The Sharing Economy and Housing Affordability: Evidence from Airbnb

Kyle Barron* Edward Kung† Davide Proserpio‡

October 20, 2017

Abstract

We assess the impact of home-sharing on residential house prices and rents. Using a dataset of Airbnb listings from the entire United States and an instrumental variables estimation strategy, we find that a 10% increase in Airbnb listings leads to a 0.42% increase in rents and a 0.76% increase in house prices. The effect is larger in zipcodes with a smaller share of owner-occupiers, a result consistent with absentee landlords reallocating their homes from the long-term rental market to the short-term rental market. A simple model rationalizes these findings.

Keywords: Sharing economy, peer-to-peer markets, housing markets, Airbnb
JEL Codes: R31, L86

*NBER; barronk@nber.org.
†Department of Economics, UCLA; ekung@econ.ucla.edu.
‡Marshall School of Business, USC; proserpi@marshall.usc.edu.

1 Introduction

Peer-to-peer markets, also referred to as the sharing economy, are online marketplaces that facilitate matching between demanders and suppliers of various goods and services. The suppliers in peer-to-peer markets are often small (mostly individuals), and they supply excess capacity that might otherwise go unutilized—hence the term "sharing economy." Proponents of the sharing economy argue that it improves economic efficiency by reducing frictions that cause capacity to go underutilized, and the explosive growth of sharing platforms (such as Uber for ride-sharing and Airbnb for home-sharing) testifies to the underlying demand for such markets.[1] Critics argue, however, that much of the growth in the sharing economy has come from skirting regulations. For example, traditional taxi drivers face more stringent regulations than Uber drivers, and traditional providers of short-term rentals (i.e., hotels, beds & breakfasts) are required to pay occupancy tax while Airbnb hosts usually are not.[2]

Beyond regulatory avoidance, home-sharing in particular has been subject to an additional source of criticism. Namely, critics argue that home-sharing platforms like Airbnb raise the cost of living for local renters, while mainly benefitting local landlords and non-resident tourists.[3] It is easy to see the economic argument. By reducing frictions in the peer-to-peer market for short-term rentals, home-sharing platforms cause some landlords to switch from supplying the market for long-term rentals—in which residents are more likely

[1] These frictions could include search frictions in matching demanders with suppliers, and information frictions associated with the quality of the good being transacted, or with the trustworthiness of the buyer or seller. See Einav et al. (2016) for an overview of the economics of peer-to-peer markets, including the specific technological innovations that have facilitated their growth.

[2] Some cities have passed laws requiring Airbnb hosts to pay occupancy tax. Enforcement, however, is difficult because there are often no systems in place for the government to keep track of who is renting on Airbnb. A key area of contention is whether Airbnb should be required to collect occupancy tax from its hosts. See The New York Times, "Lodging Taxes and Airbnb Hosts: Who Pays, and How," June 16, 2015.

[3] Another criticism of Airbnb is that the company does not do enough to combat racial discrimination on its platform (see Edelman and Luca (2014); Edelman et al. (2017)), though we will not address this issue in this paper.

to participate—to supplying the short-term market—in which non-residents are more likely to participate. Because the total supply of housing is fixed in the short run, this drives up the rental rate in the long-term market. Concern over home-sharing's impact on housing affordability has garnered significant attention from policymakers, and has motivated many cities to impose stricter regulations on home-sharing.[4]

Whether or not home-sharing increases housing costs for local residents is an empirical question. There are a few reasons why it might not. First, the market for short-term rentals may be very small compared to the market for long-term rentals. In this case, even large changes to the short-term market might not have a measurable effect on the long-term market. The short-term market could be small—even if the short-term rental rate is high relative to the long-term rate—if landlords prefer more reliable long-term tenants and a more stable income stream.

Second, the market for short-term rentals could be dominated by housing units that would have remained vacant in the absence of home-sharing. Owner-occupiers, those who own the home in which they live, may supply the short-term rental market with their spare rooms and cohabit with guests, or may supply their entire apartment during a host's vacation. These otherwise vacant rentals could also be vacation homes that would not be rented to long-term tenants because of the restrictiveness of long-term leases. In either case, such owners would not make their homes available to long-term tenants, independently of the existence of a convenient home-sharing platform. Instead, home-sharing provides them with an income stream for times when their housing capacity would otherwise be underutilized.

In this paper, we study the effect of home-sharing on the long-term rental market using data collected from Airbnb, the world's largest home-sharing

[4] For example, Santa Monica outlaws short-term, non-owner-occupied rentals of less than 30 days, as does New York State for apartments in buildings with three or more residences. San Francisco passed a 60-day annual hard cap on short-term rentals (which was subsequently vetoed by the mayor). It is unclear, however, the degree to which these regulations are enforced. We are aware of only one successful prosecution of an Airbnb host, occurring in Santa Monica in July 2016.

platform. We first develop a simple model of house prices and rental rates when landlords can choose to allocate housing between long-term residents and short-term visitors. The effect of a home-sharing platform such as Airbnb is to reduce the frictions associated with renting on the short-term market. From the model we derive three testable predictions: 1) Airbnb increases both rental rates and house prices in the long-term market; 2) the increase in house prices is greater than the increase in rental rates, thus leading to an increase in the price-to-rent ratio; and 3) the effect on rental rates is smaller when a greater share of the landlords are owner-occupiers. Intuitively, the owner-occupancy rate matters because only non-owner-occupiers are on the margin of substituting their housing units between the long and short-term rental markets. Owner-occupiers interact with the short-term market only to rent out unused rooms or to rent while away on vacation, but they do not allocate their housing to long-term tenants.

To test the model, we collect primary data sources from Airbnb, Zillow, and the Census Bureau. We construct a panel dataset of Airbnb listings at the zipcode-year-month level from data collected from public-facing pages on the Airbnb website between mid-2012 to the end of 2016, covering the entire United States. From Zillow, a website specializing in residential real estate transactions, we obtain a panel of house price and rental rate indices, also at the zipcode-year-month level. Zillow provides a platform for matching landlords with long-term tenants, and thus their price measures reflect sale prices and rental rates in the market for long-term housing. Finally, we supplement this data with a rich set of time-varying zipcode characteristics collected from the Census Bureau's American Community Survey (ACS), such as the median household income, population count, share of college graduates, and employment rate.

In the raw correlations, we find that the number of Airbnb listings in zipcode i in year-month t is positively associated with both house prices and rental rates. In a baseline OLS regression with no controls, we find that a 10% increase in Airbnb listings is associated with a 0.84% increase in rental rates and a 1.57% increase in house prices. Of course, these estimates should not be interpreted as causal, and may instead be picking up spurious correlations. For example, cities that are growing in population likely have rising rents, house prices, and numbers of Airbnb listings at the same time. We therefore exploit the panel nature of our dataset to control for unobserved zipcode level effects and arbitrary city level time trends. We include zipcode fixed effects to absorb any permanent differences between zipcodes, while fixed effects at the Core Based Statistical Area (CBSA)-year-month level control for any shocks to housing market conditions that are common across zipcodes within a CBSA.[5]

We further control for unobserved *zipcode-specific, time-varying* factors using an instrumental variable that is plausibly exogenous to local zipcode level shocks to the housing market. To construct the instrument, we exploit the fact that Airbnb is a young company that has experienced explosive growth over the past five years. Figure 1 shows worldwide Google search interest in Airbnb from 2008 to 2016. Demand fundamentals for short-term housing are unlikely to have changed so drastically from 2008 to 2016 as to fully explain the spike in interest, so most of the growth in Airbnb search interest is likely driven by information diffusion and technological improvements to Airbnb's platform as it matures as a company. Neither of these should be correlated with local zipcode level unobserved shocks to the housing market. By itself, global search interest is not enough for an instrument because we already control for arbitrary CBSA level time trends. We therefore interact the Google search index for Airbnb with a measure of how "touristy" a zipcode is in a base year, 2010. We define "touristy" to be a measure of a zipcode's attractiveness for tourists and proxy for it using the number of establishments in the food service and accommodations industry.[6] These include eating and drinking places, as well as hotels, bed and breakfasts, and other forms of short-term lodging. The identifying assumptions of our specification are that: 1) landlords in more touristy zipcodes are more (or less) likely to switch into the short-term rental

[5]The CBSA is a geographic unit defined by the U.S. Office of Management and Budget that roughly corresponds to an urban center and the counties that commute to it.

[6]We focus on tourism because Airbnb has historically been frequented more by tourists than business travelers. Airbnb has said that 90% of its customers are vacationers, but is attempting to gain market share in the business travel sector.

market in response to learning about Airbnb than landlords in less touristy zipcodes;[7] and 2) ex-ante levels of touristiness are not systematically correlated with ex-post unobserved shocks to the housing market at the zipcode level *that are also correlated in time with Google search interest for Airbnb*. We discuss the instrument in more detail in Sections 4 and 5.3. Using this instrumental variable, we estimate that a 10% increase in Airbnb listings leads to a 0.42% increase in the rental rate and a 0.76% increase in house prices. These results are consistent with our model's predictions that the effects on both rental rate and house prices will be positive, and that the effect on house prices will be larger.

The model also predicts that the effect of Airbnb will be smaller if the market has a large share of owner-occupiers. To test this, we repeat the above regressions while allowing for the effect of Airbnb to depend on the share of owner-occupiers in the zipcode. We find that the owner-occupancy rate significantly moderates the effect of Airbnb on the market for long-term housing. Going from a zipcode that is in the 25th percentile of owner-occupancy rate to a zipcode that is in the 75th percentile of owner-occupancy rate causes the rental rate impact of a 10% increase in Airbnb listings to go from 0.29% to 0.21%. We find similar results for house prices. These results are consistent with the model and suggest that Airbnb's impact on the long-term market depends on the number of landlords who are on the margin of switching between allocating their housing to long-term tenants versus short-term visitors.

