The Welfare Effects of Peer Entry: The Case of Airbnb and the Accommodation Industry[1]

Chiara Farronato[2] and Andrey Fradkin[3]

May 13, 2021

Abstract

We study the welfare effects of enabling peer supply through Airbnb in the accommodation industry. We present a model of competition between flexible and dedicated sellers - peer hosts and hotels - who provide differentiated products. We estimate this model using data from major US cities and quantify the welfare effects of Airbnb on travelers, hosts, and hotels. The welfare gains are concentrated in locations (New York) and times (New Year's) when hotels are capacity constrained. This occurs because peer hosts are responsive to market conditions, expand supply as hotels fill up, and keep hotel prices down as a result.

[1]Katie Marlowe and Max Yixuan provided outstanding research assistance. We thank Nikhil Agarwal, Susan Athey, Matt Backus, Liran Einav, Christopher Knittel, Jonathan Levin, Greg Lewis, Chris Nosko, Debi Mohapatra, Ariel Pakes, Paulo Somaini, Sonny Tambe, Dan Waldinger, Ken Wilbur, Kevin Williams, Georgios Zervas, and numerous seminar participants for feedback. We are indebted to Airbnb's employees, in particular Peter Coles, Mike Egesdal, Riley Newman, and Igor Popov, for sharing data and insights. We also thank Duane Vinson at STR and Sergey Shebalov at Sabre for sharing valuable data insights. Airbnb reviewed the paper to make sure that required confidential information was reported accurately. Farronato has no material financial relationship with entities related to this research. Fradkin was previously an employee of Airbnb, Inc. and holds stock that may constitute a material financial position.

[2]Harvard Business School and NBER, cfarronato@hbs.edu

[3]Boston University and MIT Initiative on the Digital Economy, fradkin@bu.edu

1 Introduction

The Internet has greatly reduced entry and advertising costs across a variety of industries. As an example, peer-to-peer marketplaces such as Airbnb, Uber, and Etsy currently provide a platform for small and part-time *peer* providers to sell their goods and services. Several of these marketplaces have grown quickly and become widely known brands. In this paper, we study the welfare effects of peer production in the market for short-term accommodations, where Airbnb is the main peer-to-peer platform and hotels are incumbent suppliers.

Since its founding in 2008, Airbnb has grown to list more rooms than any hotel group in the world. Yet Airbnb's growth across cities and over time has been highly heterogeneous, with supply shares ranging from over 15% to less than 1% across major US cities at the end of 2015. Airbnb's entry has also prompted policy discussion and a variety of regulatory frameworks in many cities across the world. In order to understand Airbnb's growth and its welfare effects, we present stylized facts about Airbnb room supply and its effects on hotels, which we use to motivate a demand and supply framework where accommodations can be provided by either dedicated or flexible supply – hotels vs peer hosts. A key difference between hotels and peer hosts is that while hotels have dedicated rooms that are always available to be rented by travelers, peer hosts have alternative uses for their rooms, which make them more responsive to demand and price fluctuations.

We estimate our model of competition between incumbent hotels and peer hosts, using data from top US cities to quantify the welfare effects of peer entry on travelers, incumbent hotels, and peer hosts. We find that in 2014 Airbnb generated \$305 million in consumer surplus, or about \$70 per Airbnb room-night booked, and \$112 million in peer host surplus, or about \$26 per room-night. These benefits came at the expense of hotels who experienced a 1.6% decrease in revenues and up to 2.8% decrease in variable profits. These effects were concentrated in locations (New York) and times (New Year's Eve) where hotel capacity was constrained.

Our data mainly come from two sources: proprietary data from Airbnb, and data from Smith Travel Research (STR), which tracks supply and demand metrics for the hotel industry. We obtain data on average prices, rooms sold, and rooms available at a city, day, and accommodation type level – luxury through economy – between 2011 and 2015 for the 50 largest US cities.[1] There is substantial heterogeneity in the size of Airbnb across cities and over time as measured by Airbnb supply shares, which we define as the ratio of available Airbnb rooms over the sum of hotels' room capacity and available Airbnb rooms. Cities like New York and Los Angeles have grown more quickly, reaching supply shares exceeding 15%

[1] The 50 largest US cities were selected on the basis of their total number of hotel rooms.

and 11% respectively in 2015, while cities like St. Louis and Detroit have grown more slowly, with less than 1% supply shares at the end of 2015. Within each city over time, the number of available rooms is higher during peak travel times such as Christmas and the summer. The geographic and temporal heterogeneity suggests that hosts flexibly choose when to list their rooms for rent on Airbnb, and are more likely to do so in cities and times when the returns to hosting are highest.

In Section 2 we offer additional stylized facts on differences in Airbnb supply across cities and over time. Across cities, we show that Airbnb supply shares are larger in cities where hotel prices are higher. These high prices are associated with the difficulty of expanding hotels' room capacity due to regulatory or geographic constraints. Airbnb supply is also larger in cities where residents tend to be single and have no children. These residents likely have lower costs of hosting strangers in their homes. Two other predictors of peer supply are demand trends and volatility. A city can experience periods of high and low demand due to seasonality, festivals, or sporting events. When the difference in peaks and troughs is large, the provision of accommodations exclusively by dedicated hotel rooms can be inefficiently low. We show that Airbnb's supply share is larger precisely in cities with high demand volatility, and, perhaps more intuitively, in cities where demand is trending upward.

Over time, we show that peer hosts expand and contract the number of rented rooms in response to price fluctuations more than hotels. On average, we estimate that peer supply elasticity is approximately three times as high as hotels' supply elasticity. This difference is consistent with the nature of hotels' and peers' room supply. Peer suppliers are highly responsive to market conditions, hosting travelers when prices are high, and using accommodations for private use when prices are low. In contrast, because hotels have a fixed number of rooms dedicated to travelers' accommodation, they typically choose to transact even when demand is relatively low, while they are unable to expand capacity during peaks in demand.

The heterogeneous entry of peer hosts across cities and over time has implications for their competitive effects on hotels. We estimate reduced-form regressions of hotel performance on Airbnb supply using active Airbnb listings as instruments for available listings, as well as controls for aggregate demand shocks. We find that the negative effect of Airbnb on hotel revenues is concentrated in cities with constrained hotel capacity, and that compared to other cities, hotels here experience a bigger reduction in prices than occupancy rates.

In Section 3, we rationalize the stylized facts on peer entry and hotel performance with a model of short-run competition between hotels and peer hosts. In this model, rooms for accommodations can be provided by dedicated or flexible sellers, and products are differentiated. We define the short-run horizon as one day in one city, during which the capacity of flexible and dedicated sellers is fixed, and aggregate demand is realized. Travelers choose an accommodation option among differentiated hotel and Airbnb rooms. Hotels choose quantities to maximize profits subject to their capacity constraints, while peer hosts act as a competitive fringe taking prices as given.

We use data between 2013 and 2015 from the 10 largest cities, those that have experienced the largest entry of Airbnb, to recover the primitives of our model. Our estimation strategy proceeds in three steps. First, we estimate a random coefficient multinomial logit demand model (Berry et al. (1995)). We augment our estimation with survey data regarding the preferred second choices of Airbnb travelers, which helps us identify substitution between Airbnb and hotel options. Second, we estimate hotels' cost functions assuming Cournot competition between hotels of the same scale. In order to take into account the fact that prices steeply increase when occupancy approaches 100%, we follow Ryan (2012) and rationalize these price changes with marginal costs that start increasing when hotels are close to their capacity constraint. Third, we estimate the cost distribution of peer hosts assuming that they are price takers. Together, these estimates allow us to measure consumer and peer producer surplus, as well as to quantify how surplus would change in the absence of peer supply, or if peer supply were subject to various regulations such as lodging taxes or quotas.

Section 4 presents our results. We find that consumers' mean utility for Airbnb is lower than for hotels, but that preferences for Airbnb increase between 2013 and 2015. By the end of the sample period, the mean utility from top quality Airbnb listings is close to, although still lower than the mean utility of economy and midscale hotels. We find that peer hosts often have higher marginal costs than hotels of the corresponding quality tier, and that consistent with our model, the distribution of peer costs makes peer supply highly elastic.

In the absence of Airbnb, total welfare would be lower, travelers and peer producers would be worse off, while hotels would benefit from less competition. We report the effects in 2014. In the top 10 US cities in 2014, hotels would increase profits by \$165 million but peer host surplus would go from \$112 million to zero, and consumer surplus would decrease by \$305 million. There are two ways to think about these magnitudes. On one hand, since peer production in the baseline scenario is responsible for just 3% of rooms sold in 2014, the consumer surplus loss is small relative to the revenues in the market. In particular, hotel and peer hosts' revenues in 2014 were \$27,320 million, meaning that the loss of consumer surplus is on the order of 1.1% of total revenues. On the other hand, the benefit to individual consumers and hosts is large. The consumer surplus benefit of Airbnb is \$70 per Airbnb room night and the peer surplus is \$26 per room night.

About half of consumer surplus comes from Airbnb travelers enjoying new accommodation options and lower prices, while the other half affects hotel travelers who face higher prices in the absence of Airbnb. As it turns out, because of the elastic peer supply, actual Airbnb bookings and thus surplus gains disproportionately occur in cities (New York) and times (New Year's Eve) when hotel capacity constraints bind. Indeed, 40% of the consumer surplus loss is concentrated in 19.6% of nights with high demand for accommodations. Particularly in those markets, in the absence of peer supply travelers could not easily find a substitute hotel room because hotels would be fully booked. We find that around 62% of Airbnb bookings, and 87% during nights with high demand for accommodations, would not have been hotel stays had Airbnb not existed.

The concentration of Airbnb bookings in cities and periods of peak demand suggests that in the absence of Airbnb, hotels would be limited in their ability to increase the number of booked rooms – they were already operating at or close to full capacity – but instead would be able to increase prices. Indeed, we find that without Airbnb, hotel revenues and profits increase by a higher percentage than hotel rooms sold. During periods of high demand in particular, when hotel rooms sold cannot increase, hotels would be able to increase revenues by 1.4% and profits by 2.4%

We also use our model to evaluate two policy proposals affecting peer hosts. During the time period of our sample, cities typically did not collect lodging taxes on peer hosts but over time Airbnb has negotiated agreements to collect lodging taxes on behalf of local jurisdictions. In our first policy counterfactual, we study how the market would be affected if peer hosts faced the same tax rate as hotels in each of our cities. We find that these taxes would reduce consumer and peer surplus by $95 million (which is 23% of the loss that would have occurred if Airbnb had been completely banned) but would increase lodging tax revenues by $72 million, or 1.8% increase compared to the baseline scenario. Another policy proposed by jurisdictions is to cap the number of days for which hosts could be booked. We find that a quota limiting Airbnb sales to the 90 days with the largest number of travelers in a city would decrease consumer and peer surplus by $229 million (which is 55% of the loss that would have occurred if Airbnb had been banned).

Finally, Airbnb and its peer hosts have continued growing following 2015 and have become an even larger share of the accommodations market. We use our model to investigate a counterfactual with twice as many Airbnb listings as in 2014. We find that consumer surplus and peer surplus increase by $168 million, which is 39% of the loss that would have occurred if Airbnb did not exist, while hotel profits decrease by $64 million, or by 1.1% compared to the baseline profits.

We contribute to the growing empirical literature on online peer-to-peer platforms (Einav et al., 2016). A limited number of papers have looked at the effect of online platforms on incumbents, in particular Zervas et al. (2017) for Airbnb, Seamans and Zhu (2014) and Kroft and Pope (2014) for Craigslist, and Aguiar and Waldfogel (2018) for Spotify. In this paper we not only estimate the effects on incumbent firms, but also on consumers and new producers. In addition, we highlight important dimensions of heterogeneity of the effects of Airbnb across cities and over time. A complementary paper to ours is Cohen et al. (2019), which uses discontinuities in Uber's surge pricing policy to estimate the consumer surplus from ride sharing. Both of our papers find that successful peer-to-peer platforms generate substantial consumer surplus. However, Cohen et al. (2019) ignore the impact of ride-sharing on the incumbent taxis. In particular, they assume that incumbents do not change their behavior to estimate consumer welfare from ride-sharing, while we incorporate capacity constraints and allow for hotel prices to adjust in the absence of Airbnb. This is important for our setting because even travelers who book hotel rooms benefit from Airbnb through lower prices. Similar to Cohen et al. (2019), Castillo (2020) quantifies the benefits of surge pricing from Uber while Lam and Liu (2019) extend the focus to estimate a model of competition between Uber, Lyft, and taxis using data from New York. Finally, Calder-Wang (2020) and Almagro and Dominguez-Iino (2020) estimate the externalities that Airbnb has on the rental market and on neighborhood amenities, both of which affect where local residents choose to live.

Another related stream of research studies the role of peer-to-peer markets in enabling rental markets for durable goods. Filippas et al. (Forthcoming) derive a theoretical equilibrium model for ownership and rental of durable goods, and make predictions on the existence and size of rental markets across different product categories. Fraiberger and Sundararajan (2019) calibrate a model of car usage and quantify the expected reduction in car ownership as a result of peer-to-peer rental markets.

Other work on peer-to-peer markets has focused on the market design aspects of reputation systems (Fradkin et al. (2019), Nosko and Tadelis (2019), Bolton et al. (2012)), search (Fradkin (2019), Horton (2014)), and pricing (Einav et al. (2018), Hall et al. (2019)). Those are important market design decisions affecting the welfare that Airbnb generates for peer hosts and travelers, but in our paper we do not model them and instead take them as given. Complementary work by Lewis and Zervas (2019) finds sizable benefits for hotel travelers from online reviews, which have been a feature of both Airbnb and hotels throughout our sample.

We find that host supply is highly elastic on the margin. This is consistent with analysis of suppliers on Taskrabbit (Cullen and Farronato (2019)) and Uber (Hall et al. (2019), Chen (2016)). Finally, in our analysis of growth heterogeneity across cities, we contribute to the predominantly theoretical literature on technology adoption and diffusion (e.g. Bass (1969) and Griliches (1957)).

The paper is structured as follows. In the next section, we present the data and document geographic and temporal heterogeneity in the size of Airbnb, comparing short-run elasticities of Airbnb and hotel supply, and estimating average competitive effects of Airbnb on hotel prices and occupancy rates. Section 3 presents a short-run model of demand and differentiated supply of accommodations. We also discuss our empirical strategy to structurally estimate our model parameters that determine consumer utility and supplier costs. We discuss estimation results and counterfactual scenarios in Section 4, and conclude in Section 5.

2 Data and Stylized Facts

In this section, we describe our data on Airbnb and hotels, and document some stylized facts on Airbnb entry and its effects on hotels that motivate our structural model in the next section.

We first explain why we take Airbnb as representative of peer entry in the accommodation market. Airbnb describes itself as a trusted community marketplace for people to list, discover, and book unique accommodations around the world — online or from a smartphone. The marketplace was founded in 2008 and has more than doubled in total transaction volume during every subsequent year at least until 2015, the end of our sample. Airbnb has created a market for a previously rare transaction: the short-term rental of rooms to strangers. In the past, these transactions were not commonly handled by single individuals because there were large costs to finding a match, securely exchanging money, and ensuring safety.

Airbnb plays a variety of fundamental roles in enabling peer transactions. These include marketing the platform, developing the search interface and algorithms, hosting and curating online reviews, processing payments, and providing customer service. We treat these as a black box throughout the paper, meaning that we cannot separate the share of consumer utility generated by the platform relative to the share of utility generated by peer hosts. The role of Airbnb in pricing warrants special attention. Airbnb has a two-part fee structure with a 3% fee to the host, a variable fee to the guest, and a flat booking fee. Fee rates tend to be decreasing in the total value of a booking, but they are not otherwise chosen strategically in response to specific demand or supply conditions. Airbnb has also implemented automated pricing for hosts but this primarily occurs after our main estimation sample.[2]

We use Airbnb data to study the welfare effects of facilitating peer entry in the accommo-

[2]Airbnb added an automated pricing option ('Smart Pricing') in Nov of 2015 (https://www.cnbc.com/2015/11/12/airbnb-launches-smart-pricing-for-hosts.html) and price suggestions ('Price Tips') in June of 2015.

dation market. While Airbnb is not the only company serving this market, it is the dominant platform in most US cities. Indeed, the most prominent competitor is Homeaway/VRBO, a subsidiary of Expedia. Its business has historically been concentrated in rentals of entire homes in vacation destinations, such as beach and skiing resorts.[3] Starting in Q4 of 2019, we have data on gross booking value on both companies from mandatory SEC filings (Airbnb (2020) and Expedia (2019)). Airbnb's gross booking value in Q4 of 2019 ($8.6 billion) was almost four times larger than Homeaway / VRBO ($2.3 billion).

Our proprietary Airbnb data consist of information aggregated at the level of four listing types, from luxury through economy. The variables we observe include the number of bookings, active and available listings, as well as average transacted prices. An available listing is defined as one that is either booked through Airbnb or is open to being booked on the date of stay according to a host's calendar. An active room is defined as a listing that is available to be booked or is already booked for at least one date in the future. Average transacted prices are calculated among all booked rooms on a given date, regardless of the time of booking.

We categorize Airbnb listings into four types: *Airbnb Luxury*, *Airbnb Upscale*, *Airbnb Midscale*, and *Airbnb Economy*.[4] We define listing types using the following algorithm. On the Airbnb servers, we first run a city-level hedonic regression of the transacted nightly price on listing fixed effects, date fixed effects, and bins for the number of five-star reviews and nights of trip.[5] This regression is run at the level of a listing-day pair, conditional on the listing being booked for that particular day. Second, we extract the listing fixed effects and use Bayesian shrinkage to shrink fixed effects towards the mean. Third, we compute quartiles of listing quality and categorize a listing in a given quartile if the sum of the shrunken listing fixed effect and the corresponding review count fixed effect falls into the appropriate range. For each city and day, we aggregate price and quantity information at the level of these four listing quartiles before pulling the data from the Airbnb servers for use in our study. This procedure allows us to account for heterogeneity in Airbnb listing types without specifically modeling detailed geographic and room type characteristics at a city level.[6]

The hotel data come from Smith Travel Research (STR), an accommodation industry

[3] https://vrmintel.com/airbnb-vs-homeaway-winning-race-top-vacation-rental-industry/, accessed in May 2020.

[4] These categories are defined solely for the purpose of this paper and do not correspond to any metric used by Airbnb itself. As part of our agreement with Airbnb for this project, we cannot use listing-level data. However, we could classify individual listings into four groups – luxury through economy – at our own discretion.

[5] The bins for the number of five-star reviews are: 0, 1, 2-3, 4-5, 6-10, 11-25, 26-50, 51-100, ≥ 101. The bins for nights are 1, 2, 3, 4 - 5, 6 - 7, 8 - 14.

[6] We consider the role of location within a city Appendix C.

data provider that tracks over 161,000 hotels. Our sample contains daily prices, rooms sold, and rooms available for the 50 largest US cities for the period between January 2011 and December 2015.[7] STR obtains its information by running a periodic survey of hotels, to which they ask daily revenue attributable to the sale of hotel rooms, total rooms sold, and total rooms available. For the 50 largest markets, 68% of properties are surveyed, covering 81% of available rooms. STR uses supplementary data on similar hotels to impute outcomes for the remaining hotels which are in their census but do not participate in the survey. The data is then aggregated to six hotel scales, from luxury to economy, which indicate the quality and amenities of the hotels. So the data tell us, for example, the average transacted price, the number of rooms available and the number of rooms sold on January 10^{th}, 2013 for midscale hotels in San Francisco.

Before presenting relevant facts from the data, we demonstrate how to properly measure Airbnb supply. Figure 1 displays four measures of the size of Airbnb plotted over time: active listings, two measures of available listings, and booked listings. Active listings on a given day are those listings that are available to be booked for the same day or any future date. Available listings (unadjusted) are listings that are either booked for the day or listed as available to be booked on the same day. And booked listings are listings that have been booked for the day. This figure displays three important facts. First, the share of active or available listings that are booked varies greatly over time. The booking rate is especially high during periods of high demand such as New Year's Eve and the summer. What we will show in Section 2.3 is that this is the result of a highly elastic peer supply. Second, the gap between active listings and available listings is increasing over time, suggesting that over time more and more listings active on the website are not in fact available for rent to travelers – the share of active listings that are listed as available or booked on the day of stay drops from 77% in the first month of data to 65% in the last month of data. Therefore, the meaning of an active listing does not stay constant over the entire period of our study.

The third and most relevant fact from Figure 1 is that the number of unadjusted available listings actually decreases during periods of high demand, most notably on New Year's Eve. The main reason for this is that calendar updating behavior responds to room demand. Many hosts do not pro-actively take the effort to block a date on their calendar when they are unavailable (see Fradkin (2019) for evidence). However, when they receive a request to book a room, they often reject the guest and update their calendar accordingly. Since a larger share of listings receives inquiries during high demand periods, the calendar is also

[7] The cities are ranked based on the absolute number of hotel rooms in 2014. See Census Database: http://www.str.com/products/census-database and STR Trend Reports: http://www.str.com/products/trend-reports.

more accurate during those times. Therefore, the naively calculated availability measure suffers from endogeneity and is even counter-cyclical – high when demand is low, and low when demand is high.

Since we need a measure of the size of Airbnb that stays stable over time, we create an adjusted measure of available listings. This measure includes any rooms which were listed as available for a given date or were sent an inquiry for a given date and later became unavailable. Therefore, it does not suffer from the problem of demand-induced calendar updating. It does overstate the "true" number of available rooms in the market, but as long as it overestimates true availability consistently over time we consider it to be the best measure of Airbnb size. Figure 1 displays our proposed measure (red line) against the naive measure of available listings (blue line). The new measure does not suffer from drops in availability during high demand periods. Throughout the rest of the paper we use the adjusted number of available listings as the size of Airbnb supply, and we simply call them *available listings*.

2.1 Descriptives

We first use our data to describe Airbnb's growth. Airbnb room supply has grown quickly in the aggregate, but the growth has been highly heterogeneous across geographies. Figure 2 plots the size of Airbnb measured as the number of available Airbnb listings out of the sum of available Airbnb listings and hotel room capacity.[8] Even among the top 10 cities in terms of listings, there are high growth markets like San Francisco and New York, as well as slow growth markets like Chicago and DC. This fast increase in available rooms is specific to the peer-to-peer sector and does not represent a broader growth of the supply of short-term accommodations (see Appendix Figure A6).

Within a city over time, there is heterogeneity in the number of available Airbnb listings relative to the number of hotel rooms. The fluctuations are especially prominent in New York in Figure 2, which experiences large spikes in Airbnb available rooms during New Year's Eve, and in Austin during the South by Southwest festival. The figure suggests that market conditions during these spikes are especially suited to peer-to-peer transactions.

Table 1 shows city-level descriptive statistics regarding hotels and Airbnb. An observation is a city. For every city we compute the average transacted price per room-night between January 2011 and December 2015. The table displays the mean and standard deviation

[8]Airbnb uses the terms 'available listings' and 'active listings' in financial fillings to reference metrics that do not exactly coincide with ours (Airbnb (2020)). We find a similar heterogeneity in the growth of Airbnb if we adjust for differential capacity of hotel rooms and Airbnb listings, or if we divide the number of Airbnb listings by the number of total housing units within a city.

of these average daily prices across the 50 cities in our sample, as well as other metrics analogously computed. In the average city, hotels charge \$111 per room-night and their occupancy rate, defined as the share of booked rooms out of available rooms, is 67%. Perhaps surprisingly, Airbnb has very similar transacted prices (\$114) and much lower occupancy rates (16%).[9] The within-city standard deviation of these outcomes varies greatly across cities. For example, the city at the 25th percentile has a standard deviation of hotel prices of \$10 (\$23 for Airbnb prices), while the city at the 75th percentile has a standard deviation of \$23 (\$39 for Airbnb prices). This indicates that cities differ not only in levels – some cities have consistently higher prices and occupancy rates – but also in the extent to which market conditions fluctuate over time.

During our sample period, Airbnb comprises a small share of available rooms, which is defined as the sum of available Airbnb listings plus hotel room capacity. The average Airbnb share of available rooms in the last quarter of 2015 is 4%, and in most cities it is between 1% and 6% (25th and 75th percentiles).[10] When controlling for differential guest occupancy of hotel rooms and Airbnb listings, we find that across all cities Airbnb listings can host 5% of potential guests.[11] Finally, Airbnb listings represent less than 1% of total housing units for all cities in our sample.

The rest of this section highlights three important stylized facts about peer entry. First, differences in peer entry across cities can be predicted by proxies for hotels' costs to expand room capacity, population demographics likely to affect peers' costs of hosting strangers, and proxies for travelers' growth and variability. Second, peer supply is very responsive to price, quickly expanding and contracting in response to changes in demand. In fact, peer supply is more than twice as elastic as hotel supply, which is instead capped at the maximum number of hotel rooms built in a given city. Finally, peer supply exerts competitive pressure on hotels, negatively affecting their revenues, but this negative effect is concentrated on hotel prices rather than occupancy rates, and in cities where hotels are capacity-constrained. In Sections 3 and 4 we focus on cities that experienced the largest entry of Airbnb, and thus the largest effects, to quantify the welfare benefits of peer entry.

[9] Recall that our definition of available listings underestimates occupancy by construction, since it includes rooms that turn out to be unavailable in the denominator.

