Location, location and professionalization: a multilevel hedonic analysis of Airbnb listing prices and revenue

Robbin Deboosere, Danielle Jane Kerrigan, David Wachsmuth & Ahmed El-Geneidy

Published online: 03 Apr 2019.

Submit your article to this journal ↗

Article views: 366

View Crossmark data ↗

Routledge
Taylor & Francis Group

Regional Studies Association

Location, location and professionalization: a multilevel hedonic analysis of Airbnb listing prices and revenue

Robbin Deboosere[a], Danielle Jane Kerrigan ⓘ[b], David Wachsmuth ⓘ[c] and Ahmed El-Geneidy ⓘ[d]

ABSTRACT

Hedonic modelling techniques have frequently been used to examine real estate valuation, and they have recently started to be applied to short-term rental valuation. Relying on a web-scraped data set of all Airbnb transactions in New York City (NYC) between August 2014 and September 2016, this paper presents the first hedonic regression model of Airbnb to take into account neighbourhood effects and to predict both average price per night and revenue generated by each listing. The model demonstrates that locational factors – above all, transit accessibility to jobs – and neighbourhood variation have a large impact on both price per night and monthly revenue, and further reveals how professionalization of the short-term rental market is driving more revenue to a narrower segment of hosts. Further, the findings suggest that Airbnb hosts earn a significant premium by converting long-term housing in accessible residential neighbourhoods into de facto Airbnb hotels. This premium incentivizes landlords and hosts with properties in accessible neighbourhoods to replace long-term tenants with short-term guests, forcing those in search of housing to less accessible neighbourhoods.

INTRODUCTION

In the last decade, short-term rental services – led by Airbnb – have begun to transform housing markets in cities around the world. Since its founding in 2008, Airbnb has expanded to 4.5 million listings worldwide, hosts have earned US$41 billion and Airbnb listings have accumulated over 300 million unique stays. Within a decade, Airbnb estimates its listings will host 1 billion guests per year (Airbnb, 2018). While short-term rentals existed in various forms before the advent of Airbnb and other online platforms such as VRBO, HomeAway

CONTACT

✉ robbin.deboosere@mail.mcgill.ca
[a] School of Urban Planning, McGill University, Montreal, QC, Canada.
(Corresponding author) ✉ danielle.kerrigan@mail.mcgill.ca
[b] School of Urban Planning, McGill University, Montreal, QC, Canada.
✉ david.wachsmuth@mcgill.ca
[c] School of Urban Planning, McGill University, Montreal, QC, Canada.
✉ ahmed.elgeneidy@mcgill.ca
[d] School of Urban Planning, McGill University, Montreal, QC, Canada.

and Wimdu, the sheer scale of the current services and the ability for individual homeowners and landlords to list their properties easily have created massive new revenue streams in urban property markets.

Recent research suggests that these new short-term rental revenue flows have had wide-ranging impacts. Increasingly scholars have argued that the comparatively higher potential revenue of short-term rentals has led to displacement or exacerbated housing affordability issues for local residents (Barron, Kung, & Proserpio, 2017; Gant, 2016; Horn & Merante, 2017; Lee, 2016; Mermet, 2017; Wachsmuth & Weisler, 2018; Wachsmuth, Chaney, Kerrigan, Shillolo, & Basalaev-Binder, 2018; Wachsmuth, Kerrigan, Chaney, & Shillolo, 2017). Additionally, research has shown that short-term rentals have negatively impacted the hotel industry (Guttentag, 2015; Guttentag & Smith, 2017; Oskam & Boswijk, 2016; Zervas, Proserpio, & Byers, 2017). These new revenue flows have also been shown to have a positive impact on local economies generally, and the tourism industry specifically (Gottlieb, 2013; Guttentag, 2015; Oskam & Boswijk, 2016; Zervas et al., 2017).