Finally, we consider the effect of Airbnb on housing vacancy rates. Because zipcode level data on vacancies are not available at a monthly—or even yearly—frequency, we focus on annual vacancy rates at the CBSA level. We find that annual CBSA vacancy rates have no association with the number of Airbnb listings. However, looking at type of vacancy we find that the number of Airbnb listings is positively associated with the share of homes that are vacant for seasonal or recreational use and negatively associated with the share of homes that are vacant-for-rent and vacant-for-sale. This is consistent with

[7] Landlords in more touristy zipcodes could be less likely to switch if competition from incumbent hotels is more fierce.

absentee landlords substituting away from the rental and for-sale markets for long-term residents, and towards the short-term market, which are likely then categorized as vacant-seasonal homes.[8]

Related literature

We are aware of only two other academic papers to directly study the effect of home-sharing on housing costs. Lee (2016) provides a descriptive analysis of Airbnb in the Los Angeles housing market, while Horna and Merantea (2017) use Airbnb listings data from Boston in 2015 and 2016 to study the effect of Airbnb on rental rates. They find that a one standard deviation increase in Airbnb density at the census-tract-month level is associated with a 0.4% increase in the rental rate. Our estimates are not directly comparable because we use different regressors, datasets, time periods, and geographic levels, but the estimates appear to be similar. For example, in our preferred specification, we find that one standard deviation of higher growth in Airbnb listings leads to a 0.65% increase in rental rates.

We contribute to the literature concerning the effect of home-sharing on housing costs in three ways. First, we present a model that organizes our thinking about how home-sharing is expected to affect housing costs in the long-term market. Second, we provide direct evidence for the model's predictions, highlighting especially the role of the owner-occupancy rate and of the marginal landowner. Third, we present the first estimates of the effect of home-sharing on housing costs that uses comprehensive data from across the U.S.

Our paper also contributes to the more general literature on peer-to-peer markets. One part of this literature has focused on the effect of the sharing economy on the labor market outcomes of the suppliers.[9] Another part of this literature focuses on the competition between traditional suppliers and the

[8] Census Bureau methodology classifies a housing unit as vacant even if it is temporarily occupied by persons who usually live elsewhere.

[9] See Hall and Krueger (2017) and Chen et al. (2017) for studies on the incomes and labor market outcomes for Uber drivers.

small suppliers that are enabled by sharing platforms.[10] In terms of studies on Airbnb, Zervas et al. (2017) estimate the impact of the sharing economy on hotel revenues. Our paper looks at a somewhat unique context in this literature, because we focus on the effect of the sharing economy on the reallocation of goods from one purpose to another, which may cause local externalities. Local externalities are present here because the suppliers are local and the demanders are non-local; transactions in the home-sharing market therefore involve a reallocation of resources from locals to non-locals. Our contribution is therefore to study this unique type of sharing economy in which public policy may be especially salient.

The rest of the paper is organized as follows. In Section 2, we present a simple model of house prices and rental rates where landlords can substitute between supplying the long-term and the short-term market. In Section 3, we describe the data we collected from Airbnb and present some basic statistics. In Section 4, we describe our methodology, and in Section 5 we discuss the results. Section 6 concludes.

2 Model

2.1 Basic setup

We consider a housing market with a fixed stock of housing H, which can be allocated to short-term housing S, or long-term housing L. $S + L = H$. The rental rate of short-term housing is Q and the rental rate of long-term housing is R. The two housing markets are segmented—tenants who need long-term housing cannot rent in the short-term market and tenants who need short-term housing cannot rent in the long-term market.[11]

[10]See Einav et al. (2016) for an overview of the economics of peer-to-peer markets. Horton and Zeckhauser (2016) study the effects of the sharing economy on decisions to own the underlying goods, and Gong et al. (2017) study the impact of Uber on car sales.

[11]In our view, the primary driver of this market segmentation is the length of lease and tenant rights. Local residents participating in the long-term rental market will typically sign leases of 6 months to a year, and are also granted certain rights and protections by the city. On the other hand, non-resident visitors participating in the short-term market will

For now, we assume that all housing is owned by absentee landlords and will return to the possibility of owner-occupiers later. Each landlord owns one unit of housing and decides to rent it on the short-term market or the long-term market, taking rental rates as given. A landlord will rent on the short-term market if $Q - c - \epsilon > R$, where $c + \epsilon$ is an additional cost of renting on the short-term market, with c being a common component and ϵ being an idiosyncratic component across landlords.[12] The share of landlords renting in the short-term market is therefore:

$$f(Q - R - c) = P(\epsilon < Q - R - c) \tag{1}$$

f is the cumulative distribution function of ϵ, and $f' > 0$. The total number of housing units in the short-term market are:

$$S = f(Q - R - c)H \tag{2}$$

Long-term rental rates are determined in equilibrium by the inverse demand function of long-term tenants:

$$R = r(L) \tag{3}$$

with $r' < 0$. Short-term rental rates are determined exogenously by outside markets.[13] The market is in steady state, so the house price P is equal to the

usually only rent for a few days and are not granted the same rights as resident tenants.

[12]Renting in the short-term market could be costlier than in the long-term market because the technology for matching landlords with tenants may be historically more developed in the long-term market. Landlords may have idiosyncratic preferences over renting in the long-term market vs. the short-term market if they have different preferences for the stability provided by long-term tenants.

[13]For example, they could be determined by elastic tourism demand. Relaxing this assumption and allowing for price elasticity in the short-term market would not change the qualitative results.

present value of discounted cash flows to the landlord:

$$P = \sum_{t=0}^{\infty} \delta^t E\left[R + \max\{0, Q - R - c - \epsilon\}\right]$$
$$= \frac{1}{1-\delta}\left[R + g(Q - R - c)\right] \tag{4}$$

where $g(x) = E[x - \epsilon | \epsilon < x] f(x)$ gives the expected net surplus of being able to rent in the short-term market relative to the long-term market, and $g' > 0$.

2.2 The effect of home-sharing

The introduction of a home-sharing platform reduces the cost for landlords to advertise on the short-term market, implying a decline in c. This could happen for a variety of reasons. By improving the search and matching technology in the short-term market, the sharing platform may reduce the time it takes to find short-term tenants. By providing identity verification and a reputation system for user feedback, the platform may also help reduce information costs.

We consider how an exogenous change to the cost of listing in the short-term market, c, affects long-term rental rates and house prices. Equilibrium conditions (1)-(3) imply that:

$$\frac{dR}{dc} = \frac{r' f' H}{1 - r' f' H} < 0 \tag{5}$$

So, by decreasing the cost of listing in the short-term market, the home-sharing platform has the effect of raising rental rates. The intuition is fairly straightforward: the home-sharing platform induces some landlords to switch from the long-term market to the short-term market, reducing supply in the long-term market and raising rental rates.

For house prices, we can use equation (4) to write:

$$\frac{dP}{dc} = \frac{1}{1-\delta}\left[\frac{dR}{dc} - \left(1 + \frac{dR}{dc}\right) g'\right] \tag{6}$$

We note from equation (5) that $-1 < \frac{dR}{dc} < 0$, and so $\frac{dP}{dc} < \frac{1}{1-\delta}\frac{dR}{dc} < \frac{dR}{dc} < 0$.

The latter inequality concludes that home-sharing increases house prices and that the house price response will be greater than the rental rate response. This is because home-sharing increases the value of homeownership through two channels. First, it raises the rental rate which is then capitalized into house prices. Yet if this were home-sharing's only effect, then the price response and the rental rate response would be proportional by the discount factor. Instead, the additional increase in the value of homeownership comes from the enhanced option value of renting in the short-term market. Because of this second channel, prices will respond even more than rental rates to the introduction of a home-sharing platform.

2.3 Owner-occupiers

We now relax the assumption that all homeowners are absentee landlords by also allowing for owner-occupiers. Let H_a be the number of housing units owned by absentee landlords and let H_o be the number of housing units owned by owner-occupiers. We still define L as the number of housing units allocated to long-term residents—including owner-occupiers—and therefore the number of renters is $L - H_o$. We assume that H_a is fixed, and that H_o will be determined by equilibrium house prices and rental rates.[14]

We allow owner-occupiers to interact with the short-term housing market by assuming that a fraction γ of their housing unit is excess capacity. This excess capacity can be thought of as the unit's spare rooms or the time that the owner spends away from his or her home. Owner-occupiers have the choice to either hold their excess capacity vacant, or to rent it out on the short-term market. They cannot rent excess capacity on the long-term market, due to the nature of leases and renter protections. The benefit to renting excess capacity on the short-term market is $Q - c - \epsilon$, where c and ϵ are again the cost and the idiosyncratic preference for listing on the short-term market, respectively. If

[14]If H_a is not fixed, then all of the housing stock will be owned by either absentee landlords or owner-occupiers, depending on which has the higher net present value of owning. In the Appendix, we numerically solve a model with heterogeneous agents which allows for an endogenous share of absentee landlords, and show that the qualitative results of this section still hold.

excess capacity remains unused, the owner neither pays a cost nor derives any benefit from the excess capacity. Owner-occupiers will rent on the short-term market if $Q - c - \epsilon > 0$, and thus $f(Q - c)$ is the share of owner-occupiers who rent their excess capacity on the short-term market.

Note that the choice of the owner-occupier is to either rent on the short-term market, or to hold excess capacity vacant. Thus, participation in the short-term market by owner-occupiers does not change the overall supply of housing allocated to the long-term market, L. It also does not change S, which is by definition equal to $H - L$ (we think of S as the number of units that are *permanently* allocated towards short-term housing, as determined by absentee landlords.) The equilibrium supply of short and long-term housing are therefore:

$$S = f(Q - R - c)H_a \tag{7}$$

$$L = H - f(Q - R - c)H_a \tag{8}$$

Rental rates in the long-term market continue to be determined by the inverse demand curve of residents, $r(L)$. The equilibrium response of rental rates to a change in c becomes:

$$\frac{dR}{dc} = \frac{r' f' H_a}{1 - r' f' H_a} \leq 0 \tag{9}$$

Equation (9) is similar to equation (5) except that H is replaced with H_a. Equation (9) therefore makes clear that it is the absentee landlords who affect the rental rate response to Airbnb, because it is they who are on the margin between substituting their units between the short and long-term markets. When the share of owner-occupiers is high, the rental rate response to Airbnb will be low. In fact, the response of rental rates to Airbnb could be zero if all landlords are owner-occupiers.

