Welfare Estimation in Peer-to-Peer Markets with Heterogeneous Agents: The Case of Airbnb

Abdollah Farhoodi *†

November 18, 2019

Please find the latest version here

Abstract

Peer-to-peer (P2P) markets allow small suppliers with limited capital to enter markets that were traditionally occupied by large firms. This feature provides a potential decentralized distribution of opportunities. To investigate the distribution of welfare and opportunities among agents, I study the Airbnb short-term rental market, as a successful P2P marketplace. I use daily panel data of Airbnb rentals in Chicago from August 2014 through April 2017 and apply an individual-level multinomial logit model to estimate the distribution of consumer and producer surpluses across differentiated agents and over time. I show that properties in less advantaged neighborhoods benefit the least from having access to the Airbnb market, even though these properties feature weaker competitive pressure and lower opportunity costs of renting. My results show a disproportionate concentration of welfare in neighborhoods with higher incomes and house prices. I also show evidence of a higher surplus for low-income property owners, especially for those who live in high-demand areas.

*Email: farhood2@illinois.edu; Address: Department of Economics, University of Illinois, Urbana, IL 61801

†I am indebted to Dan Bernhardt for his invaluable advice. I am grateful for valuable guidance from David Albouy, George Deltas, Jorge Lemus, and Guillermo Marshall.

1 Introduction

Peer-to-peer (P2P) markets are decentralized markets where individual buyers and sellers trade directly, usually through an online platform. Over the last decade, P2P markets have been expanding exponentially. Marketplaces such as Uber, Airbnb, Lending Club, and many others facilitate consumers' access to transportation, accommodation, and financial loans; they also let small suppliers profit from their spare capital. This feature suggests potentially large welfare gains in P2P markets for small entrepreneurs, making micro-entrepreneurship possible for people with limited access to capital and enabling "democratization of opportunities" (Sundararajan, 2016).

This paper contributes to the literature is with study of welfare distribution in P2P markets, while accounting for individual-level heterogeneities on both the supply and demand sides of the market. Specifically, this paper addresses the question whether opportunities are distributed equally among people who rent out their properties in the Airbnb short-term rental market. On the one hand, low-income suppliers are expected to value the extra income more highly than their higher-income counterparts and incur a lower opportunity cost for supplying their properties to the rental market. On the other hand, these properties are often located in low-income neighborhoods, while there is a higher demand for properties in neighborhoods with better amenities. This demand may shift the surplus toward suppliers whose properties are already in located higher-income areas.

I show that high-income neighborhoods earn a higher share of the total surplus in the Airbnb market. Moreover, properties that are located in high-demand neighborhoods earn higher surpluses per booking; even though suppliers who live in higher-income neighborhoods incur higher opportunity costs, they still benefit from higher welfare gains in the market. However, I also show evidence that suggests the presence of a higher benefit to low-income suppliers within the same neighborhoods. I show that, on average, the total surplus going to property owners who value the marginal dollar more highly is higher. My results show that the Airbnb market is more beneficial to low-income property owners whose properties are located in high-demand neighborhoods.

I use a unique dataset providing detailed information regarding Airbnb listings in

Chicago over a roughly two-and-a-half-year period running from August 2014 through April 2017 (AirDNA, 2017). The data contain around 9 million observations and include information about available Airbnb accommodation characteristics: the location, the number of bedrooms, the number of bathrooms, the type of accommodation (entire place, private or shared), hosts' ratings, number of photos, and other information available to a potential renter through the Airbnb website. The dataset also includes information about booking times and the booking status of each apartment. I use this information to estimate demand for each property in the market.

The findings show that the average consumer surplus is around 20% of the paid price for accommodations but it varies seasonally with a higher surplus during summer months. Furthermore, the average producer surplus is around 65% of hosts' revenue and is concentrated primarily in Central Chicago, in neighborhoods with higher income, and in upper-class neighborhoods. This is because people are willing to pay more for the amenities that are available in upper-class neighborhoods. My estimations show that, on average, renters are willing to pay up to $130 more for the same property located in downtown than they would pay to stay in some areas in the southern part of Chicago. Consequently, even though the opportunity cost of providing an Airbnb listing is lower for properties located in less advantaged Chicago neighborhoods, they cannot earn as much surplus per night as properties located in more touristy and business-oriented areas. For instance, the share of the total surplus in my estimations is roughly 22% for Central Side of Chicago, while this area includes only about a 1% share of the total population.

Despite the lower average surplus earn by properties located in high-demand neighborhoods, I show evidence that low-income property owners benefit the most from having access to the Airbnb market. In the absence of data indicating property owners' incomes, I use my estimations of opportunity costs as a proxy for owners' value of a marginal dollar. I show that property owners who value the marginal dollar 1% more earn, on average, 0.17% more total surplus. This effect is larger for properties that are located in the Central Side of Chicago where demand is highest. In this area, a 1% higher marginal dollar value is associated with a 0.32% higher total surplus for property owners.

The study of distribution of welfare and opportunities among agents in P2P markets requires accounting for existing heterogeneities that generate differences in opportunity costs on the supply side and preferences on the demand side of the market. Measuring welfare distribution while accounting for differences between agents is challenging. Unlike traditional markets where big firms supply masses of homogeneous goods, in P2P markets agents are usually small, they have limited capacities, and they trade differentiated products in decentralized markets with various prices that change rapidly. This decentralized structure poses a challenge when using aggregated methods of supply and demand estimation for welfare analysis.

In this paper, I tackle the problem of welfare analysis in P2P markets with heterogeneous agents. I use daily panel-data of transactions in Chicago for the Airbnb rentals, to study the welfare of both guests and hosts in the market. I estimate time-varying individual-level consumer and producer surpluses. My estimations provide a better understanding of the distribution of welfare in the Airbnb market. To the best of my knowledge, this is the first study that uses this level of dis-aggregation for both consumer and producer surplus estimations.

I estimate demand for each property (listing) in the Airbnb market using an individual-level multinomial logit model (McFadden, 1977). Bayer et al. (2007) estimate a similar individual-level demand for each house in housing market. In contrast to their paper, mine adds four features to the estimations. First, my data includes repeated renting of listings, so I estimate the model with panel data and add a time dimension to the estimations. Second, to study the distribution of welfare across suppliers, I add the supply side with strategic pricing to the model. Third, around half of the listings in the Airbnb market are not booked in a given day, and many buyers choose a hotel or another option instead of booking an Airbnb. So, I allow for presence of vacant units and the possibility of choosing an outside option. Fourth, listings that are booked are not available to potential buyers, so I allow for variations in choice sets depending on when guests choose to book a property. These are important features that are essential for modeling agents in the Airbnb market and in some cases in the housing market. The method that I use is applicable to housing sales and rental markets or, with certain adjustments, to other P2P markets.

On the supply side, I model hosts in a static, monopolistic competition environment with heterogeneity in opportunity costs and competition in prices. I derive the relationship between the producer surplus and the demand for each listing from the supply side profit maximization problem. I then use this relationship to find the producer surplus and, consequently, the opportunity costs of renting out differentiated properties in the market.

My estimation method applies an instrumental variable (IV) to identify the price elasticity of utility. I use BLP-IV as a commonly used IV in the literature (e.g, Bayer et al. (2007), Berry et al. (1995)). It includes the average characteristics of other listings located far from a specific property. The identification assumption is that characteristics listings located beyond a certain distance from those in question do not directly affect a guest's utility for renting a property. The identification assumption may be violated if unobservable demand shocks are correlated with specific characteristics of listings such as the number of bedrooms. As a robustness test, I check the results with a second IV that is created based on the share of days on which a listing was blocked. A listing is either available, reserved or blocked in the data, and I use the share of blocked days as an IV for price. The identification assumption for this new IV is that hosts respond to demand shocks only by changing their prices, and blocking depends on hosts' personal schedules. In other words, the assumption is that the hosts' personal schedules are not correlated with the demand shocks. Both IVs generate very similar estimations of parameters and support the accuracy of the identification assumptions.

Few recent papers study welfare estimation in P2P markets. In a closely related paper, Farronato and Fradkin (2017) estimate welfare in the Airbnb market. They show that welfare gains are higher where, and at times at which, hotels are capacity-constrained. For their estimation, they use aggregate market shares of Airbnbs and hotels to apply a discrete choice model of demand for and supply of listings. I diverge from their paper by applying individual-level variations in prices to estimate the distribution of welfare and account for heterogeneities among hosts. So, the choice set of a potential guest in my model includes all available listings at the time of booking while potential guests in their paper choose between hotels and Airbnbs in the aggregate.

Lam and Liu (2018) estimate real-time demand for Uber and Lyft ride sharing. Like my paper, theirs accounts for heterogeneity in surpluses over time. In the home-sharing market, in addition to time, heterogeneities among hosts' characteristics and locations play important roles in shaping consumer and producer surpluses. My model takes these differences into account. Similarly, Cohen et al. (2016) use the Uber surge-pricing algorithm to estimate the demand elasticity along several points of the aggregate demand curve and calculate total consumer surplus in the market. For further reading, Einav et al. (2016) provide a good review of the literature on P2P markets. My paper is the first that estimates the distribution of welfare among differentiated agents in a P2P market.

Several papers study the effects of P2P markets on other markets. For instance, Fraiberger et al. (2015) conclude, after calibrating their model, that car ownership declines in response to growth in the P2P car rental market. Zervas et al. (2017) study the effects of Airbnb on the hotel industry. Sheppard et al. (2016) and Barron et al. (2018) study the effect of Airbnb on housing market. Unlike these papers, mine focuses on the welfare estimation and its distribution in the Airbnb market, and it does not estimate the market's external effects on other markets.

This paper is structured as follows: In Section 2 I present the data, summary statistics, and Airbnb marketplace trends. In Section 3 I introduce the model. In section 4 I discuss the estimation and identification methodologies, and In Section 5 I discuss the estimation results and distribution of welfare in the market. I conclude in Section 6.

2 Data and Summary Statistics

This paper employs micro-level data on property listings on the Airbnb platform in the city of Chicago with around 9 million observations. The data are acquired from a private firm, Airdna.co, which collects the data on Airbnb listings across the United States (AirDNA, 2017). The data cover the period running from August 2014 through April 2017 and include an unbalanced panel of information about prices, listings' status (available, reserved, blocked), and booking times. The data also contain information about location and the characteristics of listings that are available on the Airbnb website. These characteristics consist of: the number of bedrooms, the number of bathrooms, the number of photos, the types of listings (the entire place, private or shared rooms), deposit amounts, whether the listing is for a "super host," business-ready (business ready listings provide specific accommodation standards, such as access to the internet and late entry), and so on. The analysis of this paper is restricted to Chicago, but it could easily be conducted in other markets where data are available.

