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

Hui Li, Yijin Kim, Kannan Srinivasan*

September 8, 2021

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 leading 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 how many 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 the long-term rental supply. Cities where Airbnb is more popular experience a larger rental supply reduction, but they do not necessarily have a larger percentage of switchers. Affordable units are the major sources of both the negative and positive impacts of Airbnb: they cause a larger rental supply reduction, which harms local renters; they also create a larger market expansion effect, which benefits local hosts who own affordable units and may be less economically advantaged. Policy makers need to strike a balance between local renters' affordable housing concerns and local hosts' income source needs. We also find 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 that imposes a higher tax on expensive units and show that it can outperform 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 Associate Professor of Marketing, Tepper School of Business, Carnegie Mellon University. Email: huil1@andrew.cmu.edu. Yijin Kim, consultant, LG CNS. Email: yijink@alumni.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 example, 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 because the yields can be higher with Airbnb than in the long-term rental market.[2] Such switching behavior could impact rental housing supply and affordability. Motivated by these concerns, city regulators launched various policies on short-term rentals, especially in cities where affordable housing has been an issue. For example, the City of Los Angeles approved new rules for Airbnb-type rentals in December 2018, following more than 3.5 years of debate since the law was first proposed.[3] Similarly, San Francisco saw 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] There are two prevailing types of regulations. The first type limits 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 120 days in Los Angeles).[5] The second type charges a transient occupancy tax on the listing price (e.g., 8.5% in Philadelphia and 14% in Los Angeles), which is similar to a hotel occupancy tax.[6] By 2020, many cities had imposed similar regulations on Airbnb. However, most of these policies were launched without empirical evidence. It remains unclear how Airbnb has affected 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

[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/. In addition, see https://www.forbes.com/sites/garybarker/2020/02/21/the-airbnb-effect-on-housing-and-rent/?sh=3d580ee32226.

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

[4] In a 2015 ballot measure in San Francisco, 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/.

[5] See https://www.airbnb.com/help/article/864/los-angeles-ca#nightlimits.

[6] See https://www.airbnb.com/help/article/2509/in-what-areas-is-occupancy-tax-collection-and-remittance-by-airbnb-available.

on rental supply) and the types of properties that are taken off the rental market (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. 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 rental supply reduction is larger in metro areas where Airbnb is popular, the percentage of switchers is not necessarily larger in those 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. Policy makers must take a holistic view when evaluating Airbnb's impact.

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 the highest among affordable units. However, the market expansion effect is also the largest for affordable units, as the fraction of non-switchers is the largest for 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;

these hosts are likely to be less economically advantaged than hosts who own expensive units and benefit more from additional income sources. Therefore, policy makers need to trade off between local renters' affordable housing concerns and local hosts' income source needs.

In the counterfactual analysis, we evaluate two sets of policies related to the supply and affordability of rental housing. The first set of counterfactuals is motivated by recent regulations on short-term rentals. Policy makers 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, 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 (non-switchers 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 fraction of total host profits earned by owners of luxury units; and (3) the ability to reduce the fraction of total host profits earned by economically advantaged hosts (e.g., high-income, older, or high-education hosts). The second and third measures relate to social inequality because they capture potential differential policy impacts on heterogeneous hosts. In particular, Airbnb provides the hosts an additional income source; imposing regulations can induce a redistributive effect among Airbnb hosts and affect income equality. A desirable policy should prevent the distribution of income among hosts from being skewed to those economically advantaged hosts who own expensive units and already have abundant resources. 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 focuses on rent control policy, which limits rent in the long-term rental market. Economists are virtually unanimous in concluding that rent controls are destructive because they reduce the available housing supply. 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).[7] We show that the negative effect of rent control policies is aggravated when Airbnb is available because Airbnb serves as an additional profitable option for property owners and can further motivate them to switch away from the long-term rental market.

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

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. For policy makers, assessing the impact of Airbnb is difficult, as 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 literature by addressing important policy-driven questions regarding whether and how Airbnb has impacted the rental housing market. There have been continuous concerns that hosts on the long-term rental market may switch to Airbnb, causing a reduction in rental housing supply and threatening housing affordability. These concerns motivated many cities to impose various regulations on Airbnb, including charging a transient occupancy tax and limiting the number of days that a property can be listed. However, most of these policies were launched without empirical evidence. The existing literature on the potential switching behavior of the hosts is scant. The effectiveness of current short-term rental regulations also remains unclear.

Indeed, it is challenging to answer the questions of whether there are switchers and what types of properties are switching. It is difficult to gather a comprehensive data set on both long-term and short-term rental hosts. Even if the data are collected, one cannot identify the actual potential switchers by directly examining the observed hosts' decisions; a structural model on the hosts' decisions is required. Specifically, 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 were 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, researchers must systematically model the hosts' revenue-cost trade-offs to identify switchers. A structural model allows researchers to simulate the counterfactual scenario without Airbnb and compare it with the scenario when Airbnb is present, which are required to draw conclusions about the actual switchers.

Furthermore, a structural model is required to evaluate the effectiveness of existing rental regulations. To simulate hosting behaviors under various regulations in counterfactual scenarios, one needs to model the individual hosts' decisions and recover their underlying trade-offs. In addition, the prices, occupancy rates, and supply of housing units in the counterfactuals can differ from observed ones. We need to allow them to endogenously change and solve for new equilibrium outcomes. However, existing studies mostly provide descriptive analysis or conduct regression analysis, which do not allow for counterfactual analyses to evaluate policy impacts.

We fill the gap and contribute to the literature on Airbnb and rental housing market by building a structural model of hosting behaviors. To the best of our knowledge, we are the first to systematically and formally model hosts' decisions and to recover the underlying trade-offs. This framework allows us to conduct counterfactual analysis to identify actual switchers and examine policy impacts.

Existing studies primarily focus on Airbnb's impact on housing and rental prices in a particular city using descriptive or regression analysis. For example, Lee (2016) and Gurran and Phibbs (2017) provide descriptive analyses of Airbnb and the rental housing market in Los Angeles and Sydney, Australia, respectively. 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. Marketing researchers have also recently contributed to this topic. Barron, Kung, and Proserpio (2021) use a comprehensive data set 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. These studies focus on how housing and rental prices changed after the introduction of Airbnb and do not directly address the switching behavior, the cause of changes in housing and rental prices. We formally model the hosts' supply choices. The structural model allows us to identify actual switchers by simulating the counterfactual scenario without Airbnb. We find that Airbnb mildly reduces long-term rental supply, which is consistent with findings in the literature that rental prices increased due to Airbnb. We also obtain a novel set of findings on switching behaviors across heterogeneous hosts. In particular, we find that affordable units experience a larger reduction in rental supply but have a lower fraction of switchers. These results have not been documented before and have strong policy implications for affordable housing.

In terms of the effectiveness of existing policies, existing literature on the topic is rare and adopts regression analysis. Koster et al. (2019) apply a panel regression-discontinuity design and find that the adoption of home sharing ordinances reduced housing prices by 3% and rents by 3% in Los Angeles. We use structural models to analyze individual hosts' decisions and simulate their behaviors under various counterfactual policy scenarios. We leverage data on a variety of cities and show that there is significant heterogeneity across cities, which is helpful for localized policy making. For example, we find that cities where Airbnb is more popular experience a larger reduction in the rental supply; however, these cities do not necessarily have a larger percentage of switchers. We also show that imposing a linear tax is more desirable than limiting the number of days a property can be listed.

More broadly, our paper contributes to the literature on how the sharing economy affects traditional industries and incumbent firms. For example, ride-sharing services 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 Airbnb hosts' pricing decisions 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.[8] 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. It consists of a sample of representative properties that are scientifically selected to represent all housing units. 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.[9] Each observation also includes a sampling weight, which is designed to extrapolate the sample to the full population of housing units. We use these sampling weights to present summary statistics and conduct analyses throughout the paper. 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 to 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 kitchen appliances, 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.[10] 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 properties that are potentially available

[8]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).

[9]The survey asks ex post questions about the use of the property "in the past 12 months", so the tenure information reflects the actual usage of the property.

[10]For example, 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.

for listing in the selected area.[11] A property is listed either on Airbnb (units in the Airbnb data set), the long-term rental market (rented units in the AHS data set), or neither (vacant units in the AHS data set). Hereafter, we refer to keeping the property vacant as the outside option. We distinguish between two types of units that choose the outside option, "vacant full year" and "vacant partial 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 metro-area-level population, density, employment and wage information 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 the rent-to-own ratio, unemployment rate, number of air passengers to the city, Airbnb regulation score, Airbnb history, and Google search index for Airbnb. The 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.[12] Airbnb regulation score, which measures how friendly city policies are to short-term rentals, is published by the R Street Institute.[13] 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

[11] Note that the properties in the Airbnb data set can overlap with the properties in the AHS data set. We cannot perfectly distinguish whether an AHS property is listed on Airbnb because the observed characteristics do not allow us to perfectly link a property in the Airbnb data set to a property in AHS; the AHS data set intentionally removes information that allows for identification of a property. Nevertheless, we use information in AHS that can potentially indicate listing on Airbnb and remove those overlapping units so that there is no double-counting issue. Specifically, Airbnb has two types of properties: private room listings (i.e., owners live with the guests) and entire place listings (i.e., owners do not live in the properties). First, the private room Airbnb listings may overlap with owner-occupied AHS units. However, double-counting is not a concern here because owner-occupied AHS units are not included in our analysis. Second, the entire place Airbnb listings may overlap with vacant units in the AHS data set because AHS classifies housing units as vacant if they are unoccupied or occupied by anyone who is not the usual resident (such as an Airbnb guest). To remove this type of overlapping unit, we use a categorical variable in AHS that indicates how many nights a vacant property was rented out in the past year. There are three categories, namely, "0 to 2 nights", "3 to 7 nights", and "8 or more nights", each accounting for 91.1%, 0.23%, and 8.67% of the vacant properties, respectively. Intuitively, those rented for "3 to 7 nights" and "8 or more nights" are possibly rented on Airbnb and overlap with the entire place Airbnb listings. We also find that the number of such units in AHS is comparable to the number of entire place listings in the Airbnb data set. Therefore, we remove those vacant units that were rented for "3 to 7 nights" and "8 or more nights" from the AHS data set and combine the remaining vacant and rented AHS units with the Airbnb data set. The combined data set no longer contains overlapping properties and is used for our analysis. Besides removing overlapping units, the approach of identifying overlapping units also allows us to identify Airbnb listings in the AHS data set and present switching patterns among long-term rental, Airbnb, and the outside option within the AHS data set. We present the details in the online appendix.