Since long-term residents are ex-ante homogeneous, an equilibrium with a positive share of both renters and owner-occupiers requires that house prices make residents indifferent between renting and owning:

$$P = \frac{1}{1-\delta}\left[R + \gamma g(Q - c)\right] \tag{10}$$

Equation (10) says that the price that residents are willing to pay for a home is equal to the present value of long-term rents plus the present value of renting excess capacity to the short-term market. The response of prices to a change in c is:

$$\frac{dP}{dc} = \frac{1}{1-\delta}\left[\frac{dR}{dc} - \gamma g'\right] \tag{11}$$

So, again, we see that prices are more responsive to a decrease in c than rental rates.

To summarize the results of this section, we derived three testable implications. First, rental rates should increase in response to the introduction of a home-sharing platform. This is because home-sharing causes some landowners to substitute away from supplying the long-term rental market and into the short-term rental market. Second, house prices should increase as well, but by an even greater amount than rents. This is because home-sharing affects house prices through two channels: first by increasing the rental rate, which then gets capitalized into house prices, and second by directly increasing the ability for landlords to utilize the home fully. Finally, the rental rate response will be smaller when there is a greater share of owner-occupiers. This is because owner-occupiers are not on the margin of substituting between the long-term and short-term markets, whereas absentee landlords are.[15] We now turn to testing these predictions in the data.

[15] Another class of homeowners we have yet to discuss is vacation-home owners. Owners of vacation homes can be treated either as owner-occupiers with high γ (here γ is the amount of time spent living in their primary residence), or as absentee landlords, depending on how elastic they are with respect to keeping the home as a vacation property vs. renting it to a long-term tenant. In either case, the key implications of the model will not change.

3 Data and Background on Airbnb

3.1 Background on Airbnb

Recognized by most as the pioneer of the sharing economy, Airbnb is a peer-to-peer marketplace for short-term rentals, where the suppliers (hosts) offer different kinds of accommodations (i.e. shared rooms, entire homes, or even yurts and treehouses) to prospective renters (guests). Airbnb was founded in 2008 and has experienced dramatic growth, going from just a few hundred hosts in 2008 to over three million properties supplied by over one million hosts in 150,000 cities and 52 countries in 2017. Over 130 million guests have used Airbnb, and with a market valuation of over $31B, Airbnb is one of the world's largest accommodation brands.

3.2 Airbnb listings data

Our main source of data comes directly from the Airbnb website. We collected consumer-facing information about the complete set of Airbnb properties located in the United States and about the hosts who offer them. The data collection process spanned a period of approximately five years, from mid-2012 to the end of 2016. Scrapes were performed at irregular intervals between 2012 to 2014, and at a weekly interval starting January 2015.

Our scraping algorithm collected all listing information viewable to users of the website, including the property location, the daily price, the average star rating, a list of photos, the guest capacity, the number of bedrooms and bathrooms, a list of amenities such as WiFi and air conditioning, etc., and the list of all reviews from guests who have stayed at the property.[16] Airbnb host information includes the host name and photograph, a brief profile description, and the year-month in which the user registered as a host on Airbnb.

Our final dataset contains detailed information about 1,097,697 listings and 682,803 hosts spanning a period of nine years, from 2008 to 2016. Because of

[16] Airbnb does not reveal the exact street address or coordinates of the property for privacy reasons; however, the listing's city, street, and zipcode correspond to the property's real location.

Airbnb's dominance in the home-sharing market, we believe that this data represents the most comprehensive picture of home-sharing in the U.S. ever constructed for independent research.[17]

3.3 Calculating the number of Airbnb listings, 2008-2016

Once we have collected the data, we have to define a measure of Airbnb supply. This task requires two choices: first, we need to choose the geographic granularity of our measure; second, we need to define the entry and exit dates of each listing to the Airbnb platform. Regarding the geographic aggregation, we conduct our main analysis at the zipcode level for a few reasons. First, it is the lowest level of geography for which we can reliably assign listings without error (other than user input error).[18] Second, neighborhoods are a natural unit of analysis for housing markets because there is significant heterogeneity in housing markets across neighborhoods within cities, but comparatively less heterogeneity within neighborhoods. Zipcodes will be our proxy for neighborhoods. Third, conducting the analysis at the zipcode level as opposed to the city level helps with identification. This is due to our ability to compare zipcodes within cities, thus controlling for any unobserved city level factors that may be unrelated to Airbnb but all affect neighborhoods within a city, such as a city-wide shock to labor productivity.

The second choice, how to determine the entry and exit date of each listing, comes less naturally. Unfortunately, due to a change in the way the scraping algorithm worked, our data does not allow us to identify instantaneous counts of listings until 2015.[19] Prior to 2015, if a unique listing identifier appeared

[17]To verify the accuracy of our data, we cross-checked our data with data scraped by the website Insideairbnb.com. We discuss this further in the Appendix.

[18]Airbnb does report the latitude and longitude of each property, but only up to a perturbation of a few hundred meters. So it would be possible, but complicated, to aggregate the listings to finer geographies with some error.

[19]Estimating the number of active listings is a challenge even for Airbnb. Despite the fact that Airbnb offers an easy way to unlist properties, many times hosts neglect to do so, creating "stale vacancies" that seem available for rent but in actuality are not. Fradkin (2015), using proprietary data from Airbnb, estimates that between 21% to 32% of guest

multiple times, the scraper would replace a listing's previous data with information from the most recent scrape. As of 2015, each new scrape was kept as a separate record by including a timestamp associated with it. Thus, starting with 2015 we are able to see instantaneous snapshots of all Airbnb listings in the United States at a weekly frequency. Prior to 2015, we only see the latest information collected for any one listing. Thus, to construct the number of listings going back in time, we employ a variety of methods, summarized in Table 1.

Table 1: Methods for Computing the Number of Listings

	Listing is considered active ...
Method 1	starting from host join date
Method 2	for 3 months after host join date, and after every guest review
Method 3	for 6 months after host join date, and after every guest review
Method 4	whenever it is discovered in a weekly scrape

Methods 1-3 follow Zervas et al. (2017). Method 1 is our preferred choice to measure Airbnb supply and will be our main independent variable in all the analyses presented in this paper. This measure computes a listing's entry date as the date its host registered on Airbnb and assumes that listings never exit. The advantage of using the host join date as the entry date is that for a majority of listings, this is the most accurate measure of when the listing was first posted. The disadvantage of this measure is that it is likely to overestimate the listings that are available on Airbnb (and accepting reservations) at any point it time. However, as discussed in Zervas et al. (2017), such overestimation would cause biases only if, after controlling for several zipcode characteristics, it is correlated with the error term.

Aware of the fact that method 1 is an imperfect measure of Airbnb supply, we also experiment with alternative definitions of Airbnb listings' entry and exit. Methods 2 and 3 exploit our knowledge of each listing's review dates to

requests are rejected due to this effect.

determine whether a listing is active. The heuristic we use is as follows: a listing enters the market when the host registers with Airbnb and stays active for m months. We refer to m as the listing's Time To Live (TTL). Each time a listing is reviewed the TTL is extended by m months from the review date. If a listing exceeds the TTL without any reviews, it is considered inactive. A listing becomes active again if it receives a new review. In our analysis, we test two different TTLs, 3 months and 6 months.

Finally, method 4 exploits the weekly Airbnb scrapes. The weekly scrapes obviate the need to compute listings' entry or exit dates; instead, we consider a listing active in a given month if it appears in any of that month's scrapes. The advantage of this approach is that it is the most accurate measure of point-in-time listing counts. The disadvantage is that it is only available starting in January 2015.

Despite the fact that our different measures of Airbnb supply rely on different heuristics and data, because of Airbnb's tremendous growth, all our measures of Airbnb supply are extremely correlated. The correlation between method 1 and each other measure is above 0.95 in all cases. In the Appendix, we present robustness checks of our main results to the different measures of Airbnb supply discussed above, and show that results are qualitatively unchanged.

3.4 Zillow: rental rates and house prices

Zillow.com is an online real estate company that provides estimates of house and rental prices for over 110 million homes across the U.S. In addition to giving value estimates of homes, Zillow provides a set of indexes that track and predict home values and rental prices at a monthly level and at different geographical granularities.

For house prices, we use the Zillow Home Value Index (ZHVI) which estimates the median transaction price for the actual stock of homes in a given geographic unit and point in time. The advantage of using the ZHVI is that it is available at the zipcode-month level for over 13,000 zipcodes.

For rental rates, we use the Zillow Rent Index (ZRI). Like the ZHVI, Zillow's rent index is meant to reflect the median monthly rental rate for the actual stock of homes in a geographic unit and point in time. Crucially, Zillow's rent index is based on rental *list prices* and is therefore a measure of prevailing rents for new tenants. This is the relevant comparison for a homeowner deciding whether to place her unit on the short-term or long-term market. Moreover, because Zillow is not considered a platform for finding short-term housing, the ZRI should be reflective of rental prices in the long-term market.

3.5 Other data sources

We supplement the above data with several additional sources. We use monthly Google Trends data for the search term "airbnb", which we download directly from Google. This index measures how often people worldwide search for the term "airbnb" on Google, and is normalized to have a value of 100 at the peak month. We use County Business Patterns data to measure the number of establishments in the food services and accommodations industry (NAICS code 72) for each zipcode in 2010. We collect from the American Community Survey (ACS) zipcode level 5-year estimates of median household income, population, share of 25-60 year olds with bachelors' degrees or higher, employment rate, and owner-occupancy rate. Finally, we obtain annual 1-year estimates of housing vacancy rates at the Core Based Statistical Area (CBSA) level from the same source.

3.6 Summary statistics

Figure 2 shows the geographic distribution of Airbnb listings in June 2011 and June 2016. The map shows significant geographic heterogeneity in Airbnb listings, with most Airbnb listings occurring in large cities and along the coasts. Moreover, there exists significant geographic heterogeneity in the growth of Airbnb over time. From 2011 to 2016, the number of Airbnb listings in some zipcodes grew by a factor of 10 or more; in others there was no growth at all. Figure 3 shows the total number of Airbnb listings over time in our dataset using method 1. From 2012 to 2016, the total number of Airbnb listings grew by a factor of 10, reaching over 1 million listings in 2016.

