The Distributional Impact of the Sharing Economy: Evidence from New York City

Sophie Calder-Wang *

December 20, 2021

Abstract

I estimate the welfare and distributional impact of the home-sharing platform Airbnb on New York City renters. I develop a structural model of an integrated housing market with two novel features. First, in addition to the traditional long-term rental market, absentee landlords can reallocate their housing units to the newly available short-term rental market. Second, residents can directly host short-term visitors, increasing housing utilization. Overall, renters in NYC suffer a welfare loss of $2.4 billion, where losses from increased rents dominate gains from hosting. Moreover, the increased rent burden falls most heavily on high-income, educated, and white renters. By characterizing winners and losers, this paper provides a framework for evaluating the impact of such technological innovations.

JEL Codes: D6, L1, L23, L85, L86, R21, R31, R52

*The Wharton School, University of Pennsylvania, 3733 Spruce Street, Philadelphia, PA 19104. *Email: sophiecw@wharton.upenn.edu.* I am grateful to my advisors Ariel Pakes, Edward Glaeser, Paul Gompers, Robin Lee, and Adi Sunderam for their guidance and encouragement. This paper benefited from conversations with seminar and conference participants at UC Berkeley, Columbia University, Chicago Booth, Federal Housing Finance Agency, Harvard Business School, Kellogg, MIT Sloan, NBER Summer Institute, NYU IO Day, Oxford, Stanford, Texas A&M, UT Austin, Wharton, and the Online Spatial and Urban Seminar. This research was funded in part by the Joint Center for Housing Studies at Harvard University and the Ewing Marion Kauffman Foundation. All errors are my own.

1 Introduction

Economic theory teaches that cost-reducing new technologies should improve overall welfare. By substantially reducing transaction costs, peer-to-peer platforms such as Airbnb allow existing housing units to be used by short-term visitors in exchange for payment. However, it is not necessarily Pareto improving.

Opponents of Airbnb argue that when housing units are being reallocated away from the traditional long-term rental market, it exacerbates housing affordability problems, especially in cities with a constrained housing supply such as New York City (NYC). On the other hand, proponents of Airbnb argue the additional income that hosts earn from home-sharing, especially in expensive cities, is vital to their livelihood.

Therefore, the question is empirical: For renters in NYC, does the welfare gain from home-sharing offset the welfare loss from increased housing costs? Moreover, how does the welfare impact differ across key demographic characteristics, such as income, education, race, ethnicity, and family structure?

To answer these questions, I develop and estimate the first structural model of an integrated housing market that incorporates Airbnb's two key innovations. First, the long-term and short-term rental markets become integrated on the supply side: an absentee landlord who owns a housing unit can choose between the two markets, whichever yields a greater profit. Second, the utilization of existing housing units increases when renters offer space in their homes to host short-term visitors. Importantly, the structural model captures rich heterogeneity in housing demand and heterogeneity in the cost of home-sharing, which allows me to analyze the distributional impact across demographic groups.

I model the demand for long-term rentals as a discrete choice problem featuring heterogeneous household preferences over housing attributes (McFadden, 1978). The demand model captures a rich set of housing attributes ranging from hedonic attributes (e.g., the number of bedrooms, year built, type of structure) to neighborhood attributes (e.g., commuting time and neighborhood demographics in terms of race, ethnicity, and education). Moreover, it also allows for a horizontal preference over these attributes by household demographics.

To estimate the long-term rental demand, I adopt methods from Berry, Levinsohn and Pakes (2004) and Bayer, Ferreira and Mcmillan (2007). I take advantage of the individual-level data available in the American Community Survey (ACS) Public Use Microdata Sample, where I can observe the full vector of household demographic characteristics, together with the full vector of housing attributes chosen by the household. Given that the entry of Airbnb varies greatly across space and housing types, the ability to incorporate heterogeneity into the long-term rental demand is crucial to evaluating its distributional impact.

I model the supply of short-term rentals by residents as a binary choice problem, where a resident decides whether to share her home with a short-term visitor on a given day at the prevailing market price. She makes this decision based on the trade-off between the income she makes and the cost of providing such short-term rental services. The short-term rental supply model allows households to have different costs and price sensitivities based on their demographic characteristics, such as age, education, and family structure.

To estimate the short-term rental supply, I leverage high-frequency Airbnb transaction data. Although canonical BLP methods (Berry, Levinsohn and Pakes, 1995; Nevo, 2000, 2001) are typically used to estimate demand systems, I propose a novel adaptation so they can be used to estimate a heterogeneous *supply* system. Because I observe the location of each housing unit on Airbnb, variations in the distribution of neighborhoods demographics and seasonality in the short-term rental market allow me to estimate the heterogeneity in the cost of home-sharing. Because the high-frequency daily data results in a large number of market-share equations to match, I employ the MPEC procedure developed in Dubé, Fox and Su (2012) to improve the numerical performance of the estimator.

With the estimated model parameters, I conduct counterfactual analyses to evaluate the distributional impact of Airbnb on the participants of the housing market. To evaluate the welfare impact through the rent channel from reallocation, I perform a counterfactual analysis in which all housing units made available by absentee landlords on Airbnb are returned to the long-term rental market. To evaluate the welfare impact through the host channel in the short-term rental market, I perform a counterfactual analysis in which the hosts are no longer allowed to participate in home-sharing.

The key findings are threefold: (i) The net impact of Airbnb aggregated across all renters is a loss of $2.4bn, where the losses from the rent channel at $2.7bn dominate the gains from the host channel at $0.3bn. (ii) The median renter loses $125 p.a., whereas renters who are high-income, educated, and white suffer more significant losses because they demand housing types that are more desirable in the short-term rental market. (iii) A divergence emerges between the median and the tail: the equilibrium rent increase affects all 2.1 million renter-occupied units in NYC, but host gains accrue heavily to a small fraction of households with particularly low costs of sharing, including low-income families.

From the perspective of the social planner, because the rent increase results in a welfare transfer from renters to absentee landlords, the overall welfare impact of Airbnb remains positive because it also includes the economic gains accrued to all hosts, as well as the surpluses accrued to tourists net of hotel losses. However, with 67% of the housing units being renter-occupied, the median household in the city is made worse off from having Airbnb in New York.

Contrary to what typical political or media narratives purport, welfare differences are primarily driven by the geographical patterns of Airbnb usage, where its penetration in NYC tends to be greater in educated, high-income, and predominantly white neighborhoods, leading to worse welfare consequences for them. As such, regulatory efforts against Airbnb reallocation are unlikely to be targeted at housing affordability challenges faced by vulnerable populations, where the problem is fundamentally still driven by housing supply constraints.

Related Literature

This paper is the first to build and estimate a comprehensive model of an *integrated* housing market to evaluate the economic impact of home-sharing platforms on housing market participants. The structural approach enriches the growing body of reduced-form studies documenting that Airbnb tends to increase rents in many places (Barron, Kung and Proserpio, 2021), including Boston (Horn and Merante, 2017), Los Angeles (Koster, van Ommeren and Volkhausen, 2021), New Orleans (Valentin, 2021) and Barcelona (Garcia-López, Jofre-Monseny, Martínez-Mazza and Segú, 2020).

A model-based approach is beneficial because it captures the equilibrium effects by allowing households to re-optimize when the *bundle* of housing attributes available has changed. Thus, it readily characterizes the spillover effect of Airbnb entry in one part of the city onto prices in other parts of the city. Moreover, a model-based approach featuring heterogeneous preferences can capture the distributional impact over relevant household characteristics, which is especially relevant for policy-related concerns about income or racial disparities.

This paper complements existing structural papers that focus on the impact of Airbnb on other aspects of the market. Farronato and Fradkin (2021) estimate the welfare impact of Airbnb on tourists and hotels, highlighting the differential supply elasticities of peer suppliers compared to incumbent hotels. Almagro and Domínguez-Iino (2021) estimate the value of endogenous neighborhood amenities leveraging a substantial growth of Airbnb in Amsterdam. This paper complements their work by estimating the net effect on residents as they act both as drivers of long-term rental demand and as a source of short-term rental supply. While Farronato and Fradkin (2021) estimate the overall distribution of home-sharing costs, this paper is the first to comprehensively examine how costs of peer production differ across income and demographic characteristics.

This paper finds supply constraints in the housing market (Saiz, 2010; Gyourko, 2009; Gyourko and Molloy, 2015; Murphy, 2018; Baum-Snow and Han, 2021) have a large effect on how efficiency gains from the sharing economy are distributed, adding to the literature on the economic impact of housing constraints and related housing regulations (Ganong and Shoag, 2017; Hsieh and Moretti, 2019; Favilukis, Mabille and Van Nieuwerburgh, 2021; Diamond, McQuade and Qian, 2019). In addition, since the ability to rent to short-term visitors alters the cash-flow-generating ability of the underlying housing asset, this paper is also related to studies that estimate the determinants of housing value.[1]

More broadly, this paper is an empirical analysis of the economic impact of technology adoption and diffusion (Zvi, 1957; Bass, 1969). With the proliferation of online peer-to-peer markets (Einav, Farronato and Levin, 2016), researchers have examine their economic impact in a variety of other contexts, including Seamans and Zhu (2014) for Craigslist, Aguiar and Waldfogel (2018) for Spotify, Cullen and Farronato (2021) for TaskRabbit, and for ride-sharing applications (Cohen, Hahn, Hall, Levitt and Metcalfe, 2016; Frechette, Lizzeri and Salz, 2019; Buchholz, Doval, Kastl, Matějka and Salz, 2020; Castillo, 2020). With housing costs being one of the largest items in most household budgets, it is important to understand the economic impact of home-sharing platforms on residents.

The remainder of the paper proceeds as follows. To provide intuition for the main model, I start the next section with a stylized model, highlighting the key innovations of Airbnb. Section 2 discusses the background and the data used for the analysis. Section 3 presents the main structural model, followed by Section 4 which describes how the model is estimated. Section 5 performs the counterfactual analyses examining the welfare and distributional impact of Airbnb from the rent channel and the host channel, respectively. Section 6 concludes.

A Stylized Model

In this subsection, I present a stylized model highlighting two key innovations that Airbnb brings to the housing market.

First, Airbnb allows the long-term and short-term rental markets to become integrated because absentee landlords can offer their housing units in either market. Before Airbnb, we start with two separate markets in Figure 1: a long-term rental market (Panel A) and a short-term rental market (Panel B), each in equilibrium at their respective market-clearing price of p_0^L and p_0^S. In other words, renting out a residential housing unit on a short-term basis used to be prohibitively costly for a property owner.

With the arrival of Airbnb, an absentee landlord is no longer confined to operating in the traditional long-term rental market and gains the option to participate in the newly available short-term rental market. If prices are higher in the short-term rental market $p_0^S > p_0^L$, absentee landlords will be induced to reallocate toward Airbnb and obtain higher prices, assuming for now the cost of operating in either market is zero. As more and more housing units are reallocated, the price wedge between the two markets is reduced. In equilibrium, no-arbitrage condition pins down the new market price $p_1 = p_1^L = p_1^S$ that clear both markets, as well as the equilibrium

[1] This is an extensive literature that employs a variety of methods, including both choice-based sorting (Ferreyra, 2007; Timmins, 2007; Banzhaf and Walsh, 2008; Bayer, Keohane and Timmins, 2009; Klaiber and Phaneuf, 2010; Tra, 2010; Galiani, Murphy and Pantano, 2015; Bayer, McMillan, Murphy and Timmins, 2016) and reduced-from estimates (Black, 1999; Greenstone and Gallagher, 2008; Campbell, Giglio and Pathak, 2011; Davis, 2011; Autor, Palmer and Pathak, 2014, 2019).

number of reallocated housing units S^A.

Such market integration has clear welfare consequences: the dark-blue rectangle in Panel A represents the welfare transfer to property owners from remaining renters, whereas the light-blue triangle represents the welfare loss for displaced renters who leave the city. Because I can observe the reallocated quantity S^A directly from data, I estimate the slope of the long-term rental demand D^L to quantify the magnitudes of the welfare impact from increased rents.

A second innovation of Airbnb is that it allows residents to act as peer suppliers and participate directly in the production of short-term rental services. The additional short-term rental supply provided by resident hosts is represented as an upward-sloping supply curve S^R in Panel D, where more residents will find the hassle of home-sharing worthwhile if the price p_1^S is high.

When residents share part of their home with short-term visitors on Airbnb, they do not displace themselves. As such, the expansion of short-term rental supply from resident hosts does not create a corresponding supply squeeze in the long-term rental market, in sharp contrast with housing reallocation by absentee landlords.

Such supply expansion also has clear welfare consequences: the green triangle in Panel D indicates the surplus accrued to resident hosts. Therefore, I estimate the slope of the short-term supply S^R to compute the welfare gains from the host channel, which can, in turn, be netted against the welfare losses from the rent channel.

Despite its stylized nature, this simple model clarifies the main market equilibrium conditions of the integrated housing market after the entry of Airbnb, namely, market clearing in the long-term and the short-term rental markets, together with no arbitrage for absentee landlords. It also highlights this paper's key empirical objects of interest, namely, the welfare loss through the rent channel and the welfare gain through the host channel. The main structural model in Section 3 builds on the stylized model by incorporating heterogeneity in both demand and supply.

2 Background and Data

Although "the sharing economy" does not have one single definition, several important features stand out. It allows existing asset owners to increase utilization by allowing someone else to use their asset temporarily in exchange for payment. Platform companies that facilitate such exchanges typically do not own the underlying shared assets themselves, so services are fulfilled by numerous peers on one side of the platform. In addition, the development of a well-functioning reputation system alleviates the problem of asymmetric information, thus facilitating trade.

A leading example in this category is the home-sharing platform Airbnb, an online marketplace for arranging short-term lodging, especially homestays. Founders of Airbnb started the company in 2008 when they noticed local hotels in San Francisco were sold out during a conference, so they allowed a number of guests to sleep on their air mattresses.[2] It grew rapidly over time, raising over $6bn in venture capital funding and reaching a market capitalization of over $100bn in 2021. It has over 5 million listings across the globe, significantly higher than any existing hotel group.[3]

Data

The primary data source is a full sample of Airbnb listings and transactions scraped by a third-party data vendor AirDNA, who started scraping Airbnb.com website comprehensively in 2014. For each listing in NYC, the dataset contains detailed information about the property characteristics, including the type of property and hedonic attributes such as the number of bedrooms, the number of bathrooms, and other relevant amenities. Table 1 summarizes the breakdown by listing types in NYC: 51% are entire homes, 45% are private rooms, and less than 5% are shared rooms. Importantly, the dataset also contains the latitude and the longitude of the property, allowing me to map to its corresponding neighborhood.[4]

Figure 2 shows the rapid growth of Airbnb. Between 2014 and 2018, the number of reservations made on Airbnb quadrupled in many cities. The largest metropolitan market in the U.S. measured by days reserved is New York, followed by Los Angeles and Miami. Airbnb booked 5.8 million days of stay in NYC in 2018, which is about 15% of the total number of hotel stays. The average price of an entire home on Airbnb is $224 per night, and the average price of a private room is $86 per night. In 2018, 74,963 listings had experienced active transactions on Airbnb, representing about 2.2% of all housing units. Moreover, Airbnb is the most dominant player among short-term rental platforms, capturing over 90% of the market share.

The average level of Airbnb activity in the city masks the extensive geographic heterogeneity across neighborhoods and boroughs, as illustrated in Figure 4. In Brooklyn, the neighborhood with the highest proportion of housing units ever active on Airbnb in 2018 is Greenpoint & Williamsburg (9.3%), followed by Bushwick (8.5%). In Manhattan, Chelsea, Clinton & Midtown (6.9%) is at the top, followed by Chinatown & Lower East Side (6.5%). In Queens, the most active neighborhood is Astoria & Long Island City (2.8%). In the Bronx, the most active neighborhood is Concourse, Highbridge & Mount Eden (0.4%), which has much less penetration than other boroughs.[5]

Another beneficial feature of the dataset is its high-frequency panel, available at the daily level. For each listing and every day, I observe whether the listing was available on Airbnb, its

[2] https://news.airbnb.com/about-us/

[3] The largest hotel company in the world, Marriott International, operates approximately 1.4 million rooms.

[4] Airbnb adds some perturbation to property locations to ensure host privacy by up to 500 feet (Wachsmuth and Weisler, 2018). However, it introduces minimal noise in terms of assigning a property to its neighborhood at the Public Use Microdata Area (PUMA) level.

[5] Staten Island is excluded from the analysis because it is unlikely to be in a short-term visitor's choice set.

listed price, and whether a reservation occurred on that day.[6] Figures A.13 and A.14 shows strong seasonal patterns in this market. Peak demand predictably happens on New Year's Eve each year. The trough season happens predictably in January and February. The peak-to-trough ratio is 3.5x for daily quantity and 1.6x for daily price.

Based on calendar availability and the listing type, I approximate the proportion of housing units likely to have been reallocated away from the long-term rental market. Concretely, I designate entire homes available on Airbnb for over 180 days in 2018 as "reallocated" by their property owners, whose alternative use would have been long-term rentals. It averages to 0.68% of the rental housing stock, and the distribution also exhibits heterogeneity across neighborhoods, as shown in Figure A.12.

The second dataset is the American Community Survey (ACS) Public Use Microdata Sample. Because it contains individual-level data, it is particularly helpful in estimating the housing choice model. I observe the full vector of household demographic characteristics, together with the full vector of housing attributes chosen by the household. The key demographic variables include household income, education, race, ethnicity, age, and family size. The key housing attributes include monthly rent, number of bedrooms, building age, and type of building. Moreover, I also observe the location of the home at the neighborhood level.[7]

Overall, NYC has 3.14 million occupied housing units, of which 67% are renter-occupied. Demographic disparities are substantial across neighborhoods among renters: the median renter in Manhattan makes \$67,000 a year, whereas the median renter makes \$52,000 in Queens, \$44,000 in Brooklyn, and only \$31,000 in the Bronx. Similar patterns of disparity exist for education. In terms of race and ethnicity, 47% of Manhattan residents are non-Hispanic whites, but only 7% of the Bronx residents are white.

Table 2 shows the correlations between Airbnb reallocation and neighborhood demographics, revealing more Airbnb listings in neighborhoods with more white, educated, and higher-income residents. However, a full model is required to translate this empirical pattern into distributional impact in welfare, which needs to take into account both the equilibrium price effects in the long-term rental market and the gains from hosting on Airbnb accrued to local residents.

Lastly, I augment the analysis with data from Smith Travel Research, where I obtain the daily aggregate number of hotel rooms sold in NYC from 2007 to 2018. I use them to construct a seasonality-based demand shifter to estimate the cost of peer supply.

[6] Since the scraper visits each listing multiple times per week, the number of days reserved is backed out from changes in the calendar availability.

[7] In 1-Year ACS microdata, the location of the home is available at the PUMA level. Because New York City is densely populated, it contains 55 PUMAs, and they are used as an approximation for neighborhoods.

3 Model

The main model features an integrated housing market that captures two key innovations brought about by Airbnb: Reallocation to Airbnb by absentee landlords and increased housing utilization from resident hosts.

A. Long-Term Rental Demand

I start with a model of residential housing choice using a random utility framework following McFadden (1978) and Bayer, Ferreira and Mcmillan (2007). Each resident faces a discrete choice problem among all housing types, where each housing type is defined by its neighborhood n and its physical characteristics, including the number of bedrooms, the age of the building, and the building type. It results in 1,050 housing types in NYC.

Each housing type has N_h units that are not further differentiated beyond an idiosyncratic component $\epsilon^L_{i,j}$. Hence, the long-term rental utility of household i derived from housing unit j of type h is

$$u^L_{i,j} = \alpha^L_i p^L_h + \boldsymbol{\beta}^L_i \mathbf{X}^L_h + \xi^L_h + \epsilon^L_{i,j} \tag{3.1}$$

where the superscript L indicates quantities pertaining to the long-term rental market. p^L_h refers to the price of housing type h. $\mathbf{X}^L_h$ includes both the physical and the neighborhood attributes, which includes the percentage of the households with a college degree and the percentage of African American, Hispanic, and Asian households in a given neighborhood. I also include a measure of its location amenity using the average commuting time. In addition, I allow an unobserved quality ξ^L_h that could be correlated with price. The price coefficient α^L_i and the coefficients on the housing characteristics β^L_i are determined in a flexible way based on the vector of observable demographics z_i, which includes household income, race, ethnicity, education, and family size.

Each household i makes an optimal housing choice j by maximizing utility:

$$y^L_i = j \iff u^L_{i,j} > u^L_{i,-j}.$$

The set of households who choose housing type h is the union of those who choose any housing unit j that is of type h:

$$A^L_h = \bigcup_{j:h(j)=h} \{z_i, \epsilon^L_{i,\cdot} : u^L_{i,j} > u^L_{i,-j}\}.$$

The long-term rental demand for housing type h is obtained by integrating over all households:

$$D_h^L(p_h^L, p_{-h}^L) = \int_{A_h^L} dP(\epsilon^L) dP_D^*(z)$$

where P_D^* is the empirical distribution of household demographics of all potential city residents.[8]

B. Long-Term Rental Supply

The key assumption is that the total supply of housing in NYC is fixed.[9] Prior to the arrival of Airbnb, the long-term rental market could be characterized by the market clearing price for each housing type h:

$$\forall h : \quad D_h^L(p_h^L, p_{-h}^L) = S_h^F$$

However, after the arrival of Airbnb, I also specify the model for the short-term rental market before imposing market clearing in the new integrated housing market.

