

Professional Airbnb Hosts in Mexico City: A First Approximation

Juan José Merino

Edwin Muñoz-Rodríguez

El Colegio de México

mayo de 2024

Professional Airbnb Hosts in Mexico City: A First Approximation

Anfitriones profesionales en Airbnb, CDMX: Una primera aproximación *

Juan José Merino[†] & Edwin Muñoz-Rodríguez[†]

May 2024

Abstract

We offer a first approach to the demand and supply structure of Airbnb in Mexico City using a unique panel dataset of daily listings and rentals on the platform. We find descriptive evidence of a large sector of professional hosts. We estimate various models of demand accommodating price endogeneity and taste heterogeneity, as well as a supply model under the assumption of Bertrand-Nash's competition. We consistently find that professional hosts have lower marginal costs than nonprofessional hosts. We also find that overall price-cost margins are high regardless hosts' professionalization status, and slightly higher in the case of professional ones. Finally, motivated by a worldwide regulatory trend, we investigate counterfactually the effect of producer and consumer *ad valorem* taxes that discriminate between types of hosts.

Resumen

Realizamos una primera aproximación a la estructura de la demanda y la oferta de Airbnb en la Ciudad de México usando datos panel que siguen diariamente todas las propiedades ofertadas y rentadas. Encontramos evidencia descriptiva de la existencia de un sector importante de anfitriones profesionales. Nuestras estimaciones de múltiples modelos de demanda (que acomodan endogeneidad del precio y heterogeneidad en las preferencias) y un modelo de oferta bajo el supuesto de competencia de Bertrand-Nash muestra que los anfitriones profesionales tienen menores costos marginales. También se encuentran altos márgenes precio-costos en la industria en general, y ligeramente mayores para los anfitriones profesionales. Finalmente, siguiendo una tendencia regulatoria a nivel mundial, investigamos de manera contrafactual el efecto de impuestos al consumidor y productor que discriminan entre tipos de anfitriones.

Clasificación JEL/JEL Classification: C59; L11; L83; L86; Z31

Palabras clave/keywords: Sharing economy; structural econometrics; peer-to-peer markets; market structure; professionalization

*We thank Diana Terrazas and two anonymous referees for very helpful comments. This research was funded by the Ford Foundation through a grant awarded to the PRAEM program at El Colegio de México A.C. The views and conclusions expressed in this article are solely those of the authors, and do not necessarily reflect those of the Ford Foundation.

[†]Center For Economic Studies, El Colegio de México A.C., Carretera Picacho Ajusco 20, Tlalpan, México City, C.P. 14110. Merino: jmerino@colmex.mx; Muñoz-Rodríguez: eamunoz@colmex.mx

1 Introduction

A controversial worldwide characteristic of the evolution of Airbnb in the last decade is the gradual professionalization of hosts within the platform, which has been interpreted as a shift from a disruptive peer-to-peer market towards a more traditional business-to-consumer market (Popper 2015). A flourishing multi-disciplinary literature has documented such pattern in Europe (Demir and Emekli 2021; Gyódi 2019), the United States (Dogru et al. 2020), Canada (Gibbs, Guttentag, Gretzel, Yao, et al. 2018), and South Korea (Ki and S. Lee 2019), among others. Some authors argue that this professionalization (a.k.a. hotelization) of Airbnb scalated its negative externalities on the urban landscape (Gil and Sequera 2022; D. Lee 2016). Increasingly, regulations in major touristic destinations make a clear distinction between professional and nonprofessional individual hosts, subjecting the former to a stricter set of rules or blankly banning such professional activity (D. Lee 2016; Briel and Dolnicar 2020; Iacovone 2023).¹

In this paper, we offer a first approach to the so far unknown structure of demand and supply in the Airbnb market in Mexico City, leveraging a rich panel dataset obtained from a short-term rental intelligence firm and a structural model, which we describe below. In the first step, a descriptive analysis of the daily posted properties reveals that there is heterogeneity in the type and number of properties a host manages in the platform, varying from 1 to 77 per host, as well as in the number of days a property is available to be rented in the platform during a year, ranging from 1 day to 365 days (actually rented days vary from 1 to 352 days per property). Moreover, the market share of multi-unit hosts amounts to 65.6% of the annual market despite them being only 27.93% of the total number of hosts on the platform. These facts suggest that, like some other major touristic hubs, the Mexico City market for short-term peer-to-peer rentals is a mixture of professional and nonprofessional hosts.

Motivated by these descriptive findings, we further investigate the professionalization of Airbnb's hosts in Mexico City, estimating a stylized structural model of daily demand and supply for short-

¹At the time of finishing this writing, Mexico City government had just publicly announced is preparing regulations in this spirit. See EFE (2023).

term rentals in the spirit of S. Berry, Levinsohn, and Pakes (1995). In our model, guests choose among classes of accommodations (which differ in quality, location and accommodation type) managed by professional or nonprofessional hosts.² Our stylized framework still captures four crucial features of interest in the market for short rentals: daily demand and price fluctuations, accommodations and portfolio heterogeneity across professional and nonprofessional hosts, and rich substitution patterns.

The estimation delivers marginal costs, markups and price-cost margins, which help to understand the differences between professional and nonprofessional players. Moreover, being structural, the model allows us to investigate the equilibrium effects of a differential tax targeting professional hosts, a policy in line with worldwide regulatory trends that might be of interest to local authorities.

Estimated marginal costs are in line with intuition. As expected, professionals face smaller marginal costs than nonprofessionals, consistent with economies of scale, scope, and superior management practices. Entire homes/apartments have higher associated marginal costs than private rooms. Properties in wealthier Alcaldias have higher marginal costs. Regardless of host type, price-cost margins are high, consistent with differentiated products and high fixed costs like mortgage and rent payments. Professionals have slightly higher markups and price-cost margins, which can be partially explained by their lower marginal costs and multi-product portfolio. Finally, we evaluate two counterfactual scenarios focused on differential tax targeting professional hosts. An additional consumption tax of 5% on properties owned by professionals naturally increases the total amount to be paid by the consumer, decreases the market share of professionals by 5.52%, improves the market share of nonprofessionals only marginally by 0.57%, and increases the market share of the traditional accommodation sector. On the other hand, a producer tax of 5% increases prices, decreases the market share of professionals by 5.96%, marginally improves the market share of nonprofessionals by 0.61%, and increases the market share of the traditional accommodation

²We distinguish between professional and nonprofessional hosts by the number of properties they manage. We discuss this definition and the implications and limitations of our modeling choices in sections 2 and 3. We perform a robustness test for the leading definition of professional hosts in Appendix A.4.

sector. The main difference between scenarios is that the consumer pays less in the consumer tax scenario. These results are overall robust to different estimation methodologies which we describe later.

To the best of our knowledge, we have made two original contributions: We are the first to provide systematic empirical evidence on the professionalization of Airbnb in a Latin American city and the first to take a structural approach to the professionalization issue worldwide. Our findings on professionalization for Airbnb in Mexico City are crucial because, as pointed out above, most regulations distinguish between professional and nonprofessional Airbnb hosts. Such distinction has recently been incorporated in a regulatory intent by the Mexican authorities, which might be a sensitive issue in developing countries as the sharing economy can help improve households' welfare through additional income (Trade and Development 2020).

1.1 Related literature

Only recently has the literature started to use structural models of demand and supply for differentiated products to study different aspects of Airbnb (Farronato and Fradkin 2022; Farhoodi 2021; Calder-Wang 2021). However, these are about Airbnb in the U.S., and the differences between professionals and nonprofessional hosts have yet to be studied. The closest to us is Farhoodi (2021), who studies the distribution of benefits among neighborhoods in the Chicago Airbnb market, finding that higher-income neighborhoods benefit the most from access to the platform and that a tax on platform users has redistributive effects. Our research contributes to this growing literature using structural methods to study markets for short-term rentals by offering a stylized structural framework to study the differences between professional and nonprofessional hosts, and doing so in a Latin American city.

Chen, Wei, and K. Xie (2023) is the only other paper previous to us investigating the roles of professional and nonprofessional hosts in the Airbnb market using econometric techniques. They use a quasi-experimental design, exploiting a policy change in San Francisco and New York that caps the number of properties a host can manage. This provides indirect empirical evidence of

substitution between properties managed by professional and nonprofessional hosts. On the other hand, there is a large number of papers providing descriptive evidence on the professionalization of Airbnb hosts; in addition to those cited in the Introduction, Abrate, Sainaghi, and A. G. Mauri (2022) and Ki and S. Lee (2019) document the complex pricing strategies followed by professional hosts as compared to nonprofessional ones. By implementing structural econometric techniques, we offer an alternative approach to study the relevant economic differences between professional and nonprofessional Airbnb hosts, some of which (like costs and margins) are difficult to recover otherwise.

Regarding Airbnb in Mexico City, only a few studies have been conducted so far. López Tamayo and Ramírez Álvarez (2021) conducted a hedonic price analysis for Airbnb in Mexico City, finding positive determinants of price such as the maximum number of guests allowed, number of bedrooms, number of bathrooms, number of amenities, and professional host, as well as negative determinants such as crime in the accommodation area. Ruiz-Correa et al. (2019) and Oca, Favà, and Gay (2018) have approached the issue of Airbnb in Mexico City from an Urban Studies perspective. Banco de México (2021) examined the evolution of the Airbnb market in Mexico City. We provide a state-of-the-art framework to estimate unobservable economic primitives in the industry, and a first set of estimates speaking to the differences between Airbnb hosts in Mexico City. Moreover, the proposed structural framework can be further refined to provide insights into the effect of different public policies.

We follow the literature on estimating models of product differentiation initiated by S. Berry, Levinsohn, and Pakes (1995) to address several methodological challenges. We estimated several logit and random coefficients (RC) demand models using the standard nonlinear IV technique with the modern differentiation instruments of Gandhi and Houde (2019) to account for price endogeneity. To deal with the significant heterogeneity in Airbnb listings in a computationally tractable way, we adapted an aggregation method developed by Farronato and Fradkin (2022). Finally, many listed properties are not rented in a given day, which translates into the well-known zero market share problem (Dubé, Hortaçsu, and Joo 2021). Our results are robust to the two most

common ways of handling products with zero market share in these types of models, excluding observations with zero market share and imputing to them a positive market share value close to zero (Gandhi, Lu, and Shi 2023).

2 Background, data and preliminary analysis

Airbnb was introduced in Mexico City in 2009. In 2017, the Mexico City government introduced a state tax for Airbnb guests, ranging between 3 to 5% (Airbnb 2019). In 2019, the platform collected approximately 202.8 million Mexican pesos in taxes (Airbnb 2020).

For hosts, listing an accommodation on Airbnb is free, and there is complete flexibility in setting the available rental days as desired. Airbnb provides host complimentary protection called Aircover, which covers potential damages, unexpected cleaning costs, and loss of income (Airbnb-Help Center n.d.[b]), and also extends to guests in case they suffer injuries during their stay (Airbnb-Help Center n.d.[a]). Hosts can set their prices manually, using a dynamic pricing tool offered by the platform (Airbnb-Help Center n.d.[c]), or using third-party algorithms.

Potential guests register for free on the Airbnb platform and search for accommodation options through the application. They can indicate their desired dates and length of stay and further refine their search using filtering tools.

Airbnb earns revenue by charging a fee to hosts and guests through two fee structures: the shared fee and the host-only fee. In the shared fee model, the fee is split between the host and the guest. Most hosts pay a 3% fee of the subtotal before taxes. The guest fee is typically less than 14.2% of the reservation's subtotal before taxes, varying based on several factors. In the host-only fee model (mandatory for hotels and hosts using an external property management system), the entire fee is deducted from the host's earnings, generally ranging from 14% to 16% of the subtotal before taxes.

2.1 Data

We use a private microdata set provided by AirDNA, a short-term rental intelligence firm, containing data for the 16 Alcaldias in Mexico City during 2019 scrapped from the Airbnb portal.³ It records the final status (not booked, booked, unavailable) and corresponding rental price (in US dollars) for each posted accommodation and date. Additionally, it offers detailed information about each accommodation, including exact location (latitude and longitude), number of bedrooms, bathrooms, number of photos posted, internet access, and whether the accommodation is an entire home/apartment or a shared/private room, whether it is offered by a “superhost”, and amenities offered, among others. A major drawback of this data is that it does not contain any demographic or other guests’ characteristics, which makes it difficult to identify taste variation.

Our initial microdata set consists of 6,471,668 observations corresponding to daily accommodations offered in Airbnb in Mexico City for the year 2019 (for the economy of language hereafter, we refer to one of these accommodation-date observations as a listing). We conducted the following data cleaning process: First, we removed 2,682 listings with a zero price. Subsequently, we eliminated 1,228,841 listings from 7,448 available accommodations that were never rented during 2019. Of the remaining 5,240,145 listings, shared rooms and hotel rooms represent small percentages (1.83% and 1.57% respectively, see Table 1), so we concentrate our analysis on private rooms and entire homes/apartments exclusively.