[10] Airbnb listings are typically able to host more guests than hotel rooms, and we describe adjustments to account for this in Section 3.

[11] The number of potential guests for a city-day-listing type is measured as the product of available rooms times their average occupancy. We have information on the number of guests occupying Airbnb listings for each listing type, city, and day. For hotels, we assume that the typical hotel has the same number of average guests as a 'Midscale' Airbnb listing. Once we have this metric for each listing type, we sum over all listing types to obtain the number of potential guests who can be hosted by Airbnb and hotels in a given city and day.

2.2 Predictors of Peer Entry Across Cities

In this section, we focus on differences in the size of Airbnb across the 50 cities in our sample, showing that predictors of Airbnb size include proxies for costs and demand characteristics in the accommodation industry. Awareness about the Airbnb platform grew between 2011 and 2015, and has kept growing, leading to a continuous increase in the number of hosts joining the platform (Figure 2). We assume that the last quarter in 2015, the end of our sample, provides a valid proxy for the long-run heterogeneity in Airbnb penetration across cities. In particular, if Airbnb was larger in New York than Boston in 2015 as a share of available rooms, we assume that in equilibrium Airbnb will still be larger in New York compared to Boston. So for the discussion in this section we average Airbnb share of available rooms, which is defined as in Figure 2, between October and December 2015.[12]

Figure 3 shows the correlation between Airbnb share of available rooms in 2015 and hotel daily revenues per available room in 2011, the beginning of our sample period. Not surprisingly, the size of Airbnb is positively correlated with the average revenue per available hotel room in a city, with New York being both the city with the highest hotel revenues and the one with the highest penetration of peer hosts.

One reason for high hotels' revenue per available room is the difficulty to expand hotel room capacity since hotels have high fixed costs of building and expanding their facilities. So we should expect more peer entry in cities with high hotels' fixed costs. A second reason for high hotels' revenue per available room has to do with demand trends and fluctuations. First, since hotels must pre-commit to capacity and any adjustment in the form of new hotel buildings takes 3 to 5 years to complete, unforeseen growth in demand will create an inefficiently low hotel supply. Peer hosts on the other hand can use their spare rooms to host travelers, so they can respond a lot more quickly than hotels to growth in demand for accommodations. Second, even if demand is not trending upward over time, there can be large fluctuations in demand during high and low travel seasons. It is typically inefficient for hotels to have enough dedicated capacity to absorb all potential travelers in times of peak demand, because that would lead to many unoccupied rooms most of the year. In contrast, flexible sellers are able to provide additional supply during peak times, when their rooms are especially valuable to travelers. This implies that we should expect higher demand growth and higher demand variability to be predictive of entry of peer suppliers.

In addition to factors influencing the price that hosts can expect to receive for hosting strangers in their apartment, the monetary and non-monetary costs of hosting also play an important role. Although many factors affect the costs of hosting, we focus on those

[12] Other averages of Airbnb share of available rooms, such as the average in December 2015, or the average for all of 2015 or any other year in the sample, lead to similar comparisons of Airbnb size across cities.

related to demographics.[13] Households vary in their propensities to host strangers in their homes. For example, an unmarried 30-year-old professional will likely be more open to hosting strangers than a family with children. This occurs for at least two reasons. First, children increase a host's perceived risk of the transaction. Second, unmarried professionals are more likely to travel, creating vacant space to be rented on Airbnb.

How do we measure hotel fixed costs, demand growth and volatility, and peer hosts' marginal costs? We use two proxies for hotel fixed costs. The first is the share of undevelopable area constructed by Saiz (2010). The index measures the share of a metropolitan area that is undevelopable due to geographic constraints, e.g. bodies of water or steep mountains. The second index is the Wharton Residential Land Use Regulatory Index (WRLURI), which measures regulation related to land use in each metropolitan area and is based on a nationwide survey described in Gyourko et al. (2008).[14]

We use data from air travelers to proxy for accommodation demand trends and fluctuations at the city-month level. Although we are simply interested in predicting peer entry at this point, we measure these demand characteristics during the earliest years in our sample, in order to reduce the risk that peer entry influences demand rather than vice versa. Our data come from Sabre Travel Solutions, the largest global distribution systems provider for air bookings in the US. We isolate trips entering a city as part of a round trip from a different city in order to measure the potential demand for short-term stays.[15] With the Sabre data we compute the 2012-2011 growth rate in travelers to a city and the standard deviation of incoming travelers in 2011. Finally, we proxy for peer hosts costs with the share of unmarried adults and the share of children in a Metropolitan Statistical Area using data from the Census Bureau.[16]

We use all these predictors in a linear regression of Airbnb penetration:

$$\begin{aligned} share\_airbnb_m = & \alpha_1 saiz_m + \alpha_2 wrluri_m + \alpha_3 share\_child_m + \alpha_4 share\_unmarried_m + \\ & \alpha_5 airpass\_sd_m + \alpha_6 airpass\_growt_m + \alpha_7 log(revpar)_m + \\ & \alpha_8 log(market\_size)_m + \epsilon, \end{aligned} \tag{1}$$

where m denotes one of the 46 cities for which we have complete data and *share_airbnb* is

[13]Other potential shifters of the returns to hosting include household liquidity constraints, building regulation and enforcement of short-term rentals, and the ease of vacating an apartment in high demand periods.

[14]Saiz (2010) uses these two measures to calculate the housing supply elasticity at the level of a metropolitan area.

[15]Data from Sabre include monthly number of passengers by origin and destination airport. We aggregate these observations to a Metropolitan Statistical Area-month measure of air travelers.

[16]Appendix Table A4 displays summary statistics for the cost and demand factors that we use as predictors of Airbnb penetration. Appendix Figure A7 displays raw correlation plots between each predictor and Airbnb penetration.

the Airbnb share of available rooms in the last quarter of 2015. We divide the standard deviation of incoming air travelers by 10,000 to make the coefficient comparable to the other variables. Market size, which was not defined above, is the number of available hotel and Airbnb rooms in the last quarter of 2015. We control for market size in order to isolate the component of the standard deviation of demand which is due to demand variability.

Table 2 displays regression results, where the two columns differ by the inclusion of revenue per available room as a predictor. Despite the small sample size and the inclusion of potentially redundant proxies for costs and demand, column (1) shows that all factors predict the size of Airbnb in the expected direction, and all coefficients are at least marginally significant. In column (2) we add the average hotel revenue per available room in 2011 as an additional control. The coefficient on revenue per available room is positive and statistically significant, while the coefficients on the demand and hotel investment cost proxies decrease in magnitude and some become insignificant. This result suggests that as expected, demand proxies and hotel investment costs affect peer entry mostly through price and occupancy rates. Taken altogether, our cost and demand proxies explain between 67% and 76% of the variation in Airbnb size across our cross-section of US cities.

2.3 Peer Supply Elasticity and Competitive Effects on Hotels

From Figure 1 it is clear that Airbnb bookings fluctuate over time: more rooms are booked during the peak season than in other periods. In this section, we use instrumental variable regressions to document that flexible suppliers are almost three times as elastic as dedicated suppliers. Combined with the differential entry of Airbnb across cities described in Section 2.2, this fact implies that Airbnb impacts hotels' performance differently across geographies and over time because peer hosts compete with hotels more in some cities than others, and more during certain time periods than others. We provide suggestive evidence of what this implies for hotel revenues toward the end of this section. We should note that this section is only suggestive of the directions of the effects we expect. Section 3 presents the full structural model and Section 4 its results.

To measure the average elasticity of Airbnb supply with respect to price, and compare it to that of hotels, we estimate the following equation:

$$log(Q_{mt}) = \chi log(K_{mt}) + \kappa log(p_{mt}) + \mu_{mt} + \epsilon_{mt}, \tag{2}$$

where Q_{mt} is the number of (hotel or Airbnb) bookings in city m and day t, K denotes capacity – the number of available rooms for hotels and listings for Airbnb – and p is the average transacted price. The equation is estimated separately for hotels and Airbnb. κ

is the elasticity of supply with respect to prices, and will be different between hotels and Airbnb supply. μ_{mt} includes city fixed effects, seasonality (month-year fixed effects), and day of week fixed effects. These fixed effects control for the fact that costs might change by city or over time, for example due to average differences in costs across cities or due to particular periods when hosts are less likely to occupy their residences.

Equation 2 suffers from standard simultaneity bias because the price of accommodations is correlated with demand, and with unobserved fluctuations in marginal costs. Furthermore, in the case of Airbnb, the number of available rooms K_{mt} is itself endogenous because as the beginning of Section 2 showed, hosts may update availability as a function of demand.[17] We discuss each concern in order.

We instrument for price with plausibly exogenous demand fluctuations which are typically caused by holidays or special events in a city. We use two instruments. The first is the number of arriving (not returning) flight travelers in a city-month, which we used in Section 2.2. The second comes from Google Trends, which provides a normalized measure of weekly search volume for a given query on Google. Our query of interest is "hotel(s) m", where m is the name of a US city in our sample. We de-trend each city's Google Trends series using a common linear trend to remove long-run changes in overall search behavior on Google. We use the one-week lagged search volume as an instrument. Using other lags or the contemporaneous search volume yields similar estimates. One might be worried about reverse causality, that is in fact the availability of Airbnb rooms that leads tourists to travel or search for hotels in particular destinations. While we cannot completely rule that out, the relatively small share of Airbnb bookings (across all cities less than 3%), at least until the end of our sample period, suggests that this is unlikely.

To control for the fact that room availability on Airbnb is endogenous to demand, we instrument for the number of available listings with the number of active listings, since this metric is less responsive to contemporaneous demand shocks, but is highly correlated with the number of listings that are available for rent. We leave the first stage regression results to Appendix Table A5. For the first stage of hotel supply as well as the first stage of Airbnb supply, we reject both the hypotheses of under and weak identification and we cannot reject the hypothesis that the joint set of instruments are valid.

Table 3 contains our IV estimates of Equation 2 for Airbnb and hotels separately. Turning first to column (1), a 1% increase in the average hotel daily rate increases hotel bookings by 1.3%. This elasticity is about a third as large as that of Airbnb, which is displayed in

[17]The same endogeneity issue is not important for hotels because hotel capacity is typically fixed in the short run. Indeed Appendix Figure A6 confirms that the number of hotel rooms has been fairly stable over the course of our sample period, with the exception of New York.

column (2) and is estimated to be 3.9. An important implication of this result is that smaller fluctuations in prices are needed for Airbnb supply to adjust upward or downward.[18]

We have shown that the Airbnb supply is highly responsive to price, more so than hotels: a small price increase due to high demand greatly increases the number of booked rooms on Airbnb, and this increase is almost three times as large as for hotels. The lower elasticity of hotel supply has a simple explanation, which will become clearer in our structural model. To the extent that hotels have a constant marginal cost and a fixed supply, hotel bookings cannot increase in response to increases in demand when demand is sufficiently high. The higher elasticity of flexible supply implies that there are many hosts willing to rent their rooms when prices are high, but prefer not to host when prices are just a little lower.

Where and when peer hosts decide to enter has implications for hotel outcomes, which we focus on next. Since peer hosts enter more in cities with high hotel revenues, we should expect the competitive effect of Airbnb on hotels to be greatest exactly in those geographies. To test this, we estimate the effects of peer entry on hotels' revenue, occupancy rates, and prices, and how they differ by city. Before describing our empirical strategy, we discuss the two most important challenges to identifying the effect of Airbnb. We consider the hypothetical scenario where Airbnb supply grows randomly across cities and over time. In this scenario, regressing hotel outcomes on the number of Airbnb available rooms would yield an unbiased estimate of the causal effect of Airbnb. However, as highlighted above, Airbnb does not grow randomly. In fact, Airbnb is larger in cities with high hotel revenues, and during periods of high demand within each city. Observables like the number of arriving air travelers, city fixed effects, and seasonality fixed effects, help us control for this selection, but do not completely solve the endogeneity problem. So we instrument for the currently available Airbnb supply with the number of active listings.

We need to identify the cities where we would expect the effect of Airbnb to be biggest. To do this we divide our cities into two groups. Saiz (2010) uses the WRLURI and the share of undevelopable area described in Section 2.2 to estimate the housing supply elasticity at the city level. We take that supply elasticity as a proxy for the elasticity of hotel construction, and split our sample of cities at the median level of Saiz's estimates for the cities in our sample. Cities with a low supply elasticity are likely to have a more constrained hotel capacity, hence we should expect the effect of Airbnb to concentrate in those cities.

[18] Instrumenting for prices and Airbnb available listings is important. Appendix Table A6 presents OLS estimates of Equation 2. As one would expect, with OLS we underestimate the elasticity of supply to price and overestimate the elasticity of supply to available listings. Results do not change if we adjust for the fact that Airbnb listings are typically occupied by more guests than the average Airbnb room, an adjustment we describe in Section 3.1.

Our baseline regression specification is:

$$\begin{aligned} y_{mt} =& \alpha_1 \log(airbnb_{mt}) + \alpha_2 \log(airbnb_{mt}) * constrained_m + \\ & \beta_1 \log(hotel\ rooms_{mt}) + \beta_2 \log(hotel\ rooms_{mt}) * constrained_m + \\ & \gamma \log(gtrend_{mt}) + \delta \log(travelers_{mt}) + \theta_{mt} + \nu_{mt}. \end{aligned} \tag{3}$$

Here y_{mt} is one of three hotel outcomes (log revenue per available room, log price, occupancy rate) in a city m on day t, $airbnb_{mt}$ is the number of available Airbnb listings (instrumented for with the number of active listings), $hotel\ rooms_{mt}$ is the number of available hotel rooms, $gtrend_{mt}$ is the one-week lag of Google searches for hotels in the city, $travelers_{mt}$ is the number of arriving air passengers, and $constrained_m$ is equal to 1 if the housing supply elasticity estimated by Saiz (2010) is below the median value. The vector θ_{mt} includes city fixed effects, quarter-year fixed effects and their interaction with the $constrained_m$ dummy, and day of the week fixed effects and their interaction with $constrained_m$. Importantly, the Google metric captures demand shocks at the week level, while the number of incoming air passengers captures monthly fluctuations in demand. The fixed effects capture seasonality, differences across the days of the week, and time-invariant city characteristics that affect both the size of Airbnb and hotel revenue.

The effects of interest are α_1 and α_2. α_1 is the average short-run elasticity of hotel outcomes to peers' supply over our sample period for cities with unconstrained hotel supply. α_2 is the additional effect in cities with constrained hotel supply. The coefficients are identified off of two types of variation. First, there is variation across cities and over time in the number of available listings due to increasing awareness of Airbnb. Second, there is variation in the availability of listings due to hosts' daily costs of hosting, for which we assume the instrument takes care of removing parts that might be correlated with residual daily demand for accommodations within the city.

Table 4 displays the results of the baseline specification. The coefficient on Airbnb size in column (1) is close to zero and statistically insignificant, while the coefficient on the interaction term is negative and statistically different from zero at the 5% confidence level. This coefficient implies that a 10% increase in available listings decreases the revenue per hotel room by 0.57%. The coefficient estimates for our demand proxies, Google trends and arriving air travelers, are of the correct sign and statistically significant. The same is true for the coefficient on hotel rooms. Once we break down the effect into a reduction in occupancy rates (column 2) and a reduction in prices (column 3), we see that the negative effect of Airbnb is mostly concentrated on prices in cities with constrained hotel capacity.[19]

[19] As before, we present first stage regression results in Appendix Table A7, OLS results in Table A8, and

Differences in the effect of Airbnb on hotels across constrained and unconstrained cities occur for two reasons. First, for the same level of Airbnb and hotel capacity, the effect of Airbnb is relatively larger on prices if hotel capacity constraints are more often binding (due to higher levels of demand). Second, for the same level of demand and hotel capacity, the effect on hotel revenues is larger if Airbnb listings constitutes a larger share of available rooms. Intuitively, the elasticity of hotel revenues with respect to the size of Airbnb should be higher, the higher the Airbnb share of supply because a 1 percent increase in Airbnb size is a much bigger share of market supply when Airbnb penetration is 3% then when it is 1%. Both conditions are true when we split our cities. Indeed, in December 2015 the average Airbnb supply share in hotel-constrained cities was 5.8% while it was only in 2.2% in unconstrained cities. At the same time, the average hotel occupancy rate was 62.2% in constrained cities and only 55% in unconstrained cities.

Before concluding this section, one caveat is in order. In these specifications we cannot take advantage of exogenous changes in price that would allow for a valid causal estimate of the effect of Airbnb on hotel performance, something we can do in the next sections with a structural model. However, this exercise helped us highlight a few facts from the data. We documented that the entry of peer hosts is higher where hotels' fixed costs are also high, where peers marginal costs are low, and where demand is increasing and highly variable. We have also shown that flexible supply is highly elastic, and almost three times as elastic as dedicated supply. Finally, we have shown that the entry of flexible supply has negative spillovers on the revenue of dedicated suppliers. This negative effect is concentrated in cities with binding hotel capacity constraints and predominantly impacts hotel prices rather than occupancy rates. In the rest of the paper, we focus on the 10 cities that experienced the largest entry of Airbnb, 9 of which are in the group of cities with binding hotel capacity constraints, and consider how the elastic Airbnb supply affects consumers and hotels over time.[20] This allows us to quantify how hotel capacity constraints and elastic peer supply contribute to the welfare of the agents in the market.

effects by hotel scale in Table A9. The coefficient on Airbnb listings is a statistically significant 0.021 in column (3), which suggests that some spurious correlation may still be present. Table A9 suggests that most of that correlation comes from luxury hotels, while for other hotel scales the coefficient estimate is smaller and statistically indistinguishable from zero.

[20]The 10 cities include: Austin, Boston, Los Angeles, Miami, New York, Oakland, Portland, San Francisco, San Jose, and Seattle. Austin is the only city without binding hotel capacity constraints per our definition.

3 Model and Estimation Strategy

In this section, we describe a short-run model that we use to estimate welfare gains from the entry of flexible supply. In our model, hosting services can be provided by dedicated and flexible sellers, who offer differentiated products. The equilibrium consists of daily prices and rooms sold by each accommodation type as a function of the overall demand level and the respective capacities of dedicated and flexible suppliers. We assume hotels are competing against a fringe of flexible sellers. Appendix A presents a version of this model with only one hotel type and one type of flexible hosts, but with more general demand and cost specifications. In that appendix, we prove existence and uniqueness of the equilibrium, as well as comparative statics predictions that are in line with the stylized facts from Section 2.

A market n is defined by day t and city m. On the demand side, our model is a random coefficients logit model (Petrin (2002) and Berry et al. (1995)), where rooms are differentiated across hotel scales and Airbnb listing types. On the supply side, we assume that hotels engage in Cournot competition with differentiated products across scales. Within a scale, each hotel is undifferentiated. Airbnb hosts are price takers with randomly drawn marginal costs.

Consumer Demand

Consumers make a discrete choice between hotel scales, Airbnb listing types, and an outside option for a given night. Consumer i has the following utility for room option j in market n:

$$u_{ijn} = \mu_{ijn} + \alpha_i (1 + \tau_{jn}) p_{jn} + \epsilon_{ijn}. \tag{4}$$

For consumer i, μ_{ijn} represents a mean utility for accommodation j in market n inclusive of preference heterogeneity for the inside options. The price of an accommodation is denoted p_{jn}, while τ_{jn} represents the percent difference between what the travelers pay and what the suppliers receive for accommodation j. For hotels, τ_{jn} is simply the lodging tax rate. For Airbnb rooms, it is a combination of the Airbnb commission fee and the lodging tax rate if it is collected by Airbnb.[21] Finally, ϵ_{ijn} is an idiosyncratic component with a type I extreme value distribution. We normalize the value of the outside option to 0 for all markets. This

[21] We collect the lodging tax rate from HVS Lodging Tax Reports for hotels (https://www.hvs.com/indepth/ accessed January 2021). For Airbnb, for each listing type, city, and night we have the average price paid by travelers, the average price received by hosts, the average tax collected, and the average amount kept by Airbnb as commission, from which can compute the Airbnb's commission fee and lodging tax rate if applicable.

demand specification yields the following quantities for each accommodation type:

$$Q_{jn}(p_{jn}, p_{-jn}) = D_n \int \frac{e^{\mu_{ijn}+\alpha_i(1+\tau_{jn})p_{jn}}}{1+\sum_{j'} e^{\mu_{ij'n}+\alpha_i(1+\tau_{jn})p_{j'n}}} dH(i), \tag{5}$$

where D_n is the market size, and H is the joint distribution of consumer heterogeneity. We allow for consumer heterogeneity in how travelers value the inside options (hotels and Airbnb), since this gives the model flexibility in determining what share of Airbnb travelers would substitute towards hotels in the absence of Airbnb. We also allow for consumer heterogeneity in sensitivity to price. We assume that the distribution of consumer heterogeneity is multivariate normal with a mean and variance matrix to be estimated. We do not allow for correlation across distinct components of consumer heterogeneity.

Hotel Supply

Each hotel competes with other hotels of the same scale, hotels of different scales, and peer supply. We assume that this competition takes the form of a Cournot equilibrium. Hotels of type h, where $h \in \{$*luxury, upper-upscale, upscale, upper-midscale, midscale, economy*$\}$, have aggregate room capacity K_{hn}. Since there are multiple hotels within each scale, we need to distinguish between scale-level and hotel-level quantities. We let Q_{hn} denote the scale-level number of rooms sold. We assume no differentiation in room quality within scale, so the number of rooms sold by each hotel, denoted q_{hn}, is the ratio of aggregate quantity divided by the number of hotels. Analogously, scale-level capacity is denoted K_{hn}, while hotel-level capacity is k_{hn}.[22]

We must also match the fact that prices increase sharply as the number of rooms sold approaches the number of available rooms. In practice, occupancy rates never reach 100% at the scale level, but prices start increasing before then (Figure 4). This is because, although we model hotels as homogeneous within each scale, some individual hotels may sell out before others and this may result in sharply increasing scale-level prices. In addition, if hotels face uncertainty about the actual level of demand when setting prices, increases in expected demand will increase the probability of hitting capacity constraints, thus increasing prices before realized demand reaches 100%. We allow our model to fit this increasing price profile by estimating an increasing cost function for hotels that kicks in as soon as hotel occupancy is at least 85% within a scale. The estimation of increasing marginal costs as production approaches capacity constraints was previously used by Ryan (2012) to estimate the cost structure of the cement industry.

[22]STR provides us with the number of hotels in a given scale, day, and city.

For these reasons, we assume that hotels' variable costs are made of two parts: a constant marginal cost c_{hn}, and an increasing marginal cost $\gamma_{hn}(q_{hn} - \nu k_{hn})$, which starts binding as quantity approaches the capacity constraint. Given the above discussion, we set $\nu = 0.85$. So, instead of solving a maximization problem subject to a capacity constraint, each hotel selects its quantity to maximize the following profit function:

$$\underset{q_{hn}}{Max}\ q_{hn}p_{hn}(Q_{hn}, Q_{-hn}, Q_{an}) - q_{hn}c_{hn} - \frac{\gamma_{hn}}{2}\mathbb{1}(q_{hn} > \nu k_{hn})(q_{hn} - \nu k_{hn})^2.$$

We assume that hotels observe all components of demand and competitors' costs, so that there is no uncertainty about whether $q_{hn} > \nu k_{hn}$ or not. Letting N_{hn} denote the number of hotels within scale h, we have that $q_{hn} = \frac{Q_{hn}}{N_{hn}}$. Taking advantage of the implicit function theorem, the optimization problem gives rise to the following first order condition:[23]

$$p_{hn} = -\frac{1}{N_{hn}}\frac{Q_{hn}}{Q'_{hn}} + c_{hn} + \gamma_{hn}\mathbb{1}(q_{hn} > \nu k_{hn})(q_{hn} - \nu k_{hn}), \tag{6}$$

where Q_{hn} is scale-level room demand from Equation 5, and Q'_{hn} is the derivative with respect to its own price.

Peer Supply

Peers of each quality type a, where $a \in \{$*Airbnb luxury, Airbnb upscale, Airbnb midscale, Airbnb economy*$\}$, with total available listings K_{an}, take prices as given. Hosts draw marginal costs from a normal distribution with mean ω_{an} and standard deviation σ_{an}. Each draw is iid across hosts and time. Hosts of type a choose to host only if the price p_{an} is greater than their cost. Therefore, the quantity supplied will be determined by the following equation:

$$Q_{an}(p_{an}, p_{-an}, p_{hn}) = K_{an}Pr(c \leq p_{an}) = K_{an}\Phi\left(\frac{p_{an} - \omega_{an}}{\sigma_{an}}\right). \tag{7}$$

Equilibrium

The market equilibrium consists of prices and quantities for hotels and peer hosts (p_{hn}, p_{an}, Q_{hn}, Q_{an}) such that consumers, hotels, and peer hosts make decisions to maximize their surplus, and their optimal choices are consistent with one another.

[23] The objective function is not differentiable at $q_{hn} = \nu k_{hn}$, but otherwise the first order condition holds everywhere else.

3.1 Estimation Strategy

We estimate the demand, hotel supply, and peer supply separately.

Starting first with demand, the high-level choices are the market size, the moments to match, and the instruments used. But first we need to make a normalization. Since Airbnb listings can on average host more guests than hotel rooms, we adjust quantities so that room capacity is comparable across Airbnb listings and hotel rooms. To do this, we take advantage of the fact that we have information on the average number of guests for Airbnb bookings. In addition, lower quality Airbnb listings are typically private rooms with smaller capacity than standard hotel rooms. For this reason, we assume that each hotel room is occupied by as many people as the average number of occupants of Airbnb Midscale listings in the same city. Given this adjustment, our quantities, prices, and estimates should be interpreted as referring to room-nights with standard hotel occupancy.