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

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

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

	Airbnb	Long-term Rental	Vacant Full Year	Vacant Partial Year
2015: Number of Units	155,725	8,481,448	685,332	251,240
2015: Proportion (%)	1.63	88.59	7.16	2.62
2017: Number of Units	221,730	8,409,056	603,665	233,810
2017: Proportion (%)	2.34	88.81	6.38	2.47
Number of Bedrooms	1.31 (0.87)	1.97 (0.96)	2.40 (1.06)	2.31 (1.04)
Number of Bathrooms	1.26 (0.59)	1.68 (1.17)	2.37 (1.35)	2.66 (1.31)
Apartment (%)	74.68 (43.48)	72.15 (44.83)	42.09 (49.39)	43.93 (49.68)
Kitchen (%)	90.77 (28.94)	99.14 (9.22)	98.65 (11.56)	96.66 (17.97)
Air Conditioning (%)	80.48 (39.63)	87.74 (32.79)	69.40 (46.10)	84.98 (35.76)
Heating (%)	87.22 (33.37)	99.37 (7.93)	97.88 (14.43)	96.08 (19.43)
Washer (%)	61.56 (48.65)	46.50 (49.88)	47.63 (49.96)	63.34 (48.24)
Dryer (%)	59.63 (49.06)	43.83 (49.62)	46.87 (49.92)	62.13 (48.55)
Fireplace (%)	10.91 (31.18)	11.20 (31.54)	19.08 (39.31)	13.75 (34.48)
Parking Space (%)	40.56 (49.10)	32.14 (46.70)	46.73 (49.91)	50.72 (50.04)
Private or Shared Room (%)	42.73 (49.47)			
Airbnb Daily Price ($)	148.72 (102.76)			
Airbnb Occupancy Rate (%)	31.20 (37.56)			
Monthly Rent ($)		1,263.74 (593.36)		

Note: Standard deviations are shown in parentheses. The numbers of long-term rental and vacant units in the table are extrapolated using sampling weights in the AHS data set. Similarly, the characteristics of the long-term rental and vacant units in the table are weighted averages using the sampling weights in the AHS data set.

reports and articles.[14] Lastly, Google search index for Airbnb is downloaded from Google and measures the number of Google searches for the term "airbnb" in a particular year and month. We normalize it to have a value of 100 at the peak month during the sample period.

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 partial year) units by year and the summary statistics of properties choosing each option. In 2015, 1.63% of the property owners chose Airbnb and 88.59% chose long-term rental; 7.16% of properties were vacant for the full year, and 2.62% were vacant for part of the year. These numbers changed to 2.34%, 88.81%, 6.38%, and 2.47%, respectively, in 2017 as the number of Airbnb properties increased by nearly 50% from 2015 to 2017. We find that the Airbnb units have comparable property characteristics to the long-term

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

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 may come from the same pool.

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 only part of the month. This is also supported by the histogram of the number of days 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, 51.7% 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, such as when they are away on vacation. By contrast, some properties are listed most of the time. The data show that 33.5% of the observations are listed for more than 70% of the year, 26.2% are listed for more than 80% of the year, and 18.0% 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 0.26% in Detroit to 3.46% in San Francisco. The top three metro areas with the highest proportion of Airbnb properties are San Francisco, New York, and Miami. Second, the percentage of units choosing each option differs by property characteristics. 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

Figure 1: Representative Airbnb Listing Patterns

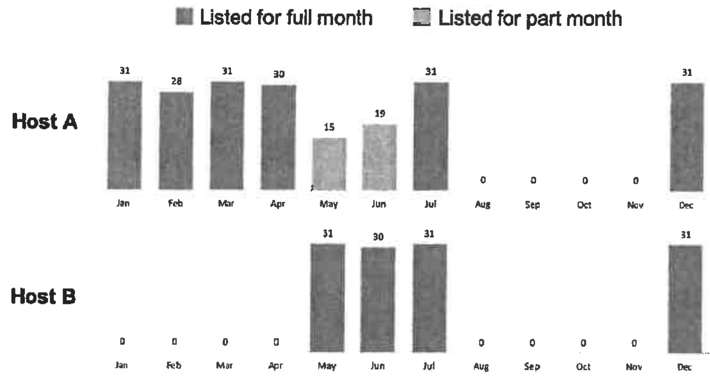


Figure 2: Histogram of Monthly Number of Days Listed

(a) 2015 (b) 2017

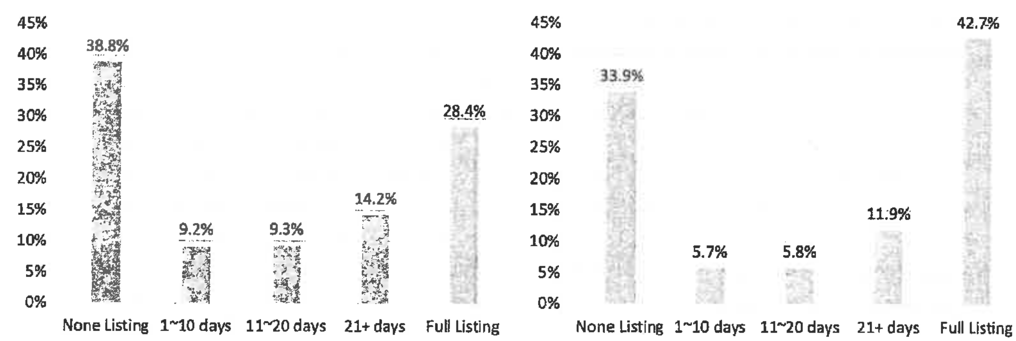


Figure 3: Histogram of Percentage of Days Listed in Each Year

(a) 2015 (b) 2017

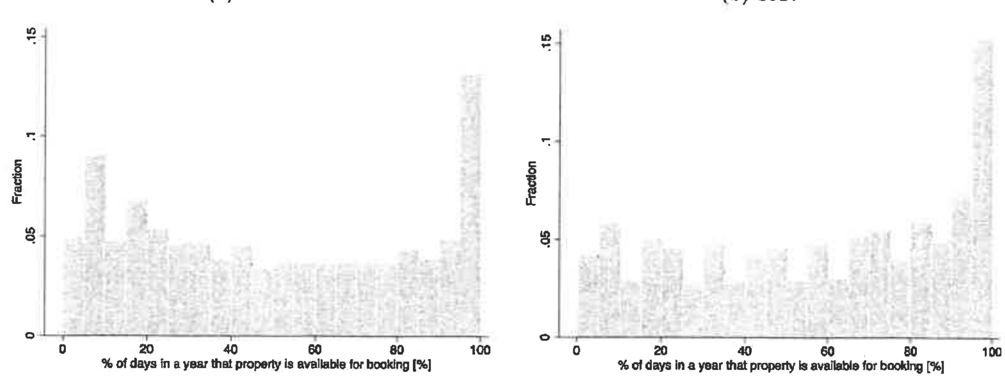
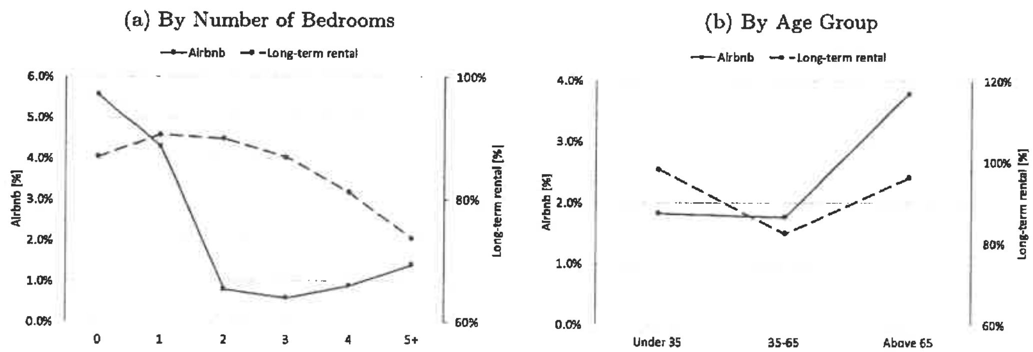


Table 2: Percentage of Units: By Metro Area

Metro Area	# Units	Percentage of Units [%]			
		Airbnb	Long-term Rental	Vacant Full Year	Vacant Partial Year
Boston-Cambridge-Newton, MA-NH	1,222,478	2.31	92.46	3.90	1.33
Chicago-Naperville-Elgin, IL-IN-WI	2,354,125	1.06	90.04	8.16	0.74
Dallas-Fort Worth-Arlington, TX	1,928,557	0.48	95.47	3.41	0.65
Detroit-Warren-Dearborn, MI	1,004,684	0.26	84.57	13.99	1.18
Miami-Fort Lauderdale-West Palm Beach, FL	2,062,516	2.46	72.16	15.98	9.40
New York-Newark-Jersey City, NY-NJ-PA	6,353,294	2.73	90.45	4.96	1.86
Phoenix-Mesa-Scottsdale, AZ	1,244,430	0.93	87.02	5.66	6.39
San Francisco-Oakland-Hayward, CA	1,352,516	3.46	91.84	3.74	0.97
Washington-Arlington-Alexandria, DC-VA-MD-WV	1,519,406	1.97	91.49	5.08	1.47

Figure 4: Percentage of Units by Property Characteristics and Demographics



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.[15]

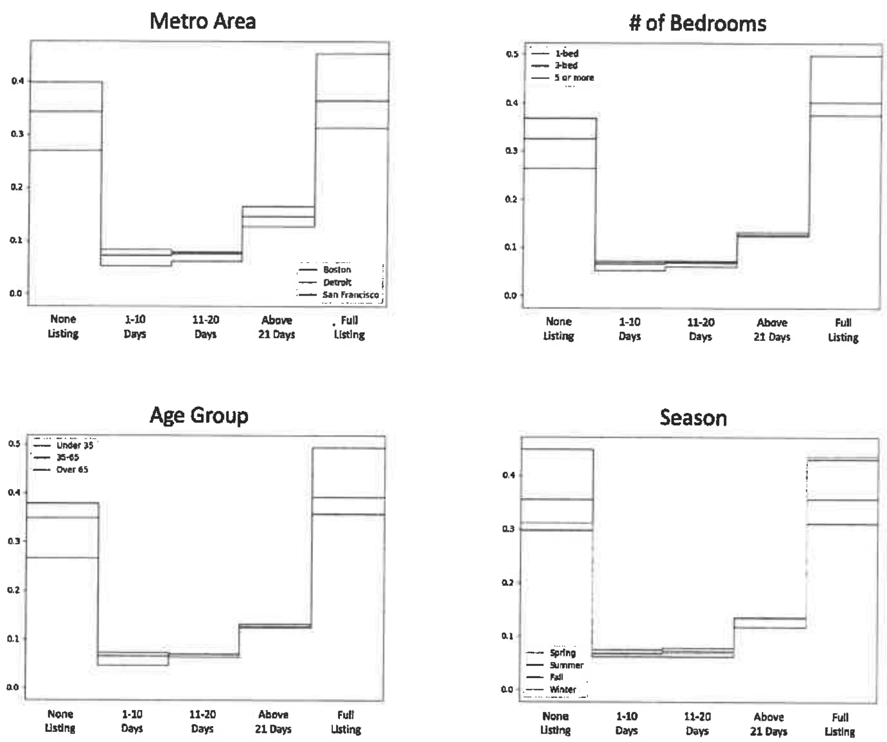
We also find that the Airbnb listing behavior varies by metro area, property characteristics, demographics, and season. Figure 5 shows a histogram of the number of days available for booking in a month by metro area, number of bedrooms, age, and season. Each observation represents a property-month combination. The overall bimodal pattern in Figure 2 holds for all subgroups, with variations across subgroups.

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. In practice, one major reason for hosts to rent out their properties on Airbnb is to earn some extra money while they

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

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



are away (e.g., when they are on vacation).[16] These hosts cannot list on Airbnb when they need to live in the property (e.g., when they are working or studying); in this case, the decision not to list in a month is due to availability, not cost-benefit trade-offs. In general, the observed hosting decisions are a result of both the endogenous decisions based on cost-benefit trade-offs and the exogenously determined host availability. We need to distinguish between the case in which a host chooses not to list because of costs and benefits and the case in which a host cannot list due to availability reasons.

We model the exogenous availability of hosts by categorizing the properties into two types. The first type is available for rent for the full year (hereafter, "fully 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, "partially 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 fully available type, the property owners select the use of their properties among three options: (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 partially available type, the property owners

[16] See https://www.cnbc.com/2019/07/03/is-running-an-airbnb-profitable-heres-what-you-need-to-know.html.

cannot choose the long-term rental option and only have options (1) and (3). Distinguishing between the two availability types is necessary to obtain unbiased estimates of hosting costs and results on switching and market expansion.[17]

Given the exogenous availability types, property owners make endogenous 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 represents 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 4.4. 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.[18] 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.

