers to the city), which can proxy for the heterogeneous tourism popularity across cities, and Airbnb history (measured as the number of months since Airbnb reached 10% of the total rooms supplied by hotels and Airbnb in a city), which can proxy for unobserved factors that relate to the length of Airbnb presence. The market-specific time fixed effects can capture market-specific seasonality patterns in Airbnb prices. The hedonic model of the occupancy rate uses the same specification except it also includes the Airbnb price as an additional regressor because the occupancy rate depends on the price.

The Airbnb supply of comparable units S_{mt}^{A} and the Airbnb price p_{it}^{A} are potentially endogenous variables. We address the endogeneity issue by instrumenting Airbnb supply S_{mt}^{A} with the Airbnb supply of noncomparable units in market m in month t and instrumenting the Airbnb price p_{it}^{A} with the average property characteristics of noncomparable Airbnb properties in market m in month t. In addition to these BLP-type instruments, we further include metro-area-level Airbnb regulation score, rent-to-own ratio, and unemployment rate as instruments for Airbnb supply. These variables are valid instruments because they serve as cost shifters and affect the hosts' incentive to list their properties, so they are correlated with Airbnb supply; they do not affect tourists' incentives, so they do not directly affect Airbnb demand. The instruments pass the weak IV test.[19] We jointly estimate the Airbnb price and occupancy equations as a system of equations using three-stage least squares (3SLS).

We summarize the covariates that enter each hedonic regression in Columns 5-8 of Table 3.

To generate the expected rent, rental occupancy rate, Airbnb price, and Airbnb occupancy rate for all properties, we first estimate the two systems of equations using the observed revenues and property attributes. Specifically, we use the observed long-term rental data from the AHS to estimate the hedonic models of rent and rental occupancy, and we use the observed Airbnb data to estimate the hedonic models of Airbnb price and occupancy. Once we obtain the estimates, we can generate the expected rent, rental occupancy rate, Airbnb price, and Airbnb occupancy rate for all properties in the Airbnb and AHS data.[20] These expected revenues are used in the property owners' decisions in Equations 1, 4 and 5.

The regression results are provided in the appendix. The coefficients have the expected signs. For example, both rent and Airbnb price increase with the number of bedrooms and bathrooms. An increase

[19]The instruments pass the weak IV test with the first-stage regression F-statistics of 1,963,222 for Airbnb supply in the Airbnb price regression, 2,694,075 for Airbnb supply in the Airbnb occupancy regression, and 10,997 for Airbnb price in the Airbnb occupancy regression.

[20]One caveat is that the hedonic models for the Airbnb price and occupancy rate contain the variable for listing type (entire place, private room, shared room), which is not available for properties in the AHS data. We assume that they will be listed on Airbnb as the entire place rather than as private or shared rooms, as most of these properties are the full-available type, meaning that the hosts do not live with the guests. In addition, entire places are the most common type on Airbnb. The results are robust if we allow the properties to be listed as private or shared rooms with a probability equal to the empirical fraction of private or shared room listings in the data.

in the aggregate rental supply is associated with a reduction in rent and rental occupancy. A higher rent decreases rental occupancy. Similarly, a higher aggregate Airbnb supply decreases the Airbnb price and a higher Airbnb price decreases the Airbnb occupancy rate. Finally, an increase in the aggregate Airbnb supply is associated with an increase in occupancy rate. This is because the Airbnb supply can have two opposite effects on the occupancy rate of a particular listing. First, a larger Airbnb supply can imply more competition and reduce the per-listing occupancy rate. Second, a larger Airbnb supply can attract more guests to the platform (i.e., the classic indirect network effect on two-sided platforms) and increase the per-listing occupancy rate. Empirically, we find that the second effect dominates the first effect; the indirect network effect appears to be a relatively stronger factor as Airbnb grew rapidly during our sample period.

5 Estimation Method

5.1 Estimation

We use the maximum likelihood estimation (MLE) method to estimate the model. The likelihood function for individual i is the joint probability of the individual's decision on the use of the property and, if the Airbnb option is chosen, the number of days to list on Airbnb:

$$l_i\left(\Theta | d_{iT}, s_{it}, \mathcal{X}_i\right) = \prod_{T} \left\{ \left[\Pr\left(d_{iT} \mid \mathcal{X}_i\right)\right]^{1(d_{iT} \neq A)} \prod_{\tau=1,2} \left[\Pr\left(d_{iT} \mid \mathcal{X}_i, \tau\right)\right]^{1(d_{iT}=A)\Pr(\tau \mid \mathcal{X}_i)} \prod_{t \in T} \left(\Pr\left(s_{it} \mid \mathcal{X}_i\right)\right)^{1(d_{iT}=A)} \right\}$$

where $\mathcal{X}_i$ contains all host demographics, metro area and property characteristics that affect the costs and revenues of individual i, $\Pr\left(d_{iT} \mid \mathcal{X}_i\right)$ is the probability of the first-stage decision, and $\Pr\left(s_{it} \mid \mathcal{X}_i\right)$ is the probability of the second-stage decision (derived below). $\Pr\left(\tau \mid \mathcal{X}_i\right)$ is the probability of owner availability types for Airbnb properties.[21] One caveat is that, as discussed in Section 3.1, we do not observe host demographics for Airbnb properties. We use the zip-code-level demographics distribution $f(\mathcal{X}_i)$ from the ACS data. For properties in the Airbnb data, the likelihood of individual i is integrated over the demographic distributions, $l_i\left(\Theta | d_{iT}, s_{it}, f(\mathcal{X}_i)\right) = \int_{\mathcal{X}_i} l_i\left(\Theta | d_{iT}, s_{it}, \mathcal{X}_i\right) f\left(\mathcal{X}_i\right) d\mathcal{X}_i$. The total log-likelihood is

$$\log \mathcal{L}\left(\Theta\right) = \sum_i \log l_i\left(\Theta | d_{iT}, s_{it}, f(\mathcal{X}_i)\right)$$

Given that we conduct a two-step estimation (i.e., estimate the hedonic regressions in the first step and the

[21] As discussed in Section 4.2, we do observe the owner availability types for two categories of Airbnb properties: 1) properties that are listed every month in the data must be the full-available type ($\tau_i = 1$); 2) properties that are listed as a "private room" must be the partial-available type because the hosts live with the guests ($\tau_i = 2$). For these properties with observed τ_i, their likelihood function has an additional component, $\Pr\left(\tau = \tau_i \mid \mathcal{X}_i\right)^{1(\tau=\tau_i)}$.

hosts' decisions in the second step), we correct the standard errors following Murphy and Topel (1985).

Derivation of the first-stage probability. $\Pr\left(d_{iT} \mid \mathcal{X}_i\right)$ is constructed based on the feasible range of the independently normally distributed error terms implied by the optimal choices in Equations 4 and 5. Define $a_\tau \equiv \sum_{t \in T} E\left[\Pi_{it}^A(s_{it}^*) - c_{iT}^{Af} \,|\tau\right]$ and $r \equiv E\left(p_{iT}^R \phi_{iT}^R - c_{iT}^R\right)$. The probabilities are:

$$
\begin{aligned}
\Pr\left(d_{iT} = O_1 \mid \mathcal{X}_i\right) &= \Pr\left(\epsilon_{iT}^{Af} > a_1, \epsilon_{iT}^R > r\right) = \left(1 - \Phi\left(\frac{a_1}{\sigma_1}\right)\right)\left(1 - \Phi\left(\frac{r}{\sigma_1}\right)\right) \\
\Pr\left(d_{iT} = O_2 \mid \mathcal{X}_i\right) &= \Pr\left(\epsilon_{iT}^{Af} > a_2\right) = \left(1 - \Phi\left(\frac{a_2}{\sigma_1}\right)\right) \\
\Pr\left(d_{iT} = R \mid \mathcal{X}_i\right) &= \Pr\left(\epsilon_{iT}^R < r, \epsilon_{iT}^R - \epsilon_{iT}^{Af} < r - a_1\right) \\
&= \Phi\left(\frac{r}{\sigma_1}\right) - \int_{-\infty}^{r} \Phi\left(\frac{\epsilon^R - r + a_1}{\sigma_1}\right)\phi\left(\frac{\epsilon^R}{\sigma_1}\right) d\epsilon^R \\
\Pr\left(d_{iT} = A \mid \mathcal{X}_i, \tau = 1\right) &= \Pr\left(\epsilon_{iT}^{Af} < a_1, \epsilon_{iT}^{Af} - \epsilon_{iT}^R < a_1 - r\right) \\
&= \Phi\left(\frac{a_1}{\sigma_1}\right) - \int_{-\infty}^{a} \Phi\left(\frac{\epsilon^{Af} - a_1 + r}{\sigma_1}\right)\phi\left(\frac{\epsilon^{Af}}{\sigma_1}\right) d\epsilon^{Af} \\
\Pr\left(d_{iT} = A \mid \mathcal{X}_i, \tau = 2\right) &= \Pr\left(\epsilon_{iT}^{Af} < a_2\right) = \Phi\left(\frac{a_2}{\sigma_1}\right)
\end{aligned}
$$

where O_1 refers to the full-available units that choose the outside option and corresponds to the "vacant full year" observations in the data. O_2 refers to the partial-available units that choose the outside option and corresponds to the "vacant part year" observations in the data. $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of the standard normal distribution. The integrals in $\Pr\left(d_{iT} = R \mid \mathcal{X}_i\right)$ and $\Pr\left(d_{iT} = A \mid \mathcal{X}_i, \tau = 1\right)$ are calculated using Gauss-Laguerre quadrature with 10 nodes.