Table 2 gives a sense of the size of Airbnb relative to the housing stock at the zipcode level. Even in 2015, Airbnb remains a very small percentage of the total housing stock: the number of Airbnb listings is only 0.13% and 1.37% of the housing stock in the median and 90th percentile zipcodes, respectively. When comparing to the stock of vacant homes, Airbnb listings in 2015 account for 1.6% of the stock of vacant homes in the median zipcode and 14% in the 90th percentile zipcode. Perhaps the most salient comparison—at least from the perspective of a potential renter—is the number of Airbnb listings relative to the stock of homes listed as vacant and for rent. This statistic reaches 8.3% in the median zipcode in 2015 and 89% in the 90th percentile zipcode. This implies that in the median zipcode, a local resident looking for a long-term rental unit will find that about 1 in 12 of the potentially available homes are being placed on Airbnb instead of being made available to long-term residents. Framed in this way, concerns about the effect of Airbnb on the housing market do not appear unfounded.

4 Methodology

Let Y_{ict} be either the price index or the rent index for zipcode i in CBSA c in year-month t, and let $AirbnbListings_{ict}$ be the number of Airbnb listings. We assume the following causal relationship between Y_{ict} and $AirbnbListings_{ict}$:

$$\ln Y_{ict} = \alpha + \beta \ln AirbnbListings_{ict} + X_{ict}\gamma + \epsilon_{ict}, \tag{12}$$

where X_{ict} is a vector of observed zipcode characteristics, and ϵ_{ict} contains unobserved factors which may causally affect Y_{ict}. If the unobserved factors are uncorrelated with the number of Airbnb listings, conditional on X_{ict}, then we can consistently estimate β by OLS. However, ϵ_{ict} and $AirbnbListings_{ict}$ may be correlated through unobserved factors at the zipcode, city, and time levels. We allow ϵ_{ict} to contain unobserved zipcode level factors δ_i, and unobserved time-varying factors at the CBSA level θ_{ct}, that affect Y_{ict} and are correlated with $\ln AirbnbListings_{ict}$. Writing: $\epsilon_{ict} = \delta_i + \theta_{ct} + \xi_{ict}$, equation (12) becomes:

$$\ln Y_{ict} = \alpha + \beta \ln AirbnbListings_{ict} + X_{ict}\gamma + \delta_i + \theta_{ct} + \xi_{ict} \tag{13}$$

Even after controlling for unobserved factors at the zipcode and CBSA-year-month level, there may still be some unobserved *zipcode-specific*, *time-varying* factors contained in ξ_{ict} that are correlated with the number of Airbnb listings. To address this issue, we construct an instrumental variable which is plausibly uncorrelated with local monthly shocks to the housing market at the zipcode level, ξ_{ict}, but likely to affect the number of Airbnb listings.

Our instrument begins with the worldwide Google Trends search index for the term "airbnb", g_t, which measures the quantity of Google searches for "airbnb" in year-month t. Such trends represent a measure of the extent to which awareness of Airbnb has diffused to the public, including both demanders and suppliers of short-term rental housing. Figure 1 plots g_t from 2008 to 2016, and shows the explosive growth of Airbnb over the time period. Crucially, the search index is *not* likely to be reflective of growth in overall tourism demand, because it is unlikely to have changed so much over this relatively short time period. Moreover, it should not be reflective of overall growth in the supply of short-term housing, except to the extent that it is driven by Airbnb.

The CBSA-year-month fixed effects θ_{ct} already absorb any unobserved variation at the year-month level. Therefore, to complete our instrument we interact g_t with a measure of how attractive a zipcode is for tourists in base year 2010, $h_{i,2010}$. We measure "touristiness" using the number of establishments in the food services and accommodations industry (NAICS code 72) in a specific zipcode. Zipcodes with more restaurants and hotels may be more attractive to tourists because these are services that tourists need to consume locally—thus, it matters how many of these services are near the tourist's place of stay. Alternatively, the larger number of restaurants and hotels may reflect an underlying local amenity that tourists value.

Our operating assumption is that landlords in more touristy zipcodes are more (or less) likely to switch from the long-term market to the short-term market in response to learning about Airbnb. Landlords in more touristy zipcodes may be more likely to switch because they can book their rooms more frequently, and at higher prices, than in non-touristy zipcodes. Conversely, landlords in more touristy zipcodes may be less likely to switch if there is much stronger competition from hotels.

In order for the instrument to be valid, $z_{ict} = g_t \times h_{i,2010}$ must be uncorrelated with the zipcode-specific, time-varying shocks to the housing market, ξ_{ict}. This would be true if either ex-ante touristiness in 2010 ($h_{i,2010}$) is independent of zipcode level shocks (ξ_{ict}), or growth in worldwide Airbnb searches (g_t) is independent of zipcode level shocks. To see how our instrument addresses potential confounding factors, consider changes in zipcode level crime rate as an omitted variable. It is unlikely that changes to crime rates across all zipcodes are systematically correlated in time with worldwide Airbnb searches. Even if they were, they would have to correlate in such a way that the correlation is systematically stronger or weaker in more touristy zipcodes. Moreover, these biases would have to be systematically present within all cities in our sample. Of course, we cannot rule this possibility out completely. However, we discuss the validity of the instrument further in Section 5.3, and present exercises that suggest the exogeneity assumption is likely satisfied.

5 Results and Extensions

5.1 The effect of home-sharing on house prices and rents

Table 3 reports the main regression results. Each panel reports the results for a different dependent variable: the log of the Zillow rent index, the log of the Zillow house price index, and the log of the price-to-rent ratio. In order to maintain our measure of touristiness, $h_{i,2010}$, as a pre-period variable, only data from 2011 to 2016 are used. This time frame covers all of the period of significant growth in Airbnb. We also include only data from the

100 largest CBSAs, in terms of 2010 population.[20] Column 1 of each panel reports the results from a simple OLS regression of the dependent variable on log listings and no controls. Column 2 includes zipcode and CBSA-year-month fixed effects, column 3 reports the 2SLS results using the instrumental variable, and column 4 adds time-varying zipcode characteristics as controls. These characteristics include the median household income, the total population, the share of 25-60 year olds with bachelors' degrees or higher, and the employment rate. Because these measures are not available at a monthly (or even annual) frequency for zipcodes, we linearly interpolate/extrapolate to the monthly level using the 2007-2011 and the 2011-2015 ACS 5-year estimates at the zipcode level. Column 4 is our preferred specification. Based on these results, we estimate that a 10% increase in Airbnb listings leads to a 0.42% increase in the rental rate, a 0.76% increase in house prices, and a 0.31% increase in the price-to-rent ratio at the zipcode level. These findings are consistent with the theoretical model, which predicts that home-sharing will increase both house prices and rental rates and that such increase is stronger for house prices.

In terms of the magnitude of the effects, we note that from 2012 to 2016, the average zipcode experienced an exogenous 6.5% per year increase in Airbnb listings, as mediated by the instrument.[21] Thus, exogenous increases to the number of Airbnb listings can explain up to 0.27% in annual rent growth and 0.49% in annual house price growth from 2012 to 2016. These effects are modest, but not trivial: the annual rent growth from 2012 to 2016 was 2.2% and the annual house price growth was 4.8%. The magnitudes are also comparable to results estimated for Boston during the period 2015-2016 by Horna and Merantea (2017), who found that a one standard deviation increase in Airbnb listings increases rental rates by 0.4%. Our results suggest that one

[20]The 100 largest CBSAs constitute the majority of Airbnb listings (over 80%). In the Appendix we show that our results are robust to the inclusion of more CBSAs.

[21]To calculate this, we first compute the predicted number of Airbnb listings from the first-stage regression using the instrumental variable. We then calculate the average annual change in the predicted number of listings across zipcodes. The average annual growth in raw Airbnb listings from 2012 to 2016 was 42%, but we do not believe it is appropriate to use this growth rate to explain house prices and rental rates because some of this growth may be endogenous.

standard deviation growth in Airbnb listings leads to a 0.65% increase in rental rates.[22]

5.2 The effect of the owner-occupancy rate

Our theoretical model predicts that the effect of Airbnb on rental rates will be smaller when the share of owner-occupiers is high. Intuitively, this is because only non-owner-occupiers are on the margin of substituting housing units between the long and short-term markets. Owner-occupiers instead use Airbnb as a way to earn rents from excess housing capacity, such as by renting out unused rooms or by renting their home out while they are away on vacation. We now use the data to explore this intuition further.

To test this prediction, we re-estimate our specification while allowing for an interaction term between the number of listings and the owner-occupancy rate.[23] The owner-occupancy rate is computed at the zipcode level in each year using ACS 5-year estimates. The regression results are reported in Table 4. The coefficient of interest, the interaction term, is negative and statistically siginficant, which suggests that the effect of Airbnb on rental rates is lower when the owner-occupancy rate is higher. We obtain similar results for house prices. These results are consistent with the theoretical model. In terms of magnitudes, the effects are economically significant. The interquartile range in the owner-occupancy rate is about 25% (57% to 82%). Thus, going from a zipcode that is in the 25th percentile of owner-occupancy rate to a zipcode that is in the 75th percentile of owner-occupancy rate causes the rental rate impact of a 10% increase in Airbnb listings to go from 0.29% to 0.21%.

Interestingly, a robust result is that the effect of Airbnb on the price-to-rent ratio is also weaker in zipcodes with a higher owner-occupancy rate. This was not necessarily predicted by the model, but could indicate differences in the character of neighborhoods with high vs. low owner-occupancy rate. For

[22] The standard deviation in monthly Airbnb growth in our data is 15%.

[23] The owner-occupancy rate itself is also included in the regression. Because there are now two endogenous regressors, we use $g_t \times h_{i,2010}$ and $g_t \times h_{i,2010} \times OOR_{ict}$ as instruments, where OOR_{ict} is the owner-occupancy rate.

instance, it could be that zipcodes with a higher owner-occupancy rate have owners who are less likely to have under-utilized housing capacity.