We use data on the 10 largest cities in terms of the share of Airbnb bookings in our sample.[24] Our estimation sample starts in 2013 and continues until July 1, 2015. We restrict the sample in this way for three practical reasons. First, in other cities and time periods market shares of Airbnb are often close to zero, which creates complications for estimation. Second, the reduced form results in Section 2 suggest that the effects of Airbnb in those markets will be limited when Airbnb market shares are close to 0. For the same reason, we also drop Airbnb options if their share of available rooms is less than 0.5% on a given day and city. Finally, we exclude the second half of 2015 and use it to validate our estimates out of sample.

One key choice we must make in the estimation is D_n, the total number of consumers considering to book accommodations. The choice of D_n will affect market shares for hotels and Airbnb, as well as the share of potential travelers choosing to stay home, to travel to other locations, or to stay in alternative accommodations, e.g. friends and family. We set D_n equal to two times the average number of rooms booked in the corresponding month in each city in 2012. This assumption allows the potential number of travelers to vary seasonally across cities, and it allows for both substitution from hotels – hotel travelers switching to Airbnb – and market expansion – travelers switching from the outside option to Airbnb. We rationalize any remaining variation over time in the total number of travelers booking accommodations with mean utilities for inside options that vary as a function of unobservable and observable characteristics.

The second choice is the set of moments that we match to the data. We construct two types of moments for the demand estimation: the standard BLP moments (market share

[24] The 10 cities include: Austin, Boston, Los Angeles, Miami, New York, Oakland, Portland, San Francisco, San Jose, and Seattle.

moments) and a moment disciplining the estimated model to match survey data on the hypothetical choice of Airbnb users if Airbnb did not exist (substitution moment).

Our market share moments are

$$m_{1jn} = \left[\delta_{jn} - \hat{\delta}_{jn}\right] Z_{jn}, \tag{8}$$

where δ_{jn} is the realized mean utility from accommodation j in market n that rationalizes the observed market shares, and $\hat{\delta}_{jn}$ is the mean utility predicted from the vector of parameters to be estimated. $\hat{\delta}_{jn}$ is the component of utility from Equation 4 that does not differ across individual travelers, and is a function of observable and unobservable characteristics of the different types of accommodations. In addition to prices, utility is a function of day of week fixed effects; city-scale-month fixed effects to account for different preferences across quality tier, location, and seasonality; city-specific and Airbnb-city-specific linear time trends; and the log of 1-week lagged Google searches for hotels in the city.

The vector Z_{jn} includes all determinants of utility described above except for prices. Given price endogeneity and consumer preference heterogeneity, we exploit supply-side variation that affects prices and substitution across options. Our first instrument takes advantage of the fact that hotel capacity constraints affect prices when they are binding but are uncorrelated with daily demand shocks. In particular, a change in demand when capacity constraints are binding will have a much bigger effect on prices than when they are not binding. We proxy for this effect by using the ratio of the log of Google searches for hotels and the available hotel rooms. Our next instrument is the lodging tax rate, which may be different between hotel and Airbnb options. The lodging tax rate changes for two reasons in our sample. First, some local authorities change the rate. Second, Airbnb starts collecting lodging taxes on behalf of certain jurisdictions. Finally, as in our reduced form, we use variation in hotel and Airbnb capacity. We use the number of hotel rooms and the number of active Airbnb listings, and we interact them with scale fixed effects.

The substitution moment comes from survey data on alternative accommodation choices of travelers booking on Airbnb. Airbnb has conducted surveys of guests in four of the sampled cities during 2013 and 2014. The surveys asked the following question: "If Airbnb had not been available, what would you have done?" Between 19% and 42% of guests across cities said that they would not have booked a hotel, effectively choosing the outside option. A simple average across cities yields a substitution share of 32% towards the outside option, which we use in our estimation.[25]

[25]In 2015, Morgan Stanley and AlphaWise conducted a representative survey of 4,116 adults in the US, UK, France, and Germany. In the survey, they asked respondents about their travel patterns. 12% of respondents had used Airbnb within the past year and when asked which travel alternative Airbnb replaced,

We match the survey responses in our model by computing the share of Airbnb travelers who would have wanted to book a hotel at the observed prices had Airbnb not been available. To predict the share of Airbnb travelers choosing hotels in the absence of Airbnb, we first note that the share of travelers choosing the outside option in market n is $s_{on} = \int \frac{1}{1+\sum_{j'} e^{\mu_{ij'n}+\alpha_i(1+\tau_{jn})p_{j'n}}} dH(i)$. Airbnb's market share, denoted $s_{airbnb,n}$, is equal to the sum of the market shares of each Airbnb option available in market n. If Airbnb listings were not available, the market share of the outside option would be $s_{on^*} = \int \frac{1}{1+\sum_{j'\in\text{hotel}} e^{\mu_{ij'n}+\alpha_i(1+\tau_{jn})p_{j'n}}} dH(i)$. Therefore in a specific market n we compare the ratio $\frac{s_{on^*}-s_{on}}{s_{airbnb,n}}$ with 32%, the survey's share of Airbnb travelers choosing the outside option:

$$m_{2n} = \left(100\frac{s_{on^*} - s_{on}}{s_{airbnb,n}} - 32\right).$$

When we sum the substitution moments across markets, we weigh each market with the same set of available Airbnb options equally. For example, markets where only 'Airbnb Luxury' options are available receive a weight equal to their share of Airbnb rooms sold $(\frac{\sum_{n' \text{ with Airbnb luxury only}} s_{airbnb,n'}D_{n'}}{\sum_n s_{airbnb,n}*D_n})$. This results in the highest weight being placed on markets where all Airbnb options are available, which is the most frequent type of markets. In the data we have 15 possible combinations of Airbnb options available. We thus get the following aggregate moment:

$$m_2 = \frac{1}{N}\sum_{i=1}^{15}\left[\frac{\sum_{n' \text{ has Airbnb options in group i}} s_{airbnb,n'}D_{n'}}{\sum_n s_{airbnb,n}D_n}\sum_{n' \text{ has Airbnb options in group i}} m_{2n'}\right]. \tag{9}$$

where N is equal to 9,110, the number of markets.

It is useful to give an intuition for how the variation in the data allows us to estimate the demand parameters. Our descriptive statistics show that the prices of hotels and Airbnb options, unadjusted for different number of occupants, are similar. This fact, together with the relatively high substitution rate between hotels and Airbnb rooms derived from survey responses, suggests that the mean utilities of hotels and Airbnb options should be fairly similar. However, in practice we also observe very different market shares, much higher for hotels than for Airbnb options. The market share and substitution moments help us rationalize these two patterns in the data. On one hand, the substitution moment helps us identify consumer preference heterogeneity (the random coefficients on price and the

58% of respondents answered something other than a hotel. See Nowak et al. (2015). We think that the major reason for the differences between the Airbnb and Morgan Stanley surveys is that Morgan Stanley sampled guests to all types of destinations including resorts and European cities. There are typically more non-Airbnb and non-hotel options for guests in these locations.

inside option). On the other, differences in market shares rationalize mean utilities that will be higher for hotels than for Airbnb options. We discuss computational details and the sensitivity of our estimates to our identifying assumptions in more detail in Appendix B.

Once we obtain demand estimates that let us compute Q_{hn} and its price derivative, we estimate the supply function from equation 6 using a linear IV approach:

$$p_{hn} + \frac{1}{N_{hn}} \frac{Q_{hn}}{Q'_{hn}} = \theta X_{hn} + \gamma_{hn} \mathbb{1}(q_{hn} > \nu k_{hn})(q_{hn} - \nu k_{hn}) + \epsilon_{hn}. \tag{10}$$

X_{hn} includes city-scale fixed effects, city-day of the week fixed effects, year-month fixed effects, and city-specific linear time trends. We allow for γ_{hn} to vary by city and by scale separately. We instrument for the increasing cost component using interactions of the lagged Google search trend with city fixed effects and hotel fixed effects. We use these instruments because they affect hotel prices only by increasing the likelihood that capacity constraints bind. The supply equation is then estimated jointly using all markets.

Finally, the supply of Airbnb can be estimated separately using another linear IV regression for the same sample period. Equation 7 implies that $\Phi^{-1}\left(\frac{Q_{an}}{K_{an}}\right) = \frac{\omega_{an}}{\sigma_{an}} + \frac{1}{\sigma_{an}} p_{an}$, where the left-hand side is the inverse of a standard normal cumulative distribution function calculated at a value equal to the share of booked rooms out of all Airbnb active listings. We estimate the following specification

$$\Phi^{-1}\left(\frac{Q_{an}}{K_{an}}\right) = \beta_a p_{an} + \gamma_a X_{an} + \epsilon_{an}, \tag{11}$$

where K_{an} is the number of active Airbnb listings of type a, p_{an} is the average transacted price of Airbnb type a in market n, and just like for the hotel supply regression, X_{an} includes city-scale fixed effects, city-day of the week fixed effects, year-month fixed effects, and city-specific linear time trends. We instrument for the transacted price with the log of Google search trends and the log of incoming air passengers.

After estimating the above equation, we can transform the coefficients into the following peer cost parameters:

$$\sigma_{an} = \frac{1}{\beta_a},\ \omega_{an} = \frac{\gamma_a X_{an} + \epsilon_{an}}{\beta_a}.$$

4 Results

In this section, we discuss the results of our estimation. We first go over our estimated parameters.[26] Then we discuss the effects of Airbnb and government regulation on consumer surplus, on hotels' and hosts' bookings, revenues, and surplus, and on lodging taxes.

4.1 Parameter Estimates

Table 5 displays the estimates of demand parameters that are common across cities and accommodation options. We first discuss the parameters governing the distribution of price sensitivity across travelers. The mean price coefficient is -.031 and the standard deviation is .004. The standard deviation is imprecisely estimated, but our estimates are consistent with existing work on hotel demand (Koulayev (2014)). Google search trends are estimated to have a positive effect on demand. We also estimate some level of heterogeneity in preferences for booking the inside option (a hotel or Airbnb room), although the coefficient is not significant at the 5% confidence level. A comparison between the first and the second columns in the table shows that coefficient estimates for mean utility parameters are not very different from standard logit estimates without consumer preference heterogeneity.

Figure 5 displays the mean willingness to pay per night for each accommodation option and city at the end of 2014. The fact that some values are negative reflects our choice of a market size that is two times the average number of booked rooms in a city-month in 2012. When looking at the mean utilities in relative terms, our estimates show that willingness to pay tends to be decreasing between luxury and economy hotels and between Airbnb luxury and economy listings. The value of the top Airbnb option is lower than the value of the lowest hotel option across all cities, with some variation in the relative differences. We cannot distinguish between alternative explanations for this difference. There are several potential explanations, including the fact that some people may not have heard of Airbnb, that business travelers were often not able to use Airbnb for their business travel, and that Airbnb did not offer complementary services such as concierge, 24-hour check-in, and daily cleaning. Within Airbnb options, distance to visitor centers in a city (a proxy of desirability for leisure travelers) is significantly correlated with consumers' willingness to pay (see Appendix C for more details).

We find that demand for accommodations is elastic on average. For example, in San Francisco, the demand elasticities range between -8.63 for luxury hotels and -2.90 for the

[26] Appendix Table A18 shows that our ability to match observed market shares is similar in and out of sample.

lowest quality of Airbnb listings.[27] The surprisingly large demand elasticity for luxury hotels is due to the fact that the limited consumer preference heterogeneity that we estimate does not completely offset the fact that demand elasticity is a function of market shares, a well-known characteristic of logit demand systems. There is also substantial variation across cities in demand elasticities, ranging between -2.58 in Portland and -6.12 in New York for midscale hotels.

Next, we turn to the estimates of hotel cost parameters. Our parameter estimates are precise and the estimation procedure explains most of the variation with an R-squared of 0.79. The interquartile range for the errors is -\$14 to \$17. Figure 6 plots the marginal cost curves for different hotel scales and different cities at the end of 2014. We find that the constant components of hotels' marginal costs have the expected relationship with hotel quality. The marginal cost for luxury hotels in New York city are \$371 on average while it is \$144 for economy hotels. We want to be careful in not interpreting these costs as actual expenditures per night-booked. Research by Kalnins (2006) suggests that due to reputational concerns, hotels tend to have a price threshold below which they will not go, and this threshold is typically higher than the cost of an additional maid- or clerk-hour. We view our estimates as reflecting this price threshold. The figure also plots the increasing component of hotels' marginal costs. We find that for all the city and hotel scale combinations, marginal costs increase relatively steeply with quantity when hotel occupancy reaches 85%. This increasing cost reflects the fact that regardless of the level of competition, hotels will increase their prices as they approach full capacity.[28] A comparison of these estimates with our reduced form results is reassuring. Indeed, the implied supply elasticities from these estimates are very close to those estimated in the reduced form section. The average supply elasticity across all markets and hotel scales is 1, which is comparable to the reduced form estimate of 1.3 from Table 3.

Finally, Figure 7 displays the mean costs over time for Airbnb listings in New York City. Costs vary over the course of the year, with higher costs during the winter season. In New York but also in other cities, costs increase monotonically in listing quality, and the mean costs exceed the mean transacted prices. These relatively high costs stem from the fact that fewer than 50% of active listings on Airbnb typically transact (Table 1).[29] With an R-squared of .42, the variation in our data can explain a little less of Airbnb costs than hotels' costs. However, we estimate economically and statistically significant dispersion in the cost distribution for all listing types, which explains the high supply elasticity of Airbnb

[27] Appendix Table A11 shows the city-specific elasticities of demand for different accommodations with respect to their own price and Appendix Table A12 shows the average cross-price elasticities.

[28] Appendix Tables A14 and A15 report the full set of cost estimates by city and scale.

[29] Appendix Table A17 displays the full set of estimates of Airbnb costs by listing type and city.

accommodations. As with the hotel estimates, the implied supply elasticities from the Airbnb cost estimates are very close to those estimated in the reduced form section. The average supply elasticity across all markets and listing types is 3.4. This is comparable to the reduced form estimate of 3.9 from Table 3. The lowest quality Airbnb listings are the most elastic, with an average supply elasticity equal to 4. Elasticity monotonically decreases as the listing quality increases, and top quality listings have an elasticity of 3.1.

4.2 Counterfactual Analysis

With the estimates in hand, we perform three types of counterfactuals and measure differences between these counterfactual and the status quo (*'Baseline'*). The first type of counterfactuals measure the welfare effects of Airbnb by removing it. The second type of counterfactuals consider the effects of proposed regulatory policies. The final counterfactual considers the implications of additional Airbnb growth. Appendix B describes how we compute the counterfactual equilibria.

Our first counterfactual scenario (*'Unconstrained'*) considers what would happen if Airbnb were removed but hotel prices remained constant and capacity constraints did not bind. In this scenario, travelers who booked on Airbnb are allowed to book any hotel option at the baseline prices, regardless of actual room availability. The unconstrained scenario allows us to measure how much better off consumers are simply because Airbnb offers a new set of options that are valued by at least some consumers. The second scenario (*'No Airbnb'*) allows hotels to adjust prices in response to the absence of competing accommodations on Airbnb. This counterfactual takes into account capacity constraints and requires computing new Cournot equilibria for each market with demand and hotel cost parameters taken from our estimates.[30]

Table 6 presents the effects of removing Airbnb on consumers, hotels, and lodging taxes for all of 2014 (Panel A) and for compression nights in 2014, i.e., nights when at least one hotel scale reaches 95% occupancy in the *'Baseline'* scenario (Panel B). Consumers would lose \$147 million in surplus in the *'Unconstrained'* scenario. Given that in the baseline scenario \$4.38 million rooms were booked on Airbnb, this loss corresponds to \$33.60 per Airbnb room-night, about 16% of the average purchase price.[31] The consumer surplus loss in this scenario only measures one channel through which Airbnb benefits consumers, i.e. product differentiation.

[30] The hotels' first order conditions (Equation 6) do not guarantee that, in the absence of Airbnb, the equilibrium quantities remain below hotel capacity. In practice however, given our parameter estimates the capacity constraints are always satisfied when Airbnb does not exist and hotels reoptimize their choices.

[31] The average purchase price that consumers face in the *'Baseline'* scenario, averaged across both hotel and Airbnb options, is \$209.

Relative to the *'Unconstrained'* counterfactual, there are two additional mechanisms through which the *'No Airbnb'* counterfactual harms consumers. First, travelers who booked on Airbnb but are considering switching to hotels now face higher hotel prices. Second, travelers who previously booked hotel accommodations also face higher prices. The consumer surplus loss in this scenario doubles, rising to $305 million. The vast majority of the difference between the *'Unconstrained'* and *'No Airbnb'* scenarios comes from inframarginal travelers who would book hotel rooms even if Airbnb were available. The price they face does not change much – it only increases by $1 on average from $211 to $212 – but 146 million travelers booked rooms in the *'Baseline'* scenario, resulting in a $155 million increase in expenditures for inframarginal travelers.[32] The remaining $3 million reduction in consumer surplus compared to the loss in the *'Unconstrained'* scenarios comes from the fact that Airbnb travelers considering to book hotels are now facing higher prices.

There are two ways to think about the magnitudes of the effects on consumer welfare. On the one hand, peer production was responsible for just 3% of rooms sold in 2014 and, as a result, the surplus is small relative to the revenue in the market. In particular, hotel and peer hosts' revenues in 2014 were $27,320 million, meaning that the loss of consumer surplus is on the order of 1.1% of aggregate revenues. On the other hand, the benefits to individual consumers are substantial. Indeed, the consumer surplus benefit of Airbnb is $70 per Airbnb room night.

We now turn to the effects of Airbnb on hotels. In the *'Unconstrained'* scenario hotels are able to increase rooms sold by 2% and revenues by 2.3%. Revenues increase by more than quantities because travelers book more Airbnb rooms when aggregate demand, and therefore average prices, are higher. If we take our cost estimates seriously, we can also look at the effect of Airbnb on hotel profits, which we calculate as hotel revenue minus the non-increasing part of the cost function. In the baseline scenario, profits constitute 21% of revenues, which seems a realistic figure. In the *Unconstrained* scenario, profits would increase by 2.5%. The ability to increase prices is what makes up for the capacity constraints and reduced occupancy in the *'No Airbnb'* counterfactual. Indeed, even if rooms sold and revenues only increase by 1.4% and 1.6% respectively, profits increase by 2.9% in the *'No Airbnb'* counterfactual, more than in the *'Unconstrained'* scenario.

Our estimates give us only a rough idea of the changes in hotel surplus for at least three reasons. First, hotels earn additional revenues through complementary services such as conferences and food sales, but also incur additional costs. Second, there are fixed costs involved in operating a hotel which we do not model. Third, our marginal cost estimates

[32]The loss is higher than $146 million because the price increases during compression nights are higher than during non-compression nights, and that's also when more rooms are booked.

correspond in part to reputation costs rather than 'true' marginal costs. These additional costs and revenues do not allow us to state with certainty whether hotel surplus is larger or smaller than our profit estimate.[33]

Not surprisingly, peer hosts would lose without Airbnb. We use the estimated cost distributions of hosts to back out the surplus that they receive from hosting on Airbnb. We truncate the cost distribution at zero, so the surplus for each day can be calculated using the following expression $PS_{an} = \int_{-Inf}^{p_{an}} (p_{an} - max(c, 0)) dF_{an}(c)$. Note that this expression ignores the variable costs of being listed for a given day, which are likely to be negligible, and the fixed costs of entry into the platform. Table 6 displays the number of rooms sold, the total revenues, and host surplus. In the aggregate, peer hosts enjoy \$112 million in producer surplus, which corresponds to \$26 in producer surplus per room-night booked.

The welfare effects are even more pronounced during compression nights (Panel B of Table 6). Although compression nights represent only 19.6% of all markets, the reduction in consumer surplus during compression nights is 40% of the aggregate reduction in consumer surplus in the *'No Airbnb'* scenario. For hotels, the increase in profits during compression nights constitutes 49% of the aggregate profit increase that they would enjoy if Airbnb did not exist. The concentration of the effects during periods of high demand is not due to travelers liking Airbnb more on compression nights—compression nights represent 26.5% of Airbnb baseline bookings and 26.7% of the reduction in consumer surplus from the *Unconstrained* counterfactual. Instead, the effect happens because of hotels' capacity constraints. In fact, during compression nights, the number of hotel rooms sold in the *No Airbnb* scenario remains 33 million, just like in the *Baseline* scenario. But without Airbnb hotel prices increase more during compression nights than during non-compression nights – \$2 versus \$0.60 price increase – with sizable increases in revenue and profits as a result.

Since cities vary in their hotel room capacity relative to demand, the effects of Airbnb exhibit geographic heterogeneity. In particular, since hotel capacity constraints are more often binding in New York and San Francisco, the reduction in consumer and peer host surplus and the increase in hotel revenues and profits if Airbnb did not exist are proportionally bigger there than in cities like Portland or Miami.[34]

Before turning to counterfactuals with regulation, we consider the extent to which Airbnb expands the market versus cannibalizes hotel demand. Appendix Table A22 displays results on the share of Airbnb travelers who would have booked a hotel room in the absence of

[33] In Appendix Table A20 we display the results assuming an alternative measure of costs for hotels imputed from the wage bill of hotels in our data and trends in the wages of maids across cities and over time. This is likely a lower bound on the true marginal cost of hotels.

[34] Appendix Tables A19 through A21 separate the effects of Airbnb on travelers, hotels, and peer hosts by city. Appendix Table A23 uses parameter estimates without consumer heterogeneity to replicate Table 6.

Airbnb. In the *'Unconstrained'* scenario, between 29% and 33% of Airbnb bookings would not have resulted in a hotel booking, which is consistent with the substitution moment used to estimate demand. However, the market expansion effect becomes much bigger when we account for capacity constraints and hotels' price responses. The share of Airbnb travelers who would have not in fact booked a hotel room increases everywhere, ranging from 49% in Austin and Portland to 70% in New York, all the way up to 87% during compression nights.

The next set of counterfactuals explore what would happen to the accommodations market if Airbnb continued to grow and if Airbnb were subject to regulation. The first and most obvious regulation is lodging taxes (*Airbnb with Lodging Taxes*). In this scenario, Airbnb guests are charged a lodging tax rate equal to the rate charged to travelers staying at hotels. Note that for some markets, this scenario is identical to the *Baseline* since Airbnb is already collecting lodging taxes. For the vast majority of the markets however, this scenario implies an increase in the wedge between what the travelers pay and what the hosts receive. Implicitly we assume that in the *Baseline* scenario hosts do not pay lodging taxes out of their share of revenues. To the extent that some hosts were already paying lodging taxes, these numbers should be considered an upper bound on the losses of peer hosts and travelers, and on the gains of hotels and local governments. Table 6 shows that in this case the reduction in consumer surplus compared to the *Baseline* is $65 million, which constitutes only 21% of the consumer surplus loss from *No Airbnb*. This would allow local governments to increase tax revenues by about $72 million – a 1.8% increase – and hotels to increase revenues and profits by $88 million (0.3%) and $31 million (0.5%) respectively. Airbnb hosts, on the other hand, would see both their revenues and surplus decreased by 27% because .9 million fewer Airbnb rooms would be sold.

The second regulatory counterfactual considers quotas. Many local governments have proposed (and some have passed) regulation limiting the number of nights a listing can be booked within a calendar year without the host present at the residence. For example, San Francisco has set the maximum number of nights to 90, while Portland requires that a host reside in a residence listed on Airbnb for at least 270 days within the year, effectively allowing for 95 days that can be booked without the host present at the residence.[35] To proxy for this regulation we consider a scenario (*Airbnb with Quotas*) in which all rooms are allowed to be booked only 90 days in a year, and we choose those days as the days with the highest number of travelers booking Airbnb or hotel accommodations in a given city. In other words, for each city the 90 days in a year with the highest demand will be considered at the

[35] For San Francisco regulation, see https://www.airbnb.com/help/article/871/san-francisco-ca. For Portland, OR, see https://www.airbnb.com/help/article/875/portland-or. France has similar regulation, https://www.airbnb.com/help/article/2108/night-limits-in-france-frequently-asked-questions. All websites were accessed January 2020.

Baseline scenario, while for the remaining days Airbnb will be banned as in the *No Airbnb* counterfactual. To the extent that Airbnb hosts cannot perfectly identify the high demand days ahead of time, this scenario may overestimate Airbnb's benefits to consumers and peer hosts, particularly on those 90 days of high demand. On the other hand, because each host can choose the days when to host travelers independently of other hosts and because there are no quotas if the host is present, this scenario is also likely to underestimate the benefits to consumers and peer hosts during the remaining 275 days in a year.

Table 6 shows that in this scenario consumers would lose 51% of the consumer surplus they would lose if Airbnb were completely banned. Because benefits are concentrated on high demand days, consumers would only lose 12% of the surplus loss from the absence of Airbnb during compression nights. Hotels would not gain as much during compression nights, but the ban on Airbnb during non-compression nights would allow them to still increase revenues and profits by 1% compared to the baseline. Local governments would experience a 1% increase in taxes – levied on the travelers who cannot book on Airbnb during the low demand days of the year – which is about half the tax revenue increase obtainable under the *Airbnb with Lodging Taxes* scenario. Peer hosts would only be allowed to sell 1.8 million rooms in this scenario, obtaining about 35% of the revenues and surplus that they would obtain without regulation.