Additionally, to prevent biases and enhance the accuracy of our estimates, for each accommodation type and Alcaldia, we filtered out all listings with prices above 99% or below 1% of the prices of booked listings during 2019.

Our final sample contains 4,844,410 listings, of which 40.72% correspond to booked ones. This information pertains to 24,956 unique accommodations scattered across the 16 Alcaldias of

³Because Airbnb data is not public, most of the time researchers have to rely on data scrapped by third parties. The AirDNA dataset has been extensively used in academic research, for instance: Chen, Wei, and K. Xie (2023), Calder-Wang (2021), Farhoodi (2021), Abrate, Sainaghi, and A. G. Mauri (2022), Banco de México (2021), Dogru et al. (2020), Gibbs, Guttentag, Gretzel, Yao, et al. (2018), Iacovone (2023), Ding, Niu, and Choo (2023), K. Xie, Heo, and Mao (2021), Sainaghi, Abrate, and A. Mauri (2021), and Tong and Gunter (2022). Based on official Airbnb quarterly earning reports, AirDNA claims their dataset has 95% accuracy. See <https://www.airdna.co/airdna-accuracy> (retrieved on May 2nd, 2024).

Table 1: Number of booked and available listings by accommodation type during the year 2019.

Listing Type	Frequency	Percentage	Cumulative Percentage
Hotel room	82,109	1.57	1.57
Shared room	96,050	1.83	3.40
Entire home/apt.	2,500,380	47.72	51.12
Private room	2,561,606	48.88	100.00
Total	5,240,145	100.00	

Mexico City. The Alcaldias with the most significant number of total listings, as well as of booked and non-booked ones, are Cuauhtémoc, Miguel Hidalgo, and Benito Juárez (see Table A.1).

Finally, in our demand models, we assume guests’ outside option is booking a room in the traditional lodging sector.⁴ We obtain data on monthly average daily hotel room occupancy from DataTur, a tool implemented by the Mexican Secretary of Tourism to monitor hotel occupancy. Table 2 compares these averages with the Airbnb’s average daily occupancy calculated by us from our final sample described in the previous paragraphs. It shows that in terms of the monthly average of daily occupancy (resp. available accommodations), Airbnb rentals represent between 16% and 20% (resp. 20% and 31%) of the occupancy (resp. availability) in Mexico City’s traditional sector.

2.2 Preliminary analysis

The 24,956 accommodations belong to 14,788 hosts. We classify these hosts as professional or not according to the number of accommodations they manage on the platform. In the body of the paper, we adopt the single-unit/multi-unit definition of nonprofessional and professional hosts, largely adopted in the literature.⁵ In Appendix A.4, we review the robustness of our results to more nuanced definitions of what it means to be a professional host, also based on the number and type of properties.

⁴We discuss the choice set in detail in section 3.1.

⁵For instance Chen, Wei, and K. Xie (2023), Abrate, Sainaghi, and A. G. Mauri (2022), Gil and Sequera (2022), Li, Moreno, and Zhang (2016), López Tamayo and Ramírez Álvarez (2021), Gibbs, Guttentag, Gretzel, Morton, et al. (2018), Magno, Cassia, and Ugolini (2018), and Tong and Gunter (2022) explicitly define professionals as those managing more than one accommodation. The distinction single-unit/multi-unit as a source of heterogeneity in managerial practices is also present in Koh, Belarmino, and Kim (2020), Kwok and K. L. Xie (2019), and Lorde, Jacob, and Weekes (2019).

Table 2: Average daily accommodations in Mexico City, 2019

Month	Hotel rooms available	Hotel rooms occupied	Airbnb listings available	Airbnb listings occupied
January	51,284	27,770	10,765	4,585
February	51,234	33,738	10,749	5,043
March	51,204	34,071	11,529	5,261
April	51,208	32,886	12,634	5,134
May	51,207	33,822	12,926	4,792
June	51,239	33,003	13,733	5,212
July	51,239	34,658	14,174	5,733
August	51,324	32,852	14,202	5,163
September	51,324	33,515	14,353	5,571
October	51,316	35,639	13,528	5,529
November	51,333	37,771	14,539	6,540
December	51,385	31,155	15,962	6,280

Sources: Hotel room occupancy data was obtained from DataTur, a public monitoring system of Mexico’s Secretary of Tourism. The Airbnb averages shown were calculated using our final sample of AirDNA data.

Of the 14,788 hosts, 26.93% were classified as professional hosts who can manage between 2 to 77 accommodations, amounting to 56.70% of the accommodations posted on Airbnb in Mexico City during 2019. More specifically, 55.22% of professional hosts offer exactly two accommodations, and 20.02% offer exactly three accommodations. On the other hand, in Table 3 we present the annual number of listings by type of host. Professional hosts offered 58.22% of the listings and 64.60% of the booked listings.

Table 3: Number of booked and non-booked listings by type of host.

	Status		Total
	Non-Booked	Booked	
Host nonprofessional	1,325,680	698,290	2,023,970
Professional	1,546,157	1,274,283	2,820,440
Total	2,871,837	1,972,573	4,844,410

Note: Listings corresponding exclusively to Entire home/apartment and Private room accommodations.

Finally, Table 4 shows descriptive evidence regarding price variation among accommodations posted on the platform. We regress prices at the listing level (i.e., accommodation-date) on reser-

vation status, host type, accommodation type, Alcaldia, and date fixed effects. The Alcaldias with the highest prices are Miguel Hidalgo, Cuauhtémoc, and Cuajimalpa. Booked listings are \$6.70 less expensive than non-booked ones. Listings offered by professional hosts are \$6.24 costlier than those from nonprofessional hosts. Lastly, private rooms are \$42.28 cheaper than entire homes/apartments.

Table 4: Regression of listing prices with respect to fixed effects of Alcaldia, reservation status, host type, accommodation type, and date.

	Listing prices	
	Coefficient	p-value
Intercept	52.16	0.00
Alcaldia (0 = Azcapotzalco)		
Miguel Hidalgo	33.28	0.00
Cuauhtémoc	20.19	0.00
Cuajimalpa de Morelos	20.06	0.00
Coyoacán	9.84	0.00
Álvaro Obregón	9.67	0.00
Benito Juárez	4.94	0.00
Milpa Alta	4.60	0.04
Tlalpan	3.62	0.00
La Magdalena Contreras	2.66	0.00
Gustavo A. Madero	-1.70	0.00
Xochimilco	-3.25	0.00
Iztapalapa	-3.41	0.00
Venustiano Carranza	-4.36	0.00
Iztacalco	-6.33	0.00
Tláhuac	-6.41	0.00
Status (0 = Non-Booked)		
Booked	-6.70	0.00
Host type (0 = nonprofessional)		
Professional	6.24	0.00
Accommodation type (0 = Entire home/apartment)		
Private room	-42.28	0.00
Date FE	Yes	
Observations	4,844,410	

Notes: Prices are in January 2019 real dollars. Recall we define listing as an accommodation-date observation.

Motivated by these preliminary findings suggesting there might be differences between professional and nonprofessional hosts, the rest of the paper is dedicated to building and estimating

models of demand and supply of differentiated products for the Airbnb market in Mexico City. As a preliminary step, the next section addresses an issue that arises due to the nature of Airbnb.

3 Creating categories and representative products

Airbnb is a market with highly heterogeneous products, which imposes nontrivial challenges in model estimation. In fact, to account for all this product heterogeneity, one would need a model accommodation-specific demand for each date, which would imply endogenous choice sets for arriving guests and capacity constraints on the supply side, which are unsolved issues at the forefront of the literature on discrete choice estimation and identification (Agarwal and Somaini 2022; Farhoodi 2021). On top of that, such a model would have a large number of products, a well-known challenge in the estimation of discrete choice models of demand (see for instance Skrainka and Judd (2011)), many of which would have zero market share (offered but not rented), which adds another layer of complexity to the estimation procedure (Dubé, Hortaçsu, and Joo 2021).

To overcome these challenges, we classify the accommodations into disjoint categories and create representative products for each category.⁶ We created 116 categories and an equal number of representative products. In this section, we give an overview of how we created these categories and representative products and their implications for modeling.⁷

We follow a procedure similar to that described by Farronato and Fradkin (2022) for classifying accommodations and restrict attention to booked listings only (which is without loss since every accommodation in our final dataset was booked at least once). We start with a hedonic regression where the natural logarithm of the prices of booked listings depends on the fixed effects of accommodation and date.⁸ The accommodation fixed effects are intended to capture the average utility a guest obtains from booking a particular accommodation.⁹

⁶Product aggregation is commonly used in discrete choice demand models. Dubé, Hortaçsu, and Joo (2021) already cited elaborates on this.

⁷Appendix A.3 contains a more detailed explanation of product categorization.

⁸Other fixed effects, such as Alcaldia, accommodation type, and type of host, are omitted due to collinearity.

⁹Following Farronato and Fradkin (2022), we apply a Bayesian shrinkage procedure to the accommodation fixed effects to mitigate the impact of sampling error.

To create categories, we first group individual accommodations by accommodation type (entire home/apartment or private room) and Alcaldia. Within each such group, we establish quality tiers using the accommodation fixed effects previously estimated; the number of quality tiers varies according to the group size. Each subgroup of individual accommodations of the same type, located in the same Alcaldia and within the same quality tier, is further split between professional and nonprofessional hosts, which enables our analysis. The resulting subgroup in this way is a product category, of which we have 116.¹⁰

For each category, a representative product is created, whose characteristics are the average of the features of the unique accommodations that make it up throughout the year; these remain constant over time. On the other hand, the daily price for each representative product is calculated by averaging the prices of the accommodations within the category reserved for that date. Similarly, to calculate the quantity of the representative product sold for each date, we sum up the reservations of accommodation within the category for that date.

3.1 The stylized choice problem

Within this stylized framework, the choice set faced by every consumer looking for accommodation on a given date is the subset of the 116 representative products offering accommodation on that date, plus an outside option of booking a hotel room outside the platform. The demand for a given representative product for a given date is the count of accommodations belonging to the corresponding category that were booked for that date. We estimate the size of the outside option with the corresponding monthly average of daily occupied hotel rooms (Table 2).

If no individual accommodation belonging to a specific category is offered on a given date, we drop that particular representative product from the corresponding date's choice set. Relatedly, there are dates when no accommodation belonging to a given category is rented despite having available accommodations, resulting in a zero market share, which might create some issues in the

¹⁰Since our research question is different from that of Farronato and Fradkin (2022), our categorization is finer than the one implemented by them. They have 4 categories per city-date (10 cities), while we have 116 potential categories for each date. The actual number varies with date as some categories might not always be offered.

estimation of the demand models. We implement the two most widespread ways of dealing with this issue (see section 4.3 for detailed discussion).

On the supply side, representative products offered by nonprofessionals are assumed to be provided by single-product hosts. In contrast, the representative products offered by professionals are assumed to be supplied by a multi-product host who maximizes joint profits.

Our framework, though highly stylized, still captures crucial features of the market for short-term rentals while keeping the estimation issues mentioned at the beginning of this section under control. Specifically, our model captures daily demand and price fluctuations; still captures variation in quality, location, and type of accommodations offered by professional and nonprofessional hosts, and complex substitution patterns; as well as the multi-product nature of professional hosts' portfolio, which enables economies of scope and internalization of within portfolio substitution effects.

Naturally, this stylized aggregation abstracted some aspects of the Airbnb market for short rentals. The most critical elements lost by our modeling choices (on top of the choice set endogeneity) are the variation across professional hosts' portfolios and an even more extensive product heterogeneity, which we leave for future research.

4 Model and Estimation

In this section we present our theoretical models of demand and supply, as well as our estimation strategy. It is important to emphasize that we model and estimate daily demand and supply, so in total we have 365 markets.

4.1 Models of demand

We model consumer demand using logit and mixed logit (a.k.a. random coefficients) models. On a given date, guests demand one out of a subset of the 116 categories offered by professional hosts (who offer more than one accommodation type) and nonprofessional hosts (who offer only

one accommodation type) or the outside option of renting a hotel room in the traditional lodging sector.¹¹ Next, we formalize these ideas.