A fundamental question underlying this research is what factors are responsible for the performance of short-term rentals. What predicts a frequently booked listing, and what predicts a high-revenue listing? Hedonic modelling techniques, developed by Lancaster (1966) and Rosen (1974), have frequently been used to examine real estate valuation. These techniques use regression analysis to disentangle the bundle of different factors that affect real-estate prices, for example, by measuring the independent impact on prices of a property being located near a park, or having an additional bedroom. These techniques have recently started to be applied to short-term rentals, principally to identify the factors that predict higher nightly prices. Several scholars have examined the relationship between Airbnb listing price from the perspective of a hedonic analysis. but have been limited in important respects related to data availability and model sophistication (Gibbs, Guttentag, Gretzel, Morton, & Goodwill, 2018; Teubner, Hawlitschek, & Dann, 2017; Wang & Nicolau, 2017).

Relying on a complete data set of all Airbnb transactions within the boundaries of the municipality of NYC between August 2014 and September 2016, this paper presents the first hedonic regression model of Airbnb to take into account neighbourhood effects and to predict both average price per night and revenue generated by each listing. The model demonstrates that locational factors – above all transit accessibility to jobs – and neighbourhood variation have a large impact on both price per night and monthly revenue, and further reveals how professionalization of the short-term rental market is driving more revenue to a narrower segment of hosts. The findings suggest that Airbnb hosts earn a significant premium by converting long-term housing in transit-accessible residential neighbourhoods into de facto Airbnb hotels. This premium incentivizes landlords with properties in these neighbourhoods to replace long-term tenants with short-term guests – forcing those in search of housing into less accessible neighbourhoods.

LITERATURE REVIEW

Hedonic regression models are often used in real estate analysis to separate the effects of individual housing characteristics on home prices, under the assumption that the price of housing reflects the composition of the underlying bundle of attributes. Hedonic models, therefore, describe the price of housing through a number of objectively measured characteristics that together are thought to constitute the full package of housing attributes (Lancaster, 1966; Rosen, 1974). From the perspective of real estate markets, these hedonic modelling techniques have allowed scholars to determine the impacts of highly location-specific characteristics on real estate prices (Glaesener & Caruso, 2015; Li & Brown, 1980). They have also allowed scholars to measure the impact on house prices of different forms of public transit (Armstrong & Rodriguez,

2006; Benjamin & Sirmans, 1996; Chatman, Tulach, & Kim, 2012; Lewis-Workman & Brod, 1997), accessibility (Franklin & Waddell, 2003), walkability, (Rauterkus & Miller, 2011), and even the impact of bike-sharing stations (El-Geneidy, Lierop, & Wasfi, 2016).

While typically applied to housing prices, hedonic models have also been used to measure the impact of certain factors on hotel room pricing (Bull, 1994; Chen & Rothschild, 2010; Herrin & Carvell, 1990; Lee & Jang, 2011; Zhang, Ye, & Law, 2011). Key factors examined have been the location of hotels in relation to certain amenities, star rating and service quality. Hotel location can be conceptualized in relation to the city centre (Bull, 1994) or to certain business or tourist amenities (Herrin & Carvell, 1990; Zhang et al., 2011). In general, the findings of this research have been mixed. Herrin and Carvell (1990) note that prices drop the further a hotel is located from San Francisco's major tourist destination of Fisherman's Wharf, whereas Zhang et al. (2011) do not observe a significant effect from the presence of scenic spots, but find a price drop the further a hotel is located from transportation hubs.

With the rapid growth of short-term rentals over the last several years, scholars have begun to examine price determinants of Airbnb listings (Gibbs et al., 2018; Hill, 2015; Lee et al., 2015; Li, Moreno, & Zhang, 2015; Teubner et al., 2017; Wang & Nicolau, 2017). Airbnb has its own hedonic pricing algorithm, the development of which is described by Hill (2015). It uses three major elements in suggesting a listing price: similarity, recency and location. The similarity element predicts the successful price of a new listing by comparing it with existing listings which are similar in a number of different features, including the listing type (private room, entire home or apartment, and shared room), how many people it sleeps, the type of property (condo, house or even 'cave'), and the number of reviews. The recency element adjusts projected listing prices for seasonality and non-cyclical pricing changes. Finally, the location element predicts the impact of location on pricing, given that Airbnb listings are more broadly distributed than hotels, and given the importance of neighbourhood amenities which cannot be determined by a simple distance from the city centre (Hill, 2015).