C. Short-Term Rental Supply

The supply of short-term rentals on Airbnb is categorized into two types: those provided by absentee landlords and those provided by residents.

C1. Short-Term Rental Supply from Absentee Landlord

Absentee landlords will reallocate housing units from the traditional long-term rental market to the newly available short-term rental market if doing so is profitable. In this case, for every unit reallocated, one fewer unit will be available for traditional long-term tenants.

Let the price in the short-term rental market be $p_{h,t}^A$ for type h on day t. The utility of accepting a short-term rental visitor on day t in housing unit j of type h is

$$u_{j,t}^A = p_{h,t}^A + \nu_{j,t}^A \tag{3.2}$$

where $\nu_{j,t}^A$ represents the cost of operating the short-term rental.

Then, an absentee landlord chooses to reallocate if he can make more money in the short-term

---

[8] The market size is the relevant metro area, which includes NYC and surrounding areas within a commutable distance. In practice, I focus on all contiguous counties that surround NYC, namely, Hudson, Nassau, Westchester, and Bergen.

[9] Housing construction in NYC has been depressed over the past three decades. Figure A.16 shows the distribution of building ages, where 41% of the homes today were built prior to 1940, 88% were built prior to 1990, and only 2.9% were built after 2010.

rental market than in the long-term rental market:

$$y_j^A = 1 \iff \frac{1}{T}\sum_{t=0}^{T} \max\{\, p_{h,t}^A + \nu_{j,t}^A,\, 0\,\} > p_h^L + \nu_j^L$$

where ν_j^L is the cost associated with operating property j in the long-term rental market and p_h^L is the corresponding long-term rental rate. Note the decision to reallocate is made with the consideration of a relatively long period T (e.g., a year), which allows for predictable variations in prices and vacancies due to seasonality in the short-term rental market.

Hence, the short-term supply of type h from absentee landlord reallocation is as follows:

$$S_h^A(p_h^L, p_{h,\cdot}^A) = \int_{A_h^A} dP(\nu^A, \nu^L), \quad A_h^A = \bigcup_{j:h(j)=j} \{\nu_{j,\cdot}^A, \nu_j^L : \quad y_j^A = 1\}.$$

On a given day t, the short-term supply from reallocation is the integral over those who can operate profitably at the market rate, given that it has already been reallocated to Airbnb:

$$S_{h,t}^A(p_h^L, p_{h,\cdot}^A) = \int_{A_{h,t}^A} dP(\nu^A, \nu^L), \quad A_{h,t}^A = \bigcup_{j:h(j)=j} \{\nu_{j,\cdot}^A, \nu_j^L : \quad y_j^A = 1, \quad u_{j,t}^A > 0\}.$$

C2. Short-Term Rental Supply from Residents

As the second type of short-term rental supply, city residents can directly host short-term visitors in their homes, perhaps more accurately reflecting the spirit of the "sharing economy." In this case, the utility derived from the supply of a short-term rental room by household i in neighborhood n on day t is

$$u_{i,t}^R = \alpha_i^R p_{n,t}^A + \boldsymbol{\beta}_i^R \mathbf{X}_{n,t}^R + \xi_{n,t}^R + \epsilon_{i,t}^R \tag{3.3}$$

where $p_{n,t}^A$ denotes the price of an Airbnb private room and $\mathbf{X}_{n,t}^R$ captures the cost of providing a private room. β_i^R is a function of household demographics, including age, education, income, and family structure. The price coefficient α_i^R is a function of household income. The model allows for an unobserved cost that may be correlated with the prevailing market price. A household-specific idiosyncratic taste for home-sharing on a given day is included as $\epsilon_{i,t}^R$. Because the product is a private room, it is only differentiated at the neighborhood and day level. In other words, a private room in a two-bedroom home is not further differentiated from one in a three-bedroom home.

Then, a resident chooses to supply a room on a given day if the utility of hosting is greater than the value of the outside option (e.g., personal use), which is normalized to zero. The total short-term rental supply by residents in neighborhood n on day t is the integral over such households in the neighborhood:

$$S^R_{n,t}(p^A_{n,t}) = \int_{A^R_{n,t}} dP(\epsilon^R) dP^*_{D_n}(z), \quad A^R_{n,t} = \{z_i, \epsilon^R_{i,t} : u^R_{i,t} > 0\}$$

where $P^*_{D_n}$ denotes the empirical distribution of household demographics in neighborhood n.

D. Short-Term Rental Demand

I denote the short-term rental demand for Airbnb (after hotels) from short-term visitors as $D^A_{h,t}(p^A_{h,t}, p^A_{-h,t})$, where $h \in \mathcal{H} \cup \mathcal{H}_0$. The choice set includes both entire homes of all types $\mathcal{H}$ and private rooms in all neighborhoods $\mathcal{H}_0$.

E. Market Equilibrium

The market is characterized by a sorting equilibrium. In the long-term rental market, the equilibrium price vector p^L_h ensures the demand for each housing type h equals its supply, which is the total number of underlying housing units *less* what have been reallocated:

$$\forall h \in \mathcal{H} : \quad D^L_h(p^L_h, p^L_{-h}) = S^F_h - S^A_h(p^L_h, p^A_{h,\cdot}). \tag{3.4}$$

In the short-term rental market, every day, the equilibrium price vector $p^A_{h,t}$ ensures the short-term rental demand equals the short-term rental supply, which is composed of both entire homes and private rooms:

$$\forall t, h \in \mathcal{H} : \quad D^A_{h,t}(p^A_{h,t}, p^A_{-h,t}) = S^A_{h,t}(p^L_h, p^A_{h,\cdot}) \tag{3.5}$$

$$\forall t, h \in \mathcal{H}_0 : \quad D^A_{h,t}(p^A_{h,t}, p^A_{-h,t}) = S^R_{h,t}(p^A_{h,t}). \tag{3.6}$$

4 Estimation

The estimation of the structural model is composed of two parts. In the first part, I estimate the long-term rental demand model using individual-level data from the cross-section of housing choices. In the second part, I estimate the short-term rental supply model using aggregate data across multiple neighborhoods and time periods.

4.1 Estimating Long-Term Rental Demand

To estimate the long-term rental demand, I use a two-step procedure. In Step 1, I estimate the heterogeneous preference coefficients. In Step 2, I estimate the linear coefficients.

Moment Conditions for Long-Term Rental Demand

Recall that the utility for household i considering home j of type h depends on its price p^L_h and its housing attributes X^L_h. I parameterize the preference coefficients as the sum of two components:

the first component is linear coefficients common to all households, and the second component is heterogeneous coefficients applied to observable household demographics z_i:

$$u_{i,j}^L \quad = \quad \alpha_i^L\, p_h^L \;+\; (\boldsymbol{\beta}_i^L)^T\, \mathbf{X}_h^L \;+\; \xi_h^L \;+\; \epsilon_{i,j}^L$$

$$\begin{bmatrix} \alpha_i^L \\ \boldsymbol{\beta}_i^L \end{bmatrix} = \underbrace{\begin{bmatrix} \alpha^L \\ \boldsymbol{\beta}^L \end{bmatrix}}_{\substack{\text{1. linear coefficients,} \\ \text{common to all;} \\ \text{part of the mean utility } \delta_h^L}} + \underbrace{\begin{bmatrix} \pi_{\alpha,1}^L \dots \pi_{\alpha,K}^L \\ \pi_{\boldsymbol{\beta},1}^L \dots \pi_{\boldsymbol{\beta},K}^L \end{bmatrix}}_{\substack{\text{2. heterogeneous coefficients,} \\ \text{household-specific;} \\ \text{part of heterogeneous utility } \lambda_{i,h}^L}} \begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,K} \end{bmatrix}$$

With logit errors, the probability of household i choosing type h is

$$Pr(y_i^L \in h;\, \delta^L, \pi^L) = \frac{N_h \exp\left(\delta_h^L + \lambda_{i,h}^L\right)}{\sum_{h'} N_{h'} \exp\left(\delta_{h'}^L + \lambda_{i,h}^L\right)}$$

where $N_h = S_h^F - S_h^A$ denotes the number of available housing units of type h. δ_h^L is the mean utility from housing type h that is common to all households, and $\lambda_{i,h}^L$ is the heterogeneous part of the utility specific to household i:

$$\begin{aligned} \lambda_{i,h}^L &= \Big(\sum_k \pi_{\alpha,k}^L z_{i,k}\Big) p_h^L + \Big(\sum_k \pi_{\boldsymbol{\beta},k}^L z_{i,k}\Big)^T \mathbf{X}_h^L \\ \delta_h^L &= \alpha^L p_h^L + (\boldsymbol{\beta}^L)^T \mathbf{X}_h^L + \xi_h^L \end{aligned}$$

Step 1 Moments

Because I observe the individual-level choices $\mathbb{1}\{y_i^L \in h\}$ from the data, I construct moment conditions that match the market shares, as well as moment conditions that match the covariance between the housing attribute $X_{b,h}^L$ and the average characteristics z_k of households who choose h. For instance, I match the covariance between the number of bedrooms and the average size of the families who choose that type of house. Thus, the set of moment conditions that pin down δ^L and π^L are as follows:

$$\forall h: \quad \mathbb{E}\big[\, Pr(y_i^L \in h;\, \delta^L, \pi^L)\,\big] \quad = \quad s_h^L \tag{4.1}$$

$$\forall b, k: \quad \text{Cov}\big(\, \mathbb{E}[z_k | y_i^L \in h;\, \delta^L, \pi^L],\ X_{b,h}^L\,\big) \quad = \quad \text{Cov}\big(\, \bar{z}_{k,h},\ X_{b,h}^L\,\big) \tag{4.2}$$

where the left-hand side denotes the model predictions and the right-hand side is estimated from its empirical counterparts:

$$\hat{s}_h^L = \frac{1}{N} \sum_i \mathbb{1}\{y_i^L \in h\}, \quad \hat{\bar{z}}_{k,h} = \frac{1}{N} \sum_i \mathbb{1}\{y_i^L \in h\} z_{i,k}$$

where $N = \sum_h N_h$. Market share moments (Eq 4.1) take the expectation over all households i. Attribute covariance moments (Eq 4.2) take the expectation over all housing types h.

The identification relies only on individual rationality, together with the fact that the housing prices and the attributes p_h^L, X_h^L are exogenous from the perspective of a single household making the choice. Given that each household can reasonably assumed to be infinitesimal, this condition holds.

Step 2 Moments

Because I allow market price p_h^L to be correlated with unobserved quality ξ_h^L, I employ an identification strategy that takes advantage of the shape of the housing characteristics space, following Berry, Levinsohn and Pakes (1999) and Bayer, Ferreira and Mcmillan (2007).

Intuitively, a home situated in a crowded part of the housing attribute space has a low equilibrium price, regardless of its own unobserved quality. As such, I construct a price instrument using the characteristics of other homes in the city. Specifically, to characterize the impact of the attribute space on market prices, I compute an alternative vector of equilibrium prices p^{IV} as an instrument for the observed prices p^L by setting the unobserved quality to zero $\xi_h^L = 0$ and resolving the market-clearing conditions across all home types:

$$\forall h: \quad \mathbb{E}\big[Pr\big(y \in h;\, (p^{IV}, \alpha^L, \beta^L, \xi^L = 0;\, \pi^L, X^{\text{exog}})\big)\big] = s_h^L \tag{4.3}$$

$$\forall h: \quad \mathbb{E}\big[Pr\big(y \in h;\, (p^L, \alpha^L, \beta^L, \xi^L;\, \pi^L, X)\big)\big] = s_h^L \tag{4.4}$$

$$\mathbb{E}\big[\xi^L p^{IV}\big] = 0 \tag{4.5}$$

$$\mathbb{E}\big[\xi^L X_b^{\text{exog}}\big] = 0 \tag{4.6}$$

Importantly, the construction of the alternative equilibrium price p^{IV} as an instrument utilizes an explicit supply-side pricing equation. In the housing contexts, it boils down to a market clearing condition with a fixed housing supply in Eq (4.3).[10] Note that for each guess of the demand parameter α^L, β^L, ξ^L, there exists a corresponding p^{IV}, so demand parameters and the price instrument are estimated jointly, where all moment conditions listed above are satisfied.

The identification assumption is that there exists a subset of housing characteristics $\mathbf{X}^{\text{exog}}$ that is independent of unobserved quality:

$$\xi_h^L \perp\!\!\!\perp \mathbf{X}^{\text{exog}}.$$

To find a subset of the housing attributes that may be considered reasonably independent of the unobserved ξ^L, I use a vector of immutable physical attributes of the housing stock,

[10] Although BLP-style instruments can be formed in various ways, using a single price instrument p^{IV} to capture the equilibrium price impact of the attribute space is an approximation of the optimal instrument in the sense of Chamberlain (1987), Reynaert and Verboven (2014), and Berry, Levinsohn and Pakes (1999).

following Bayer, Ferreira and Mcmillan (2007).[11] Meanwhile, I exclude average neighborhood demographics, because education, race, and ethnicity at the neighborhood level are likely to enter into unobserved quality.

Parameter Estimates for Long-Term Rental Demand

Step 1 Estimates

Moment conditions constructed in Step 1 pin down the heterogeneous preference coefficients. The heterogeneous WTP for each housing and neighborhood attributes by households demographics is calculated as $-\pi^L_{b,k}/\alpha^L$, summarized in Table A.9.

Among housing attributes, I find larger households have higher WTP for more bedrooms. Among neighborhood attributes, I find households have strong preferences for shared race and education background. African American, Hispanic, and Asian households have significantly higher WTP for neighborhoods with higher percentages of their own race and ethnicity. For example, an Asian household is willing to pay \$410 per month to live in a neighborhood that is one-standard-deviation higher in its percentage of Asian.[12] Educated families prefer neighborhoods with higher proportions of educated families. In addition, the outside option is more valuable for higher-income and larger families, whereas racial and ethnic minorities prefer to live in the city, all else being equal. The heterogeneity in front of the price coefficient reflects differential price sensitives, where higher-income and more educated households tend to be less price sensitive.

Step 2 Estimates

Moment conditions constructed in Step 2 pin down the linear coefficients, including the price coefficient and the average WTP for each housing attribute.

Table 3 column (2) shows the price coefficient α^L is -2.04, using the instrumented estimator. The F-stat is 15.7. The use of the price instrument has a significant impact, compared with the OLS specification shown in column (1). As a placebo test, Figure A.10 shows that the price instrument has the desired property of being uncorrelated with neighborhood attributes.

Table 3 column (3) shows the average household has a higher WTP for more bedrooms. It also prefers pre-war structures and fewer units in the building. The average household is willing to pay \$383 for a one-standard-deviation reduction in commuting time. Because the city's housing stock is divided into 1,050 housing types based on its neighborhood and housing attributes, the estimated demand elasticities vary by housing type. On average, the price response to a 1%

[11] The attributes included are indicators of the number of bedrooms, type of building, age of building, average commuting time to work centers, and an indicator for being inside the city.

[12] I do not attempt to explain whether it is an Asian household's preference for its neighbor's ethnicity or for neighborhood businesses that cater to its preferences. I also do not differentiate between whether it is due to housing market preference or discrimination. These parameters simply reflect the underlying choice patterns in the data and are assumed to be unchanged in the counterfactuals.

reduction in the supply of all housing types in NYC is estimated to be 1%, which implies a price elasticity of the aggregate demand of approximately -1.0.[13]

4.2 Estimating Short-Term Rental Supply

Although BLP methods are widely used by researchers to estimate demand systems (Berry, Levinsohn and Pakes, 1995; Nevo, 2000, 2001), I propose an adaptation so that it can be used to estimate a random-coefficient *supply* system. The key insight is that because the neighborhood location of each Airbnb listing is observed, it allows me to match the aggregate "market shares" of Airbnb supply in each neighborhood every day. As such, variations in the distribution of demographic characteristics across neighborhoods and short-term demand seasonality allow me to estimate the heterogeneity in cost.

Moment Conditions for Short-Term Rental Supply

Recall that the utility that a resident host i living in neighborhood n derives from sharing a private room on Airbnb on day t depends on how she values the income from sharing $p^A_{n,t}$ relative to the cost of providing such short-term rental services. To capture heterogeneity, I parameterize the preference coefficients α^R_i and $\boldsymbol{\beta}^R_i$ as the sum of a common component and a component that depends on its demographic characteristics:

$$\begin{aligned} u^R_{i,t} &= \alpha^R_i p^A_{n,t} + (\boldsymbol{\beta}^R_i)^T \mathbf{X}^R_{n,t} + \xi^R_{n,t} + \epsilon^R_{i,t} \\ \begin{bmatrix} \alpha^R_i \\ \boldsymbol{\beta}^R_i \end{bmatrix} &= \underbrace{\begin{bmatrix} \alpha^R \\ \boldsymbol{\beta}^R \end{bmatrix}}_{\text{common to all}} + \underbrace{\begin{bmatrix} \pi^R_{\alpha,1} \dots \pi^R_{\alpha,K} \\ \pi^R_{\boldsymbol{\beta},1} \dots \pi^R_{\boldsymbol{\beta},K} \end{bmatrix}}_{\text{household-specific}} \begin{bmatrix} z_{i,1} \\ \vdots \\ z_{i,K} \end{bmatrix} \end{aligned}$$

where $p^A_{n,t}$ denotes the prevailing price of an Airbnb private room in neighborhood n on day t. $\mathbf{X}^R_{n,t} = [1,\ t,\ t^2, Z_{month}, Z_{dow}, Z_{holiday}]^T$ captures features that contribute to the cost of hosting. It includes a constant term, a quadratic time trend, and dummies for month fixed effects, day-of-the-week fixed effects, and holiday fixed effects. Because the constant term captures the cost of sharing a room, such as the time and the hassle, I parametrize its coefficient β^R_i as a function of household income, age, education, and family structure. I also allow the price coefficient α^R_i to be a function of household income, permitting one's price sensitivity to differ by income.

Because each resident faces a binary choice between sharing and not sharing, assuming logit

[13]Compared to the literature, Albouy, Ehrlich and Liu (2016) uses a completely different set of assumptions and models whereby they find that the price elasticity of renters to be approximately -0.83 in U.S. metropolitan areas.

error, the quantity of supplied is

$$s^R_{i,n,t}(\delta^R_{n,t}, \pi^R) = \frac{\exp(\delta^R_{n,t} + \lambda^R_{i,n,t})}{1 + \exp(\delta^R_{n,t} + \lambda^R_{i,n,t})}$$
$$\lambda^R_{i,n,t} = \Big(\sum_k \pi^R_{\alpha,k} z_{i,k}\Big) p^A_{n,t} + \Big(\sum_k \pi^R_{\boldsymbol{\beta},k} z_{i,k}\Big)^T \mathbf{X}^R_{n,t}$$
$$\delta^R_{n,t} = \alpha^R p^A_{n,t} + (\boldsymbol{\beta}^R)^T \mathbf{X}^R_{n,t} + \xi^R_{n,t}$$

where the market share $S^R_{n,t} = \sum_{i \in n} s^R_{i,n,t}$ is based on the cost of sharing among all residents currently residing in neighborhood n.

As the unobservable cost at the neighborhood-time level $\xi^R_{n,t}$ is allowed to be correlated with the price $p^A_{n,t}$, I instrument the short-term rental price using a measure of tourist demand seasonality. Specifically, I use the number of hotel visits to the entire city of New York on the same day but lagged by seven years. Because Airbnb had a minuscule market share at that time,[14] there is little scope for reverse causality: unobserved structural errors in the cost of home-sharing should not affect the total number of NYC hotel visits from seven years ago. Moreover, if one is concerned that seasonality in the supply cost remains correlated with lagged hotel demand, I add a host of calendar-related controls, including month fixed effects, day-of-the-week fixed effects, and holiday fixed effects. What kind of variation does the price instrument leave us with? For example, it captures the impact of foreign holidays. They are persistent over time and affect hotel demand in NYC through increased tourism demand, but they are plausibly uncorrelated with the cost of home-sharing by city residents.

The moment conditions match market shares in each neighborhood n every day:

$$\forall n, t: \quad \mathbb{E}_{D^*_n}[\, s^R_{i,n,t}(\delta^R_{n,t}, \pi^R)\,] = s^{R,\text{o}}_{n,t} \tag{4.7}$$
$$\mathbb{E}[\xi^R Z] = 0 \tag{4.8}$$

where $s^{R,\text{o}}_{n,t}$ denotes the observed market share and $Z = [\mathbf{X}^{\mathbf{R}}, \mathbf{p}^{\mathbf{R,IV}}]$ includes the lagged hotel visits as the price instrument.