Logit Guest n chooses an accommodation belonging to category $j \in \mathcal{J}_t$, where \mathcal{J}_t is the set of categories available in day t on the platform. The utility guest n receive from an accommodation in category j , with observed features \vec{x}_j and price p_{jt} , including the tax rate τ_j on day t is:

$$U_{njt} = \vec{\beta}' \vec{x}_j - \alpha \log(p_{jt}(1 + \tau_j)) + \xi_{jt} + \varepsilon_{njt} \quad (1)$$

$\vec{\beta}$ and α characterize the preferences of consumers for accommodations' observed attributes. ξ_{jt} captures preferences for unobserved features that do not vary among guests. ε_{njt} corresponds to preferences for unobserved aspects that do vary among guests. As usual, we assume that ε_{njt} are i.i.d. Gumbel. For simplicity, we denote $\delta_{jt} = \vec{\beta}' \vec{x}_j - \alpha \log(p_{jt}(1 + \tau_j)) + \xi_{jt}$ to represent the average utility. There is an outside option (such as renting a hotel room outside the platform) whose utility we normalize to 0 in any day t . Thus, the probability of a guest choosing the accommodation in category j for day t is:

$$s_{jt}(\vec{\delta}_t(\vec{p}_t, \vec{x}_t, \vec{\beta}, \alpha)) = \frac{e^{\delta_{jt}}}{1 + \sum_{i \in \mathcal{J}_t} e^{\delta_{it}}}, \quad (2)$$

Mixed logit Also known as the random coefficients model, it is the most popular extension of the logit model in demand estimation because of its flexibility to accommodate heterogeneity in tastes among consumers, which allows flexible substitution patterns (Einav and Levin 2010; S. T. Berry and Haile 2021; Gandhi and Nevo 2021). In relation to the baseline model, the choice problem for guest n remains unchanged: to choose $j \in \mathcal{J}_t$. However, the random utility model in equation (1) becomes now:

$$U_{njt} = \vec{\beta}'_n \vec{x}_j - \alpha_n \log(p_{jt}(1 + \tau_j)) + \xi_{jt} + \varepsilon_{njt} \quad (3)$$

¹¹See section 3.1.

with $\vec{\beta}_n$ a random vector and α_n a random variable such that $\beta_{nk} \sim N(\bar{\beta}_k, \sigma_k^2)$ for the k th characteristic, mutually independent, and $\alpha_n \sim N(\bar{\alpha}, \sigma_p^2)$, independent from each component of $\vec{\beta}_n$. As before, these coefficients characterize the guest n 's preferences for accommodations' observed attributes. Also, as before, ξ_{jt} captures preferences for unobserved features that do not vary among guests, and ε_{njt} captures unobserved taste variation among guests, assumed to be distributed i.i.d Gumbel.

It is useful to separate the average and individual taste shocks. Rewrite (3) as,

$$U_{njt} = \delta_{jt} + \mu_{njt} + \varepsilon_{njt} \quad (4)$$

where $\delta_{jt} = \sum_k \bar{\beta}_k x_{jk} - \bar{\alpha} \log p_{jt} + \xi_{jt}$ is the part of utility which varies across alternatives but not consumers. On the other hand, $\mu_{njt} = \sum_k \sigma_k \nu_{nk} x_{jk} - \nu_{np} \sigma_p \log p_{jt}$ is the part of utility varying across consumers, with ν_{nk} , $k = 1, \dots, \dim(\vec{x}_j)$ and ν_{np} mutually independent normal standard taste shocks. Thus, the unconditional probability of a guest choosing the accommodation in category j for day t is now:

$$s_{jt}(\vec{\delta}_{jt}, \vec{p}_t, \vec{x}_j, \{\sigma_k\}_k, \sigma_p) = \int \frac{e^{\delta_{jt} + \mu_{njt}}}{1 + \sum_{i \in \mathcal{J}_t} e^{\delta_{it} + \mu_{nit}}} dF(\nu_{n1}, \dots, \nu_{n \dim(\vec{x}_j)}, \nu_p) \quad (5)$$

with F a multivariate standard normal distribution with dimension $\dim(\vec{x}_j) + 1$.

4.2 Supply model

We assume Nash-Bertrand competition. Hosts choose prices to maximize daily profits. The crucial difference between professional and nonprofessional hosts is that the former offer accommodations in different categories and maximize the joint daily profit, while nonprofessionals only offer accommodations in a single category. Let $\mathcal{J}_t^\Omega \subset \mathcal{J}_t$ the set of categories belonging to professional hosts and $\mathcal{J}_t^\omega \subset \mathcal{J}_t$ the ones belonging to nonprofessional hosts. Let \mathcal{M}_t the market size in day t . For the categories of products belonging to professional hosts, we set up a unique decision-maker who faces the following profit maximization problem:

$$\max_{\{p_{jt}\}_{j \in \mathcal{J}_t^\Omega}} \sum_{j \in \mathcal{J}_t^\Omega} (p_{jt} - c_{jt}) \mathcal{M}_t s_{jt}(\vec{\delta}_t(\vec{p}_t, \vec{x}_t, \vec{\beta}, \alpha)) \quad (6)$$

On the other hand, for the representative products belonging to nonprofessional hosts, for each category $j \in \mathcal{J}_t^\omega$, we set up a decision maker who faces the following profit maximization problem:

$$\max_{p_{jt}} (p_{jt} - c_{jt}) \mathcal{M}_t s_{jt}(\vec{\delta}_t(\vec{p}_t, \vec{x}_t, \vec{\beta}, \alpha)) \quad (7)$$

The **market equilibrium** for the day t consists of the seller's prices and consumer's choices, such that both hosts and guests make the decision that maximizes their profits and utilities, respectively, and that their optimal decisions are consistent with those of others.

4.3 Estimation

We follow the standard procedure for estimating models of demand and supply for differentiated products with aggregate data (S. Berry, Levinsohn, and Pakes 1995). Demand is estimated using the Generalized Method of Moments (GMM) with the moments obtained from the IV orthogonality conditions and the expected market shares. For mixed logit, we leverage state-of-the-art techniques implemented in PyBLP, a freely available Python library developed by leading researchers in empirical IO and increasingly used in the discipline to estimate BLP type models (Conlon and Gortmaker 2020). Demand estimates coupled with hosts' optimality conditions pin down their marginal costs, which are, in turn, used to perform counterfactuals.¹²

¹²The code to recover marginal costs and perform counterfactuals was written in Python, and it is available upon request.

4.3.1 Demand

To estimate our **logit** model, we analytically invert the expressions for the choice probabilities (2) as in S. T. Berry (1994) to obtain an estimate of the δ_j 's which are afterward regressed on our variables of interest to obtain the corresponding coefficients for the utility. Specifically, since the mean utility of the outside option is normalized to zero, the expression (S. T. Berry 1994, p. 250):

$$\delta_{jt} = \ln s_{jt} - \ln s_{0t} \quad (8)$$

resulting from inverting the logit probabilities pin downs analytically the δ_{jt} for each product j and market t . These δ_{jt} 's are afterward regressed against product characteristics according to the following specification:

$$\delta_{jt} = \vec{\beta}' \vec{x}_j - \alpha \log(p_{jt}(1 + \tau_j)) + \xi_{jt} \quad (9)$$

The results are reported in the Table 5, column (1).

The previous specification assumes p_{jt} is exogenous. To correct for price endogeneity, we employ the methodology of S. Berry, Levinsohn, and Pakes (1995), using differentiation instruments in the spirit of Gandhi and Houde (2019).¹³ Once the δ_j 's are obtained from equation 8, moment conditions obtained from the instruments' orthogonality conditions are used to obtain estimates for β and α using GMM and PyBLP. The results are reported in Table 5, column (2).

Finally, for the estimation of the **mixed logit**, we employed the two-step GMM method as implemented in PyBLP and used differentiation instruments to correct for endogeneity. To estimate

¹³The differentiation instruments proposed by Gandhi and Houde (2019) are a variation of the broadly-known BLP instruments, which are functions of product characteristics and leverage the fact that the price of product j depends on characteristics of other products and that dependence differs if these are their products or competitors' products. The differentiation instruments are intended to improve empirical performance and avoid weak IV challenges that can arise in practice by leveraging distances among products in the characteristics space. In particular, their "local" version captures localized competition through the number of "close-by" products in the characteristics space. Specifically, these instruments are constructed as follows: Given the matrix of characteristics X , x_{jtk} is the characteristic k in X for product j in market t , which is produced by firm f , that is, $j \in J_{ft}$. The instruments are constructed using the function $g_{jtk}(X) = \left[\sum_{j' \in J_{ft}} \mathbb{1}(|d_{jtkj'}| < SD_k), \sum_{j' \notin J_{ft}} \mathbb{1}(|d_{jtkj'}| < SD_k) \right]$, where $j \neq j'$, $d_{jtkj'} = x_{jtk} - x_{j'tk}$, and SD_k is the standard deviation of these pairwise differences computed across all markets. We leverage the PyBLP implementation of these instruments, together with their own product characteristics, which are assumed to be exogenous and, therefore, instruments by themselves. On top of Gandhi and Houde (2019) see also Gandhi and Nevo (2021)

the value of the integral in (5), we employed 1000 Halton draws. From integral (5), for any given values of $(\{\sigma_k\}_k, \sigma_p)$, the $\vec{\delta}$ values such that the predicted share equals the observed market share are calculated, which implies solving a fixed point problem, which is solved using a SQUAREM routine. The results are reported in Table 5, column (3).

Zero-valued market shares It is important to emphasize that certain representative products have no bookings for some dates despite being offered in the market. Such products, therefore, exhibit a market share of zero in those dates, an issue that happens in 778 day-category pairs (out of 40,493 day-category pairs). This is a well-known issue in estimating discrete choice models because it rejects any multinomial choice probabilities (McFadden 1974; Dubé, Hortaçsu, and Joo 2021). In this research, we consider the two most common solutions to this problem, which are either to eliminate observations with a zero market share (hereafter “Drop”) or to impute an extremely small value (hereafter “Imputation”); we use a value of 10^{-12} for “Imputation”.

Remarks on data Recall our final estimation dataset consists of market shares of the representative products and the outside option for every day in 2019 (section 3). The features of the representative products are the logarithm of the total amount paid by guests to rent an Airbnb accommodation (“ $\ln(\text{price}(1 + \tau))$ ”) with τ the lodging tax (3% in 2019) and the price in US dollars of January 2019; number of amenities in tens (“amenities”);¹⁴ the number of bathrooms (“bathrooms”); the number of bedrooms (“bedrooms”); the maximum number of guests allowed (“maxguests”); “professional”, a dummy variable indicating whether the host type for a given category has multiple listings on the platform; and “private room”, a dummy variable indicating whether the representative product is a private room or an entire home/apartment. Additionally, we employ fixed effects for the 16 Alcaldias of Mexico City and for each day of the year 2019. To create the differentiation instruments, we use “amenities”, “bathrooms”, “bedrooms”, and “maxguests”. To measure the size of the outside option for each day, we used the corresponding monthly average of hotel

¹⁴The AirDNA dataset registers a large variety of amenities each accommodation offers. The variable “amenities” was constructed by counting the number of amenities recorded in the database for each accommodation.

bookings per day reported in DataTur (see Table 2).

4.3.2 Supply

For each day, to obtain marginal costs that are consistent with market equilibrium, the first-order conditions for professional hosts and for each category of nonprofessional hosts,

$$s_{jt} + (p_{jt} - c_{jt}) \left(\frac{ds_{jt}}{dp_{jt}} \right) + \sum_{it \in \mathcal{J}_t^\Omega \setminus \{jt\}} \left[(p_{it} - c_{it}) \frac{ds_{it}}{dp_{jt}} \right] = 0, \forall j \in \mathcal{J}_t^\Omega \quad (10)$$

$$s_{jt} + (p_{jt} - c_{jt}) \left(\frac{ds_{jt}}{dp_{jt}} \right) = 0 \quad (11)$$

are inverted and evaluated using the demand estimates.

Table 5: Demand results

	(1) Logit			(2) Logit with endogeneity			(3) Mixed logit with endogeneity					
	Drop	Imputation		Drop	Imputation		Drop	Imputation				
	θ	SE	θ	θ	SE	θ	θ	SE	θ	SE		
β												
$\ln(\text{price}(1 + \tau))$	-1.09	0.01	-1.57	0.05	-1.97	0.02	-2.60	0.05	-1.90	0.12	-1.45	0.68
amenities	1.87	0.03	2.96	0.11	2.31	0.04	3.52	0.10	-0.87	5.87	-0.45	1.63
bathrooms	-0.18	0.03	-0.04	0.11	-0.21	0.03	-1.01	0.08	-1.80	0.33	-15.06	3.05
bedrooms	1.15	0.04	2.61	0.20	0.74	0.05	2.36	0.17	0.24	0.16	-1.41	0.64
maxguests	-0.55	0.02	-1.34	0.07	-0.01	0.02	-0.30	0.06	-0.05	0.07	-3.65	1.16
professional (0 = nonprofessional)	0.28	0.01	0.71	0.03	0.05	0.01	0.13	0.03	-0.10	0.03	-0.69	0.19
private room (0 = entire home/apt.)	-0.84	0.02	-1.85	0.06	-0.45	0.02	-0.44	0.04	-0.71	0.10	-4.73	0.92
intercept absorbed	Yes		Yes		Yes		Yes		Yes		Yes	
Date FE	Yes		Yes		Yes		Yes		Yes		Yes	
Alcaldia FE	Yes		Yes		Yes		Yes		Yes		Yes	
σ												
$\ln(\text{price}(1 + \tau))$									0.00	25.70	0.01	9.53
amenities									3.25	2.62	8.28	1.42
bathrooms									1.52	0.15	8.87	1.59
bedrooms									0.00	3.77	0.00	11.97
maxguests									0.00	2.26	2.62	0.76
intercept									3.19	4.93	5.63	1.13
N	39,715		40,493		39,715		40,493		39,715		40,493	

Notes: SE symbolizes robust standard errors. Prices are in dollars in January 2019 real values. The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} .