Derivation of the second-stage probability. $\Pr(s_{it} \mid \mathcal{X}_i)$ is constructed based on the feasible range of the normally distributed error term $\{\epsilon_{it}^{Av}\}$ implied by the optimal choices in Equation 2:

$$
\begin{aligned}
\Pr(s_{it} = 0 \mid \mathcal{X}_i) &= \Pr\left(\epsilon_{it}^{Av} > p_{it}^A \phi_{it}^A - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right) \\
&= 1 - \Phi\left(\frac{1}{\sigma_2}\left(p_{it}^A \phi_{it}^A - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right)\right) \\
\Pr(s_{it} = s, 0 < s < \bar{s} \mid \mathcal{X}_i) &= \Pr\left(\epsilon_{it}^{Av} = \frac{p_{it}^A \phi_{it}^A}{\exp(s/\bar{s})} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right) \\
&= \phi\left(\frac{1}{\sigma_2}\left(\frac{p_{it}^A \phi_{it}^A}{\exp(s/\bar{s})} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right)\right) \\
\Pr(s_{it} = \bar{s} \mid \mathcal{X}_i) &= \Pr\left(\epsilon_{it}^{Av} < \frac{p_{it}^A \phi_{it}^A}{\exp(1)} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right) \\
&= \Phi\left(\frac{1}{\sigma_2}\left(\frac{p_{it}^A \phi_{it}^A}{\exp(1)} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})\right)\right)
\end{aligned}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of the standard normal distribution.

5.2 Identification

We observe three types of information in the data: the revenues, the first-stage decision of the hosts, and the second-stage decision of the hosts. They are used to identify the revenue-side parameters in the hedonic regressions and the cost-side parameters in the hosts' decisions. The revenue-side parameters are directly identified and obtained by regressing the observed revenues on the observed characteristics. Given the revenues, the cost-side parameters are identified by the hosts' decisions.

The cost-side parameters in the second-stage decision include the Airbnb variable cost parameters $\{\bar{c}^{Av}, \beta^{Av}, \xi_0^{Av}, \xi_1^{Av}, \sigma_2\}$, which are identified by the Airbnb listing pattern. In particular, the average number of days listed and its variation across host demographics, properties, metro areas, and seasons identify the Airbnb variable cost parameters $\left\{\bar{c}^{Av}, \beta^{Av}\right\}$. The time-related parameters $\left\{\xi_0^{Av}, \xi_1^{Av}\right\}$ are identified by the listing pattern differences across years and markets with different lengths of Airbnb history.

The cost-side parameters in the first-stage decision include the owner availability type parameters β^a, the Airbnb fixed cost parameters $\{\bar{c}_\tau^{Af}, \beta^{Af}, \xi_0^{Af}, \xi_1^{Af}\}$, and the standard deviation of the idiosyncratic shocks σ_1. The fraction of properties that choose Airbnb and its variation across metro areas, demographics, properties, and over time identify the Airbnb fixed cost parameters $\{\bar{c}_\tau^{Af}, \beta^{Af}, \xi_0^{Af}, \xi_1^{Af}\}$. The owner availability type parameters β^a are identified from the two groups of Airbnb properties with observed owner availability types (i.e., listed every month in a year; private room listings) through variations in their host demographics, metro areas, and property characteristics.

Note that the cost-side parameters entering the second-stage decision can also be identified by the data on the first-stage decision, and vice versa, as the first and second stages are linked. The expected Airbnb profit from the second stage enters the first-stage decision; thus, the data on the first-stage decision impose over-identifying restrictions on the parameters in the second stage. Similarly, the identification of the parameters in the first stage is also affected by the data on the second-stage decision.

Exclusion restrictions. Property characteristics, host demographics, metro area characteristics enter both the cost-side components and the revenue-side regressions. Exclusion restrictions come from the nonoverlapping variables. As summarized in Table 3, each cost component or revenue regression has exclusive variables that do not enter other components. For instance, the aggregate supply of Airbnb and rental units affects the revenues but not the costs: they affect the prices and occupancy rates through competition in the market; however, they do not directly influence the hosting costs of an individual host. Additionally, tourism (number of air passengers to a city) affects only the revenue side because it captures the demand of tourists, whereas mortgage affects only the cost side because it influences the incentives of hosts. Finally, Airbnb-related variables (e.g., Airbnb history) affect only Airbnb and not the long-term rental market. In general, the exclusion restrictions stem from the fact that renters (demand) and hosts (supply) face different trade-offs when making their decisions and that Airbnb and the long-term rental market serve different consumers (tourists and local renters).

For the overlapping variables that appear in both the cost and the revenue sides, they are separately identified because we observe three types of information (revenues, first-stage decision of whether to list, and second-stage decision of how long to list) and how they vary by the overlapping variables.[22] Consider as an example, the number of bedrooms. It enters both the cost-side components and the revenue-side regressions. First, the revenue-side parameters on the number of bedrooms are directly identified by how the observed prices and occupancy rates change with the number of bedrooms. Second, conditional on the revenues, the cost-side parameters are identified by the variation in the observed first- and second-stage decisions with respect to the number of bedrooms. Specifically, the parameters in the Airbnb fixed cost and variable cost are separately identified if, for instance, properties with more bedrooms are more likely to choose Airbnb in the first stage but list shorter in the second stage; in this case, the coefficient on the number of bedrooms is negative in Airbnb fixed cost and positive in Airbnb variable cost.

6 Estimation Results

6.1 Model Fit

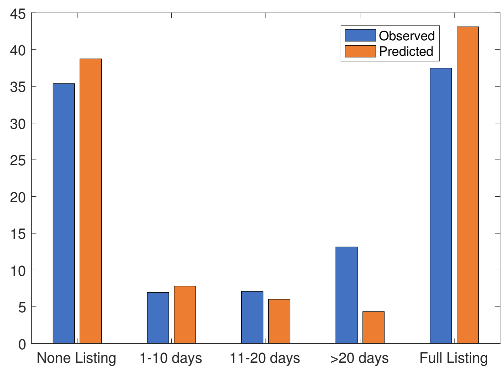
Table 4 shows the observed and predicted percentages of Airbnb, long-term rental, and outside option properties, and Figure 6 shows the observed and predicted Airbnb listing patterns. The model fits the first- and second-stage decisions well. It is also capable of fitting the heterogeneity for both decisions. Figure 7 presents the percentage of Airbnb properties for the first-stage model fit (left) and the percentage of unit-month observations with no listing for the second-stage model fit (right), by property characteristics, metro area, and demographics. The model captures the data pattern: as the number of bedrooms increases, the percentage of Airbnb properties decreases except for units with 4 or 5+ bedrooms, and the percentage of no-listing months decreases. The estimated percentages of the Airbnb properties and no-listing months are also comparable to the observed percentages for each metro area and demographic group. Overall, these results suggest that the model can recover the heterogeneous hosting costs across property characteristics, metro areas, and demographics.

[22]Note that there is no overlap between the covariates in owner availability type (γ_{iT}) and Airbnb fixed cost (c_{iT}^{Af}). This is because owner availability type τ also determines the baseline Airbnb fixed cost $\bar{c}_{\tau}^{Af}$. Therefore, the covariates in owner availability type affect Airbnb fixed cost c_{iT}^{Af} through $\bar{c}_{\tau}^{Af}$ and do not need to be duplicated in c_{iT}^{Af}.

Table 4: Model Fit: First-Stage Decision

[%]	Airbnb	Rental	Vacant full year	Vacant part year
Observed	2.19	87.70	7.13	2.98
Predicted	2.47	87.16	7.00	2.95

Figure 6: Model Fit: Second-Stage Decision



6.2 Parameter Estimates

Tables 5a and 5b report the parameter estimates for the first-stage and second-stage decisions.

Airbnb variable cost. The estimates of $\bar{c}^{Av}$ and β^{Av} suggest that the average variable cost of Airbnb hosting is \$29.6 per day, with a 25 percentile of \$11.6 and a 75 percentile of \$49.5. The estimates suggest that additional bedrooms, bathrooms, and facilities increase the variable cost of hosting. The daily cost for an entire place listing is \$32.4 greater than that of a private or shared room listing.