Finally, motivated by Airbnb's continued growth after our sample, we consider what would happen if Airbnb had twice as many Airbnb active listings drawn from the same cost distribution that we estimated under the baseline scenario (*Double Airbnb Rooms*). This counterfactual estimates the effect of increasing Airbnb supply without changing the utility for these options. The effect of these additional rooms will be smaller than the removal of Airbnb because their main effect is to lower the prices of Airbnb rooms rather than adding additional options. The lower prices would attract travelers who don't have as high a preference for Airbnb as the first Airbnb guests and would put additional pricing pressure on hotels.

Table 6 shows that doubling Airbnb rooms would increase consumer surplus by $130 million. Comparing this to the $305 million loss in consumer surplus if Airbnb did not exist, it implies that the additional Airbnb supply would be about 43% as valuable as the initial supply. The further reduction in hotel revenues and profits is also around 40% of the effect that the initial Airbnb supply had on hotels. For peer hosts, doubling Airbnb supply would bring about 30% as much surplus as the baseline supply level. In this counterfactual, Airbnb rooms do not completely replace hotels as the most common accommodations option, a result of the much lower mean utilities that we estimated for Airbnb compared to hotels.

5 Conclusion

The spread of digital technology has enabled peer production in the accommodations industry. We have studied the welfare implications of this new mode of production for consumers, incumbent providers (hotels), and peer hosts.

We began by describing the industry. The returns to peer hosts vary across cities. Factors affecting hotel room capacity, demand trends and volatility, and and peers' costs of hosting strangers in their homes are all strong predictors of Airbnb penetration across cities. The returns to peer hosts also vary over time. We find that peer host supply is three times as elastic as hotel supply, rapidly expanding when demand and prices increase. The highly elastic host supply implies that the largest effects of Airbnb occur in markets where hotels are often near full capacity, which we confirm with reduced form regressions. In particular, we show that Airbnb entry negatively affects hotel revenues in cities where hotels are more likely to be capacity-constrained, and that the effect is more concentrated on price than on quantity, at least compared to non-capacity-constrained cities.

Our descriptive facts provide intuition for the mechanisms at play when peer supply is allowed to host travelers and compete with hotels. To quantify the welfare effects of peer supply, we then presented and estimated a model of competition between peer hosts and hotels. In addition to confirming the results from our reduced form analysis, our estimates point to sizable benefits of peer supply. The availability of peer hosts generates \$305 million in consumer surplus in 2014 for the 10 largest US cities. About half of that surplus comes from consumers' heterogeneous preferences for accommodations, while the other half comes from competition that reduces prices and expands capacity when it is most needed. In addition, Airbnb generates \$112 million in peer host surplus in 2014, or \$26 per room-night.

We showed that hotels experience the competitive effects of peer hosts. Without Airbnb hotel revenues would be 1.6% higher, even if between 49% and 87% of nights booked on Airbnb would not have resulted in a hotel booking in the absence of Airbnb. These travelers would have instead chosen an alternative option, which could represent staying with friends or family or not traveling at all.

Our analysis informs the active policy debate regarding whether and how to regulate peer-to-peer accommodations. Proposed policies include fees and taxes, mandated registrations, quotas, caps on the number of nights per listing, and outright bans.[36] Our analysis suggests that Airbnb is especially beneficial to consumer and host welfare during peak demand periods in hotel constrained cities. In fact, allowing Airbnb rooms to be booked only 90 days out of a year would recoup 49% of the consumer surplus loss that would occur in the absence

[36] See https://www.airbnb.com/help/article/1376/responsible-hosting-in-the-united-states.

of Airbnb. This result favors a regulatory framework that preserves the benefits of peer production during peak demand days. We also showed that parity in lodging taxes between peer hosts and hotels would raise an additional $72 million in tax revenues while reducing consumer and peer hosts surplus by an amount equal to 23% of the loss that would occur if Airbnb were banned.

Airbnb has continued its rapid growth in both active listings and global awareness in the time after our data sample. Our model suggests that doubling Airbnb supply in 2014, holding everything else constant, could have about 40% of the effect that the baseline level of Airbnb supply had on consumers and hotels, and about 30% of the effect for peer hosts. There are many aspects of Airbnb's growth that such a counterfactual does not capture. In particular, consumer utility for Airbnb listings may have changed over time due to changes in the composition of rooms available and changes in the Airbnb platform.

In this paper we have documented two fundamental reasons why peer production is valuable in the accommodation industry, which can be generalized to cities that have experienced a sizable growth in Airbnb listings. First, peers offer a differentiated product that is not a perfect substitute to hotel rooms and is valued by at least some consumers. Second, the hotel sector in many cities is frequently constrained by a limited number of available rooms, which leads to high prices during demand peaks because hotels cannot accommodate all potential travelers. Peer production expands available supply at exactly these times of peak demand, thus reducing hotel pricing power and increasing consumer surplus. To the extent that supply of rooms on Airbnb has become more professionalized and fixed over time, our distinction of flexible versus dedicated capacity can be made not just between hotels and peer hosts, but within Airbnb across occasional and professional hosts.

We have focused on the short-run effects of a peer-to-peer platform on the agents directly involved – hotels, peer hosts, and travelers – without highlighting the platform's own costs and revenues. In the longer run, the number of hotel rooms is likely to adjust in response to peer entry. Peer production can also have externalities and spillovers into other markets, including the labor and housing markets (Horton (2019), Barron et al. (2018)). We leave the study of these important effects for future work.

References

Aguiar, Luis, and Joel Waldfogel. 2018. "As streaming reaches flood stage, does it stimulate or depress music sales?" International Journal of Industrial Organization, 57: 278 – 307.

Airbnb. 2020. "Airbnb S1 Report."

Almagro, Milena, and Tomas Dominguez-Iino. 2020. "Location Sorting and Endogenous Amenities: Evidence from Amsterdam." Working Paper.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M Shapiro. 2017. "Measuring the sensitivity of parameter estimates to estimation moments." The Quarterly Journal of Economics, 132(4): 1553–1592.

Barron, Kyle, Edward Kung, and Davide Proserpio. 2018. "The Sharing Economy and Housing Affordability: Evidence from Airbnb." EC '18, 5. New York, NY, USA:Association for Computing Machinery.

Bass, Frank. 1969. "A New Product Growth for Model Consumer Durables." Management Science, 15(5): 215–227.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. "Automobile Prices in Market Equilibrium." Econometrica, 63(4): 841–890.

Bolton, Gary, Ben Greiner, and Axel Ockenfels. 2012. "Engineering Trust: Reciprocity in the Production of Reputation Information." Management Science, 59(2): 265–285.

Bulow, Jeremy I., John D. Geanakoplos, and Paul D. Klemperer. 1985. "Multimarket Oligopoly: Strategic Substitutes and Complements." Journal of Political Economy, 93(3): 488–511.

Calder-Wang, Sophie. 2020. "The Distributional Impact of the Sharing Economy on the Housing Market." Working Paper.

Castillo, Juan Camilo. 2020. "Who Benefits from Surge Pricing?" Working Paper.

Chen, M. Keith. 2016. "Dynamic Pricing in a Labor Market: Surge Pricing and Flexible Work on the Uber Platform." EC '16, 455. New York, NY, USA:Association for Computing Machinery.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2019. "Using Big Data to Estimate Consumer Surplus: The Case of Uber." , (22627).

Conlon, Christopher, and Jeff Gortmaker. 2020. "Best practices for differentiated products demand estimation with pyblp." The RAND Journal of Economics, 51(4): 1108–1161.

Cullen, Zoë, and Chiara Farronato. 2019. "Outsourcing Tasks Online: Matching Supply and Demand on Peer-to-Peer Internet Platforms."

Einav, Liran, Chiara Farronato, and Jonathan Levin. 2016. "Peer-to-peer markets." Annual Review of Economics, 8: 615–635.

Einav, Liran, Chiara Farronato, Jonathan Levin, and Neel Sundaresan. 2018. "Auctions versus Posted Prices in Online Markets." Journal of Political Economy, 126(1): 178–215.

Expedia. 2019. "Expedia 10-k Report 2019."

Filippas, Apostolos, John J. Horton, and Richard J. Zeckhauser. Forthcoming. "Owning, Using, and Renting: Some Simple Economics of the "Sharing Economy"." Management Science.

Fradkin, Andrey. 2019. "Search, Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb." SSRN Electronic Journal.

Fradkin, Andrey, Elena Grewal, and David Holtz. 2019. "The Determinants of Online Review Informativeness: Evidence from Field Experiments on Airbnb." SSRN Electronic Journal.

Fraiberger, Samuel, and Arun Sundararajan. 2019. "Peer-to-Peer Rental Markets in the Sharing Economy." SSRN Electronic Journal.

Friedman, James W. 1971. "A non-cooperative equilibrium for supergames." The Review of Economic Studies, 38(1): 1–12.

Friedman, James W. 1977. Oligopoly and the Theory of Games. Vol. 8, North-Holland.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." Econometrica, 25(4): 501–522.

Gyourko, J., A. Saiz, and A. Summers. 2008. "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." Urban Studies, 45(3): 693–729.

Hall, Jonathan, Cory Kendrick, and Chris Nosko. 2019. "The Effects of Uber's Surge Pricing: A Case Study."

Horton, John J. 2014. "Misdirected Search Effort in a Matching Market: Causes, Consequences and a Partial Solution." EC '14, 357. New York, NY, USA:Association for Computing Machinery.

Horton, John J. 2019. "The Tragedy of Your Upstairs Neighbors: Is the Airbnb Negative Externality Internalized?"

Kalnins, Arturs. 2006. "Markets: The U.S. Lodging Industry." The Journal of Economic Perspectives, 20(4).

Koulayev, Sergei. 2014. "Search for Differentiated Products: Identification and Estimation." The RAND Journal of Economics, 45(3): 553–575.

Kroft, Kory, and Devin G. Pope. 2014. "Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist." Journal of Labor Economics, 32(2): 259–303.

Lam, Tom, and Meng Liu. 2019. "Demand and Consumer Surplus in the On-Demand Economy: The Case of Ride Sharing."

Lewis, Gregory, and Georgios Zervas. 2019. "The Welfare Impact of Consumer Reviews: A Case Study of the Hotel Industry."

Nosko, Chris, and Steven Tadelis. 2019. "The Limits of Reputation in Platform Markets: An Empirical Analysis and Field Experiment." , (20830).

Nowak, Brian, Thomas Allen, Jamie Rollo, Vaughan Lewis, Lin He, Ananda Chen, Wilson W. Ng, Michael Costantini, Owen Hyde, Kevin Liu, Mark Savino, Baset A. Chadhry, Anne M. Grube, and Ed Young. 2015. "Global Insight: Who Will Airbnb Hurt More - Hotels or OTAs?"

Petrin, Amil. 2002. "Quantifying the benefits of new products: The case of the minivan." Journal of political Economy, 110(4): 705–729.

Ryan, Stephen P. 2012. "The Costs of Environmental Regulation in a Concentrated Industry." Econometrica, 80(3): 1019–1061.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." The Quarterly Journal of Economics, 125(3): 1253–1296.

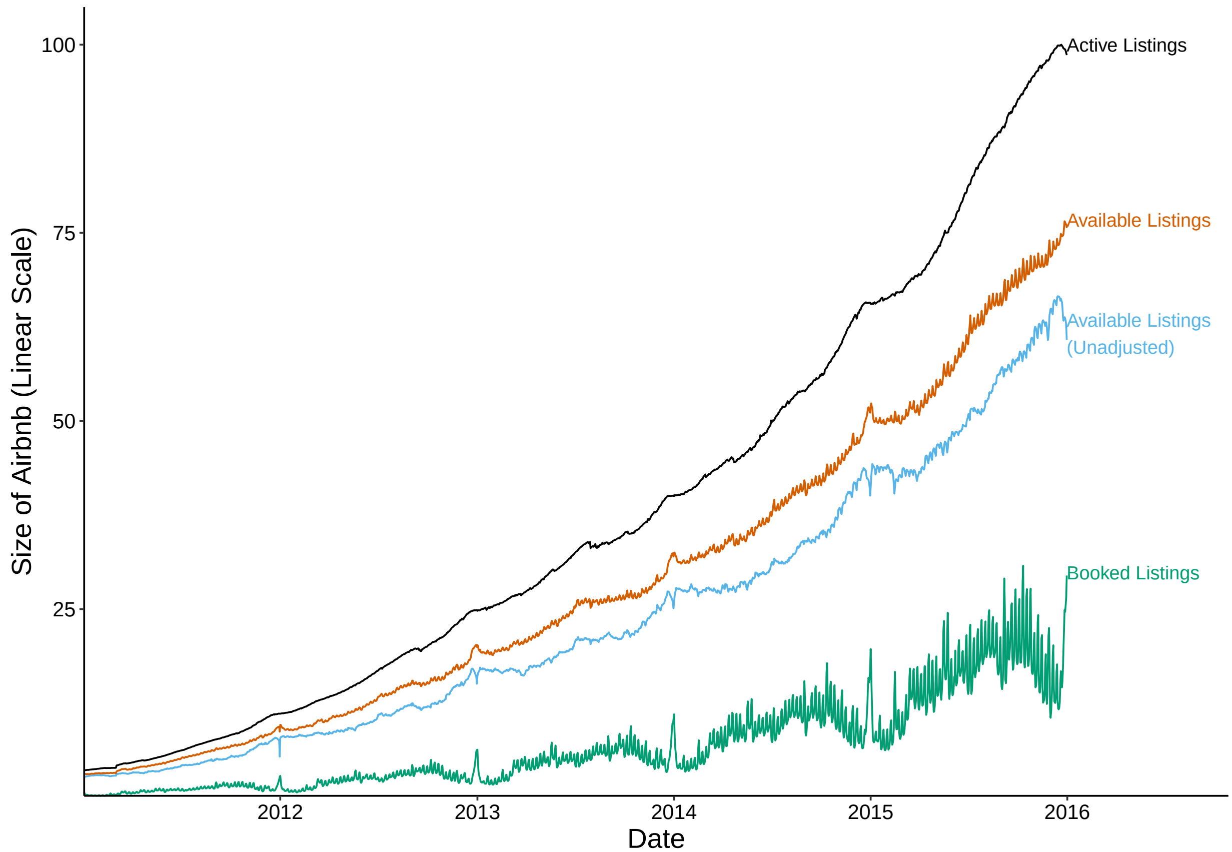
Seamans, Robert, and Feng Zhu. 2014. "Responses to Entry in Multi-Sided Markets: The Impact of Craigslist on Local Newspapers." Management Science, 60(2): 476–493.

Shapiro, Carl. n.d.. "Chapter 6 Theories of Oligopoly Behavior." In Handbook of Industrial Organization. Vol. 1, 329–414. Elsevier.

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

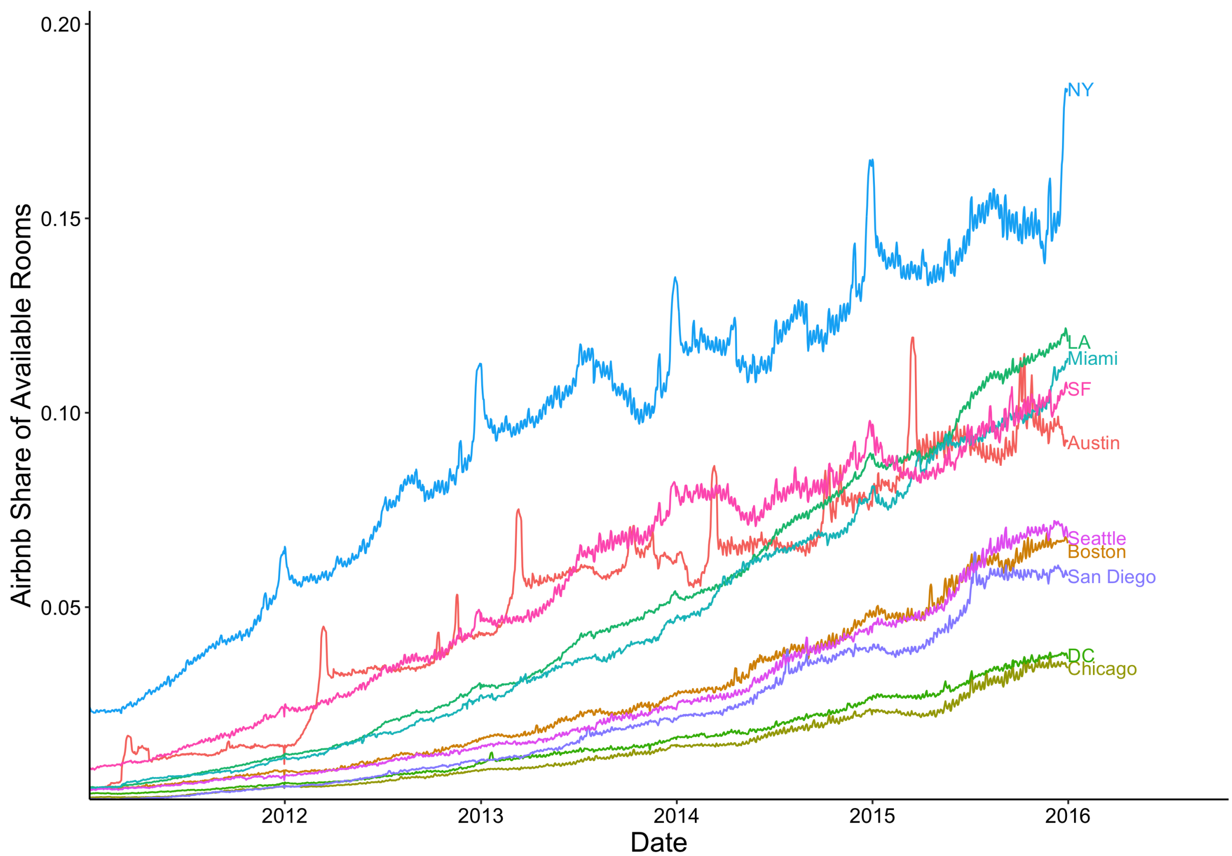
Figures

Figure 1: Measures of Airbnb Supply



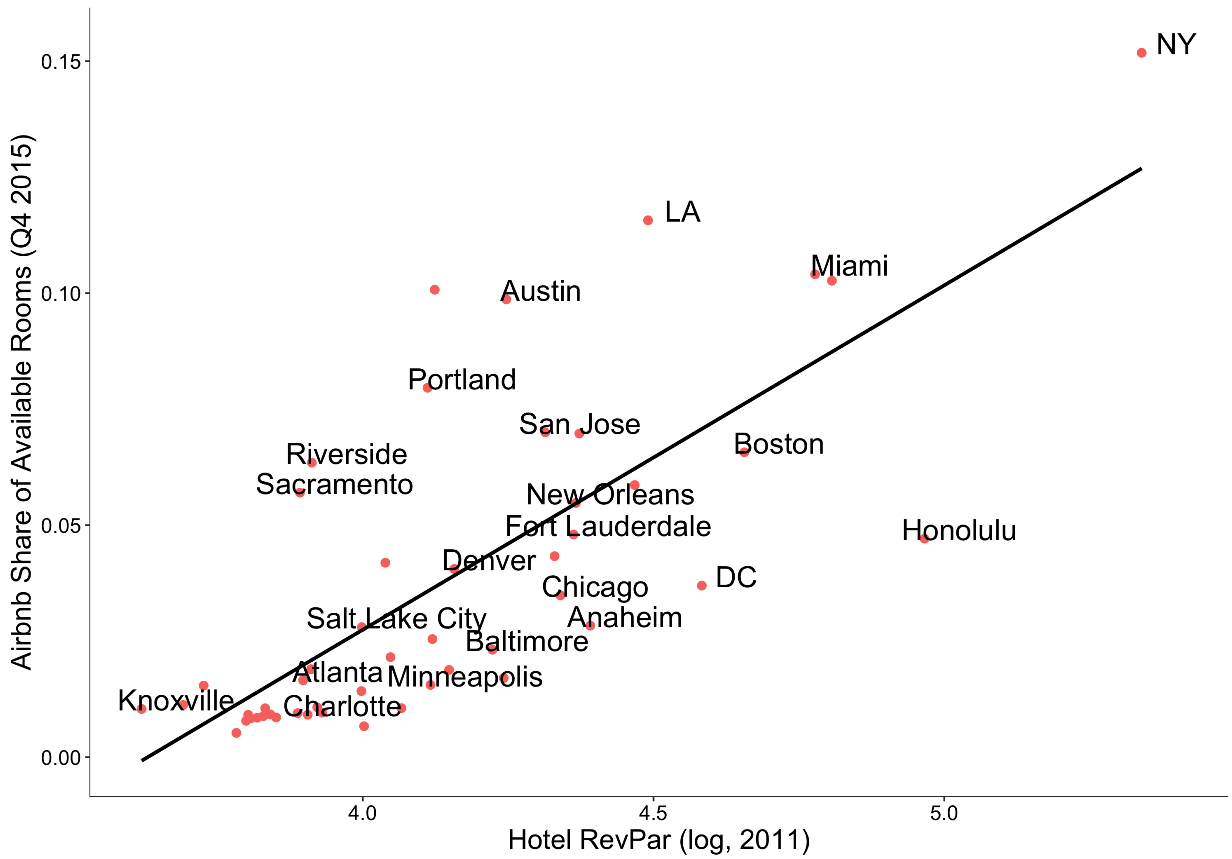
This figure plots four measures of the size of Airbnb. An active listing is defined as a listing available to be booked or booked for any future date. An (unadjusted) available listing is one that is either booked or has an open calendar slot on the date of stay. Available listings augment the unadjusted measure with listings that were contacted for a particular date of stay and were later updated to be unavailable for that date. A booked listing is one that has been booked for that date. The y-axis is normalized by the maximum number of active listings during our sample period to protect the company's proprietary data.

Figure 2: Growth of Airbnb



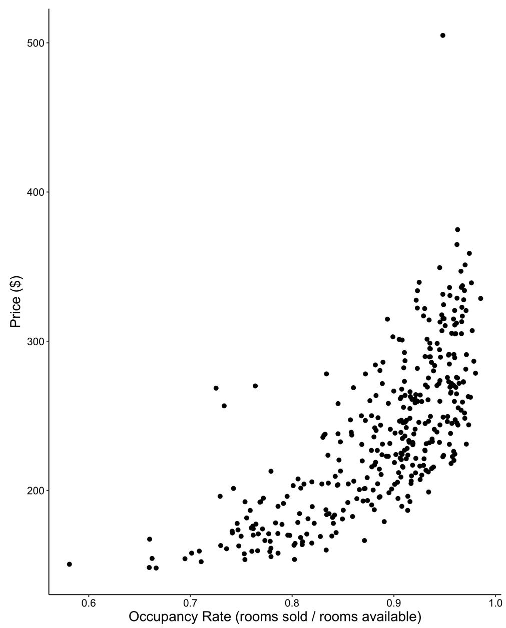
The figure plots the size of Airbnb over time in 10 selected cities. The y-axis is the share of Airbnb listings out of all Airbnb listings and hotel room capacity on a given day. The 10 selected cities are those with the largest number of listings on Airbnb at the end of our sample period among the 50 US major cities. Appendix Figure A6 shows that hotel room capacity has been fairly stable over the same time period.

Figure 3: Airbnb Penetration and Hotel Revenues per Available Room



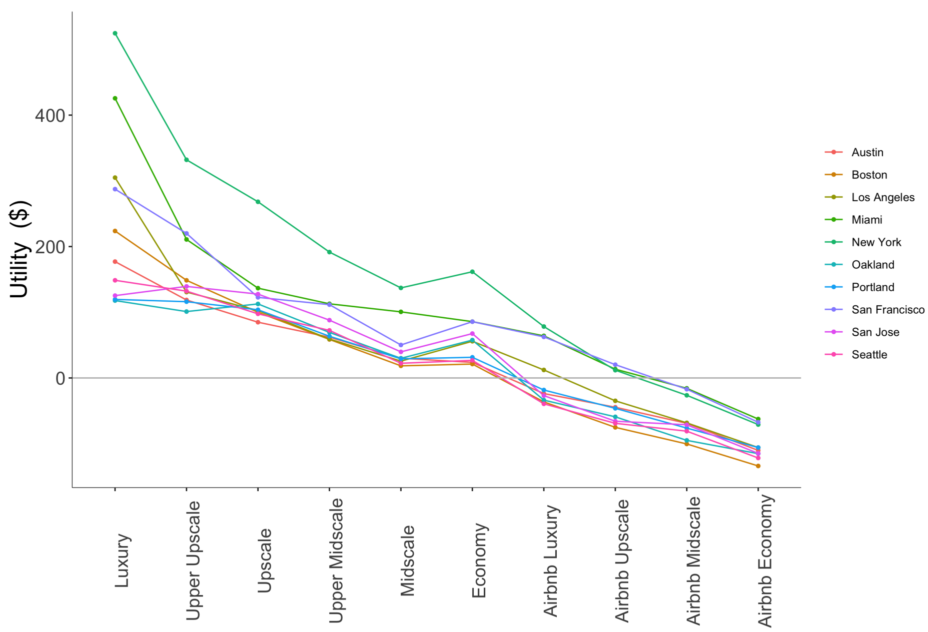
This figure plots the size of Airbnb against hotels' average revenue per available room for each of the 50 cities in our sample. The size of Airbnb is measured as the average daily share of Airbnb listings out of all (hotel and Airbnb) rooms available for short-term accommodation. The average is computed over the last quarter of 2015. The hotels' revenue per available room is the daily ratio of total hotel revenues divided by the number of available hotel rooms, averaged over the course of 2011. The fitted line weighs each city equally.