We make several assumptions. First, we assume that the set of potential properties that can be rented

[17]To see why, suppose we do not distinguish between the two types and allow all hosts to choose among Airbnb, long-term rental, and the outside option in the model. If property i is actually the "partially available" type and is listed on Airbnb in the data, the model that forces all hosts to choose among all three options will conclude that property i's Airbnb profit must be higher than its long-term rental profit. However, this is not necessarily true in reality because the fact that property i is not listed on the long-term rental market is due to availability reasons and not because of cost-revenue trade-offs; property i's host can only compare profits from Airbnb versus vacant. Therefore, for a "partially available" type property on Airbnb, a model that forces all hosts to choose among all three options will overestimate its Airbnb profit and underestimate its Airbnb hosting cost. Distinguishing between the two types is also important for analyzing switching and market expansion, which we detail in Section 7.2.

[18]We need to model the costs in addition to the revenues because the observed revenues alone cannot explain the observed hosting behaviors. For example, we observe that properties with more bedrooms have higher revenues on either the long-term rental market or Airbnb, suggesting that it is more beneficial for the hosts to rent out these properties. However, by contrast, we observe that properties with more bedrooms have lower probabilities of listing on either the long-term rental market or Airbnb, as shown in Figure 4. It means that revenue itself cannot rationalize the hosting decisions in the data. We need to incorporate cost-side components in the hosts' decision problem. Our cost-side estimates show that properties with more bedrooms have higher hosting costs, which can explain why these properties are less likely to be listed on either market in the data.

or vacant is fixed as exogenously given in the data. We do not consider the case in which hosts purchase or build new properties because of the introduction of Airbnb.[19] Second, we assume away the case in which long-term rental tenants sublet on Airbnb because we do not observe whether an Airbnb listing is from a sublease or not. This type of case may be relatively rare, as lease agreements often include clauses that prohibit sublets. Services such as SubletAlert.com and SubletSpy also help landlords find tenants who have violated the agreement. Our model focuses on the property owners' decisions and does not consider the renters' decisions to sublet. Future research may extend the proposed model to incorporate cases in which tenants sublet on Airbnb if such data are available and such cases become more common.[20] Third, we model the long-term rental choice as a yearly decision because the long-term rental lease is usually one year long in practice. Therefore, the model does not allow the host to choose long-term rental for part of the year and Airbnb for the rest of the year.

4.1 Second Stage: Continuous Decision of Listing on Airbnb

We first describe the model setup for the second-stage decision, because the profits from the second stage are nested into the first-stage decision and property owners must form expectations about the second stage before making first-stage decisions. Note that although in the data we observe the second-stage monthly decisions only for Airbnb properties, in the model we need to solve for the optimal second-stage monthly decisions for every property. This is because hosts in the model choose among Airbnb, long-term rental, and vacant options in the first stage. To make this decision, they need to formulate and compare the expected profits from each option, regardless of which option they chose in the data. The expected profits of the Airbnb option are calculated by summing over the expected monthly Airbnb profits in the second stage. Therefore, we need to solve for the second-stage monthly decisions for each host in the model.

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.[21] In the counterfactual analysis, we allow $\bar{s}$ to reflect the maximum listing length imposed by

[19]New properties built or purchased because of Airbnb count towards the market expansion effect of Airbnb; they would not have been listed on the long-term rental market without Airbnb. Allowing for these new properties will increase the size of the market expansion effect, so our estimated size in this paper can serve as a lower bound in this case.

[20]In the case where renters of long-term rental properties sublease on Airbnb, these properties are still part of the long-term rental supply as property owners still rent these properties to renters on the long-term rental market. Therefore, the property owners are not switchers from long-term rental to Airbnb and do not count towards cannibalization. However, the renters earn additional income on Airbnb; they would not have had this additional income source without Airbnb. Therefore, the renters are non-switchers and count towards the market expansion effect of Airbnb. Overall, considering renter sublease does not affect the size of cannibalization, while it increases the size of market expansion. Our estimated size of cannibalization is not affected while our estimated size of market expansion can serve as a lower bound in this case.

[21]In practice, there are government regulations that limit the maximum number of days on 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

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 online appendix.

The optimal monthly 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 4.4. 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 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 variable 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. Specifically, X_{it}^{Av} includes property characteristics (number of bedrooms/bathrooms/amenities, listing type), host demographics (age, education, income, marital status, gender), and a set of metro-area-level characteristics that relate to the cost of hosting.[22] The metro-area-level characteristics include population, density,

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.

[22]The host demographics are captured by dummy variables, one for each demographic group. For example, there are three age groups, "under 35", "35 to 65", and "over 65", which are captured by three dummy variables. For AHS properties, we observe individual host demographics, so the demographic dummy variables take the value of 0 or 1. For Airbnb properties, we do not

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 variable is specified as $\xi_{mt}^{Av} = \xi_t + \xi_0^{Av} \cdot 1\{T = 2017\} + \xi_1^{Av} \cdot (T - T_m^0)$, where ξ_t are season fixed effects (spring, summer, fall, winter) and 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 monthly unobservables that vary across seasons. The second component captures any yearly unobservables that affect all metro areas (e.g., because of Airbnb's national marketing) relative to those in the baseline year 2015. The third component captures any market-specific time trend related to Airbnb history or how long Airbnb has been present in a city. For example, 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 3a.

Derivation of the second-stage probability. Let $\Pr(s_{it} \,|\, \mathcal{X}_i)$ denote the second-stage choice probabilities, where $\mathcal{X}_i$ contains all host demographics, metro area and property characteristics that affect the costs and revenues of host i. We can construct $\Pr(s_{it} \,|\, \mathcal{X}_i)$ based on the feasible range of the normally distributed error term $\{\epsilon_{it}^{Av}\}$ in c_{it}^{Av} implied by the optimal choices in Equation 2:

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

where $\Phi(\cdot)$ is the CDF of the standard normal distribution. The derivation of $\Pr(s_{it} = s, 0 < s < \bar{s} \,|\, \mathcal{X}_i)$ accounts for the fact that the number of days listed s_{it} is an integer. We detail the derivation in the online appendix.

observe individual host demographics and impute them using the local zip-code-level demographics. The demographic dummy variables take the value of the percentage of a specific demographic group in the zip code. For example, the dummy variable "age under 35" takes the value of 30% if the host lives in a zip code where 30% of the population is under 35.

Note that we do not restrict c_{it}^{Av} to be nonnegative for mathematical reasons. The optimal number of days listed s_{it} can take values from 0 to $\bar{s}$. Allowing c_{it}^{Av} to take any real values guarantees that the probabilities of s_{it} taking each possible value sum up to 1, $\Pr(s_{it}=0\,|\,\mathcal{X}_i)+\Pr(0<s_{it}<\bar{s}\,|\,\mathcal{X}_i)+\Pr(s_{it}=\bar{s}\,|\,\mathcal{X}_i)=1$.[23] To interpret the value of c_{it}^{Av}, a lower value of c_{it}^{Av} suggests that the property is more likely to be listed longer. A negative value of c_{it}^{Av} suggests that the property is more likely to be listed for the full month. 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.

4.2 First Stage: Discrete Decision of Where to List

In the first stage, property owners decide whether and where to list their properties given their availability types. The fully 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 fully 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 partially 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

[23] Specifically, we can re-write the second-stage probabilities as functions of c_{it}^{Av}: $\Pr(s_{it}=0\,|\,\mathcal{X}_i)=\Pr\left(c_{it}^{Av}>p_{it}^{A}\phi_{it}^{A}\right)$, $\Pr(0<s_{it}<\bar{s}\,|\,\mathcal{X}_i)=\Pr\left(\frac{p_{it}^{A}\phi_{it}^{A}}{\exp(1)}<c_{it}^{Av}<p_{it}^{A}\phi_{it}^{A}\right)$, $\Pr(s_{it}=\bar{s}\,|\,\mathcal{X}_i)=\Pr\left(c_{it}^{Av}<\frac{p_{it}^{A}\phi_{it}^{A}}{\exp(1)}\right)$. Given that s_{it} can take any value from 0 to $\bar{s}$, these three probabilities needs to sum up to 1. Allowing c_{it}^{Av} to be in the whole real line guarantees that the requirement is satisfied. However, if we restrict c_{it}^{Av} to be non-negative, the first two probabilities remain the same, while the third probability will be smaller than before because now $\Pr(s_{it}=\bar{s}\,|\,\mathcal{X}_i)=\Pr\left(0<c_{it}^{Av}<\frac{p_{it}^{A}\phi_{it}^{A}}{\exp(1)}\right)$ and is smaller than $\Pr\left(c_{it}^{Av}<\frac{p_{it}^{A}\phi_{it}^{A}}{\exp(1)}\right)$. In this case, the sum of the three probabilities will be smaller than 1, which violates the requirement.

Table 3: Summary of Covariates

(a) 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 characteristics		yes	yes	yes	yes	yes	yes
Demographics	yes		yes	yes	yes	yes	yes
Metro area characteristics	yes		yes	yes (metro-	yes (metro-	yes	yes
Season fixed effect			yes	year-month)	year-month)	(metro-year)	(metro-year)
Mortgage		yes	yes				
Wage and employment in accommodations industry		yes	yes				
Airbnb regulation score			yes				
Airbnb history		yes	yes	yes	yes		
Tourism (air passengers)				yes	yes		
Google search index for Airbnb				yes	yes		
Total Airbnb supply				yes	yes		
Airbnb price					yes		
Total rental supply						yes	yes
Rent							yes

(b) Data Description of Covariates

Covariate	Data Description
Property characteristics	Number of bedrooms/bathrooms/amenities, property type: house (dummy)
Demographics	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 under 35, education below high school, income below 50k, female, and married.
Metro area characteristics	Population, density
Mortgage	Mortgage affordability index from Zillow
Wage and employment in accommodations industry	Average wage in the accommodation industry, percentage of the population who work in the accommodation industry
Airbnb regulation score	A score that measures how friendly city regulations are to short-term rental
Airbnb history	Time since Airbnb reached 10% of total rooms supplied by hotels and Airbnb in a city
Tourism	Number of air passengers to a city
Google search index for Airbnb	Number of searches for "airbnb" on Google
Total Airbnb supply	Number of days listed on Airbnb by units that are comparable to the focal property in the same city in a month
Airbnb price	Average daily price of the focal property on Airbnb in a month
Total rental supply	Number of units listed on the long-term rental market that are comparable to the focal property in the same city in a year
Rent	Rent of the focal property on the long-term rental market in a year

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). We discuss how the expectations are formed in Section 4.4. 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}:

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

We show in the online appendix how we compute the ex ante monthly profit, accounting for s_{it} being integers. Note that we are able to calculate $E\left[\Pi_{it}^A(s_{it}^*)\right]$ for every host without observing the actual monthly decisions in the data. This is because the optimal number of days listed s_{it}^* and the ex ante monthly Airbnb profits $E\left[\Pi_{it}^A(s_{it}^*)\right]$ are functions of the hosting costs and expected prices and occupancy rates. The hosting costs are functions of the observable characteristics as shown in Equation 3; the expected monthly prices and occupancy rates are functions of the observable characteristics as shown in Section 4.4. Therefore, s_{it}^* and $E\left[\Pi_{it}^A(s_{it}^*)\right]$ are functions of the observable characteristics as well, which are observed for every property, regardless of whether the property is listed on Airbnb in the data. We can use the observed characteristics to compute s_{it}^* and $E\left[\Pi_{it}^A(s_{it}^*)\right]$ for every host without observing their actual second-stage monthly decisions.