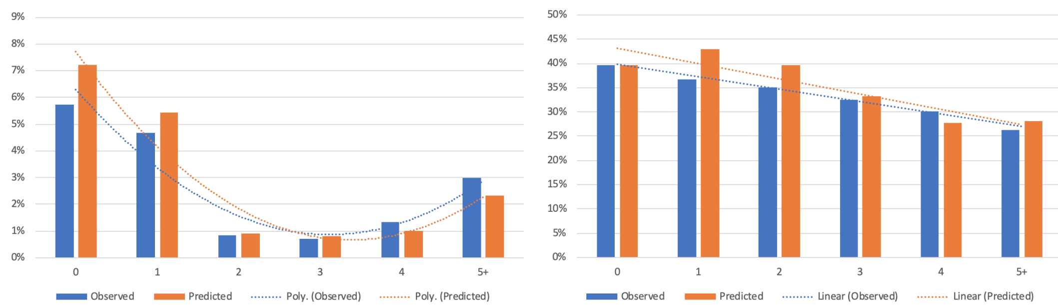
Note that these estimates are very comparable to the prices charged by third-party short-term rental cleaning services, which serve as an out-of-sample validation for our estimates. For example, Tidy charges between \$40 and \$45 for cleaning a one-bedroom unit, which is comparable to our estimate of \$33.9.[23]

Host demographics and metro area characteristics also affect the Airbnb variable cost and how long hosts list their properties. Hosts have a lower estimated variable cost and list longer if they are younger, have a high school education level and medium income, are male, and are married. Hosts list longer in cities with larger population, lower density, lower mortgage pressure, more favorable Airbnb regulation scores, and a longer Airbnb history. Hosts also list longer if there are more employment and lower wages in the accommodation industry, as resources in this industry such as room cleaning can also be used for Airbnb hosting and may facilitate Airbnb hosting. The estimated variable cost is lower in 2017 than in 2015 and in winter than in fall.

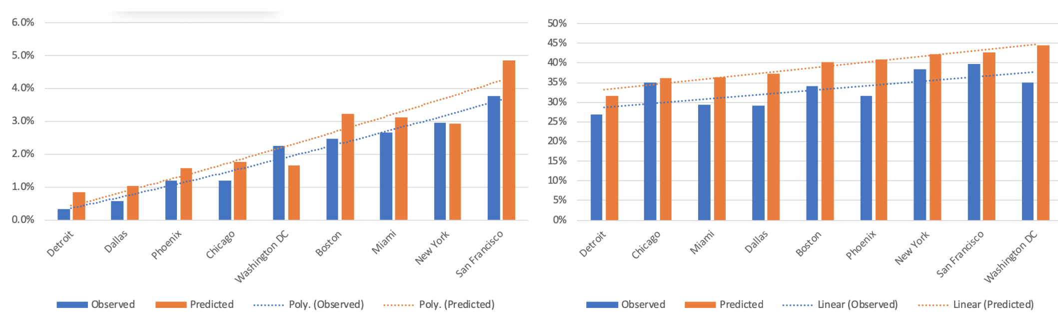
[23]See https://www.tidy.com/compare-house-cleaning-prices

Figure 7: Model Fit: By Property Characteristics, Metro Area, and Demographics

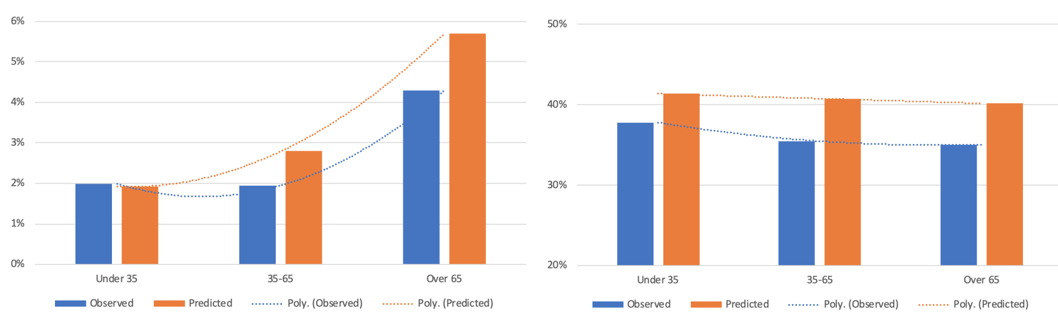
(a) By Property Characteristics (Number of Bedrooms)



(b) By Metro Area



(c) By Demographics (Age)



Note: The y-axis of the plots in the left column represents the percentage of Airbnb properties for the first-stage decision. The y-axis of the plots in the right column represents the percentage of unit-month observations with no listing for the second-stage decision. In Panel (b), metro areas are sorted by the observed percentages.

Table 5: Parameter Estimates

(a) First-Stage Decision

		Est.	Std.			Est.	Std.
$\bar{c}_1^{Af}$		50.866**	0.653		age: 35-65	1.207e-4	1.254e-4
$\bar{c}_2^{Af}$		24.794**	0.599		age: over 65	1.178**	0.0094
σ_1		8.138**	0.0213		edu: high school	1.384**	0.0054
	# of bedrooms	7.823**	0.104		edu: bachelor	1.255**	0.0124
	# of bathrooms	9.320**	0.122		income: 50K-100K	-1.189**	0.0228
	# of amenities	-2.223**	0.0196		income: over 100K	-0.372**	0.0036
	type: house (dummy)	0.117**	0.0123	β^a	gender: male	0.756**	0.0098
	studio (dummy)	1.413**	0.237		marital status: never	-0.050**	8.416e-4
β^{Af}	mortgage	-9.953**	0.212		population	0.0995**	7.570e-4
	employment in accomm.	-0.0064**	2.834e-4		density	-0.408**	0.0216
	wage in accomm.	0.0063**	2.154e-4		single bedroom (dummy)	-1.772**	0.0128
	year fixed effect	1.831**	0.0295		constant	-2.286**	0.0115
	Airbnb history	-2.168**	0.0326				

(b) Second-Stage Decision

		Est.	Std.		Est.	Std.
$\bar{c}^{Av}$		-0.095**	6.676e-4	σ_2	56.768**	0.144
	age: 35-65	26.740**	0.0633	population	-0.964**	0.022
	age: over 65	34.133**	0.194	density	0.106**	0.011
	edu: high school	-151.840**	0.702	mortgage	32.398**	0.067
	edu: bachelor	-82.188**	0.534	employment in accomm.	-4.261**	0.037
	income: 50K-100K	-10.105**	0.050	wage in accomm.	0.0554**	5.386e-4
β^{Av}	income: over 100K	4.060**	0.313	Airbnb regulation score	-0.393**	5.211e-3
	gender: male	-1.088**	0.475	year fixed effect	-15.763**	0.084
	marital status: never	59.658**	0.422	Airbnb history	-1.791**	0.184
	# of bedrooms	6.879**	0.065	summer	13.990**	0.081
	# of bathrooms	3.187**	0.057	fall	23.364**	0.101
	# of amenities	1.445**	0.084	winter	-3.808**	0.142
	entire room (dummy)	32.417**	0.147			

Note: * and ** represent significance at the 5% and 1% levels. The following variables are in logged form: population, density, mortgage, Airbnb history. Employment in the accommodation industry is in percentages, and wages in the accommodation industry are in $10k. The baseline demographics group is age below 35, education below high school, income below 50k, female, and married.

Owner availability type. The probabilities of being the full-available and partial-available types vary by host demographics and metro area. The estimates of β^a suggest that hosts are more likely to be the full-available type if they are senior, have a high school education, have a low income, are male, married and live in cities with a larger population and a lower density. Hosts are more likely to be the partial-available type if they own a single-bedroom property. Note that the owner availability type also affects the baseline Airbnb fixed cost $\bar{c}_\tau^{Af}$, which further determines whether the host chooses Airbnb.

Airbnb fixed cost. The estimates of $\bar{c}_\tau^{Af}$ and β^{Af} suggest that the median Airbnb fixed cost of Airbnb properties is \$646 per month (\$21.5 per day), with a 25th percentile of \$322 per month (\$10.7 per day) and a 75th percentile of \$1660 per month (\$55.3 per day). Note that the average daily price is \$199.8 according to Table 1. The fixed cost can include the psychological cost of embracing the new platform technology and renting out property to transient guests and other tangible costs such as learning how to set up the technology and earn higher profits on Airbnb (e.g., set prices) and preparing property photos and descriptions and furnishings and amenities. The fixed cost can be quite large when, for example, property owners are reluctant to learn the new technology and find it uncomfortable to rent their home to complete strangers, or when they must procure more furnishings and amenities to set up their properties as Airbnb listings. In fact, the psychological cost for the hosts is one of the major obstacles that Airbnb needs to overcome to "convince people to open up their home and allow guests to stay," especially after cases of hosts reporting that their properties were trashed after hosting guests or that they faced safety issues.[24] The learning cost of hosting is another major factor that Airbnb needs to compensate the hosts for, as evident by Airbnb's significant spending on technology and administrative costs associated with the hosts.[25]

We find that the Airbnb fixed cost is higher for properties with more bedrooms and bathrooms and higher for a house than an apartment. The fixed cost is lower and property owners are more likely to choose Airbnb in cities with higher employment and a lower wage in the accommodation industry, and with a longer Airbnb presence. Property owners are also more likely to choose Airbnb in cities where mortgages are high, which might be because property owners leverage Airbnb as an additional income source to pay their mortgages.[26] In fact, the primary use of the hosting income is to pay mortgages according to a survey conducted by Airbnb.[27] Airbnb hosts can even use Airbnb income as proof of worth when applying for mortgage refinancing.[28] Finally, the full-available type of hosts have higher Airbnb fixed costs and are less

[24] See https://www.growthmanifesto.com/airbnb-growth-strategy and https://www.vox.com/2020/2/12/21134477/airbnb-loss-profit-ipo-safety-tech-marketing.