5.3 Discussion: validity of the instrumental variable

As discussed in Section 4, the instrument based on the Google search index and the "touristiness" of each zipcode will not be valid if there is some confounding factor that is both 1) correlated in time with the Google search index for Airbnb, g_t, and 2) correlated with differences in touristiness across zipcodes within CBSAs, $h_{i,2010}$. While we cannot rule out the possibility of such a confounding factor, a number of results suggest that the exogeneity assumption is likely to be satisfied.

First, we note that adding time-varying zipcode level characteristics to the instrumental variables regressions does not affect our results (see columns (3) and (4) of Tables 3 and 4.) This suggests that the instrument is *not* highly correlated with zipcode level population, employment rate, college share, or income. These variables are fairly basic measurements of zipcode level economic outcomes, and are likely to be highly correlated with other unobserved factors that affect zipcode level housing markets. Thus, it is unlikely that the instrument is correlated with other unobserved zipcode level factors that affect housing markets. This also rules out possible endogeneity concerns related to the possibility that touristy zipcodes may have been gentrifying faster than non-touristy zipcodes.

Second, we point out that the effect of the owner-occupancy rate in reducing the impact of Airbnb is consistent across specifications. So, if one wanted to argue that the effects we estimate are the result of spurious correlation, one would have to find a confounder that is not only correlated with Google searches for Airbnb and with zipcode level touristiness, but that also affects zipcodes with high owner-occupancy rate less than zipcodes with low owner-occupancy rate. Moreover, such confounder should not be strongly correlated with zipcode level population, employment rate, college share, or income. We cannot think of any obvious candidates. In fact, the owner-occupancy rate in standard asset market models of house prices and rental rates is indeterminate, as residents are indifferent between owning and renting.[24] Changes to the amenity level of a particular neighborhood will be reflected in both rents and prices, but there is no specific prediction about the owner-occupancy rate. In contrast, our model of home-sharing has a very clear prediction about the effect of owner-occupancy rate on the price response to Airbnb, and all of our empirical results are consistent with this prediction.

Finally, we present two additional exercises to test the validity of the instrument. First, we test whether there are differential pre-trends in the house prices of zipcodes of different levels of touristiness in the period before Airbnb became popular. If the instrument is exogenous, there should be no differential pre-trends. Figure 4 plots the Zillow house price index for zipcodes in different quartiles of 2010 touristiness, from 2009 to the end of 2016.[25] The figure shows there are no differential pre-trends in the Zillow Home Value Index (ZHVI) for zipcodes in different quartiles of touristiness until after 2012, which also happens to be when interest in Airbnb began to grow according to Figure 1. This is true when computing the raw averages for the ZHVI within quartile (top panel), as well as when computing the average of the residuals after controlling for zipcode and CBSA-year-month fixed effects (bottom panel). The lack of differential pre-trends suggests that zipcodes with different levels of touristiness do *not* generally have different house price trends, but they only began to diverge after 2012, when Airbnb started to become well known.[26]

Second, we test whether the instrument is positively correlated with house prices and rental rates in zipcodes that were never observed to have any Airbnb listings. If the instrument is valid, then it should only be correlated to house prices and rental rates through its effect on Airbnb listings. Therefore, in areas

[24] See Poterba (1984).

[25] We cannot repeat this exercise with rental rates because Zillow rental price data did not begin until 2011 or 2012 for most zipcodes.

[26] Unfortunately, 2012 also happens to be the year that house prices began to recover from the Great Recession. It is possible that touristy zipcodes have a different recovery pattern than non-touristy zipcodes. However, even if this were the case, it is not clear why the differential recovery should be uncorrelated with zipcode level demographics, and why it should affect zipcodes with different owner-occupancy rates differently.

with no Airbnb, we should not see a positive relationship between the instrument and house prices and rental rates. To test this, we regress the Zillow rent index, house price index, and price-to-rent ratio on the instrumental variable directly, using only data from zipcodes in which we never observed any Airbnb listings. Table 5 reports the results of these regressions and shows that we do not find any statistically significant relationship between the instrument and house prices/rental rates in zipcodes without Airbnb. If anything, we find that there is a negative relationship between the instrument and house prices/rental rates in zipcodes without Airbnb, though the estimates are imprecise and the sample size is considerably reduced when considering only such zipcodes.[27] Thus, there does not seem to be any evidence that the instrument would be positively correlated with house prices/rental rates, except through its effect on short-term rentals.

5.4 The effect of home-sharing on housing reallocation

We now provide some direct evidence that home-sharing affects rental rates and house prices through the reallocation of housing stock. To do this, we will investigate the effect of Airbnb on housing vacancies. Because vacancy data is not available at the zipcode level at a monthly or annual frequency, we focus on annual CBSA level vacancies. We regress vacancy rates at the CBSA-year level on the number of Airbnb listings, year fixed effects, and CBSA fixed effects. Data on vacancies come from annual ACS 1-year estimates at the CBSA level.[28] Table 6 reports the results.

The first thing to note in Table 6 is that the number of Airbnb listings at the CBSA level appears uncorrelated with the total number of vacancies, once controlling for CBSA and year fixed effects (column 1). However, when we break the vacancy rate down by the type of vacancy, we find a positive (though statistically insignificant) association with the share of homes classi-

[27] If we regress house prices and rental rates on the instrument for zipcodes *with* Airbnb, we find a positive and statistically significant relationship.

[28] We compute the total number of vacancies as sum of the number of vacant seasonal units, vacant-for-rent units, and vacant-for-sale units. We ignore vacant units that are for migrant workers, and we ignore vacant units for which the reason for vacancy is unknown.

fied as vacant for seasonal or recreational use and a negative (and statistically significant) association with the share of homes that are vacant-for-rent and vacant-for-sale.

It is important to note that the Census Bureau classifies homes as vacant even if they are temporarily occupied by persons who usually live elsewhere. Thus, homes allocated permanently to the short-term market are supposed to be classified as vacant, and will likely also be classified as seasonal or recreational homes by their owners and/or neighbors.[29] The positive association of Airbnb with vacant-seasonal homes, and the negative association with vacant-for-rent and vacant-for-sale homes is therefore consistent with absentee landlords substituting away from the rental and for-sale markets for long-term residents and allocating instead to the short-term market.

6 Conclusion

Our results suggest that Airbnb growth can explain 0.27% in annual rent growth and 0.49% in annual house price growth from 2012 to 2016. The increases to rental rates and house prices occur through two channels. In the first channel, home-sharing increases rental rates by inducing some landlords to switch from supplying the market for long-term rentals to supplying the market for short-term rentals. The increase in rental rates through this channel is then capitalized into house prices. In the second channel, home-sharing increases house prices directly by enabling homeowners to generate income from excess housing capacity. This raises the value of owning relative to renting, and therefore increases the price-to-rent ratio directly.

Our paper contributes to the debate surrounding home-sharing policy. Critics of home-sharing argue that it raises housing costs for local residents, and we find evidence confirming this effect. On the other hand, we also find evidence that home-sharing increases the value of homes by allowing owners to better utilize excess capacity. In our view, regulations on home-sharing should

[29]When a home is vacant, Census workers will interview neighbors about the occupancy characteristics of the home.

(at most) seek to limit the reallocation of housing stock from the long-term to the short-term markets, without discouraging the use of home-sharing by owner-occupiers. One regulatory approach could be to only levy occupancy tax on home sharers who rent the entire home for an extended period of time, or to require a proof of owner-occupancy in order to avoid paying occupancy tax.

To summarize the state of the literature on home-sharing, researchers have found that home-sharing 1) raises local rental rates by causing a reallocation of the housing stock; 2) raises house prices through both the capitalization of rents and the increased ability to use excess capacity; and 3) induces market entry by small suppliers of short-term housing who compete with traditional suppliers (Zervas et al. (2017)). More research is needed, however, in order to achieve a more complete welfare analysis of home-sharing. For example, home-sharing may have positive spillover effects on local businesses if it drives a net increase in tourism demand. On the other hand, home-sharing may have negative spillover effects if tourists create negative amenities, such as noise or congestion, for local residents. Moreover, home-sharing introduces an interesting new mechanism for scaling down the local housing supply in response to negative demand shocks—a mechanism that was not possible when all of the residential housing stock was allocated to the long-term market.

References

Chen, M. Keith, Judith A. Chevalier, Peter E. Rossi, and Emily Oehlsen, "The Value of Flexible Work: Evidence from Uber Drivers," *NBER Working Paper 23296*, 2017.

Edelman, Benjamin G and Michael Luca, "Digital discrimination: The case of airbnb.com," *Harvard Business School Working Paper*, 2014.

Edelman, Benjamin, Michael Luca, and Dan Svirsky, "Racial discrimination in the sharing economy: Evidence from a field experiment," *American Economic Journal: Applied Economics*, 2017, *9* (2), 1–22.

Einav, Liran, Chiara Farronato, and Jonathan Levin, "Peer-to-Peer Markets," *Annual Review of Economics*, October 2016, *8*, 615–635.

Fradkin, Andrey, "Search frictions and the design of online marketplaces," *Working Paper, Massachusetts Institute of Technology*, 2015.

Gong, Jing, Brad N. Greenwood, and Yiping Song, "Uber Might Buy Me a Mercedez Benz: An Empirical Investigation of the Sharing Economy and Durable Goods Purchase," *SSRN Working Paper*, 2017.

Hall, Jonathan V. and Alan B. Krueger, "An Analysis of the Labor Market for Uber's Driver-Partners in the United States," *ILR Review*, 2017.

Horna, Keren and Mark Merantea, "Is Home Sharing Driving Up Rents? Evidence from Airbnb in Boston," *Journal of Housing Economics*, 2017.