Property owners are heterogeneous in the 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 (normalized to zero for identification reasons) 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 $\tau = 1$ denotes the fully available type and $\tau = 2$ denotes the partially available type. The baseline cost $\bar{c}_\tau^{Af}$ can take different values for fully available owners ($\tau = 1$) and partially available owners ($\tau = 2$). 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 variable 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 variable is specified as $\xi_{mT}^{Af} = \xi_0^{Af} \cdot 1\{T = 2017\} + \xi_1^{Af} \cdot (T - T_m^0)$, 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 related to Airbnb's history in a city.

Derivation of the first-stage probability. We can construct the first-stage choice probabilities $\Pr(d_{iT} | \tau_i, \mathcal{X}_i)$ based on the feasible range of the independently normally distributed error terms $\left\{\epsilon_{iT}^{Af}, \epsilon_{iT}^{R}\right\}$ in c_{iT}^{Af} implied by the optimal choices in Equations 4 and 5. Define $a_\tau \equiv \sum_{t\in T} E\left[\Pi_{it}^A(s_{it}^*) - c_{iT}^{Af} | \tau_i\right]$ and $r \equiv E\left(p_{iT}^R \phi_{iT}^R - c_{iT}^R\right)$. For the fully available hosts ($\tau_i = 1$), they choose among Airbnb, long-term rental, and the outside option and have the following choice probabilities:

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

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of the standard normal distribution. The integrals in $\Pr(d_{iT} = R | \tau_i = 1, \mathcal{X}_i)$ and $\Pr(d_{iT} = A | \tau_i = 1, \mathcal{X}_i)$ are calculated using Gauss-Laguerre quadrature with 10 nodes. For the partially available hosts ($\tau_i = 2$), they choose between Airbnb and the outside option and have the following choice probabilities:

$$
\begin{aligned}
\Pr(d_{iT} = O \,|\, \tau_i = 2, \mathcal{X}_i) &= \Pr\left(\epsilon_{iT}^{Af} > a_2\right) = \left(1 - \Phi\left(\frac{a_2}{\sigma_1}\right)\right) \\
\Pr(d_{iT} = A \,|\, \tau_i = 2, \mathcal{X}_i) &= \Pr\left(\epsilon_{iT}^{Af} < a_2\right) = \Phi\left(\frac{a_2}{\sigma_1}\right)
\end{aligned}
$$

4.3 Owner Availability Types

In the data, we observe the owner availability types τ_i 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 fully 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 fully available type, whereas owners of properties that are kept vacant for the some part of the year due to occasional self-use are the partially available type. However, we do not observe the owner availability types for properties on Airbnb in the data.[24] There are two exceptions: 1) properties that are listed every month throughout the year in the data must be the fully 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 partially 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 fully available or partially available.

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 fully available type ($\tau_i = 1$) in year T is

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

The probability that it is the partially available type is $\Pr(\tau_i = 2) = 1 - \gamma_{iT}$. 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 3a.

4.4 Expectation of Revenue

Property owners' decisions in Equations 1, 4, and 5 contain 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 properties that were listed on the long-term rental market in a year and observe Airbnb price and occupancy rate only for properties that were listed on Airbnb

[24] Note that one cannot conclude that Airbnb properties that are listed for some part of the year must be the partially 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 fully available host may choose to list for only part of the year because the costs exceed the benefits for the rest of the year.

in a year. 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 7.1. 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 market m in year T are

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

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

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_{imT}^{R} (measured as the number of units that are comparable to property i and choose to list on the long-term rental market), and metro-area-specific year fixed effects ψ_{mT}^{Rp}.[25] 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.[26] We exclude outliers with rents below the 5th percentile and above the 95th percentile when running the regressions.

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

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

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

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_{imt}^{A} (measured as the number of days listed by all units that are comparable to property i on Airbnb), Airbnb-related variables

[25] "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.

[26] 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.

x^{A}_{mt}, and metro-area-specific year and month fixed effects ψ^{Ap}_{mt}. The market-specific year and month fixed effects can capture market-specific seasonality patterns in Airbnb prices. The Airbnb-related variables include tourism (measured as the number of air passengers to the city), Airbnb history (measured as the number of months since Airbnb reached 10% of the total rooms supplied by hotels and Airbnb in a city), and Google search index for Airbnb (measured as the number of Google searches for "airbnb"). In particular, tourism can proxy for the heterogeneous tourism popularity across cities. Airbnb history can proxy for unobserved factors that relate to the length of Airbnb's presence in a city. Google search index for Airbnb can proxy for unobserved demand shocks that are common across cities due to Airbnb growth. 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. We exclude outliers with Airbnb price below the 5th percentile and above the 95th percentile when running the regressions.

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

The rental and Airbnb supply of comparable units $\{S^{R}_{imT}, S^{A}_{imt}\}$ and the rental and Airbnb price $\{p^{R}_{iT}, p^{A}_{it}\}$ in the hedonic regressions are potentially endogenous variables. We address the endogeneity issue using instruments and discuss the details in Section 5.2. We jointly estimate Equations 9 and 10 as a system of equations and Equations 11 and 12 as another system of equations using three-stage least squares (3SLS) to allow for the correlation of the error terms within each equation system. The regression results are provided in the online appendix. The coefficients have the expected signs.[27]

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.[28] These expected revenues are used in the property owners' decisions in Equations 1, 4 and 5.

[27]For example, both rent and Airbnb price increase with the number of bedrooms and bathrooms. An increase 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 Airbnb price and occupancy. A higher Airbnb price decreases Airbnb occupancy.

[28]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 fully 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.

5 Estimation Method

5.1 Estimation

We use the maximum likelihood estimation (MLE) method to estimate the model. The likelihood function for host i is the joint probability of the first-stage decision on the use of the property and, if the Airbnb option is chosen, the second-stage decision on the number of days to list on Airbnb. The choice set of the first-stage decision depends on the hosts' availability types. There are three sets of hosts in the data:

(1) Hosts that are known to be the fully available type ($\tau_i = 1$). They choose among Airbnb, long-term rental, and the outside option. Let Ω_1 denote this set of hosts. It contains hosts of long-term rental properties, properties that are vacant for the full year, and Airbnb properties that are known to be the fully available type.

(2) Hosts that are known to be the partial available type ($\tau_i = 2$). They choose between Airbnb and the outside option. Let Ω_2 denote this set of hosts. It contains hosts of properties that are vacant for part of the year and Airbnb properties are are known to be the partial available type.

(3) Hosts that we do not observe the types. They can be the fully available type with probability $\Pr(\tau_i = 1 \,|\, \mathcal{X}_i)$ and the partially available type with probability $\Pr(\tau_i = 2 \,|\, \mathcal{X}_i)$. Let Ω_U denote this set of hosts. It contains hosts of Airbnb properties that are of unknown types.

For a host in Set Ω_1, his first-stage decision has three options. His contribution to the likelihood function is:

$$l_{1i}(\Theta | d_{iT}, s_{it}, \mathcal{X}_i) = \prod_T \{\Pr(d_{iT} = R \,|\, \tau_i = 1, \mathcal{X}_i)^{1(d_{iT}=R)} \cdot \Pr(d_{iT} = O \,|\, \tau_i = 1, \mathcal{X}_i)^{1(d_{iT}=O)} \cdot \left[\Pr(d_{iT} = A \,|\, \tau_i = 1, \mathcal{X}_i) \prod_{t \in T} (\Pr(s_{it} \,|\, \mathcal{X}_i))\right]^{1(d_{iT}=A)} \} \tag{13}$$

where $\mathcal{X}_i$ contains all host demographics, metro area and property characteristics that affect the costs and revenues of host i, $\Pr(d_{iT} \,|\, \tau_i, \mathcal{X}_i)$ are the probabilities of the first-stage decision, and $\Pr(s_{it} \,|\, \mathcal{X}_i)$ is the probability of the second-stage decision. For a host in Set Ω_2, his first-stage decision has two options. His contribution to the likelihood function is:

$$l_{2i}(\Theta | d_{iT}, s_{it}, \mathcal{X}_i) = \prod_T \{\Pr(d_{iT} = O \,|\, \tau_i = 2, \mathcal{X}_i)^{1(d_{iT}=O)} \cdot \left[\Pr(d_{iT} = A \,|\, \tau_i = 2, \mathcal{X}_i) \prod_{t \in T} (\Pr(s_{it} \,|\, \mathcal{X}_i))\right]^{1(d_{iT}=A)} \} \tag{14}$$

For a host in Set Ω_U, his contribution to the likelihood function is:

$$l_{Ui}\left(\Theta|d_{iT}, s_{it}, \mathcal{X}_i\right) = \prod_{T}\left\{\left[\sum_{\tau_i=1,2} \Pr\left(\tau_i \,|\, \mathcal{X}_i\right) \cdot \Pr\left(d_{iT} = A \,|\, \tau_i, \mathcal{X}_i\right)\right] \cdot \prod_{t\in T}\left(\Pr\left(s_{it} \,|\, \mathcal{X}_i\right)\right)\right\}^{1(d_{iT}=A)} \tag{15}$$

where $\Pr\left(\tau_i \,|\, \mathcal{X}_i\right)$ is the probability of owner availability types. The full likelihood function combines the probabilities from the three sets of hosts:[29]

$$\mathcal{L}\left(\Theta|d, s, \mathcal{X}, \tau\right) = \prod_{i}\left[l_{1i}\left(\Theta|d_{iT}, s_{it}, \mathcal{X}_i\right)^{1\{i\in\Omega_1\}} l_{2i}\left(\Theta|d_{iT}, s_{it}, \mathcal{X}_i\right)^{1\{i\in\Omega_2\}} l_{Ui}\left(\Theta|d_{iT}, s_{it}, \mathcal{X}_i\right)^{1\{i\in\Omega_U\}}\right] \tag{16}$$

One caveat is that, as discussed in Section 3.1, we do not observe host demographics for Airbnb properties. For Airbnb properties, the likelihood function is integrated over the zip-code-level demographics distribution $f(\mathcal{X}_i)$ from the ACS data. Given that we conduct a two-step estimation (i.e., estimate the hedonic regressions in the first step and the hosts' decisions in the second step), we correct the standard errors following Murphy and Topel (1985).

5.2 Identification

General identification strategies. We observe three types of information in the data: the revenues, the hosts' first-stage decision, and the hosts' second-stage decision. 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 $\{\bar{c}^{Av}, \beta^{Av}, \xi_t, \xi_0^{Av}, \xi_1^{Av}, \sigma_2\}$, which are identified by the Airbnb listing pattern. In particular, the average number of days listed and its variation across host demographics, properties, and metro areas identify the Airbnb variable cost parameters $\{\bar{c}^{Av}, \beta^{Av}\}$. The time-related parameters $\{\xi_t, \xi_0^{Av}, \xi_1^{Av}\}$ are identified by the listing pattern differences over time and across markets with different lengths of Airbnb history.

The cost-side parameters in the first-stage decision include owner availability type parameters β^a, 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

[29] As discussed in Section 4.3, we observe the owner availability types for two categories of Airbnb properties: 1) properties that are listed every month in the data must be the fully available type ($\tau_i = 1$); 2) properties that are listed as a "private room" must be the partially available type because the hosts live with the guests ($\tau_i = 2$). For these Airbnb properties with observed τ_i, their likelihood function has an additional component, $\Pr\left(\tau_i = \tau \,|\, \mathcal{X}_i\right)^{1(\tau_i=\tau)}$.