[25] See https://www.vox.com/2020/2/12/21134477/airbnb-loss-profit-ipo-safety-tech-marketing.

[26] The estimates suggest that mortgage pressure reduces the Airbnb fixed cost but increases the Airbnb variable cost. This means that hosts in cities with a high mortgage pressure are more likely to choose Airbnb in the first stage but to list for less time in the second stage. This may be because hosts in these cities are more likely to use Airbnb to pay their mortgage while they are still living in the properties; although they are willing to list, their cost of managing the listing is high.

[27] See https://www.airbnbcitizen.com/the-airbnb-community-in-seattle/

[28] See https://www.cnbc.com/2018/02/22/homeowners-are-using-airbnb-rental-income-to-refinance-mortgages.html

likely to choose Airbnb than the partial-available type of hosts. This may be because these full-available type hosts have long-term rental as their default option and are reluctant to overcome the inertia and adopt the new technology of Airbnb.

6.3 Cannibalization and Market Expansion

Airbnb can create both a negative impact of cannibalization and a positive impact of market expansion on the rental housing market. To evaluate the impact of Airbnb, we use the estimates to simulate the property owners' choices if there was no Airbnb. Intuitively, hosts' decisions can be different with and without Airbnb. Some hosts choose the outside option when Airbnb is not present and choose Airbnb when it becomes available. These hosts represent the market expansion effect of Airbnb: they would not have listed on the long-term rental market and benefit from having Airbnb as an additional income source. Some hosts choose the long-term rental market when Airbnb is not present and choose Airbnb when it becomes available. These hosts are switchers from the long-term rental market and represent the cannibalization effect of Airbnb or the reduction in the long-term rental supply due to Airbnb. Note that the full-available type hosts have the long-term rental option while the partial-available type hosts do not. Therefore, Airbnb hosts of the full-available type can come from both cannibalization and market expansion, whereas Airbnb hosts of the partial-available type only come from market expansion.

Let D^{R0} and D^{R1} denote the number of long-term rental units without and with Airbnb. Let D^A denote the number of Airbnb units when Airbnb is present. Among Airbnb units, the number of switchers or cannibalization units is $D^{R0} - D^{R1}$ and the number of nonswitchers or market expansion units is $D^A - (D^{R0} - D^{R1})$. We use two measures to evaluate Airbnb's impact. The first is the percentage change in rental supply due to Airbnb ($\frac{D^{R1}-D^{R0}}{D^{R0}}$), which captures the negative impact of Airbnb on the long-term rental market. The second is the percentage of Airbnb units that come from cannibalization, or the cannibalization rate, $\frac{D^{R0}-D^{R1}}{D^A}$. This represents the percentage of switchers among all (switchers and nonswitchers) Airbnb units, which captures the relative sizes of the negative and positive impacts of Airbnb. The measures are linked to the cost estimates of our model, as hosts with a high (low) Airbnb hosting cost are more likely to remain in (leave) the long-term rental market when Airbnb is introduced.

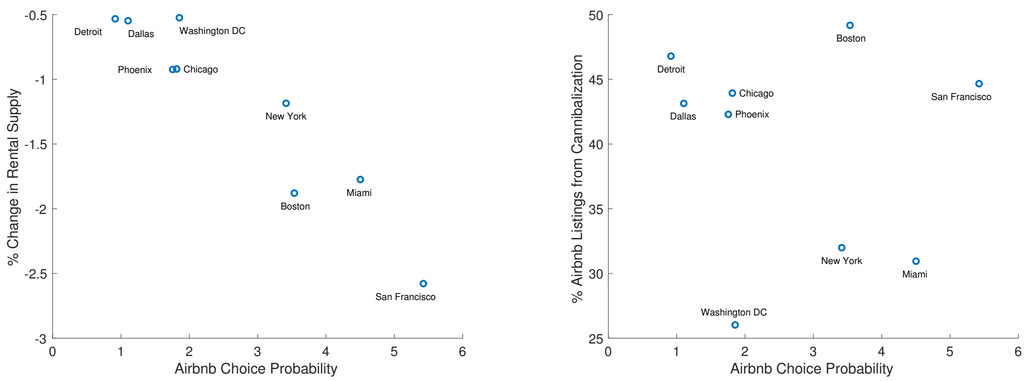
We first plot the percentage change in rental supply across metro areas in Figure 8a. We find that Airbnb causes a mild reduction in the rental supply, ranging from -0.52% in Washington D.C. to -2.58% in San Francisco. The reduction in the rental supply tends to be larger in metro areas where Airbnb is a popular choice for property owners.

However, the percentage change in the rental supply alone does not provide a holistic view of Airbnb's

Figure 8: Cannibalization and Market Expansion: By Metro Area

(a) Change in Rental Supply

(b) Cannibalization Rate



impact. We must also consider the market expansion effect created by Airbnb. We plot the cannibalization rate, or the percentage of switchers, across metro areas in Figure 8b. We find that the percentage of switchers varies significantly, ranging from 26.0% in Washington D.C. to 49.2% in Boston.

Interestingly, although the reduction in the rental supply is greater in metro areas where Airbnb is popular, the cannibalization rate is not necessarily larger in these areas. For example, Miami and New York are among the cities with the highest Airbnb popularity and the largest rental supply reduction; however, their percentages of switchers are among the lowest. This suggests that most of the Airbnb listings in Miami and New York are from market expansion rather than cannibalizing the rental supply. City regulators must thoroughly evaluate both the positive and negative impacts of Airbnb.

Table 6 presents the two measures (the percentage change in the rental supply $\frac{D^{R1}-D^{R0}}{D^{R0}}$ and the percentage of Airbnb units from cannibalization $\frac{D^{R0}-D^{R1}}{D^{A}}$) by property characteristics and demographics. In terms of property characteristics, the reduction in rental supply is largely concentrated among lower priced, affordable units rather than among higher priced, luxurious units. A basic studio or one-bedroom apartment originally on the long-term rental market is more likely to be taken off than a house with multiple bedrooms and more amenities. However, the market expansion effect is also larger for affordable units, leading to a lower cannibalization rate for these units. In terms of demographics, the reduction in rental supply and the cannibalization rate are higher for senior, lower education, medium-income, male, and never married hosts.

Importantly, the results speak to how Airbnb affects housing affordability. We find suggestive evidence that Airbnb does raise affordable housing concerns, as rental supply reduction is larger among affordable units. However, the market expansion effect is also larger for affordable units, as the fraction of nonswitchers is larger among affordable units on Airbnb. This suggests that, interestingly, affordable units are the major sources of both the negative cannibalization impact and the positive market expansion impact of Airbnb.

Table 6: Cannibalization and Market Expansion: By Property Characteristics and Demographics

(a) Property Characteristics

[%]		$\frac{D^{R1}-D^{R0}}{D^{R0}}$	$\frac{D^{R0}-D^{R1}}{D^{A}}$	[%]		$\frac{D^{R1}-D^{R0}}{D^{R0}}$	$\frac{D^{R0}-D^{R1}}{D^{A}}$
# of Bedrooms	0	-4.12	44.09	# of Amenities	1	-1.55	10.42
	1	-1.94	26.88		2	-1.10	33.18
	2	-0.62	51.54		3	-0.98	39.89
	3	-0.79	74.52		4	-1.09	28.13
	4	-1.06	65.64		5	-1.40	35.58
	5+	-2.58	48.23		6	-1.62	43.56
# of Bathrooms	1	-1.49	34.59	Property Type	Apt	-1.23	36.09
	2	-0.81	27.39		House	-1.13	37.80
	3	-0.60	75.48	Listing Type	Entire Place	-1.20	46.68
	4	-0.68	88.21		Private Room	-0.19	0.05

(b) Demographics

[%]		$\frac{D^{R1}-D^{R0}}{D^{R0}}$	$\frac{D^{R0}-D^{R1}}{D^{A}}$	[%]		$\frac{D^{R1}-D^{R0}}{D^{R0}}$	$\frac{D^{R0}-D^{R1}}{D^{A}}$
Age	under 35	-0.71	34.41	Income	under 50K	-1.02	32.21
	35-65	-0.99	29.94		50K-100K	-1.34	47.57
	over 65	-3.16	51.08		over 100K	-1.55	34.98
Education	under	-1.59	50.18	Gender	male	-1.46	45.35
	high school	-1.07	34.15		female	-0.98	29.50
	bachelor's	-0.46	12.01	Marital Status	never married	-1.46	43.91
					other	-1.04	32.18

Although Airbnb harms local renters by reducing affordable rental supply, it also serves as a valuable income source and benefits local hosts who own affordable units and are more likely to be economically disadvantaged. Therefore, policy makers need to trade off between local renters' affordable housing concerns and local hosts' income source needs.