The hedonic analyses undertaken by other scholars confirm the importance of Hill's (2015) factors of similarity, although not always with the same direction of the effect. Elements such as 'superhost' status (a distinction granted by Airbnb to hosts who rent frequently, receive high reviews and are responsive in their communications), and better reviews were identified as having a positive impact on price (Teubner et al., 2017; Wang & Nicolau, 2017). The number of reviews received by a listing was identified by Hill (2015) as having a positive impact on price. In direct contrast, Gibbs et al. (2018), looking at five Canadian cities, found that a higher number of reviews resulted in a lower average price. Regarding the impact of location, these studies only examined the Euclidian distance from either the city centre or city hall (Gibbs et al., 2018; Teubner et al., 2017; Wang & Nicolau, 2017), and thus were unable to take account of heterogeneity of neighbourhoods and neighbourhood amenities.

Li et al. (2015) examined the impact of professional hosts – those who manage more than one Airbnb listing – in Chicago. These professional hosts were found to have 16.9% higher daily revenue on the properties they manage, and a 15.5% higher occupancy rate. Gibbs et al. (2018), in their hedonic analysis, by contrast, only found professional hosts (using the same definition) earning a premium in Montreal, and this amount was only 3.5% greater than non-professional hosts.

In sum, existing hedonic research on Airbnb in relation to the existing Airbnb price literature has identified a set of important factors impacting prices and listing performance, but with at times contradictory findings. A major limitation of this research, moreover, has been a lack of sensitivity to a listing's location, which is typically either not addressed or is addressed with a simple straight-line distance model. The present study overcomes this limitation by incorporating a more sophisticated set of neighbourhood and location factors into its model, including transit accessibility to jobs and land-use diversity. The study thus offers important new evidence

about the determinants of home-sharing performance through multilevel, mixed-effects modelling techniques, which account for variance in the model at different levels to minimize external bias in the coefficients generated in the model.

DATA AND METHODOLOGY

This study uses a hedonic regression analysis to determine what factors influence individual Airbnb listings' average price per night and average revenue per month. Building on Gibbs et al. (2018), Wang and Nicolau (2017) and Teubner et al. (2017), it explores similarity factors (Hill, 2015) concerning a listing's physical attributes, such as the number of bedrooms, and concerning the hosts themselves, such as their status as a 'superhost'. Importantly, and unlike previous research, the study also includes multiple location and neighbourhood variables, including if the listing is within 800 m of a subway station (a commonly used threshold to estimate whether a resident will walk to access rail transportation), accessibility to jobs by public transit, land-use diversity, population density, median household income, the percentage of non-Hispanic White people and the number of Airbnb listings per census tract. Seasonality was also incorporated into the analysis by examining month-to-month variation in the average price per night and revenue per month. The core idea of using a hedonic analysis is housing units, which consist of a series of structural and locational variations that are inseparable; therefore, a model needs to incorporate more than just individual transactions of various structural and locational attributes (Orford, 2000). Using a hedonic analysis allows more accurate modelling of price by allowing bundles of housing attributes to be examined.