The market includes 23,485 distinct properties in the time span of the data, covering most neighborhoods in the city. A median property is booked 65 times and earns about $6800 over 2.5 years. These properties earned a total of $195 million in the time period of the data that runs from August 2014 through April 2017. Total revenue is not uniformly distributed in the market. Figure 1 shows some block-groups earn more than $3 million in just 2.5 years of the data while other neighborhoods have not benefited much from the Airbnb platform.

The Airbnb market is highly dynamic, and its scope has increased exponentially over time. Figure 2 shows that total monthly revenue exhibits strong seasonal variation and almost doubled every year. It is important to take this variation in the scope of market into account for welfare analysis. Figure 3 zooms in on the last month of the data and shows large high-frequency daily variations, peaking over the weekends, in total revenue.

The market experiences seasonal variation in average prices. Figure 4 plots the trends in the average prices of booked and all listings in the market. Both trends follow seasonal variations with higher prices over the summer. An average booked property costs around $120 while an average unbooked or booked property costs $140 per day for renters.

These trends show wide variations in the market over time and across geography. They also point out the importance of heterogeneity in listing characteristics (such as the number of bedrooms and listing types) when analysing the Airbnb market. In the following section I show how I account for these heterogeneities in welfare estimation.

3 Model

In this section, I introduce a static model of supply and demand for home-sharing. On the demand side, guests face a discrete set of options and choose their best available option. Guests are price takers, and they maximize utility based on observed prices and listing attributes. On the other hand, suppliers compete in a monopolistic competition environment with competition over prices. Suppliers' decisions are based on their profit functions and the probability that a property will be booked.

I first start with the demand side of the model and guests' rental decisions on each day. Potential guests choose among available listings in the city with a specific capacity each day. The utility function is based on Berry et al. (1995) and Bayer et al. (2007). Guests choose among a discrete set of available listings at the time of booking to maximize their indirect utility. The indirect utility of guest i from choosing listing h is as follows:

$$u_h^i = \delta_{ht} + \epsilon_h^i = \alpha P_{ht} + \xi_h + \gamma_t + \epsilon_h^i \tag{1}$$

where

$$\xi_h = \beta X_h + \gamma_l + \nu_h,$$

δ_{ht} is the mean utility for listing h at time t, and ϵ_h^i is the individual specific utility for the listing. I follow the discrete choice literature and assume that the error term ϵ_h^i is independently and identically distributed and draws from an extreme value type-I distribution function. In the absence of more information about guest characteristics, I assume that the utility parameters are constant, and ϵ_h^i is the only heterogeneity among consumers.

The mean utility for each listing at a given time, δ_{ht}, is divided into three parts; αP_{ht} is the disutility from a given price, ξ_h is the listing-specific utility, and γ_t is the time-specific utility. γ_t is a day fixed effect and captures the average utility from booking an Airbnb listing on a specific day. The listing-specific utility, ξ_h, includes the utility from a listing's observable characteristics, X_h, such as the number of bedrooms, bathrooms, and reviews as well as location fixed effects, γ_l, and a control, ν_h, for unobservable characteristics. ν_h captures the utility derived from unobservable characteristics such as utility from the exact location of the listing or information that is available to guests by looking at photos of a listing but is not available to econometricians.

Every potential guest chooses among a discrete set of available listings at the time of booking to maximize her utility. Thus, the probability of choosing a listing depends on the utilities from all available options at the time of booking. Next, I normalize the utility from the outside option to zero to find the probability that any particular listing is chosen. One can consider the outside option as the option to book a hotel instead of an Airbnb or even staying home. Using this normalization, s_{iht} is the probability that potential guest i chooses listing h at time t:

$$s^i_{ht} = \frac{\exp(\delta_{ht})}{1 + \sum_{k \in M^i_t} \exp(\delta_{kt})} \tag{2}$$

where, M^i_t is set of all Airbnb options available to the guest at the time of booking. This set has a superscript i and varies for guests depending on time of a booking. Here, I allow for variations in the choice sets depending on guests' booking time. The variable choice set takes into account cases in which a listing is already booked and is no longer available in other guests' choice set anymore.

Variable choice set is an important feature of the model that should be taken into account when suppliers have limited capacity and goods are short-term rented, as in P2P markets. As is true of many P2P markets, the housing market also includes suppliers that have only one property to sell. In demand estimation in these markets, an already sold property should not appear in the buyers' choice set. Having information about the exact timing of each booking allows me to include variable choice sets in my estimations. I exclude an early booked listing from the choice set of a guest who decides to book.

Next, I model the profit maximization of hosts. On the supply side, hosts or suppliers compete in a static monopolistic competing environment with competition over prices. They choose their prices to maximize their expected profits for a specific day, as follows:

$$\max_{P_{jt}} (P_{jt} - c_{jt}) Pr_{jt} \implies P_{jt} = c_{jt} - \frac{Pr_{jt}}{dPr_{jt}/dP_{jt}} \tag{3}$$

where P_{jt} is the price of listing j at time t, Pr_{jt} is the probability that the listing is booked, and c_{jt} is the opportunity cost of renting. This cost may include the opportunity cost of not using the unit or the risk of renting a room to a stranger. It varies with both observable and un-observable listing characteristics and is estimated as the residual of the difference between the price and the producer surplus. The opportunity costs are not observable. One can, however, estimate the probability of booking from equation (2) and recover the marginal costs and producer surplus through their relationship with prices shown in equation (3).

As is shown in equation (3), hosts set their prices to their opportunity costs plus a markup, $PS_j = -\frac{Pr_{jt}}{dPr_{jt}/dP_{jt}}$. This markup is a host's surplus from renting out her property at a given time. To estimate the property owners' surplus, I first find the probability that a listings is booked for a specific day as follows:

$$Pr_{jt} = 1 - \prod_{i \in N_t} (1 - s^i_{jt}) \tag{4}$$

The probability that listing j is booked for a specific day equals one minus the probability that nobody selects the listing. s^i_{jt} is the probability that the potential guest i selects listing j on day t and is derived from the demand-side estimation of mean utilities. Using equation (3) and (4), producer surplus is calculated as follows:

$$PS_{jt} = \frac{1 - \prod_{i \in N_t}(1 - s^i_{jt})}{\alpha \sum_{i \in N_t} s^i_{jt} \prod_{i \in N_t}(1 - s^i_{jt})} \tag{5}$$

where N_t is the market size or set of all potential guests in the market at time t. The first term in equation (5) shows the relationship between the probability that listing j is booked and the probabilities that each potential guest chooses that listing. The product term is the probability that nobody in the market chooses listing j. The second term shows the relationship between the producer surplus and the probabilities that each listing is chosen, which I estimate from the demand side of the model.

4 Estimation Methodology

I estimate consumer and producer surplus in two steps. In the first step I estimate the mean utilities for each listing and calculate the probability that each potential guest chooses each listing. In the second step I calculate the probability that each listing is booked, the producer surplus, and the opportunity costs.

I follow Bayer et al. (2007) in the first-step estimation of demand parameters. There are, however, four important differences between mine and their estimation method. First, I add a time dimension to the model. In their paper they estimate a cross-sectional model of demand for housing options, but in Airbnb the market is highly dynamic and varies every day. So, I use information on repeated rents of each property and estimate property-level demand. Second, I allow cases in which guests choose outside option in addition to possibility of having vacant units. In their paper, everybody who wants to buy a house will buy a house, the market clears, and every seller sells her house. In the Airbnb market, though, a big share of the listings are not booked on a given day and many potential guests choose hotels or other option besides Airbnb. So, it is important to account for the possibility that a guest chooses an outside option in addition to having excess supply in the market.

The third difference here is that I account for variability in choice sets based on guests' decision times (booking times). Guests choose the best option that maximizes their utility at the time of booking. If a listing is already booked it does not show up in the guests' choice set. This is an important feature in many P2P markets where suppliers capacities are limited. Fourth, to estimate the distribution of welfare among property owners, I add a supply side factor and include for strategic pricing over time, which is absent from static model of Bayer et al. (2007). All these extensions are essential in the Airbnb market where there are wide dynamics in market size, choice sets, and pricing.

Next, I estimate the utility parameters by maximizing the likelihood of observed choices. The maximum likelihood function is as follows:

$$ll = \sum_t \sum_i \sum_h I^i_{ht} \log(s^i_{ht}) + \sum_t \sum_i I^i_{ot} \log(s^i_{ot}) \tag{6}$$

where I^i_{ht} is one if a potential guest i chooses listing h; I^i_{ot} is one if she chooses the outside option and zero otherwise. Here, s^i_{ht}, and s^i_{ot} are the probabilities that guest i chooses listing h and the outside option, respectively. The first step is to find the utility parameters $(\alpha, \xi_h, \gamma_t)$ that maximize the log-likelihood function (ll). Here, I use the first-order conditions to search over ξ_h and γ_t, and apply an instrumental variable for price to find α for a given ξ_h, γ_t.

The first-order derivatives of the log-likelihood provide intuitive conditions for pinning down the utility parameters:

$$\begin{aligned}
\frac{\partial ll}{\partial \xi_h} = 0 &\implies N_h = \sum_t \sum_i s^i_{ht} \\
\frac{\partial ll}{\partial \gamma^t} = 0 &\implies N_{ot} = \sum_i s^i_{ot}
\end{aligned} \tag{7}$$

where N_h and N_{ot} are the total number of times that listing h is booked and the total number of potential guests who choose an outside option at time t, respectively. The first condition shows that sum of the probabilities of booking a listing whenever the listing is available equals the total number of times that the listing is booked. This condition pins down the property-specific mean utility (ξ_h). If a listing experiences more bookings over its lifetime the model predicts higher utility for that listing. The second condition concludes that the sum of probabilities of choosing the outside option over all potential guests on a given day equals the total number of potential guests who choose the outside option. If a larger share of potential guests choose the outside option, the model predicts lower time-specific utility (γ_t) for all listings on that day.