To estimate the supply system with 75,895 market-share equations, I cast the problem as a minimization routine over the GMM objective, using the Mathematical Programming with

[14] My data spans 2014 to 2018. When lagged by seven years, in 2007, Airbnb had not been founded yet. In 2011, Airbnb had only 300 listings in NYC. By contrast, the hotels in NYC sold over 25 million nights in 2007.

Equilibrium Constraints (MPEC) specification developed in Dubé, Fox and Su (2012):

$$\begin{aligned}
\min_{\delta^R_{n,r},\alpha^R,\beta^R,\pi^R_{\cdot},\eta} \quad & \eta^T W \eta \\
s.t. \quad & \forall n,t: \quad S^R_{n,t}(\delta^R_{n,t},\pi^R_{\cdot}) = S^{R,o}_{n,t} \\
& \eta = Z'(\delta^R - \alpha^R p^A - \boldsymbol{\beta}^R \mathbf{X}^R).
\end{aligned}$$

The main advantage of this estimation method is the sparsity structure of the Jacobian and the Hessian. Namely, the mean utility $\delta^R_{n,t}$ only affects the equilibrium in market (n,t) and does not affect other markets.[15] It allows the optimizer to perform better numerically, especially when the number of markets is large.

Parameter Estimates for Short-Term Rental Supply

Table 4 summarizes the supply coefficients of providing private Airbnb rooms. In column (4), using the lagged hotel visit as the price instrument, I find the price coefficient is 0.056, implying an average supply elasticity of 5.96. In comparison, without instrumenting, the price coefficients are biased downward significantly, shown in columns (1) and (2). Column (3) shows the inclusion of calendar fixed effects has a moderate effect on the price coefficient.

The average cost of home-sharing is high, estimated at \$224 per night, reflecting the fact that the overall share of hosts as a percentage of all residents is still low. In terms of heterogeneity, I find lower-cost suppliers tend to be young and educated households with no children. Having a college degree is associated with a reduction of \$59 per night in the cost of home-sharing. Having children increases the sharing cost by \$47 per night. Being 10 years younger is associated with a reduction of \$18. Moreover, I find a negative interaction between household income and price, suggesting lower-income households are more price sensitive. The average supply elasticity increases to 6.70 for those with one-standard-deviation lower income.

5 Counterfactuals

In this section, I provide estimates of the welfare and distributional impact of Airbnb. First, I discuss the welfare losses via the rent channel due to housing reallocation by absentee landlords. Second, I discuss the welfare gains via the host channel as residents act as peer suppliers. Finally, I discuss the net effects and implications for the social planner.

5.1 The Welfare and Distributional Impact of the Rent Channel

5.1.1 Aggregate Welfare Impact of the Rent Channel

To estimate the impact of Airbnb through the rent channel, I conduct a counterfactual analysis in which all housing units reallocated to Airbnb by absentee landlords are now returned

[15] See the appendix for the derivations of the relevant sparse matrices.

to the long-term rental market. Specifically, I recompute the vector of long-term rental prices across all housing types p_h^L by ensuring that housing demand equals housing supply:

$$\forall h : D_h^L(p_h^{L,\text{ No Airbnb}}, p_{-h}^{L,\text{ No Airbnb}}) = S_h^F. \tag{5.1}$$

With logit errors, the compensating variation is:

$$CV_i^L = \frac{1}{\alpha_i^L}\Big(\ln \sum_{j\in\mathcal{S}^F\setminus\mathcal{S}^A} \exp\left(V_{i,j}^L\right) - \ln \sum_{j\in\mathcal{S}^F} \exp\left(V_{i,j}^{L,\text{No Airbnb}}\right)\Big) \tag{5.2}$$

where $V_{i,j}^L = \alpha_i^L p_h^L + \beta_i^L X_h^L + \xi_h^L$ and $V_{i,j}^{L,\text{ No Airbnb}} = \alpha_i^L p_h^{L,\text{ No Airbnb}} + \beta_i^L X_h^L + \xi_h^L$ represent the non-idiosyncratic component of household utility at the actual and the counterfactual prices, respectively. $\mathcal{S}^F$ denotes the entire set of housing units, and $\mathcal{S}^A$ denotes the set of housing units observed to have been reallocated by absentee landlords.

Overall, housing reallocation to Airbnb results in a material welfare impact for all renters in the city. About 0.68% of the rental housing stock is reallocated to Airbnb. When all such units are returned to the long-term rental market, overall rent changes by 0.71%. The average compensating variation of the rent channel is a loss of $138 p.a. per renter, whereas the median loss is $128 p.a. and the standard deviation is $29.

The median renter's annual household income is about $47,000. Although the magnitude translates to only 25 basis points of the annual income, the increase in equilibrium rents affects *all* 2.1 million renters in the city. When aggregated across all renters, it amounts to a direct transfer to property owners of $200mm p.a., or $2.7bn in NPV terms.[16]

5.1.2 Drivers of Aggregate Welfare Changes

Connecting directly with the model, I illustrate that the welfare impact of the rental channel is driven by an inelastic housing supply. Moreover, the welfare impact is shouldered primarily by remaining renters, as opposed to displaced renters.

Role of Housing Supply Restrictions

When a supply-constrained city experiences a housing supply squeeze from Airbnb, the model illustrates the overall price effect can be decomposed into (i) the direct effect of a supply reduction on a given housing type, (ii) the spillover effect of a supply reduction from other housing types, and (iii) the indirect price impact when the supply of substitute housing types is also restricted.

[16] The NPV calculation assumes a capitalization rate of 7.5% based on NYC hotel REITs as of 2018, based on data compiled by CBRE Research. This assumption likely produces a lower bound on the NPV because the cap rate of the integrated market will likely fall in between the cap rates of the traditional hotel market and the traditional residential market.

Consider the market clearing conditions of the long-term rental market:

$$\forall h: \quad D_h^L(p_h^L,\, p_{-h}^L,\, s^L) - s_h^L = 0 \tag{5.3}$$

where s^L denotes the entire vector of supply of each housing type. Taking the total derivative of the market-clearing condition with respect to s_h and s_h' yields the following relationships:

$$\frac{dp_h^L}{ds_h^L} = \left(\frac{\partial D_h^L}{\partial p_h^L}\right)^{-1} \Bigg(\underbrace{1}_{\substack{\text{(i). direct impact of}\\ \text{supply reduction}}} - \underbrace{\sum_{k \neq h} \frac{\partial D_h^L}{\partial p_k^L} \frac{dp_k^L}{ds_h^L}}_{\substack{\text{(iii). indirect impact from}\\ \text{price increases of other home types}}} - \frac{\partial D_h^L}{\partial s_h^L} \Bigg) < 0 \tag{5.4}$$

$$\frac{dp_h^L}{ds_{h'}^L} = \left(\frac{\partial D_h^L}{\partial p_h^L}\right)^{-1} \Bigg(0 - \underbrace{\sum_{k \neq h} \frac{\partial D_h^L}{\partial p_k^L} \frac{dp_k^L}{ds_{h'}^L}}_{\substack{\text{(ii). spillover from}\\ \text{reductions of other home types}}} - \frac{\partial D_h^L}{\partial s_{h'}^L} \Bigg) < 0 \tag{5.5}$$

Eq (5.4) shows that the overall price impact from a supply reduction is a combination of the direct impact from the supply change s_h^L and the indirect impact from price changes of other housing types dp_k^L/ds_h^L. In particular, because housing supply is *constrained* across all housing types except for the outside option, a supply reduction in a given housing type h will lead to a price increase in other housing types k because of substitution. However, this price increase in p_k^L, in turn, creates additional upward pressure on the original housing type p_h^L.

Intuitively, an inelastic housing supply is akin to a setting of imperfect competition with quantity fixing: when a competitor's price increases, the optimal response is to *increase* one's own price when the total quantity is fixed. In the housing context, even though each absentee landlord does not have market power and participates in the market competitively, the overall difficulty in expanding the housing supply acts as the quantity-fixing mechanism.

In general equilibrium, the presence of such supply restrictions leads to a more exacerbated price response than the partial equilibrium effect alone, which is especially true for neighborhoods with little direct Airbnb penetration, as illustrated in Figure A.18.

Displaced vs. Remaining Renters

To decompose the price impact into displaced and remaining renters, I first quantify the welfare impact because of changes in the equilibrium price versus changes in the choice set.

Denote $W_i(p, \mathcal{S}) = \big(\ln \sum_{j \in \mathcal{S}} \exp V_{i,j}(p_h)\big)/\alpha_i$. The compensating variation in Eq (5.2) can be separated into one part driven by price changes and another part driven by choice set changes:

$$CV_i^L = \underbrace{\Big(W_i(p^L, \mathcal{S}^F \setminus \mathcal{S}^A) - W_i(p^L, \mathcal{S}^F)\Big)}_{\substack{\text{Welfare change due to} \\ \text{changes in the choice set}}} + \underbrace{\Big(W_i(p^L, \mathcal{S}^F) - W_i(p^{CF}, \mathcal{S}^F)\Big)}_{\substack{\text{Welfare change due to} \\ \text{changes in the equilibrium prices}}} \tag{5.6}$$

I find the increase in equilibrium housing prices contributed to a welfare impact of \$201mm p.a. and the choice set reduction contributed to a welfare impact of \$58mm p.a.

The welfare transfer from renters to property owners can be computed by summing over the price changes $\sum_h \Delta p_h^L \times (s_h^F - s_h^A)$, which amounts to \$200mm p.a. and is the first-order term (i.e., the dark-blue rectangle in Figure 1). The welfare loss from displaced renters amounts to \$0.9mm p.a., which is the second-order term (i.e., the light-blue triangle in Figure 1).

Hence, the vast majority of Airbnb's welfare impact through the long-term rental market amounts to a transfer from the remaining renters to property owners, whereas a small portion is through a reduction in the choice set. Less than 1% is due to the welfare loss of renters displaced outside of the city.

5.1.3 Distributional Implications of the Rent Channel

To understand the differential welfare impact along demographic lines, I compute the compensating variation for each household CV_i^L and aggregate them into their respective categories.

Household Size

I find Airbnb results in larger welfare losses from the rent channel for smaller households. Figure 6a shows the median welfare loss for households of size one is \$134 p.a. with a standard deviation of \$32. The median welfare loss for households of size four is \$116 p.a. Even though there remains significant variation within each household size category, the most negative welfare losses are still concentrated in small households.

In the Airbnb data, smaller housing units are disproportionately more prevalent in the short-term rental market relative to its underlying availability. Over 80% of entire-home listings on Airbnb (with over 180+ days availability) have fewer than two bedrooms in the NYC market. In comparison, 46% of the housing stock available for long-term rental have fewer than two bedrooms, as shown in Figure 3. This underlying pattern of Airbnb usage is one of the key drivers for why the welfare impact is more concentrated on smaller households. Nonetheless, despite the large differences in the bedroom-count distribution, the differences in the welfare impact across households are reduced because a reduction in one housing type creates equilibrium price effects on all other housing types.

Race and Ethnicity

I find that Airbnb results in larger welfare losses from the rent channel for white renters than African American, Hispanic, or Asian renters. Figure 6b shows the median welfare loss for white households is \$152 p.a. with a standard deviation of \$35 and a significant left tail. The median welfare loss for African American renters is \$134 p.a. The median welfare loss for Hispanic renters is \$113 p.a. The median welfare loss for Asian renters is \$127 p.a. Overall, when measured in dollar terms, the increased rent due to Airbnb reallocation hurts white renters the most, compared with minority renters.

Figure 5 shows significant sorting along demographic lines. Compared with Figure 4, it shows that neighborhoods with more Airbnb rentals tend to have greater percentages of white households. Hispanic and Asian households appear to be somewhat less segregated and are generally not concentrated in regions of heightened Airbnb activities. Although a municipal planner may choose to place different social welfare weights on different households, this analysis highlights that welfare implications depend heavily on the geographic patterns of Airbnb activity.

Education

I find Airbnb results in larger welfare losses from the rent channel for those with college degrees. Figure 7a shows the median welfare loss for college-educated renters is \$156 p.a. with a standard deviation of \$31. The median welfare loss for those without college a degree is \$120 p.a.

Besides the geography of Airbnb activities, the distributional difference is also driven by sorting preferences in housing demand. Specifically, given the preference for those with college degrees to live in neighborhoods that have more educated households, when renters in city center are affected by Airbnb, they are more likely to substitute to other educated neighborhoods, creating a spillover price impact to those who are demographically more similar to them.

Indeed, the structural model effectively captures the fact that sorting preferences can lead to a compounded welfare impact along demographic lines. Table 5 shows a subset of the cross-elasticities, where a price increase in a given neighborhood leads to higher rates of substitution to other neighborhoods that are closer in the demographic characteristics space.

For example, Forest Hills and Jackson Heights are located in close geographical proximity to each other in Queens, both far from the city center. Neither neighborhood experienced much Airbnb penetration, with less than 0.5% of the housing stock being affected. However, Forest Hills has a much higher proportion of educated households at 63%, whereas Jackson Heights has only 28%. Meanwhile, Forest Hills has a white majority, whereas Jackson Heights has a Hispanic majority. As a result, a price increase in Chelsea and Midtown, which has the highest level of Airbnb reallocation in the city, generates a much higher rate of substitution towards Forest Hills (0.7%) than Jackson Heights (0.1%), because both Chelsea and Forest Hills have high proportions of white and educated households. Consequently, in the counterfactual analysis, the equilibrium price impact in Forest Hills is much higher than in Jackson Heights.

Household Income

I find Airbnb results in larger welfare losses from the rent channel for those with higher income. Figure 7b shows the median welfare loss for renters in the top income quintile is \$167 p.a. with a standard deviation of \$38. The median welfare loss for renters in the bottom income quintile is \$123 p.a.

Consistent with the positive correlation between Airbnb activity and median neighborhood income shown in Table 2, the counterfactual analysis quantifies the extent to which the welfare losses are greater for higher-income renters. However, besides geography, another relevant factor is that the price coefficient α_i^L is smaller in magnitude for higher-income households, indicating their WTP for housing amenities is higher.

How much of the welfare differences across income groups is driven by the geographic pattern of Airbnb activity versus the fact that these households tend to have a higher WTP for housing? To disentangle these two different channels, I conduct an alternative counterfactual analysis if Airbnb listings were distributed *uniformly* across geographic space and housing types. Table A.11 shows that less than a quarter of the distributional differences are attributable to their higher WTP, whereas about three quarters of the differences are attributable to the geographical patterns of Airbnb activity. Therefore, geography still plays an outsized role in determining the distributional impact.

To summarize this section, the squeeze on the long-term rental market due to Airbnb results in a moderate yet material welfare transfer from renters to property owners at \$200mm p.a. or \$2.7bn in NPV terms. The median renter loses about \$128 p.a. The general equilibrium price effect is elevated across all renters because housing supply is difficult to expand, which acts as a quantity-fixing mechanism of the market. Across renters, I find higher-income, educated, and white renters suffer the most significant losses, when measured in dollar terms. The distributional differences are driven primarily by the geographical patterns of Airbnb activity and exacerbated by the preference for demographic sorting in housing choices.

5.2 The Welfare and Distributional Impact of the Host Channel

5.2.1 Aggregate Welfare Impact of the Host Channel

To estimate the welfare gains derived from a resident's ability to host on Airbnb,[17] I perform a counterfactual analysis in which the option to host on Airbnb is no longer available. The compensating variation of household i residing in neighborhood n is computed as follows

$$CV_i^R = \frac{1}{\alpha_i^R} \sum_t \ln\left(1 + \exp(V_{i,t}^R)\right) \tag{5.7}$$

[17] I focus on the supplier surplus of renters, to make it directly comparable to the previous section.

where $V_{i,t}^R = \alpha_i^R p_{n,t}^A + \beta_i^R X_{n,t}^R + \xi_{n,t}^R$ and is summed over the course of the year.

Overall, the distribution of the supplier surplus is centered close to zero but with a heavy right tail, suggesting the bulk of the benefits are accrued to a concentrated few. Figure 8 shows the surplus distribution of all renters in the city, where the median supplier surplus is only \$0.4 p.a. At the 75th percentile, the supplier surplus remains immaterial at \$5.9 p.a. However, the expected surplus on the very right tail above the 99th percentile amounts to \$307 p.a. When integrated over all renters, the total surplus produced by direct home-sharing amounts to \$23mm a year, or \$300mm in NPV terms.

5.2.2 Distributional Implications of the Host Channel

One of the key benefits of estimating a random-coefficient supply system is its ability to analyze the supplier surplus by observable demographic characteristics.

Household Income

Even though the home-sharing surplus is immaterial for the majority of the households, higher-income households are still expected to have a larger surplus on average. Meanwhile, the lowest-income group enjoys the largest benefits in the right tail, suggesting that a few low-income household benefit substantially from home-sharing.

Figure 9 shows the median host surplus for renters in the top income quintile is \$5.2 p.a., and the median host surplus for renters in the bottom income quintile is \$0.1 p.a. By contrast, the average surplus above the 99th percentile is significant at \$232 p.a. for the top income quintile, but it is even higher at \$454 p.a. for the bottom income quintile.

The difference between the median and the tail outcome results from two countervailing forces. Intuitively, the decision to share is based on a comparison of the market price $p_{n,t}^A$ and the cost to share $|\beta_i^R|/\alpha_i^R$. On the one hand, Airbnb demand tends to be stronger in centrally located high-income neighborhoods. On the other hand, higher household income also lowers one's price sensitivity α_i^R, suggesting the sharing income may not be as valuable.

Because there is significant demand seasonality in the short-term rental market, the income from hosting during peak demand times could be particularly valuable for lower-income households. The cross-partial of the compensating variation with respect to the market price $p_{n,t}^A$ and the price coefficient α_i^R is positive:

$$\frac{\partial CV_{i,t}^R}{\partial p_{n,t}^A} = \frac{\exp\left(\alpha_i^R p_{n,t}^A + \beta_i^R X_{n,t}^R + \xi_{n,t}^R\right)}{1 + \exp\left(\alpha_i^R p_{n,t}^A + \beta_i^R X_{n,t}^R + \xi_{n,t}^R\right)} \quad \Rightarrow \quad \frac{\partial^2 CV_{i,t}^R}{\partial p_{n,t}^A \partial \alpha_i^R} > 0.$$

where α_i^R is decreasing in household income. The supply elasticity for the top income quintile is 5.0, but much higher at 7.4 for the bottom income quintile. It explains the data pattern where more Airbnb rooms are booked in lower-income neighborhoods during periods of high tourism demand.

Household Demographics

I find that households that are younger and have no children obtain greater supplier surpluses as resident hosts. For the youngest households with no children, the average surplus is expected to be $80 p.a. For households with children, the average surplus is immaterial across all age groups. In terms of education, I find that educated households accrue greater supplier surplus as resident hosts. On average, the expected surplus is $36 p.a. for educated households and immaterial for households without a college degree.

5.3 Net Welfare Impact

5.3.1 Net Welfare Impact on Renters

To estimate the net welfare impact of Airbnb on renters, I combine welfare losses from the rent channel and welfare gains from the host channel. Because losses from the rent channel are diffused and gains from the host channel are concentrated, Figure 10 shows the net welfare impact for the median renter is a loss of $125 p.a. In fact, the net welfare for over 97% of renters is negative. Nonetheless, the long right tail in the distribution of host gains results in a similarly long tail in the net welfare impact, averaging to $164 p.a. above the 99th percentile.

Net Welfare by Household Income and Demographics

Overall, higher-income renters experience larger net losses on average and experience smaller net gains in the tail. Table 6 shows that renters in the top income quintile lose $146 p.a. on average. For renters in the bottom income quintile, the average net welfare is a loss of $114, but in the right tail, they accrue a net welfare gain of $319. Hence, for a small proportion of households, especially low-income ones, they can not only make up for the higher rents but also obtain significant surpluses from becoming a host.

In terms of education, the median net welfare impact is worse for more educated renters because the net welfare impact is dominated by the rent channel for the median renter. However, because education is also associated with a lower cost of home-sharing, educated renters are more likely to be in the right tail of the host surplus distribution. The divergence of the median and the tail by education is a result of the interaction between the demand and cost parameters in the short-term rental market.

In terms of race and ethnicity, the median white renter loses more than the median African American or Hispanic renter. This difference between the group medians is primarily driven by the differential welfare losses from the rent channel. The net welfare impact on the median white renter is -$152 p.a. In comparison, the net impact on the median African American renter is -$134 p.a., and for the median Hispanic renter, it is -$113 p.a.

Net Welfare by Geography

In addition to decomposing the net welfare impact by demographics, it is also instructive to decompose the results by neighborhood location, especially because the short-term rental demand varies substantially across the geographic space.

Interestingly, neighborhoods that experience heavier losses for the median renter also tend to experience large net gains in the right tail. For example, the net welfare impact for the median renter in Chelsea, which has high levels of Airbnb penetration, is -$146 p.a., while the tail of its host surplus is above $600 p.a., illustrated in Figures A.19 and A.20. The irony of the divergence is explained by the overall spatial patterns of short-term rental demand: high short-term demand for the neighborhood drives more reallocation of the housing units away from the long-term rental market, thereby raising rents for everyone in the neighborhood. At the same time, high short-term rental demand also increases the prices that resident hosts can benefit from when they share their homes. However, as low-cost resident hosts are relatively few in number, their large gains affect only the tail of the distribution, not the median.