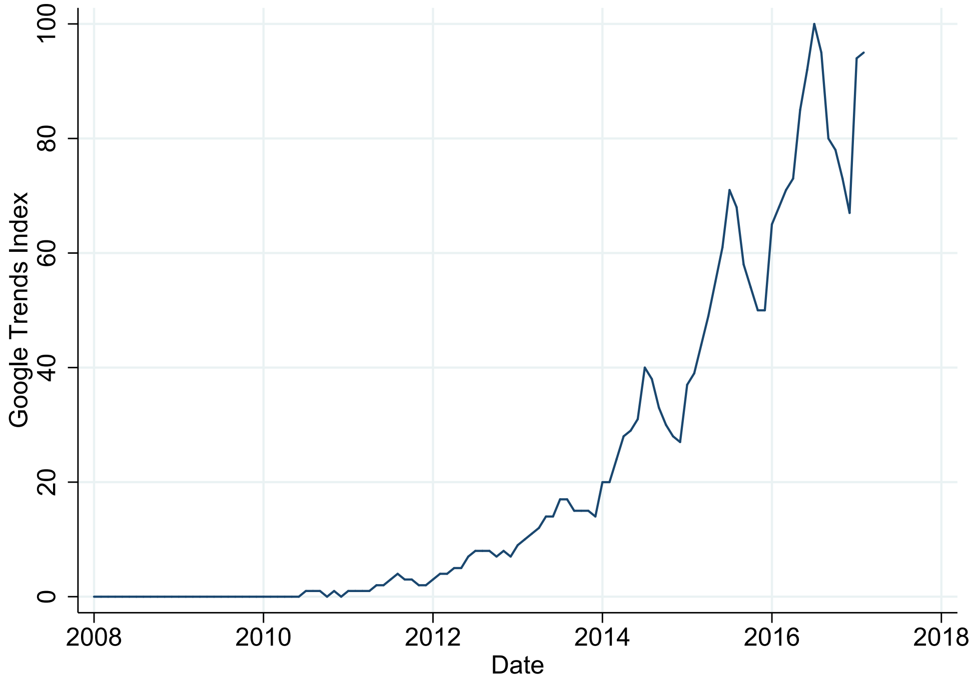
Horton, John J. and Richard J. Zeckhauser, "Owning, Using and Renting: Some Simple Economics of the "Sharing Economy"," *NBER Working Paper 22029*, 2016.

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

Poterba, James M., "Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach," *The Quarterly Journal of Economics*, 1984, *99* (4), 729–752.

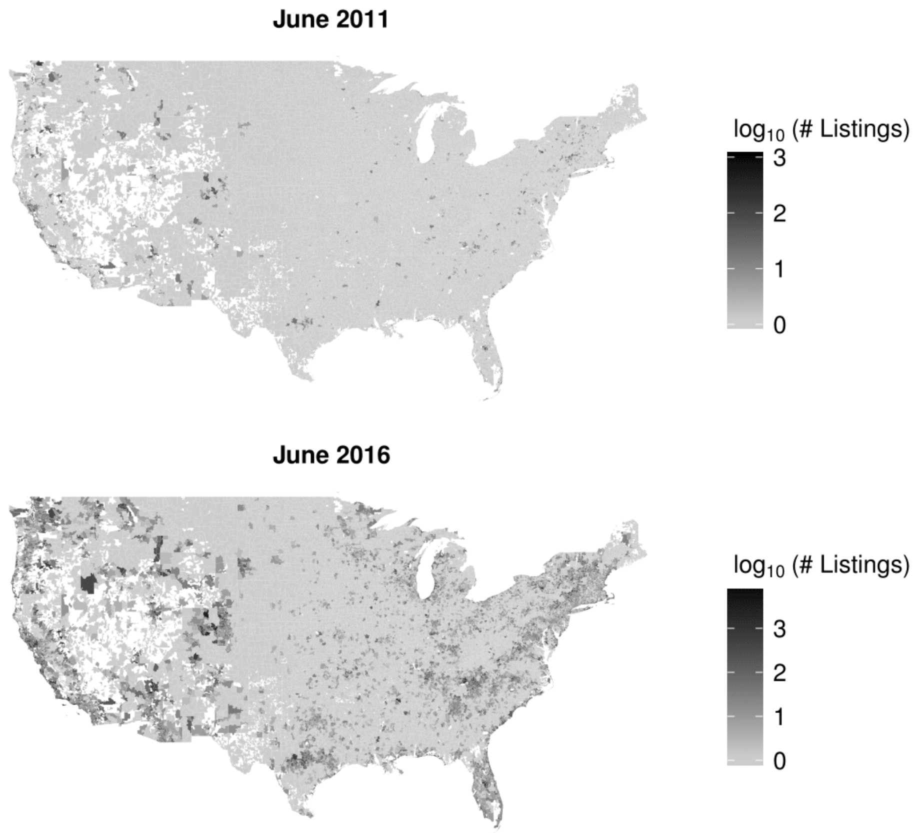
Zervas, Georgios, Davide Proserpio, and John W. Byers, "The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry," *Journal of Marketing Research*, 2017, *54* (5), 687–705.

Figure 1: Google Trends Search Index for Airbnb (Worldwide, 2008-2017)



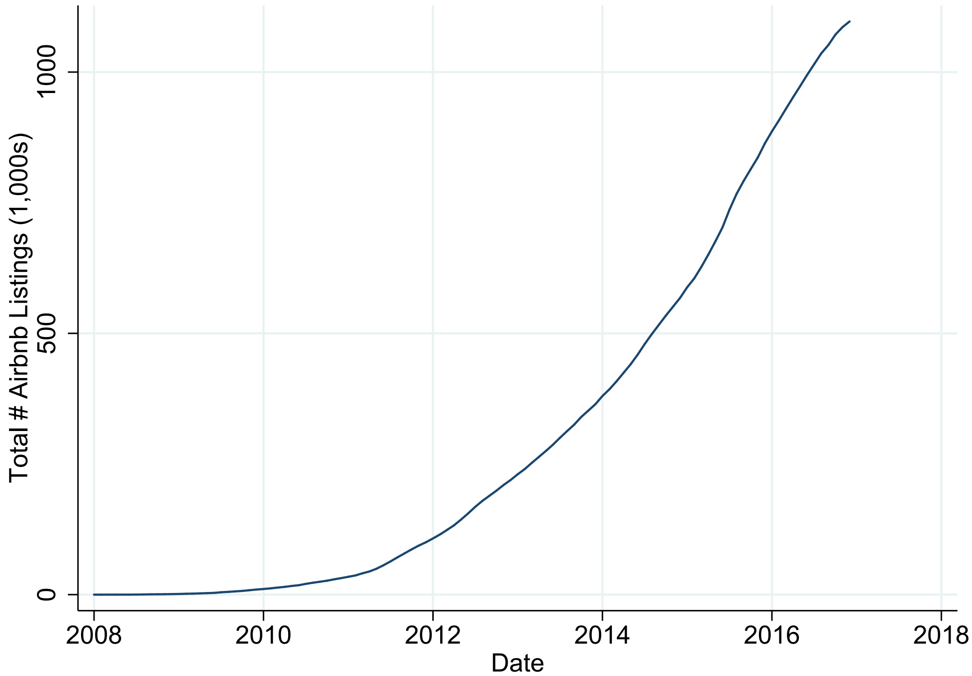
Note: Weekly Google Trends index for the single English search term "Airbnb", from any searches worldwide. Google Trends data are normalized so that the date with the highest search volume is given the value of 100.

Figure 2: Map of Airbnb Listings by Zipcode, 2011-2016



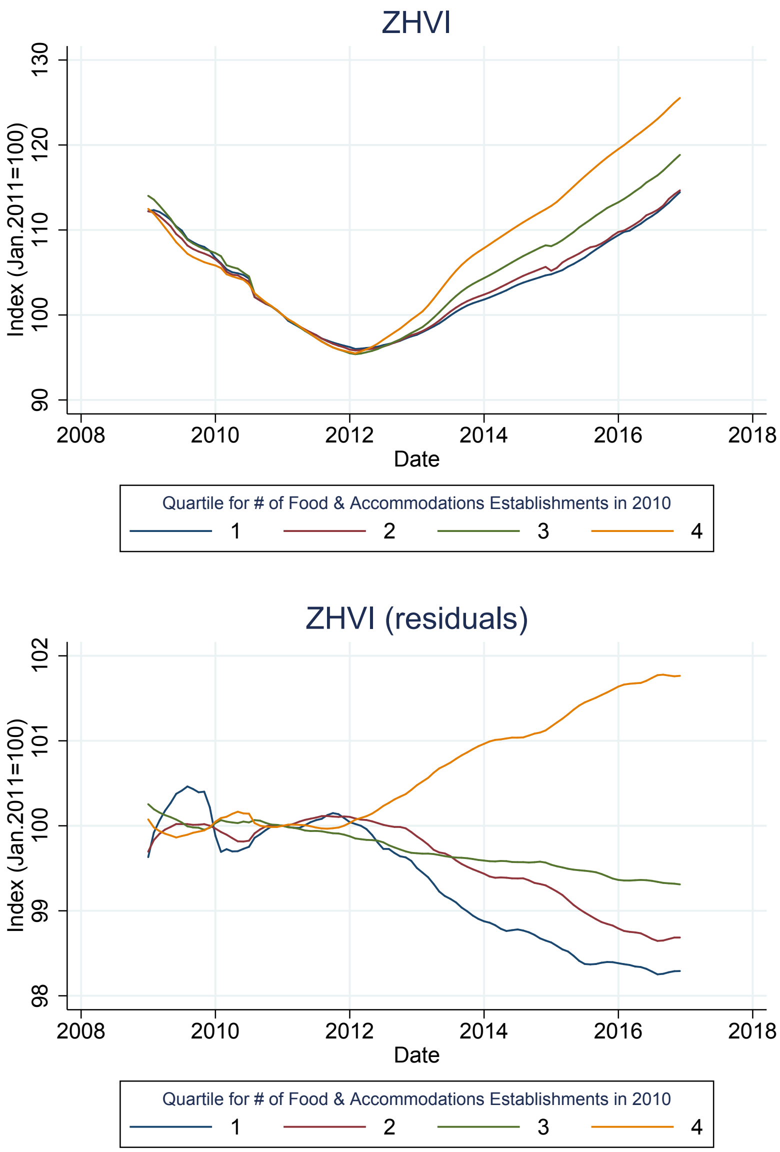
Note: The number of listings is calculated using method 1 in Table 1. Log listings is set to zero if there are zero listings. Geographic areas without zipcode boundary information are colored white.

Figure 3: Total Number of Airbnb Listings (US, 2008-2016)



Note: The number of listings is calculated using method 1 in Table 1.