Figure 4: Prices and Occupancy Rates



This figure plots prices and occupancy rates of upscale hotels in New York in 2014.

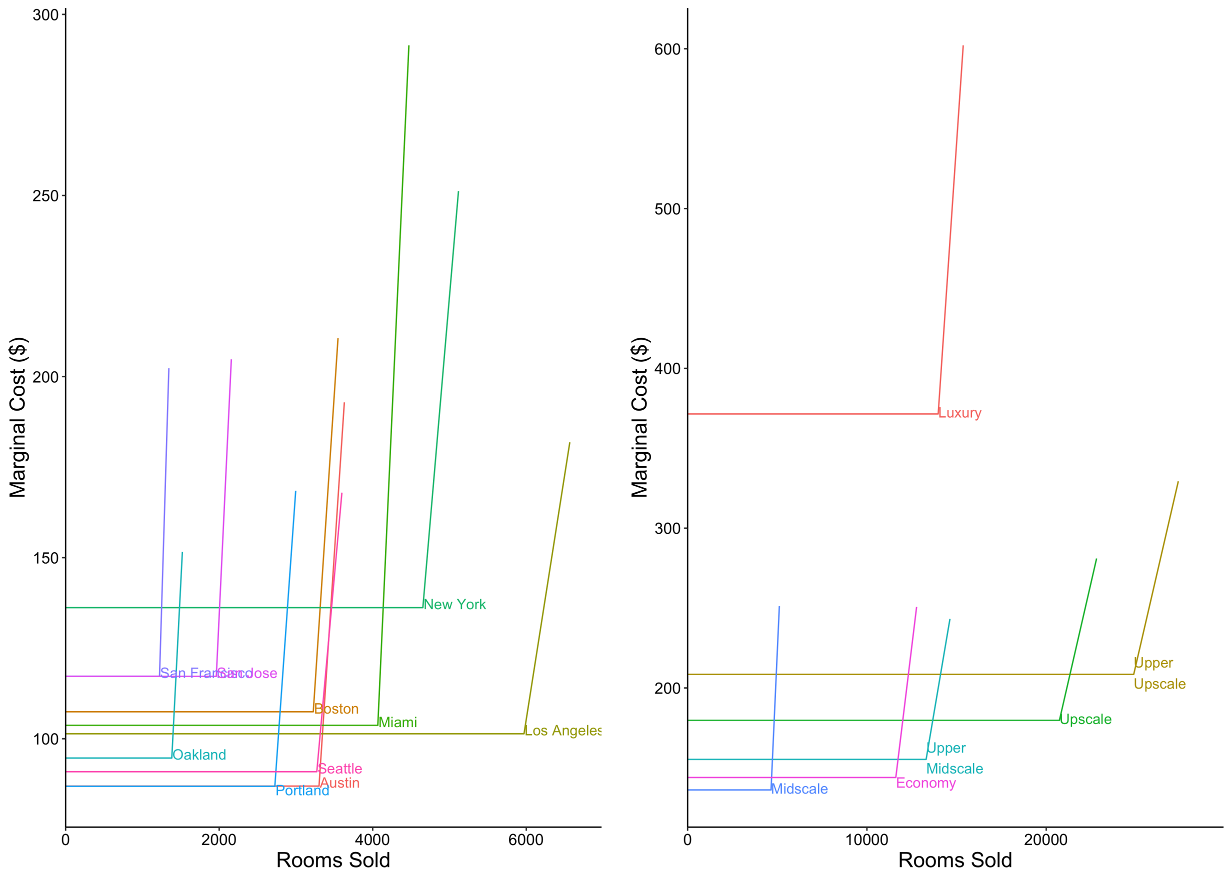
Figure 5: Estimated Utilities for Accommodation Options Across Cities



This figure plots the estimated mean utilities for accommodation options across the 10 cities used in our estimation. The values are computed as averages over December 2014.

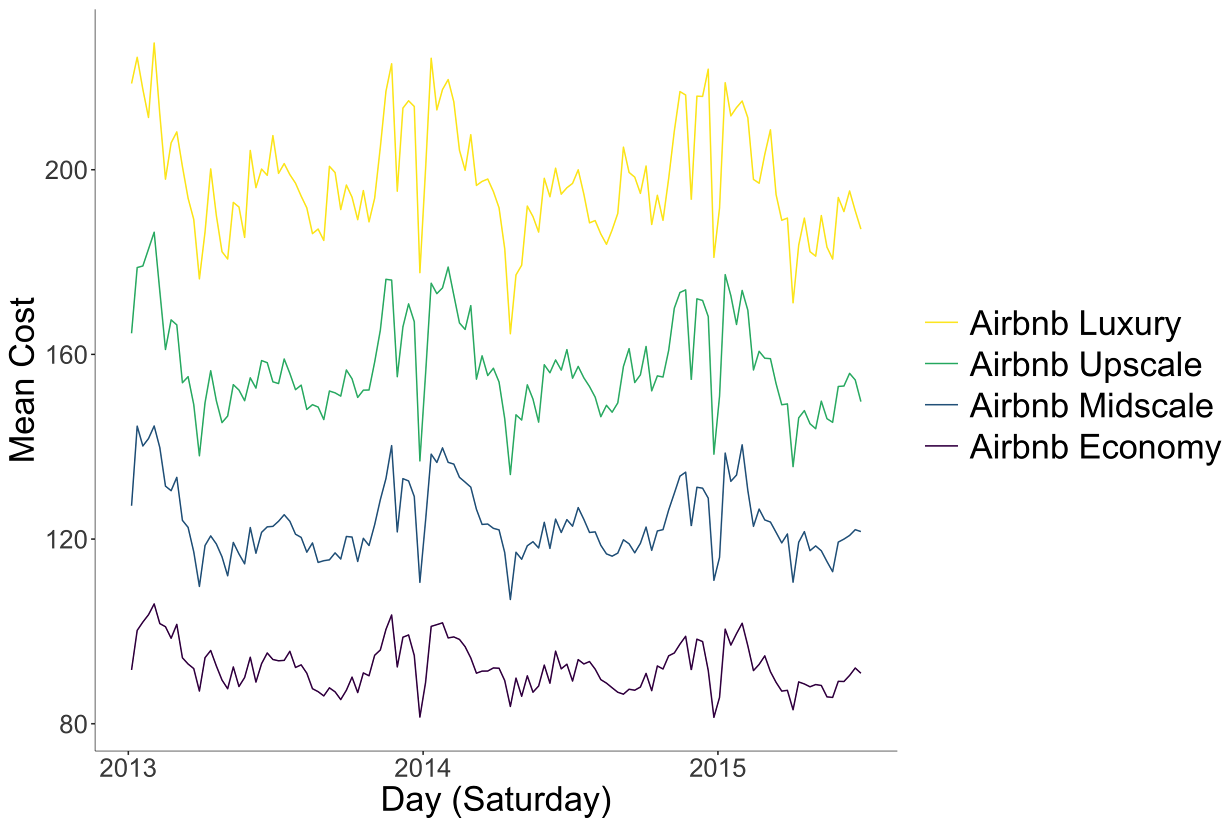
Figure 6: Estimated Hotel Costs

(a) Costs by City – Midscale Hotels (b) Costs by Hotel Scale – New York City



These figures plot the estimated marginal cost curves of hotels across cities (left panel) and across scales (right panel). The values are computed as averages over December 2014. Appendix Tables A14 and A15 display the cost estimates by city and hotel scale.

Figure 7: Mean Costs of Airbnb Hosts in New York City



The figures plot the estimated mean costs of Airbnb hosts in New York over time. Appendix Table A17 displays all the estimated means and standard deviations.

Tables

Table 1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Mean Hotel Occupancy	50	0.67	0.07	0.62	0.66	0.72
Std Dev Hotel Occupancy	50	0.14	0.03	0.12	0.14	0.15
Mean Hotel Price in $	50	110.77	35.66	88.45	100.90	124.29
Std Dev Hotel Price	50	17.97	9.84	10.42	14.98	23.22
Mean Hotel Revenue (Thousand $)	50	3,945.50	3,696.93	1,638.89	2,630.30	5,012.71
Mean Airbnb Occupancy	50	0.16	0.05	0.13	0.14	0.20
Std Dev Airbnb Occupancy	50	0.10	0.02	0.08	0.10	0.11
Mean Airbnb Price in $	50	113.93	25.80	97.76	103.77	125.78
Std Dev Airbnb Price	50	32.09	12.43	23.18	29.41	38.66
Mean Airbnb to Hotel Price Ratio	50	1.08	0.28	0.94	1.03	1.16
Std Dev Price Ratio	50	0.32	0.16	0.21	0.29	0.37
Airbnb Share of Available Rooms (Q4 2015)	50	0.04	0.03	0.01	0.02	0.06
Airbnb Share of Potential Guests (Q4 2015)	50	0.05	0.04	0.01	0.03	0.08
Airbnb Share of Housing Units (Q4 2015)	50	0.001	0.001	0.0004	0.001	0.002

This table shows hotel and Airbnb descriptive statistics for the 50 cities in our sample. For each city, we compute the mean and standard deviation of daily metrics for hotels and Airbnb listings between January 2011 and December 2015. The metrics we consider are occupancy rates, prices per room-night, revenues, ratio of Airbnb to hotel prices. The last three rows show Airbnb size as a share of available rooms, potential guests, and housing units in the last quarter of our sample, October - December 2015. The Airbnb share of available rooms is computed as the average daily share of available rooms (Airbnb listings divided by the sum of Airbnb listings and hotel rooms) between October and December 2015. The Airbnb share of potential guests is computed as the average share of available rooms adjusted for their realized capacity, i.e. number of guests occupying a room. To make this adjustment, we have data on Airbnb realized number of guests per room at the city-day-listing type level. Since we do not have the same metric for hotels, we assume that the typical hotel has the same number of average guests as a Midscale Airbnb listing. The Airbnb share of potential guests is typically higher than the Airbnb share of available rooms because an Airbnb listing is on average occupied by more guests than hotel rooms. Finally, the Airbnb share of housing units is the average of the ratio of available Airbnb listings divided by the number of housing units in the Metropolitan Statistical Area.

Table 2: City Characteristics and Size of Airbnb

	Airbnb Share of Rooms (Q4 2015)	
	(1)	(2)
Undevelopable Area	0.036**	0.026*
	(0.017)	(0.015)
Wharton Residential Land Use Index (WRLURI)	0.009*	0.005
	(0.004)	(0.004)
SD. Incoming Air Passengers (2011)	0.002*	0.00003
	(0.001)	(0.001)
% Growth in Air Passengers (2012-2011)	0.125*	0.107*
	(0.067)	(0.058)
% Never Married	0.504***	0.308**
	(0.164)	(0.152)
% Children	−0.399*	−0.218
	(0.203)	(0.182)
Log(Rev. Per Room (2011))		0.061***
		(0.017)
Log(Market Size)	−0.008	−0.012
	(0.008)	(0.007)
Constant	0.052	−0.133
	(0.091)	(0.093)
Observations	46	46
R^2	0.668	0.758

This table shows linear regressions of the size of Airbnb on market characteristics linked to supply constraints, demand volatility, and the costs of hosting (Equation 1). The size of Airbnb is the average of daily share of rooms in the last quarter of 2015. The standard deviation of incoming passengers is divided by 10,000 to make the coefficient comparable to the other variables. Descriptive statistics are shown in Table 1 (for the outcome variable) and Appendix Table A4 (for the predictors). Market size is measured as the average number of rooms available in the last quarter of 2015. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 3: The Supply Elasticity of Hotels and Peer Hosts

	Log(Hotel Rooms Boooked + 1)	Log(Airbnb Rooms Booked + 1)
	(1)	(2)
log(Hotel Rooms + 1)	0.543**	
	(0.215)	
log(Hotel Price)	1.289***	
	(0.103)	
log(Airbnb Available Listings + 1)		0.385***
		(0.142)
log(Airbnb Price)		3.893***
		(0.288)
IV	Yes	Yes
City FE	Yes	Yes
Year-Month FE	Yes	Yes
Day of Week FE	Yes	Yes
Observations	90,900	84,959
R^2	0.954	0.774

The table shows results of IV regressions of the log of hotel and Airbnb bookings on the corresponding price and room availability (Equation 2). In column 1 we instrument for hotel prices with the one week lag of the log of the Google Search Trends and the log of arriving (not returning) flight travelers. In column 2 we instrument for both Airbnb prices and the number of available listings. We use three instruments: the one week lag of the log of the Google Search Trends and the log of arriving flight travelers as in column 1, plus the number of Airbnb active listings. First stage results are reported in Appendix Table A5 and OLS results are reported in Appendix Table A6. Standard errors are clustered at the city level. Adding the city-day observations with no Airbnb bookings (and using hotel prices in column 2) does not change the results. Results would not change if in Column 2 we included the log of departing air travelers and the one-week lag of the log of local Google Search Trends for hotels outside of the city as additional controls. *$p<0.1$; **$p<0.05$; ***$p<0.01$.

Table 4: Hotel Revenues and Airbnb Entry

	Log(RevPAR) (1)	Occupancy Rate (2)	Log(Price) (3)
log(Incoming Air Passengers)	1.104***	0.371***	0.482***
	(0.063)	(0.041)	(0.040)
log(Google Search Trend)	0.246***	0.076***	0.109***
	(0.042)	(0.012)	(0.024)
log(Hotel Rooms + 1)	−0.936***	−0.521***	−0.089
	(0.326)	(0.137)	(0.168)
log(Hotel Rooms + 1)*Inelastic Housing Supply	−0.475	0.055	−0.612**
	(0.370)	(0.174)	(0.281)
log(Airbnb Available Listings + 1)	0.020	−0.002	0.021**
	(0.016)	(0.007)	(0.010)
log(Airbnb Available Listings + 1)*Inelastic Housing Supply	−0.057**	−0.002	−0.054**
	(0.025)	(0.010)	(0.022)
IV	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Observations	90,900	90,900	90,900
R^2	0.740	0.591	0.856

This table shows results of IV estimates of Equation 3, where the size of Airbnb is measured as the number of available listings. The Google search trend is a one-week lag. The instruments for available listings and its interaction with the dummy for inelastic housing supply are the number of active listings and its corresponding interaction with the dummy. The dependent variable is log revenue per available room in column (1), occupancy rate in column (2), and log price in column (3). First stage results are reported in Appendix Table A7 and OLS results are reported in Appendix Table A8. Results for different hotel scales are presented in Appendix Table A9, and results using different measures of the size of Airbnb are in Appendix Table A10. Standard errors are clustered at the city level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 5: Estimates of Selected Demand Parameters

	Random Coefficients Logit		Standard Logit	
Parameter	Estimate	Std. Error	Estimate	Std. Error
Log Google Trend	2.355	0.281	1.783	0.059
Price	-0.031	0.002	-0.025	0.001
Std. Deviation on Inside Option	1.725	1.060	.	.
Std. Deviation on Price	0.004	0.004	.	.

This table displays the estimates and standard errors for selected parameters in travelers' utility. High scale includes luxury and upper upscale hotels.

Table 6: Aggregate Surplus (MM)

	Consumers	Hotels			Peer Hosts			Government
	Change in Consumer Surplus	Rooms Sold	Revenues	Profits	Rooms Sold	Revenues	Peer Surplus	Lodging Taxes
Panel A: All markets in 2014								
Baseline		146	26,803	5,687	4.38	517	112	3,986
No Airbnb (Unconstrained)	-147	149	27,412	5,833				4,071
No Airbnb	-305	148	27,238	5,852				4,045
Airbnb With Lodging Tax	-65	146	26,891	5,718	3.47	377	82	4,058
Airbnb With Quotas	-157	147	27,106	5,754	1.49	181	40	4,027
Double Airbnb Rooms	130	145	26,630	5,623	6.18	672	146	3,962
Panel B: Compression Nights in 2014 (19.6% of all markets)								
Baseline		33	7,240	3,414	1.16	145	31	1,084
No Airbnb (Unconstrained)	-39	33	7,435	3,505				1,111
No Airbnb	-121	33	7,341	3,495				1,097
Airbnb With Lodging Tax	-24	33	7,258	3,429	0.94	107	23	1,103
Airbnb With Quotas	-14	33	7,256	3,424	1.01	125	28	1,086
Double Airbnb Rooms	53	33	7,199	3,380	1.65	190	41	1,078

This table displays how consumers, hotels, peer hosts, and lodging taxes change under the baseline scenario and five alternative scenarios, two scenarios without Airbnb and three scenarios with Airbnb and regulation. 'Unconstrained' refers to the counterfactual scenario in which Airbnb options do not exist, hotels do not adjust prices and can accommodate any additional bookings regardless of their actual capacity. In the 'No Airbnb' counterfactual, we let hotel prices readjust in response to the absence of Airbnb and accounting for hotel capacity constraints. The next counterfactuals consider new equilibrium prices and quantities under different regulation. The 'Airbnb with Lodging Tax' counterfactual keeps Airbnb availability at the baseline level, but Airbnb travelers are charged the same lodging tax rate as hotel travelers. The 'Airbnb with Quotas' counterfactual allows for Airbnb rooms to exist during the 90 days in a year with the largest number of travelers choosing to book accommodations in a particular city in the baseline scenario. For the other 275 days, Airbnb rooms are not allowed and so the equilibrium prices and quantities mirror those in the 'No Airbnb' counterfactual. Note that those 90 days when Airbnb is allowed are not the same across all cities, but rather are determined independently for each city. The 'Double Airbnb Rooms' counterfactual holds lodging taxes while doubling the number of active Airbnb listings. Panel A displays metrics aggregated across all cities and nights in 2014 while Panel B focuses on the markets when at least one hotel option has an occupancy rate of 95% or more (19.6% of markets in 2014 are considered compression nights). All variables are in millions. For heterogeneity of the effects across cities, see Appendix Tables A19 through A21. For counterfactuals without consumer heterogeneity, see Table A23.

APPENDIX

The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb

Chiara Farronato and Andrey Fradkin

For Online Publication

A Appendix: Theoretical Framework and Proofs

We present a theoretical model for understanding market structure with dedicated supply (hotels) and flexible supply (peer hosts) in the accommodation industry. It is a version of the model presented in Section 3 with more general demand and cost specifications, but with only one type of hotels and one type of Airbnb hosts. We prove existence and uniqueness of the equilibrium under certain conditions, as well as some comparative statics predictions that are corroborated by the stylized facts from Section 2.

In our model, hosting services can be provided by dedicated and flexible sellers, who offer differentiated products. The model has a short and long-run component. The short-run equilibrium consists of daily prices and rooms sold of each accommodation type as a function of the overall demand level and the respective capacities of dedicated and flexible suppliers. We assume hotels are competing against a fringe of flexible sellers. The long-run component determines the entry condition of flexible sellers as a function of fixed hotel capacity and the distribution of demand states.

The Short-Run

We start with the short-run equilibrium representing daily market outcomes. We simplify the exposition by assuming that there is one single hotel and one undifferentiated type of Airbnb listings. Let h denote the hotel and a denote Airbnb. Further, let K_h denote the mass of existing dedicated capacity (number of available hotel rooms), and K_a the existing flexible capacity (available Airbnb rooms). Demand state, s, is drawn from a distribution $F(.)$, which can be interpreted as the distribution of demand states over the course of a year. Hotel rooms and Airbnb rooms are differentiated products. We denote $p_i^s(Q_i, Q_j)$ the inverse demand function for product i, which depends on the quantity for product i, product j, and demand state s. We assume that products i and j are substitutes, so $p_i^s(Q_i, Q_j)$ is decreasing in both Q_i and Q_j, and the prices of both products are increasing in the demand state.

The short-run sequence of events is as follows. Capacities K_h and K_a are given, demand state s is realized, the hotel sets quantities and at the same time Airbnb sellers choose whether to host at the prevailing prices. We assume that the hotel faces marginal cost c_h to book one room for one night, and it sets its quantity to maximize profits subject to its capacity constraint:

$$\begin{aligned} &\underset{Q_h}{Max} \quad Q_h\left(p_h^s(Q_h, Q_a) - c_h\right) \\ &s.t. \qquad\quad Q_h \le K_h \end{aligned} \tag{A1}$$

Flexible sellers have unit capacity and variable marginal costs of renting their room. We assume that marginal costs of peers are randomly drawn from a known distribution $G(.)$.

When choosing whether to rent out their room for a night, flexible producers take prices as given, and sell their unit if and only if the market clearing price is greater than their cost. The choices of individual hosts are aggregated to determine the total number of flexible rooms rented:

$$Q_a = K_a G\left(p_a^s(Q_h, Q_a)\right), \tag{A2}$$

where K_a is the mass of peer hosts, and $G\left(p_a^s(Q_h, Q_a)\right)$ is the share of hosts with costs lower than $p_a^s(Q_h, Q_a)$.

The market equilibrium consists of prices and quantities for hotel rooms and peer rooms that equate flexible and dedicated room demand with flexible and dedicated supply. For the proofs about the short-run equilibrium, we remove the superscript s, and denote hotel profits $\Pi = Q_h\left(p_h(Q_h, Q_a) - c_h\right)$.

The result of existence and uniqueness of the equilibrium is based on Friedman (1971) and Friedman (1977), and is equivalent to the stability requirement in Bulow et al. (1985). Also see Shapiro (n.d.) for an overview of equilibrium in Cournot models.

Proposition 1 *There is a unique equilibrium if the hotel's profit function is twice continuously differentiable, the hotel's marginal revenue curve does not rise with its own or its competitors' output, and* $\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a)\frac{\partial p_a}{\partial Q_a} - 1\right] \ge \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a} K_a g(p_a)\frac{\partial p_a}{\partial Q_h}$.

Proof. Equilibrium is given by the intersection of the hotel and peer hosts' reaction curves (Fig. A1), determined by Equations A1 and A2. Under the condition in the proposition, both reaction curves are downward sloping, and we can prove that the hotel's reaction curve always has a slope smaller than the slope of the reaction curve of peer hosts. This ensures that the curves intersect at most once, either along one of the axes – where one type of supply sells zero rooms – or at an interior point where both suppliers sell rooms.

We first consider the case in which hotel capacity is not binding. The reaction curves are given by the following system of equilibrium equations:

$$\frac{\partial \Pi(Q_h, Q_a)}{\partial Q_h} = 0$$
$$K_a G\left(p_a(Q_h, Q_a)\right) - Q_a = 0.$$

Totally differentiating the system of equations leads to

$$\frac{\partial^2 \Pi}{\partial Q_h^2} dQ_h + \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a} dQ_a = 0 \tag{A3}$$
$$K_a g(p_a) \frac{\partial p_a}{\partial Q_h} dQ_h + \left(K_a g(p_a) \frac{\partial p_a}{\partial Q_a} - 1\right) dQ_a = 0. \tag{A4}$$

Equation A3 implies that the hotel's reaction curve has slope equal to $\frac{dQ_a}{dQ_h} = -\frac{\partial^2 \Pi}{\partial Q_h^2} \Big/ \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}$, while equation A4 implies that peer hosts' reaction curve has slope equal to $\frac{dQ_a}{dQ_h} = -K_a g(p_a) \frac{\partial p_a}{\partial Q_h} \Big/ \Big(K_a g(p_a) \frac{\partial p_a}{\partial Q_a} - 1$

The numerator and the denominator in both slopes are negative. For the slope of the hotel's reaction curve, this is because we have constant marginal costs and the hotel's marginal revenue curve does not increase with its own or its competitors' output. For the slope of the peer hosts' reaction curve, it is because for normal goods we have $\frac{\partial p_a}{\partial Q_a} \leq 0$, and substitutatibility implies $\frac{\partial p_a}{\partial Q_h} \leq 0$.

The slope of the hotel's reaction curve is smaller than the slope of its competitors' reaction curve whenever $-\frac{\partial^2 \Pi}{\partial Q_h^2} \Big/ \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a} \leq -K_a g(p_a) \frac{\partial p_a}{\partial Q_h} \Big/ \Big[K_a g(p_a) \frac{\partial p_a}{\partial Q_a} - 1\Big]$. Reordering, the condition is equivalent to the condition in the proposition.

We now consider the case in which hotel capacity is binding, so $K_h - Q_h = 0$. In this case the hotel's reaction curve is vertical and crosses the x-axis (0 demand for Airbnb) at K_h. Regardless of whether the hotel's reaction curve hits its maximum room capacity before crossing the x-axis, the reaction curves of the hotel and of peer hosts will cross at most once at an interior point.