Meanwhile, the variation in terms of how much host gains could offset rent increases is driven by the neighborhood's cost of home-sharing, which favors centrally located low-income neighborhoods. For instance, the median household income among renters in Bushwick is only 60% of the median household income of the neighboring Park Slope and Williamsburg. As such, residents there are more likely to find home-sharing worthwhile. Despite being a relatively lower-income area, Bushwick is still conveniently located in Brooklyn and attracts short-term visitors. As a result, the surplus from home-sharing is relatively more valuable to its residents.

5.3.2 Implications for the Social Planner

To evaluate the aggregate welfare impact on all participants of the market beyond just renters, I combine the model estimates for renters with a back-of-the-envelope approximation for property owners, tourists, and hotels, summarized in Figure 11.

First, the net impact on renters is a welfare loss of $2.4bn. The main component is a transfer of $2.7bn in NPV terms to absentee landlords. The welfare loss from displaced renters is about $0.01bn. On the other hand, the supply model estimates a gain of $0.3bn from home-sharing.

Second, the net impact on all housing market participants is positive. To include all housing market participants, one needs to take into account the impact on property owners, including both absentee landlords and owner-occupiers. Absentee landlords are better off because of increased rents: they gain $2.7bn because the welfare loss suffered by renters from the rent channel is a direct transfer to existing property owners. Absentee landlords are also better off because of their host surplus: although I do not estimate the cost parameters directly for them, a back-of-the-envelope estimation based on their actual Airbnb revenue less the forgone rents results in a surplus of $0.7bn per year. Owner-occupiers are weakly better off because they can be viewed as paying rent to themselves: the increase in the user-cost of their home is offset by the increase in property value.

Lastly, the net welfare impact on society is also positive. A back-of-the-envelope approximation produces a substantial gain of $1.6bn for tourists, net of hotel losses.[18] Combined with the welfare impact on housing market participants in NYC, the total efficiency gain due to Airbnb from the social planner's perspective is positive and substantial. Despite an overall increase in efficiency, because 67% of the housing units in NYC are renter-occupied, the median person in the city is a renter and is made worse off by Airbnb in NYC.

As NYC proceeds to enforce restrictions on Airbnb reallocation,[19] the welfare transfer to property owners will be reversed. Incumbent renters and hotels will be better off, at the expense of erasing most tourist gains and host surpluses. However, the distributional analysis from the previous section suggests such restrictions will lead to a bigger reduction in housing costs for higher-income, educated, and white renters. As such, an Airbnb ban is not particularly effective at targeting housing affordability problems faced by lower-income or minority households.

5.3.3 Robustness and Limitations

This section discusses the robustness of the results against a number of key simplifying assumptions and clarifies the limitations that could affect the interpretation of the findings.

A simplifying assumption of the model is that the long-term rental demand is unchanged after the entry of Airbnb. In theory, households should take into account their expected host surpluses in the short-term rental market when they make long-term housing choices. Such a joint model would predict that households with low costs of sharing will move closer toward neighborhoods that are popular among tourists. However, the empirical impact of such additional shifts in the long-term rental demand is likely muted because the average cost of home-sharing turns out to be high, with fewer than 1% of residents becoming hosts in the data.

Relatedly, assuming long-term rental demand is unchanged also implicitly abstracts away the impact of Airbnb on neighborhood amenities. Conceptually, it can be an important issue if tourism is a city's dominant industry. For example, Almagro and Domínguez-Iino (2021) show that over 20% of the rental housing stock in certain parts of Amsterdam have become Airbnbs, which resulted in a proliferation of tourism amenities such as restaurants and bars and a decline of amenities valuable to long-term residents. By contrast, in NYC, even the most active neighborhood, Chelsea, sees a reallocation of 3.4% of its housing stock to Airbnb, where the average level of reallocation is only 0.68%. Thus, simplifying away the impact of Airbnb on endogenous neighborhood amenities is unlikely to alter the main findings of the

[18]The approximation is based on an average consumer surplus of $42 per room night and a decline of 5% in hotel profitability (Farronato and Fradkin, 2021). Hotel usage data are from Smith Travel Research.

[19]Under the Multiple Dwelling Law, short-term rentals of Class A properties are not allowed unless its permanent resident is present. If fully enforced, it effectively rules out housing reallocation by absentee landlords.

paper substantially.

Next, both the long-term rental demand and the short-term rental supply model are static. In the long-term rental model, I have assumed away moving costs. In the short-term rental model, I have abstracted away adoption costs, where resident hosts may exert a one-time effort to list on Airbnb. As such, the estimated welfare impact should be thought of as an approximation where the impact of fixed costs is smoothed out.

Lastly, rent regulations are not modeled explicitly. Rent stabilization is a prominent feature of the NYC housing market, accounting for 32% of housing. Interestingly, about a third of them charges "preferential rents" allowing landlords to bypass rent increase limits set by the city.[20] However, it still implies about 20% of housing cannot freely adjust its rent beyond the annual city limit of 1.5%. Even though the average estimated price impact of Airbnb on rent does not exceed 1.0%, the omission of explicit rent stabilization can still create a bias when other market forces are generating upward price pressure at the same time.

Although this paper does not estimate a full model of residential demand with a rent-stabilized segment, taking the estimated coefficients as given, a back-of-the-envelope calculation suggests rent regulation causes the impact of Airbnb to be further concentrated on remaining market-price units, which tend to be occupied by higher-income tenants. Thus, the estimated distributional differences likely provide a lower bound, meaning higher-income renters may have shouldered an even larger welfare loss.

6 Conclusion

In this paper, I build and estimate the first integrated housing model to evaluate the aggregate and distributional impact of the home-sharing platform Airbnb on NYC renters.

The structural model highlights two important features of the sharing economy: the reallocation of resources and the increased utilization of resources. In a supply-constrained market, the reallocation of housing to Airbnb leads to an increase in equilibrium rents across all housing units in the city, not just for the specific units removed. It results in a widespread loss for all renters, aggregating to a transfer of $2.7bn to property owners. Moreover, the heterogeneity allowed in the structural model shows more significant losses are shouldered by renters who are higher-income, more educated, and white, because they tend to value housing and neighborhood amenities that are highly desirable to short-term visitors as well.

The utilization channel allows residents to provide short-term rental services in their existing homes. The estimated supply model finds the cost of home-sharing remains high for most households, suggesting the median host gain is immaterial and host surpluses aggregate to only $0.3bn. Nonetheless, a small fraction of people with particularly low costs of sharing obtain substantial surpluses, including a few enterprising low-income families taking advantage of peak

[20] *Source*: New York State's Division of Housing and Community Renewal.

short-term rental demand.

As NYC proceeds to enforce restrictions against Airbnb reallocation, it will likely reduce housing costs for all renters, especially higher-income and educated renters. However, it will be welfare-reducing overall, and it does not address the fundamental problems of housing affordability created by housing supply constraints. It is a challenge that afflicts not only NYC but also many other productive cities across the United States.

More broadly, the welfare impact of the sharing economy arises when innovation reduces existing market and regulatory frictions (Buchak, Matvos, Piskorski and Seru, 2018; Koijen and Yogo, 2016). The reallocation of housing facilitated by Airbnb effectively removes the trade barrier between the long-term and short-term rental markets, reducing the price wedge and increasing welfare.

The utilization aspect of the sharing economy is an intriguing technological feature because it allows enterprising residents, including low-income ones, to engage in business activities that would otherwise be costly to start. These peer-to-peer platforms reduce the barriers to entry for enterprising individuals. The growth and behavior of a whole class of platform entrepreneurs is an exciting avenue for future research.

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.

Albouy, David, Gabriel Ehrlich, and Yingyi Liu. 2016. "Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis." *NBER Working Paper.*

Almagro, Milena and Tomás Domínguez-Iino. 2021. "Location Sorting and Endogenous Amenities: Evidence from Amsterdam *."

Autor, David H., Christopher J. Palmer, and Parag A. Pathak. 2019. "Ending Rent Control Reduced Crime in Cambridge." *AEA Papers and Proceedings*, *109*, 381–384.

Autor, David H., Christopher Palmer, and Parag A. Pathak. 2014. "Housing Market Spillovers: Evidence from the End of Rent Control in Cambridge Massachusetts." *Journal of Political Economy*, *122* (3): 661–717.

Banzhaf, H. Spencer and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism." *American Economic Review*, *98* (3): 843–863.

Barron, Kyle, Edward Kung, and Davide Proserpio. 2021. "The effect of home-sharing on house prices and rents: Evidence from Airbnb." *Marketing Science*, *40* (1): 23–47.

Bass, Frank M. 1969. "A New Product Growth for Model Consumer Durables." *Management Science.*

Baum-Snow, Nathaniel and Lu Han. 2021. "The Microgeography of Housing Supply." *Working Paper*, (2010): 1–59.

Bayer, Patrick and Christopher Timmins. 2007. "Estimating Equilibrium Models for Sorting Across Destinations." *The Economic Journal*, *117*, 353–374.

Bayer, Patrick, Fernando Ferreira, and Robert Mcmillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy*, *115* (4).

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. "Migration and Hedonic Valuation: The Case of Air Quality." *Journal of Environmental Economics and Management*, *58* (1): 1–14.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2016. "A Dynamic Model of Demand For Houses and Neighborhoods." *Econometrica*, *84* (3): 893–942.

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

Berry, Steven, James Levinsohn, and Ariel Pakes. 1999. "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy." *American Economic Review*, *89* (3): 400–430.

Berry, Steven, James Levinsohn, and Ariel Pakes. 2004. "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." *Journal of Political Economy*, *112* (1): 68–105.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, *114* (2): 577–599.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru. 2018. "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." *Journal of Financial Economics*, *130* (3): 453–483.

Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matějka, and Tobias Salz. 2020. "The Value of Time: Evidence from Auctioned Cab Rides." *Working Paper.*

Campbell, John, Stefano Giglio, and Parag Pathak. 2011. "Forced Sales and House Prices." *American Economic Review*, *53* (9): 1689–1699.

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

Chamberlain, Gary. 1987. "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions." *Journal of Econometrics*, *34* (3): 305–334.

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2016. "Using Big Data to Estimate Consumer Surplus : The Case of Uber." *NBER Working Paper.*

Cullen, Zoë and Chiara Farronato. 2021. "Outsourcing tasks online: Matching supply and demand on peer-to-peer internet platforms." *Management Science*, *67* (7): 3985–4003.

Davis, Lucas W. 2011. "The Effect of Power Plants on Local Housing Values and Rents." *Review of Economics and Statistics*, *93* (4): 1391–1402.

Diamond, Rebecca, Tim McQuade, and Franklin Qian. 2019. "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco." *American Economic Review*, *109* (9): 3365–3394.

Dubé, Jean-Pierre, Jeremy T. Fox, and Che-Lin Su. 2012. "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation." *Econometrica*, *80* (5): 2231–2267.

Einav, Liran, Chiara Farronato, and Jonathan Levin. 2016. "Peer-to-Peer Markets." *Annual Review of Economics*, *8* (1): 615–635.

Farronato, Chiara and Andrey Fradkin. 2021. "The Welfare Effects of Peer Entry in the Accommodation Market: The Case of Airbnb." *American Economic Review (Conditional Acceptance).*

Favilukis, Jack Y, Pierre Mabille, and Stijn Van Nieuwerburgh. 2021. "Affordable Housing and City Welfare." *Review of Economic Studies (Conditional Acceptance).*

Ferreyra, Maria Marta. 2007. "Estimating the Effects of Private School Vouchers in Multidistrict Economies." *American Economic Review*, *97* (3): 789–817.

Frechette, Guillaume R., Alessandro Lizzeri, and Tobias Salz. 2019. "Frictions in a Competitive, Regulated Market Evidence from Taxis." *American Economic Review.*

Galiani, Sebastian, Alvin Murphy, and Juan Pantano. 2015. "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment." *American Economic Review*, *105* (11): 3385–3415.

Ganong, Peter and Daniel Shoag. 2017. "Why Has Regional Income Convergence in the U. S. Declined?" *Journal of Urban Economics*, *102*, 76–90.

Garcia-López, Miquel Àngel, Jordi Jofre-Monseny, Rodrigo Martínez-Mazza, and Mariona Segú. 2020. "Do short-term rental platforms affect housing markets? Evidence from Airbnb in Barcelona." *Journal of Urban Economics*, *119* (June 2019).

Greenstone, Michael and Justin Gallagher. 2008. "Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program." *Quarterly Journal of Economics*, (August): 951–1003.

Gyourko, Joseph. 2009. "Housing Supply." *Annual Review of Economics*, *1*, 295–318.

Gyourko, Joseph and Raven Molloy. 2015. "Regulation and Housing Supply." *Handbook of Regional and Urban Economics*, *5*, 1289–1337.

Horn, Keren and Mark Merante. 2017. "Is Home Sharing Driving up Rents? Evidence from Airbnb in Boston." *Journal of Housing Economics*, *38*, 14–24.

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics.*

Klaiber, H. Allen and Daniel J. Phaneuf. 2010. "Valuing Open Space in a Residential Sorting Model of the Twin Cities." *Journal of Environmental Economics and Management*, *60* (2): 57–77.

Koijen, Ralph S. J. and Motohiro Yogo. 2016. "Shadow Insurance." *Econometrica*, *84* (3): 1265–1287.

Koster, Hans R.A., Jos van Ommeren, and Nicolas Volkhausen. 2021. "Short-term rentals and the housing market: Quasi-experimental evidence from Airbnb in Los Angeles." *Journal of Urban Economics*, *124* (April): 103356.

McFadden, Daniel. 1978. "Modeling the Choice of Residential Location." *Spatial Interaction Theory and Planning Models*, pp. 72–77.

Murphy, Alvin. 2018. "A dynamic model of housing supply." *American Economic Journal: Economic Policy*, *10* (4): 243–267.

Nevo, Aviv. 2000. "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry." *The RAND Journal of Economics*, *31* (3): 395.

Nevo, Aviv. 2001. "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, *69* (2): 307–342.

Reynaert, Mathias and Frank Verboven. 2014. "Improving the Performance of Random Coefficients Demand Models: The Role of Optimal Instruments." *Journal of Econometrics*, *179* (1): 83–98.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *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.

Timmins, Christopher. 2007. "If You Cannot Take the Heat, Get out of the Cerrado... Recovering the Equilibrium Amenity Cost of Nonmarginal Climate Change in Brazil." *Journal of Regional Science*, *47* (1): 1–25.

Tra, Constant I. 2010. "A Discrete Choice Equilibrium Approach to Valuing Large Environmental Changes." *Journal of Public Economics*, *94* (1-2): 183–196.

Valentin, Maxence. 2021. "The Effects of Regulating the Housing Short-Term Rental Market: Evidence From New Orleans." *Real Estate Economics*, *49.*

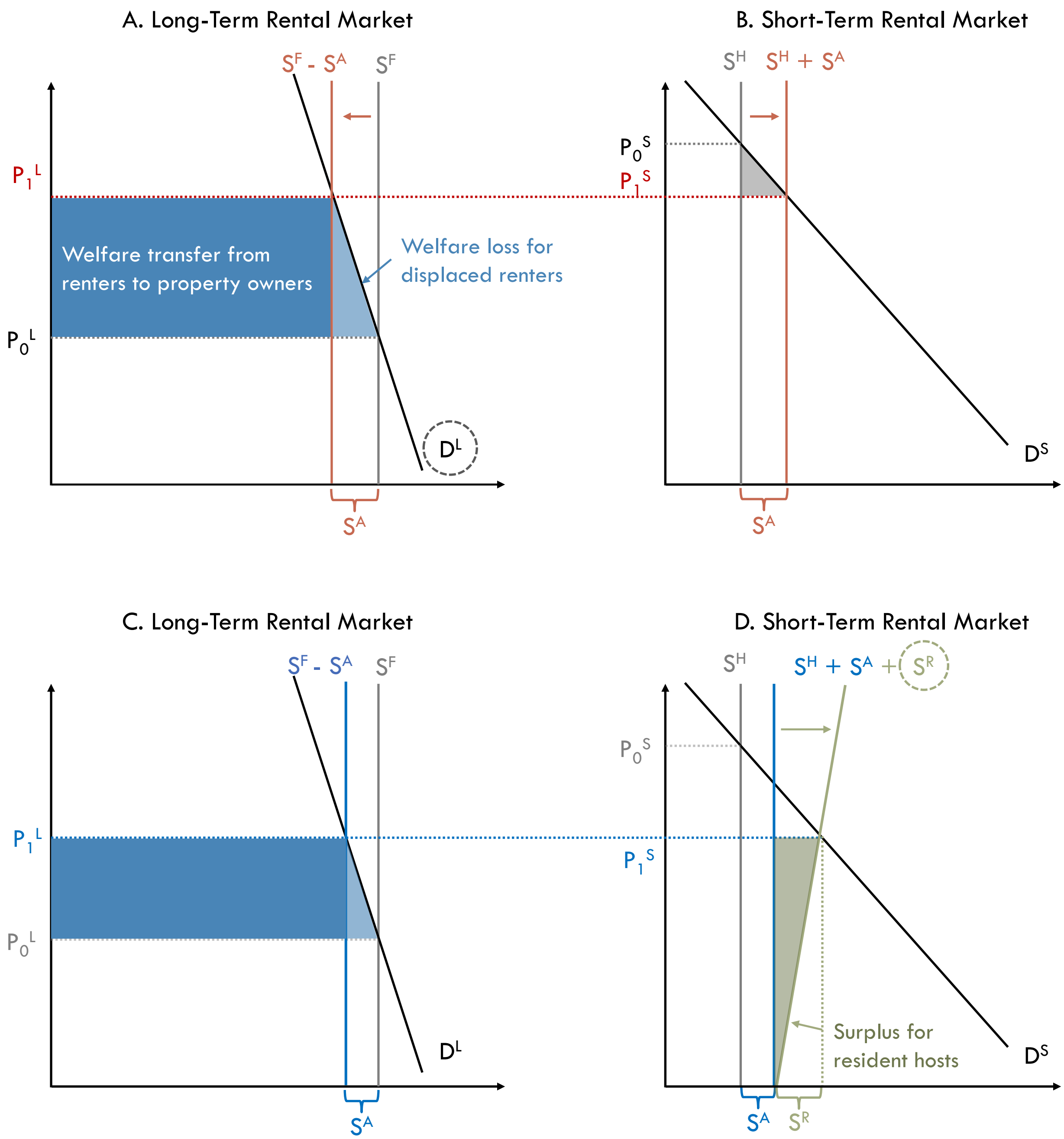
Wachsmuth, David and Alexander Weisler. 2018. "Airbnb and the Rent Gap: Gentrification through the Sharing Economy." *Environment and Planning A: Economy and Space*, *50* (6): 1147–1170.

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

7 Figures

Figure 1: An Illustrative Model of an Integrated Housing Market

Panels A and B illustrate the market equilibrium when Airbnb allows absentee landlords to reallocate from the long-term rental to the short-term rental market. Panels C and D illustrate the market equilibrium when Airbnb also allows residents to share their homes with directly short-term visitors.



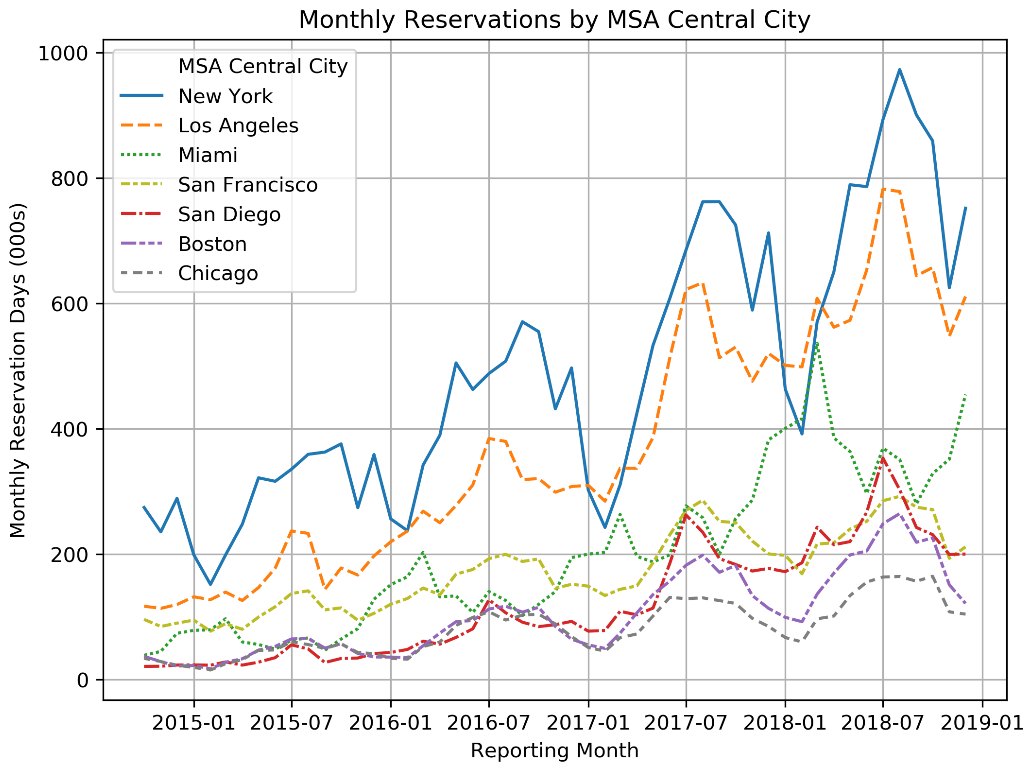


Figure 2: Growth of Airbnb across the US: The time series plot shows the rapid growth in the number of monthly reservations across select MSAs. New York is the largest metropolitan market for Airbnb in the U.S. Between 2015 and 2018, the number of reservations quadrupled.