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.

The identification of switching from long-term rental to Airbnb mainly relies on properties with similar observable characteristics. Specifically, the revenue-side prices and occupancy rates and the cost-side fixed and variable costs of hosting in our model are functions of property, host, and metro area characteristics. Therefore, properties with similar characteristics have similar revenues and costs and make similar decisions in a particular setting. As the setting exogenously changes over time, similar properties make different decisions over time in the data, which identifies switching in the model estimation.

More generally, there are exogenous factors in the data that vary over time and across markets that help identify switches from long-term rental to Airbnb. First, on the revenue side, the overall demand for Airbnb grows over time as Airbnb penetrates the market. Cities in which Airbnb entered earlier experienced larger growth. Cities also have different levels of tourism popularity before Airbnb is introduced. Hosts in cities with high tourism popularity are more likely to earn higher revenues on Airbnb when Airbnb is introduced and switch away from long-term rentals. Second, on the cost side, markets differ in conditions that affect Airbnb hosts. For example, cities have different mortgage pressures, which affect hosts' motivations to list on Airbnb. Cities also differ in resources in the accommodations industry before Airbnb is introduced and differ in levels of support in policies on Airbnb after Airbnb is introduced, both of which affect the costs of hosting. Hosts in cities with more favorable conditions are more likely to switch away from long-term rentals. Finally, there is seasonality in both demand-side tourism patterns and supply-side hosting costs. These seasonal fluctuations also affect short-term hosting decisions.

Endogeneity in hedonic regressions. We use instruments to address the endogeneity issue of the rental and Airbnb supply of comparable units $\{S^R_{imT}, S^A_{imt}\}$ and the rental and Airbnb prices $\{p^R_{iT}, p^A_{it}\}$.

First, we instrument the rental supply S^R_{imT} using the rental supply of comparable rental units in other markets in year T. The supply of comparable units in other markets is a valid instrument because (1) it is correlated with the supply of comparable units; hosts of these units have similar characteristics and are affected by similar cost shifters as comparable units in the focal market; (2) it is not correlated with the prices of the focal unit because comparable units in other markets do not directly compete with the focal unit and do not affect the focal unit's prices.

Second, we instrument Airbnb supply S^A_{imt} using the 12-month lagged Airbnb supply of comparable units in the focal market. It is a valid instrument because (1) the 12-month lagged units and the current-period comparable units share similar cost shifters that are related to the time of the year, so the 12-month lagged supply and the current-period supply are correlated; (2) the 12-month lagged unobservables and the current-period unobservables are unlikely to be serially correlated given the relatively long time gap, so the 12-month lagged Airbnb supply is uncorrelated with the current-period demand shocks and is uncorrelated with the focal unit's prices. In addition to these 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.

Third, we instrument the rent p^R_{iT} using average property characteristics of comparable rental properties in other markets in year T. Similarly, we instrument the Airbnb price p^A_{it} using the average property characteristics of comparable Airbnb properties in other markets in month t. The average property characteristics of comparable properties in other markets are valid instruments because (1) characteristics affect hosting costs and in turn prices; comparable properties in other markets have similar characteristics and thus comparable hosting costs as the focal property, so characteristics of comparable units in other markets are correlated with the prices of the focal unit; (2) comparable units in other markets do not directly compete with the focal unit and do not affect the focal unit's occupancy rate, so characteristics of comparable units in other markets do not correlate with the focal unit's occupancy rate.

The instruments pass the weak IV tests.[30]

Exclusion restrictions. Property characteristics, host demographics, and metro area characteristics enter both the cost-side components and the revenue-side regressions. Exclusion restrictions come from the non-overlapping variables. As summarized in Table 3a, each cost component or revenue regression has exclusive variables that do not enter other components. For example, 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. In addition, 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 hosts' incentives. Finally, Airbnb-related variables (e.g., Airbnb history) affect only Airbnb and not the long-term rental

[30]The first-stage regression F-statistics are 864,055 for Airbnb supply, 22,573 for Airbnb price, 119,837 for rental supply and 196 for rent. The incremental R-squared of the first-stage regression when instruments are added is 0.092 for the rental regressions and 0.275 for the Airbnb regressions.

Table 4: Model Fit: First-Stage Decision

[%]	Airbnb	Rental	Vacant full year	Vacant partial year
Observed	1.98	88.70	6.77	2.55
Predicted	2.27	89.74	5.63	2.37

market. In general, the exclusion restrictions stem from the facts 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.[31] Consider the number of bedrooms as an example. 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 example, owners of 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.

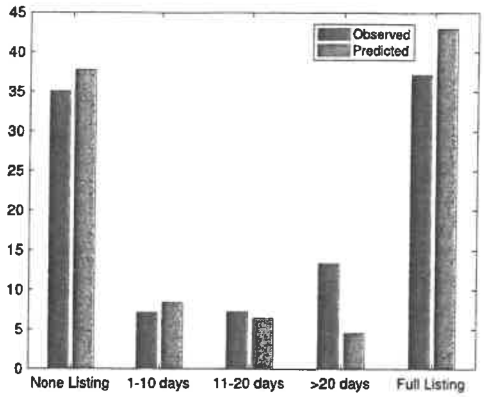
6 Estimation Results

6.1 Model Fit

Table 4 shows the observed and predicted percentages of Airbnb, long-term rental, and outside option properties. 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 patterns across heterogeneous groups. Overall, these results suggest that the model can recover the heterogeneous hosting costs across property characteristics, metro areas, and demographics.

[31] 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}.

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. We use the estimates of $\{\bar{c}^{Av}, \beta^{Av}\}$ and Equation 3 to calculate the predicted variable costs of Airbnb hosting for each property in the data. The values of the predicted costs can be interpreted relative to the hosts' decisions. A smaller predicted variable cost c_{it}^{Av} means that the property is more likely to be listed longer in a month on Airbnb. Across properties in the data, the median predicted variable cost of Airbnb hosting is \$27.9 per day, with a 25 percentile of \$12.1 and a 75 percentile of \$45.2. The estimates suggest that additional bedrooms and facilities increase the variable cost of hosting. The daily cost for an entire place listing is \$29.5 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 \$33.5.[32]

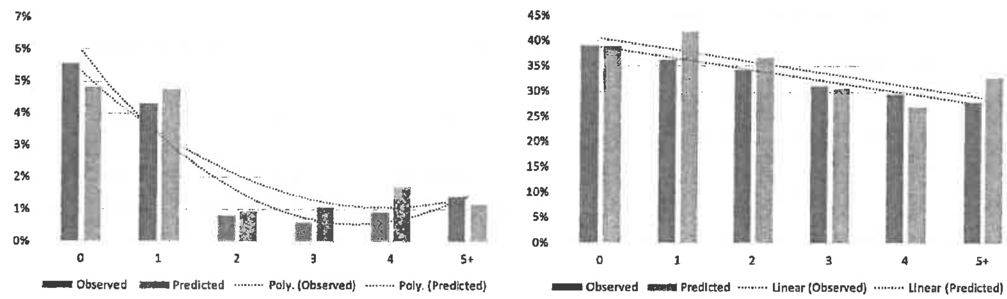
Host demographics and metro area characteristics also affect the Airbnb variable cost and how long hosts list their properties. Hosts have a lower estimated variable cost and list longer if they are younger, have a high school education level and medium income, are female, and are never married. Hosts list longer in cities with a smaller population, a 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.

Owner availability type. The probability of being fully available or partially available varies by host

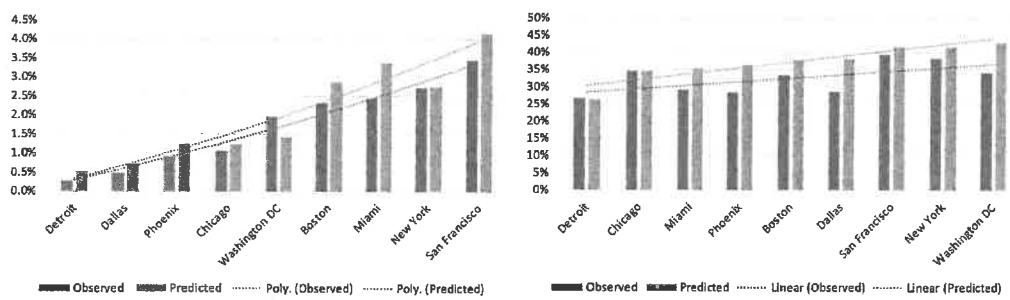
[32]See https://www.tidy.com/compare-house-cleaning-prices

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

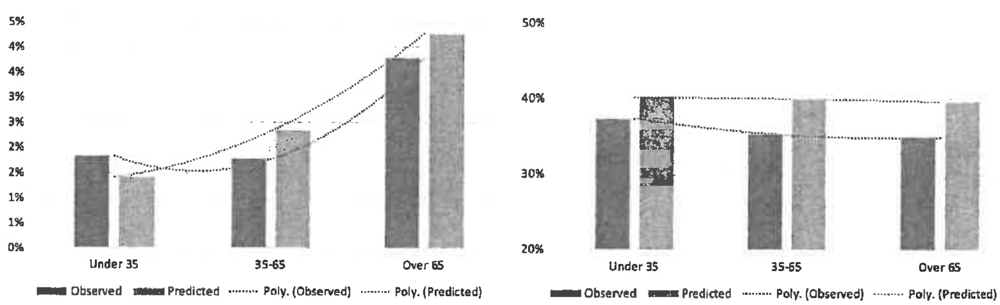
(a) By Property Characteristics (Number of Bedrooms)



(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.

Table 5: Parameter Estimates

(a) First-Stage Decision

		Est.	Std.			Est.	Std.
$\bar{c}_1^{Af}$		61.31**	0.984		age: 35-65	0.0685**	0.0035
$\bar{c}_2^{Af}$		21.49**	1.001		age: over 65	0.6783**	0.0805
σ_1		7.273**	0.0184		edu: high school	1.278**	0.0351
β^{Af}	# of bedrooms	5.571**	0.0984	β^{a}	edu: bachelor	3.689**	0.0255
	# of bathrooms	6.319**	0.0983		income: 50K-100K	-2.986**	0.0751
	# of amenities	-1.420**	0.0085		income: over 100K	-3.033**	0.0830
	type: house (dummy)	4.605**	0.0522		gender: male	2.231**	0.0051
	studio (dummy)	0.588**	0.191		marital status: never	-0.390**	0.0563
	mortgage	-9.059**	0.3257		population	0.231**	0.0039
	employment in accomm.	-1.173e-4	1.489e-4		density	-0.713**	0.0328
	wage in accomm.	6.953e-4**	2.720e-4		single bedroom (dummy)	-3.086**	0.0201
	year fixed effect	-0.648**	0.0701		constant	-3.887**	0.0240
	Airbnb history	1.389**	0.0381				

(b) Second-Stage Decision

		Est.	Std.		Est.	Std.
$\bar{c}^{Av}$		59.41**	1.275	σ_2	51.55**	0.140
β^{Av}	age: 35-65	8.195**	0.145	population	0.778**	0.067
	age: over 65	9.619**	0.310	density	3.516**	0.424
	edu: high school	-217.0**	0.700	mortgage	37.05**	0.121
	edu: bachelor	-127.8**	0.666	employment in accomm.	-2.889**	0.140
	income: 50K-100K	-10.063**	0.691	wage in accomm.	1.135e-4	7.921e-4
	income: over 100K	5.087**	0.157	Airbnb regulation score	-23.79**	0.109
	gender: male	4.228**	0.391	year fixed effect	-8.574**	0.215
	marital status: never	-4.447**	0.0458	Airbnb history	-11.52**	0.083
	# of bedrooms	5.197**	0.146	summer	12.26**	0.320
	# of bathrooms	-0.738**	0.039	fall	21.00**	0.046
	# of amenities	1.258**	0.062	winter	-4.454**	0.162
	entire room (dummy)	29.53**	0.224			