5 Results

In this section, we present our estimates for the models of demand and supply. These estimates are used to explore the equilibrium effects of a hypothetical policy taxing professional hosts within Airbnb.

5.1 Demand

Table 5 presents our estimations of the different demand specifications, i.e., the logit, logit corrected for endogeneity, and the mixed logit corrected for endogeneity, for both ways of addressing the zero market share problem.

Beginning with the logit model without addressing endogeneity (Table 5, column (1)), most coefficients are statistically significant at 95%, and their signs are consistent between the “Drop” and “Imputation” methodologies of handling the zero market share problem (with the exception of “bathrooms” in the “Imputation” method).

Controlling for endogeneity in prices using the IV approach outlined in the previous section (Table 5, column (2)), most of the estimates remain significant at 95% (with the exception of “maxguests” in the “Drop” method) and preserve the directions obtained in the simplest specification. Most coefficients increase in absolute value relative to the estimation, which does not correct the endogeneity issue, emphasizing the importance of correcting the endogeneity issue.

With the inclusion of heterogeneity in consumer preferences using the mixed logit, and correcting for endogeneity (Table 5, column (3)), we find that the parameters corresponding to the variance of the random taste shocks are only significant for “bathrooms” in both methodologies and “amenities”, “maxguests”, and the intercept in the imputation methodology. This is not surprising given the lack of demographic or other information varying at the guest level.

In our analysis, the most relevant estimates are those associated with “ $\ln(\text{price}(1 + \tau))$ ” due to their implications on the price elasticity of demand, which is crucial for the supply model estimations. Values close to 0 would correspond to inelastic demands, which are incompatible

with the profit maximization proposed in our supply model. Relative to these coefficients on “ $\ln(\text{price}(1 + \tau))$ ”, we observe they are consistent across models and ways of handling the zero market share, that they have the expected sign, and the values for logit models for the “Drop” versions are more negative in magnitude compared to their “Imputation” counterparts. For the mixed logit specification, the variances for “ $\ln(\text{price}(1 + \tau))$ ” are insignificant. The fact that the coefficients on the taste shock parameters are not different from zero offers statistical evidence in favor of the simpler logit model (Train 2009). Therefore, our preferred specification overall is the logit with endogeneity, which we use for subsequent analyses.

The discrepancy across models regarding the preferences for the physical characteristics of accommodations (“amenities”, “bathrooms”, “bedrooms”, and “maxguests”) or the insignificance of the estimates suggests that possibly their linear specification in consumer utility is not the most appropriate way to include them. However, since the analyses of the supply and counterfactual primarily depend on the price elasticities of demand, which are consistent across specifications and ways of handling the zero market share problem, we leave this subject matter for future research.

5.2 Supply

For the estimation of the model of the supply, we used the coefficients of our preferred specification for the demand model: logit with endogeneity. Table 6 presents regressions of our estimates of daily marginal costs (c_{jt}), equilibrium markups ($p_{jt} - c_{jt}$), and price-cost margins ($\frac{p_{jt} - c_{jt}}{p_{jt}}$) against host and accommodation type, Alcatraz, and date. The full distributions by host and accommodation type can be seen in the Appendix (Figures A.3 y A.4).

It is meaningful to note that both methods of addressing the zero market share issue yield similar results, as shown in Table 6.¹⁵ As a result, we will focus on the “Drop” method.

¹⁵There are some minor differences due to the differences in the magnitude of the coefficient associated with the price elasticity in the demand model.

Table 6: Results of the Supply Model

	MC				P-MC				(P-MC)/P			
	Drop		Imputation		Drop		Imputation		Drop		Imputation	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Intercept	22.39	0.00	28.22	0.00	23.22	0.00	17.65	0.00	0.5100	0.00	0.3854	0.00
Host type (0 = nonprofessional)												
Professional	-1.76	0.00	-1.60	0.00	3.67	0.00	2.31	0.00	0.0921	0.00	0.0702	0.00
Accommodation type (0 = Entire home/apartment)												
Private room	-20.72	0.00	-25.56	0.00	-21.49	0.00	-15.98	0.00	0.0301	0.00	0.0222	0.00
Alcaldia (0 = Azcapotzalco)												
Álvaro Obregón	5.42	0.00	6.80	0.00	5.63	0.00	4.25	0.00	-0.0201	0.00	-0.0152	0.00
Benito Juárez	7.75	0.00	9.95	0.00	8.04	0.00	6.22	0.00	-0.0141	0.00	-0.0108	0.00
Coyoacán	17.60	0.00	23.40	0.00	18.23	0.00	14.62	0.00	-0.0209	0.00	-0.0159	0.00
Cuajimalpa de Morelos	12.20	0.00	15.29	0.00	12.65	0.00	9.56	0.00	-0.0302	0.00	-0.0228	0.00
Cuauhtémoc	29.29	0.00	36.92	0.00	30.39	0.00	23.10	0.00	-0.0269	0.00	-0.0203	0.00
Gustavo A. Madero	-0.30	0.74	-0.38	0.74	-0.31	0.74	-0.24	0.74	0.0002	0.90	0.0001	0.90
Iztacalco	-2.25	0.01	-2.82	0.01	-2.33	0.01	-1.76	0.01	0.0093	0.00	0.0070	0.00
Iztapalapa	-1.85	0.04	-2.34	0.04	-1.92	0.04	-1.46	0.04	0.0066	0.00	0.0050	0.00
La Magdalena Contreras	3.90	0.00	4.93	0.00	4.03	0.00	3.08	0.00	-0.0121	0.00	-0.0093	0.00
Miguel Hidalgo	28.87	0.00	36.71	0.00	29.93	0.00	22.95	0.00	-0.0293	0.00	-0.0222	0.00
Milpa Alta	3.62	0.04	4.34	0.05	3.59	0.05	2.61	0.06	-0.0175	0.00	-0.0132	0.00
Tlalpan	2.48	0.01	3.12	0.01	2.58	0.01	1.95	0.01	-0.0077	0.00	-0.0058	0.00
Tláhuac	0.29	0.80	0.33	0.80	0.27	0.82	0.19	0.81	0.0416	0.00	0.0277	0.00
Venustiano Carranza	0.41	0.66	0.51	0.65	0.43	0.65	0.32	0.65	-0.0076	0.00	-0.0057	0.00
Xochimilco	0.71	0.45	0.76	0.51	0.74	0.45	0.47	0.51	0.0056	0.00	0.0035	0.00
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,715	40,493	39,715	40,493	39,715	40,493	39,715	40,493	39,715	40,493	40,493	40,493

Notes: The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} . Prices and marginal costs are in dollars in January 2019 real values.

Regarding host type, compared to nonprofessional hosts, professional hosts' accommodations have lower daily cost by \$1.76 ($p\text{-value} < 0.05$) and earn \$3.67 more ($p\text{-value} < 0.05$) per reserved accommodation, also have 9 percentage points ($p\text{-value} < 0.05$) higher price-cost margin. On the one hand, the lower marginal costs estimated for professional hosts are consistent with intuition that they might take advantage of economies of scale, scope, and superior management practices, which in turn explains the higher markups professional hosts are able to capture. On the other, it is important to notice that price-cost margins are overall high regardless of host type (the baseline in the most conservative estimation is 38%), which is consistent with the existence of large fixed costs (rent, mortgage, certain amenities like internet or cable service) and differentiated products.

Compared to renting an entire home/apartment, renting a private room costs \$20.72 less ($p\text{-value} < 0.05$) and earns \$21.49 less ($p\text{-value} < 0.05$), but has a higher price-cost margin of 0.0301 ($p\text{-value} < 0.05$). Therefore, while more dollars are earned by renting an entire home/apartment, renting by room yields a higher price-cost margin.

Regarding the different Alcaldias in Mexico City, Cuauhtémoc, Miguel Hidalgo, and Coyoacán have the highest daily costs but also the highest dollar gains. In contrast, Iztacalco, Iztapalapa, and Gustavo A. Madero have the lowest daily costs but also the lowest markups. Concerning price-cost margins, the highest are those of Tláhuac, Iztacalco, and Iztapalapa, and the lowest are those of Cuajimalpa, Miguel Hidalgo, and Cuauhtémoc.

5.3 Counterfactual Exercises

In 2019, guests making a reservation on Airbnb paid a 3% lodging ad valorem tax. As mentioned in the Introduction, governments around the world have implemented differential regulations for professional and nonprofessional hosts, with the former being subject to stricter rules. Consequently, we examine two counterfactual ad valorem tax policies that differentiate professional and nonprofessional hosts, one targeting guests and another targeting hosts.¹⁶

¹⁶While it has been shown that physical incidence of per unit taxes is irrelevant under very general conditions (Weyl and Fabinger 2013), it is a matter of elementary algebra to corroborate that even in the textbook examples physical incidence of ad valorem taxes is not irrelevant (Pauwels and Schroyen 2024).

Importantly, the counterfactual exercises were performed for both methodologies addressing the zero market share issue, and similarly to the demand and supply analyses, no major differences were found between the two approaches, which is reassuring. Therefore, as in the case of the supply model, we will focus on the estimates using the “Drop” method.

The first counterfactual examines the effects of a **consumption tax**, τ , of 5% (additional to the existing 3%) on accommodations belonging to professional hosts only. Table 7 presents the average changes (in percentage) relative to the current equilibrium estimates in subsection 5.2. As expected, the additional tax for booking with professional hosts decreases the price charged by professional hosts (by 1.53%), increases the total amount paid by a guest for staying in a professional host’s property (by 3.24%) and decreases professionals’ market share (by 5.52%). Naturally, markups and price-cost margins for professionals decrease. The market share lost by professionals is only marginally captured by nonprofessional hosts within the platform (0.57%), the remaining being captured by the traditional lodging sector (outside option). Considering the accommodation type, the effects for the entire home/apartment and private rooms are identical in signs, but those for private rooms are larger in absolute magnitude.

Table 7: Percentage comparison: Counterfactual vs. Current Equilibrium, Consumption Tax

	Professional				Accommodation type				Total	
	Yes		No		Entire home/apt.		Private room			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Drop										
$\Delta\%Price$	-1.53	1.07	0.00	0.00	-0.51	0.72	-1.04	1.30	-0.77	1.08
$\Delta\%Price(1+\tau)$	3.24	1.12	0.00	0.00	1.86	1.96	1.40	1.61	1.63	1.81
$\Delta\%Share$	-5.52	2.06	0.57	0.05	-2.91	3.64	-2.07	3.01	-2.49	3.37
$\Delta\%(P-MC)$	-2.40	1.40	0.00	0.00	-0.86	1.13	-1.57	1.83	-1.21	1.56
$\Delta\%((P-MC)/P)$	-0.88	0.39	0.00	0.00	-0.35	0.43	-0.54	0.58	-0.44	0.52
Imputation										
$\Delta\%Price$	-1.22	0.85	0.00	0.00	-0.41	0.58	-0.80	1.02	-0.60	0.85
$\Delta\%Price(1+\tau)$	3.58	0.89	0.00	0.00	1.97	2.04	1.58	1.72	1.78	1.90
$\Delta\%Share$	-7.97	2.11	0.81	0.08	-4.00	4.97	-3.09	4.23	-3.54	4.64
$\Delta\%(P-MC)$	-2.51	1.47	0.00	0.00	-0.90	1.19	-1.59	1.91	-1.25	1.63
$\Delta\%((P-MC)/P)$	-1.32	0.66	0.00	0.00	-0.50	0.63	-0.81	0.92	-0.65	0.81

Note: The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} .