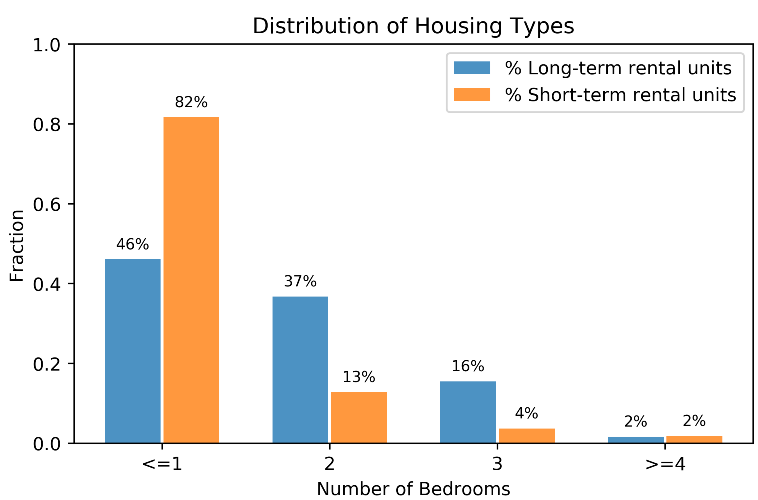


Figure 3: A comparison of housing types available on the long-term and short-term rental markets, respectively. Notably, smaller housing types are much more prevalent in the short-term rental market than in the long-term rental market.

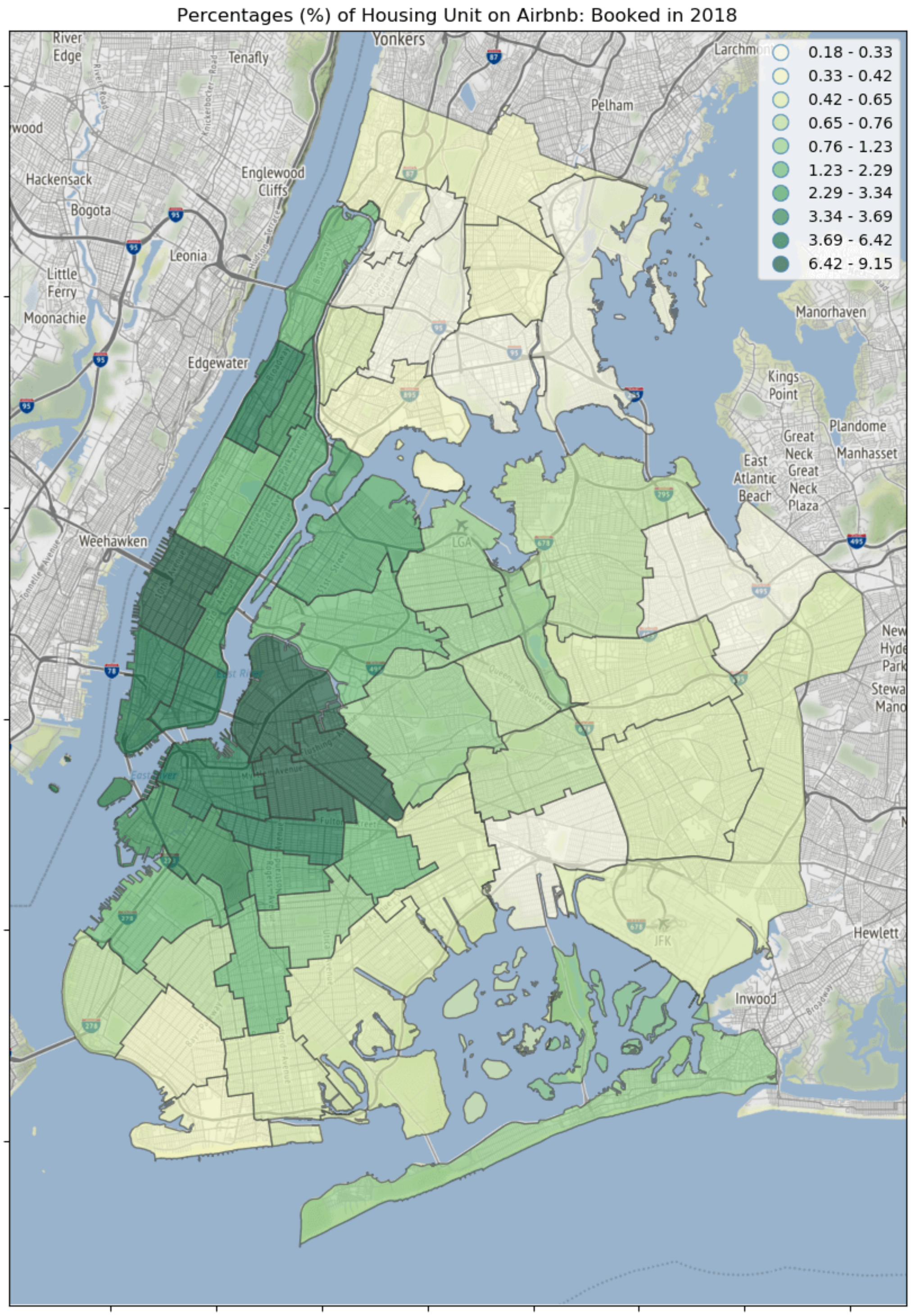


Figure 4: Percentage (%) of Housing Units on Airbnb
(Having at Least One Reservation in 2018)

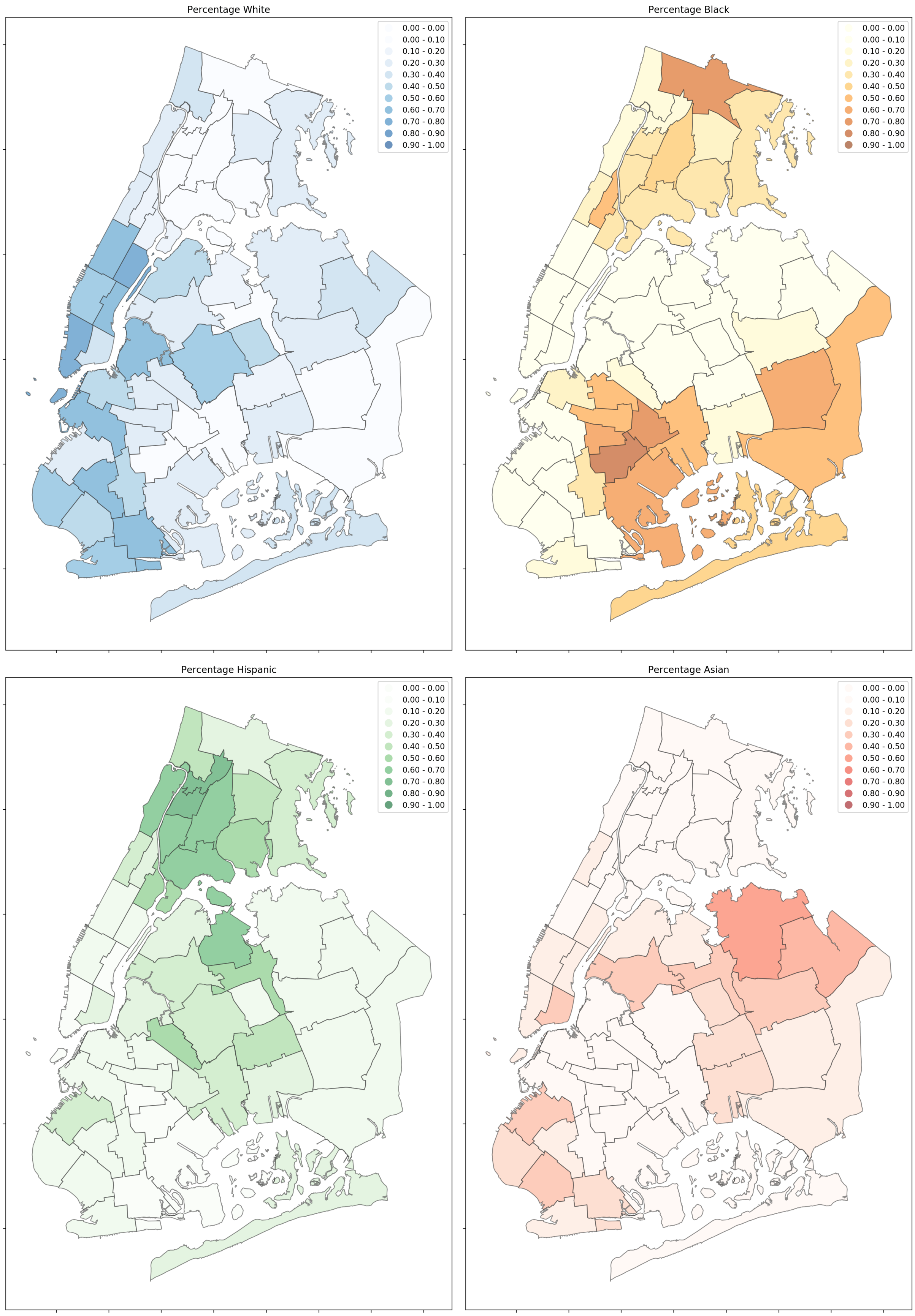
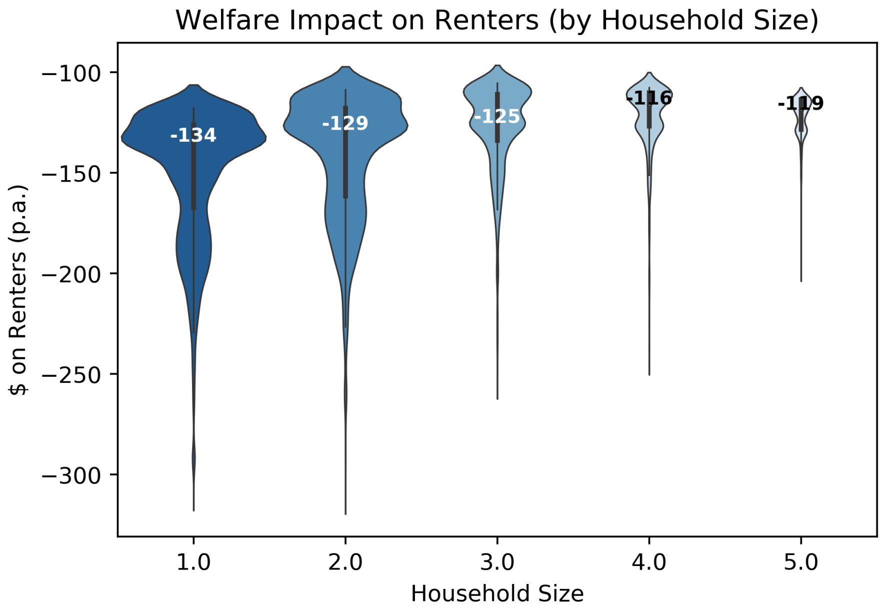
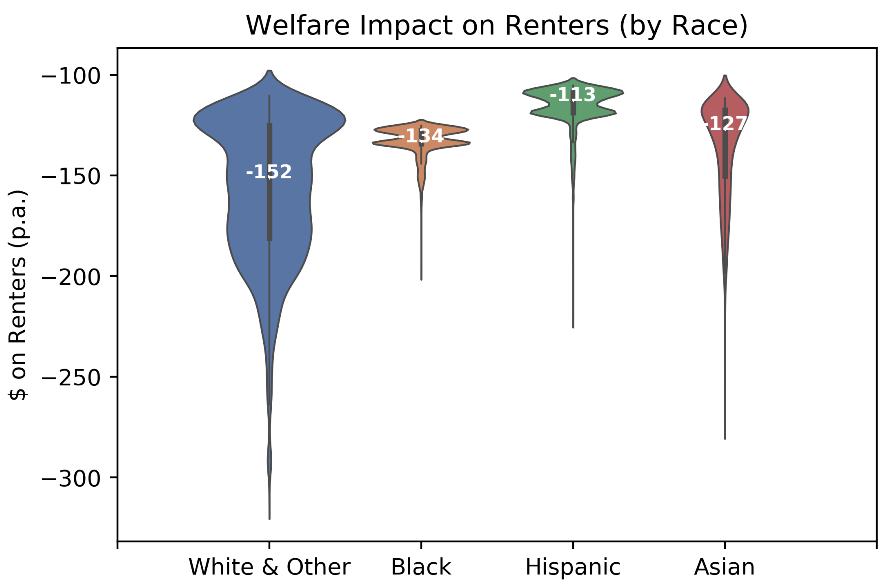


Figure 5: Race and Ethnicity across Neighborhoods

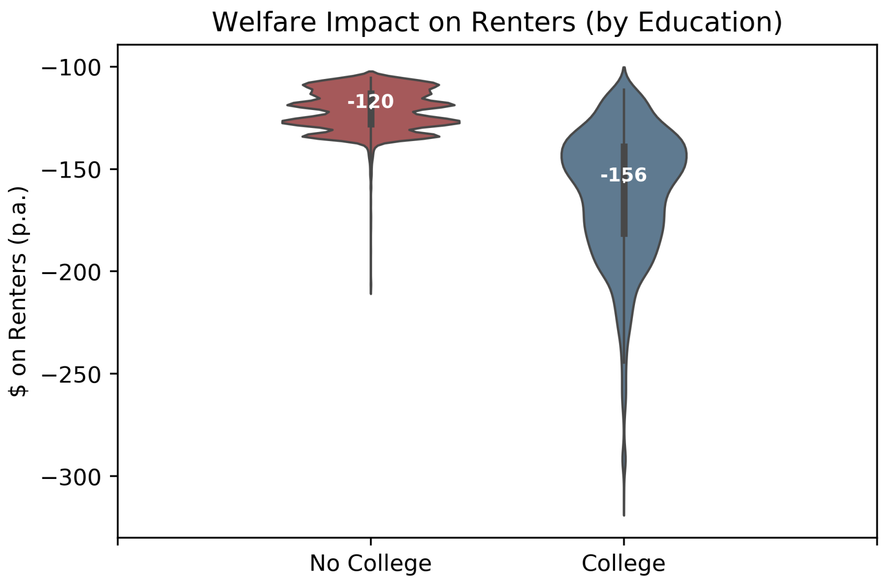


(a) Welfare impact via the rent channel by household size

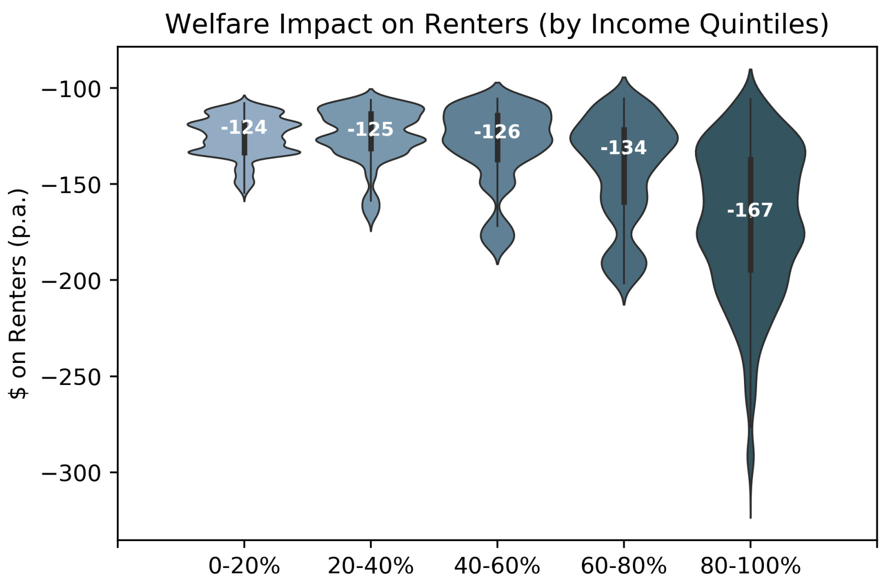


(b) Welfare impact via the rent channel by race and ethnicity

Figure 6: The welfare impact of Airbnb on renters *via the rent channel* by various demographic characteristics. Smaller households suffer more on average. In terms of race and ethnicity, white renters suffer more. The labeled numbers indicate the category median. The width of each kernel density plot corresponds to the frequency of the category in the population. The mini-box plot in the center indicates the inter-quartile range using the thick black line, and 1.5x the inter-quartile range using the thin black line.



(a) Welfare impact via the rent channel by education



(b) Welfare impact via the rent channel by household income

Figure 7: The welfare impact of Airbnb on renters *via the rent channel* by various demographic characteristics. Educated and higher-income renters suffer more on average. The labeled numbers indicate the category median. The width of each kernel density plot corresponds to the frequency of the category in the population. The mini-box plot in the center indicates the inter-quartile range using the thick black line, and 1.5x the inter-quartile range using the thin black line.

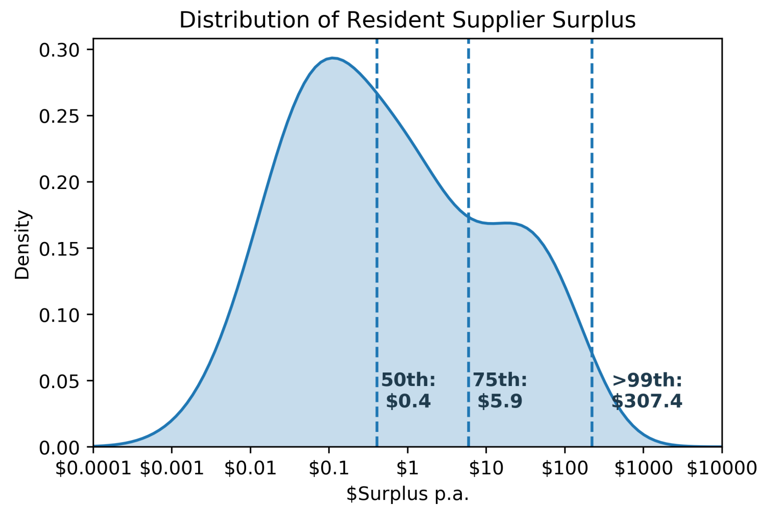


Figure 8: The kernel density plot of the supplier surplus through Airbnb, namely, the welfare impact on renters via the *host* channel. Note this plot is over the logarithm of the surplus. For the median household, the surplus from being an Airbnb supplier is immaterial at only $0.4 p.a., reflecting the fact that most households do not participate in the sharing platform. The 70th percentile surplus is at $5.9 p.a. The 90th percentile surplus is at $45.6 p.a. However, in the far-right tail (>99%), the surplus is substantial at $307 p.a.

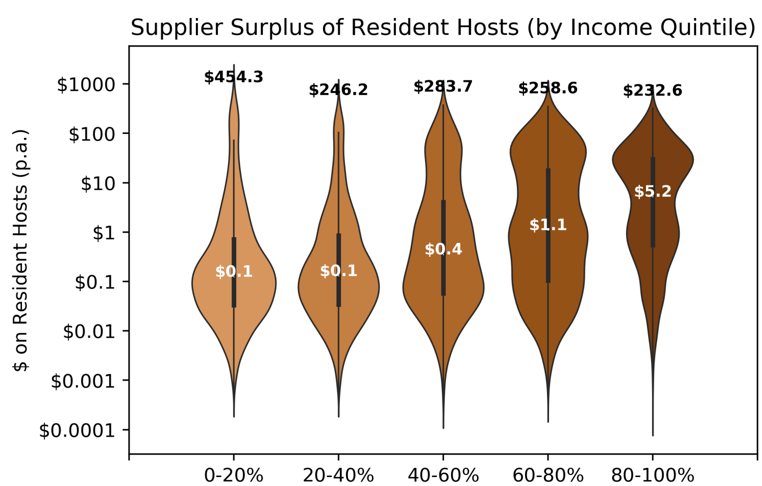


Figure 9: The welfare impact of Airbnb on renters via the *host* channel by income quintiles. Note this plot is over the logarithm of the surplus on the vertical axis. The numbers in the middle of the density plot represent the category median, whereas the numbers at the top represent the conditional average above the 99th percentile. The median surplus from home-sharing is immaterial across all income levels, but higher-income groups still have a higher mean. In the tail, the lowest-income quintile accrue larger benefits.