Note that an observed "full-time" ("part-time") listing does not necessarily imply cannibalization (market expansion). In other words, it is not appealing to assume, without modeling the hosts' decisions, that all full-time hosts on Airbnb are switchers and should have been listed on the long-term rental market. Therefore, our structural model framework is helpful in recovering the underlying decision-making process of the hosts and identifying the actual potential switchers. Specifically, cannibalization occurs when property owners switch from long-term rentals to Airbnb. Even if hosts list their properties on Airbnb full time, it would not be cannibalization if they would not have chosen the long-term rental option in the absence of Airbnb. They could have chosen to keep their properties vacant in the absence of Airbnb if their costs (revenues) of long-term rental are high (low). In contrast, part-time listings can be from cannibalization if they would have been in the long-term rental market in the absence of Airbnb. This is possible if the Airbnb profit is large enough to allow property owners to list part time and still earn more than listing in the long-term rental market.

7 Counterfactuals

This section evaluates the impact of a series of policies intended to ensure the supply and affordability of rental housing. First, we consider a set of short-term rental regulations on Airbnb such as imposing taxes and limiting the maximum number of days that a property can be listed. Second, we investigate rent control policy on long-term rental, in particular, how its impact can be affected by the presence of Airbnb.[29]

7.1 Policy Implementation

Short-term rental regulation. We consider three types of short-term rental regulations on Airbnb. The first two regulations impose taxes. Currently, in the U.S., occupancy taxes are levied on facilities that provide transient rental rooms, for example, hotels, motels, and Airbnb. We consider two types of tax: linear and convex. First, the linear tax mimics the current policy, which imposes a fixed percentage of the listing price as an occupancy tax. To account for tax pass-through, let p_{it}^{A} denote the listing price paid by consumers, and let $p_{it}^{A,host}$ denote the price received by hosts. The price paid by consumers p_{it}^{A} enters the hedonic regressions in Equations 11 and 12 while the price received by hosts $p_{it}^{A,host}$ enters the hosts' decisions in Equations 1, 4 and 5. The prices and occupancy rates are determined such that $p_{it}^{A,host} = p_{it}^{A} - t_1 \cdot p_{it}^{A}$ in equilibrium, where t_1 is the tax rate and $0 < t_1 < 1$. Second, we propose a convex tax that charges a higher tax on expensive units and a lower tax on affordable units, which is motivated by our finding that the cannibalization rate is higher for expensive units. Similar to the linear tax, we operationalize it as $p_{it}^{A,host} = p_{it}^{A} - t_2 \cdot (p_{it}^{A})^2$ in equilibrium.

The third short-term rental regulation is a listing restriction, which limits the maximum number of days that a property can be listed on short-term rental platforms. San Francisco, for example, allows hosts to rent out their unit as an entire place for up to 90 days per year. We simulate the case in which hosts are able to list up to a certain number of months in a year. The counterfactual analysis is operationalized as follows: in the second stage, we calculate the optimal number of days to be listed per month. We allow the hosts to choose the months that have the highest expected profits up to the pre-specified maximum number of months. Based on the total ex ante expected profit from the chosen months, they choose among Airbnb, long-term rental, and the outside option in the first stage.

Rent control. Rent control is a system of laws placing a maximum price, or a "rent ceiling," on what

[29] In practice, Airbnb can affect rental housing affordability by changing rental supply (i.e., the number of switchers) and rent, both of which are allowed to be endogenous in our counterfactual analysis. We focus on presenting the changes in rental supply in this section because the changes in rent are found to be very small (less than 1%). This is because the number of Airbnb properties, compared to long-term rental and vacant properties, is still very small in both the data and the counterfactual analysis. Given the current market landscape, Airbnb's impact on long-term rent is limited; Airbnb mainly affects the long-term rental market by reducing rental supply rather than raising rental prices. The impact on rent could become significant if Airbnb accounts for a larger share of the market in the future.

landlords may charge tenants. It covers a spectrum of regulations that can vary from setting the absolute amount of rent that can be charged with no allowed increases to placing different limits on the amount that rent can increase. These restrictions may continue between tenancies or may be applied only within the duration of a tenancy. As of March 2019, the states of California, Maryland, New Jersey, New York, and Oregon, and the city of Washington D.C. have some rent control or stabilization policies on the books while 37 states prohibit or ban rent control outright.[30]

Economists have concluded that rent controls are destructive. According to a 1990 poll of 464 economists, 93% of U.S. respondents agreed, either completely or with provisos, that "a ceiling on rents reduces the quantity and quality of housing available" (Alston, Kearl, and Vaughan 1992). We argue that the negative impact of rent control policy can be exacerbated when another profitable option for hosts, Airbnb, is available. We illustrate how the presence of Airbnb affects the impact of rent control policy by simulating policy outcomes with and without Airbnb. To operationalize the rent control policy, we assume that the rent is capped at $r\%$ below the observed rent where $r\%$ can mimic the type of rent control that limits the maximum percentage of rent increase from the previous year.

7.2 Equilibrium

When implementing the counterfactual policies, it is important to allow the rent, rental occupancy rate, Airbnb price and occupancy rate to endogenously change in the new equilibrium according to the hedonic models outlined in Section 4.3. Specifically, given different counterfactual policies, the number of properties and the types of properties that choose long-term rental and Airbnb can change. The new characteristics and the new aggregate Airbnb and rental supply enter the hedonic models and generate a new set of expectations on rent, rental occupancy rate, Airbnb price and occupancy rate. The equilibrium is defined as a fixed point of the Airbnb price, Airbnb occupancy rate, aggregate Airbnb supply, rent, rental occupancy rate, and aggregate rental supply $\{p^A_{it}, \phi^A_{it}, S^A_{mt}, p^R_{iT}, \phi^R_{iT}, S^R_{mT}\}$. The numerical algorithm to solve for the equilibrium is as follows:

1. Let superscript (k) denote the k-th iteration. Begin with the aggregate Airbnb supply $S^{A(k)}_{mt}$ and aggregate rental supply $S^{R(k)}_{mT}$. Given $S^{A(k)}_{mt}$ and $S^{R(k)}_{mT}$, construct the expected rent $p^{R(k+1)}_{iT}$, rental occupancy rate $\phi^{R(k+1)}_{iT}$, Airbnb price $p^{A(k+1)}_{it}$, and Airbnb occupancy rate $\phi^{A(k+1)}_{it}$ for each property using the hedonic models in Equations 9, 10, 11, and 12.

2. Given the updated $p^{A(k+1)}_{it}$, $\phi^{A(k+1)}_{it}$, $p^{R(k+1)}_{iT}$, and $\phi^{R(k+1)}_{iT}$, solve the property owners' problem under each counterfactual policy. Compute the updated aggregate Airbnb supply $S^{A(k+1)}_{mt}$ and aggregate

[30] See https://www.curbed.com/2019/3/8/18245307/rent-control-oregon-housing-crisis

rental supply $S_{mT}^{R(k+1)}$.

3. Check for the convergence of $\left|p_{it}^{A(k+1)} - p_{it}^{A(k)}\right|$, $\left|\phi_{it}^{A(k+1)} - \phi_{it}^{A(k)}\right|$, $\left|S_{mt}^{A(k+1)} - S_{mt}^{A(k)}\right|$, $\left|p_{iT}^{R(k+1)} - p_{iT}^{R(k)}\right|$, $\left|\phi_{iT}^{R(k+1)} - \phi_{iT}^{R(k)}\right|$, and $\left|S_{mT}^{R(k+1)} - S_{mT}^{R(k)}\right|$. If convergence is not achieved, return to Step 1.

We initialize the algorithm using the observed aggregate Airbnb supply and aggregate rental supply. Varying the initialization point produces robust results.

7.3 Result Analysis

7.3.1 Short-Term Rental Regulations

Overall policy impact. Figure 9a shows the effect of the short-term rental regulations by plotting the number of switchers (cannibalization) on the x-axis and the number of nonswitchers (market expansion) on the y-axis. Each line represents one type of regulation, and each point on the line represents a particular level of regulation. For example, the level of regulation for the maximum month limit varies from 12 months to 3 months, and the level of regulation for the linear tax rate varies from 0% to 90%. Arrow (a) indicates the direction of stricter regulation, for example, a higher tax rate and lower number of months allowed to list. Comparing different levels of regulation within each policy, we find that there is a trade-off in terms of choosing the level of regulation: stricter regulations help reduce the number of switchers (cannibalization); however, they also reduce the number of nonswitchers (market expansion).

A desirable policy should reduce the negative impact of Airbnb (switcher or cannibalization) while maintaining the positive impact of Airbnb (nonswitcher or market expansion). The cannibalization rate is a measure that accounts for both impacts. Therefore, we examine the following measure when comparing policies:

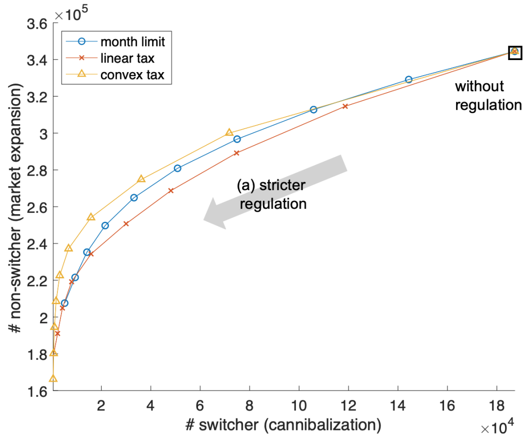
(1) The cannibalization rate, or the fraction of switchers among all (switchers and nonswitchers) listings.