The second counterfactual examines the effects of a **producer tax**, ρ , of 5% of the price of accommodations belonging to professional hosts only. Table 8 presents the average changes (in percentage) relative to the current equilibrium estimates in subsection 5.2. Not surprisingly, the tax on renting imposed on professional hosts increases the price charged by professional hosts (by 3.52%), increases the total amount paid by a guest for staying at a professional host’s property in the same proportion (3.52%), and decreases the market share of professionals (by 5.96%). As expected, markups and price-cost margins for professionals decrease. The market share lost by professionals is only marginally captured by nonprofessional hosts within the platform (0.61%), the remaining being captured by the traditional lodging sector (outside option). Considering accommodation type, the effects for entire homes/apartments and private rooms are identical in signs.

Table 8: Percentage comparison: Counterfactual vs. Current Equilibrium, Tax on hosts

	Professional				Accommodation type				Total	
	Yes		No		Entire home/apt.		Private room			
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Drop										
$\Delta\%Price$	3.52	1.21	0.00	0.00	2.02	2.12	1.52	1.74	1.77	1.96
$\Delta\%Price(1+\tau)$	3.52	1.21	0.00	0.00	2.02	2.12	1.52	1.74	1.77	1.96
$\Delta\%Share$	-5.96	2.22	0.61	0.05	-3.14	3.93	-2.24	3.26	-2.70	3.64
$\Delta\%(P-MC)$	-2.59	1.51	0.00	0.00	-0.93	1.22	-1.69	1.98	-1.30	1.68
$\Delta\%((P-MC)/P)$	-5.91	0.40	0.00	0.00	-2.83	2.87	-3.12	3.05	-2.97	2.97
Imputation										
$\Delta\%Price$	3.88	0.96	0.00	0.00	2.14	2.21	1.72	1.87	1.93	2.06
$\Delta\%Price(1+\tau)$	3.88	0.96	0.00	0.00	2.14	2.21	1.72	1.87	1.93	2.06
$\Delta\%Share$	-8.60	2.27	0.88	0.09	-4.32	5.36	-3.34	4.57	-3.83	5.00
$\Delta\%(P-MC)$	-2.71	1.59	0.00	0.00	-0.97	1.28	-1.71	2.06	-1.34	1.76
$\Delta\%((P-MC)/P)$	-6.35	0.68	0.00	0.00	-2.98	3.04	-3.32	3.36	-3.15	3.21

Note: The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} .

Comparing both counterfactual scenarios, from the consumer’s perspective, there is an overall higher cost to rent a property belonging to a professional host when the tax is directed at the host; however, it is important to note that the difference is marginal (0.28%). From the producer’s side, the market share of professional hosts experiences a greater percentage decrease when the tax is imposed on them for renting out an accommodation. However, the overall effect of either tax on

a professional's price-cost margins is a reduction of only about 3 percentage points relative to the baseline scenario.

6 Conclusion

Using a unique detailed panel dataset containing information on daily accommodations for Airbnb in Mexico City, we found descriptive evidence consistent with professional hosts playing a significant role in the platform, similar to what has been found elsewhere. Motivated by this fact, we estimated multiple demand and supply models to determine market structure indicators like marginal costs, markups, and price-cost margins.

Consistent with intuition, professional hosts' accommodations are found to be associated with lower marginal costs and larger markups, which indicates professional hosts have more ability to capture benefits, probably due to economies of scale, scope, and improved management practices. However, overall, price-cost margins are high and only slightly higher for professionals, which is consistent with the existence of large fixed costs (like mortgages and rents) and the differentiated products nature of the market.

Also, in line with intuition, marginal costs and markups for entire homes/apartments are larger than those for private rooms, but the latter has a larger per-dollar yield.

We examined two counterfactual tax policies that differentiate between professional and non-professional hosts. The counterfactual results are in line with fundamental economic intuition.

A remarkable and reassuring fact is that there are no substantial differences between the estimates obtained with the two most widespread methods of addressing the zero market share issue across all the exercises.

Overall, on top of the structural differences between professional and nonprofessional hosts in the Airbnb market in Mexico City, our findings highlight the usability of structural models for the counterfactual evaluation of policies for Airbnb. However, there is still room for improvement. In this first approach, we have abstracted away capacity constraints and endogenous choice sets,

issues that remain to be resolved on the research frontier. We also abstracted from entry and exit issues, whose impact varies across hosts and accommodation types, as well as the variation in the portfolios offered by professionals. All these are key avenues for future research. Last but not least, while the dataset used is on par with the used in the literature, being scrapped data, it lacks important information, such as guest-specific demographics, which is crucial to identifying taste variation.

References

- Abrate, Graziano, Ruggero Sainaghi, and Aurelio G Mauri (2022). “Dynamic pricing in Airbnb: Individual versus professional hosts”. In: *Journal of Business Research* 141, pp. 191–199.
- Agarwal, Nikhil and Paulo J Somaini (2022). *Demand analysis under latent choice constraints*. Tech. rep. National Bureau of Economic Research.
- Airbnb (2019). *Recaudación y Liquidación de Impuestos por Parte de Airbnb en México*. <https://www.airbnb.mx/help/article/2288/recaudacin-y-liquidacin-de-impuestos-por-parte-de-airbnb-en-mxico>. accessed: January 19, 2022.
- (2020). *Airbnb recolectó y pagó más de 200 mdp en Impuestos de hospedaje en México*. <https://news.airbnb.com/ea/airbnb-mexico-impuesto-hospedaje-2019/>. Accessed: January 19, 2022.
- Airbnb-Help Center (n.d.[a]). *AirCover for guests*. https://www.airbnb.com/help/article/3218?_set_bev_on_new_domain=1712857371_MzFjYWRlN2ViMTlk&locale=en. Accessed: April 10, 2024.
- (n.d.[b]). *Getting protected through AirCover for Hosts*. https://www.airbnb.com/help/article/3142?_set_bev_on_new_domain=1712857371_MzFjYWRlN2ViMTlk&locale=en. Accessed: April 10, 2024.
- (n.d.[c]). *Smart Pricing*. <https://www.airbnb.com/help/article/1168>. Accessed: April 10, 2024.

- Banco de México (2021). *El Mercado de Propiedades Residenciales de Alquiler Temporal en México*. Tech. rep., 28–33.
- Berry, Steven, James Levinsohn, and Ariel Pakes (1995). “Automobile Prices in Market Equilibrium”. In: *Econometrica* 63.4, pp. 841–890. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/2171802> (visited on 11/09/2023).
- Berry, Steven T (1994). “Estimating discrete-choice models of product differentiation”. In: *The RAND Journal of Economics*, pp. 242–262.
- Berry, Steven T and Philip A Haile (2021). “Foundations of demand estimation”. In: *Handbook of industrial organization*. Vol. 4. 1. Elsevier, pp. 1–62.
- Briel, Dorine von and Sara Dolnicar (May 2020). *THE EVOLUTION OF AIRBNB REGULATION - An International Longitudinal Investigation 2008-2020*. SocArXiv t4nqs. Center for Open Science. DOI: 10.31219/osf.io/t4nqs. URL: <https://ideas.repec.org/p/osf/socarx/t4nqs.html>.
- Calder-Wang, Sophie (2021). “The distributional impact of the sharing economy on the housing market”. In: *Available at SSRN 3908062*.
- Chen, Wei, Zaiyan Wei, and Karen Xie (2023). “Regulating professional players in peer-to-peer markets: Evidence from Airbnb”. In: *Management Science* 69.5, pp. 2893–2918.
- Conlon, Christopher and Jeff Gortmaker (2020). “Best practices for differentiated products demand estimation with pyblp”. In: *The RAND Journal of Economics* 51.4, pp. 1108–1161.
- Demir, Ersin and Gözde Emekli (2021). “Is Airbnb no longer a sharing economy platform? Evidence from Europe’s top 10 Airbnb destinations”. In: *Anatolia* 32.3, pp. 470–488.
- Ding, Kai, Yue Niu, and Wei Chong Choo (2023). “The evolution of Airbnb research: A systematic literature review using structural topic modeling”. In: *Heliyon* 9.6, e17090. ISSN: 2405-8440. DOI: <https://doi.org/10.1016/j.heliyon.2023.e17090>. URL: <https://www.sciencedirect.com/science/article/pii/S2405844023042986>.
- Dogru, Tarik et al. (2020). “Airbnb 2.0: Is it a sharing economy platform or a lodging corporation?” In: *Tourism Management* 78, p. 104049.

- Dubé, Jean-Pierre, Ali Hortaçsu, and Joonhwi Joo (2021). “Random-coefficients logit demand estimation with zero-valued market shares”. In: *Marketing Science* 40.4, pp. 637–660.
- EFE (Nov. 2023). *¿Airbnb en la mira? CDMX prepara reformas a la ley de alojamientos por plataformas digitales*. <https://www.elfinanciero.com.mx/cdmx/2023/11/13/airbnb-en-la-mira-cdmx-prepara-reformas-a-la-ley-de-alojamientos-por-plataformas-digitales/>. Accessed: November 13, 2023.
- Einav, Liran and Jonathan Levin (2010). “Empirical industrial organization: A progress report”. In: *Journal of Economic Perspectives* 24.2, pp. 145–162.
- Farhoodi, Abdollah (2021). “Democratizing the Opportunities: Who Benefits from the Airbnb Market?” In: *Available at SSRN 3480518*.
- Farronato, Chiara and Andrey Fradkin (2022). “The welfare effects of peer entry: the case of Airbnb and the accommodation industry”. In: *American Economic Review* 112.6, pp. 1782–1817.
- Gandhi, Amit and Jean-François Houde (Oct. 2019). *Measuring Substitution Patterns in Differentiated-Products Industries*. NBER Working Papers 26375. National Bureau of Economic Research, Inc. URL: <https://ideas.repec.org/p/nbr/nberwo/26375.html>.
- Gandhi, Amit, Zhentong Lu, and Xiaoxia Shi (2023). “Estimating demand for differentiated products with zeroes in market share data”. In: *Quantitative Economics* 14.2, pp. 381–418.
- Gandhi, Amit and Aviv Nevo (2021). “Empirical models of demand and supply in differentiated products industries”. In: *Handbook of industrial organization*. Vol. 4. 1. Elsevier, pp. 63–139.
- Gibbs, Chris, Daniel Guttentag, Ulrike Gretzel, Jym Morton, et al. (2018). “Pricing in the sharing economy: A hedonic pricing model applied to Airbnb listings”. In: *Journal of Travel & Tourism Marketing* 35.1, pp. 46–56.
- Gibbs, Chris, Daniel Guttentag, Ulrike Gretzel, Lan Yao, et al. (2018). “Use of dynamic pricing strategies by Airbnb hosts”. In: *International Journal of Contemporary Hospitality Management* 30.1, pp. 2–20.

- Gil, Javier and Jorge Sequera (2022). “The professionalization of Airbnb in Madrid: Far from a collaborative economy”. In: *Current Issues in Tourism* 25.20, pp. 3343–3362.
- Gyódi, Kristóf (2019). “Airbnb in European cities: Business as usual or true sharing economy?” In: *Journal of Cleaner Production* 221, pp. 536–551.
- Iacovone, Chiara (2023). “Debate on regulation and professionalisation in the short-term rental housing market”. In: *Geoforum* 146, p. 103870.
- Ki, Donghwan and Sugie Lee (2019). “Spatial distribution and location characteristics of Airbnb in Seoul, Korea”. In: *Sustainability* 11.15, p. 4108.
- Koh, Yoon, Amanda Belarmino, and Min Gyung Kim (2020). “Good fences make good revenue: An examination of revenue management practices at peer-to-peer accommodations”. In: *Tourism Economics* 26.7, pp. 1108–1128.
- Kwok, Linchi and Karen L Xie (2019). “Pricing strategies on Airbnb: Are multi-unit hosts revenue pros?” In: *International Journal of Hospitality Management* 82, pp. 252–259.
- Lee, Dayne (2016). “How Airbnb short-term rentals exacerbate Los Angeles’s affordable housing crisis: Analysis and policy recommendations”. In: *Harv. L. & Pol’y Rev.* 10, p. 229.
- Li, Jun, Antonio Moreno, and Dennis J. Zhang (2016). “Pros vs Joes: Agent Pricing Behavior in the Sharing Economy”. In: *Behavioral & Experimental Economics eJournal*. URL: <https://api.semanticscholar.org/CorpusID:867036>.
- López Tamayo, Diego Alberto and Aurora Alejandra Ramírez Álvarez (2021). *Análisis de Precios Hedónicos para Airbnb en la CDMX*. Serie documentos de trabajo del Centro de Estudios Económicos. El Colegio de México, Centro de Estudios Económicos. URL: <https://EconPapers.repec.org/RePEc:emx:ceedoc:2021-07>.
- Lorde, Troy, Jadon Jacob, and Quinn Weekes (2019). “Price-setting behavior in a tourism sharing economy accommodation market: A hedonic price analysis of AirBnB hosts in the caribbean”. In: *Tourism Management Perspectives* 30, pp. 251–261.