Figure 4: Trends in Zillow Home Value Index by “Tourstiness” of Zipcode



Note: The top panel plots the ZHVI index, normalized to January 2011=100, averaged within different groups of zipcodes based on their level of “touristiness” in 2010. Touristiness is measured as the number of establishments in the food services and accommodations sector (NAICS code 72) in 2010, and the zipcodes are separated into four equally sized groups. The bottom panel plots the residuals from a regression of the ZHVI on zipcode fixed effects and CBSA-month fixed effects.

Table 2: Size of Airbnb Relative to the Housing Stock (zipcodes, 100 largest CBSAs)

	p10	p20	p50	p75	p90
Year 2011					
Airbnb Listings	0.00	0.00	0.00	1.83	7.50
Housing Unites	1,058.00	2,812.50	7,438.00	12,829.00	18,037.00
Airbnb as a Percentage of					
Total Housing Units	0.00	0.00	0.00	0.02	0.10
Renter-occupied Unites	0.00	0.00	0.001	0.09	0.39
Vacant Units	0.00	0.00	0.01	0.26	1.02
Vacant-for-rent Units	0.00	0.00	0.13	1.30	5.58
Year 2015					
Airbnb Listings	0.58	2	7.92	28.50	98.90
Housing Unites	1,089.00	2,894.50	7,582.00	13,128.00	18,282.00
Airbnb as a Percentage of					
Total Housing Units	0.01	0.05	0.13	0.40	1.37
Renter-occupied Unites	0.05	0.18	0.54	1.66	5.26
Vacant Units	0.13	0.52	1.60	4.76	14.00
Vacant-for-rent Units	0.67	2.45	8.26	27.00	89.00

Note: This table reports the size of Airbnb relative to the housing stock, by zipcodes for the 100 largest CBSAs as measured by 2010 population. The number of Airbnb listings is calculated using method 1 in Table 1. Data on housing stocks, occupancy characteristics, and vacancies come from ACS zipcode level 5-year estimates. We report data for the year 2015 instead of 2016 because data from the 2016 ACS are not yet available.

Table 3: The Effect of Airbnb on Rental Rates and House Prices

	Panel A Dep var: ln Rent Index				Panel B Dep var: ln Price Index				Panel C Dep var: ln Price/Rent			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
ln Airbnb Listings	0.0843***	0.00622***	0.0442***	0.0421***	0.157***	0.00702***	0.0788***	0.0761***	0.0737***	0.000749	0.0309***	0.0312***
	(0.00213)	(0.000522)	(0.00326)	(0.00324)	(0.00382)	(0.000749)	(0.00621)	(0.00619)	(0.00184)	(0.000775)	(0.00442)	(0.00451)
ln Median HH Income				0.0261***				0.0152				-0.0205
				(0.00850)				(0.0140)				(0.0137)
ln Population				0.0363***				0.0680***				0.0284**
				(0.00901)				(0.0152)				(0.0141)
College Share				0.0656***				0.0696**				0.00887
				(0.0195)				(0.0297)				(0.0283)
Employment Rate				0.0461**				0.0323				-0.00930
				(0.0204)				(0.0341)				(0.0311)
Zipcode FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
CBSA-year-month FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
Instrumental variable			Yes	Yes			Yes	Yes			Yes	Yes
Observations	592,439	592,439	592,439	592,007	525,241	525,241	525,241	524,972	496,663	496,648	496,648	496,451
R^2	0.128	0.991	0.990	0.990	0.153	0.996	0.994	0.994	0.142	0.979	0.978	0.978

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: The number of Airbnb listings is calculated using method 1 in Table 1. To avoid taking the log of a zero, one is added to the number of Airbnb listings before taking logs. Because zipcode demographic characteristics are not available at the monthly (or even annual level), zipcode-month measures for household income, population, college share, and employment rate are interpolated from the 2011 and 2015 ACS 5-year estimates. Clustered standard errors at the zipcode level are reported in parenthesis.

Table 4: The Effect of Airbnb on Rental Rates and House Prices, by Owner-Occupancy Rate

	Panel A Dep var: ln Rent Index				Panel B Dep var: ln Price Index				Panel C Dep var: ln Price/Rent			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
ln Airbnb Listings	0.189***	0.0198***	0.0505***	0.0483***	0.356***	0.0293***	0.0737***	0.0698***	0.173***	0.00762***	0.0207***	0.0196***
	(0.00622)	(0.00129)	(0.00319)	(0.00321)	(0.0116)	(0.00202)	(0.00491)	(0.00491)	(0.00583)	(0.00181)	(0.00377)	(0.00384)
... × Owner-Occupancy Rate	-0.111***	-0.0223***	-0.0357***	-0.0336***	-0.217***	-0.0357***	-0.0492***	-0.0453***	-0.114***	-0.0108***	-0.0106**	-0.00968**
	(0.0102)	(0.00178)	(0.00364)	(0.00362)	(0.0182)	(0.00279)	(0.00567)	(0.00561)	(0.00884)	(0.00248)	(0.00425)	(0.00426)
ln Median HH Income				0.0113				0.00463				-0.0139
				(0.00926)				(0.0144)				(0.0148)
ln Population				0.0588***				0.121***				0.0665***
				(0.00907)				(0.0155)				(0.0153)
College Share				0.0694***				0.0798***				0.0197
				(0.0211)				(0.0293)				(0.0307)
Employment Rate				0.0681***				0.119***				0.0511
				(0.0217)				(0.0347)				(0.0339)
Zipcode FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
CBSA-year-month FE		Yes	Yes	Yes		Yes	Yes	Yes		Yes	Yes	Yes
Instrumental Variable			Yes	Yes			Yes	Yes			Yes	Yes
Observations	492,119	492,119	492,119	491,759	437,691	437,691	437,691	437,470	412,565	412,550	412,550	412,389
R^2	0.223	0.992	0.991	0.991	0.251	0.997	0.996	0.997	0.227	0.982	0.981	0.981

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: The number of Airbnb listings is calculated using method 1 in Table 1. To avoid taking the log of a zero, one is added to the number of Airbnb listings before taking logs. Because zipcode demographic characteristics are not available at the monthly (or even annual level), zipcode-month measures for household income, population, college share, and employment rate are interpolated from the 2011 and 2015 ACS 5-year estimates. The owner-occupancy rate is calculated as the number of owner-occupied housing units divided by the sum of owner-occupied units and renter-occupied units, using ACS 5-year estimates. Clustered standard errors at the zipcode level are reported in parenthesis.

Table 5: IV Validity Check: Correlation Between Instrument and Rents/Prices in Zipcodes Without Airbnb

	Panel A Dep var: ln Rent Index			Panel B Dep var: ln Price Index			Panel C Dep var: ln Price/Rent		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
IV	6.70e-06	-3.05e-06	-2.85e-06	-3.24e-05	-2.91e-06	-2.83e-06	-5.43e-05***	-3.22e-06	-3.06e-06
	(1.56e-05)	(3.40e-06)	(3.36e-06)	(3.20e-05)	(5.06e-06)	(5.07e-06)	(1.75e-05)	(5.32e-06)	(5.34e-06)
ln Median HH Income			-0.00458			-0.00718			0.00184
			(0.0146)			(0.0213)			(0.0272)
ln Population			0.0199			0.0417**			0.0142
			(0.0160)			(0.0194)			(0.0232)
College Share			0.0552			0.165***			0.100
			(0.0396)			(0.0465)			(0.0608)
Employment Rate			0.0356			-0.0162			-0.0568
			(0.0384)			(0.0410)			(0.0577)
Zipcode FE		Yes	Yes		Yes	Yes		Yes	Yes
CBSA-year-month FE		Yes	Yes		Yes	Yes		Yes	Yes
Observations	55,890	55,116	55,116	46,105	45,044	44,972	39,270	38,083	38,083
R^2	0.000	0.979	0.979	0.001	0.993	0.994	0.008	0.964	0.964

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: This table reports regression results when dependent variable is regressed on the insrumental variable (Google Trends index for "airbnb" interacted with the number of food service and accommodations establishments in 2010) directly, for zipcodes that were never observed to have any Airbnb listings. Because zipcode demographic characteristics are not available at the monthly (or even annual level), zipcode-month measures for household income, population, college share, and employment rate are interpolated from the 2011 and 2015 ACS 5-year estimates. Clustered standard errors at the zipcode level are reported in parenthesis.

Table 6: The Effect of Airbnb on Vacancy Rates

	(1) All Vacant Units	(2) Seasonal Homes	(3) Vacant-for-Rent	(4) Vacant-for-Sale
ln Airbnb Listings	-5.45e-06	0.00612	-0.00462***	-0.00151**
	(0.00485)	(0.00444)	(0.00151)	(0.000752)
CBSA FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: These regressions are at the CBSA-year level. The number of Airbnb listings is calculated using method 1 in Table 1. To avoid taking the log of a zero, one is added to the number of Airbnb listings before taking logs. The dependent variable is the number of vacant units divided by the total number of housing units. Data on vacancies comes from annual ACS 1-year estimates. Seasonal homes are housing units described as being for seasonal, recreational, or occassional use. Note that according to Census methodology, housing units occupied temporarily by persons who usually live elsewhere are classified as vacant units.

For Online Publication: Appendix

A Model with Endogenous Owner-Occupiers

The model in Section 2 can be extended to allow the share of owner-occupiers to be endogenous. However, ex-ante heterogeneity in potential buyers needs to be introduced or else an equilibrium with all three of renters, owner-occupiers, and absentee landlords would require that equations (4) and (10) both be equal. If they were not, then either long-term residents will outbid absentee landlords to own all the housing, or the opposite will happen.