The hedonic model was generated from data for all Airbnb listings in NYC that were available at any point from August 2014 to September 2016 (the maximal timeframe for which all NYC data were available at the time the study was conducted). These data were scraped from the official Airbnb webpage by AirDNA, a third-party data and analytics firm that tracks Airbnb listings. The data set provides structural information about each listing, including the number of bedrooms and bathrooms, the maximum number of guests, if the host was a superhost, if the unit could be instantly booked, and the overall rating of the listing based on guest feedback (between 0 and 5 stars). The data also provide daily estimates of reservation activity – whether the listing was reserved, available or blocked – along with listed nightly prices. (Until late 2015, activity information was directly scrapable from Airbnb, after which point the firm began obfuscating whether an unavailable property was reserved or blocked. Using machine learning on their existing canonical performance data, AirDNA developed an occupancy model to estimate reservations since late 2015.) While each Airbnb listing has precise latitude–longitude coordinates associated with it, these coordinates do not correspond to the exact location of a listing but rather a random point within a 200 m radius of the listing's actual location. This is the result of spatial obfuscation Airbnb applies to its listings, and while there are strategies for dealing with this obfuscation in the context of aggregated spatial analysis (e.g., at the scale of census tracts or neighbourhoods), for the purposes of point-level analysis the lowest-error option is to use the reported coordinates without modification. (See Wachsmuth & Weisler, 2018, for a further discussion of Airbnb's spatial obfuscation, along with a broader discussion of the challenges of working with web-scraped data.)

The hedonic model used only apartment listings that had generated revenue in at least two months of the 26-month period over which the data were available. The model excluded non-apartments because of the difficulty in accounting for comparisons across the wide variety of listing types in NYC (which included 'caves', 'taxi cabs' and 'boats') and because apartments are by far the most common listing type, accounting for 81.7% of all listings. In traditional hedonic analyses, repeat sales have been shown to be an important factor in increasing the efficiency and accuracy of the model (Bailey, Muth, & Nourse, 1963). Limiting the study sample to listings that had

generated revenue in at least two separate months allows this model to be more accurate. Finally, to ensure the analysis only includes listings actively reserved during this period, listings with no user reviews were discarded. The resulting population for this model was 386,153. This population is equal to the number of unique listings multiplied by the number of months that the listing had revenue. Using a geographic information system (GIS) the x,y-coordinates of each listing were mapped for further spatial analysis.

The model has two separate dependent variables and 20 independent variables, presented in Table 1. The first dependent variable is the average price per night, which for the study population was US$153.16. This is the same dependent variable used by Gibbs et al. (2018), Wang and Nicolau (2017) and Teubner et al. (2017). The second dependent variable is the average revenue per month, which was US$2287.19. There are three categories of independent variables: structural, host, and location and neighbourhood. All data on an individual listing's structural and host variables come from the AirDNA data set. Structural variables included are the number of bedrooms (mean = 1.1), the number of bathrooms (mean = 1.1), the maximum number of guests (mean = 3.0), a dummy variable that equals 1 if the listing is an entire

Table 1. Model variables.

Variable	Description	Mean[a]	SD
Price	Average price per night (US$)	153.16	113.25
Revenue	Revenue per month (US$)	2287.19	2352.45
Structural variables			
Entire apartment	Dummy variable if the entire apartment is rented	0.6109	0.4875
Shared room	Dummy variable if the room is shared	0.0274	0.1633
Bedrooms	Number of bedrooms	1.1112	0.6606
Bathrooms	Number of bathrooms	1.0850	0.3262
Guests	Maximum number of guests	2.9785	1.7281
Host variables			
Superhost	Dummy variable if the host is a superhost	0.0868	0.2816
Multihost – 2–10	Dummy variable if the host has between two and 10 listings	0.3655	0.4816
Multihost – >10	Dummy variable if the host has more than 10 listings	0.0165	0.1272
Number of photos	Number of photographs for a listing	14.2015	10.8734
Instant booking	Dummy variable for is an instant booking is available	0.1292	0.3355
Reviews	Number of reviews	31.5926	36.1101
Overall rating	Overall star rating of listing	4.5738	0.3934
Location and neighbourhood variables			
Subway	Dummy variable if the listing is within 800 m of a subway stop	0.9038	0.2948
Transit accessibility	Accessibility to jobs by transit in 30 min (100,000 jobs)	13.1812	8.1446
Dist Times Square	Network distance to Times Square (km)	6.4343	3.9046
Diversity	Land-use diversity, measured by an entropy index	0.9557	0.2867
Population density	Population density in the census tract (1000 residents/km^2)	29.1037	15.4984
Median household income	Median household income in the census tract (US$1000)	74.1427	36.4930
Percentage white	Percentage of residents who identify as white in the census tract	47.5045	26.6700
Listing density	Number of Airbnb listings in the census tract	322.2075	261.7637