- Magno, Francesca, Fabio Cassia, and Marta Maria Ugolini (2018). “Accommodation prices on Airbnb: effects of host experience and market demand”. In: *The TQM Journal* 30.5, pp. 608–620.
- McFadden, Daniel (1974). “Conditional Logit Analysis of Qualitative Choice Behavior”. In: *Frontiers in Econometrics*.
- Oca, Álvaro Madrigal Montes de, Joan Sales Favà, and Antonio López Gay (2018). “El auge de Airbnb en la Ciudad de México : implicaciones espaciales y sociodemográficas”. In: URL: <http://www.recercat.cat/handle/2072/510837>.
- Pauwels, Wilfried and Fred Schroyen (2024). “The legal incidence of ad valorem taxes matters”. In: *Journal of Economics*, pp. 1–13.
- Popper, Ben (Dec. 2015). *Airbnb’s worst problems are confirmed by its own data*. <https://www.theverge.com/2015/12/4/9849242/airbnb-data-new-york-affordable-housing-illegal-hotels>. Accessed: November 13, 2023.
- Rosen, Sherwin (1974). “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”. In: *Journal of Political Economy* 82.1, pp. 34–55. ISSN: 00223808, 1537534X. URL: <http://www.jstor.org/stable/1830899> (visited on 11/11/2023).
- Ruiz-Correa, Salvador et al. (Feb. 2019). “Mi Casa Es Su Casa? Examining Airbnb Hospitality Exchange Practices in a Developing Economy”. In: *Trans. Soc. Comput.* 2.1. ISSN: 2469-7818. DOI: 10.1145/3299817. URL: <https://doi.org/10.1145/3299817>.
- Sainaghi, Ruggero, Graziano Abrate, and Aurelio Mauri (2021). “Price and RevPAR determinants of Airbnb listings: Convergent and divergent evidence”. In: *International Journal of Hospitality Management* 92, p. 102709. ISSN: 0278-4319. DOI: <https://doi.org/10.1016/j.ijhm.2020.102709>. URL: <https://www.sciencedirect.com/science/article/pii/S0278431920302619>.
- Skranka, Benjamin and Kenneth Judd (Feb. 2011). “High Performance Quadrature Rules: How Numerical Integration Affects a Popular Model of Product Differentiation”. In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.1870703.

- Tong, Baldwin and Ulrich Gunter (2022). “Hedonic pricing and the sharing economy: How profile characteristics affect Airbnb accommodation prices in Barcelona, Madrid, and Seville”. In: *Current Issues in Tourism* 25.20, pp. 3309–3328.
- Trade, United Nations Conference on and Development (Feb. 2020). *Digital platforms and value creation in developing countries: Implications for national and international policies*. https://unctad.org/system/files/official-document/tdb_ede4d2_en.pdf.
- Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge University Press. DOI: 10.1017/CBO9780511805271.
- Weyl, E Glen and Michal Fabinger (2013). “Pass-through as an economic tool: Principles of incidence under imperfect competition”. In: *Journal of political economy* 121.3, pp. 528–583.
- Xie, Karen, Cindy Yoonjung Heo, and Zhenxing Eddie Mao (2021). “Do professional hosts matter? Evidence from multi-listing and full-time hosts in Airbnb”. In: *Journal of Hospitality and Tourism Management* 47, pp. 413–421. ISSN: 1447-6770. DOI: <https://doi.org/10.1016/j.jhtm.2021.04.016>. URL: <https://www.sciencedirect.com/science/article/pii/S1447677021000668>.

Appendix

A.1 Tables

Table A.1: Listings during the year 2019 by Alcaldia

	Status		Total
	Non-Booked	Booked	
Alcaldia			
Cuauhtémoc	1,037,084	988,007	2,025,091
Miguel Hidalgo	450,836	324,680	775,516
Benito Juárez	472,477	270,237	742,714
Coyoacán	273,913	135,903	409,816
Álvaro Obregón	165,778	59,247	225,025
Tlalpan	139,861	49,241	189,102
Cuajimalpa de Morelos	65,499	26,115	91,614
Venustiano Carranza	55,282	34,284	89,566
Gustavo A. Madero	51,831	22,110	73,941
Iztacalco	40,873	22,697	63,570
Azcapotzalco	36,125	15,272	51,397
Iztapalapa	40,612	10,637	51,249
La Magdalena Contreras	21,746	8,132	29,878
Xochimilco	18,196	4,764	22,960
Tláhuac	1,698	985	2,683
Milpa Alta	26	262	288
Total	2,871,837	1,972,573	4,844,410

Notes: The listings correspond only to entire homes/apartments and private rooms. The number of listings corresponds to our final sample after cleaning the data.

A.2 Figures

Figure A.1: Listings price distribution before filtering by the interaction between the accommodation type and type of host.

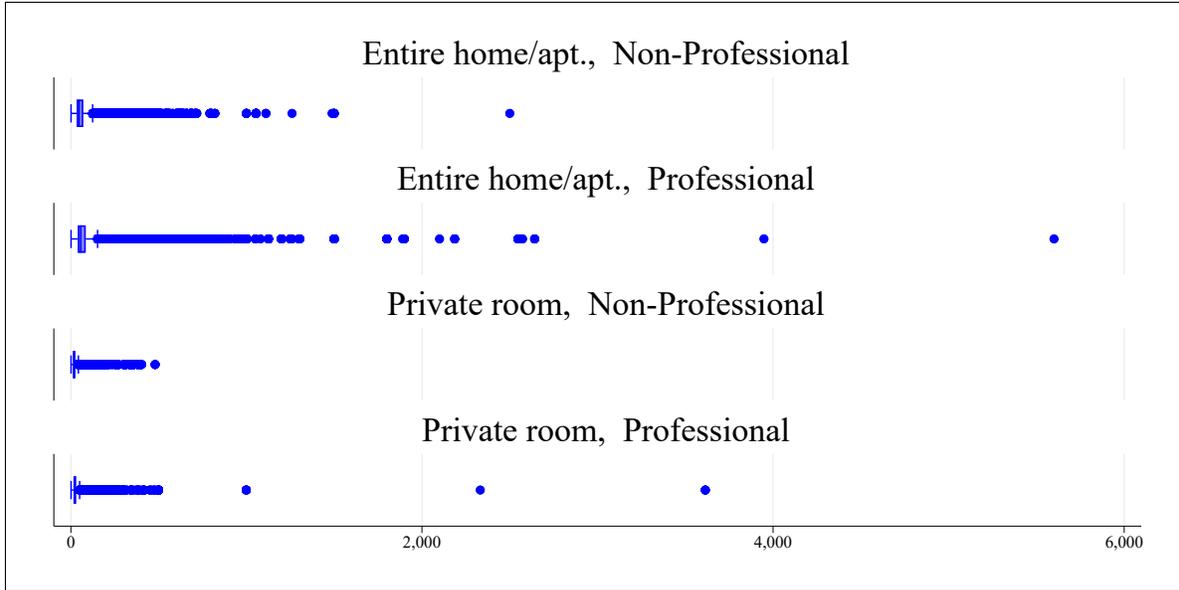


Figure A.2: Listings price distribution after filtering by the interaction between the accommodation type and type of host.

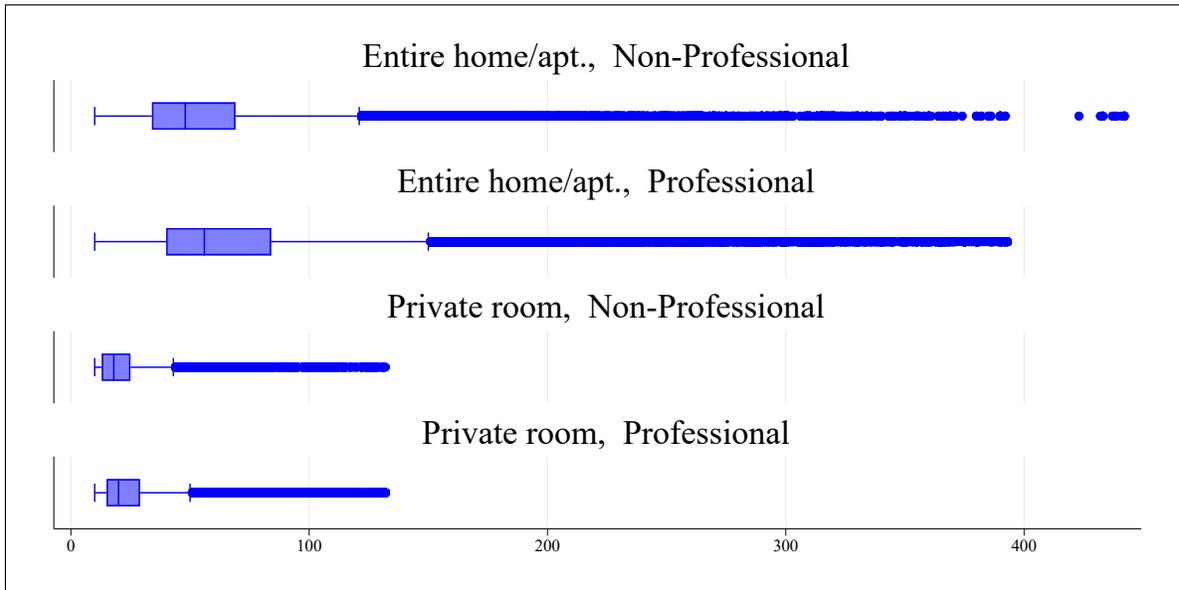


Figure A.3: Distribution of the marginal costs by the interaction between the accommodation type and type of host, drop method.

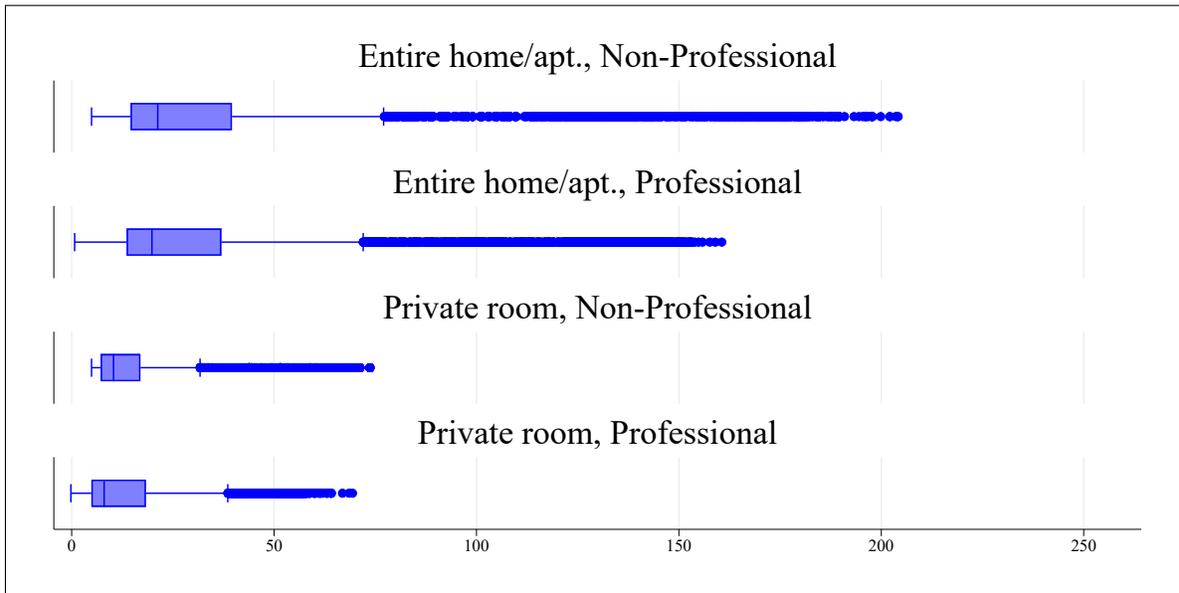
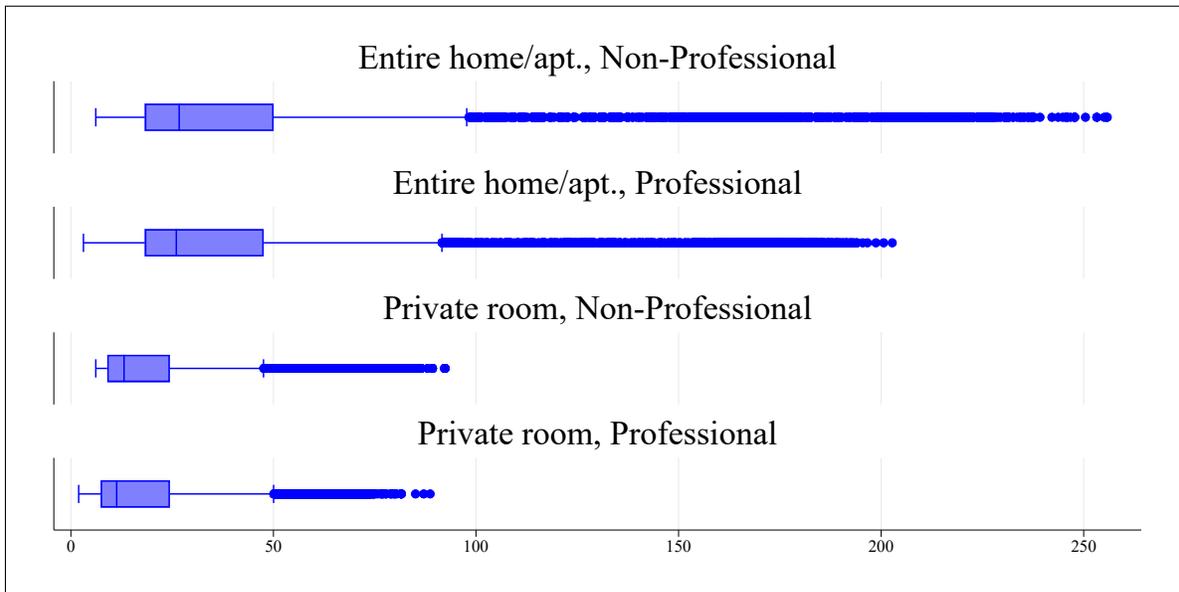


Figure A.4: Distribution of the marginal costs by the interaction between the accommodation type and type of host, imputation method.