The unique equilibrium can be characterized as one of four options, as Fig. A1 shows. If the hotel's reaction curve is always below the reaction curve of peer hosts, the equilibrium will be along the y-axis, where the hotel sells no rooms. If the hotel's reaction curve is always above peer hosts' reaction curve, the equilibrium will be along the x-axis, where peers sell no rooms. Otherwise the equilibrium will be at the crossing point of the two reaction curves, where both suppliers sell some rooms, and the hotel can be either capacity-constrained or unconstrained. ■

The short-run model offers some comparative statics predictions. Under standard conditions, hotel profits per available room, as well as both prices and occupancy rates, are lower if K_a is higher. The separate effect of an increase in K_a on hotel prices is higher if hotel capacity constraints are more often binding, but the opposite is true for the effect on occupancy. Intuitively, this occurs because the increase in flexible capacity affects hotels through a reduction in their residual demand, and when hotels are capacity constrained, their supply curve is vertical. A marginal downward shift in residual demand will have no effect on quantity and a large effect on price if supply is perfectly inelastic (Figure A2). We present the propositions and the proofs below.

Proposition 2 *Hotel profits and quantities weakly decrease in K_a. Hotel prices decrease in K_a if and only if $\frac{\partial^2 \Pi}{\partial Q_h^2}\frac{\partial p_h}{\partial Q_a} - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}\frac{\partial p_h}{\partial Q_h} \geq 0$.*

Proof. In order to prove Proposition 2 it is useful to separately consider markets where the hotel capacity constraint binds and markets where it does not. In markets where the hotel constraint binds the two equilibrium conditions are $Q_h = K_h$ and $Q_a = K_a G\left(p_a(Q_h, Q_a)\right)$. By totally differentiating the system of equilibrium equations we find the total derivatives of the hotel's and peer hosts' quantities with respect to peer hosts' capacity:

$$\left[\frac{dQ_h}{dK_a}\right]^c = 0 \tag{A5}$$

$$\left[\frac{dQ_a}{dK_a}\right]^c = \frac{-G(p_a)}{K_a g(p_a)\frac{\partial p_a}{\partial Q_a} - 1}. \tag{A6}$$

In markets where the hotel constraint does not bind the two equilibrium conditions are $\partial\Pi(Q_h,Q_a)/\partial p_h = 0$ and $Q_a = K_a G\left(p_a(Q_h, Q_a)\right)$. By totally differentiating the system of equilibrium equations we find the total derivatives of the hotel's and peer hosts' quantities with respect to peer hosts' capacity:

$$\left[\frac{dQ_h}{dK_a}\right]^u = \frac{G(p_a)\frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}}{\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a)\frac{\partial p_a}{\partial Q_a} - 1\right] - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a} K_a g(p_a)\frac{\partial p_a}{\partial Q_h}} \tag{A7}$$

$$\left[\frac{dQ_a}{dK_a}\right]^u = \frac{-G(p_a)\frac{\partial^2 \Pi}{\partial Q_h^2}}{\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a)\frac{\partial p_a}{\partial Q_a} - 1\right] - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a} K_a g(p_a)\frac{\partial p_a}{\partial Q_h}}. \tag{A8}$$

We start by proving that hotel quantities are a decreasing function of flexible capacity in both constrained and unconstrained equilibria. Since in the constrained equilibrium the hotel quantity is fixed at its maximum capacity, its derivative with respect to flexible capacity is zero (equation A5). We simply need to prove that the derivative in equation A7 is negative. Again, this is directly implied by the conditions for existence and uniqueness of the equilibrium from Proposition 1. Indeed, the numerator is negative because hotel's marginal revenues are decreasing in peer hosts' quantity. The denominator is positive because of the condition in Proposition 1.

So far, we have proved that an increase in flexible capacity decreases hotel quantities by showing that $\frac{dQ_h}{dK_a} \leq 0$ whether or not the hotel is operating at capacity. Now we prove that an increase in flexible capacity also decreases hotel profits. An increase in K_a affects hotel profits $\Pi = Q_h(p_h - c_h)$ through changes in Q_h and Q_a:

$$\frac{d\Pi}{dK_a} = \frac{\partial \Pi}{\partial Q_h}\frac{dQ_h}{dK_a} + \frac{\partial \Pi}{\partial Q_a}\frac{dQ_a}{dK_a}.$$

Regardless of whether the hotel capacity constraint is binding, the first term in the summation is zero. If the hotel's capacity constraint is binding, it is because $\frac{dQ_h}{dK_a} = 0$ from equation A5. If the hotel's capacity constraint is not binding, it is because the hotel's first order condition holds with equality, so $\frac{\partial \Pi}{\partial Q_h} = 0$. The second term in the summation has the same sign as $\frac{\partial \Pi}{\partial Q_a}$ since $\frac{dQ_a}{dK_a}$ is positive regardless of whether the hotel's capacity constraint is binding (equations A6 and A8). Since $\frac{\partial \Pi}{\partial Q_a} = Q_h \frac{\partial p_h}{\partial Q_a}$ is negative because hotel and flexible rooms are substitutes, so is the derivative of hotel profits with respect to flexible capacity.

We are left with proving that an increase in flexible capacity also decreases hotel prices whenever $\frac{\partial^2 \Pi}{\partial Q_h^2}\frac{\partial p_h}{\partial Q_a} - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}\frac{\partial p_h}{\partial Q_h} \geq 0$. An increase in K_a affects hotel prices through changes in Q_h and Q_a:

$$\frac{dp_h}{dK_a} = \frac{\partial p_h}{\partial Q_h}\frac{dQ_h}{dK_a} + \frac{\partial p_h}{\partial Q_a}\frac{dQ_a}{dK_a}.$$

In the case where the hotel capacity constraint is binding, the quantity derivatives with respect to K_a are given by equations A5 and A6. So the derivative of hotel prices with respect to flexible capacity simplifies to $\frac{dp_h}{dK_a} = -\frac{G(p_a)\partial p_h/\partial Q_a}{K_a g(p_a)\partial p_a/\partial Q_a - 1}$, which is always negative. In the case where the hotel capacity constraint is not binding, the quantity derivatives with respect to K_a are given by equations A7 and A8. After substitution, the derivative of hotel prices with respect to flexible capacity becomes

$$\frac{dp_h}{dK_a} = \frac{G(p_a)}{\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a)\frac{\partial p_a}{\partial Q_a} - 1\right] - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}K_a g(p_a)\frac{\partial p_a}{\partial Q_h}}\left[\frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}\frac{\partial p_h}{\partial Q_h} - \frac{\partial^2 \Pi}{\partial Q_h^2}\frac{\partial p_h}{\partial Q_a}\right].$$

The first ratio is always positive, so hotel prices decrease with flexible capacity if and only if $\frac{\partial^2 \Pi}{\partial Q_h^2}\frac{\partial p_h}{\partial Q_a} - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}\frac{\partial p_h}{\partial Q_h} \geq 0$, which is the condition stated in the proposition. ∎

Proposition 3 *The reduction in hotel rooms sold when flexible capacity increases is larger when hotel capacity constraints do not bind. The reduction in hotel prices when flexible capacity increases is larger when hotel capacity constraints bind.*

The first part of proposition 3 is a trivial comparison of equation A5, which is always zero, and equation A7, which is never positive, and strictly negative when hotel's marginal revenue is strictly decreasing in competitors' quantity.

For the second part of the proposition, when hotel prices are a decreasing function of flexible capacity we want to prove that $\left[\frac{dp_h}{dK_a}\right]^c = \frac{\partial p_h}{\partial Q_a}\left[\frac{dQ_a}{dK_a}\right]^c \leq \left[\frac{dp_h}{dK_a}\right]^u = \frac{\partial p_h}{\partial Q_h}\left[\frac{dQ_h}{dK_a}\right]^u + \frac{\partial p_h}{\partial Q_a}\left[\frac{dQ_a}{dK_a}\right]^u$. Substituting equations A6, A7, and A8 gives

$$\frac{-\frac{\partial p_h}{\partial Q_a}G(p_a)}{K_a g(p_a)\frac{\partial p_a}{\partial Q_a}-1} \leq \frac{\frac{\partial p_h}{\partial Q_h}G(p_a)\frac{\partial \Pi^2}{\partial Q_h \partial Q_a} - \frac{\partial p_h}{\partial Q_a}G(p_a)\frac{\partial^2 \Pi}{\partial Q_h^2}}{\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a)\frac{\partial p_a}{\partial Q_a}-1\right] - \frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}K_a g(p_a)\frac{\partial p_a}{\partial Q_h}}.$$

After some algebra the inequality simplifies to $K_a g(p_a)\left(\frac{\partial p_h}{\partial Q_h}\frac{\partial p_a}{\partial Q_a} - \frac{\partial p_h}{\partial Q_a}\frac{\partial p_a}{\partial Q_h}\right) - \frac{\partial p_h}{\partial Q_h} \geq 0$, which is always true because $\frac{\partial p_h}{\partial Q_h}\frac{\partial p_a}{\partial Q_a} - \frac{\partial p_h}{\partial Q_a}\frac{\partial p_a}{\partial Q_h} \geq 0$ when hotel's and peer hosts' rooms are substitutable,[37] and because $\frac{\partial p_h}{\partial Q_h} \leq 0$. ■

The Long-Run

In the long-run, entry of flexible suppliers, i.e. K_a, is endogenous. We assume that K_h is fixed.[38] We now let the price of each room option be a function of room quantities and the demand state, $p_a^s(Q_h, Q_a)$ and $p_h^s(Q_h, Q_a)$. Recall that we assume that both prices are an increasing function of the demand state and a decreasing function of both quantities. We simplify notation by writing p_a^s and p_h^s. We define the expected daily benefit of joining Airbnb as $v_a = \int_s E_c\left(max\{0, p_a^s - c\}\right) dF(s)$, and the one-time cost of joining as C, randomly drawn for each potential host. The expression $E_c\left(max\{0, p_a^s - c\}\right)$ denotes the expected per-period profit of a flexible seller given demand state s, where the expectation is taken over the distribution of marginal costs. We also let T denote the number of days a peer host will be available to host on Airbnb after joining the platform, so that the net benefit is $Tv_a - C$. We let K_a denote the mass of potential hosts who find it profitable to join Airbnb, i.e. all those hosts with $C \leq Tv_a$.

A peer-to-peer platform enables the entry of flexible sellers. Flexible sellers decide

[37] When marginal costs are constant, substitutability between hotel's and peer hosts' rooms is defined equivalently as $\frac{\partial p_a}{\partial Q_h} \leq 0$ or $\frac{\partial Q_h}{\partial p_a} \geq 0$ (and analogously for hotel prices and peer hosts' quantities). Applying the implicit function theorem to the demand system $p_h(Q_h, Q_a) - p_h = 0$ and $p_a(Q_h, Q_a) - p_h = 0$ implies that the two are equivalent definitions of substitutability if and only if $\frac{\partial p_h}{\partial Q_h}\frac{\partial p_a}{\partial Q_a} - \frac{\partial p_h}{\partial Q_a}\frac{\partial p_a}{\partial Q_h} \geq 0$.

[38] Our model does not allow hotels to adjust dedicated capacity K_h in response to peer entry. Over many years, peer entry could partially crowd out dedicated sellers. Since our data only spans the first few years of Airbnb diffusion and hotel construction projects take between 3 and 5 years to complete, we are unable to empirically capture hotels' capacity adjustments. Exploring the entry and exit decisions of dedicated producers would be a valuable extension of our work.

whether to join the peer-to-peer platform and start producing as a function of expected demand and expected marginal costs.

Proposition 4 *Entry of flexible sellers is larger (K_a increases) if the distribution of peers' marginal costs c decreases in the sense of first-order stochastic dominance. K_a increases if K_h decreases. K_a increases if $F(s)$ increases in the sense of first order stochastic dominance. K_a also increases in response to a mean-preserving spread in $F(s)$ if peer hosts' prices are a convex function of s.*

It is intuitive that if the distribution of flexible marginal costs c shifts to the left, $E_c\left[max\{0, p_a^s - c\}\right]$ weakly increases in every demand state, so v_a increases and more flexible sellers enter.

It is also straightforward to see that if $F(s)$ shifts to the right, $E_c\left[max\{0, p_a^s - c\}\right]$ will not change for any demand state, but higher demand states are more likely so v_a increases, inducing more flexible entry.

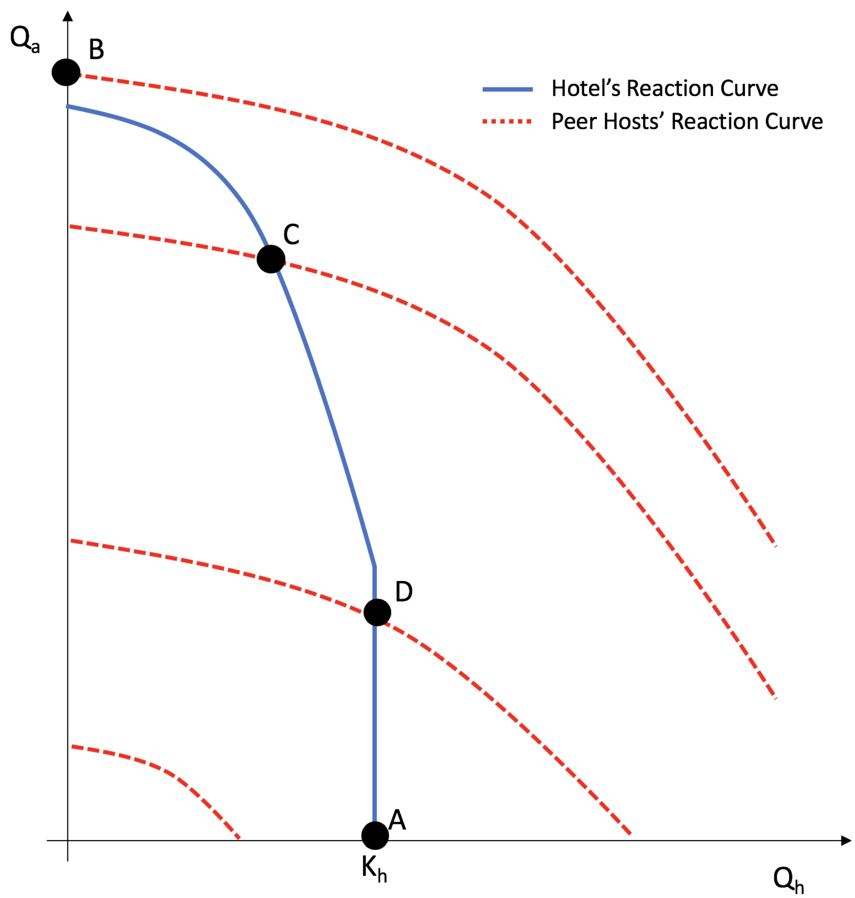
Proving that a reduction in K_h induces more flexible entry requires a little more explanation. Assume K_h decreases on the margin. For demand states for which K_h was not binding, the decrease in hotel capacity has no effect, so p_a^s does not change for s lower than a certain threshold. For demand states in which K_h was binding the two equilibrium conditions are $Q_h = K_h$ and $Q_a = K_a G(p_a^s)$. We proved above (for Propositions 2 and 3) that an increase in flexible capacity decreases both hotel and peer prices. An analogous proof is valid for a change in hotel capacity. So for high demand states a decrease in hotel capacity increases flexible prices. So far we showed that in unconstrained demand states flexible prices do not change if K_h decreases, while in constrained demand states they increase. This is a shift in the distribution of flexible prices in the sense of first order stochastic dominance. So $\frac{dv_a}{dK_h} \leq 0$ and a decrease in hotel capacity induces more flexible entry.

Finally, a mean-preserving spread of $F(s)$ induces more entry of flexible sellers if p_a^s is a convex function of s.[39] The utility function for demand state s, $E_c\left[max\{0, p_a^s - c\}\right]$, is a convex function of p_a^s, so the result is a direct implication of Jensen's inequality. Intuitively, flexible sellers lose nothing from low demand periods since they can choose not to host, and gain high profits in periods of high demand. In order to verify whether p_a^s is a convex function of s, as before we totally differentiate the system of equilibrium equations $Q_a = K_a G(p_a^s)$ and $\frac{\partial \Pi^s}{\partial Q_h} = 0$ (which is $Q_h = K_h$ if hotels are capacity-constrained) with respect to the demand state and the quantity variables. We then note that $\frac{dp_a^s}{ds} = \frac{\partial p_a^s}{\partial s} + \frac{\partial p_a^s}{\partial Q_h}\frac{dQ_h}{ds} + \frac{\partial p_a^s}{\partial Q_a}\frac{dQ_a}{ds}$. When the hotel is not capacity constrained, the total derivative is equal to $\frac{dp_a^s}{ds} =$

[39]Note that the sufficient condition that p_a^s is a convex function of s does not hold in general since it depends on both the shape of the demand curves as well as the distribution of peer costs.

$\frac{-\frac{\partial p_a^s}{\partial s}\frac{\partial^2 \Pi}{\partial Q_h^2}+\frac{\partial p_a^s}{\partial Q_h}\frac{\partial^2 \Pi}{\partial Q_h \partial s}}{\frac{\partial^2 \Pi}{\partial Q_h^2}\left[K_a g(p_a^s)\frac{\partial p_a^s}{\partial Q_a}-1\right]-\frac{\partial^2 \Pi}{\partial Q_h \partial Q_a}K_a g(p_a^s)\frac{\partial p_a^s}{\partial Q_h}}$, while when the hotel is capacity constrained the total derivative simplifies to $\frac{dp_a^s}{ds} = \frac{-\frac{\partial p_a^s}{\partial s}}{K_a g(p_a^s)\frac{\partial p_a^s}{\partial Q_a}-1}$. Convexity of p_s^a in s requires that $\frac{dp_a^s}{ds}$ is non-decreasing in s, which depends on the shape of the demand curves and the distribution of peer costs. ■

Figure A1: Equilibrium Quantities

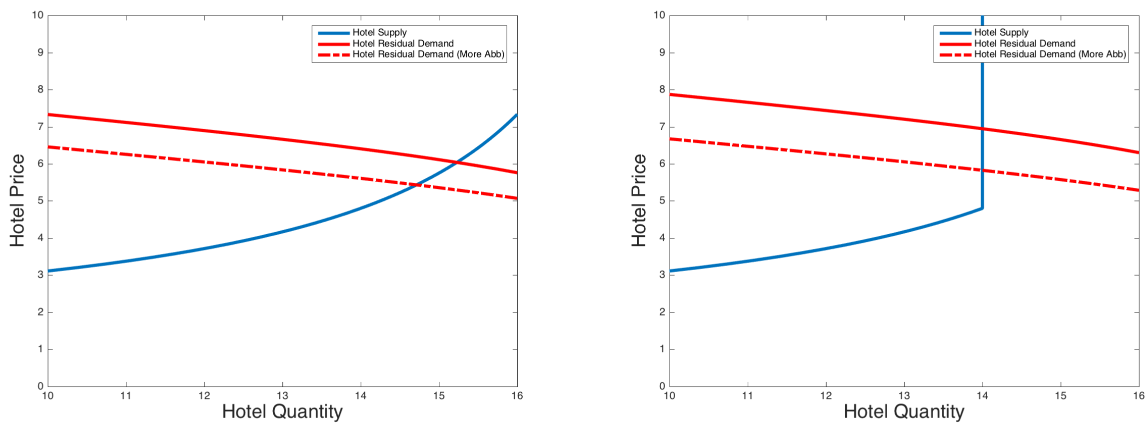


The figure plots the reaction curves, or best response functions, of the hotel and peer hosts. The hotel's reaction curve (solid line) is determined by the first order condition of its profit maximization problem, unless the optimum quantity is the maximum hotel capacity – vertical part of the solid line. The peer hosts' reaction curve (one of the four dotted lines) is determined by the equilibrium condition $K_a G(p_a) = Q_a$, where peer hosts take prices as given. Depending on the position of the peer hosts' reaction curve relative to the hotel's curve, the equilibrium is one of the four points denoted by A, B, C, or D.

Figure A2: Predictions on the Effect of Peer Supply on Hotels

(a) Unconstrained Equilibrium

(b) Constrained Equilibrium



The figures plot the supply and demand curve for hotel rooms in two scenarios. The hotel supply curve is drawn holding constant the price of peer rooms p_a, varying the demand state d, and letting the hotel set the price to maximize its profits as in Equation A1. The left panel displays an unconstrained equilibrium, while the right panel displays an equilibrium where the hotel capacity constraint is binding. Peer entry represents a downward shift in demand for hotel rooms. This downward shift will affect hotel quantity relatively more when the hotel supply curve is more elastic. The opposite is true for the effect on hotel prices, which is higher in the capacity-constrained equilibrium.

B Appendix: Computational Details and Sensitivity Analysis

In this section we describe our computational procedure to estimate the model and compute the equilibrium, as well as a sensitivity analysis of our demand estimates presented in Table 5.

To estimate the demand model, we use the PyBLP package (Conlon and Gortmaker (2020)) with some minor modifications to allow for our substitution moment. We use the KNITRO solver with a tolerance of 10^{-5} and the product rule of degree 7 to estimate the main specification. The estimates are robust to a variety of starting points. We use a 1-step procedure in order to set the relative weight on the substitution moment relative to the market share moments.[40]

We use several strategies for conducting sensitivity analysis. The first strategy is to compute estimates from our model using alternative values of the substitution moment (Equation 9). This would be useful if one were concerned that our estimate of the substitution to the outside option, which we set at 32% given Airbnb survey data, was biased. We choose two comparison values, one value that implies that an additional 10% of Airbnb travelers would substitute towards the outside option if Airbnb did not exist and another value that implies that 10% fewer Airbnb travelers would choose the outside option. We report the results in Table A1. With less substitution to the outside option, the random coefficient on the inside option would increase, as expected. Similarly, with more substitution to the outside option, the random coefficient on the inside option would decrease. As the table shows, changes in the substitution moments affect the estimates of other parameters such as price sensitivity, albeit less.

We can do a similar analysis to measure how our instruments help us identify the linear coefficient on price. To do so, we estimate the logit specification without consumer preference heterogeneity in four ways: with all our instruments combined, with each type of instrument separately, and without any instruments. The results from these estimates are displayed in Figure A3. Each set of instruments yield negative estimates for the price coefficient, in sharp contrast to the non-IV estimate, but the magnitude and precision vary. The price estimate from using all instruments jointly lies in the middle of the range of the IV estimates and is more precisely estimated than any of them.

[40] The weighing matrix is block diagonal, with $(\hat{Z}'\hat{Z})^{-1}$ in the top left and the substitution moments in the bottom right. $\hat{Z}$ is the set of instruments, with the fixed effects for month/city/scale and day of week partialed out. PyBLP does this by using the Python package 'pyhdfe' to run a regression of each column on the fixed effects and producing the residuals. This is a computational trick used to increase the speed of computation when there are many fixed effects.

Andrews et al. (2017) propose an approach to evaluate the sensitivity of our estimates to assumptions based on local perturbations. To measure sensitivity, we compute the sensitivity matrix $\Lambda = -(G * W * G)^{-1} G * W$ where G is the Jacobian of the moments with respect to the non-linear parameters and W is the weighing matrix. There are two types of moments. The first set, standard in Berry et al. (1995), includes the linear IV moments (Equation 8). The second set includes the substitution moment from Equation 9. We multiply Λ by $(\hat{Z}'\hat{Z})^{-1}$ for the linear IV moments and by $1/N$, where N is the number of markets, for the substitution moment.

We report this sensitivity measure in Figure A2 denoted as 'Raw.' We follow Andrews et al. (2017) in scaling the raw estimates by the inverse of the standard deviation of $\hat{Z}$ for each instrument. For the substitution moment, we scale it so that the sensitivity represents the response to an increase of .1 in the share of Airbnb travelers who would substitute to the outside option if Airbnb did not exist. All of the moments help to identify the non-linear parameters, as expected. We see that the implied change in the random coefficient on the constant due to a change in the substitution moment is within an order of magnitude of the changes we observed by re-estimating the model with different values of the substitution moment. We also see that the price coefficient is highly sensitive to the ratio of the Google Search Trends to capacity.

Finally, we describe some details to compute the equilibrium with our parameter estimates under different scenarios. We use a trust region reflective algorithm within the fsolve function in Matlab. We find a price vector that solves a system of hotel and Airbnb equations. The hotel equations, one per hotel scale, come from their first order conditions (Equation 6). The Airbnb equations, one per Airbnb options, come from the equilibrium condition in Equation 7, which allows us to find the price that rationalizes a particular number of booked listings. For a candidate price vector, we first find demand, which allows us to compute the markup in Equation 6 and Airbnb rooms sold in Equation 7. The algorithm then minimizes the difference between the candidate price vector and the prices implied by the supply equations. We use the baseline prices as our initial starting values. When the sum of the absolute price deviations across all options is bigger than 1e-7 we try two more starting values: one with prices that are 6.7% higher than the baseline prices, and the other with prices that are 13.3% higher. Out of these attempts, we select the solution with the lowest sum of absolute price deviations across all options.