Note: [a]The mean of a dummy variable represents the percentage of all listings that have a value of 1 for that dummy variable.

apartment, and a dummy variable that equals 1 if the listing is a shared room. These structural variables were included based on previous studies (Gibbs et al., 2018; Hill, 2015; Lee et al., 2015; Teubner et al., 2017; Wang & Nicolau, 2017).

The host variables used were: a dummy variable equal to 1 if the host has superhost status; a dummy variable equal to 1 if the host has between two and 10 listings; a dummy variable equal to 1 if the host has more than 10 listings; and a dummy variable equal to 1 if the instant book feature was available. Additionally, the model used variables for the number of photographs included with each listing (mean = 14.2), the number of reviews (mean = 31.6) and the overall rating of the listing (mean = 4.6). While Teubner et al. (2017) looked at details such as the host's gender, age and if they were smiling in their profile photograph, they found no significant price effects based on these details. Accordingly, such variables were not included in this study, and host variables, in general, were chosen on previous studies (Gibbs et al., 2018; Hill, 2015; Lee et al., 2015; Wang & Nicolau, 2017).

Multiple location and neighbourhood variables were used. These included: a dummy variable equal to 1 if a listing is located within 800 m of a subway station (approximately a 10-min walk); the number of jobs that can be reached within 30 min by transit from each listing ('job accessibility' per 100,000 jobs; mean = 13.2); the network distance to Times Square (mean = 6.4 km); the land-use diversity measured by an entropy index (mean = 0.96); the population density per census tract (1000 residents/km^2; mean = 29.1); the median household income (in US$1000s; mean = 74.1); and the percentage of residents who identify as non-Hispanic White (mean = 47.5). Accessibility to jobs is used as a proxy for accessibility to tourist amenities based on the idea that the majority of Airbnb guests are tourists. Ideally there would be a direct variable for accessibility to tourist amenities such as restaurants, entertainment venues, local landmarks and museums. As this information is not widely available, accessibility to jobs is a proxy as most typical tourist venues and attractions are employers and contribute to jobs. Further, Ashworth and Page (2011) note the fact that tourist landscapes coexist with employment, dispersed and grouped into zones such as the 'entertainment district'. This 'employment-based attraction' is especially true as tourism in major urban centres tends to be based around attractions such as cultural venues, including museums and theatres, as well as shopping streets and restaurant and bar districts (Bull & Church, 2001). In NYC, Midtown Manhattan is one of the preeminent centres of employment, and also the location of many tourist amenities. Therefore, high access to jobs in NYC reflects a high degree of access to Midtown Manhattan and other tourist amenities, but also reflects other parts of the city that would be attractive to tourists. Transit accessibility is used as tourists are less likely to have access to the use of a personal vehicle. Cervero and Kockelman (1997) argue that land-use diversity is likely representative of other attributes including density and walkability. Studies have shown that higher degrees of walkability (albeit based on 'walk score' data) are related to higher property values (Pivo & Fisher, 2011; Rauterkus & Miller, 2011).

All the distances involved were calculated as network distances, based on the nearest census block group centroid. The distance from the closest block group centroid was used instead of the distances from the listing *x*,*y*-coordinates to reduce computational complexity. The relevant shapefiles used in the calculations were provided by the City of New York's Planning Department. Data regarding accessibility to jobs were gathered from the Access Across America project. This data set measures the number of jobs accessible within 30 min per block and was aggregated to the census block group to be on the same scale as the other location and neighbourhood variables. Owing to the overwhelming number of jobs in Midtown Manhattan, this measure acts as an approximate proxy of the centrality of any individual block group. Land-use diversity – a proxy for neighbourhood walkability – was measured through an entropy index based on the

work of Cervero and Kockelman (1997), as follows:

$$\text{Land use diversity} = \sum_j \alpha_j \cdot p_j \ln(p_j)$$

where j depicts a land-use class; p_j is the percentage of land-use j within 800 m of network distance (10 min of walking) from the centroid of the census block group nearest to the Airbnb listing; and α_j is −1 if the land use in question is negatively associated with walking (industrial uses, vacant lots, parking lots, airports, gas stations), and 1 otherwise (e.g., residential, commercial, park uses). Data were used from the NYC Department of Planning's MapPLUTO™ database. The data set gave the use of every building. These uses were then simplified and aggregated from 25 to 21 categories by collapsing the different residential uses into one use. Median household income, population density and the percentage of the population that is non-Hispanic White were provided by the 2015 American Community Survey (five-year estimates) at the census tract level.

To account properly for the variation within and among neighbourhoods, a multilevel regression model with three levels was used. The levels were at the listing itself, the census tract in which the listing was located, and the borough in which the listing was located. This model differs from previous hedonic analyses of Airbnb listings which rely on ordinary least squares (OLS) regression using the listing's reported location (Gibbs et al., 2018; Li et al., 2015; Teubner et al., 2017; Wang & Nicolau, 2017). The core difference is a multilevel model 'explicitly models spatial autocorrelation and heterogeneity' (Orford, 2000, p. 1649). Housing in spatial proximity is more likely to share similar characteristics (Orford, 2000). For example, one factor in home prices is that real estate agents set prices based on nearby housing. Therefore, unlike models relying on OLS, a multilevel model allows for spatial autocorrelation to be taken into account (Orford, 2000). Additionally, this multilevel model accounted for micro-neighbourhood effects that could otherwise not be captured by ordinary variables, such as the desirability of certain neighbourhoods. The result is a more sophisticated hedonic model than previous OLS-based studies.

RESULTS

Table 2 shows the results of the multilevel hedonic regression models divided between the two dependent variables: average price per night and average revenue per month. Both variables are logarithmically transformed.

The tourism and accommodation industries tend to be highly seasonal. Unsurprisingly, therefore, seasonality has a significant impact on the price per night and revenue of Airbnb listings in NYC. The results in Table 2 are compared with December and demonstrate significant increases in both price and revenue in the summertime, along with some more fine-grained variation which suggests that hosts adapt to seasonal patterns of visitor demand. For example, prices in December are higher than prices in April, but listings generate less revenue. A plausible interpretation of this pattern is that high holiday-time demand at the end of December is only partially met with new supply from residents leaving for their own holidays. The equilibrium price increases because demand outstrips supply, but the part-time hosts introducing new supply only list their homes for a short period of the month, and hence drag down average revenue.

Entire-home listings comprise 61.1% of the population in the model, and they earn a decisive premium in both price and revenue. Entire-home listings command a 51.2% higher price than private-room listings (a room in a private home or apartment, usually sharing bathrooms and kitchen facilities with the host) and generate 51.2% more revenue. Shared rooms (sharing a room with a host or others, such as a hostel dorm) are priced 36.7% lower than private rooms and generate 42.4% less monthly revenue.

Extra bathrooms command a larger price and revenue premium than extra bedrooms. One additional bedroom is associated with a price increase of 13.2% and a 11.0% increase in monthly revenue. By contrast, each additional bathroom is associated with a 14.2% increase in listing price per night and 14.8% higher revenue per month. The higher performance associated with additional bathrooms could reflect the extent to which they indicate more luxurious listings or serve as a proxy for a larger home at a given number of bedrooms. An extra place for a guest predicts a 6.4% higher listing price and a 7.7% increase in monthly revenue.

Table 2. Model results.