I use these two first order maximization conditions to form the following contraction mapping:

$$\delta^{T+1}_{ht} = \delta^T_{ht} + \left(log(N_h) - log(\sum_t \sum_i \hat{s}^i_{ht}) \right) - \left(log(N_{ot}) - log(\sum_i \hat{s}^i_{ot}) \right). \tag{8}$$

Equation (8) is a contraction mapping and converges on a single solution. The fixed point is where both terms in parentheses are zero. This is where both the first-order maximization conditions in equation (7) are satisfied. For a given α, this contraction mapping converges on a unique property and time mean utility fixed effects (ξ_h, γ_t). To estimate α, I apply an IV for price and estimate α for given property and time fixed effects. Then, for a given α, I estimate ξ_h and γ_t using contraction mapping and then re-estimate α. This process is repeated until it converges on a solution for all utility parameters $(\alpha, \xi_h, \gamma_t)$.

4.1 Instrument for Price

An instrumental variable regression approach is used to identify the slope of demand and, in turn, consumer and producer surpluses. I use two instruments for price. The results are very similar using either of these instruments.

Estimating the coefficient of price requires using an instrument to avoid the possible correlation between the price and unobserved property characteristics such as exact locations. This correlation may bias the calculation toward a zero estimation of α. Because an un-observable characteristic is most likely correlated with price and utility in the same direction. For instance, a nice apartment in a good location provides higher utility, and heavier demand leads the supplier to increase prices. This potential positive correlation between utility and price yields a less negative estimation of α.

The first instrument that I use is the BLP instrument (Berry et al., 1995). This is a common instrument in the literature and is used in multiple similar studies (e.g. Bayer et al. (2007), Nevo (2001)). The instrument represents the average characteristics of competitors located far from a given listing. The intuition is that the characteristics of other listings do not directly affect guests' utility. The identification assumption is therefore that, after controlling for characteristics of other listings inside the neighborhood, the characteristics of far-away properties affect utility only through prices. Equation (1) confirms this assumption and shows that the characteristics of rival options does not appear in the guests' utility function.

The relevance of IV is through the relationship between price and utility derived from all options in the market. Hosts set their prices to reflect demand for their properties. Demand for each listing is a function of the utility derived from all options in the market. Thus, it is a function of other listings' characteristics. Equation (5) shows the relationship between price and utility derived from other options in the market and confirms the relevance condition of the BLP instrument.

The second instrument that I use to check the robustness of the results is the share of days for which a listing has been blocked. In the data, a listing is either available, booked, or blocked. Hosts block their properties if they are not willing to rent them out on a specific day. This decision depends on the personal schedule of the host. The identification assumption is that such a personal schedule is not correlated with unobservable demand shocks. The intuition is that if a host expects a drop-in-demand shock, she sets a minimum price for her listing and does not respond to this shock by blocking her property. The identification assumption can be rejected if the host changes her schedule in response to a demand shock. For instance, she might postpone blocking and using her place for personal purposes when demand is high. In the next section, I show that the results derived from both instruments are very similar. Thus, I believe that personal schedule is often fixed and hosts' blocking decisions are not correlated with demand shocks.

4.2 Definition of the Market

The information that first requested on the Airbnb website include the destination, the time of travel, and the number of guests. I follow the same rule and limit the choice sets to all available listings for a day with specific number of guests in Chicago. I assume that renters choose among listings are are available on the same day. This assumption ignores the possibility of choosing among listings that are available over several days. Because I model the decisions in a static setting, every day represents a separate market.

I also assume that renters search for listings with specific capacities for number of guests. I divide the listings in five categories comprising 1, 2, 3-4, 5-6, and more than 6 as a maximum number of guests. So, a property with capacity for one guest is not in the choice set of a person who is looking for a place for two or more guests. Similarly, properties with higher capacities are not in the choice sets of those who are looking for places for fewer maximum number of guests. Therefore, in my estimations potential guests choose among all available listings in a specific day with a specific capacity.

4.3 Market Size

The demand estimation requires information about the number of potential renters in the market (N_t). The discrete choice literature usually relies on an assumption about the size of a market. For instance, Berry et al. (1995) assume that every household in the U.S. is a potential buyer in the car market. In the context of the housing market, Bayer et al. (2007) assume that there is no excess demand and that market size equals total number of houses sold in a year. About the Airbnb market, Farronato and Fradkin (2017) assume that market size is three times the total number of bookings in the corresponding month the previous year. The assumption about market size affects demand estimation and the probability that each listing is booked. As is shown in Figure 4 and discussed in the summary statistics section, the Airbnb market experienced exponential growth and seasonal variations in the total transactions. A proper choice of market size should enable the analysis to reveal these variations in the market.

Ideally, I need information about all potential guests who search on the Airbnb website for booking on a specific day as a measure of market size. In the absence of this information, I use Google search trends for "Airbnb Chicago" as a proxy for the total number of potential guests. Google reports a monthly index (between 0 to 100) of total searches for a keyword. I scale this index to find an average booking rate of 60% in the first three months of the data. This scaling coefficient comes from Fradkin (2017). Using sample data about search activities of users in an Airbnb market from September 2013 through September 2014, he shows that around 40% of users who send inquiries through the Airbnb website do not book a listing and instead choose an outside option. I use this result to scale the index for Google search trends. Figure 5 shows the trends in calculated numbers of potential guests and total numbers of booked listings. It confirms that the scaled index captures growth and seasonal variations in the market.

One complication is that using the scaled index of Google trends provides a monthly and city-level sizes of the market. As discussed in the section 4.2, however, my estimation requires information about potential guests on a given day who are looking for listings with specific capacities. To find the day-capacity-level size of the market, I interpolate the scaled measure of overall market size from Google trends. To do so, I multiply the scaled index by the share of booked listings with specific capacities times the share of booked listings for each day as is shown in Appendix 8. This interpolation provides the desired level of market size.

5 Results

In this section, I begin with reporting the results of the mean utility estimations. I then report the results of instrumental variable regressions and estimates of marginal utility for the various characteristics of a listing. I then estimate the consumer and producer surpluses, the costs of renting, and their distributions in the market.

5.1 Parameter Estimates

In Table 1 I report the estimation results for property- and time-specific mean utilities and the overall estimated mean utility in dollar terms. The mean utility estimates are all negative, with an average of $-294. The negative sign is because of the large share of the outside option in the market.

Next, I find the willingness to pay for each characteristic of a listing by conducting an IV regression of mean utilities on price and the listing's characteristics. Table 2 shows the estimates of the mean price coefficient (α) using the BLP-IV regressions. The coefficient of an OLS regression is -0.014. This coefficient is biased toward zero because hosts set higher prices for listings with higher utilities. The mean price coefficient in the IV regression is -0.026 and is higher in absolute value than the OLS estimate. The instrument is relatively strong and yields first-stage F-statistics (excluding the instrument) of around 14. The estimated mean price coefficient is consistent with the estimates of Farronato and Fradkin (2017) for the Airbnb market. The IV-regression estimations obtained with BLP-IV are robust to including listing characteristics in the regression. BLP-IV is a widely used IV in the literature but the identification assumption may be violated if listings' characteristics are correlated with demand shocks. I apply another IV for price based on the share of days on which a listing is blocked to check the robustness of the results. The second IV results are shown in table 3. The second IV is stronger than BLP-IV, with first-stage F-statistics (excluding the instrument) of 86. The estimates are very similar to the estimations of the BLP-IV regressions. Although, one can imagine scenarios in which it make sense to reject the identification assumption for both IVs, as discussed in section 4.1, finding consistent results with two distinct IVs is promising.

Table 4 shows the full regression coefficients. In this table each coefficient represents the willingness to pay for a specific characteristic of a property. To find the willingness to pay in dollar values, I divide the coefficients by the coefficient of price. Using the Blocked days-IV regression, the willingness to pay for an extra bedroom and bathroom are $32 and $29, respectively. Similarly, Figure 6 shows the willingness to pay across Chicago neighborhoods. An average guest is willing to pay up to $130 for the same property in Central Chicago compared with what she is willing to pay in some neighborhoods in the southern part of the city. This figure illustrates the importance of location in Airbnb booking decisions. These results are intuitive because most touristic attractions and business activities in Chicago are concentrated in the Downtown areas. This result shows that renters are willing to pay less to stay in properties located in the southern parts of Chicago, where the average income is lower than in Downtown. This result suggests the presence of a demand-side source of unequal distribution of welfare across neighborhoods in the market. I study the distribution of welfare in greater detail in the next section.

5.2 Welfare Estimates

In this section I estimate consumer and producer surpluses in terms of compensating variations. I therefore ask by how much a guest should be compensated to maintain the same level of utility if she is banned from booking an Airbnb listing in Chicago.

The compensating variation in dollars for consumers is:

$$CS^i = \frac{1}{\alpha} \log \left(1 + \sum_{j \in M_t^i} \exp(\delta_j) \right). \tag{9}$$

Equation (9) expresses the difference between the expected maximum utility of guest i in the market and her utility when she uses only the outside option (Train, 2009). This difference is a measure of consumer surplus in terms of compensating variation. In the above equation, δ_j is the mean utility for listing j. Because the only source of heterogeneity on the consumer side is guest-specific utility shock (ϵ_j^i), mean utility is assumed to be the same across all guests in the market. As discussed in section 4, my estimation method is easily applicable to a data set with information about guests. This information could potentially provide heterogeneities across mean utilities for which I do not account in the model. It should be noted, however, that M_j^i varies across guests depending on when renters make booking decision.

Similarly, on the producer side, I ask by how much a host in Chicago should be compensated to maintain the same level of profit if she is banned from renting out her property through Airbnb. I use equation (5) to estimate the producer surplus for each host in the market.

I estimate the producer surplus while assuming that agents are rational and apply strategic pricing with full information about demand. In practice, some agents do not respond to variations in demand and do not change their prices as expected. This is a potential problem in most structural estimations. I believe, thought, following the rationality assumption is still reasonable to rely on average estimates of surpluses that derive from the rationality assumption. Studying agents' pricing decisions and how they are different from the pricing of rational agents with full information, and whether there is learning in the market, are interesting topics for future researches.

The rationality assumption affects my estimations of the producer surplus. In my estimations, around 10% of producer surplus estimates are higher than price. This means that around 10% of the estimates of opportunity costs are negative. To address this estimation problem, I truncate the marginal cost estimates that are less than zero to zero. The overall patterns and estimations are not affected whether I winsorize the estimates or drop the negative opportunity costs from the data.