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 host demographics are captured by dummy variables, one for each demographic group. The baseline demographics group is age under 35, education below high school, income below 50k, female, and married. For Airbnb properties, we do not observe individual host demographics and impute them using the local zip-code-level demographics. The demographic dummy variables take the value of the percentage of a specific demographic group in the zip code.

demographics and metro area. The estimates of β^a suggest that hosts are more likely to be the fully available type if they are seniors, have a bachelor degree, 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 partially 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. We use the estimates of $\{\bar{c}_{\tau}^{Af}, \beta^{Af}\}$ and Equation 7 to calculate the predicted fixed costs of Airbnb hosting for individual properties in the data. The values of the predicted costs can be interpreted relative to the hosts' decisions. A smaller predicted fixed cost c_{iT}^{Af} means that the property is more likely to be listed on Airbnb. We find that the median predicted Airbnb fixed cost of Airbnb properties is \$633 per month (\$21.1 per day), with a 25th percentile of \$338 per month (\$11.3 per day) and a 75th percentile of \$2368 per month (\$78.9 per day). Note that the average daily price is \$148.7 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, as well as 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, descriptions, 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.[33] The learning cost of hosting is another major factor for which Airbnb needs to compensate the hosts, as evident by Airbnb's significant spending on technology and administrative costs associated with the hosts.[34]

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.[35] In fact, the primary use of the hosting income is to pay mortgages according to a survey

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

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

[35]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.

conducted by Airbnb.[36] Airbnb hosts can even use Airbnb income as proof of worth when applying for mortgage refinancing.[37] Finally, the fully available type of hosts have higher Airbnb fixed costs and are less likely to choose Airbnb than the partially available type of hosts. This may be because these fully available type of hosts have long-term rental as their default option and are reluctant to overcome the inertia and adopt the new technology of Airbnb.

7 Counterfactuals

Given the model estimates, we conduct a series of counterfactual analyses.[38] The first set of analyses evaluate the impact of Airbnb on rental market and housing affordability. We simulate the property owners' choices when Airbnb is present versus when it is not present. We compare the two sets of equilibrium outcomes to evaluate how many Airbnb properties would have been listed on the long-term rental market without Airbnb (i.e., cannibalization effect of Airbnb) and how many properties would not (i.e., market expansion effect of Airbnb). The second set of counterfactual analyses evaluate the impact of a series of policies intended to ensure the supply and affordability of rental housing. We consider the two prevailing short-term rental regulations on Airbnb in practice, imposing taxes and limiting the maximum number of days that a property can be listed. We further propose a new policy and assess the desirability of different policies. Finally, we investigate rent control policy on long-term rental, particularly how its impact can be affected by the presence of Airbnb. In all counterfactual analyses, we assume that the set of properties is exogenously given in the data and abstract away from the case in which hosts purchase or build new properties because of the introduction of Airbnb.

7.1 Equilibrium

When solving for new equilibrium under counterfactual scenarios, we allow the rent, rental occupancy rate, Airbnb price and occupancy rate to endogenously change according to the hedonic regressions in Section 4.4. Specifically, given different counterfactual scenarios, 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

[36] See https://www.airbnbcitizen.com/the-airbnb-community-in-seattle/
[37] See https://www.cnbc.com/2018/02/22/homeowners-are-using-airbnb-rental-income-to-refinance-mortgages.html
[38] 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 endogenously change 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 rentals 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 market share in the future.

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 $\{p_{it}^A, \phi_{it}^A, S_{imt}^A, p_{iT}^R, \phi_{iT}^R, S_{imT}^R\}$. 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 $S_{imt}^{A(k)}$ and aggregate rental supply $S_{imT}^{R(k)}$. Given $S_{imt}^{A(k)}$ and $S_{imT}^{R(k)}$, construct the expected rent $p_{iT}^{R(k+1)}$, rental occupancy rate $\phi_{iT}^{R(k+1)}$, Airbnb price $p_{it}^{A(k+1)}$, and Airbnb occupancy rate $\phi_{it}^{A(k+1)}$ for each property using the hedonic regressions in Equations 9, 10, 11, and 12.

2. Given the updated $p_{it}^{A(k+1)}$, $\phi_{it}^{A(k+1)}$, $p_{iT}^{R(k+1)}$, and $\phi_{iT}^{R(k+1)}$, solve the property owners' problem under each counterfactual policy. Compute the updated aggregate Airbnb supply $S_{imt}^{A(k+1)}$ and aggregate rental supply $S_{imT}^{R(k+1)}$.

3. Check for the convergence of $\left|p_{it}^{A(k+1)} - p_{it}^{A(k)}\right|$, $\left|\phi_{it}^{A(k+1)} - \phi_{it}^{A(k)}\right|$, $\left|S_{imt}^{A(k+1)} - S_{imt}^{A(k)}\right|$, $\left|p_{iT}^{R(k+1)} - p_{iT}^{R(k)}\right|$, $\left|\phi_{iT}^{R(k+1)} - \phi_{iT}^{R(k)}\right|$, and $\left|S_{imT}^{R(k+1)} - S_{imT}^{R(k)}\right|$. 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.

Note that we use the hedonic regression coefficients estimated from the observed data in the counterfactual analyses. These coefficients capture how hosts form expectations about prices and occupancy rates in the data. The underlying assumption is that hosts in the counterfactuals form expectations in the same way as they do in the observed scenario. We believe that it can be a reasonable assumption in the short run in our setting.[39] This assumption is also similar to the assumptions made in the existing literature on durable product demand (e.g., Nair 2007, Gowrisankaran and Rysman 2011).[40] A limitation of the hedonic regression approach is that it captures how prices and occupancy rates are determined in a simplified and non-structural way. The regression coefficients are estimated using data in the current stage of the market where Airbnb is relatively small. The results may not apply to the long term if Airbnb's presence becomes significant relative to long-term rental market or if the ways that prices and occupancy rates are determined systematically change in the long run.

[39] First, the coefficients in Equations 9-10 capture how rent and rental occupancy are determined in the long-term rental market. Given that the size of Airbnb is relatively small (2.5% of the long-term rental market), Airbnb's presence and the policies on Airbnb are unlikely to systematically change the way rent and rental occupancy are determined in the long-term rental market in the short run. Second, the coefficients in Equations 11-12 capture how price and occupancy are determined on Airbnb. In practice, Airbnb hosts set prices by accounting for property characteristics, location, and seasonal demand; some hosts use third-party pricing services, which account for similar pricing factors (Li and Srinivasan 2019). These factors are captured in Equations 11-12. The ways these factors affect Airbnb prices and occupancy rates may not systematically change in the short run when certain regulations are introduced.

[40] Specifically, literature on durable product demand assumes that consumers expect that the prices follow an AR (1) process. First, the coefficients of the AR (1) process are estimated using the observed prices. Second, fixing these coefficient estimates, consumers make new product adoption decisions in the counterfactual analyses. As another example, Li and Srinivasan (2019) first use the observed price and supply to estimate how Airbnb price and supply are determined by observed characteristics and supply. Then, they use the estimated coefficients to obtain new price and supply in the equilibrium in the counterfactual analysis.

7.2 Cannibalization and Market Expansion

We first evaluate the impact of Airbnb on the long-term rental market and housing affordability. Airbnb can create both a negative impact of cannibalization and a positive impact of market expansion in the rental housing market. To evaluate the impact of Airbnb, we use the model estimates to simulate the property owners' choices when Airbnb is present versus when it is not present. We allow prices and occupancy rates to endogenously change when solving for new equilibrium outcomes under both scenarios.

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.

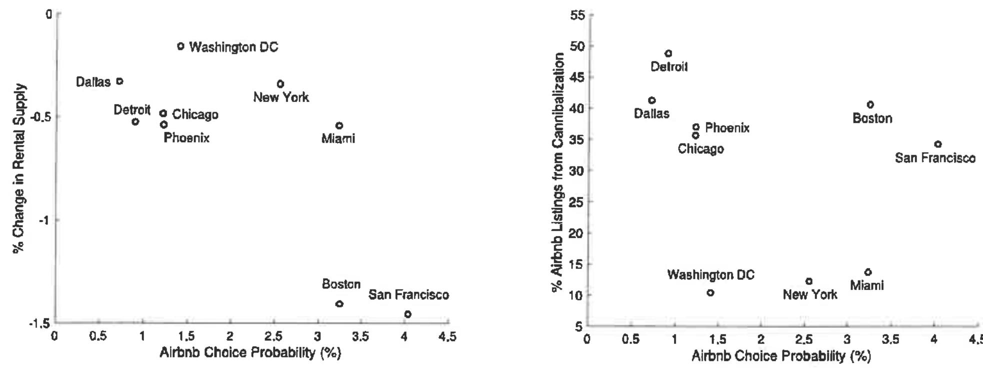
Note that the fully available hosts have the long-term rental option, whereas the partially available hosts do not. Therefore, only the fully available hosts can switch from the long-term rental market; the partially available hosts cannot. In general, the fully available hosts choose among all three options and can thus create both cannibalization and market expansion. The partially available hosts only choose between Airbnb versus vacant and can only create market expansion. Specifically, cannibalization can come from one situation: a fully available host would have listed the property on the long-term rental market and chooses to list on Airbnb when Airbnb is present. Market expansion can come from two situations: (1) a fully available host who has an unoccupied unit would have kept the entire unit vacant without Airbnb and chooses to list on Airbnb when Airbnb is present; (2) a partially available host would not have rented out without Airbnb and chooses to list on Airbnb when Airbnb is present.

Let D^{R0} and D^{R1} denote the equilibrium number of long-term rental units without and with Airbnb. Let D^A denote the equilibrium 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 non-switchers 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 Airbnb units (switchers and non-switchers), 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

Figure 8: Cannibalization and Market Expansion by Metro Area

(a) Change in Rental Supply

(b) Cannibalization Rate



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 -0.16% in Washington D.C. to -1.46% 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 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 10.4% in Washington D.C. to 48.8% in Detroit.

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. Thus, 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, more affordable units rather than among higher priced luxury 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

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	-1.47	35.36	# of Amenities	1	-0.82	7.56
	1	-0.70	13.53		2	-0.54	23.66
	2	-0.45	44.73		3	-0.33	21.64
	3	-0.42	39.12		4	-0.56	18.58
	4	-0.38	21.08		5	-0.72	23.55
	5+	-0.04	3.13		6	-0.78	26.10
# of Bathrooms	1	-0.69	22.97	Property Type	Apt	-0.56	22.37
	2	-0.45	16.38		House	-0.49	22.43
	3	-0.23	31.80	Listing Type	Entire Place	-0.54	31.15
	4	-0.15	16.83		Private Room	-0.01	0.01

(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	-0.29	19.69	Income	under 50K	-0.33	14.80
	35-65	-0.43	17.06		50K-100K	-0.70	35.87
	over 65	-1.68	38.44		over 100K	-0.94	25.83
Education	under high school	-0.76	39.15	Gender	male	-0.76	33.39
	high school	-0.45	18.93		female	-0.35	14.12
	bachelor's	-0.12	3.53	Marital Status	never married	-0.68	28.13
					other	-0.46	19.03

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 non-switchers is larger among affordable units on Airbnb. This suggests that, interestingly, affordable units are major sources of both the negative cannibalization impact and the positive market expansion impact of Airbnb. Although Airbnb harms local renters by reducing the 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 less economically advantaged. Therefore, policy makers need to strike a balance 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 hosts switch from long-term rentals to Airbnb. Even if hosts list their properties on Airbnb full time, it is not 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 due to 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 hosts to list part time and still earn more than listing in the long-term rental market.