	Price				Revenue			
	Coefficient	*p*	Confidence interval[a]		Coefficient	*p*	Confidence interval[a]	
Time variables								
January	−0.1063	***	−0.1085	−0.1040	−0.5841	***	−0.5958	−0.5725
February	−0.1406	***	−0.1431	−0.1380	−0.3920	***	−0.4049	−0.3790
March	−0.0874	***	−0.0898	−0.0850	−0.0919	***	−0.1041	−0.0797
April	−0.0296	***	−0.0319	−0.0273	0.0879	***	0.0762	0.0997
May	0.0019		−0.0003	0.0042	0.2917	***	0.2803	0.3031
June	0.0402	***	0.0379	0.0424	0.2706	***	0.2591	0.2822
July	0.0215	***	0.0193	0.0238	0.1971	***	0.1856	0.2086
August	0.0121	***	0.0099	0.0142	0.2288	***	0.2180	0.2396
September	0.0413	***	0.0392	0.0434	0.3691	***	0.3585	0.3797
October	0.0346	***	0.0322	0.0369	0.3041	***	0.2923	0.3159
November	−0.0069	***	−0.0093	−0.0046	−0.0621	***	−0.0742	−0.0501
Structural variables								
Entire apartment	0.5115	***	0.5052	0.5178	0.5121	***	0.4985	0.5256
Shared room	−0.3670	***	−0.3815	−0.3525	−0.4235	***	−0.4549	−0.3922
Bedrooms	0.1321	***	0.1270	0.1372	0.1097	***	0.0988	0.1206
Bathrooms	0.1421	***	0.1340	0.1502	0.1481	***	0.1307	0.1656
Guests	0.0637	***	0.0615	0.0659	0.0768	***	0.0720	0.0815
Host variables								
Superhost	0.0565	***	0.0460	0.0670	0.1507	***	0.1289	0.1726
Multihost − 2–10	−0.0037		−0.0092	0.0018	0.0660	***	0.0542	0.0778
Multihost − > 10	−0.0924	***	−0.1135	−0.0713	0.0887	***	0.0446	0.1328
Number of photos	0.0036	***	0.0033	0.0038	0.0019	***	0.0013	0.0024
Instant booking	−0.0177	***	−0.0256	−0.0099	0.1424	***	0.1256	0.1591
Reviews	−0.0005	***	−0.0006	−0.0004	0.0084	***	0.0082	0.0086
Overall rating	0.0826	***	0.0771	0.0880	0.0433	***	0.0311	0.0555
Location and neighbourhood variables								
Subway	−0.0201	***	−0.0338	−0.0063	−0.0350	***	−0.0593	−0.0107
Transit accessibility	0.1220	***	0.1090	0.1350	0.1320	***	0.1120	0.1520
Dist Times Square	−0.0147	***	−0.0168	−0.0126	−0.0245	***	−0.0277	−0.0213
Diversity	0.0093		−0.0055	0.0240	0.0065		−0.0211	0.0341
Population density	−0.0011	***	−0.0016	−0.0006	−0.0011	***	−0.0017	−0.0005
Median household income	0.0012	***	0.0009	0.0015	0.0009	***	0.0006	0.0013
Percentage white	0.0019	***	0.0015	0.0022	0.0010	***	0.0005	0.0014
Listing density	0.0001	***	0.0001	0.0002	0.0001	***	0.0001	0.0002
Constant	3.3377	***	−0.0338	−0.0063	5.6364	***	5.5323	5.7404

Notes: Dependent variables: ln(price per night) and ln(monthly revenue).
[a]95% confidence interval.
*95% significance level; **99% significance level; ***99.9% significance level.

Table 2. Continued.

Random-effects parameters	Price			Revenue		
	Estimate	Confidence interval[a]		Estimate	Confidence interval[a]	
Borough level						
sd(Constant)	0.0632	0.0332	0.1205	0.0695	0.0334	0.1450
Census tract level						
sd(Constant)	0.0891	0.0826	0.0960	0.0775	0.0693	0.0868
Listing level						
sd(Constant)	0.2641	0.2624	0.2659	0.4958	0.4913	0.5002
sd(Residual)	0.1436	0.1433	0.1440	0.7373	0.7355	0.7390

Note: [a]95% confidence interval.