5.2.1 Time Variations

The total estimated consumer surplus in the time span of my data (around 2.5 years) is \$21 million. This is equivalent to an average consumer surplus of \$15 per day or 20% of each dollar spent. The producer surplus is three times higher than the consumer surplus at a total of \$67 million over 2.5 years. The average producer surplus is \$46 per day or 65% of each dollar spent. One potential explanation of a higher share of producer surplus is that consumers have easy access to an outside option such as a hotel. On the other hand, the outside option for hosts is not using their property for personal use. They cannot gain from their extra capacities if they do not rent out their properties on Airbnb.

This is a potential explanation of the relatively larger producer surplus. It should be noted that hosts face the cost of renting only when someone books their property. So, I assume that properties in the Airbnb market are absent from long-term rental market and I do not consider the opportunity cost of long-term rentals in my estimations. A reasonable extension to my estimations that is a work in progress is adding a fixed cost to the model that captures the long-term opportunity cost of participating in the Airbnb market.

Figure 7 shows the trends in estimated consumer and producer surpluses as percentages of price. They both follow seasonal variations with larger shares taken by surplus over the summer and periods of high demand. Even though the cost of entry in P2P markets is much lower than it is in traditional markets, both seasonal and slightly upward trends in the producer surplus suggest that the supply side does not perfectly adjust to demand. These results suggest that entry is not cost-less, and that, in contrast to a perfect competition market, hosts gain more market power when demand is high.

5.2.2 Geographic Variations

An important question is how welfare is distributed across neighborhoods and whether welfare is dis-proportionally distributed across areas. The answer to this question provides insight into who benefits more from the short-term Airbnb rental market. Figure 8 illustrates the distribution of the total producer surplus across Chicago neighborhoods. This map shows the extend to which each neighborhood has benefited from Airbnb over approximately 2.5 years of the sample period. It is apparent that the total surplus is disproportionately distributed across neighborhoods. Some neighborhoods, mostly near the downtown and northern parts of Chicago, generate a surplus of more than $2 million while there are neighborhoods located mostly in the southern parts of Chicago that generate a surplus of less than $50 thousand. Considering this disproportional distribution of welfare is important for any regulation of the Airbnb market.

This distribution of the surplus in part reflects different numbers of bookings in different neighborhoods. There is less demand for booking an Airbnb in less touristy areas or in neighborhoods located further from main attractions, and there are fewer active listings in these areas. Figure 10 shows the geographic distribution of the average surplus per booking. Again, properties in the downtown and northern parts of Chicago earn more surplus through Airbnb per booking. This is mainly due to higher willingness to pay for staying in these locations.

Figure 9 shows the distribution of estimated hosts' opportunity costs of renting out their properties. Opportunity cost measures the risk of renting, the cost of not reserving for personal use, and any potential loss incurred by renting out a property. It also includes the value of money for the host. The more highly hosts value the money, the lower is the estimated opportunity cost of renting. The results show that property owners located in the southern part of Chicago value more the gains from Airbnb and have lower opportunity costs of renting out their properties.

Comparing the map of opportunity costs in Figure 9 with the map of the total surplus in Figure 8, one sees that, even though the opportunity cost of providing accommodation services is lower in south Chicago, the surplus is concentrated in the downtown and northern areas. This results is mainly from higher demand for more touristy and business-focused areas, which shifts away the potential surplus from disadvantaged neighborhoods to more expensive areas.

In Table 5 I report the distribution of the surplus in nine "Sides" of Chicago. I show how much each Side would have lost if Airbnb were to be banned from operating in the city. The North Side, followed by the West and Central Sides of Chicago have benefited the most from access to Airbnb. In the third and fourth columns in this table I report the shares of surplus and population for each Side. As can be seen, the share of the producer surplus is not distributed comparable to the share of the population. For instance, Central Chicago earned roughly 30% of the total surplus but it includes only 1% of the total population.

In addition to the concentration of total welfare in upper-class neighborhoods, properties in these areas earn more surplus per booking. This shows that the greater willingness to pay for accommodations in these neighborhoods dominates the higher opportunity cost of properties in these areas. Figure 10 shows the distribution of average producer surplus per booking and indicated that properties in Central Chicago benefit from greater market power. Figure 11 illustrates correlations between the median income and the average surplus per booking across neighborhoods. This figure suggests a strong positive correlation between neighborhood income and producer surplus per booking. In fact, a 1% increase in median income is associated with around a 0.2% higher producer surplus per booking.

5.2.3 Property-level Variations

In this section I consider the distribution of surpluses across properties in the market. The results reported in the last section show the concentration of welfare in more advantaged neighborhoods. This distribution is mostly demand-driven and benefits areas where people are willing to pay more to access the existing amenities. It is, however, important to determine the distribution of welfare across property owners who rent out their properties in the Airbnb market to obtain a more accurate view of the distribution of welfare in the market.

Property owners who live in the same neighborhood may nevertheless have different characteristics and valuations for a marginal dollar. I show that, even though the surplus is concentrated in advantaged neighborhoods, on average, property owners who value money more highly earn a higher surplus than others in each neighborhood. This result suggests that Airbnb is more beneficial to property owners with lower incomes and those who value a marginal dollar more highly. In the absence of detailed information about hosts and their income in my data, I use my estimations of opportunity costs as a measure of property owners' valuation an extra dollar. I explain the details in the following.

My estimations allow for an indirect estimation of the opportunity cost of renting out properties. I estimate opportunity costs as the residual of the difference between price and the estimated producer surplus, as is shown in equation (3). The following equation is a modification of equation (3), where hosts have different valuations of money:

$$\max_{P_{jt}}(\tau_j P_{jt} - c_{jt})Pr_{jt} \implies P_{jt} = \frac{1}{\tau_j}c_{jt} - \frac{Pr_{jt}}{dPr_{jt}/dP_{jt}}, \tag{10}$$

where τ_j is how much host j values an extra dollar. This parameter may vary with property owner's income or her implicit wage/time costs. The varying valuations for money show up in my estimation of opportunity costs. In fact, one should expect lower estimates of opportunity costs in low-income neighborhoods and in locations where people value the extra dollar more highly (as reflected in higher τ_j). Figure 12 confirms the positive correlation of estimated opportunity costs with median neighborhood income. This correlation partly explains the correlation between the opportunity cost of renting out a unit and the rental price of the unit in the long-term rental market. Figure 13 illustrates the strong positive correlation between opportunity costs and average rents across neighborhoods.

To determine the relationship between the property owners' surplus and the marginal value of a dollar, I regress the surplus on my estimation of the opportunity cost for each property and control for neighborhood fixed effects. Controlling for neighborhood fixed effects captures the average geographic differences in opportunity costs that reflects different rental values. The first two columns of Table 6 show the negative correlation of the producer surplus with opportunity costs. The results indicate that a 1% decrease in opportunity costs is associated with 0.17% increase in property owners' total surplus. This result suggests that property owners who value money more highly benefit to greater extent from having access to the Airbnb market.

This result is more significant among the subset of owners who are located in high-demand areas. Column (3) in the table 6 shows the correlations between property owners' surpluses and their average opportunity costs in the Central Side of Chicago. In this location, a 1% decrease in the opportunity cost is associated with a 0.31% increase in property owners' surpluses. In Column (4) of this I use share of days on which a listing was blocked as a measure of the importance of extra dollars to property owners. As is the case with owners who incur lower opportunity costs, those who block their listings more often earn, on average, lower total surplus. These findings suggest that having access to the Airbnb market is more beneficial to property owners who value a marginal dollar more highly (low-income owners) and the Airbnb market is especially beneficial to the low-income owners who live in high-demand areas.

6 Summary

In this paper I introduce a framework for studying welfare distribution in P2P markets in which agents trade highly differentiated goods. In these markets, accounting for heterogeneity over time, location, and certain characteristics is important to ensure accurate welfare analysis. These factors play key roles in shaping consumer and producer surpluses. Among them are, for example, the times and locations at which services are booked in ride-sharing markets; the risk associated with and amounts of loans in online funding markets; and the times, locations, and characteristics of listings in home-sharing markets.

This paper utilizes micro-level data on Airbnb rentals along with an individual-level multinomial logit model to account for heterogeneities across agents. The findings show, on average, a consumer surplus of 20%, and a producer surplus of 65% for each dollar spent in the Chicago Airbnb market. These surpluses are not distributed homogeneously across various groups of listings or in distinct times and locations.

The estimated average producer and consumer surpluses follow seasonal variations in prices. The producer surpluses are disproportionately concentrated in high-income and upper-class neighborhoods. This concentration is mostly demand-driven and is attributed to higher willingness to pay for areas with more attractive amenities for potential renters. My results show that renters are willing to pay up to $130 more for accommodations in downtown areas than they will to pay to stay in some low-demand neighborhoods, mostly in the southern parts of Chicago. Even though surplus is concentrated in upper-class neighborhoods, I show evidence of higher surpluses for property owners who value the marginal dollar more highly. My results suggest that there are high welfare gains from having access to the Airbnb market for low-income owners, especially for those who live in areas with higher demand for short-term accommodations.

My estimation methodology is applicable to other highly dynamic P2P markets in which agents with limited capacities trading in a decentralized market. My study also enhances the existing methodology for demand estimation in the housing market by enabling econometricians to use repeated sales and panel data and accounts for seasonal variations in demand, possible vacancies, and variable choice sets of buyers. These are important features of housing sales or rental markets that my estimation methodology takes into account.

References

AirDNA (2017), 'Short-term rental data and analytics — airbnb and vrbo'.
URL: *https://www.airdna.co/*

Barron, K., Kung, E. and Proserpio, D. (2018), The sharing economy and housing affordability: Evidence from airbnb, *in* 'Proceedings of the 2018 ACM Conference on Economics and Computation', ACM, pp. 5–5.

Bayer, P., Ferreira, F. and McMillan, R. (2007), 'A unified framework for measuring preferences for schools and neighborhoods', *Journal of political economy* **115**(4), 588–638.

Berry, S., Levinsohn, J. and Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Cohen, P., Hahn, R., Hall, J., Levitt, S. and Metcalfe, R. (2016), Using big data to estimate consumer surplus: The case of uber, Technical report, National Bureau of Economic Research.

Einav, L., Farronato, C. and Levin, J. (2016), 'Peer-to-peer markets', *Annual Review of Economics* **8**, 615–635.