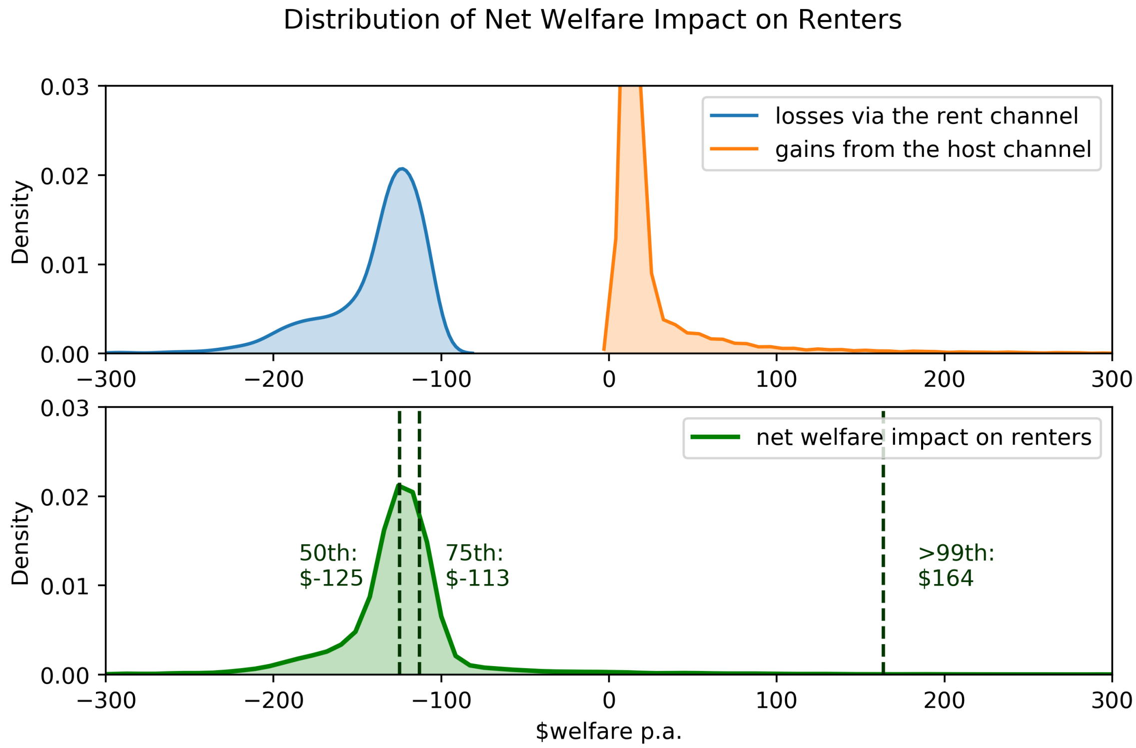


Figure 10: The net welfare impact of Airbnb on renters, combining losses from the rent channel and gains from the host channel.

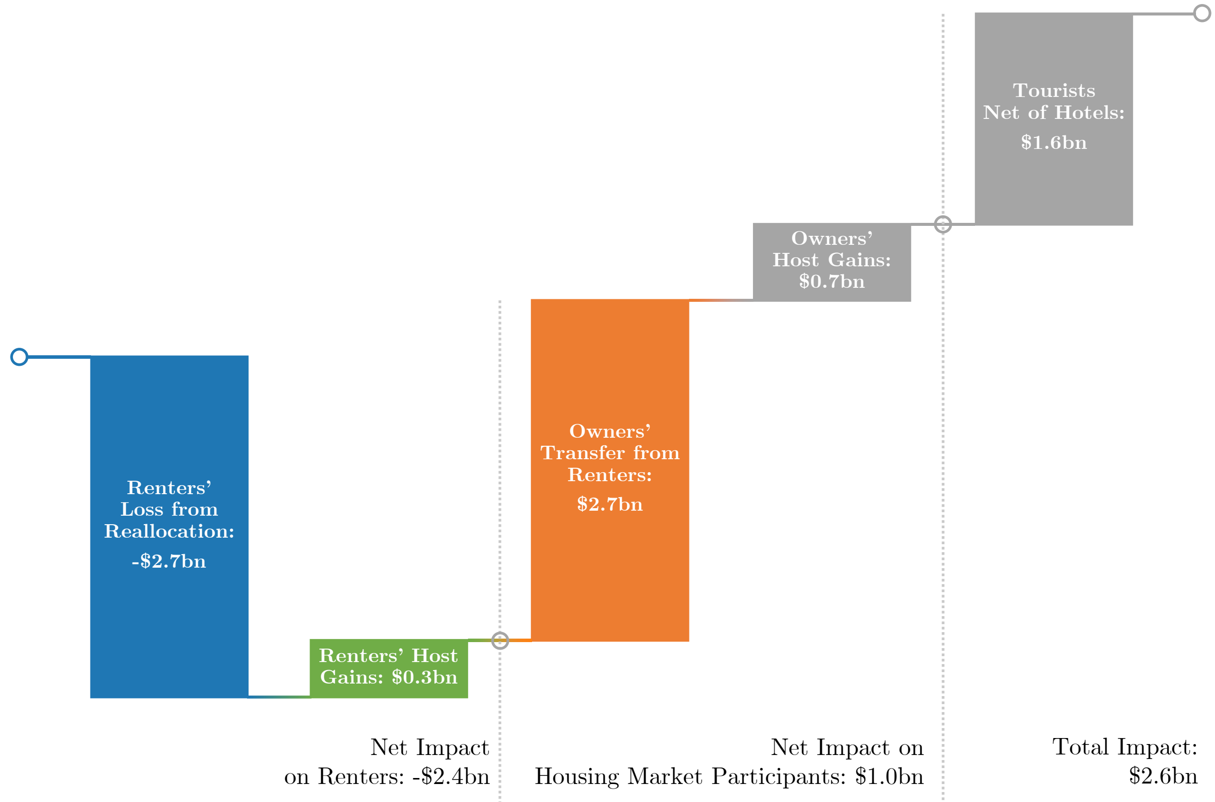


Figure 11: Aggregate welfare impact for all relevant participants. The colored bars represent estimates from the model. The gray bars represent back-of-the-envelope approximations.

8 Tables

Table 1: Breakdown of Airbnb Activity in NYC over 2018

Notes: This table summarizes the transactions on Airbnb in NYC over 2018. On Airbnb, a property is listed as one of three types: entire home/apt, private room, or shared room. Over 95% of the properties are listed as either "entire home/apt" or "private room," which are the focus of this paper. Overall, over 74,963 properties on Airbnb have experienced at least one reservation, accounting for 2.2% of the housing units in the city.

Listing Type	Bedroom(s)	Num Days Reserved (000s)	Num Properties (000s)	Average Daily Rate
Entire home/apt	All	3,204	38	$224
Entire home/apt	1	1,968	25	$178
Entire home/apt	2	838	9	$257
Entire home/apt	3	289	3	$350
Entire home/apt	4	82	1	$530
Private room	-	2,526	34	$86
Shared room	-	128	2	$59
Total		5,858	75	$156

Table 2: Single-Variate Regression of Airbnb Penetration on Neighborhood Characteristics

The *dependent variable* is the log of Airbnb reallocation.

	(1)	(2)	(3)	(4)	(5)	(6)
Pct White	2.834***					
	(0.780)					
Pct Black		-0.336				
		(0.822)				
Pct Hispanic			-3.438***			
			(0.852)			
Pct Asian				-0.211		
				(1.493)		
Pct College					5.244***	
					(0.784)	
ln(Median Income)						1.803***
						(0.382)
N	52	52	52	52	52	52

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Parameter Estimates for the Long-Term Rental Demand Model

Notes: This table provides the linear coefficients estimated for the long-term rental demand model. The *dependent variable* is the mean utility of each housing type δ_h^L estimated from *Step 1* using individual-level data. Column (2) assumes exogenous prices. Column (2) shows the IV regression where I use the alternative equilibrium price vector at $\xi^L = 0$ as the instrument for the observed prices. In Column (3), I show the WTP for each of the housing and neighborhood attributes for the average household.

	(1) OLS	(2) Instrumented	(3) ($) WTP Mo.
Monthly Rent ($k)	0.0213	-2.044***	
	(0.0341)	(0.609)	
One-Bedroom	0.425***	0.929***	454.5***
	(0.0447)	(0.188)	(78.2)
Two-Bedroom	0.528***	1.325***	648.2***
	(0.0465)	(0.280)	(93.5)
Three-Bedroom	0.271***	1.392***	681.0***
	(0.0555)	(0.393)	(76.7)
Four-Bedroom	-0.179***	0.904*	442.3*
	(0.0668)	(0.505)	(162)
Built After 1980	-0.114***	0.139	68.2
	(0.0402)	(0.145)	(60.9)
Built 1940 to 80	-0.00917	-0.242**	-118.4**
	(0.0337)	(0.105)	(43.9)
5+ Units	0.00182	-0.209**	-102.3**
	(0.0282)	(0.0974)	(41.2)
Commuting Time (Std)	0.119***	-0.782***	-382.6***
	(0.0215)	(0.279)	(28.2)
Inside NYC	-1.026***	2.536**	1240.7***
	(0.0683)	(1.036)	(147)
N	1050	1050	1050

Table 4: Parameter Estimates for the Short-Term Rental Supply Model

Notes: This table provides the estimated parameters for the short-term rental supply by resident hosts, using the MPEC procedure. The standard errors are clustered at the neighborhood level. In columns (1) and (2), prices are assumed to be exogenous and not instrumented. In columns (3) and (4), the total hotel bookings in the city (lagged by seven years) are used as the instrument for prices in the short-term rental market. The instrument is strong, with an F-stat of 25.4 in column (4) when controlling for month FE, day-of-week FE and holiday FE (including Christmas and New Year's Eve) in the corresponding linear specification. Column (5) provides the costs in dollar terms (per diem). Overall, the average cost to host is high, and the low-cost suppliers are those with a college degree, young, and have no children. The average price elasticity is 5.96. Because the interaction between household income and price is negative, low-income households are more elastic, averaging 6.70 for one-standard-deviation lower income.

		(1) Naïve	(2) Naïve	(3) IV	(4) IV	(5) ($) per diem
Linear Coef.	*Non-Linear Coef.*					
Price		0.006	0.007	0.052	0.056	
		(0.002)	(0.001)	(0.002)	(0.002)	
	x ln(income)	-0.018	-0.018	-0.018	-0.011	
		(0.001)	(0.002)	(0.003)	(0.006)	
Cost		15.44	15.51	22.07	21.36	224.3
		(0.10)	(0.09)	(0.12)	(0.11)	(12.7)
	x Has College	-1.17	-2.55	-3.47	-3.27	-58.9
		(0.68)	(0.24)	(0.27)	(0.25)	(4.8)
	x Has Children	2.40	2.58	1.95	2.60	46.7
		(0.42)	(0.36)	(0.53)	(0.44)	(8.1)
	x Age (yr)	0.094	0.093	0.091	0.097	1.8
		(0.005)	(0.005)	(0.006)	(0.006)	(0.1)
	x ln(income)	0.24	-0.14	-0.39	-0.29	-5.1
		(0.09)	(0.13)	(0.26)	(0.48)	(8.7)
Quad. Time		Yes	Yes	Yes	Yes	
Month FE		No	Yes	No	Yes	
Day of Week FE		No	Yes	No	Yes	
Holiday FE		No	Yes	No	Yes	
N		75,895	75,895	75,895	75,895	

Table 5: Estimated Own- and Cross- Price Semi-Elasticities

Notes: Entry (i, j) corresponds to the percentage (%) of share changes in neighborhood j when the price of housing in neighborhood i increases by $1,000 per month. This table shows only a subset of the neighborhoods for illustration, whereas the full matrix is 52 x 52. For example, this table shows that when price increases in Chelsea, Clinton, & Midtown, households there will substitute away toward similar neighborhoods such as the Upper West Side. However, the substitution towards Forest Hills (an educated and predominantly white neighborhood) is much larger Jackson Heights (a less-educated and predominantly Hispanic neighborhood), even though they are located closely in Queens.

	The Bronx Hunts Point, Longwood & Melrose	*Manhattan* Central Harlem	*Manhattan* Upper West Side & West Side	*Manhattan* Chelsea, Clinton & Midtown	*Queens* Jackson Heights & North Corona	*Queens* Forest Hills & Rego Park
Hunts Point, Longwood & Melrose	(21.9)	0.55	0.09	0.15	0.32	0.20
Central Harlem	0.58	(23.7)	0.20	0.32	0.18	0.30
Upper West Side & West Side	0.10	0.20	(21.0)	1.02	0.10	0.69
Chelsea, Clinton & Midtown	0.12	0.21	0.70	(21.6)	0.12	0.71
Jackson Heights & North Corona	0.62	0.33	0.19	0.32	(18.8)	0.35
Forest Hills & Rego Park	0.15	0.21	0.54	0.81	0.14	(37.5)

Table 6: Net Welfare Impact by Household Income

Notes: This table compares the welfare impact of Airbnb on renters by household income quintiles. Overall, higher-income households experience larger losses on average and experience smaller gains in the tail.

Welfare Impact by Renter Income Quintile ($ p.a.)

Loss via the Rent Channel	Mean	Median	P25	P75	>P99
0-20%	-125	-124	-134	-118	-115
20-40%	-124	-125	-131	-113	-106
40-60%	-130	-126	-137	-114	-106
60-80%	-142	-134	-159	-122	-105
80-100%	-169	-167	-195	-137	-106

Gain via the Host Channel	Mean	Median	P25	P75	>P99
0-20%	12	0.1	0.0	0.7	454
20-40%	7	0.1	0.0	0.8	246
40-60%	14	0.4	0.1	4.0	284
60-80%	20	1.1	0.1	17.9	259
80-100%	24	5.2	0.5	30.0	233

The Net Welfare Impact	Mean	Median	P25	P75	>P99
0-20%	-114	-122	-131	-114	319
20-40%	-117	-118	-128	-111	101
40-60%	-116	-119	-130	-109	137
60-80%	-122	-125	-139	-112	109
80-100%	-146	-144	-174	-122	73

Table 7: Net Welfare Impact by Household Demographics

Notes: This table compares the median welfare impact of Airbnb on renters by household characteristics. Overall, educated and white renters experience greater losses.

	Median Impact $p.a.			Tail Impact $p.a.		
	Loss via Rent	Gain via Host	Net Impact	Loss via Rent	Gain via Host	Net Impact
Overall	-128	0.4	**-125**	-109	307	164
Education						
Without College	-120	0.1	**-120**	-105	16	-98
With College	-156	10.8	**-136**	-112	393	253
Race						
Asian	-127	0.7	**-119**	-112	381	245
Black	-134	0.2	**-129**	-125	227	85
Hispanic	-113	0.1	**-111**	-105	232	107
White & Other	-152	1.9	**-130**	-111	326	179

A Appendix For Online Publication

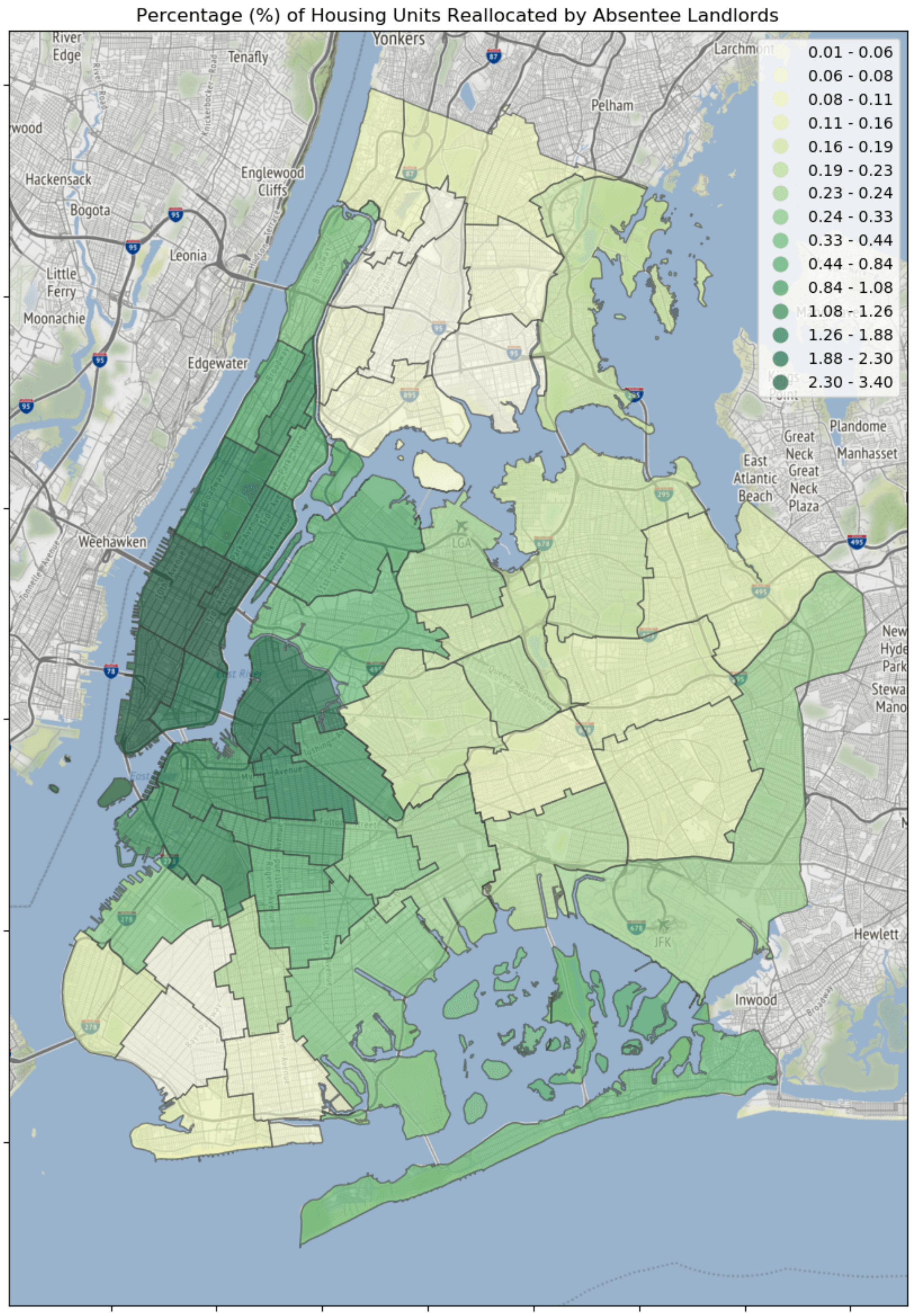


Figure A.12: Percentage (%) of Rental Housing Units Reallocated
(Available on Airbnb for 180+ Days in 2018)

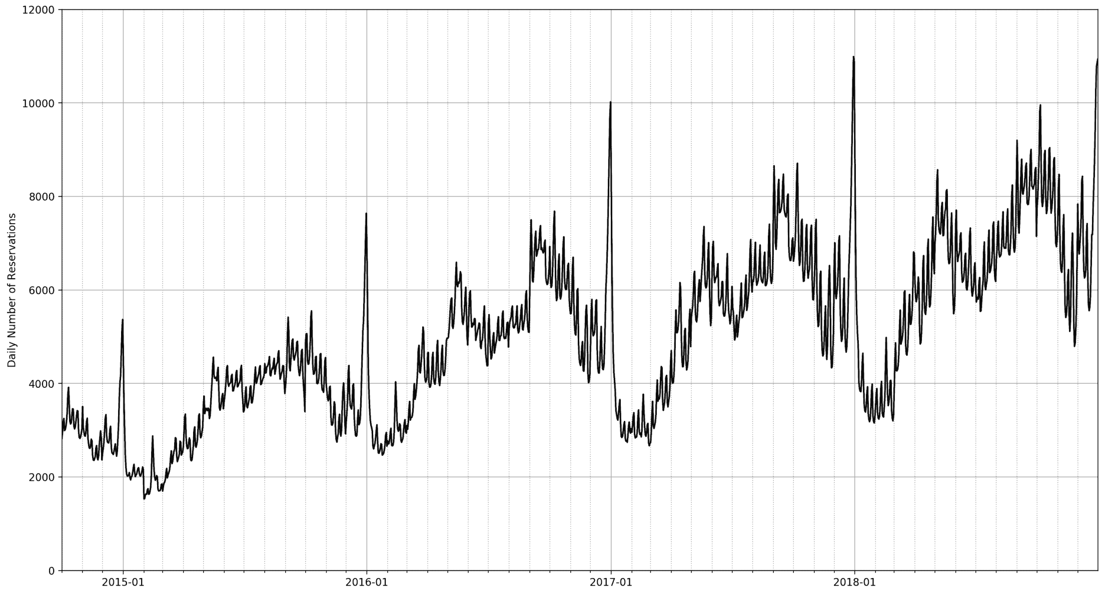


Figure A.13: The daily number of reservations of private rooms sold on Airbnb in NYC.