We introduce heterogeneity in the most parsimonious way possible. Consider a set of N individuals who potentially interact with a local housing market. Each individual can choose to be a renter, an owner-occupier, an absentee landlord, or none of the above. Let us normalize the utility for "none of the above" to zero. The present value of utility that person i gets from being a renter is:

$$\begin{aligned} u_{i,r} &= U - \frac{1}{1-\delta} R + \epsilon_{i,r} \\ &= u_r + \epsilon_{i,r} \end{aligned}$$

Here, U is the present value of amenities that the individual gets from being a resident in this market. $\frac{1}{1-\delta}R$ is the present value of rents. $\epsilon_{i,r}$ is an idiosyncratic utility shock which is known ex-ante. The present value that person i gets from being an owner is:

$$\begin{aligned} u_{i,o} &= U - P + \frac{1}{1-\delta} \gamma g(Q-c) + \epsilon_{i,o} \\ &= u_o + \epsilon_{i,o} \end{aligned}$$

Here, U is again the present value of amenities, P is the purchase price of housing, and $\frac{1}{1-\delta}\gamma g(Q-c)$ is the present value of rents received from selling excess capacity on the peer-to-peer market. Finally, the present value that person i gets from being an absentee landlord is:

$$\begin{aligned} u_{i,a} &= -P + \frac{1}{1-\delta}\left[R + g(Q-R-c)\right] + \epsilon_{i,a} \\ &= u_a + \epsilon_{i,a} \end{aligned}$$

For analytical tractability, let the utility shocks ϵ_i be distributed i.i.d. type 1 extreme value. The share of individuals that choose option j out of $j = \{r, o, a\}$ is:

$$s_j = \frac{\exp u_j}{1 + \sum_{k \in \{r,o,a\}} \exp u_k}$$

The equilibrium conditions determining R and P are:

$$(s_a + s_o)N = H$$

and:

$$\left[1 - f(Q-R-c)\right] s_a N = s_r N$$

The first condition is the market clearing condition for the housing market as a whole; i.e. the number of absentee landlords plus owner-occupiers is equal to the housing stock. The second condition is the market clearing condition for the long-term rental market; i.e. the number of renters is equal to the number of absentee landlords allocating housing to the long-term market.

We leave the derivation of analytical results for this model to future work or enterprising students. For this Appendix, we will simply present some numerical results which are consistent with all the key predictions in Section 2. Choosing $N = 10$, $H = 2$, $U = \$500,000$, $\delta = 0.95$, $\gamma = 0.1$, $Q = \$25,000$, and letting the distribution of idiosyncratic costs to listing in the short-term market be uniform from \$0 to \$100,000, we consider a change of c from ∞ (no home-sharing) to $c = 0$ (costless home-sharing). Table 7 below shows the results. Consistent with the model, the introduction of home-sharing under these model parameters results in a modest increase in both rental rates and house prices, and the increase in house prices is larger than the increase in rental rate. The qualitative results are robust to different parameter choices.

Table 7: Simulation Results

	$c = \infty$	$c = \$50k$	Δ
Rent	\$25,069	\$25,193	0.49%
Price	\$502,773	\$507,702	0.98%

B Comparison to Insideairbnb.com Data

To validate the accuracy of our dataset, in this section we compare our Airbnb listing information with that obtained by Insideairbnb.com, a website that keeps track of Airbnb data in a few key cities. Data from Insideairbnb have been featured in USA Today and have been used for policy research by the city of San Francisco. Because Insideairbnb.com does not collect data all over the U.S., but rather for a handful of specific cities, we compare data for the city of Los Angeles. The Insideairbnb scrape of Los Angeles with timestamp July 3, 2016 contains 15,958 listings. Out of 15,958 listings, we are able to exacly match 15,768 listings, or approximately 99% of the Insideairbnb.com listings (our snapshot data contains a total of 15,808 listings for the city of Los Angeles for the month of June 2016—the closest period to the Insidearibnb.com data). Results are similar when comparing to Insideairbnb data for other cities. Due to the high degree of match between our data and Insideairbnb, we are reassured of the accuracy of our data.

C Robustness Checks

In this section, we show that our main results are robust to the alternative methods of calculating Airbnb supply, as discussed in Section 3. Table 8 replicates the full specification as in Table 4, with zipcode demographic controls, using the methods for calculating Airbnb supply listed in Table 1. Columns (1), (2), and (3) of each panel in Table 8 correspond to methods 2, 3, and 4 of Table 1, respectively. The results when using methods 2 and 3 are very similar in magnitude to the results using method 1. The results when using method 4 are somewhat different, but we note that this is primarily driven by an imprecise estimate of the effect of Airbnb on rents. Otherwise, the results are qualitatively similar.

In Table 9, we show that our results are robust to the choice of CBSAs to include in our estimation sample. The main results used the 100 largest CBSAs, but Table 9 shows that the results are not particularly sensitive to this choice.

Table 8: Robustness Checks: Alternative Methods of Measuring Airbnb Supply

	Panel A Dep var: ln Rent Index			Panel B Dep var: ln Price Index			Panel C Dep var: ln Price/Rent		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
ln Airbnb Listings	0.0472***	0.0496***	-0.00634	0.0678***	0.0716***	0.119***	0.0179***	0.0193***	0.135***
	(0.00335)	(0.00351)	(0.0141)	(0.00521)	(0.00545)	(0.0234)	(0.00371)	(0.00389)	(0.0303)
... × Owner-Occupancy Rate	-0.0318***	-0.0352***	-0.0210*	-0.0422***	-0.0474***	-0.131***	-0.00773	-0.00945*	-0.108***
	(0.00440)	(0.00440)	(0.0118)	(0.00684)	(0.00684)	(0.0195)	(0.00521)	(0.00505)	(0.0255)
ln Median HH Income	0.00622	0.00851	0.105***	-0.00236	8.13e-05	0.0126	-0.0164	-0.0155	-0.0887
	(0.00923)	(0.00925)	(0.0386)	(0.0144)	(0.0144)	(0.0442)	(0.0148)	(0.0148)	(0.0594)
ln Population	0.0625***	0.0612***	-0.0239	0.127***	0.125***	-0.124***	0.0694***	0.0686***	-0.142**
	(0.00894)	(0.00902)	(0.0407)	(0.0157)	(0.0157)	(0.0460)	(0.0153)	(0.0153)	(0.0563)
College Share	0.0637***	0.0661***	0.0121	0.0713**	0.0737**	0.0553	0.0178	0.0184	0.0723
	(0.0206)	(0.0207)	(0.0812)	(0.0292)	(0.0292)	(0.0815)	(0.0307)	(0.0307)	(0.113)
Employment Rate	0.0694***	0.0688***	0.0950	0.116***	0.117***	-0.0957	0.0512	0.0512	-0.268*
	(0.0213)	(0.0215)	(0.0926)	(0.0343)	(0.0345)	(0.0995)	(0.0337)	(0.0337)	(0.142)
Method for Calculating # Listings	2	3	4	2	3	4	2	3	4
Zipcode FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA-year-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrumental Variable	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	491,759	491,759	77,868	437,470	437,470	69,099	412,389	412,389	66,781
R^2	0.991	0.991	0.998	0.996	0.996	0.999	0.981	0.981	0.994

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: Columns (1), (2), and (3) calculate Airbnb listings according to methods 2, 3 and 4 of Table 1, respectively. To avoid taking the log of a zero, one is added to the number of Airbnb listings before taking logs. Because zipcode demographic characteristics are not available at the monthly (or even annual level), zipcode-month measures for household income, population, college share, and employment rate are interpolated from the 2011 and 2015 ACS 5-year estimates. The owner- occupancy rate is calculated as the number of owner-occupied housing units divided by the sum of owner-occupied units and renter-occupied units, using ACS 5-year estimates. Standard errors are clustered at the zipcode level.

Table 9: Robustness Checks: Alternative Samples of CBSAs

	Panel A Dep var: ln Rent Index			Panel B Dep var: ln Price Index			Panel C Dep var: ln Price/Rent		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
ln Airbnb Listings	0.0527***	0.0483***	0.0457***	0.0781***	0.0698***	0.0646***	0.0232***	0.0196***	0.0182***
	(0.00376)	(0.00321)	(0.00298)	(0.00570)	(0.00491)	(0.00452)	(0.00437)	(0.00384)	(0.00357)
... × Owner-Occupancy Rate	-0.0353***	-0.0336***	-0.0319***	-0.0483***	-0.0453***	-0.0434***	-0.0105**	-0.00968**	-0.00969**
	(0.00397)	(0.00362)	(0.00349)	(0.00611)	(0.00561)	(0.00539)	(0.00459)	(0.00426)	(0.00413)
ln Median HH Income	0.0156	0.0113	0.00886	0.0257	0.00463	0.00758	0.00170	-0.0139	-0.00897
	(0.0107)	(0.00926)	(0.00863)	(0.0170)	(0.0144)	(0.0131)	(0.0161)	(0.0148)	(0.0137)
ln Population	0.0449***	0.0588***	0.0608***	0.113***	0.121***	0.112***	0.0797***	0.0665***	0.0588***
	(0.0105)	(0.00907)	(0.00858)	(0.0187)	(0.0155)	(0.0140)	(0.0177)	(0.0153)	(0.0141)
College Share	0.0640***	0.0694***	0.0711***	0.0629*	0.0798***	0.0754***	0.0211	0.0197	0.00775
	(0.0248)	(0.0211)	(0.0195)	(0.0352)	(0.0293)	(0.0268)	(0.0354)	(0.0307)	(0.0285)
Employment Rate	0.0750***	0.0681***	0.0672***	0.117***	0.119***	0.102***	0.0418	0.0511	0.0400
	(0.0256)	(0.0217)	(0.0201)	(0.0409)	(0.0347)	(0.0308)	(0.0379)	(0.0339)	(0.0308)
Sample: *N* largest CBSAs	50	150	200	50	150	200	50	150	200
Zipcode FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CBSA-year-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrumental Variable	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	388,780	491,759	547,202	351,285	437,470	491,466	332,705	412,389	459,122
R^2	0.991	0.991	0.991	0.996	0.997	0.997	0.982	0.981	0.981

Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Note: The samples in columns (1), (2), and (3) are the 50, 150, and 200 largest CBSAs, respectively (the baseline was 100 CBSAs as reported in Table 4. The number of Airbnb listings is calculated using method 1 in Table 1. To avoid taking the log of a zero, one is added to the number of Airbnb listings before taking logs. Because zipcode demographic characteristics are not available at the monthly (or even annual level), zipcode-month measures for household income, population, college share, and employment rate are interpolated from the 2011 and 2015 ACS 5-year estimates. The owner-occupancy rate is calculated as the number of owner-occupied housing units divided by the sum of owner-occupied units and renter-occupied units, using ACS 5-year estimates. Standard errors are clustered at the zipcode level.