Farronato, C. and Fradkin, A. (2017), The welfare effects of peer entry in the accommodation market: The case of airbnb, Technical report, Technical report, Working paper.

Fradkin, A. (2017), 'Search, matching, and the role of digital marketplace design in enabling trade: Evidence from airbnb', *Matching, and the Role of Digital Marketplace Design in Enabling Trade: Evidence from Airbnb (March 21, 2017)* .

Fraiberger, S. P., Sundararajan, A. et al. (2015), 'Peer-to-peer rental markets in the sharing economy', *NYU Stern School of Business research paper* **6**.

Lam, C. and Liu, M. (2018), 'More than taxis with an app: How ride-hailing platforms promote market efficiency'.

McFadden, D. (1977), Modelling the choice of residential location, Technical report, Cowles Foundation for Research in Economics, Yale University.

Nevo, A. (2001), 'Measuring market power in the ready-to-eat cereal industry', *Econometrica* **69**(2), 307–342.

Sheppard, S., Udell, A. et al. (2016), 'Do airbnb properties affect house prices', *Williams College Department of Economics Working Papers* **3**.

Sundararajan, A. (2016), *The sharing economy: The end of employment and the rise of crowd-based capitalism*, Mit Press.

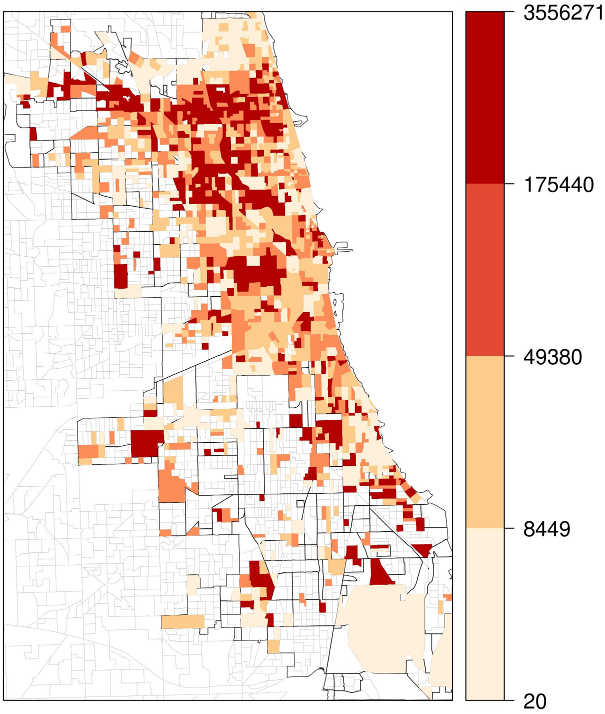
Train, K. E. (2009), *Discrete choice methods with simulation*, Cambridge university press.

Zervas, G., Proserpio, D. and Byers, J. W. (2017), 'The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry', *Journal of marketing research* **54**(5), 687–705.

7 Figures and Tables

7.1 Figures

Figure 1. Block-Group Level Distribution of Total Revenue from Aug 2014 to Apr 2017



Note: Block-groups are defines in city of Chicago data portal.

Figure 2. Total Revenue Over Time

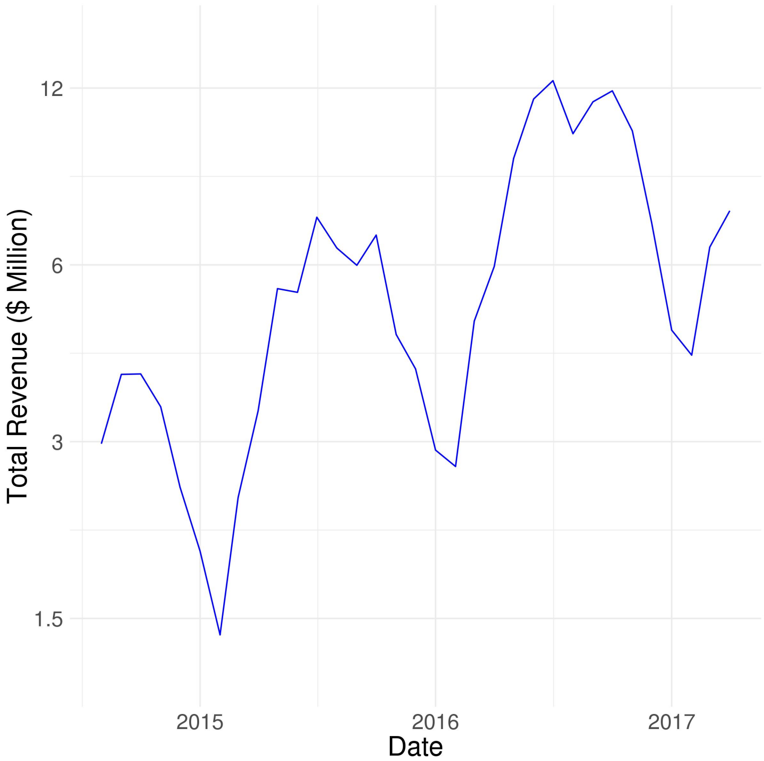


Figure 3. Total Daily Revenue in Apr-2017

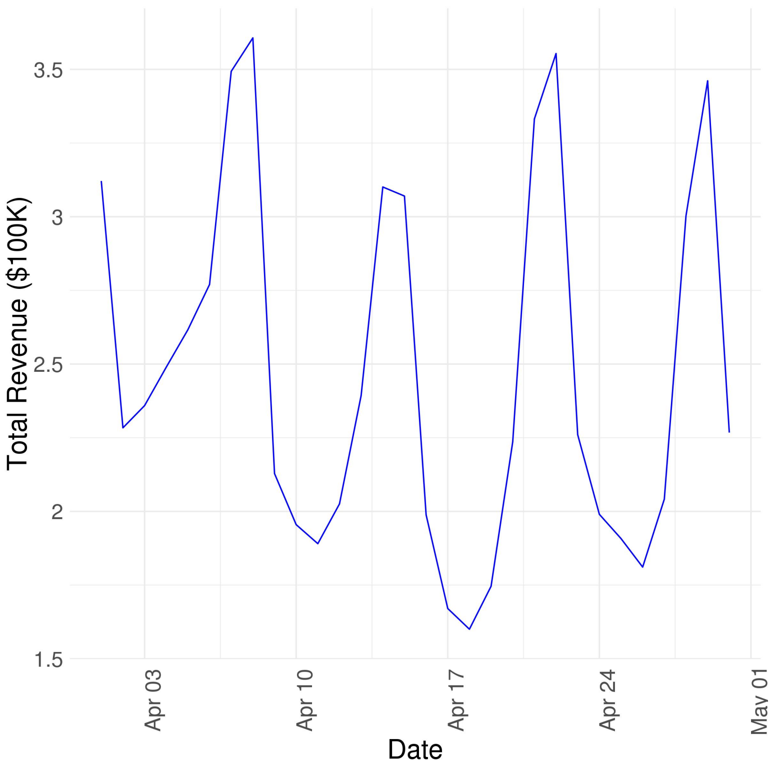


Figure 4. Average Price of All (Booked or Available) and Booked Listings Over Time