We find that our proposed policy of a convex tax is the most desirable among the three short-term rental regulations. As shown in Figure 9b, the convex tax induces a lower cannibalization rate than the other two policies. The linear tax is the second-best policy, and the month limit is the worst.

Differential impact on hosts. In addition to the overall policy impact, we also examine how the policies differentially affect heterogeneous host groups. In particular, the positive impact of Airbnb is that it benefits the nonswitchers by providing an alternative income source, which is especially valuable for economically disadvantaged hosts with lower priced units. Those nonswitchers would not have listed on the long-term rental market in the absence of Airbnb. Imposing the regulations can induce a redistributive effect among hosts and impact who can continue to benefit from Airbnb, thereby influencing social inequality.

Figure 9: Short-Term Rental Regulations

(a) Number of Switchers and Nonswitchers



(b) Cannibalization Rate

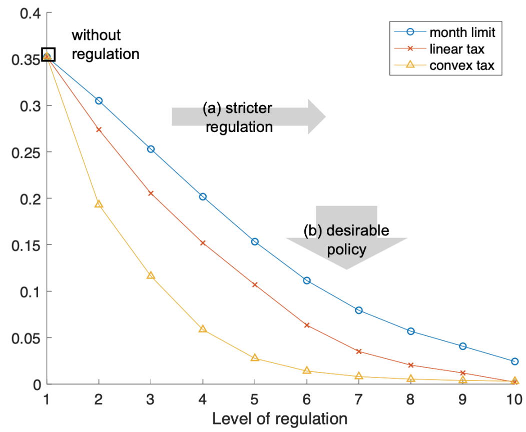
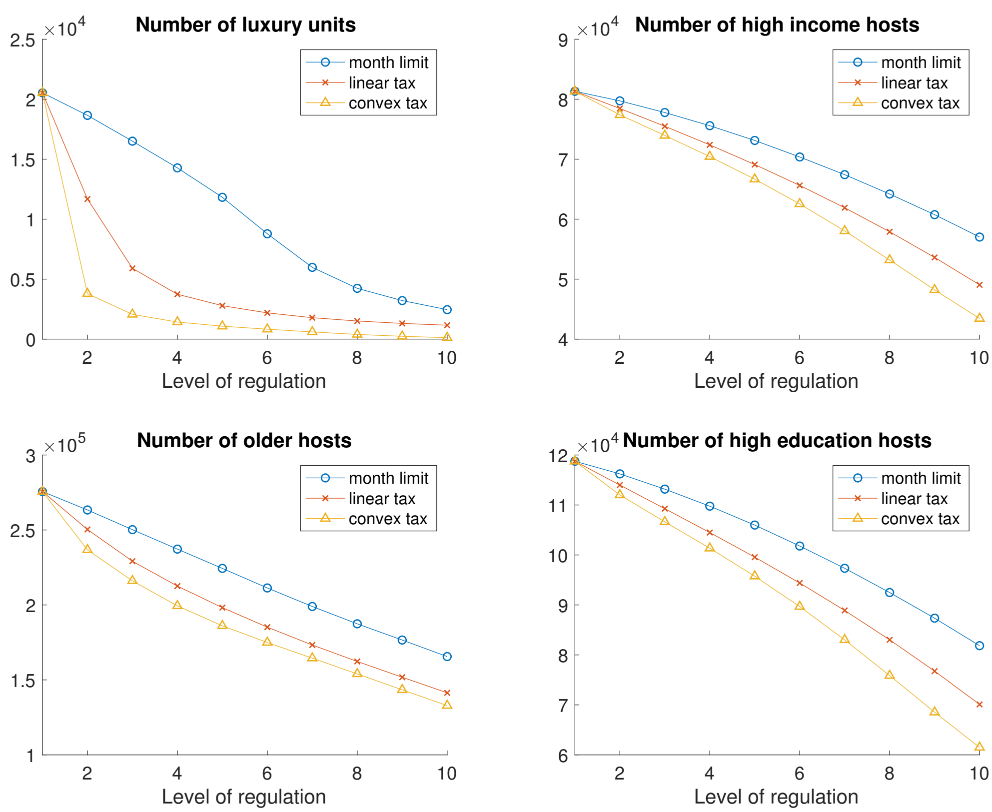


Figure 10: Short-Term Rental Regulations: Differential Impact on Nonswitching Hosts



A desirable policy, in addition to that defined by measure (1), should also prevent economically advantaged hosts who already have abundant resources from using Airbnb as an additional income source, which can exacerbate social inequality. We define two additional measures of policy desirability:

(2) The number of luxury units among nonswitchers;

(3) The number of economically advantaged hosts among nonswitchers.

In Figure 10, we plot the numbers of luxury units (3 bedrooms or above), high-income hosts (income more than 100k), older hosts (age above 35), and high-education hosts (bachelor's degree or higher) among nonswitchers. As the level of regulation increases or the regulation becomes stricter, there are fewer nonswitchers under all policies, which is consistent with our previous finding in Figure 9a. However, in comparing the policies, we find that the convex tax again performs the best in terms of having the smallest number of nonswitchers with luxury units, high income, age, or education. The linear tax again performs the second best, and the month limit performs the worst.

Overall, our proposed policy of a convex tax outperforms the other two policies in terms of all three measures: (1) reducing the cannibalization rate, (2) reducing the number of luxury units among nonswitchers, and (4) reducing the number of economically advantaged hosts among nonswitchers. The linear tax appears to perform better than the month limit. The convex tax performs the best because the percentage of switchers is larger among higher priced, luxury units rather than among lower priced, affordable units. The convex tax discourages higher price properties from being taken off the long-term rental market, which helps limit cannibalization, but has less influence on lower priced properties, which helps maintain market expansion.

7.3.2 Long-Term Rental Regulations: Rent Control

To examine how Airbnb and rent control policy affect each other, we simulate market outcomes under four scenarios: (a) there is no rent control policy, and Airbnb is not available; (b) rent is controlled, and Airbnb is not available; (c) there is no rent control policy, and Airbnb is available; and (d) rent is controlled, and Airbnb is available. Comparing a (c) and b (d) suggests a negative impact of rent control in the absence (presence) of Airbnb. Comparing a (b) and c (d) suggests a negative impact of Airbnb in the absence (presence) of rent control. Importantly, we find that Airbnb and rent control can exacerbate each other's negative impact.

First, we find that the presence of Airbnb can amplify the negative impact of rent control. In Table 7, the first column shows the percentage decrease in the rental supply due to rent control in the absence of Airbnb, and the second column shows the percentage when Airbnb is present. Consistent with the near-consensus among economists discussed above, we find that rent control policy reduces rental supply. Importantly, this

Table 7: Impact of Airbnb on the Negative Effect of Rent Control Policy

Level of Rent Control (r)	% Change in Rental Supply Due to Rent Control without Airbnb	with Airbnb
2.5%	-0.257	-0.299
5.0%	-0.519	-0.602
7.5%	-0.784	-0.909
10.0%	-1.054	-1.220
12.5%	-1.327	-1.533
15.0%	-1.603	-1.848
17.5%	-1.882	-2.166
20.0%	-2.162	-2.485

Table 8: Impact of Rent Control Policy on the Negative Effect of Airbnb

Level of Rent Control (r)	None	2.5%	5.0%	7.5%	10.0%	12.5%	15.0%	17.5%	20.0%
% Change in Rental Supply Due to Airbnb	-1.10	-1.14	-1.18	-1.22	-1.26	-1.30	-1.34	-1.38	-1.42

negative impact of rent control policy is exacerbated when Airbnb is available; the reduction in rental supply due to rent control is larger with than without Airbnb. This exacerbating effect is more prominent when a rent control policy is stricter. This is because Airbnb provides property owners with an alternative option in addition to listing on the long-term rental market. When faced with a rent control policy, more property owners quit long-term rental and switch to Airbnb.

Second, we find that the presence of rent control can also amplify the negative impact of Airbnb. Table 8 shows the percentage decrease in rental supply induced by Airbnb under varying strictness of rent control. The percentage reduction in rental supply due to Airbnb is larger when a rent control policy is in effect and increases as the rent control policy becomes stricter.

Overall, the presence of Airbnb and rent control policy can each have a negative impact on the long-term rental supply. We find that when they are jointly present, they can exacerbate each other's negative impact. Policymakers must exercise caution when implementing rent control policy in the presence of Airbnb.

8 Conclusion

We investigate how Airbnb affects rental supply and affordability and provide policy implications for short-term rental regulations and long-term rent control. We model property owners' decisions in two stages: (1) the yearly decision on the usage of their properties among Airbnb, long-term rental, and neither and (2) the monthly decision on how many days to list on Airbnb if they choose to list on Airbnb in the first stage. Given the revenue data on rent, rental occupancy rate, Airbnb price, and Airbnb occupancy rate, we estimate the hosting costs of property owners.