7.3 Policy Evaluation

This subsection evaluates the impact of various policies intended to ensure rental housing supply and affordability: regulations on Airbnb (e.g., tax and day limit) and long-term rentals (e.g., rent control).

7.3.1 Policy Implementation

Short-term rental regulations on Airbnb. We focus on three types of regulations on Airbnb. The first and second types are the most prevalent policies in practice. The third is a new policy we propose based on our findings about Airbnb and housing affordability.

Specifically, the first type of regulation limits the maximum 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 120 days in Los Angeles). The second type charges a transient occupancy tax as a fixed percentage of the listing price (e.g., 8.5% in Philadelphia and 14% in Los Angeles), which is similar to a hotel occupancy tax. Both the first and second types of regulations are motivated by concerns about switchers from the long-term rental market to Airbnb, which can hurt the rental housing supply and affordability. By 2020, many cities have imposed these types of regulations on Airbnb.[41]

The third type of regulation is a convex tax that charges a higher tax on expensive units and a lower tax on affordable units. We propose this new policy because it shares a similar goal with the existing two policies and can help reduce the proportion of switchers. Heterogeneous properties have different proportions of switchers. To reduce switching, policy makers can consider charging a higher tax rate on properties with a larger proportion of switchers. As shown in Section 7.2, we find that affordable units have a lower proportion of switchers while expensive units have a higher proportion of switchers. Therefore, it can be helpful to charge a higher tax rate on the expensive units among which the proportion of switchers is larger. This constitutes a convex tax for which the tax rate increases as the listing price increases.

We need to operationalize the three types of regulations in the counterfactual analyses. To operationalize the first regulation, day limit, we simulate the case in which hosts are able to list up to a certain number of

[41] See https://www.airbnb.com/help/article/864/los-angeles-ca#nightlimits and https://www.airbnb.com/help/article/2509/in-what-areas-is-occupancy-tax-collection-and-remittance-by-airbnb-available.

months in a year. Specifically, we calculate the optimal number of days to be listed per month in the second stage. 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. To operationalize the second regulation, occupancy tax, we use a linear tax as a fixed percentage of the listing price. To account for tax pass-through, let p_{it}^{A} denote the listing price paid by consumers and $p_{it}^{A,host}$ denote the price received by hosts. The price paid by consumers p_{it}^{A} enters the hedonic regressions in Equations 11 and 12, whereas 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 $0 < t_1 < 1$. To operationalize the third regulation, convex tax, we use $p_{it}^{A,host} = p_{it}^{A} - t_2 \cdot (p_{it}^{A})^2$ such that the implied average tax rate as a fraction of the listing price $t_2 p_{it}^{A}$ increases with the listing price.

Long-term rental regulations: rent control. We focus on one type of regulation on long-term rental, rent control, which is commonly observed in practice and is intended to ensure rental housing affordability. Rent control is a system of laws placing a maximum price, or a "rent ceiling," on what 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, and 37 states prohibit or ban rent control outright.[42]

The rent control policy impacts hosts' incentives to rent and was implemented before Airbnb was introduced. The introduction of Airbnb further impacts hosts' incentives to rent and can interfere with the rent control policy. Economists have concluded that rent control policies 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 policies 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.

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

7.3.2 Short-Term Rental Regulations

Overall policy impact. Figure 9a shows the effect of short-term rental regulations by plotting the number of switchers (cannibalization) on the x-axis and the number of non-switchers (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 non-switchers (market expansion).

A desirable policy should reduce the negative impact of Airbnb (switcher or cannibalization) while maintaining the positive impact of Airbnb (non-switcher 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 listings (switchers and non-switchers).

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 examine how the policies differentially affect heterogeneous host groups. In particular, Airbnb provides hosts an alternative income source, which is especially valuable for less economically advantaged hosts. If the economically advantaged hosts earn more profits, the distribution of income among hosts will be more unequal and social inequality will be exacerbated. In practice, there have been continuing concerns that Airbnb exacerbates income disparity as the gains from Airbnb are disproportionately skewed to those with more wealth.[43] Imposing the regulations can induce a redistributive effect among hosts and affect income and social equality.

In addition to that defined by measure (1), a desirable policy should prevent the distribution of income among hosts from being skewed to those economically advantaged hosts who own expensive units and already have abundant resources. We define two additional measures of policy desirability:

(2) The fraction of total host profits earned by owners of luxury units.

(3) The fraction of total host profits earned by economically advantaged hosts.

[43] For instance, see https://www.usnews.com/news/cities/articles/2019-05-02/airbnbs-controversial-impact-on-cities and https://www.epi.org/publication/the-economic-costs-and-benefits-of-airbnb-no-reason-for-local-policymakers-to-let-airbnb-bypass-tax-or-regulatory-obligations/.

Figure 9: Short-Term Rental Regulations

(a) Number of Switchers and Non-switchers (b) Cannibalization Rate

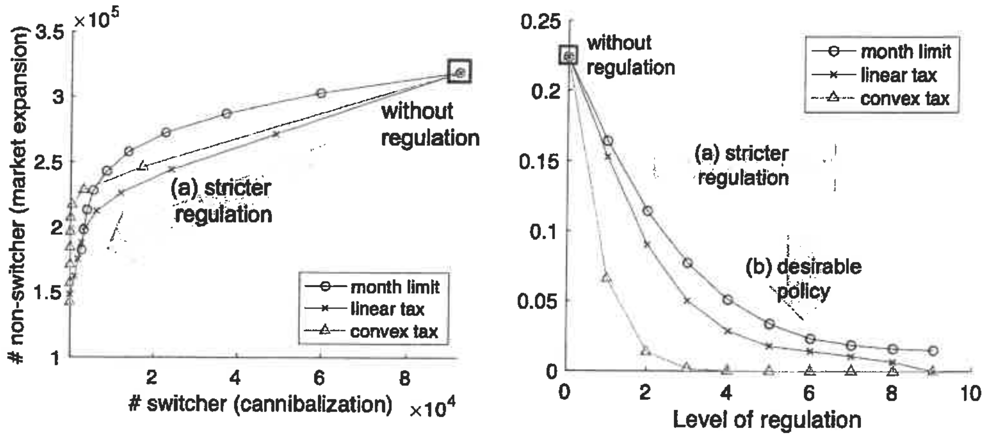
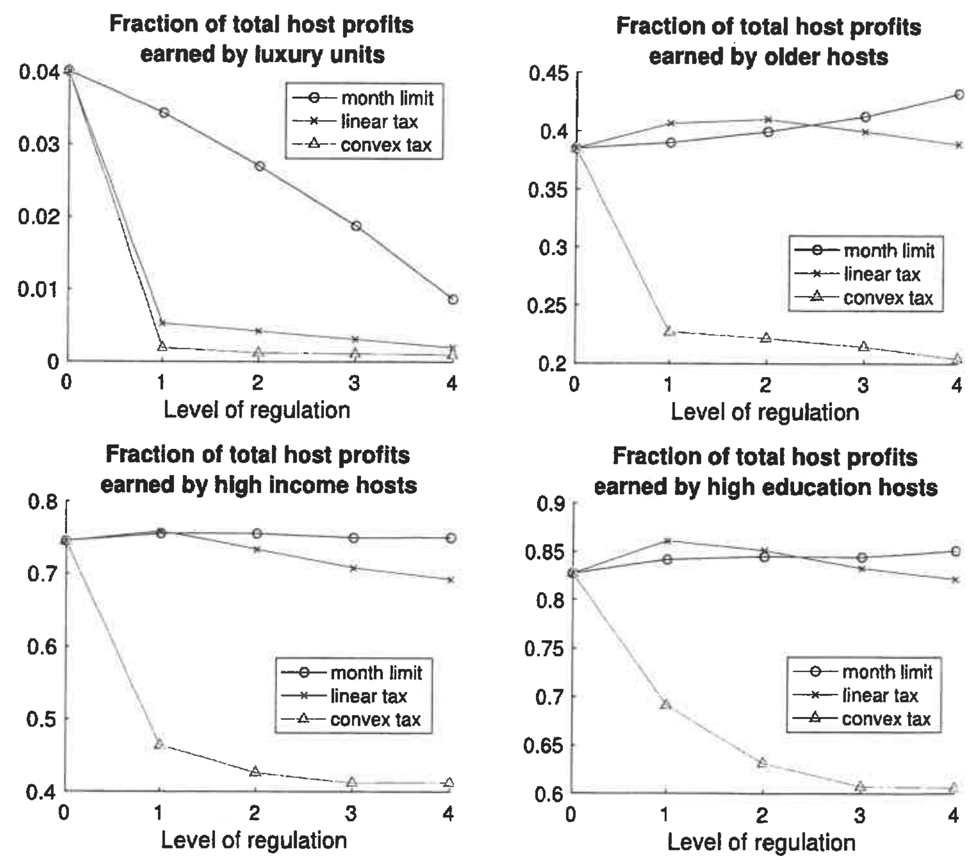


Figure 10: Short-Term Rental Regulations: Differential Impact on Hosts



Specifically, host profit equals (after-tax) revenue subtracts hosting cost. We calculate the host profit earned by each host and examine the fraction of total host profits earned by specific host groups. In Figure 10, we plot the fraction of total host profits earned by owners of luxury units (4 bedrooms or above), high-income hosts (income more than 100k), older hosts (age above 65), and high-education hosts (bachelor's degree or higher). We find that the convex tax again performs best in terms of having the smallest fraction of total host profits earned by owners with luxury units, high income, age, or education.

Overall, our proposed policy of a convex tax outperforms the other two policies in all three measures: (1) reducing the cannibalization rate, (2) reducing the fraction of total host profits earned by owners of luxury units, and (3) reducing the fraction of total host profits earned by economically advantaged hosts. The linear tax appears to perform better than the month limit. The convex tax performs best because the percentage of switchers is larger among higher priced luxury units than among lower priced affordable units. The convex tax discourages taking higher price properties off the long-term rental market, which helps limit cannibalization, but has less influence on lower priced properties, which helps maintain market expansion.

7.3.3 Long-Term Rental Regulations: Rent Control

To examine how Airbnb and rent control policies 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. The difference between a (c) and b (d) represents the negative impact of rent control in the absence (presence) of Airbnb. The difference between a (b) and c (d) represents the 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 Airbnb's presence 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 the rental supply. Importantly, this negative impact of rent control policy is exacerbated when Airbnb is available: the reduction in rental supply due to rent control is larger with Airbnb than that without Airbnb. This exacerbating effect is even more prominent with stricter rent control policies. 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 the long-term rental market and switch to Airbnb.

Second, we find that the presence of rent control can also amplify Airbnb's negative impact. Table 8 shows the percentage decrease in rental supply induced by Airbnb under varying strictness of rent control.