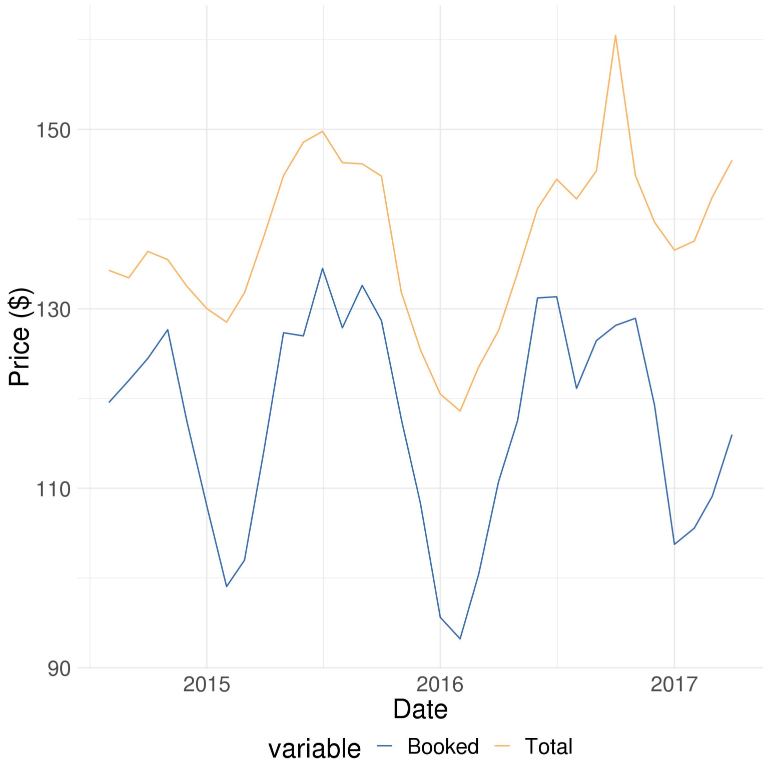


Figure 5. Monthly market size and total number of bookings

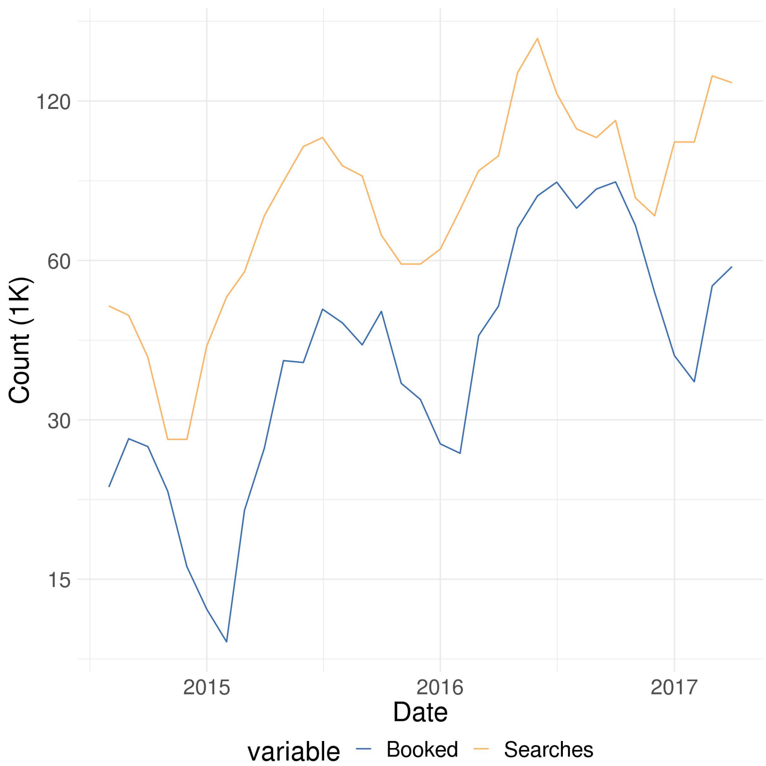
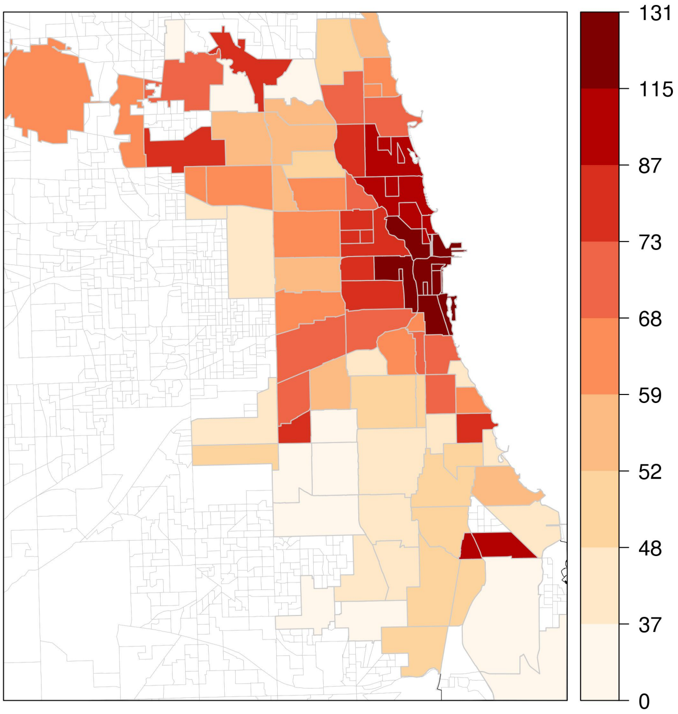


Figure 6. Average Consumer Willingness to Pay for Chicago Neighborhoods in Dollars



Note: This map plots neighborhood fixed effects in the mean utility regressions. The fixed effects are adjusted with coefficient of price and are presented in dollar values. The borders are community neighborhoods as defined in City of Chicago data portal.

Figure 7. Average Consumer and Producer Surplus Per Dollar Spent

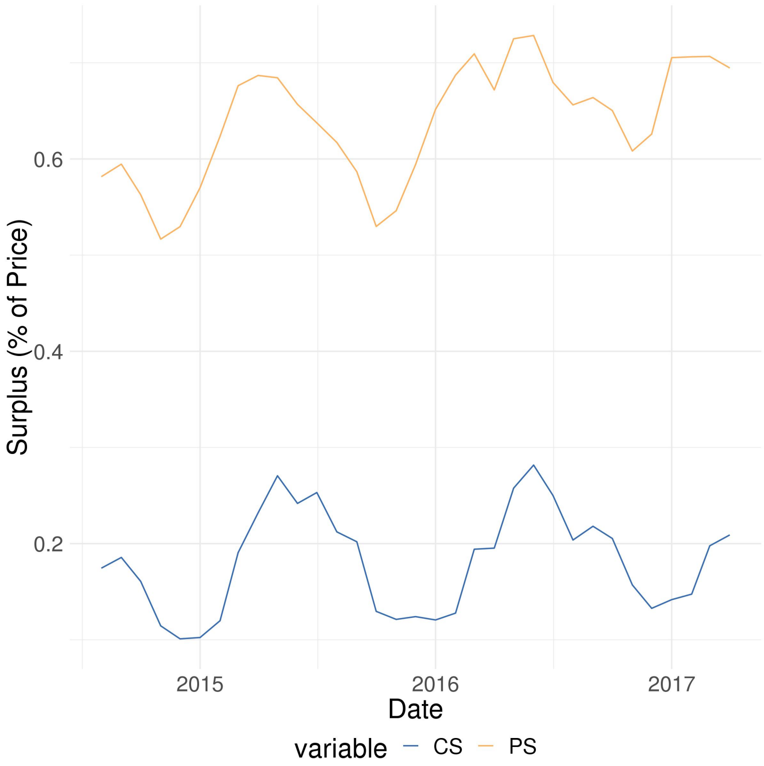
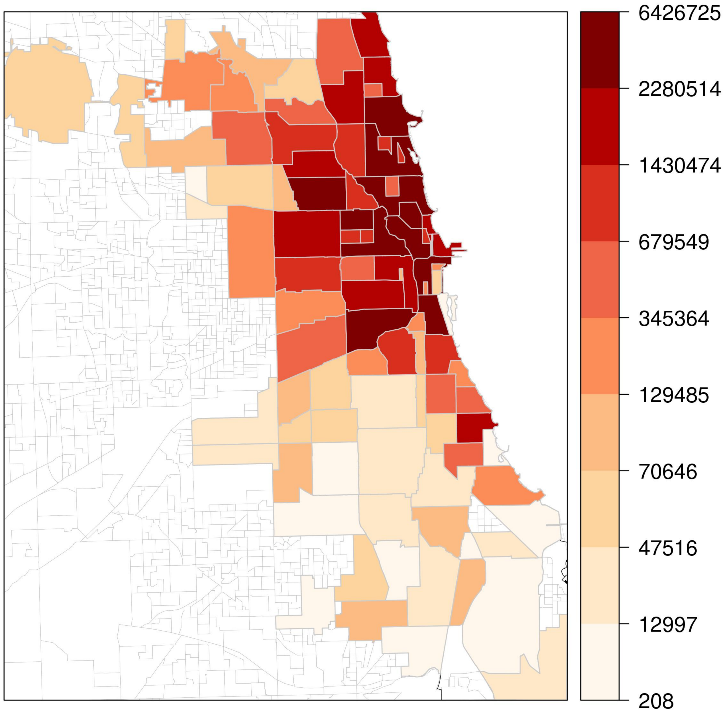
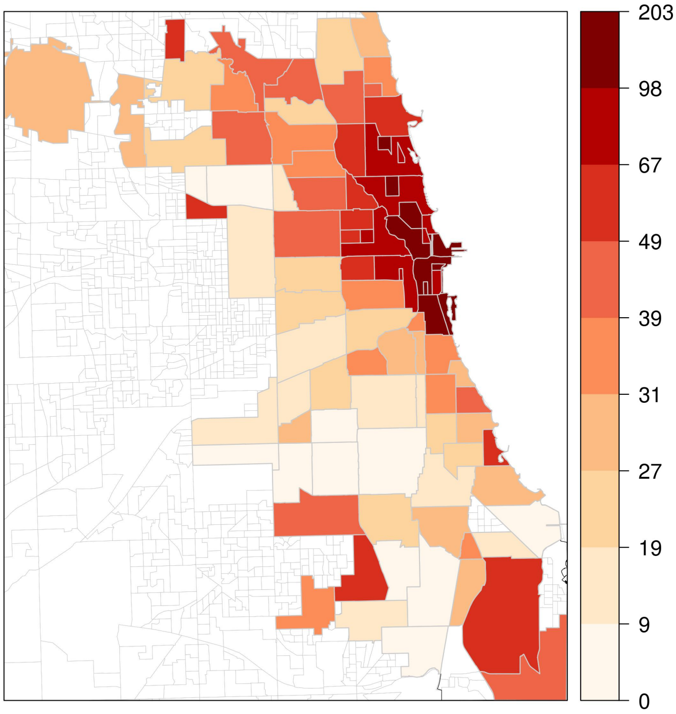


Figure 8. Geographic Distribution of Total Producer Surplus



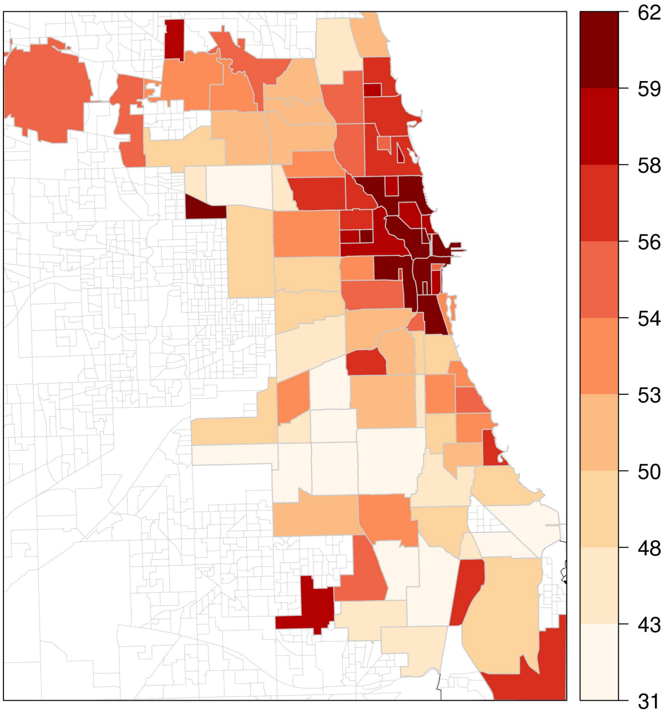
Note: The borders are community neighborhoods as defined in City of Chicago data portal.