A.3 Classification and aggregation

In section 3, we briefly described the procedure for generating product categories and corresponding representative products. In this subsection, we will provide more details.

To classify accommodations, we follow a procedure similar to the one described by Farronato and Fradkin (2022),¹⁷ and restrict our attention to reserved listings only (which is without loss since every accommodation in our final dataset was reserved at least once). Although the reason behind using only the information from reserved listings is not explicitly mentioned in Farronato and Fradkin (2022), this idea is supported by the hedonic prices literature (Rosen 1974).

The process we follow to categorize accommodations comprises a sequence of five steps in the following order: First, we run a hedonic regression where the natural logarithm of the prices of reserved listings depends on the fixed effects of accommodation and date.¹⁸ Second, we use Bayesian shrinkage to shrink the accommodation fixed effects towards the mean to mitigate the impact of sampling error.¹⁹ Third, we group individual accommodations by type (entire home/apartment or private room) and Alcaldia. Fourth, to establish quality tiers,²⁰ the range of shrunken accommodation fixed effects values of the accommodations for each group was evenly divided. Finally, each subgroup was split into two, distinguishing between accommodations belonging to professional hosts and those belonging to nonprofessional hosts. We refer to the product categories as the grouping of accommodations based on their Alcaldia, accommodation type (entire home/apartment or private room), quality tier, and host type (professional or nonprofessional).

For each category, a representative product is created, whose characteristics are the average of the features of the unique accommodations that make it up throughout the year; these remain

¹⁷Since our research question is different from that of Farronato and Fradkin (2022), our categorization is finer than the one implemented by them. They have 4 categories per city-date, while we have 116 potential categories for each date. The actual number varies with date as some categories might not always be offered.

¹⁸The hedonic regression has the following form: $\ln(\text{price}_{hd}) = \Gamma_h + \iota_d + \eta_{pd}$, where the *price* correspond to reserved listings, Γ represent the accommodation fixed effects of accommodation h , ι represent the date fixed effects of date d , and η is the error term. Other fixed effects, such as Alcaldia, accommodation type, superhost, and host type (professional or nonprofessional), are omitted due to collinearity.

¹⁹We use the term “shrunken accommodation fixed effect”, denoted as γ , to refer to the “accommodation fixed effects”, Γ , that have undergone the Bayesian shrinkage procedure.

²⁰Higher shrunken fixed effects of accommodation were used as a proxy for higher quality.

constant over time. On the other hand, the daily price for each category is calculated by averaging the prices of the accommodations within the category reserved for that day. Similarly, to calculate the quantity of the representative product sold for each date, we sum up the reservations of accommodations within the category for that date.²¹

Using the procedure mentioned earlier, we created 116 categories and representative products. It should be noted, however, that the number of categories per date tends to fluctuate throughout the year. Certain categories do not have any accommodations available on specific dates, and at times, even though certain categories have available accommodations, none may be rented, resulting in a market share of zero. On average, there are 108.81 categories with reserved accommodations per date. However, if we include categories with zero market share, the average daily number of categories increases to 110.94.

Quality tiers Ideally, for creating quality tiers, we would use quantiles to divide the shrunken accommodation fixed effects (which we denote as γ) to maintain tiers with an equal number of accommodations. However, firstly, as seen in tables A.2 and A.3, the number of accommodations—and thus shrunken accommodation fixed effects—by Alcaldía and accommodation type is generally low. Secondly, as observed in figures A.5 and A.6, in most cases the values are clustered around the mean. Therefore, if we used quantiles, we would have quality tiers with considerable dispersion of γ values, especially in the higher and lower quantiles, due to the general presence of outliers. For all of the above, we opted to divide the range of values into equal parts at the cost of having tiers with different numbers of accommodations, and this approach results in tiers with more homogeneous γ values. The number of qualities per Alcaldía was chosen based on the number of accommodations offered. We established four tiers for Benito Juárez, Coyoacán, and Miguel Hidalgo; five tiers for Cuauhtémoc; for the remaining Alcaldías, only one quality tier was designated.

²¹The average price of the available accommodations is used only for categories with zero market share on a specific date.

Table A.2: Statistics of shrunken fixed effects of accommodations corresponding to entire homes/apartments.

Alcaldía	Mean	SD	Median	Min	Max	N
Azcapotzalco	0.41	0.36	0.35	-0.22	1.44	102
Benito Juárez	0.64	0.45	0.63	-0.60	1.92	1698
Coyoacán	0.68	0.60	0.64	-0.70	2.68	715
Cuajimalpa de Morelos	1.12	0.50	1.13	-0.01	2.26	286
Cuauhtémoc	0.99	0.56	0.94	-0.36	2.96	6088
Gustavo A. Madero	0.32	0.51	0.30	-0.70	2.07	118
Iztacalco	0.11	0.44	0.06	-0.77	1.24	139
Iztapalapa	0.16	0.61	0.06	-0.78	1.73	92
La Magdalena Contreras	0.55	0.66	0.60	-0.63	1.85	59
Miguel Hidalgo	1.21	0.60	1.16	-0.20	3.02	2453
Milpa Alta	0.62	0.93	0.40	-0.79	1.96	9
Tlalpan	0.56	0.66	0.46	-0.65	2.21	327
Tláhuac	0.26	1.06	-0.04	-0.76	2.12	17
Venustiano Carranza	0.35	0.46	0.28	-0.55	1.80	225
Xochimilco	0.39	0.65	0.29	-0.78	2.98	64
Álvaro Obregón	0.80	0.56	0.80	-0.72	2.46	491

Table A.3: Statistics of shrunken fixed effects of accommodations corresponding to private rooms.

Alcaldía	Mean	SD	Median	Min	Max	N
Azcapotzalco	-0.30	0.44	-0.40	-0.83	1.79	184
Benito Juárez	-0.17	0.42	-0.21	-0.87	1.32	2251
Coyoacán	-0.12	0.56	-0.23	-0.89	1.75	1342
Cuajimalpa de Morelos	0.18	0.55	0.16	-0.80	1.73	218
Cuauhtémoc	0.06	0.51	-0.01	-0.89	1.92	3983
Gustavo A. Madero	-0.38	0.35	-0.43	-0.85	0.58	249
Iztacalco	-0.27	0.47	-0.38	-0.86	1.53	195
Iztapalapa	-0.41	0.41	-0.55	-0.89	0.83	207
La Magdalena Contreras	-0.11	0.50	-0.20	-0.80	1.53	96
Miguel Hidalgo	0.13	0.54	0.09	-0.82	1.86	1675
Milpa Alta	-0.14	0.39	-0.21	-0.58	0.54	8
Tlalpan	-0.26	0.53	-0.35	-0.89	2.01	643
Tláhuac	-0.65	0.19	-0.74	-0.90	-0.15	21
Venustiano Carranza	-0.23	0.46	-0.29	-0.89	1.14	230
Xochimilco	-0.32	0.43	-0.37	-0.82	1.24	99
Álvaro Obregón	-0.02	0.57	-0.09	-0.86	1.74	672

Figure A.5: Distribution of shrunken fixed effects corresponding to entire homes/apartments by Alcaldia.

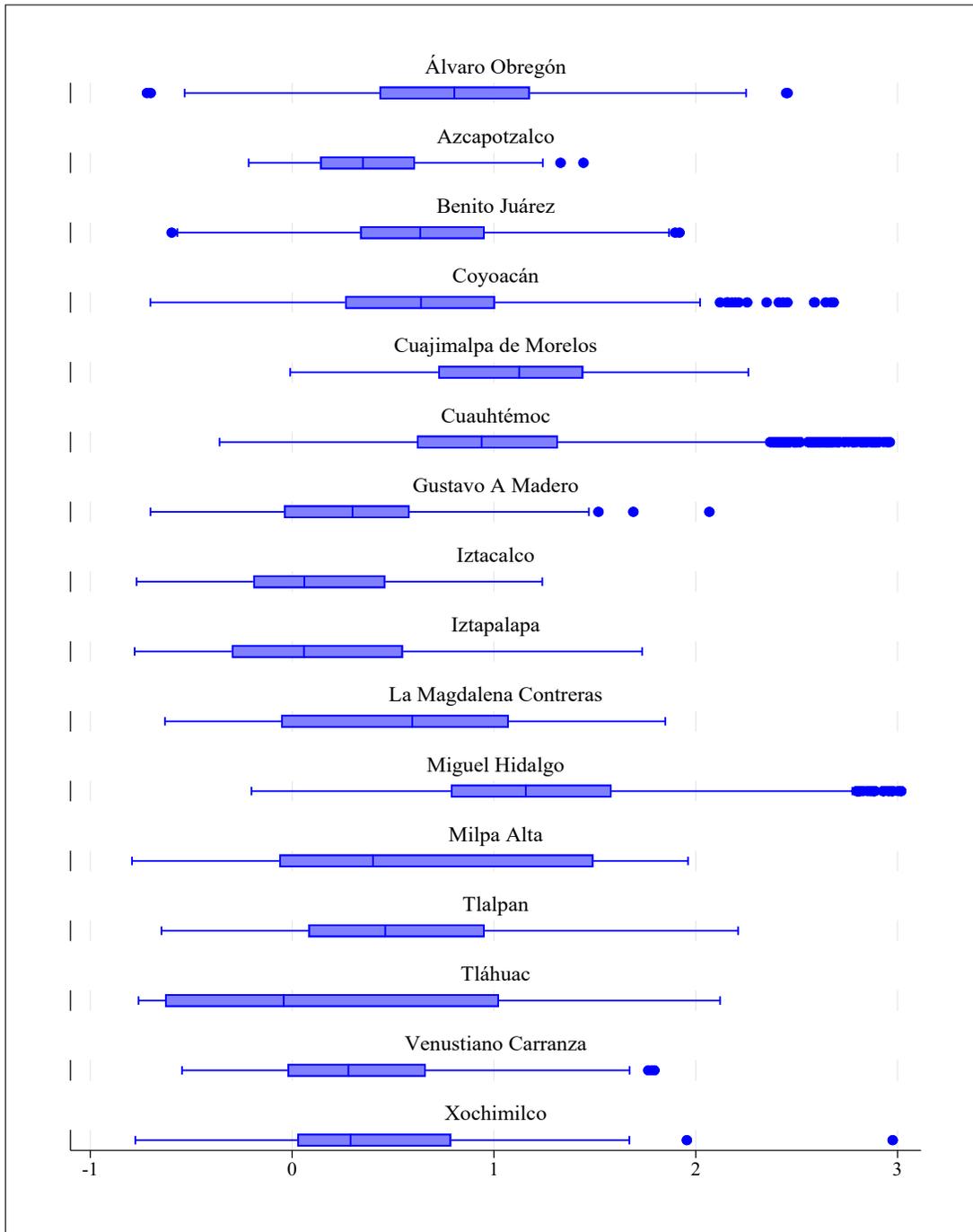
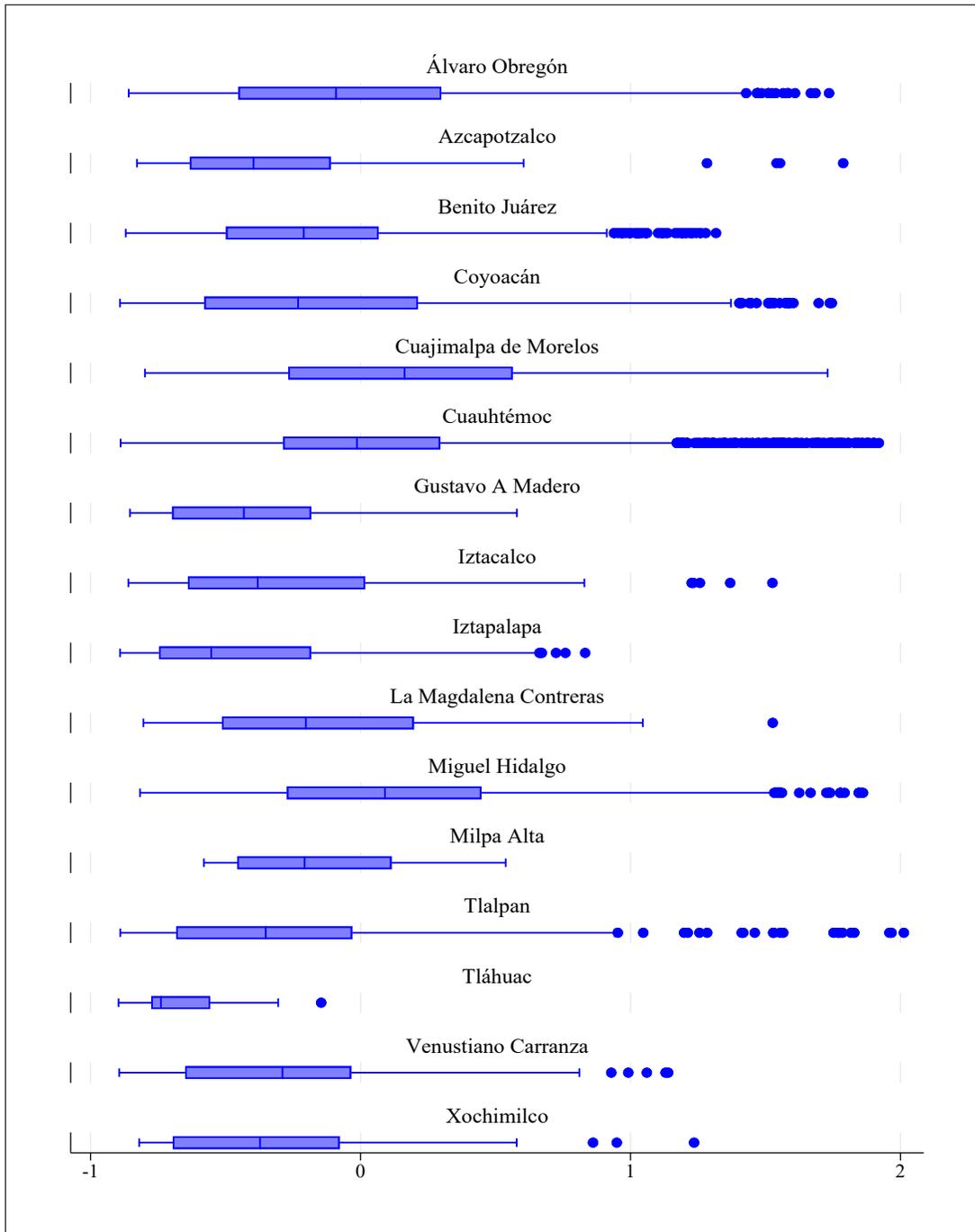