Table A1: Estimates of Selected Demand Parameters - Varying Substitution to Outside Option Moment

	Baseline: Share to OO = .32		Share to OO = .22		Share to OO = .42	
Parameter	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Log Google Trend	2.355	0.281	2.646	0.585	2.151	0.156
Price	-0.031	0.002	-0.029	0.003	-0.031	0.002
Std. Deviation on Inside Option	1.725	1.060	2.902	2.617	0.789	0.756
Std. Deviation on Price	0.004	0.004	0.002	0.007	0.004	0.004

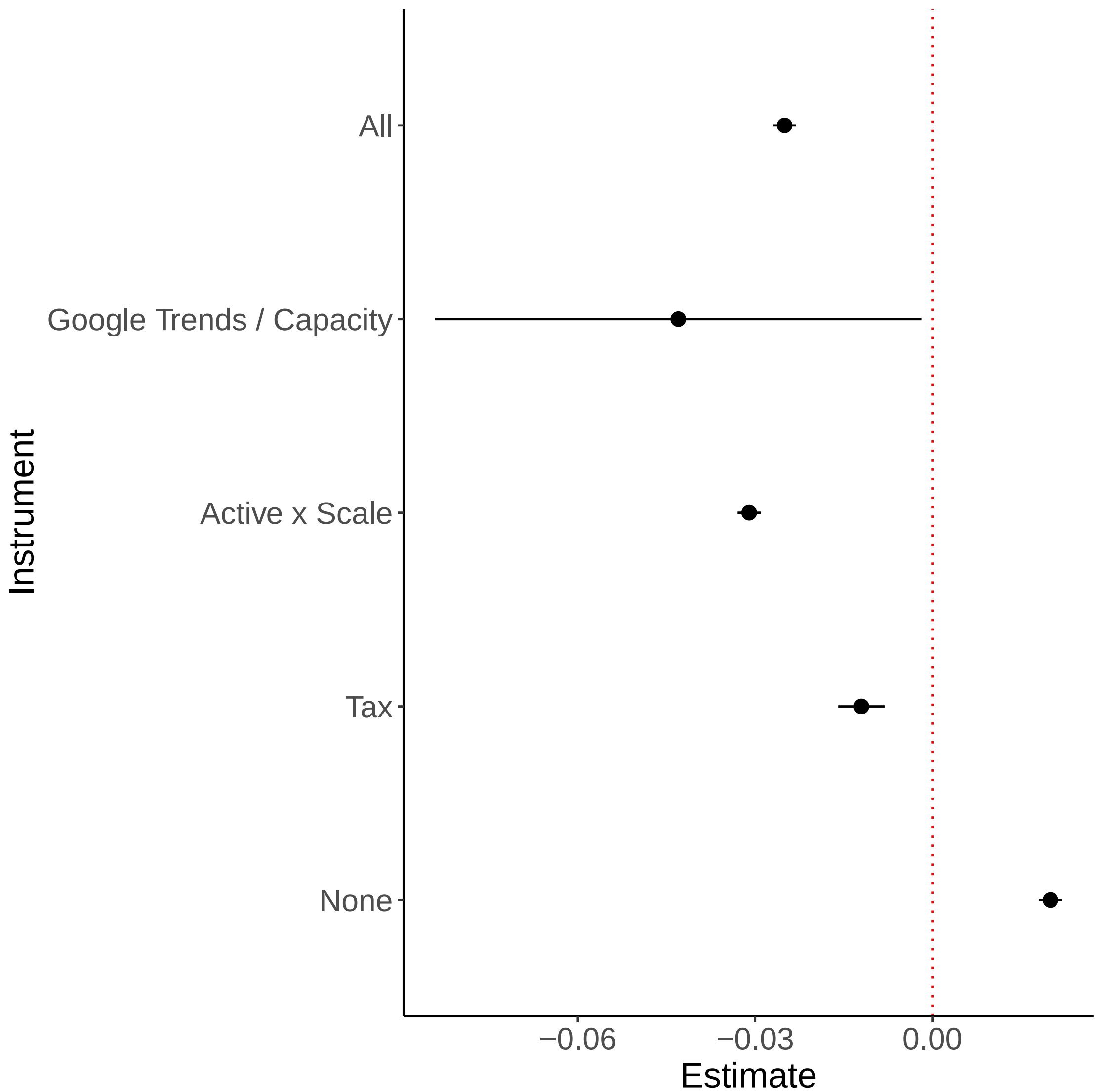
This table displays the estimates and standard errors for selected parameters in travelers' utility. The first reproduces the estimates in column 1 of Table 5, which assume that 32% of Airbnb bookings would have gone to the outside option if Airbnb did not exist. The second column assumes the substitution to the outside option is instead 22% and the third column assumes the substituion to be 42%.

Table A2: Sensitivity of Random Coefficient to Estimation Moments

Moment	Raw - Constant	Scaled - Constant	Raw - Price	Scaled - Price
Survey Moment	57.31	5.73	-0.63	-0.06
Google Trend / Capacity	-0.00	-10.70	0.00	0.27
Tax Rate	-0.02	-2.07	-0.00	-0.06
Active: Luxury	25.94	0.37	-6.60	-0.09
Active: Airbnb Midscale	-307.06	-2.17	-4.89	-0.03
Active: Airbnb Economy	-289.37	-2.67	-1.67	-0.02
Active: Upper Upscale	2.81	0.03	-6.37	-0.08
Active: Upscale	-60.51	-0.36	-13.86	-0.08
Active: Upper Midscale	-119.58	-1.53	-5.26	-0.07
Active: Midscale	20.03	0.89	-1.02	-0.05
Active: Economy	31.16	0.62	1.50	0.03
Active: Airbnb Luxury	-107.59	-0.36	-23.68	-0.08
Active: Airbnb Upscale	-235.12	-1.45	-8.49	-0.05
Log Google Trend	-1.44	-17.43	0.00	0.05
Austin/TX Hotel X Time	-37.90	-4.01	-0.66	-0.07
Austin/TX Airbnb X Time	-14.72	-2.77	-0.08	-0.02
Boston/MA Hotel X Time	-55.08	-5.82	0.43	0.05
Boston/MA Airbnb X Time	-5.72	-1.05	-0.21	-0.04
Los Angeles/Long Beach/CA Hotel X Time	-80.13	-8.47	1.06	0.11
Los Angeles/Long Beach/CA Airbnb X Time	-13.35	-1.74	-0.30	-0.04
Miami/Hialeah/FL Hotel X Time	-14.26	-1.51	-0.42	-0.04
Miami/Hialeah/FL Airbnb X Time	1.28	0.18	-0.41	-0.06
New York/NY Hotel X Time	-12.47	-1.32	-0.98	-0.10
New York/NY Airbnb X Time	-22.84	-2.96	-0.36	-0.05
Oakland/CA Hotel X Time	-71.75	-7.59	0.36	0.04
Oakland/CA Airbnb X Time	-10.19	-2.48	-0.04	-0.01
Portland/OR Hotel X Time	-92.35	-9.77	0.82	0.09
Portland/OR Airbnb X Time	-20.15	-3.77	0.02	0.00
San Francisco/San Mateo/CA Hotel X Time	-16.85	-1.78	-0.26	-0.03
San Francisco/San Mateo/CA Airbnb X Time	14.02	1.82	-0.79	-0.10
San Jose/Santa Cruz/CA Hotel X Time	-33.18	-3.51	-0.33	-0.03
San Jose/Santa Cruz/CA Airbnb X Time	-3.46	-0.81	-0.19	-0.04
Seattle/WA Hotel X Time	-69.64	-7.36	0.41	0.04
Seattle/WA Airbnb X Time	-9.98	-2.19	-0.08	-0.02

This table displays the sensitivity values of Andrews et al. (2017) for our substitution moment and for the instruments residualized by the fixed effects included in the demand model. The column 'Scaled' scales the values of 'Raw'. This is scaled by .1 in the case of the survey moment, which represents a change of .1 in the substitution to the outside option. It is scaled by the standard deviation of the residualized instruments for the other rows.

Figure A3: Sensitivity of Logit Estimates to Instruments



The figure plots the estimated price coefficients and the 95% confidence intervals for the demand model without consumer preference heterogeneity under five instrumentation options: using all our instruments, using only the Google search trends divided by room capacity, using only room capacity interacted with accommodation option fixed effects, using only tax rates, and using no instruments for price. 'All' refers to the logit specification in the paper, which is shown in column 2 of Table 5.

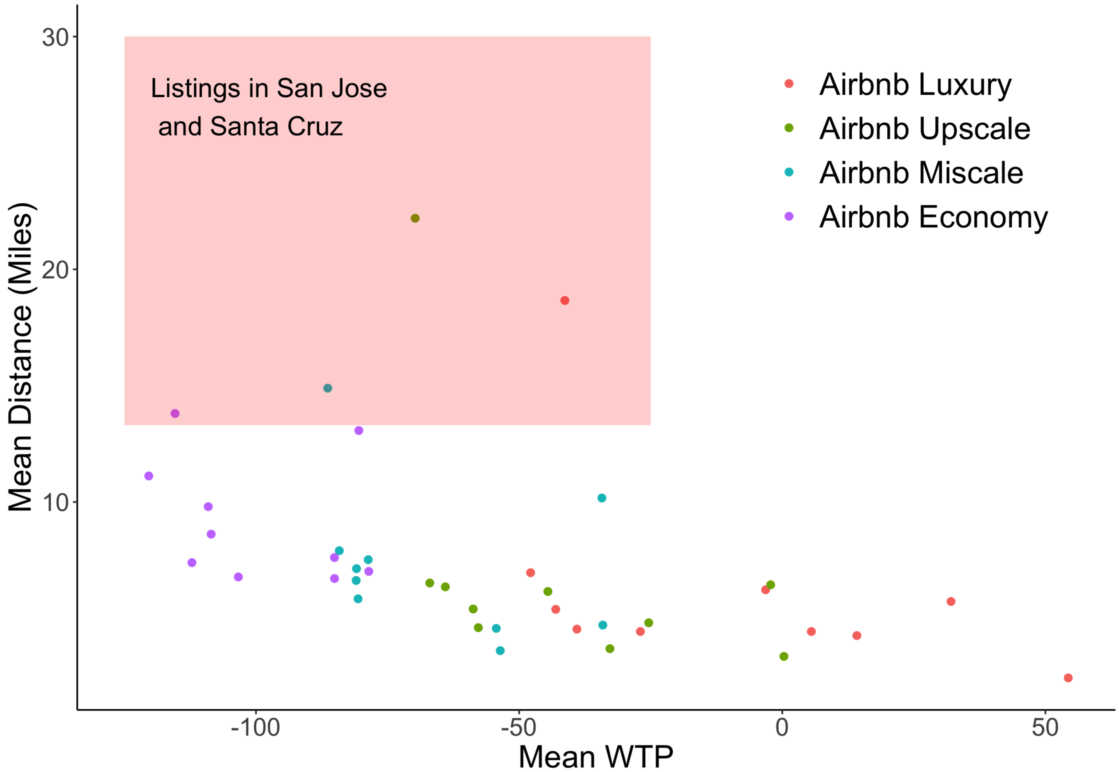
C Appendix: The Role of Location in Determining Mean Utilities

In this section we describe some results regarding the importance of location in determining the value consumers place on each accommodation type. In order to measure the 'quality' of a location, we measure its distance to visitor centers in a city. The implicit assumption is that visitor centers are located close to points of interest for leisure travelers. For each city, we find up to three top listed visitor centers according to searches on Google for 'visitor centers' in that city. Separately for each listing and hotel we calculate its distance from the closest visitor center. Finally, we aggregate the data to the scale level by taking a weighted average across all options within that accommodation category. Note that for hotels, since we do not have sales at the hotel level, we weigh each hotel by the number of available listings. For Airbnb, since we compute this distance before aggregating, we weigh listings by transactions.

We report two main findings. First, the willingness to pay of travelers for Airbnb options is decreasing in their distance from the closest visitor center. Figure A4 plots the average willingness to pay in December of 2014 against the average distance in miles. We see that better Airbnb options ('Airbnb Luxury') have a higher willingness to pay and a lower distance to visitor centers. There are several outliers, which are listings in San Jose and Santa Cruz, likely caused by the fact that leisure travel demand in this market comes from rural locations rather than urban areas. To measure the correlation, we can run a simple linear regression where the outcome variable is willingness to pay and the explanatory variable is distance to the closest visitor center. We find a negative and statistically significant relationship (Table A3). Each additional mile is associated with a $4.17 decrease in willingness to pay.

Our second finding is that hotels are often further from visitor centers than Airbnb listings. In Figure A5 we plot the average distance between hotels and the closest visitor center (triangles) as well as the average distance between Airbnb options and the closest visitor center (circles). We find that hotel options are often much further away from the visitor center than booked Airbnb options, although this varies by city and hotel type. Higher quality hotel options are located closer to the visitor center. Part of the difference in location is explained by the different weights (available versus booked rooms), but most of it is likely explained by the fact that many hotels cater to business travelers, who may want to stay close to airports, and in business districts that are not always close to tourist attractions. At the same time, luxury hotels are more likely to serve leisure travelers and are therefore located close to the most desirable places for visitors in a city.

Figure A4: Willingness to Pay and Proximity to Tourist Attractions



The figure plots the estimated willingness to pay for an option against the average distance between booked listings and the closest visitor center. Options for San Jose/Santa Cruz are highlighted in red as outliers.

Table A3: Willingness to Pay and Proximity to Tourist Attractions

	WTP
Constant	-16.050
	(15.848)
Mean Distance (miles)	-4.174**
	(1.613)
N	40
R2	0.130

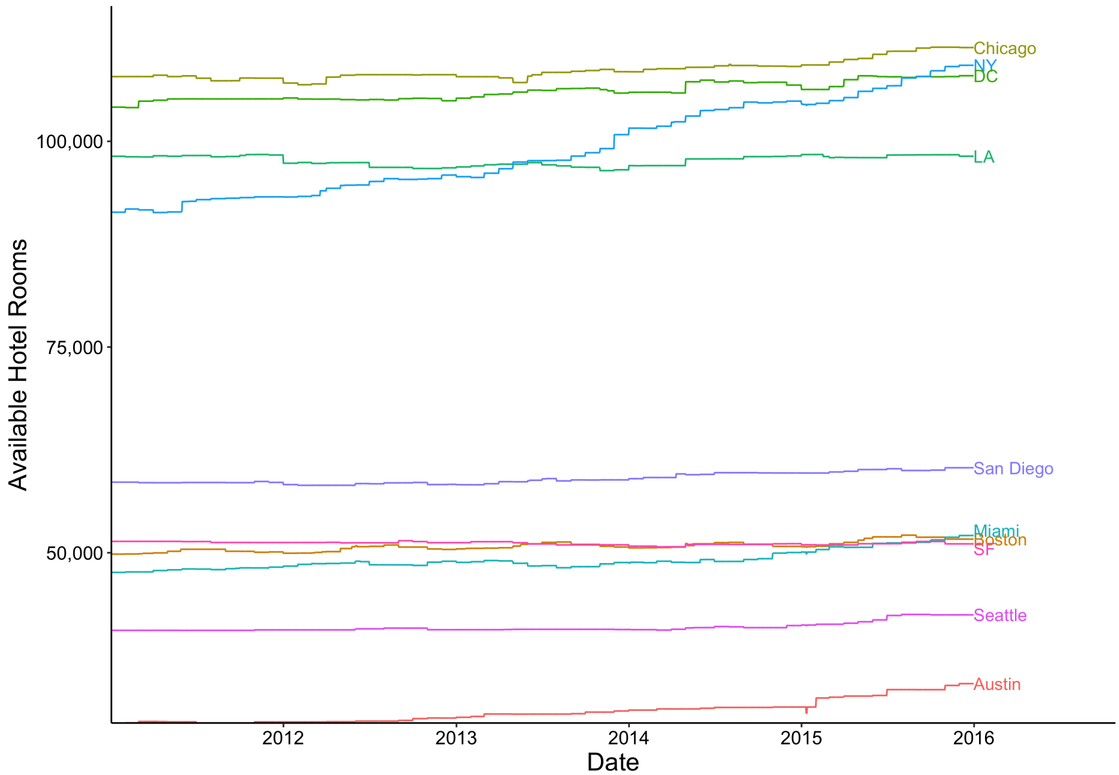
This table displays a regression of the willingness to pay for each Airbnb option on the average distance to the closest visitor center.

Figure A5: Proximity to Tourist Attractions by City and Scale

The figure plots the average distance to the closest visitor center for each accommodation option and city.

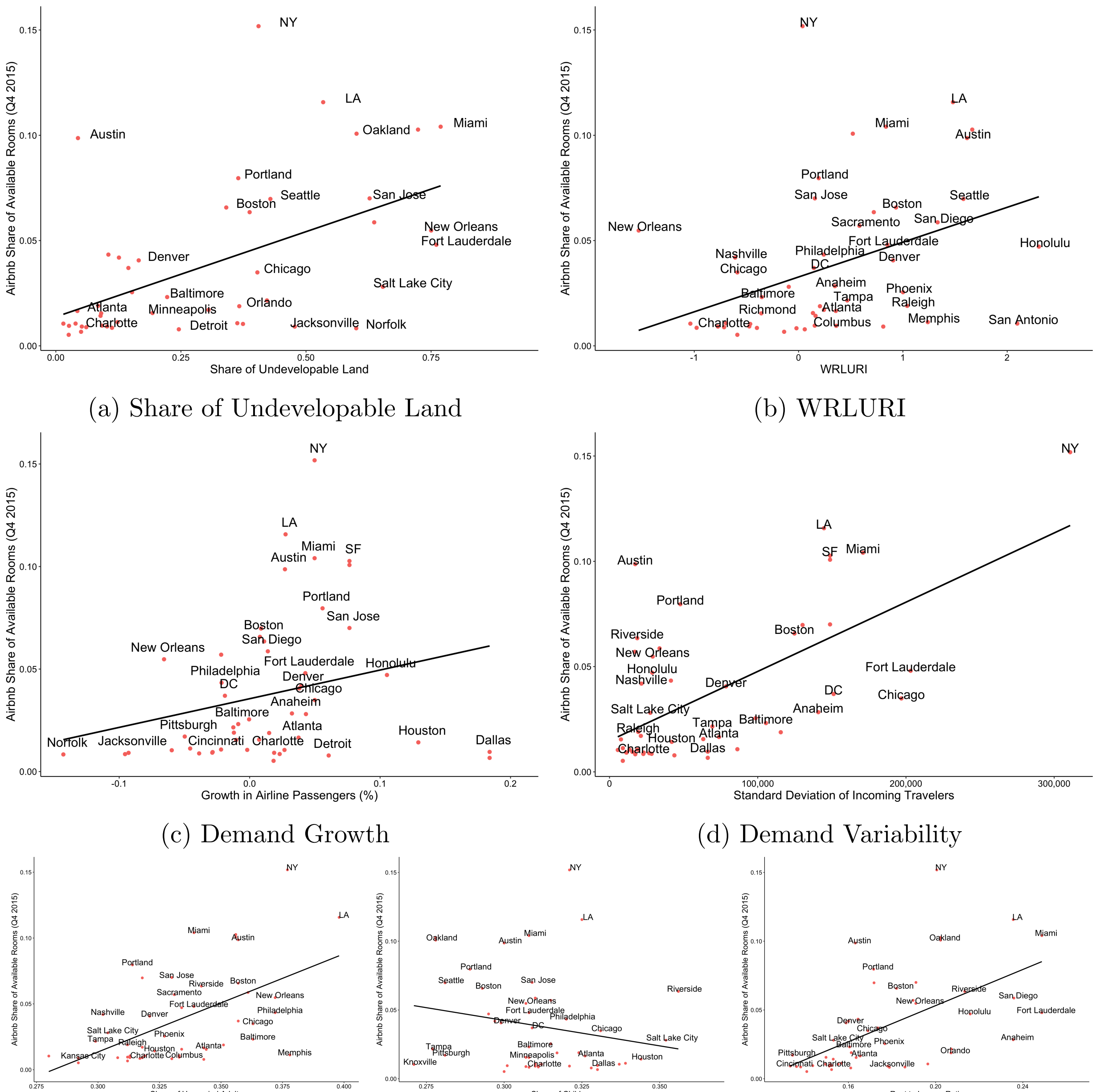
D Additional Figures and Tables

Figure A6: Hotel Rooms



The figure plots hotel room capacity over time for the top 10 cities. In contrast to the growth of Airbnb (Figure 2), the number of hotel rooms has been relatively stable over this time period.

Figure A7: Predictors of Peer Production



(a) Share of Undevelopable Land

(b) WRLURI

(c) Demand Growth

(d) Demand Variability

(e) Share of Unmarried Adults

(f) Share of Children

(g) Housing Costs

The figures plot the size of Airbnb against proxies for accommodation costs and demand characteristics. The size of Airbnb is measured as the average share of available rooms in the last quarter of 2015. Panel (a) focuses on the share of undevelopable area developed by Saiz (2010), which measures the share of a city that is undevelopable due to geographic constraints, like steep mountains or the ocean. Panel (b) uses the Wharton Residential Land Use Regulation Index, which measures the stringency of the local regulatory environment for housing development, which we consider to be similar for commercial buildings. Panel (c) plots Airbnb share against the growth rate in incoming air passengers to an MSA between June 2011 and June 2012. Panel (d) uses the standard deviation of air travelers from 2011 monthly data on arriving passengers at major US airports. Panel (e) proxies for peers' marginal costs with the share of unmarried adults in the Metropolitan Statistical Area (MSA) and Panel (f) with the share of children. Finally, Panel (g) plots the size of Airbnb against the ratio of median rent to household income in the MSA in 2010. The fitted lines weigh each city equally.

Table A4: Descriptive Statistics of Predictors of Airbnb Penetration

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
WRLURI	50	0.30	0.84	−0.36	0.20	0.85
Share of Undevelopable Area	46	0.30	0.24	0.09	0.23	0.43
Percent Never Married	48	0.33	0.03	0.31	0.33	0.36
Share of Children	48	0.31	0.02	0.30	0.31	0.32
Rent to Income Ratio	50	0.18	0.03	0.15	0.17	0.20
Std Dev of Google Trend (2011)	50	12.05	4.22	9.62	11.51	13.70
Std Dev of Incoming Passengers (2011) / 10,000	50	6.62	6.31	1.80	3.93	10.82
Passengers' Growth (2012-2011)	50	0.02	0.06	−0.02	0.01	0.04

The table shows descriptive statistics on market characteristics for the 50 cities in our sample. The WRLURI and Saiz's share of undevelopable area are proxies for constraints to hotel supply (see Gyourko et al. (2008) and Saiz (2010)). The share of children and unmarried adults proxy for hosting costs of Airbnb hosts, and are retrieved from the Census Bureau (https://www.census.gov/data.html). The standard deviation of Google trends and incoming passengers are two measures of demand volatility and are obtained from Google Trends (https://trends.google.com) and Sabre Travel Solutions, the largest global distribution systems provider for air bookings in the US.