Figure 9. Geographic Distribution of Average Opportunity Cost Per Property Per Day



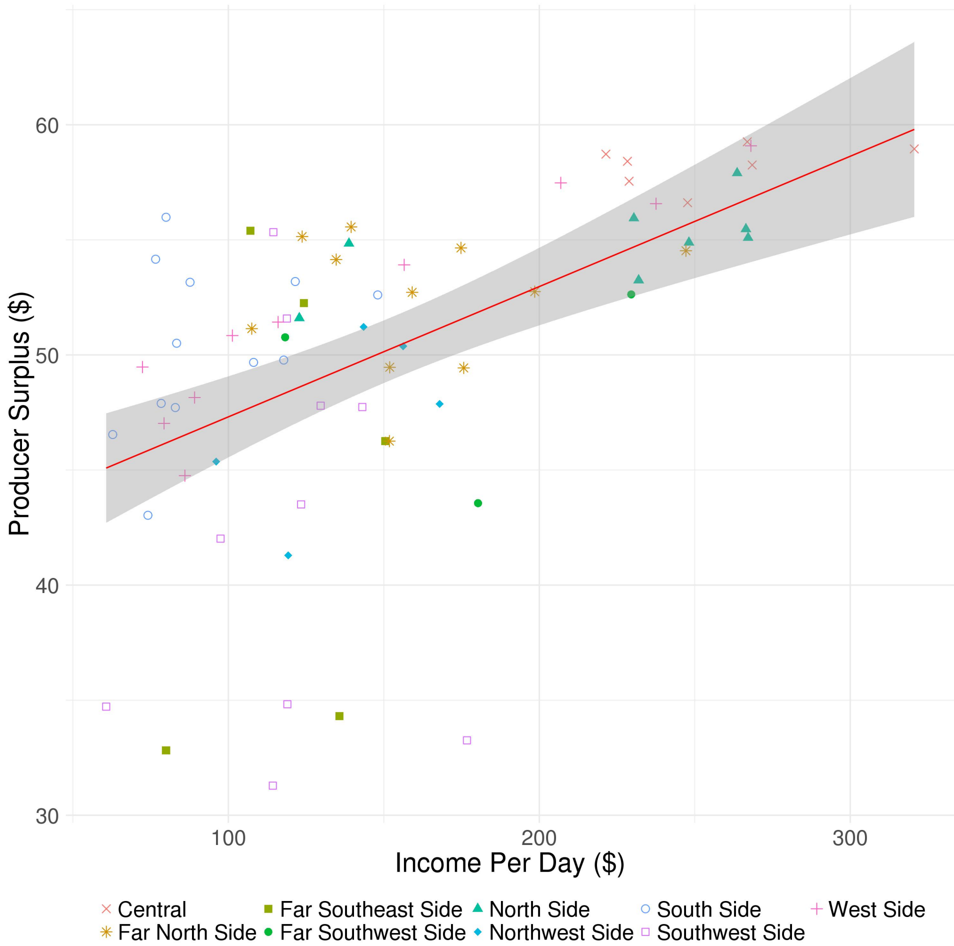
Note: The borders are community neighborhoods as defined in City of Chicago data portal.

Figure 10. Geographic Distribution of the Average Producer Surplus Per Property Per Day



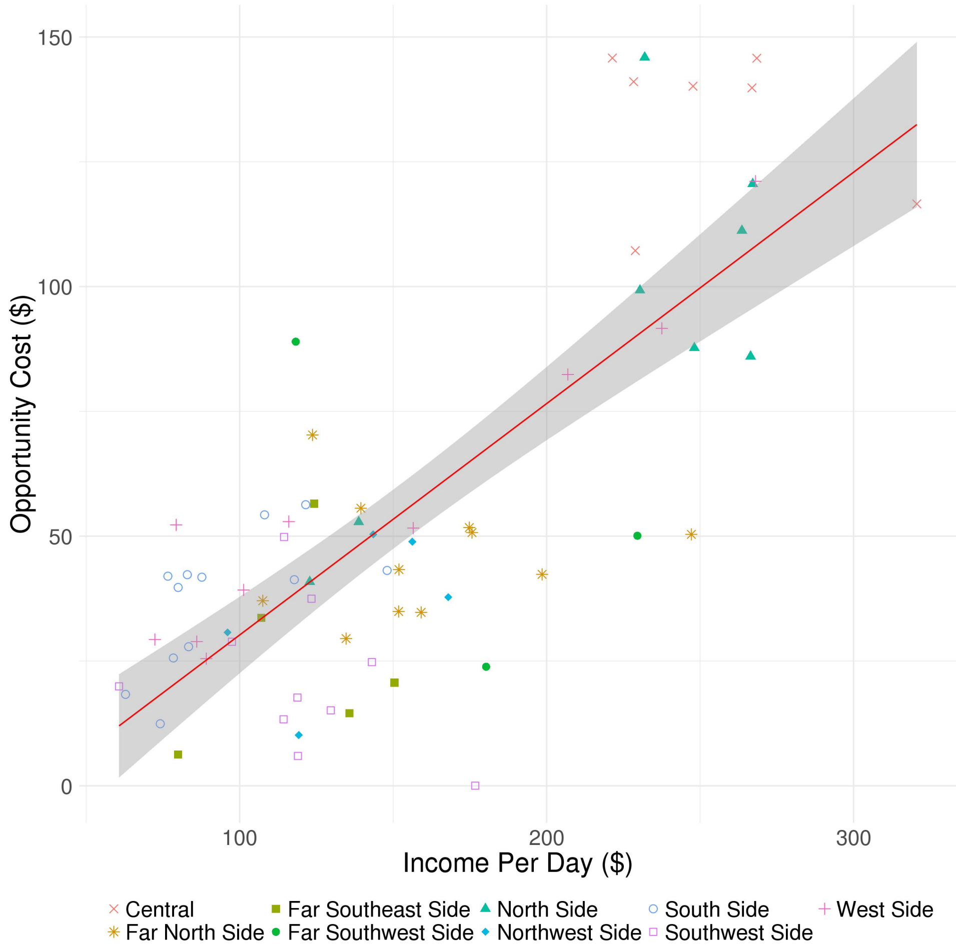
Note: The borders are community neighborhoods as defined in City of Chicago data portal.

Figure 11. Correlations between average income and average producer surplus per booking across neighborhoods



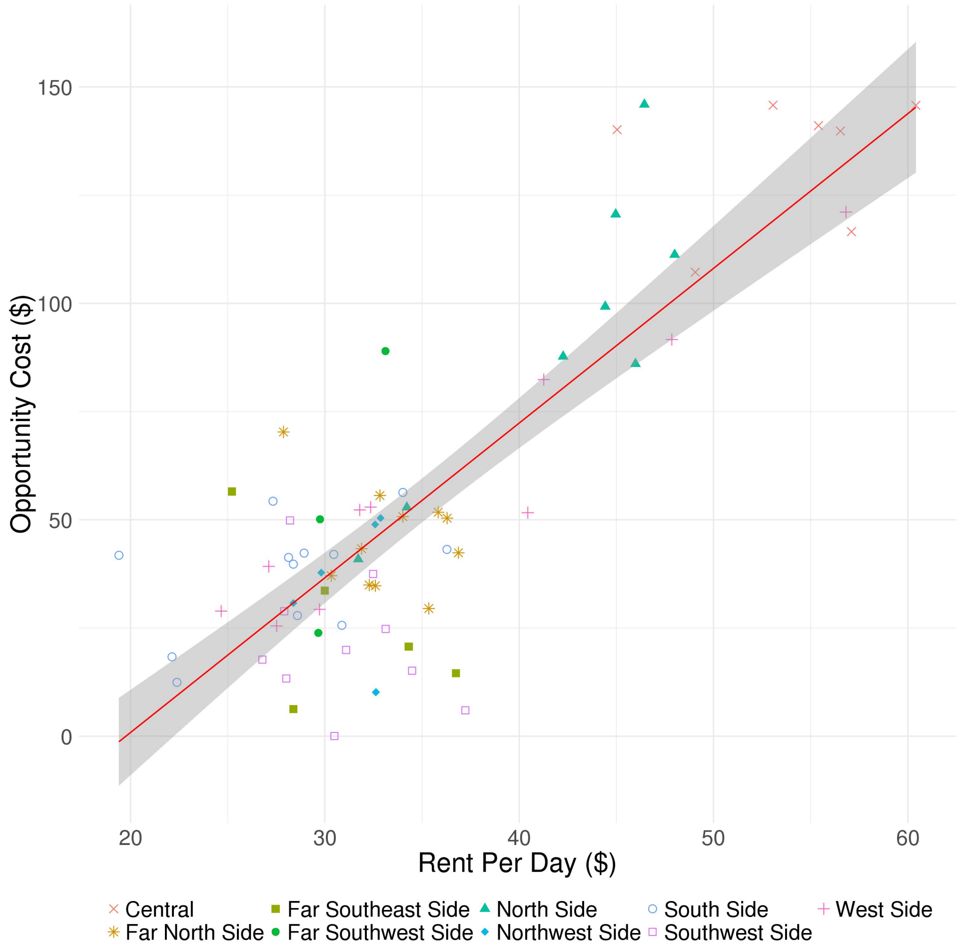
Note: Neighborhoods and Sides of Chicago are defined on the City of Chicago data portal.

Figure 12. Correlations between average income per day and the average opportunity cost of renting out listings on a given day across neighborhoods



Note: Neighborhoods and Sides of Chicago are defined on the City of Chicago data portal.

Figure 13. Correlations between average rent per day and the opportunity cost of renting out on a given day across neighborhoods



Note: Neighborhoods and Sides of Chicago are defined on the City of Chicago data portal.

7.2 Tables

Table 1. Distribution of Mean Utility Parameters

Statistic	Pctl(25)	Mean	Pctl(75)	St. Dev.
δ_{ht}/α	−334.004	−294.465	−244.134	78.763
γ_t/α	−156.871	−120.914	−81.263	56.773
ξ_h/α	−82.545	−33.341	−1.887	84.805
Outside Option	0.311	0.457	0.592	0.157

Note: Dollar value estimations of mean utility parameters. The last row shows shares of outside options in the market. The negative mean utility estimations show high shares of outside options in the market.