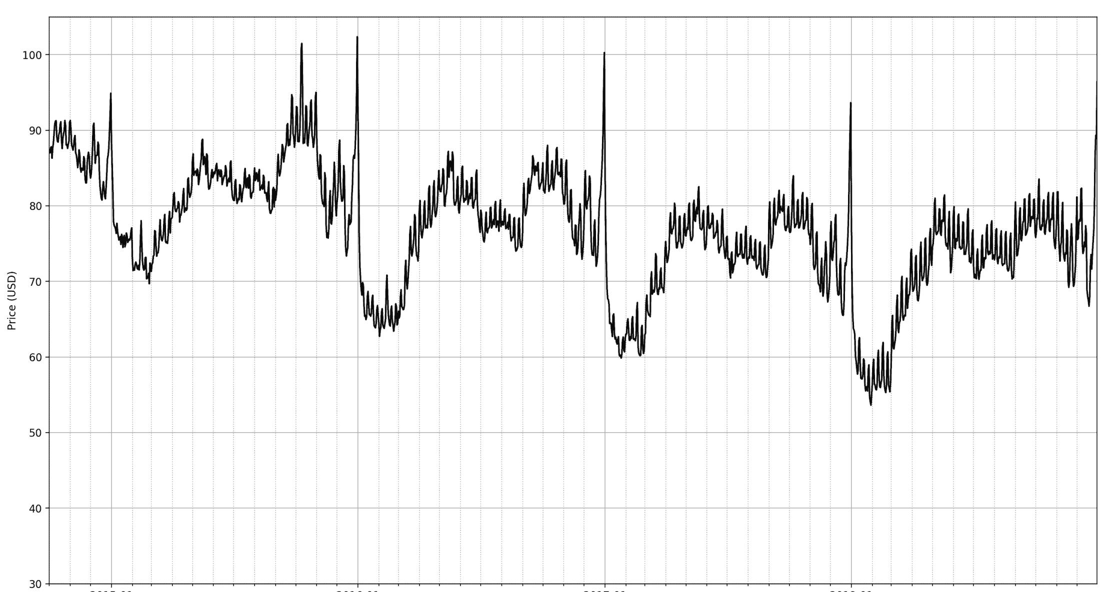


Figure A.14: The daily average price of private rooms sold on Airbnb in NYC.

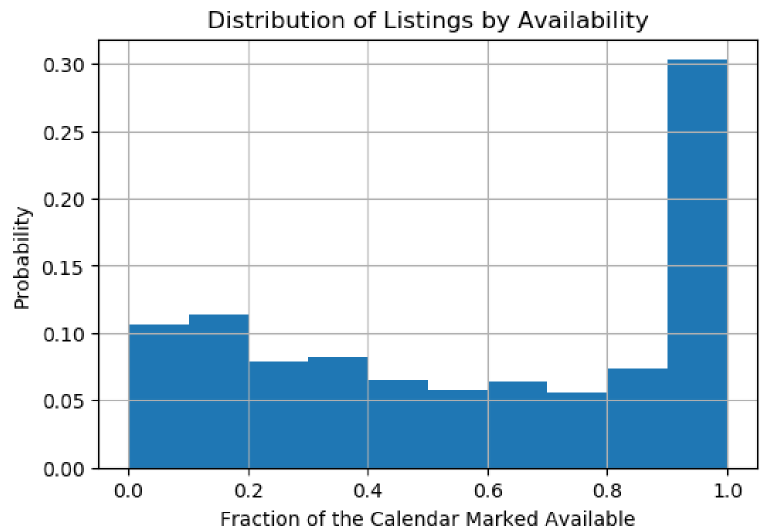


Figure A.15: Distribution of Airbnb Listings by Calendar Availability

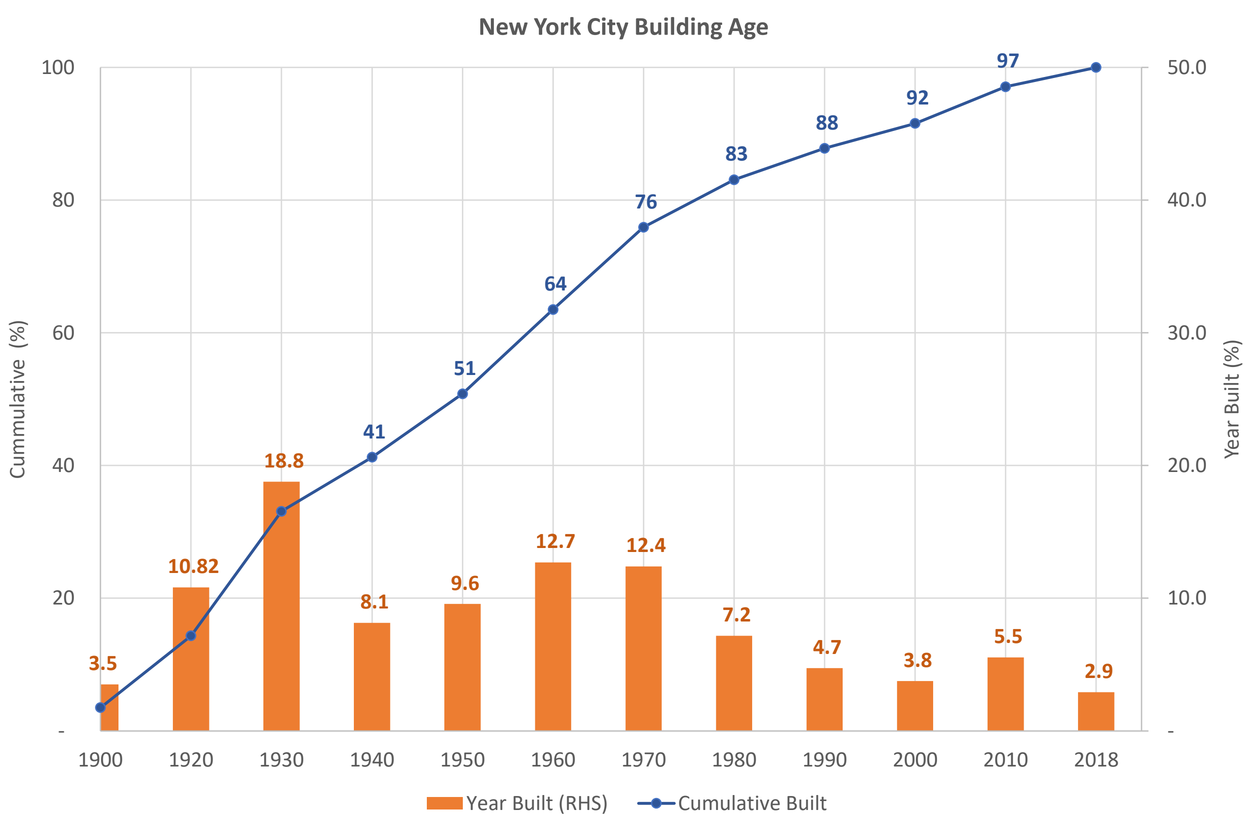


Figure A.16: Building Age of NYC Housing Units

Data based on ACS 2017 and New York City Housing and Vacancy Survey (2017). Over 40% of the housing units were built prior to 1940. Housing construction since the 1980s has remained depressed. 88% of the housing were built prior to 1990. Only 2.9% of the housing stock was built post-2010, whereas 3.5% of the units were built prior to 1900.

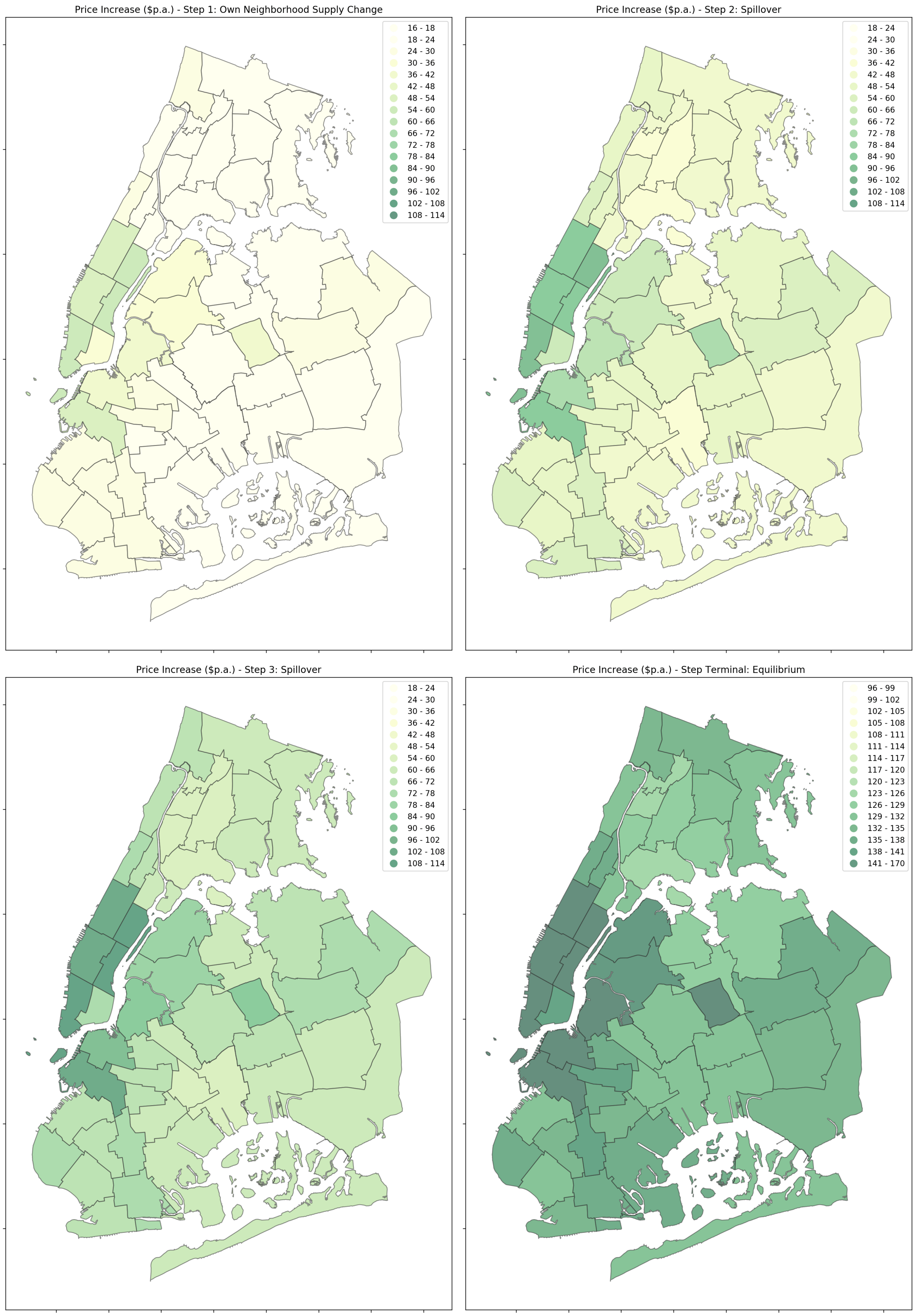


Figure A.18: The equilibrium effects of supply restrictions (using successive steps of best responses to illustrate the equilibrating process)

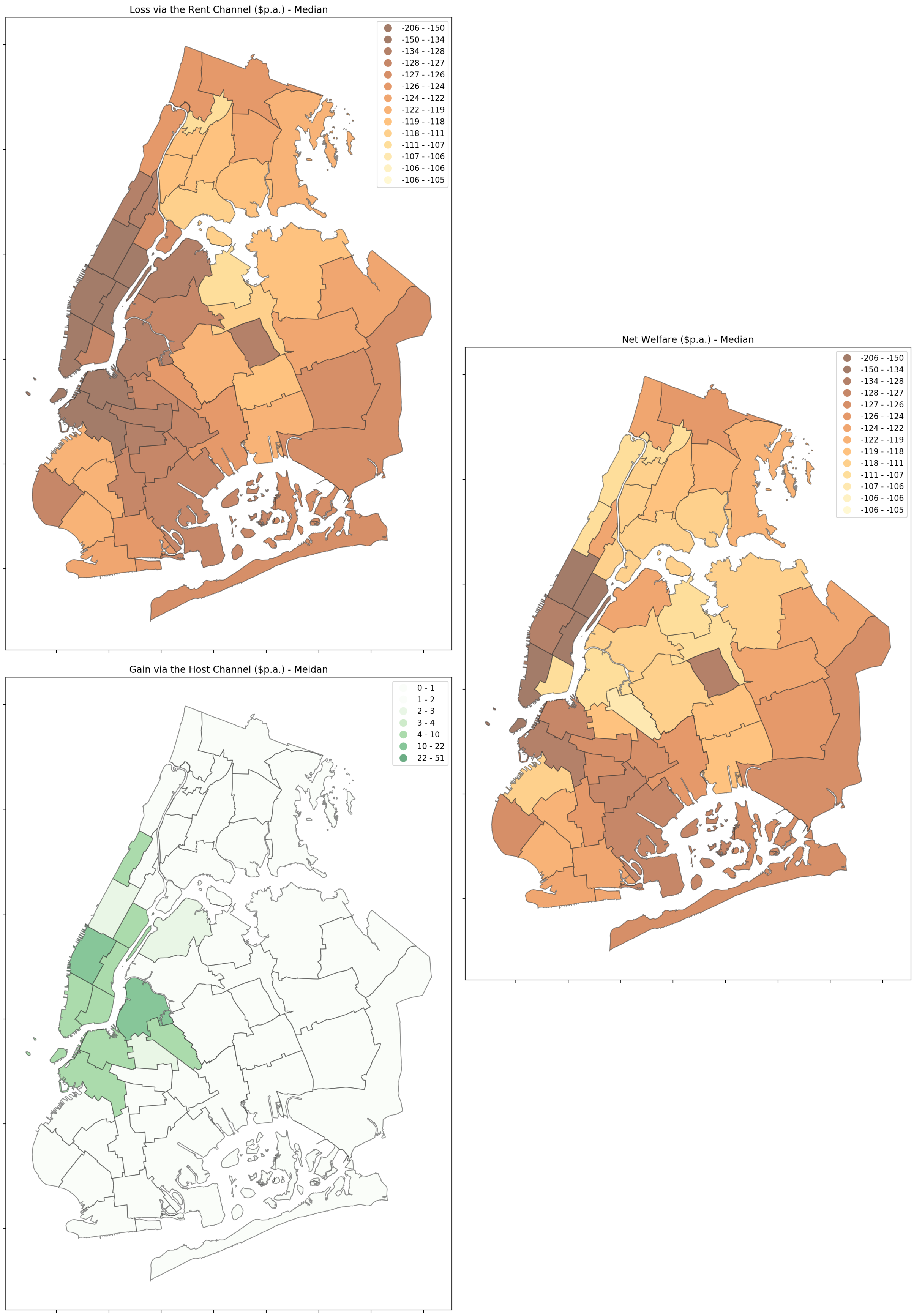


Figure A.19: Net Welfare Impact by Neighborhood (Median): The median renter in all neighborhoods experiences a loss due to Airbnb as the rent channel dominates.

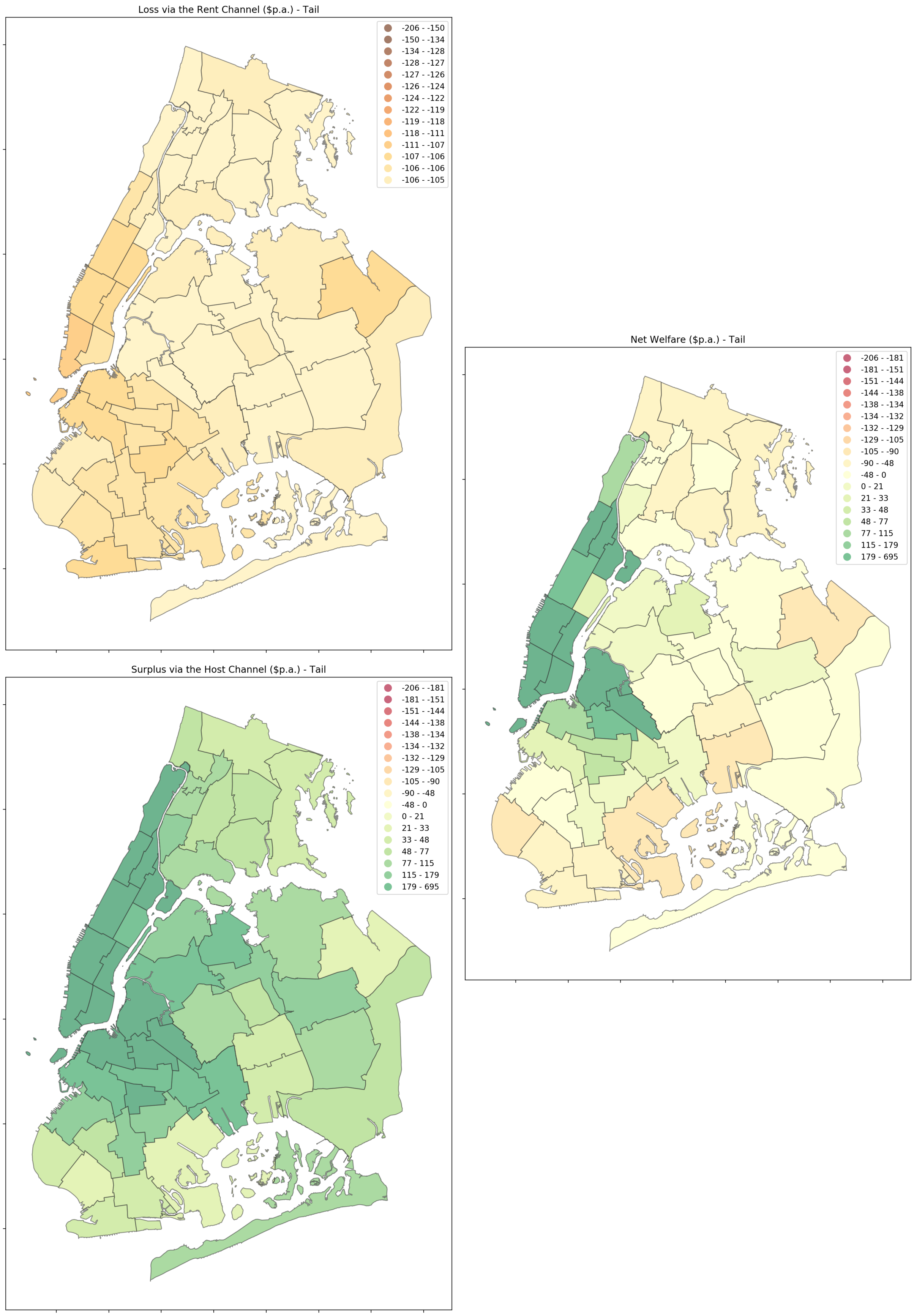


Figure A.20: Net Welfare Impact by Neighborhood (Right Tail): For a small fraction of the renters, the gains from home-sharing can easily make up for the rent losses, especially in neighborhoods with high short-term rental demand.

Table A.8: Income and Demographic Characteristics of NYC Renters

Notes: Based on the American Community Survey 2017 1-Yr estimates, the table summarizes the income and demographic characteristics of NYC renters across all boroughs, excluding Staten Island. Note that there are substantial variations in household income, education, race and ethnicity across the boroughs.

	All	**The Bronx**	**Brooklyn**	**Manhattan**	**Queens**
Annual Household Income ($k)					
0-20%	8	6	7	10	12
20-40%	24	17	22	31	32
40-60%	47	31	44	67	52
60-80%	83	54	78	126	83
80-100%	164	99	156	265	142
Education					
With College	38%	18%	36%	58%	32%
Race / Ethnicity					
White (non-Hispanic)	32%	7%	35%	47%	28%
African American	27%	36%	35%	16%	17%
Hispanic	32%	58%	21%	25%	31%
Asian	11%	2%	8%	12%	22%

Table A.9: Heterogeneous Parameter Estimates for the Long-Term Rental Demand (Step 1)

Notes: This table provides the heterogeneous coefficients estimated from *Step 1* of the long-term rental demand model. The coefficients on housing attributes are presented as WTP in terms of monthly dollars. The coefficient on price is applied to the monthly rent ($k). The omitted categories are studios, those built prior to 1940, and buildings with fewer than 5 units. Neighborhood characteristics are standardized to variance 1. The standard errors for the ratios are computed using the delta method. I highlight the most significant demographic characteristics for each housing and neighborhood attribute.