We find that Airbnb mildly cannibalizes the rental market but has a market expansion effect. The percentage of switchers varies significantly across cities. The rental supply reduction is larger for lower priced, affordable units than for higher priced, luxurious units, suggesting that Airbnb can raise concerns about housing affordability. However, the market expansion effect is also larger for affordable units, suggesting that owners of affordable units benefit more from having Airbnb as an income source. Metro areas where Airbnb is popular (e.g., San Francisco, New York, and Miami) experience a larger reduction in long-term rental supply due to Airbnb; however, some of them benefit more from a larger market expansion effect, suggesting that the percentage of switchers is not necessarily larger in those cities.

The counterfactual results suggest that short-term rental regulations help reduce cannibalization; however, they also reduce market expansion. We assess commonly used regulations such as limiting the number of days that a property can be listed and a linear tax and propose a new convex tax that charges a higher tax on expensive units. We show that the proposed convex tax outperforms the linear tax, which further outperforms the day limit according to three measures of policy desirability: (1) reducing the cannibalization rate, (2) preventing luxury units from leveraging Airbnb as an additional income source, and (3) preventing economically advantaged hosts (e.g., high-income, older, or high-education hosts) who may already have abundant resources from leveraging Airbnb, which can exacerbate social inequality. Finally, rent control must be implemented with greater caution when Airbnb is available, as lower profits from long-term rentals can cause landlords to switch to Airbnb and exacerbate the side effect of a rent control policy.

There are a few limitations of this study that represent directions for future research. First, we assume away the case in which long-term rental tenants sublet on Airbnb because we cannot observe whether an Airbnb host is a tenant or a property owner. This type of case may be relatively rare, as lease agreements often include clauses that restrain sublets. Services such as SubletAlert.com and SubletSpy also help landlords find tenants who have violated the agreement. Future research may extend the proposed model to incorporate cases in which tenants sublet on Airbnb if such data are available.

Second, we do not explicitly model the competition between hotels and Airbnb. The hedonic models of the Airbnb price and occupancy rate are estimated conditional on the observed competitive landscape between hotels and Airbnb. The implicit assumption is that hotels in the counterfactual analysis follow the same strategy as they do in the observed scenario. The equilibrium we solve for can be regarded as a partial equilibrium without hotel responses. We believe that the trade-off between long-term rental and Airbnb is the first-order effect for property owners who are the key players in studying the impact of Airbnb on the long-term rental market. In addition, hotels do not appear to have responded to Airbnb in practice according to Li and Srinivasan (2019) who study hotel and Airbnb competition. In the future, when hotels have systematically responded to Airbnb, researchers can incorporate hotel responses into our framework.

References

[1] Alston, R. M., Kearl, J. R., & Vaughan, M. B. (1992). Is There a Consensus Among Economists in the 1990's? *American Economic Review*, 82(2), 203-209.

[2] Barron, K., Kung, E., & Proserpio, D. (2019). The Sharing Economy and Housing Affordability: Evidence from Airbnb. Forthcoming, *Marketing Science.*

[3] Berger, T., Chen, C., & Frey, C. B. (2018) Drivers of Disruption? Estimating the Uber Effect. *European Economic Review*, 110, 197-210.

[4] Burtch, G., Carnahan, S., & Greenwood, B. N. (2018). Can You Gig It? An Empirical Examination of the Gig Economy and Entrepreneurial Activity. *Management Science*, 64(12), 5461-5959.

[5] Dowling, K., Manchanda, P. and Spann, M. (2019) The Existence and Persistence of the Pay-Per-Use Bias in Car Sharing Services. Available at SSRN: https://ssrn.com/abstract=3204233.

[6] Einav, L., Farronato, C., & Levin, J. (2016). Peer-to-Peer Markets. *Annual Review of Economics*, 8(1), 615-635.

[7] Gong, J., Greenwood, B. N., & Song, Y. (2017). Uber Might Buy Me a Mercedes Benz: An Empirical Investigation of the Sharing Economy and Durable Goods Purchase. Available at SSRN: https://ssrn.com/abstract=2971072.

[8] Greenwood, B. & Wattal, S. (2017). Show Me the Way to Go Home: An Empirical Investigation of Ride-sharing and Alcohol Related Motor Vehicle Fatalities. *MIS Quarterly*, 41(1), 163-187.

[9] Gurran, N. & Phibbs, P. (2017). When Tourists Move In: How Should Urban Planners Respond to Airbnb? *Journal of the American Planning Association*, 83(1), 80-92.

[10] Horn, K. & Merante, M. (2017). Is Home Sharing Driving Up Rents? Evidence From Airbnb in Boston. *Journal of Housing Economics*, 38, 14-24.

[11] Jiang, B. & Tian, L. (2018). Collaborative Consumption: Strategic and Economic Implications of Product Sharing. *Management Science*, 64(3), 983-1476.

[12] Koster, H., van Ommeren, J. and Volkhausen, N., (2019) Short-Term Rentals and the Housing Market: Quasi-Experimental Evidence From Airbnb in Los Angeles. CEPR Discussion Paper No. DP13094. Available at SSRN: https://ssrn.com/abstract=3226869

[13] Lee, D. (2016). How Airbnb Short-Term Rentals Exacerbate Los Angeles's Affordable Housing Crisis: Analysis and Policy Recommendations. *Harvard Law and Policy Review*, 10(1), 229-253.

[14] Li, J., Moreno, A., & Zhang, D. (2016). Pros vs Joes: Agent Pricing Behavior in the Sharing Economy (August 28, 2016). Ross School of Business Paper No. 1298. Available at SSRN: https://ssrn.com/abstract=2708279.

[15] Li, H. & Srinivasan, K. (2019). Competitive Dynamics in the Sharing Economy: An Analysis in the Context of Airbnb and Hotels. *Marketing Science*, 38(3), 2019; 365-391

[16] Murphy, K. and R. Topel (1985). Estimation and Inference in Two Step Econometric Models. *Journal of Business and Economic Statistics* 3, 370–9.

[17] Sheppard, S. & Udell, A. (2018). Do Airbnb Properties Affect House Prices? Working paper.

[18] Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, 13(1), 1-44.

[19] Tian, L. & Jiang, B. (2017). Effects of Consumer-to-Consumer Product Sharing on Distribution Channel. P*roduction and Operations Management*, 27(2), 350-367.

[20] Zervas, G., Proserpio, D., & Byers, J. W. (2017) The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. *Journal of Marketing Research*, 54(5), 687-705.

[21] Zhang, S., Mehta, N., Singh, P., & Srinivasan, K. (2018). Do Lower-Quality Images Lead to Greater Demand on Airbnb? Working paper.

Appendix

A. Restricting the Number of Days to be Integers

Let $u(s) \equiv \frac{p_{it}^{A}\phi_{it}^{A}}{\bar{s}(\exp(s/\bar{s})-\exp((s-1)/\bar{s}))} - (\bar{c}^{Av} + \beta^{Av}X_{it}^{Av} + \xi_{mT}^{Av})$ and $l(s) \equiv \frac{p_{it}^{A}\phi_{it}^{A}}{\bar{s}(\exp((s+1)/\bar{s})-\exp(s/\bar{s}))} - (\bar{c}^{Av} + \beta^{Av}X_{it}^{Av} + \xi_{mT}^{Av})$. Taking into account the fact that the number of days to list the property on Airbnb is integer, the second-stage probabilities are

$$
\begin{aligned}
\Pr(s_{it}=0) &= \Pr\left(\Pi_{it}^{A}(0) > \Pi_{it}^{A}(1)\right) = \Pr\left(\epsilon_{it}^{Av} > l(0)\right) = 1 - \Phi\left(\frac{l(0)}{\sigma_2}\right) \\
\Pr(s_{it}=s\ (s=1,2,...,\bar{s})) &= \Pr\left(\Pi_{it}^{A}(s) > \Pi_{it}^{A}(s-1) \text{ and } \Pi_{it}^{A}(s) > \Pi_{it}^{A}(s+1)\right) \\
&= \Pr\left(l(s) < \epsilon_{it}^{Av} < u(s)\right) = \Phi\left(\frac{u(s)}{\sigma_2}\right) - \Phi\left(\frac{l(s)}{\sigma_2}\right) \\
\Pr(s_{it}=\bar{s}) &= \Pr\left(\Pi_{it}^{A}(\bar{s}) > \Pi_{it}^{A}(\bar{s}-1)\right) = \Pr\left(\epsilon_{it}^{Av} < u(\bar{s})\right) = \Phi\left(\frac{u(\bar{s})}{\sigma_2}\right)
\end{aligned}
$$

Given the optimal number of days to list the property on Airbnb, the ex ante monthly profit from Airbnb hosting is

$$
E\left[\Pi_{it}^{A}(s_{it}^{*})\right] = \left[\int_{-\infty}^{\infty} \Pi_{it}^{A}\left(s_{it}^{*}\right) f\left(\epsilon_{it}^{Av}\right) d\epsilon_{it}^{Av}\right]
$$

where the integral is expanded as

$$
\begin{aligned}
\int_{-\infty}^{\infty} \Pi_{it}^{A}\left(s_{it}^{*}\right) f\left(\epsilon_{it}^{Av}\right) d\epsilon_{it}^{Av} &= \int_{l(0)}^{\infty} \Pi_{it}^{A}(0) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&+ \int_{l(1)}^{u(1)} \Pi_{it}^{A}(1) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&+ ... \\
&+ \int_{l(\bar{s}-1)}^{u(\bar{s}-1)} \Pi_{it}^{A}((\bar{s}-1)) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&+ \int_{-\infty}^{u(\bar{s})} \Pi_{it}^{A}(\bar{s}) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av}
\end{aligned}
$$