Table 2. BLP-IV Regression of Mean Utilities

	Delta - Mean Utility Regression			
	(OLS)	(IV)	(IV)	(First Stage)
Price	-0.014^{***}	-0.026^{***}	-0.026^{***}	$--$
	(0.0003)	(0.005)	(0.005)	
Date FE	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	No	Yes	Yes
Other Controls	Yes	No	Yes	Yes
Clustered	Yes	Yes	Yes	Yes
F-Statistics	5672.82	1526.17	2068.35	3973.53
F-Stat(excl instr.)				13.98
Observations	4,233,376	4,233,376	4,233,376	4,233,376

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows the price coefficients of IV-regressions of mean utility on price and listing characteristics. The first column shows the results of an OLS regression. The second and third columns show the result of regressions of BLP-IV. In all regressions I control for the maximum number of guests, the types of listings, and average characteristics of other options in the neighborhood. Regressions are all clustered at the neighborhood level. Other controls include the number of reviews, number of photos, and dummy variables for instant bookings, being super hosts, asking for deposits, and being business-ready

Table 3. Blocked Days-IV Regression of Mean Utilities

	Delta - Mean Utility Regression			
	(OLS)	(IV)	(IV)	(First Stage)
Price	-0.015^{***}	-0.026^{***}	-0.025^{***}	
	(0.0003)	(0.003)	(0.002)	
IV				-27.327^{***}
				(2.932)
Date FE	Yes	Yes	Yes	Yes
Neighborhood FE	Yes	No	Yes	Yes
Other Controls	Yes	No	Yes	Yes
Clustered	Yes	Yes	Yes	Yes
F-Statistics	6221.01	1159	2036.18	4048.96
F-Stat(excl instr.)				86.85
Observations	4,233,376	4,233,376	4,233,376	4,233,376

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows the price coefficient of IV-regressions of mean utility on price and characteristics of listings. The first column shows the results of an OLS regression. The second and third columns show the result of regressions of BLP-IV. In all regressions I control for the maximum number of guests, the types of listings, and average characteristics of other options in the neighborhood. Regressions are all clustered at the neighborhood level. Other controls include the number of reviews, number of photos, and dummy variables for instant bookings, being super hosts, asking for deposits, and being business-ready

Table 4. Full Table - IV Regression of Mean Utilities

	Delta - Mean Utility Regression		
	(OLS)	(Blocked-IV)	(BLP-IV)
Max Guests	0.170***	0.250***	−0.057
	(0.018)	(0.030)	(0.117)
Entire Place	0.682***	1.070***	1.301***
	(0.034)	(0.104)	(0.220)
Bedrooms	0.502***	0.791***	0.676***
	(0.034)	(0.078)	(0.168)
Bathrooms	0.316***	0.716***	0.712***
	(0.080)	(0.131)	(0.211)
Overall Rating	0.339***	0.471***	0.511***
	(0.035)	(0.059)	(0.075)
Number of Photos	0.007***	0.010***	0.012***
	(0.002)	(0.002)	(0.003)
Business Ready	0.012	−0.039	−0.056
	(0.028)	(0.040)	(0.047)
Super-host	0.406***	0.358***	0.357***
	(0.031)	(0.042)	(0.039)
Deposit	0.062*	0.097**	0.083**
	(0.033)	(0.038)	(0.042)
Instant-book	0.316***	0.183***	0.185***
	(0.033)	(0.045)	(0.069)
Price	−0.015***	−0.025***	−0.026***
	(0.0003)	(0.002)	(0.005)
Date FE	Yes	Yes	Yes
Neighborhood FE	Yes	Yes	Yes
Neighborhood Controls	No	No	Yes
Clustered	Yes	Yes	Yes
Observations	4,233,376	4,233,376	4,233,376

Note: *p<0.1; **p<0.05; ***p<0.01

Note: This table shows the price coefficients of IV-regressions of mean utility on price and listing characteristics. In the first column I report the results of an OLS regression. In the second and third columns I report the results of regressions of the share of blocked days and BLP-IV, respectively. All regressions are clustered at the neighborhood level. Neighborhood controls include the average characteristics of rival listings in the neighborhood. Each coefficient shows the mean willingness to pay for that characteristic.

Table 5. Distribution of the Producer Surplus Across Sides of Chicago

Side	Producer Surplus (PS) ($)	Share of PS	Share of Population
North Side	22,472,189	0.276	0.190
West Side	21,858,242	0.268	0.174
Central	18,709,368	0.229	0.096
Far North Side	10,368,850	0.127	0.236
South Side	5,494,908	0.067	0.125
Northwest Side	1,657,773	0.020	0.091
Southwest Side	584,932	0.007	0.057
Far Southeast Side	266,063	0.003	0.013
Far Southwest Side	155,063	0.002	0.018

Note: The first column lists nine Sides Chicago as defined on the City of Chicago Data Portal. The second column shows the total producer surplus from August 2014 through April 2017 in each side. The third column shows the share of the total producer surplus and the fourth column shows share of population in each Side.

Table 6. Property-level Correlations of the Producer Surplus with Opportunity Costs and the Share of Blocked Days

	Logarithm of Total Produce Surplus Per Property			
	(1)	(2)	(3)	(4)
Opportunity Cost	-0.15^{***}	-0.17^{***}	-0.31^{***}	
	(0.02)	(0.02)	(0.05)	
Share of Blocked				-1.30^{***}
				(0.05)
Neighborhood FE	No	Yes	Yes	Yes
Clustered	Yes	Yes	Yes	Yes
Subset	Total	Total	Central	Total
Controls	Yes	Yes	Yes	Yes
Observations	13,481	13,481	3,266	13,481
Adjusted R^2	0.02	0.03	0.02	0.09
Note:		*p<0.1; **p<0.05; ***p<0.01		

Note: In the first through third columns I show the correlations between the total producer surplus per property and the average opportunity cost of each property. In the third column I report the results of a regression on the subset of the Central Side of Chicago. In the fourth column I report the correlations between the producer surplus of each property and the share of days that the property was blocked. In all regressions I control for types of listings, number of bedrooms, and number of bathrooms.

8 Appendix

Estimating Market Size

The size of the market or the number of potential guests on each day who are looking for properties with specific number of guests to rent is estimated as follows:

First, I use Google trend searches for "Airbnb Chicago" for each month. Google reports an index that ranges between 0-100 for numbers of searches. I use this index to capture seasonal variations and overall trends in market size. To construct a measure of market size, I scale the index such that it generates 60% booking rates in the first three months of the sample period as follows:

$$\text{Monthly market size} = \text{Google trends} \times \frac{\text{total reservations in the first three months}}{\text{Google trend searches in the first three months} \times 0.6}.$$

To find the market-level number of potential guests, I interpolate the monthly measure of market size using the share of bookings with a specific number of maximum number of guests and the share of bookings on a given day as follows:

$$\begin{aligned}\text{Daily market size} =& \text{Monthly market size} \times \\ & \text{share of bookings with a specific maximum number of guests in the month} \times \\ & \text{share of bookings on a given day}\end{aligned}$$

Proof of Equation (5)

$$\begin{aligned}
&\text{Producer surplus of listing } h \; = -\frac{Pr_{ht}}{\partial Pr_{ht}/\partial P_h} \\
&Pr_{ht} = 1 - \prod_{i \in N_t} \left(1 - \frac{\exp(\delta_{ht})}{1 + \sum_{j \in M_{it}} \exp(\delta_{ht})}\right) = 1 - \prod_{i \in N_t} (1 - s^i_{ht}) \\
&-\frac{\partial Pr_{ht}}{\partial P_h} = \alpha \sum_{i \in N_t} \frac{\exp(\delta_{ht})}{1 + \sum_{j \in M_t} \exp(\delta_{ht})} \prod_{i \in N_t} \left(1 - \frac{\exp(\delta_{ht})}{1 + \sum_{j \in M_{it}} \exp(\delta_{ht})}\right) \\
&= \alpha \sum_{i \in N_t} s^i_{ht} \prod_{i \in N_t} (1 - s^i_{ht}) \implies \\
&PS_{ht} = \frac{1 - \prod_{i \in N_t} (1 - s^i_{ht})}{\alpha \sum_{i \in N_t} s^i_{ht} \prod_{i \in N_t} (1 - s^i_{ht})}
\end{aligned}$$

Proof of Equation (7)

- Pr^i_{ht}: probability that guest i prefers listing h over all options at time t
- Pr_{ht}: probability that listing h is booked at time t
- $\bar{Pr}^i_{ot}$: probability that guest i prefers the outside option over unbooked listings at time t

- δ_h^t: utility of listing h at time t
- ξ_h, γ_t: listing and time mean-utility fixed effects
- M_t^i : set of available listings for potential guest i in the market (the market is defined as day - number of guests)
- N_{ot} : number of potential guests who choose the outside option at time t
- N_t : set of potential guests at time t
- B_t : set of all renters who booked a listing
- NB_t : set of all potential renters who did not book a listing

$$Pr_{ht} = 1 - \prod_{i \in N_t} (1 - Pr_{ht}^i) = 1 - (1 - Pr_{ht}^i)^{N_t}$$

$$\delta_h^t = \xi_h + \gamma_t$$

$$Pr_{ht}^i = \frac{\exp \delta_h^t}{1 + \sum_{j \in M_t^i} \exp \delta_j^t}, \quad Pr_{ot}^i = \frac{1}{1 + \sum_{j \in M_t^i} \exp \delta_j^t}$$

$$\frac{\partial Pr_{ht}^i}{\partial \gamma_t} = \frac{exp(\delta_h^t)(1 + \sum_{j \in M_t^i} \exp \delta_j^t) - exp(\delta_h^t) \sum_{j \in M_t^i} \exp \delta_j^t}{(1 + \sum_{j \in M_t^i} \exp \delta_j^t)^2} = Pr_{ht}^i - Pr_{ht}^i(1 - Pr_{ot}^i) = Pr_{ht}^i Pr_{ot}^i$$

$$\frac{\partial Pr_{ht}^i}{\partial \delta_h} = Pr_{ht}^i(1 - Pr_{ht}^i)$$

$$ll = \sum_t \sum_i \sum_h I_{ht}^i \log(Pr_{ht}^i) + \sum_t \sum_i I_{ot}^i \log(Pr_{ot}^i) = \sum_t \sum_i \sum_h I_{ht}^i \log(Pr_{ht}^i) + \sum_t \sum_{i \in NB} I_{ot}^i \log(Pr_{ot}^i)$$

$$\frac{\partial ll}{\partial \xi_h} = 0 \implies$$

$$\sum_t (\sum_{i=h} (1 - Pr_{ht}^i) + \sum_{i \neq h, i \in B} -Pr_{ht}^i) + \sum_t \sum_{i \in NB} -Pr_{ht}^i = 0 \implies N_h = \sum_t \sum_i Pr_{ht}^i$$

$$\frac{\partial ll}{\partial \gamma^t} = 0 \implies$$

$$\sum_{i \in B} Pr_{o_t}^i - \sum_{i \in NB} (1 - Pr_{ot}^i) = 0 \implies N_o^t = \sum_{i \in B} Pr_o^i + \sum_{i \in NB} Pr_o^i = \sum_i Pr_o^i$$