Figure A.6: Distribution of shrunken fixed effects corresponding to private rooms by Alcaldia.



A.4 Robustness

Our investigation compares nonprofessional hosts to professional hosts (which, following the literature, we define as multi-product hosts). We test the robustness of our results by modifying our definition of professional hosts. We divide multi-product hosts into six segments that are created from the interaction between the number of accommodations they offer (2 to 3, or at least 4) and the accommodation type they offer (only entire homes/apartments, only private rooms, or both). We use the same procedure for creating categories of products described in the section 3 (and detailed in the Appendix A.3) with the difference that we use the division of products by the number of accommodations of the host (1, 2 to 3, at least 4) instead of dividing products by single-product hosts and multi-product hosts.

This creates 275 categories and representative products in this robustness test, more than double the number of products used in our leading estimation (5). This means that we have more observations for both the “Drop” and “Imputation” methods of handling market shares of 0. The “Drop” method has 89,619 observations, while the “Imputation” method has 96,574 observations. In contrast, the dataset from section 5 had only 39,715 and 40,493 observations for the “Drop” and “Imputation” methods, respectively.

Consistent with this new product categorization, we modify our representative hosts as follows: Instead of a unique representative host for all categories belonging to multi-product hosts, we set up six decision-makers that originate from the interaction between the number of accommodations they offer (2 to 3, or at least 4) and the accommodation type they offer (only entire homes/apartments, only private rooms, or both).²² We keep a representative host for each category belonging to single-product hosts.

Demand results Table A.4 shows the estimation results. As in 5.1 the variance of the taste shocks in the mixed logit models are mostly statistically non-significant. Therefore, as in the body of the text we prefer the logit corrected for endogeneity.

²²In a sense, this heterogeneity of multi-unit hosts can be regarded as more realistic than our baseline specification.

The coefficients of main interest are those related to variable “ $\ln(\text{price}(1 + \tau))$ ”. These coefficients have a negative impact on consumer utility, which is as expected, and their magnitudes are similar to those reported in subsection 5.1. The coefficient for the “Drop” method is -1.53, while for the “Imputation” method, it is -2.87. In comparison, the coefficient for the “Drop” method in the subsection 5.1 is -1.97, and for the “Imputation” method, it is -2.60.

Supply results As in subsection 5.2, we used the logit coefficients corrected by endogeneity, and we will focus on the “Drop” method. The results are shown in Table A.5.

While subsection 5.2 estimates suggest that multi-product hosts generally have lower marginal costs by \$1.76 ($p\text{-value} < 0.05$), higher markups by \$3.67 ($p\text{-value} < 0.05$), and slightly higher price-cost margins by 0.0921 ($p\text{-value} < 0.05$) than single-product hosts, this is not always the case when we segment hosts by the number of accommodations. When comparing hosts who offer only one accommodation to those who offer 2 to 3 accommodations, the latter have lower costs by \$2.08 ($p\text{-value} < 0.05$), lower markups by \$2.98 ($p\text{-value} < 0.05$), and only slightly higher price-cost margin by 0.0101 ($p\text{-value} < 0.05$). Conversely, hosts with at least 4 accommodations have statistically equal costs than single-product hosts, but have higher markups by \$2.04 ($p\text{-value} < 0.05$) and higher price-cost margins by 0.0209 ($p\text{-value} < 0.05$).

For a host, when compared to renting out an entire home or apartment, renting out just a private room has a lower marginal cost of \$13.74 ($p\text{-value} < 0.05$), lower markups by \$26.75 ($p\text{-value} < 0.05$), and a lower price-cost margin of 0.0022 ($p\text{-value} < 0.05$). These estimates align in sign and magnitudes with those previously reported in subsection 5.2.

Table A.4: Robustness - Demand results

	Logit			Logit with endogeneity			Mixed logit with endogeneity												
	Drop	Imputation		Drop	Imputation		Drop	Imputation											
	θ	θ	SE	θ	θ	SE	θ	θ	SE										
β																			
$\ln(\text{price}(1 + \tau))$	-0.74	0.01		-2.08	0.04		-1.53	0.03		-2.87	0.12		-1.51	0.04		-4.14	5.17		
amenities	1.32	0.02		2.07	0.09		1.56	0.02		2.57	0.10		-0.38	0.15		3.06	0.40		
bathrooms	0.01	0.02		-0.87	0.09		-0.10	0.02		-0.54	0.08		-0.71	0.13		-2.59	1.14		
bedrooms	0.09	0.02		2.30	0.11		0.22	0.02		2.05	0.11		0.02	0.12		-6.69	4.95		
maxguests	-0.18	0.01		-0.77	0.04		0.10	0.01		-0.65	0.05		0.17	0.02		-0.05	0.55		
Host by number of accommodations (0 = 1 accommodation)																			
2 to 3	-1.15	0.01		-2.55	0.04		-1.25	0.01		-2.39	0.04		-1.26	0.01		-2.95	0.14		
at least 4	-1.04	0.01		-1.66	0.04		-1.22	0.01		-1.77	0.04		-1.24	0.02		-2.74	0.32		
Private room (0 = Entire home/apt.)	-0.62	0.01		-1.59	0.05		-0.42	0.01		-1.72	0.05		-0.45	0.02		-2.73	0.10		
intercept absorbed	Yes		Yes	Yes		Yes	Yes		Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes		
Date FE	Yes		Yes	Yes		Yes	Yes		Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes		
Alcaldia FE	Yes		Yes	Yes		Yes	Yes		Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes		
σ																			
$\ln(\text{price}(1 + \tau))$																0.00	18.20	1.45	2.53
amenities																1.55	0.09	0.00	30.00
bathrooms																0.61	0.10	0.84	1.87
bedrooms																0.19	0.23	4.31	1.19
maxguests																0.00	4.64	0.00	22.49
intercept																0.00	229.51	0.02	226.20
N	89,619		96,574	89,619		96,574	89,619		96,574	89,619		96,574	89,619		96,574				

Notes: SE symbolizes robust standard errors. The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} .

Table A.5: Robustness - Supply results

	MC				P-MC				(P-MC)/P			
	Drop		Imputation		Drop		Imputation		Drop		Imputation	
	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Intercept	15.99	0.00	30.22	0.00	30.63	0.00	16.32	0.00	0.6599	0.00	0.3521	0.00
Host by number of accommodations (0 = 1 accommodation)												
2 to 3	-2.08	0.00	-1.97	0.00	-2.98	0.00	-0.80	0.00	0.0101	0.00	0.0052	0.00
at least 4	0.03	0.84	0.32	0.25	2.04	0.00	0.74	0.00	0.0209	0.00	0.0113	0.00
Accommodation type (0 = Entire home/apartment)												
Private room	-13.74	0.00	-26.02	0.00	-26.75	0.00	-14.18	0.00	-0.0022	0.00	-0.0012	0.00
Alcaldía (0 = Azcapotzalco)												
Álvaro Obregón	4.14	0.00	7.35	0.00	7.72	0.00	3.93	0.00	-0.0079	0.00	-0.0040	0.00
Benito Juárez	5.13	0.00	10.41	0.00	9.60	0.00	5.57	0.00	-0.0043	0.00	-0.0023	0.00
Coyoacán	10.53	0.00	22.52	0.00	19.70	0.00	12.04	0.00	-0.0063	0.00	-0.0034	0.00
Cuajimalpa de Morelos	10.15	0.00	18.90	0.00	19.03	0.00	10.10	0.00	-0.0099	0.00	-0.0052	0.00
Cuauhtémoc	19.25	0.00	37.45	0.00	36.14	0.00	20.04	0.00	-0.0087	0.00	-0.0045	0.00
Gustavo A. Madero	-0.19	0.67	-1.04	0.19	-0.45	0.58	-0.57	0.18	0.0004	0.19	0.0008	0.00
Iztacalco	-1.04	0.02	-2.96	0.00	-2.09	0.01	-1.60	0.00	0.0007	0.02	0.0018	0.00
Iztapalapa	-0.75	0.10	-1.89	0.02	-1.48	0.08	-1.01	0.02	0.0013	0.00	0.0008	0.00
La Magdalena Contreras	4.33	0.00	7.52	0.00	8.04	0.00	4.01	0.00	-0.0064	0.00	-0.0024	0.00
Miguel Hidalgo	18.66	0.00	35.71	0.00	35.01	0.00	19.10	0.00	-0.0094	0.00	-0.0049	0.00
Milpa Alta	2.12	0.07	3.90	0.08	3.78	0.09	2.04	0.08	-0.0068	0.00	-0.0035	0.00
Tlalpan	2.38	0.00	4.06	0.00	4.41	0.00	2.17	0.00	-0.0035	0.00	-0.0017	0.00
Tláhuac	0.27	0.70	0.23	0.84	0.37	0.78	0.11	0.86	0.0030	0.00	0.0017	0.00
Venustiano Carranza	0.62	0.15	0.83	0.28	1.12	0.16	0.44	0.29	-0.0037	0.00	-0.0018	0.00
Xochimilco	1.06	0.03	0.88	0.29	1.81	0.06	0.46	0.31	-0.0049	0.00	-0.0019	0.00
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	89,619	96,574	96,574	89,619	89,619	96,574	89,619	96,574	89,619	96,574	89,619	96,574

Notes: The Drop method excludes observations with zero market share, while the Imputation method keeps these observations but imputes them with a market share value of 10^{-12} . Prices and marginal costs are in dollars in January 2019 real values.