Here, the first term for the interval with $s_{it}^{*}=0$ is zero:

$$
\int_{l(0)}^{\infty} \Pi^{A}(0) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} = \int_{l(0)}^{\infty} 0 \cdot f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} = 0
$$

The terms for the intervals with $s_{it}^{*}=s\ (s=1,2,...,\bar{s})$ is computed as:

$$
\begin{aligned}
&\int_{l(s)}^{u(s)} \left[p_{it}^A \phi_{it}^A s - c_{it}^{Av} \bar{s} \left(\exp\left(\frac{s}{\bar{s}}\right) - 1\right)\right] f(\epsilon_{it}^{Av}) d\epsilon_{it} \\
&= \left[p_{it}^A \phi_{it}^A s - \bar{s}\left(\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av}\right)\left(\exp\left(\frac{s}{\bar{s}}\right) - 1\right)\right] \int_{l(s)}^{u(s)} f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&\quad - \left[\bar{s}\left(\exp\left(\frac{s}{\bar{s}}\right) - 1\right)\right] \int_{l(s)}^{u(s)} \epsilon_{it}^{Av} f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&= \left[p_{it}^A \phi_{it}^A s - \bar{s}\left(\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av}\right)\left(\exp\left(\frac{s}{\bar{s}}\right) - 1\right)\right] \left[\Phi\left(\frac{u(s)}{\sigma_2}\right) - \Phi\left(\frac{l(s)}{\sigma_2}\right)\right] \\
&\quad - \left[\bar{s}\left(\exp\left(\frac{s}{\bar{s}}\right) - 1\right)\right] \left[-\frac{\sigma_2}{\sqrt{2\pi}}\left(\exp\left(-\frac{u(s)^2}{2\sigma_2^2}\right) - \exp\left(-\frac{l(s)^2}{2\sigma_2^2}\right)\right)\right]
\end{aligned}
$$

For the last term for the interval with $s^* = \bar{s}$, recall that $\Pi_{it}^A(\bar{s})$ is bounded by the maximum possible profit, $p_{it}^A \bar{s}$.

$$
\begin{aligned}
&\int_{-\infty}^{u(\bar{s})} \Pi_{it}^A(\bar{s}) f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&= \int_{-\infty}^{\frac{p_{it}^A(\phi_{it}^A - 1)}{\exp(1)-1} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})} \left[p^A \bar{s}\right] f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av} \\
&\quad + \int_{\frac{p_{it}^A(\phi_{it}^A - 1)}{\exp(1)-1} - (\bar{c}^{Av} + \beta^{Av} X_{it}^{Av} + \xi_{mT}^{Av})}^{u(\bar{s})} \left[p_{it}^A \phi_{it}^A \bar{s} - c_{it}^{Av} \bar{s}\left(\exp(1) - 1\right)\right] f(\epsilon_{it}^{Av}) d\epsilon_{it}^{Av}
\end{aligned}
$$

where the integrals are computed similarly as in the other terms for the intervals with $s_{it}^* = s$ $(s = 1, 2, ..., \bar{s})$.

B. Hedonic Regression Results

Table 9 shows the hedonic regression results for rent and rental occupancy rate. Table 10 shows the hedonic regression results for Airbnb price and occupancy rate. The R-squared value for Airbnb occupancy rate is relatively low, which may be due to a large variation in occupancy rate over time even within the same property. In fact, Airbnb occupancy rate seems to be quite random at the individual property level. Analysis of within- and across-property variation shows a large within-property variation across months, and including market-specific month fixed effects in the regression does not explain the large within-property variation. However, the model-predicted Airbnb occupancy rate is consistent with average occupancy rate for each property. In other words, hosts are on average correct in predicting their occupancy rates, which is more important when they make the first-stage decisions.

Table 9: Hedonic Regression: Rent and Rental Occupancy

DV:	Rent		Rental Occupancy	
Constant	1,532***	(100.1)	0.340***	(0.0390)
Rental Supply	-171.6***	(39.22)	-0.0172***	(0.00640)
Rent	–		-9.48e-05***	(2.41e-05)
Metro Area - Year FE	Yes		Yes	
Demographics				
Age				
35-65	-76.36***	(22.40)	-0.0199***	(0.00334)
Over 65	-47.84	(36.05)	0.0509***	(0.00464)
Education				
Bachelor's	-25.32	(27.98)	0.131***	(0.00354)
High School Grad	338.6***	(32.67)	0.164***	(0.00912)
Marital Status				
Never Married	1.706	(26.04)	0.0664***	(0.00325)
Married Now	1.462	(26.12)	0.0559***	(0.00326)
Gender				
Male	28.26	(19.56)	0.0264***	(0.00253)
Race				
Asian	-256.1***	(38.59)	0.554***	(0.00783)
Black	-602.9***	(39.68)	0.545***	(0.0154)
White	-533.8***	(48.75)	0.534***	(0.0142)
Origin				
Hispanic	-232.9***	(24.32)	0.0457***	(0.00638)
Household Income				
50K-100K	31.04	(23.31)	0.0148***	(0.00300)
Over 100K	439.9***	(29.93)	0.0477***	(0.0112)
Property Characteristics				
# of Bedrooms				
1	290.0***	(77.86)	0.0330***	(0.0120)
2	460.4***	(85.21)	0.0541***	(0.0154)
3	573.5***	(82.71)	0.0810***	(0.0173)
4	521.6***	(100.3)	0.0779***	(0.0177)
5+	919.7***	(133.4)	0.120***	(0.0277)
# of Bathrooms	99.65***	(9.550)	0.00633**	(0.00268)
# of Rooms	7.446	(14.21)	-0.00222	(0.00178)
# of Amenities	25.76***	(8.378)	0.00609***	(0.00122)
Property Type				
House	-325.1***	(29.84)	-0.0259***	(0.00868)
Other	-632.8***	(119.1)	-0.0244	(0.0213)
# of Units in the Structure				
2	-291.1***	(37.85)	-0.0225***	(0.00846)
3-4	-263.0***	(34.25)	-0.0180**	(0.00765)
5-9	-193.3***	(31.75)	-0.00814	(0.00611)
10+	-266.2***	(31.52)	-0.0174**	(0.00753)
Unit Age	-12.69***	(1.428)	-4.79e-05	(0.000354)
Unit Age Squared	0.0910***	(0.0126)	-1.28e-06	(2.70e-06)
N	15,670		15,670	
R^2	0.2217		0.6641	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

Table 10: Hedonic Regression: Airbnb Price and Occupancy Rate

DV:	Logged Airbnb Price		Airbnb Occupancy	
Constant	3.911***	(0.0169)	0.799***	(0.0489)
Airbnb Supply	-0.0702***	(0.00176)	0.0166***	(0.00143)
Logged Airbnb Price	–		-0.100***	(0.0122)
Metro Area - Year FE	Yes		Yes	
Metro Area - Month FE	Yes		Yes	
Demographics				
Age	0.0142***	(9.71e-05)	0.00145***	(0.000184)
Household Income	0.00199***	(1.37e-05)	5.04e-05*	(2.58e-05)
Education				
Bachelor's	0.0271***	(0.00676)	-0.213***	(0.00442)
High School Grad	-0.349***	(0.00869)	-0.449***	(0.00708)
Marital Status				
Never Married	-0.977***	(0.0116)	-0.106***	(0.0141)
Married Now	0.297***	(0.0125)	0.206***	(0.00890)
Gender				
Male	0.225***	(0.00739)	-0.0305***	(0.00554)
Race				
Asian	0.230***	(0.00644)	0.0104**	(0.00505)
Black	-0.179***	(0.00697)	0.0132***	(0.00504)
White	0.113***	(0.00737)	-0.0215***	(0.00499)
Origin				
Hispanic	-0.0947***	(0.00335)	0.00802***	(0.00247)
Property Characteristics				
# of Bedrooms				
1	0.151***	(0.00156)	-0.0301***	(0.00210)
2	0.402***	(0.00140)	-0.0116**	(0.00498)
3	0.672***	(0.00172)	0.00384	(0.00826)
4	0.957***	(0.00226)	0.0217*	(0.0117)
5+	1.163***	(0.00292)	0.0444***	(0.0143)
# of Bathrooms	0.170***	(0.000542)	-0.0193***	(0.00210)
# of Amenities	0.0101***	(0.000201)	0.0228***	(0.000180)
Property Type				
House	0.00910***	(0.000908)	0.0247***	(0.000602)
Other	0.0618***	(0.000814)	0.0594***	(0.000921)
Room Type				
Private/Shared	-0.604***	(0.000786)	-0.124***	(0.00737)
Airbnb-related metro variables				
Airbnb history	0.00226***	(0.000381)	0.00587***	(0.000250)
Air passengers (in millions)	0.00480	(0.00640)	0.0247***	(0.00417)
N	3,219,447		3,219,447	
R^2	0.5685		0.1182	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.