Table A5: The Supply Elasticity of Hotels and Peer Hosts – First Stage Estimates

	log(Hotel Price) (1)	log(Airbnb Price) (2)	log(Airbnb Available Listings + 1) (3)
log(Incoming Air Passengers)	0.479***	0.38	-0.09***
	(0.041)	(1.142)	(0.014)
log(Incoming Google Searches)	0.11***	0.062	-0.009
	(0.025)	(0.785)	(0.009)
log(Hotel Rooms + 1)	-0.252*		
	(0.134)		
log(Airbnb Active Listings + 1)		0.11	0.962***
		(0.807)	(0.015)
Observations	90,900	84,959	84,959
R^2	0.854	0.578	0.999

First stage results of Table 3. Column (1) is the first stage of column (1) from Table 3. Columns (2) and (3) are the first stage of column (2). Standard errors are clustered at the city level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. *In both cases, we reject the hypotheses of under and weak identification, and reject that the joint set of instruments is not valid. For the first column, the Kleibergen-Paap LM statistic is 10.6 (p-value of 0.0049), the Kleibergen-Paap Wald F statistic is 109.5, and the Hansen J statistic is 0.11 (p-value of 0.7443). For the second and third column jointly, the Kleibergen-Paap LM statistic is 16.4 (p-value of 0.0003), the Kleibergen-Paap Wald F statistic is 44.44, and the Hansen J statistic is 0.995 (p-value of 0.3185).*

Table A6: The Supply Elasticity of Hotels and Peer Hosts – OLS

	Log(Hotel Rooms Boooked + 1)	Log(Airbnb Rooms Booked + 1)
	(1)	(2)
log(Hotel Rooms + 1)	0.628***	
	(0.187)	
log(Hotel Price)	1.063***	
	(0.058)	
log(Airbnb Available Listings + 1)		0.696***
		(0.050)
log(Airbnb Price)		0.689***
		(0.071)
City FE	Yes	Yes
Year-Month FE	Yes	Yes
Day of Week FE	Yes	Yes
Observations	91,250	85,146
R^2	0.956	0.950

OLS regression results of Equation 2. Otherwise the table is identical to Table 3. The number of observations is higher than in Table 3 because of our instrumentation strategy that uses lagged values. Standard errors are clustered at the city level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A7: Hotel Revenues and Airbnb Entry – First Stage Estimates

	log(Airbnb Available Listings + 1)	log(Airbnb Available Listings + 1)* Inelastic Housing Supply
	(1)	(2)
log(Airbnb Active Listings + 1)	0.943***	0.001
	(0.023)	(0.022)
log(Airbnb Active Listings + 1)*Inelastic Housing Supply	0.066**	1.007***
	(0.026)	(0.31)
log(Incoming Air Passengers)	-0.084***	-0.04
	(0.012)	(0.35)
log(Incoming Google Searches)	-0.017*	-0.003
	(0.009)	(0.143)
log(Hotel Rooms + 1)	-0.063	0.069
	(0.299)	(0.537)
log(Hotel Rooms + 1)*Inelastic Housing Supply	0.047	-0.126
	(0.317)	(2.825)
Observations	90,900	90,900
R^2	0.999	1

First stage results of Table 4. All columns in Table 4 have the same first stage regressions. Standard errors are clustered at the city level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$. *With a Kleibergen-Paap LM statistic of 27.2 and Wald F statistic of 1,802.3, we reject the hypotheses of under and weak identification.*

Table A8: Hotel Revenues and Airbnb Entry – OLS

	Log(RevPAR)	Occupancy Rate	Log(Price)
	(1)	(2)	(3)
log(Incoming Air Passengers)	1.103***	0.370***	0.481***
	(0.063)	(0.041)	(0.040)
log(Google Search Trend)	0.246***	0.077***	0.109***
	(0.042)	(0.012)	(0.024)
log(Hotel Rooms + 1)	−0.933***	−0.520***	−0.088
	(0.326)	(0.138)	(0.168)
log(Hotel Rooms + 1)*Inelastic Housing Supply	−0.562	0.021	−0.660**
	(0.371)	(0.168)	(0.287)
log(Airbnb Available Listings + 1)	0.022	−0.004	0.027**
	(0.023)	(0.009)	(0.012)
log(Airbnb Available Listings + 1)*Inelastic Housing Supply	−0.086***	−0.011	−0.074***
	(0.032)	(0.012)	(0.023)
City FE	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes
Observations	90,900	90,900	90,900
R^2	0.740	0.592	0.856

OLS regression results of Equation 3. Otherwise the table is identical to Table 4. Standard errors are clustered at the city level. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A9: Hotel Revenues and Airbnb Entry – Heterogeneity by Hotel Scale

	Log(Price)			
	(1)	(2)	(3)	(4)
log(Incoming Air Passengers)	0.650***	0.498***	0.446***	0.425***
	(0.071)	(0.039)	(0.038)	(0.034)
log(Google Search Trend)	0.144***	0.096***	0.114***	0.113***
	(0.051)	(0.021)	(0.024)	(0.020)
log(Hotel Rooms + 1)	−0.019	−0.097	−0.195	0.049
	(0.503)	(0.155)	(0.211)	(0.346)
log(Hotel Rooms + 1)*Inelastic Housing Supply	−0.563	−0.742***	−0.573	−0.809*
	(0.541)	(0.242)	(0.405)	(0.461)
log(Airbnb Available Listings + 1)	0.090***	0.020	0.023*	0.017
	(0.029)	(0.012)	(0.012)	(0.011)
log(Airbnb Available Listings + 1)*Inelastic Housing Supply	−0.076**	−0.063***	−0.065**	−0.072***
	(0.037)	(0.023)	(0.033)	(0.023)
Hotel Scale	Luxury	Upscale	Midscale	Economy
City FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Observations	58,176	90,900	90,900	89,082
R^2	0.796	0.739	0.814	0.913

The table shows the estimates of Equation 3 split by hotel scales. Otherwise the table is identical to the last column in Table 4. The number of observations varies because not every geography and time period has luxury or economy hotels. Standard errors are clustered at the city level. $^*p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table A10: Hotel Revenues and Airbnb Entry – Different Measures of Airbnb

	Log(Price)			
	(1)	(2)	(3)	(4)
log(Incoming Air Passengers)	0.482***	0.466***	0.481***	0.378***
	(0.040)	(0.039)	(0.040)	(0.036)
log(Google Search Trend)	0.108***	0.110***	0.109***	0.095***
	(0.024)	(0.024)	(0.024)	(0.022)
log(Hotel Rooms + 1)	−0.094	−0.064	−0.088	−0.012
	(0.171)	(0.174)	(0.168)	(0.157)
log(Hotel Rooms + 1)*Inelastic Housing Supply	−0.604**	−0.803**	−0.660**	−0.348
	(0.282)	(0.359)	(0.287)	(0.378)
log(Active Listings + 1)	0.020**			
	(0.010)			
log(Active Listings + 1)*Inelastic Housing Supply	−0.053**			
	(0.022)			
log(Available Listings Raw + 1)		−0.001		
		(0.016)		
log(Available Listings Raw + 1)*Inelastic Housing Supply		−0.098***		
		(0.029)		
log(Available Listings + 1)			0.027**	
			(0.012)	
log(Available Listings + 1)*Inelastic Housing Supply			−0.074***	
			(0.023)	
log(Booked Listings + 1)				0.060***
				(0.009)
log(Booked Listings + 1)*Inelastic Housing Supply				0.028
				(0.018)
City FE	Yes	Yes	Yes	Yes
Year-Quarter FE	Yes	Yes	Yes	Yes
Day of Week FE	Yes	Yes	Yes	Yes
Observations	90,900	90,900	90,900	90,900
R^2	0.856	0.858	0.856	0.867

The table shows results of OLS estimates of Equation 3, where the size of Airbnb is measured as the number of active listings (column 1), the number of available listings unadjusted for demand-induced calendar updates (column 2), the adjusted number of available listings (column 3), and the number of booked listings (column 4). Otherwise the table is identical to Table A8.

Table A11: Demand Own-Price Elasticities by City and Accommodation Type

	Austin	Boston	Los Angeles	Miami	New York	Oakland	Portland	San Francisco	San Jose	Seattle
Luxury	-8.35	-9.14	-9.36	-8.00	-11.03	-6.02	-5.67	-8.63	-6.13	-6.69
Upper Upscale	-5.01	-4.87	-4.42	-5.40	-7.38	-4.23	-4.54	-5.25	-5.13	-4.94
Upscale	-4.01	-4.21	-4.28	-4.51	-6.63	-3.66	-3.43	-5.04	-4.37	-4.01
Upper Midscale	-3.34	-3.85	-3.73	-4.09	-6.00	-3.46	-2.83	-5.04	-3.93	-3.42
Midscale	-2.96	-3.53	-3.24	-4.07	-6.12	-2.96	-2.58	-4.91	-3.69	-3.08
Economy	-1.66	-2.78	-2.24	-3.38	-5.45	-2.03	-1.88	-3.71	-2.69	-2.13
Airbnb Top	-4.80	-4.39	-4.45	-5.33	-5.21	-3.03	-3.36	-5.08	-3.77	-3.77
Airbnb Upper Mid	-3.98	-3.70	-3.63	-4.38	-4.34	-2.74	-3.02	-4.63	-3.29	-3.35
Airbnb Lower Mid	-3.30	-3.12	-3.02	-3.81	-3.55	-2.43	-2.52	-3.90	-2.90	-2.87
Airbnb Low	-2.52	-2.33	-2.19	-3.00	-2.56	-1.90	-2.05	-2.90	-2.12	-2.21

This table displays the own-price elasticities of demand implied by our structural estimates, computed as averages at the city and accommodation type level.

Table A12: Demand Cross-Price Elasticities by Accommodation Type

	Luxury	Upper Upscale	Upscale	Upper Midscale	Midscale	Economy	Airbnb Top	Airbnb Upper Mid	Airbnb Lower Mid	Airbnb Low
Luxury	-7.90	1.18	0.74	0.41	0.17	0.35	0.03	0.01	0.01	0.01
Upper Upscale	0.73	-5.12	0.77	0.44	0.18	0.39	0.03	0.02	0.01	0.01
Upscale	0.67	1.13	-4.42	0.45	0.19	0.41	0.03	0.02	0.01	0.01
Upper Midscale	0.63	1.12	0.77	-3.97	0.19	0.42	0.03	0.02	0.01	0.01
Midscale	0.61	1.11	0.77	0.46	-3.71	0.43	0.03	0.02	0.01	0.01
Economy	0.59	1.10	0.77	0.47	0.20	-2.79	0.03	0.02	0.01	0.01
Airbnb Top	0.68	1.16	0.78	0.47	0.20	0.44	-4.39	0.02	0.01	0.01
Airbnb Upper Mid	0.65	1.16	0.78	0.48	0.20	0.45	0.04	-3.77	0.01	0.01
Airbnb Lower Mid	0.64	1.16	0.78	0.48	0.21	0.46	0.04	0.02	-3.22	0.01
Airbnb Low	0.64	1.22	0.79	0.50	0.21	0.49	0.04	0.02	0.02	-2.45

This table displays the average own and cross-price demand elasticities across the 10 hotel scales in our estimation sample, computed as averages across the cities.

Table A13: Hotel Cost Estimates - IV versus OLS

	Hotels' Cost Function	
	(1)	(2)
$\gamma(Austin)$	13.127***	3.610***
	(0.429)	(0.067)
$\gamma(Boston)$	9.874***	3.474***
	(0.285)	(0.052)
$\gamma(LA)$	9.228***	2.747***
	(0.430)	(0.078)
$\gamma(Miami)$	21.587***	7.915***
	(0.483)	(0.104)
$\gamma(NY)$	10.399***	4.618***
	(0.246)	(0.055)
$\gamma(Oakland)$	7.920***	2.390***
	(0.421)	(0.065)
$\gamma(Portland)$	9.120***	3.067***
	(0.299)	(0.074)
$\gamma(SF)$	8.469***	3.863***
	(0.268)	(0.059)
$\gamma(SanJose)$	13.651***	3.491***
	(0.711)	(0.082)
$\gamma(Seattle)$	7.930***	3.194***
	(0.227)	(0.056)
$\gamma(UpperUpscale)$	−5.980***	−1.848***
	(0.233)	(0.043)
$\gamma(Upscale)$	−5.124***	−1.214***
	(0.282)	(0.054)
$\gamma(UpperMidscale)$	−3.457***	−0.526***
	(0.395)	(0.078)
$\gamma(Midscale)$	1.960***	0.310***
	(0.655)	(0.109)
$\gamma(Economy)$	3.210***	1.670***
	(0.839)	(0.161)
Regression Type	IV	OLS
Observations	54,660	54,660
R^2	0.786	0.902

This table displays the coefficient estimates for $\gamma_{hn} = \gamma_h + \gamma_{m(n)}$ from Equation 10, where m denotes the city and h denotes the hotel scale. City estimates correspond to Luxury hotels. Column (1) displays IV estimates where $(q_{hn} - \nu k_{hn})$ is instrumented for with the Google Search trends, while column (2) shows OLS estimates.

Table A14: Hotel Cost Estimates - Linear Component

Market	Luxury	Upscale	Upper Midscale	Midscale	Economy
Austin	221.53	103.82	100.71	86.93	54.16
Boston	240.04	137.65	122.03	107.44	90.31
Los Angeles	331.62	136.61	117.12	101.37	85.98
Miami	309.17	115.02	98.04	103.71	100.23
New York	371.49	179.75	155.30	136.20	143.97
Oakland	153.16	129.43	114.19	94.67	79.39
Portland	155.94	122.05	101.55	86.87	70.83
San Francisco	269.50	159.17	151.16	117.24	110.92
San Jose	178.65	121.50	128.67	117.22	93.43
Seattle	164.15	130.86	111.63	90.90	70.04

This table displays the average 2014 costs of hotels across cities and scales that are implied by our structural estimates. The costs shown here are the linear part of the hotel cost functions from Equation 6.

Table A15: Hotel Cost Estimates - Increasing Component

Market	Luxury	Upscale	Upper Midscale	Midscale	Economy
Austin	13.13	8.00	9.67	15.09	16.34
Boston	9.87	4.75	6.42	11.83	13.08
Los Angeles	9.23	4.10	5.77	11.19	12.44
Miami	21.59	16.46	18.13	23.55	24.80
New York	10.40	5.27	6.94	12.36	13.61
Oakland	7.92	2.80	4.46	9.88	11.13
Portland	9.12	4.00	5.66	11.08	12.33
San Francisco	8.47	3.35	5.01	10.43	11.68
San Jose	13.65	8.53	10.19	15.61	16.86
Seattle	7.93	2.81	4.47	9.89	11.14

This table displays the costs of hotels across cities and scales that are implied by our structural estimates. The costs shown here are the increasing part of the hotel cost functions from Equation 6.

Table A16: Peer Hosts Cost Estimates - IV versus OLS

	Peer Hosts' Cost Function	
	(1)	(2)
$\beta(AirbnbLuxury)$	0.017***	0.007***
	(0.0004)	(0.0001)
$\beta(AirbnbUpscale)$	0.024***	0.010***
	(0.001)	(0.0002)
$\beta(AirbnbMidscale)$	0.032***	0.014***
	(0.001)	(0.0002)
$\beta(AirbnbEconomy)$	0.037***	0.023***
	(0.001)	(0.0004)
Regression Type	IV	OLS
Observations	28,801	28,801
R^2	0.416	0.689

This table displays the coefficient estimates for β_a from Equation 11, where a denotes the Airbnb option. Column (1) displays IV estimates where the price is instrumented for with the Google Search trends, while column (2) shows OLS estimates.

Table A17: Peer Hosts Mean Costs and Standard Deviation of Costs by City

	Mean Cost			
	Airbnb Economy	Airbnb Midscale	Airbnb Upscale	Airbnb Luxury
Austin	93.63	118.89	154.15	219.80
Boston	81.44	107.04	131.88	182.25
Los Angeles	85.70	110.54	135.57	184.48
Miami	100.01	129.55	165.83	232.79
New York	92.35	123.52	157.29	197.70
Oakland	71.47	93.61	110.80	146.34
Portland	69.75	84.20	100.10	129.07
San Francisco	97.48	127.99	158.49	191.27
San Jose	78.09	102.26	120.47	155.43
Seattle	76.49	95.56	118.56	158.92
Standard Deviation	26.84	30.86	41.43	58.84

This table displays the mean costs for Airbnb options by city in 2014 implied by our structural estimates (Equation 11). The last line displays the estimated standard deviation of costs within each option type.

Table A18: In and Out of Sample Model Fit

	Average Share	Avg. Deviation		Avg. Absolute Deviation	
		In Sample	Out of Sample	In Sample	Out of Sample
Overall	0.06	-0.006	-0.007*	0.019	0.022*
Luxury	0.05	-0.015	-0.014	0.031	0.037*
Upper Upscale	0.13	-0.009	-0.01	0.053	0.065*
Upscale	0.11	-0.005	0.005*	0.032	0.04*
Upper Midscale	0.08	-0.005	-0.002*	0.015	0.018*
Midscale	0.04	-0.006	-0.011*	0.01	0.014*
Economy	0.11	-0.005	-0.019*	0.022	0.029*
Airbnb Luxury	0.01	-0.002	-0.004*	0.004	0.007*
Airbnb Upscale	0.00	-0.002	-0.004*	0.003	0.005*
Airbnb Midscale	0.00	-0.002	-0.004*	0.002	0.005*
Airbnb Economy	0.00	-0.002	-0.004*	0.002	0.004*
Austin	0.07	-0.007	-0.008	0.023	0.026*
Boston	0.06	-0.005	-0.006	0.019	0.023*
Los Angeles	0.05	-0.005	-0.005	0.013	0.017*
Miami	0.05	-0.006	-0.013*	0.017	0.023*
New York	0.06	-0.005	-0.006	0.02	0.027*
Oakland	0.06	-0.006	-0.004	0.021	0.02
Portland	0.06	-0.005	-0.003*	0.015	0.017*
San Francisco	0.05	-0.006	-0.011*	0.021	0.025*
San Jose	0.06	-0.006	-0.006	0.026	0.028*
Seattle	0.06	-0.006	-0.005	0.017	0.02*

This table displays the average market share across the markets and accommodation options (first column) and how well our model can replicate these market shares. To do this we simulate each market 200 times. For each simulation, we randomly draw a demand shock from the iid shocks implied by our model estimates and we compute simulated market shares for each accommodation option in a given market. We then compute the average deviation from the realized market share as well as the average absolute deviation. We do this separately for the first and second half of 2015. The first half of 2015 is considered "in sample" because it is used in the estimation, while the second half is "out of sample." The first row implies that the average market share is 0.06, and the average deviation from the realized market share is -0.006 for the first half of 2015 and -0.007 for the second half. This means that on average we are underestimating market shares by about 10%. The average absolute deviation in sample is 0.019, while out of sample it is 0.022. This means that on average our implied market shares are off by a third approximately. The stars denote whether the difference between in sample and out of sample deviations is significant at the 5% confidence level.

Table A19: Change in Consumer Surplus By Markets

City	Airbnb Rooms Sold (Baseline) (Thousands)	Change in CS (Million)				
		Unconstrained	No Airbnb	Airbnb with Lodging Tax	Airbnb with Quotas	Double Airbnb
Austin	149	-5.12	-9.65	-2.51	-3.17	4.51
Boston	210	-6.96	-14.04	-2.45	-6.17	5.78
Los Angeles	772	-26.04	-43.51	-9.14	-21.28	18.00
Miami	273	-9.41	-15.26	-3.51	-6.89	5.94
New York	1,776	-58.86	-141.25	-31.56	-79.65	61.86
Oakland	113	-3.81	-6.77	-0.93	-2.93	2.79
Portland	166	-5.67	-9.02	-0.68	-3.89	3.93
San Francisco	635	-21.51	-47.64	-10.83	-25.89	20.08
San Jose	115	-3.88	-6.51	-0.96	-3.30	2.85
Seattle	175	-5.98	-11.03	-2.24	-3.96	4.69
All	4,381	-147.23	-304.70	-64.79	-157.12	130.43
Non-Compression Nights	3,218	-107.99	-183.26	-41.09	-143.38	76.95
Compression Nights	1,163	-39.24	-121.44	-23.70	-13.74	53.48

This table displays the change in consumer surplus from five alternative scenarios, two scenarios without Airbnb and three scenarios with Airbnb and regulation. The first column shows the number of Airbnb rooms sold in 2014 . The next columns splits the results in column 1 of Table 6 by city and compression nights.

Table A20: Competitive Effects on Hotels By Markets

City	Rooms Sold (MM)						Revenues ($MM)					
	Baseline	Unconstr.	No Airbnb	Lodg. Tax	Quotas	2*Airbnb	Baseline	Unconstr.	No Airbnb	Lodg. Tax	Quotas	2*Airbnb
Austin	8	8	8	8	8	8	1,043	1,059	1,056	1,046	1,050	1,038
Boston	14	14	14	14	14	14	2,477	2,502	2,495	2,480	2,489	2,469
Los Angeles	28	29	29	28	29	28	4,162	4,239	4,227	4,176	4,207	4,136
Miami	14	14	14	14	14	14	2,643	2,678	2,673	2,649	2,664	2,631
New York	33	34	33	33	33	32	8,830	9,148	9,037	8,871	8,979	8,750
Oakland	5	6	6	5	6	5	635	644	643	636	640	632
Portland	7	7	7	7	7	7	795	808	807	796	802	790
San Francisco	16	16	16	16	16	16	3,259	3,347	3,318	3,272	3,301	3,234
San Jose	10	10	10	10	10	10	1,410	1,421	1,419	1,411	1,416	1,406
Seattle	11	11	11	11	11	11	1,550	1,567	1,563	1,552	1,557	1,544
All	146	149	148	146	147	145	26,803	27,412	27,238	26,891	27,106	26,630
All (Non Compression)	113	116	115	114	115	113	19,563	19,977	19,898	19,633	19,850	19,431
All (Compression)	33	33	33	33	33	32	7,240	7,435	7,341	7,258	7,256	7,199

City	Profits ($MM)						Alternative Profits ($MM)					
	Baseline	Unconstr.	No Airbnb	Lodg. Tax	Quotas	2*Airbnb	Baseline	Unconstr.	No Airbnb	Lodg. Tax	Quotas	2*Airbnb
Austin	289	295	294	290	289	287	594	604	603	596	598	591
Boston	580	587	587	581	582	577	1,345	1,359	1,356	1,347	1,351	1,340
Los Angeles	510	520	527	513	513	503	2,154	2,194	2,193	2,162	2,177	2,139
Miami	543	552	550	544	543	540	1,823	1,847	1,845	1,827	1,837	1,813
New York	2,264	2,350	2,354	2,281	2,310	2,230	5,440	5,633	5,591	5,470	5,540	5,380
Oakland	89	91	92	90	90	88	149	152	153	150	151	148
Portland	125	127	128	125	125	123	313	318	318	313	315	310
San Francisco	615	633	640	621	627	605	1,705	1,751	1,744	1,714	1,730	1,689
San Jose	390	393	393	390	391	389	594	598	599	595	597	592
Seattle	282	286	287	283	283	280	752	761	760	754	756	749
All	5,687	5,833	5,852	5,718	5,754	5,623	14,869	15,216	15,162	14,928	15,050	14,750
All (Non Compression)	2,272	2,327	2,357	2,289	2,330	2,243	10,454	10,682	10,661	10,498	10,624	10,373
All (Compression)	3,414	3,505	3,495	3,429	3,424	3,380	4,414	4,535	4,501	4,430	4,426	4,377

This table displays hotel bookings, revenue, and profits from the baseline and five alternative scenarios, two scenarios without Airbnb and three scenarios with Airbnb and regulation. The table splits the hotel results of Table 6 by city and compression nights. The costs used in the profit calculation are those estimated from Equation 6, except that we exclude the increasing cost component from the computed costs. The costs used in the alternative profit calculation are derived from imputed accounting costs combining the wage bill in the STR data and trends in the wages of maids. This is likely a lower bound on the true marginal cost of hotels.

Table A21: Peer Producer Surplus

City	Rooms Sold (Thousands)				Revenues ($Thousands)				Total Peer Surplus ($Thousands)			
	Baseline	Lodg. Tax	Quotas	2*Airbnb	Baseline	Lodg. Tax	Quotas	2*Airbnb	Baseline	Lodg. Tax	Quotas	2*Airbnb
Austin	149	111	64	206	20,576	13,916	10,819	26,936	3,604	2,420	1,765	4,604
Boston	210	167	77	296	22,053	16,298	8,270	28,494	5,121	3,812	1,936	6,675
Los Angeles	772	606	251	1,088	82,794	60,009	26,929	106,710	19,413	14,212	6,671	25,424
Miami	273	211	90	378	35,076	25,254	12,875	45,475	6,135	4,386	2,217	7,842
New York	1,776	1,400	585	2,498	225,359	164,190	76,004	294,644	46,722	33,895	16,359	60,979
Oakland	113	96	41	159	8,792	6,957	3,160	10,838	2,769	2,242	1,019	3,590
Portland	166	151	62	240	13,435	11,778	5,066	17,101	4,411	3,896	1,725	5,831
San Francisco	635	492	210	900	81,221	58,480	27,714	106,600	16,567	11,858	5,869	21,718
San Jose	115	94	34	162	10,736	8,172	3,291	13,634	2,830	2,193	845	3,704
Seattle	175	139	75	250	16,537	12,074	7,320	21,168	4,477	3,307	2,071	5,840
All	4,381	3,468	1,491	6,177	516,577	377,128	181,446	671,600	112,048	82,220	40,477	146,206
Non-Compression Nights	3,218	2,527	482	4,528	371,852	270,032	55,957	482,062	80,602	58,941	12,860	105,188
Compression Nights	1,163	942	1,009	1,649	144,725	107,096	125,489	189,538	31,446	23,280	27,617	41,018

This table displays Airbnb bookings, revenue, and profits from the baseline and three alternative scenarios, two scenarios without Airbnb and three scenarios with Airbnb and regulation. The table splits the Airbnb results of Table 6 by city and compression nights. Airbnb costs are taken from the distribution by the parameter estimates of Equation 7 truncated at zero (i.e., negative costs are considered equal to zero).

Table A22: Airbnb Bookings: Market Expansion versus Business Stealing

City	Share New Bookings	
	Unconstrained	No Airbnb
Austin	0.29	0.49
Boston	0.33	0.60
Los Angeles	0.32	0.51
Miami	0.33	0.51
New York	0.33	0.70
Oakland	0.32	0.55
Portland	0.32	0.49
San Francisco	0.33	0.69
San Jose	0.33	0.51
Seattle	0.33	0.57
All	0.33	0.62
Non-Compression Nights	0.33	0.53
Compression Nights	0.31	0.87

This table shows the share of Airbnb bookings in the 'Baseline' scenario that would not have been hotel bookings in the two counterfactual scenarios without Airbnb. This represents the share of Airbnb bookings constituting market expansion. The two counterfactual scenarios are defined as in Table 6. All calculations are for 2014.

Table A23: Aggregate Surplus (MM) – Standard Logit Estimates

	Consumers	Hotels			Peer Hosts			Government
	Change in Consumer Surplus	Rooms Sold	Revenues	Profits	Rooms Sold	Revenues	Peer Surplus	Lodging Taxes
Panel A: All markets in 2014								
Baseline		146	26,804	4,847	4.38	517	112	3,986
No Airbnb (Unconstrained)	-175	148	27,306	4,959				4,055
No Airbnb	-274	147	27,117	4,954				4,027
Airbnb With Lodging Tax	-55	146	26,863	4,866	3.53	385	84	4,055
Airbnb With Quotas	-153	147	27,026	4,890	1.49	181	40	4,015
Double Airbnb Rooms	107	146	26,689	4,809	6.01	649	141	3,971
Panel B: Compression Nights in 2014 (19.6% of all markets)								
Baseline		33	7,241	2,871	1.16	145	31	1,084
No Airbnb (Unconstrained)	-47	33	7,410	2,942				1,108
No Airbnb	-98	33	7,310	2,923				1,093
Airbnb With Lodging Tax	-19	33	7,251	2,880	0.95	108	24	1,103
Airbnb With Quotas	-12	33	7,252	2,878	1.01	125	28	1,086
Double Airbnb Rooms	39	33	7,215	2,851	1.61	184	40	1,081

This table is the same as Table 6 except that it uses the demand and supply estimates without consumer preference heterogeneity. The demand parameter estimates used for this counterfactual scenarios are presented in the last two columns of Table 5. All variables are in millions.