WTP ($)	Ln Income	HH Size	Black	Hispanic	Asian	College
Housing Characteristics						
One Bedroom	75.8	255.1	-40.3	-67.0	-177.1	-86.4
	(31.2)	(78.6)	(55.1)	(54.5)	(75.6)	(48.3)
Two Bedroom	59.1	520.5	98.1	-24.2	-273.4	-212.6
	(28.8)	(156.4)	(63.8)	(54.1)	(101.0)	(76.9)
Three Bedroom	32.4	717.6	143.8	-37.2	-329.1	-214.6
	(28.7)	(214.9)	(80.7)	(64.8)	(121.0)	(82.6)
Four Bedroom	85.0	884.9	-206.5	-328.3	-297.9	-244.9
	(66.7)	(266.0)	(172.3)	(171.9)	(170.2)	(134.0)
Built After 1980	22.4	-35.9	157.9	55.1	42.3	-27.8
	(14.7)	(13.7)	(56.4)	(33.3)	(36.7)	(25.0)
Built 1940-1980	-102.6	6.1	137.0	125.4	-84.8	58.6
	(34.3)	(10.1)	(55.6)	(50.6)	(47.5)	(32.6)
5+ Units	9.7	58.6	-3.8	-118.4	-110.6	-6.7
	(9.5)	(18.6)	(25.9)	(41.3)	(40.1)	(17.6)
Nbhd. Characteristics						
Pct Black (Std)	56.7	-47.9	774.4	330.6	272.1	67.6
	(20.3)	(15.9)	(232.6)	(101.2)	(86.6)	(27.9)
Pct Hispanic (Std)	56.3	-22.8	376.9	469.3	221.5	94.8
	(19.9)	(9.4)	(115.8)	(141.5)	(71.5)	(33.7)
Pct Asian (Std)	47.8	-14.3	98.2	138.0	410.0	-37.7
	(16.9)	(7.1)	(39.0)	(44.9)	(123.9)	(19.2)
Pct College (Std)	145.9	-54.0	185.9	37.0	93.8	260.2
	(45.7)	(18.5)	(68.1)	(32.2)	(44.7)	(81.7)
Inside NYC	-337.8	-421.2	120.0	29.1	299.0	-2.3
	(106.9)	(128.6)	(97.6)	(83.9)	(129.4)	(68.6)
Commuting Time (Std)	38.7	-6.3	127.7	50.7	210.4	8.4
	(19.9)	(11.9)	(56.9)	(43.0)	(76.8)	(33.1)
Monthly Rent	0.33	-0.03	-0.36	-0.23	-0.18	0.21
	0.10	0.02	0.13	0.10	0.10	0.08

Table A.10: Price Instrument Placebo Test

Notes: This table provides a placebo test showing that the price instrument in the long-term rental market is indeed not correlated with neighborhood attributes, which are likely correlated unobserved quality. In contrast, the actual prices are correlated with these attributes, as expected.

	(1) Price Actual	(2) Price IV	(3) Price Actual	(4) Price IV	(5) Price Actual	(6) Price IV
Pct College (Std)	0.248***	0.00131				
	(0.0220)	(0.00124)				
Pct Black (Std)			-0.0738***	0.0000462		
			(0.0159)	(0.000747)		
Pct Hispanic (Std)					-0.0891***	-0.000889
					(0.0144)	(0.000719)
Number of Bedrooms = 1	0.219***	0.236***	0.219***	0.236***	0.226***	0.236***
	(0.0418)	(0.00328)	(0.0456)	(0.00328)	(0.0447)	(0.00328)
Number of Bedrooms = 2	0.420***	0.423***	0.430***	0.423***	0.430***	0.423***
	(0.0453)	(0.00298)	(0.0477)	(0.00298)	(0.0473)	(0.00298)
Number of Bedrooms = 3	0.609***	0.604***	0.605***	0.604***	0.610***	0.604***
	(0.0547)	(0.00364)	(0.0577)	(0.00364)	(0.0572)	(0.00364)
Number of Bedrooms = 4	0.465***	0.630***	0.455***	0.630***	0.460***	0.630***
	(0.103)	(0.00748)	(0.103)	(0.00747)	(0.104)	(0.00748)
Commuting Time (Std)	-0.155***	-0.148***	-0.317***	-0.149***	-0.324***	-0.149***
	(0.0296)	(0.00198)	(0.0260)	(0.00183)	(0.0246)	(0.00177)
Built After 1980	0.0346	0.135***	0.0199	0.135***	0.0223	0.135***
	(0.0439)	(0.00227)	(0.0458)	(0.00227)	(0.0458)	(0.00227)
Built 1940 to 80	-0.116***	-0.129***	-0.112***	-0.129***	-0.112***	-0.129***
	(0.0382)	(0.00241)	(0.0403)	(0.00241)	(0.0398)	(0.00241)
5+ Units	0.0145	-0.0628***	0.0148	-0.0628***	0.0245	-0.0626***
	(0.0365)	(0.00185)	(0.0384)	(0.00185)	(0.0383)	(0.00186)
Inside NYC	1.572***	1.508***	1.592***	1.508***	1.578***	1.508***
	(0.0511)	(0.00369)	(0.0534)	(0.00369)	(0.0530)	(0.00369)
N	1050	1050	1050	1050	1050	1050

Table A.11: Counterfactual using Actual Airbnb Penetration vs. Uniform Airbnb Penetration

Notes: This table compares the counterfactual analysis using the actual Airbnb listing data with an alternative counterfactual analysis using a hypothetical scenario where the penetration of Airbnb is uniform across space and housing types while holding the total supply of Airbnb listings the same. This comparison illustrates that the primary driver of the welfare differences is due to the geography of actual Airbnb listings. The remaining difference in the counterfactual welfare with uniform Airbnb activity captures the higher willingness-to-pay for housing attributes by higher-income and more educated households.

	Actual Penetration		**Uniform Penetration**	
($ p.a.)	Median	Difference	Median	Difference
Household Size				
Household Size = 1	(134)		(142)	
Household Size = 5	(119)		(141)	
Small vs. Large Households		**(15)**		**(1)**
Race / Ethnicity				
White	(152)		(137)	
Black	(134)		(150)	
Hispanic	(113)		(130)	
Asian	(127)		(132)	
White vs. Hispanic		**(39)**		**(7)**
Education				
With College	(156)		(141)	
Without College	(120)		(132)	
With vs. Without College		**(36)**		**(9)**
Household Income				
Highest Quintile	(167)		(144)	
Lowest Quintile	(124)		(135)	
Highest vs. Lowest Income Quintile		**(43)**		**(9)**

B More on Long-Term Rental Demand Estimation

The main idea is that one can use a supply-side pricing equation to produce a price instrument. The key intuition is that the availability of similar housing characteristics in the market has an impact on the equilibrium price of a given house, but is uncorrelated with the unobserved quality. Moreover, if I have two or more housing characteristics, then their relative availability allows this identification strategy to work even with only one cross-section of the market.

In this section, I first describe the model, followed by the identification strategy, and end with some findings from a simulation study.

B.1 Model

Consider the following model with N households indexed by i and M homes indexed by j in a given housing market. Each home is endowed with a vector of physical features $X_{j,k}$ with $k = 0, \ldots, K$, with at least two features $K \geq 2$. Let $X_{j,0}$ denote the indicator for the outside option, with its price normalized to zero. Moreover, each home has an unobserved quality component ξ_j. There is a random coefficient $\nu_{i,k}$ associated with each of the K features, drawn randomly from a normal distribution $\sigma \times \mathcal{N}(0, I)$. Here, I assume σ is known to simplify the exposition.[21] The utility for household i renting home j is as follows:

$$U_{i,j} = \alpha p_j + \sum_k \beta_{i,k} X_{j,k} + \xi_j + \epsilon_{i,j} \tag{B.1}$$

where $\beta_{i,k} = \beta_k + \nu_{i,k}$. Given that each household maximizes its utility, the choice probability of a household i over home j becomes:

$$P_{i,j}(p, \mathbf{X}; \Theta) = \frac{exp(\delta_j + \lambda_{i,j'})}{\sum_{j'} exp(\delta'_j + \lambda_{i,j'})} \tag{B.2}$$

$$\delta_j(p, \mathbf{X}; \Theta) = \alpha p_j + \sum_k \beta_k X_{j,k} + \xi_j \tag{B.3}$$

$$\lambda_{i,j}(p, \mathbf{X}) = \sum_k \nu_{i,k} X_{j,k} \tag{B.4}$$

where $\Theta = (\alpha, \vec{\beta}_k, \xi)$. The model is closed by a sorting equilibrium where price p^e clears the market: The demand for each home equals the observed supply, namely, $\forall j : s_j^F = 1$:

$$\forall j : \sum_i P_{i,j}(p^e, \mathbf{X}; \Theta) = 1 \tag{B.5}$$

[21] In the actual model, it is captured by the coefficients in front of household demographics, which are estimated offline first using individual-level choice data.

B.2 Estimation

The key identification assumption is that the unobserved quality ξ is *independent* of the physical features of the home $\mathbf{X}$. As a result, the unobserved quality ξ is uncorrelated with the price instrument z, which is constructed as the market clearing price with ξ set to zero:

$$\forall k : \mathbb{E}[\xi X_k] = 0 \tag{B.6}$$

$$\mathbb{E}[\xi z] = 0 \tag{B.7}$$

$$\forall j : \sum_i P_{i,j}\big(z, \mathbf{X}; (\alpha, \vec{\beta}, \xi = 0)\big) = 1 \tag{B.8}$$

$$\forall j : \sum_i P_{i,j}\big(p^e, \mathbf{X}; (\alpha, \vec{\beta}, \xi)\big) = 1 \tag{B.9}$$

Remark B.1. Why does the instrument work? The key intuition is that the supply-side model, namely, that the demand clears the fixed supply of homes, implies homes with rarer characteristics are going to have higher equilibrium prices, *compared with* homes with housing characteristics that are more common. And this component of variation (the scarcity premium) is uncorrelated with the unobserved quality of the particular home.

Suppose that there are only two relevant housing characteristics in the market: X_1 indicates a home was built within the last 5 years, which is relatively rare in NYC, and X_2 indicates a home has two bedrooms (as opposed to just one), which is more common. Both features are desirable. Without loss of generality, assume that the mean utilities for the two features are the same $\beta_1 = \beta_2$.

The key identification assumption is that the price premium observed for homes in brand-new buildings is higher than the price premium for homes with an extra bedroom. It is not because homes with two bedrooms have unobservably lower quality but rather because it is a much more common housing characteristic than being brand-new.

Remark B.2. It is essential that there be a distribution of preferences over the physical features of the homes. Because for the IV strategy to work, the model has to produce equilibrium prices that are higher for rarer features, even if the valuations for it by the average household (namely, the household with $\nu_i = 0$) might be completely identical. In other words, the distribution of preferences over characteristics means those who care a lot about a rare characteristic will bid up its price in equilibrium. In practice, households of different income levels are likely to have different price sensitivities, there will be a distribution over the WTP $\beta_{i,k}/\alpha_i$ for housing attributes.

Remark B.3. It is also essential that there are at least two relevant housing characteristics $K \geq 2$ in the market for the estimation to work with just one cross-section. Otherwise, the concept of the rarity of a characteristic is undefined. Because I can compare the rarity of one housing characteristic with another housing characteristic in just one cross-section, it is the reason the instrument can work with just a cross-section of the market.

B.3 Simulation Study

In this section, I generate a simple dataset with only two binary features X_1, X_2 with known parameters, where the first housing characteristic X_1 is much rarer than the second one X_2, even though the mean utilities on them are the same.

I show the price instrument constructed in Eq (B.8) indeed recovers the true parameter, unlike the OLS. Moreover, I also show the typical hedonic regression will result in estimated WTP for amenities significantly different from the true model parameters, whereas the proposed IV estimation strategy does recover the WTP for the average household as specified by the model.

B.3.1 Simulation Set-Up

More specifically, I have $N = M = 200$. Here is the breakdown of their characteristics with one home being the outside option (indicator by X_0):

(i) 20 homes have $X_1 = 1, X_2 = 1$

(ii) 20 homes have $X_1 = 1, X_2 = 0$

(iii) 120 homes have $X_1 = 0, X_2 = 1$

(iv) Remaining 39 homes have $X_1 = 0, X_2 = 0$

The true parameters of the simulation are set as follows:

$$\alpha = -1, \quad \beta_0 = 3, \quad \beta_1 = 1, \quad \beta_2 = 1 \tag{B.10}$$

Moreover, the unobserved quality ξ is drawn from a normal distribution with a standard deviation of 1/200. ν_i are drawn from a normal distribution with a standard deviation of 1 for both characteristics. With the model fully specified as above, I can solve it by computing the equilibrium price $p^e(\mathbf{X}; \alpha, \vec{\beta}, \xi)$. Then, the mean utility δ is computed as follows:

$$\delta_j = \alpha p_j^e + \beta_0 X_{j,0} + \beta_1 X_{j,1} + \beta_2 X_{j,2} + \xi_j \tag{B.11}$$

B.3.2 Simulation Results

I compute the price instrument that clears the market. I checked that the price instrument is indeed uncorrelated with the unobserved quality with $cov(z, \xi) = 0$. I find the IV recovers the true price coefficient, whereas the OLS produces a biased estimate of them, as shown in B.1. I also find the estimation strategy correctly recovers the true value of both amenities $(-\beta_k/\alpha)$, whereas a hedonic regression of price on amenities greatly overestimates the value of the rarer amenity X_1 and underestimates the value of the more common amenity X_2, as illustrated in Table B.2. Note that the result *does not* require $\beta_1 = \beta_2$. I can vary the parameter vector to different values and the estimation strategy will still work, as long as at least two characteristics are present.

Table B.1: Regression on Mean Utility

The *dependent variable* is the mean utility δ_j. The price instrument is constructed using the market clearing prices assuming $\xi = 0$. The F-stat of the first stage is well above 10 (at 306.0). The OLS produces a biased estimate of the price coefficient, whereas the IV recovers it.

	(1) OLS	(2) First Stage	(3) IV
Price	-0.635***		-1.042***
	(0.0363)		(0.0589)
Price Instrument		0.960***	
		(0.0549)	
X1	0.498***	0.0549	1.057***
	(0.0499)	(0.0754)	(0.0810)
X2	0.715***	0.0316	1.033***
	(0.0283)	(0.0428)	(0.0460)
Inside Option	1.907***	0.120	3.125***
	(0.109)	(0.164)	(0.176)
N	200	200	200

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B.2: Estimation of the Willingness to Pay for Amenities

In this table, I compare the amenity value estimated by the model $(-\beta_k/\alpha)$ with the hedonic regression. In particular, even though the mean utility of the two characteristics X_1 and X_2 are identical, the relative scarcity of X_1 pushes the hedonics to produce a much greater coefficient.

	(1) Hedonic	(2) Sorting
X1	1.374***	1.015***
	(0.00139)	(0.0204)
X2	0.781***	0.991***
	(0.00122)	(0.0120)
Inside Option	2.994***	2.999***
	(0.00110)	(0.000760)
N	200	200

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C More on Short-Term Rental Supply Estimation

In this section, I provide a number of relevant computational details about the estimation procedure used to estimate the supply system. Following Dubé, Fox and Su (2012), I formulate the GMM objective function as a mathematical program with equilibrium constraints (MPEC). With the problem appropriately defined, I provide the analytical Jacobian and Hessian used to estimate the parameters, where I highlight the sparsity features of the problem that make the computation reasonably efficient.

C.1 Problem Formulation

Without loss of generality, let q index a total of Q markets. Let $Z_q = [\, p_q^{IV} \;\; X_q^R \,]^T$ be a vector of exogenous shifters and $X_q = [\, p_q \;\; X_q^R \,]^T$ be the endogenous shifters. Let $Z = [Z_1, \ldots, Z_Q]^T$ and let $X = [X_1, \ldots, X_Q]^T$ be the data matrix of size $Q \times (X + 1)$. Let D_q denote the empirical distribution of the demographic characteristics in market q.

The requisite moment condition is

$$\mathbb{E}\big[(\delta_q - \theta_1^T X_q)^T Z_q\big] = 0 \tag{C.1}$$

To solve for the supply system coefficients, I formulate the problem as

$$\min_{(\delta, \theta_1, \theta_2, \eta)} \quad \eta^T W \eta \tag{C.2}$$

$$\text{s.t.} \quad \forall q: \; S_q(\delta_q, \theta_2; X_q, D_q) = S_q^o \tag{C.3}$$

$$\eta = Z'(\delta - \theta_1^T X) \tag{C.4}$$

As such, I denote the Lagrangian of the problem as

$$f_q(\delta_q, \theta_2; X_q, D_q) = S_q(\delta_q, \theta_2; X_q, D_q) - S_q^o \tag{C.5}$$

$$g(\delta, \theta_1; Z, X) = \eta - Z^T(\delta - \theta_1^T X) \tag{C.6}$$

$$G(\eta; W) = \eta^T W \eta \tag{C.7}$$

$$\mathcal{L}(\delta, \theta_1, \theta_2, \eta) = G(\eta; W) + \sum_q \lambda_f^q f_q(\delta, \theta_2; X_q) + \lambda_g^T g(\delta, \theta_1; Z, X) \tag{C.8}$$

Note θ_1 denotes the linear coefficients of the model, whereas θ_2 denotes the non-linear coefficients of the model. Further, let $\theta_2 = [\pi_b^k]$ where $b = 1, \ldots, (X + 1)$ indexes the product characteristics and $k = 1, \ldots, K$ indexes the demographic characteristics. As such, the heterogeneous component of the home sharing is

$$\lambda_{i,q} = \sum_b \big(\sum_k \pi_k^b z_{i,k} \big) X_q^b \tag{C.9}$$

where $z_{i,k} \sim P^*_{D_q}$ is drawn from the empirical distribution of the demographics in market q. With logit error, the market share S_q is thus computed as

$$S_q(\delta_q, \theta_2; X_q, D_q) = \frac{1}{N_q} \sum_q P_{i,q}(\delta_q, \theta_2; X_q, D_q) = \frac{1}{N_q} \sum_q \frac{\exp(\delta_q + \lambda_{i,q})}{1 + \exp(\delta_q + \lambda_{i,q})} \tag{C.10}$$

C.2 Analytical Derivations

Analytical Derivatives of MPEC

The gradient of the objective function is

$$\nabla_{(\delta,\theta_1,\theta_2,\eta)} G(\eta; W) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2W\eta \end{bmatrix} \tag{C.11}$$

The Jacobian for the constraints is

$$\nabla_{(\delta,\theta_1,\theta_2,\eta)} (f, g) = \begin{bmatrix} \frac{\partial f}{\partial \delta} & 0 & \frac{\partial f}{\partial \theta_2} & 0 \\ \\ \frac{\partial g}{\partial \delta} & \frac{\partial g}{\partial \theta_1} & 0 & I_g \end{bmatrix} \tag{C.12}$$

where

$$\frac{\partial f_q}{\partial \delta_q} = \frac{1}{N_q} \sum_i P_{i,q}(1 - P_{i,q}), \quad \frac{\partial f_q}{\partial \delta'_q} = 0, \quad q \neq q' \tag{C.13}$$

$$\frac{\partial f_q}{\partial \pi_k^b} = \frac{1}{N_q} \sum_i P_{i,q}(1 - P_{i,q}) z_{i,k} X_q^b \tag{C.14}$$

$$\frac{\partial g}{\partial \delta} = -Z^T, \quad \frac{\partial g}{\partial \theta_1} = Z^T X \tag{C.15}$$

Note that the upper-left $Q \times Q$ block of $\frac{\partial f}{\partial \delta}$ contains only diagonal terms $\frac{\partial f_q}{\partial \delta_q}$. When the number of markets Q is large, this results in a sparse Jacobian, which is particularly attractive computationally.

Analytical Hessians of MPEC Lagrangian

Next, I derive the analytical Hessian of the MPEC Lagrangian:

$$\nabla^2 \mathcal{L}(\delta, \theta_1, \theta_2, \eta) = \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \delta^2} & 0 & \frac{\partial^2 \mathcal{L}}{\partial \delta \partial \theta_2} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\partial^2 \mathcal{L}}{\partial \theta_2 \partial \delta} & 0 & \frac{\partial^2 \mathcal{L}}{\partial \theta_2^2} & 0 \\ 0 & 0 & 0 & \frac{\partial^2 \mathcal{L}}{\partial \eta^2} \end{bmatrix} \tag{C.16}$$

where

$$\frac{\partial^2 \mathcal{L}}{\partial \delta_q^2} = \lambda_f^q \frac{\partial f_q^2}{\partial \delta_q^2} = \lambda_f^q \frac{1}{N_q} P_{i,q}(1 - P_{i,q})(1 - 2P_{i,q}) \tag{C.17}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \delta_q \delta_q'} = 0, \quad q \neq q' \tag{C.18}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \delta_q \partial \pi_k^b} = \lambda_f^q \frac{\partial f_q^2}{\partial \delta_q \partial \pi_k^b} = \lambda_f^q \frac{1}{N_q} P_{i,q}(1 - P_{i,q})(1 - 2P_{i,q}) z_{i,k} X_q^b \tag{C.19}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \pi_k^b \partial \pi_{k'}^{b'}} = \sum_q \lambda_f^q \frac{\partial f_q^2}{\partial \pi_k^b \partial \pi_{k'}^{b'}} = \sum_q \lambda_f^q \frac{1}{N_q} P_{i,q}(1 - P_{i,q})(1 - 2P_{i,q}) z_{i,k} X_q^b z_{i,k'} X_q^{b'} \tag{C.20}$$

$$\frac{\partial^2 \mathcal{L}}{\partial \eta^2} = 2W \tag{C.21}$$

Again, note the upper-left $Q \times Q$ block representing $\frac{\partial^2 \mathcal{L}}{\partial \delta^2}$ has non-zero entries only on the diagonal term. Thus, even with a large Q, the Hessian remains sparse, making it computationally more tractable.