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

Level of Rent Control (r)	% Change in the Rental Supply Due to Rent Control without Airbnb	with Airbnb
5.0%	-0.60	-0.64
10.0%	-1.22	-1.29
15.0%	-1.85	-1.96
20.0%	-2.49	-2.63

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

Level of Rent Control (r)	None	5.0%	10.0%	15.0%	20.0%
% Change in the Rental Supply Due to Airbnb	-0.54	-0.58	-0.61	-0.65	-0.68

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 a rent control policy can each have a negative impact on the long-term rental supply. We find that when both are present, they can exacerbate each other's negative impact. Thus, policy makers must exercise caution when implementing rent control policies 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 of choosing among Airbnb, the long-term rental market, and the outside option and (2) the monthly decision on how many days to list on Airbnb if they choose 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 luxury units, suggesting that Airbnb can raise concerns about housing affordability. However, the market expansion effect is also greater 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 the 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 greater 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) reducing the fraction of total host profits earned by owners of luxury units, and (3) reducing the fraction of total host profits earned by economically advantaged hosts (e.g., high-income, older, or high-education hosts). Finally, rent control must be implemented with greater caution when Airbnb is available, as lower profits from long-term rentals can lead landlords to switch to Airbnb and exacerbate the side effect of a rent control policy.

This study has a few limitations that represent directions for future research. First, the set of policy desirability measures we examine cannot capture every aspect of the policy effects. We focus on the effects on hosts' decisions given that our data set allows us to model hosting behaviors. However, many other potential effects are important for policy makers but could not be addressed in this paper, for example, the cascading effects of hosts' behaviors, effects on renters, and long-term effects on new home purchases and construction. Exploring these effects offers promising directions for future research.

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 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 Airbnb's impact 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 competition between hotels and Airbnb. 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. (2021). The Sharing Economy and Housing Affordability: Evidence from Airbnb. *Marketing Science*, 40(1):23-47.

[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.

Table 9: Switching Patterns in the AHS Data Set

(a) Switching Patterns: Cell Percentages (%)

		2017 Airbnb	2017 Vacant	2017 Long-term rental	*Row Total*
2015	Airbnb	0.15	0.37	0.10	*0.63*
2015	Vacant	0.30	5.63	1.74	*7.67*
2015	Long-term rental	0.35	1.47	89.87	*91.70*
	Column Total	*0.80*	*7.48*	*91.72*	*100.00*

(b) Switching Patterns: Row Percentages (%)

		2017 Airbnb	2017 Vacant	2017 Long-term rental	*Row Total*
2015	Airbnb	24.19	59.56	16.25	*100.00*
2015	Vacant	3.86	73.46	22.68	*100.00*
2015	Long-term rental	0.39	1.61	98.01	*100.00*

(c) Switching Patterns: Column Percentages (%)

		2017 Airbnb	2017 Vacant	2017 Long-term rental
2015	Airbnb	18.92	4.99	0.11
2015	Vacant	36.87	75.30	1.90
2015	Long-term rental	44.21	19.71	97.99
	Column Total	*100.00*	*100.00*	*100.00*

Online Appendix

A. Evidence of Switching

As discussed in Footnote 11 of the paper, vacant units that are rented for "3 to 7 nights" and "8 or more nights" in the AHS data are potentially Airbnb listings and overlap with entire place listings on Airbnb. We flag these potential Airbnb listings in the AHS data set as choosing the Airbnb option in a year. The remaining vacant properties chose the vacant option. The long-term rental properties chose the long-term rental option. This means that we can observe properties that choose all three options (long-term rental, Airbnb, and vacant) within the AHS data set. Given the longitudinal nature of the AHS data set, we can present some switching patterns over time at the property level within the AHS data set.

Table 9a presents the switching patterns for properties that we observe data for two years. The rows represent the option that a property chose in 2015 and the columns represent the option that the same property chose in 2017. The number in each cell represents the percentage of properties that have the corresponding option combination. For instance, the first cell suggests that 0.15% of the properties chose Airbnb in 2015 and chose Airbnb in 2017. The percentages of all cells add up to 100%. The table shows that there were switching behaviors among all three options, suggesting that switching exists in our data set.

Besides presenting the switching patterns in terms of the cell percentages in Table 9a, we also present the switching patterns in terms of the row percentages in Table 9b and the column percentages in Table 9c. The row percentages of cells in the same row add up to 100%. The column percentages of cells in the same column add up to 100%. The row and column percentages are useful because these switching patterns in the data can relate to the two measures of switching behaviors we use in Section 7.2 of the paper:

(1) Percentage of reduction in rental supply due to Airbnb. This measure corresponds to the row percentage of the cell "2015 long-term rental -> 2017 Airbnb" in Table 9b because it represents the percentage of long-term rental units that switched to Airbnb. We find that this number is 0.39% in the table. This suggests that the percentage of rental reduction due to Airbnb is -0.39% in the data, which is comparable to the model-predicted ones in the paper (between -0.16% and -1.46%).

(2) Cannibalization rate or the fraction of Airbnb listings that come from switching from long-term rental. This measure corresponds to the column percentage of the cell "2015 long-term rental -> 2017 Airbnb" in Table 9c because it represents the percentage of Airbnb units that come from switching from long-term rental. We find that this number is 44.21% in the data, which is comparable to the model-predicted ones in the paper (between 10.4% and 48.8%).

The fact that the data-generated switching patterns are comparable to our model-predicted ones boosts our confidence that our model is able to obtain reasonable estimates of cannibalization effect and switching. It serves as a validity check on the effect magnitudes we obtain from the model. One caveat of the data-generated switching patterns is that they only reflect switching to/from Airbnb entire place listings and do not include Airbnb private room listings. The reason is that the approach we use to identify potential Airbnb units in AHS can only identify entire place listings as discussed in Footnote 11. Given that the data patterns only include part of Airbnb listings, the implied relative sizes of switchers and non-switchers may be inaccurate. However, the absolute number of switchers is accurate: while non-switchers can be either entire place listings or private room listings, switchers can only be entire place listings (as private room listings cannot choose long-term rental and would not be switchers from long-term rental). The data patterns include all entire place listings. Therefore, they include all the switchers and serve the goal of showing that switching exists in the data, which helps the identification of switching.

B. 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

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

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

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

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 $s_{it}^{*}=s\ (s=1,2,...,\bar{s}-1)$ 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 $s_{it}^* = s$ $(s = 1, 2, ..., \bar{s} - 1)$.

C. Hedonic Regression Results

Table 10 shows the hedonic regression results for rent and rental occupancy rate. Table 11 shows the hedonic regression results for Airbnb price and occupancy rate. The error terms are clustered at the metro area level. The R-squared value for Airbnb occupancy rate is relatively low, which may be due to a large variation in the 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 making first-stage decisions.

Table 10: Hedonic Regression: Rent and Rental Occupancy

DV:	Rent		Rental Occupancy	
Constant	1,186***	(42.36)	0.848***	(0.137)
Rental Supply	-100.6***	(29.81)	-0.00182***	(0.00060)
Rent	–		-0.000178***	(0.000015)
Metro Area - Year FE	Yes		Yes	
Demographics				
Age				
35-65	-55.88***	(9.402)	-0.0346***	(0.00731)
Over 65	-147.0***	(15.14)	0.0733***	(0.0178)
Education				
High School Grad	-31.83***	(11.05)	0.224***	(0.00549)
Bachelor's	155.3***	(12.96)	0.275***	(0.0185)
Marital Status				
Never Married	-22.79**	(10.71)	0.112***	(0.00476)
Married Now	-33.68***	(10.65)	0.118***	(0.00553)
Gender				
Male	-16.61**	(8.181)	0.0483***	(0.00358)
Race				
Black	-184.2***	(10.06)	0.00603	(0.0215)
Other	34.96***	(11.77)	-0.254***	(0.00593)
Origin				
Hispanic	-118.9***	(10.13)	0.0661***	(0.0142)
Household Income				
50K-100K	121.1***	(9.705)	0.0444***	(0.0144)
Over 100K	366.2***	(12.82)	0.0776*	(0.0423)
Property Characteristics				
# of Bedrooms				
1	13.28	(30.21)	-0.00994	(0.0113)
2	72.94**	(33.65)	0.0166	(0.0150)
3	114.0***	(34.60)	0.0545***	(0.0183)
4	80.86*	(42.66)	0.0519***	(0.0183)
5+	122.0**	(58.12)	0.0746***	(0.0257)
# of Bathrooms	65.85***	(5.018)	0.00590	(0.00779)
# of Rooms	19.23***	(6.078)	-0.00495	(0.00316)
# of Amenities	37.47***	(3.565)	0.0132***	(0.00450)
Property Type				
House	-131.5***	(12.52)	-0.0176	(0.0158)
Other	-478.6***	(49.28)	-0.0337	(0.0580)
# of Units in the Structure				
2	-145.5***	(15.91)	-0.0306*	(0.0177)
3-4	-111.3***	(14.38)	-0.0118	(0.0139)
5-9	-97.88***	(13.30)	-0.00866	(0.0123)
10+	-138.1***	(13.23)	-0.0116	(0.0166)
Unit Age	-6.892***	(0.608)	0.000196	(0.000823)
Unit Age Squared	0.0507***	(0.00532)	-3.35e-06	(6.15e-06)
N	15,670		15,670	
R^2	0.396		0.578	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. The rental supply is in million units. The baseline demographics group is age under 35, education below high school, income below 50k, female, white, non-Hispanic origin, and widowed/divorced/separated.

Table 11: Hedonic Regression: Airbnb Price and Occupancy Rate

DV:	Logged Airbnb Price		Airbnb Occupancy	
Constant	4.844***	(0.0146)	2.645***	(0.120)
Airbnb Supply	-0.0196***	(0.00166)	-0.0207***	(0.00147)
Logged Airbnb Price	–		-0.425***	(0.0247)
Metro Area - Year FE	Yes		Yes	
Metro Area - Month FE	Yes		Yes	
Demographics				
Age	0.00862***	(8.51e-05)	0.00363***	(0.000225)
Household Income	0.00195***	(1.21e-05)	0.000706***	(4.93e-05)
Education				
Bachelor's	-0.135***	(0.00591)	-0.306***	(0.00594)
High School Grad	-0.438***	(0.00761)	-0.654***	(0.0125)
Marital Status				
Married Now	-1.308***	(0.00952)	-0.702***	(0.0333)
Never Married	-0.185***	(0.0103)	-0.0188*	(0.00969)
Gender				
Male	0.314***	(0.00628)	0.0813***	(0.00935)
Race				
Black	-0.364***	(0.00235)	-0.129***	(0.00922)
Other	-0.0924***	(0.00282)	-0.0697***	(0.00328)
Origin				
Hispanic	-0.241***	(0.00242)	-0.0873***	(0.00629)
Property Characteristics				
# of Bedrooms				
1	0.118***	(0.00129)	0.00443	(0.00311)
2	0.375***	(0.00114)	0.116***	(0.00933)
3	0.578***	(0.00145)	0.207***	(0.0143)
4	0.733***	(0.00207)	0.284***	(0.0182)
5+	0.692***	(0.00297)	0.252***	(0.0173)
# of Bathrooms	0.0965***	(0.000513)	0.00371	(0.00243)
# of Amenities	0.0132***	(0.000170)	0.0271***	(0.000356)
Property Type				
House	0.0201***	(0.000775)	0.0329***	(0.000815)
Other	0.0477***	(0.000678)	0.0764***	(0.00131)
Room Type				
Private/Shared	-0.515***	(0.000645)	-0.290***	(0.0127)
Airbnb-related metro variables				
Airbnb history (months)	0.00207***	(0.000340)	0.000610**	(0.000288)
Air passengers (in millions)	-0.0191***	(0.00536)	0.0531***	(0.00448)
Google search trend	0.00107***	(8.89e-05)	-0.00349***	(7.85e-05)
N	2,917,491		2,917,491	
R^2	0.525		0.139	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are in parentheses. The baseline demographics group is education below high school, female, white, non-Hispanic origin, and widowed/divorced/separated.