A set of characteristics that collectively suggest a listing is a commercial operation, as opposed to a part-time rental, are associated with substantially higher monthly revenue than the baseline, but not consistently higher nightly rates. Commercial Airbnb operators, in other words, appear to set pricing more aggressively in order to maximize occupancy and, hence, revenue. The first of these characteristics is 'superhost' status, which Airbnb confers on hosts based on a high frequency of hosting, positive reviews from guests and a low cancellation rate. According to the model, this superhost status allows hosts to charge a slight premium (5.6% more per night) and results in a substantial increase of 15.1% in monthly revenue. The differential between these two figures implies that superhosts conduct considerably more bookings per month than other hosts. The significant impact of superhost status on both price and revenue supports a similar finding by Wang and Nicolau (2017), but contradicts Teubner et al. (2017) who did not identify superhost status as having statistical significance.

Airbnb allows hosts to activate an instant booking feature, which allows prospective guests to bypass the usual host approval process and have their requested booking confirmed immediately. Listings with the instant booking feature activated have 1.8% lower average nightly prices, but have much larger monthly average revenue of 14.2%. Much like superhost status, instant booking is characteristic of high-volume listings, which provide a more hotel-like experience than less commercial listings whose hosts prefer to approve booking requests manually.

The final factor associated with commercial short-term rental operations that predicts higher revenue is the number of listings a given host controls. Hosts with between two and 10 listings do not charge a significantly different price per night from hosts with a single listing, but they earn 6.6% more in monthly revenue, reflecting higher occupancy rates. By contrast, hosts managing more than 10 properties charge a substantially lower price per night than single-listing hosts (9.2% less), but earn substantially *more* revenue (8.9%). Noting the same result, Li et al. (2015) suggest that this pattern could reflect learning effects among high-volume hosts, especially responses to change in demand (either seasonal or event based).

Listings with more photographs charge marginally higher prices and earn marginally more revenue. One additional review is associated with a 0.05% decrease in price, but a 0.8% increase in revenue. This additional revenue is likely due to the fact that lower priced listings are booked more frequently and therefore they will have more reviews (Gibbs et al., 2018). One additional star in average rating is associated with a price increase of 8.3% and a revenue increase of 4.3%.

Proximity to mass transit is a strong predictor of listing price and revenue, but only when it facilitates access to the central city. In fact, controlling for accessibility, listings in proximity to subway stations charge less per night (2.0% less) and have less monthly revenue (3.5% less). However, listings with high accessibility (defined here as access to jobs via transit) charge an average of 12.2% more per night and earn an average of 13.2% more monthly revenue. This finding

suggests that Airbnb guests value transit for what destinations (represented here by the proxy of number of jobs) it can take them to, not for proximity to transit itself. In the case of Airbnb listings, the subway is valuable if it takes one quickly to Midtown Manhattan.

Beyond accessibility, location per se appears to matter. Listings charge higher prices per night (1.5%) and earn more monthly average revenue (2.4%) for every 1 km closer to Times Square they are located – an effect that is independent of the measures of accessibility discussed above. Moreover, the more listings there are in the same census tract, the higher the price and revenue the listing generates. (Each additional listing in the census tract increases price per night and revenue per month by 0.01%.) This suggests the presence of agglomeration effects in the Airbnb marketplace: areas that are currently attractive to tourists are more promising locations for future listings growth. This finding further suggests that areas with high listing density have generally not reached their saturation point, and instead could continue to profitably absorb new listings.

Divergent results were found for other location characteristics. Land-use diversity does not command higher average prices, nor does it generate significantly greater amounts of revenue per month. However, and related to this variable, listings in areas with higher density, all else being equal, have lower prices per night and lower revenue per month (both decline by 0.1% with a 1000 person/km^2 increase in density). Combined with the findings on transit accessibility, this suggests that guests might not value actually staying in mixed-use, walkable neighbourhoods, so long as they can access such neighbourhoods conveniently with transit.

At the same time, guests appear to value affluent, non-Hispanic White neighbourhoods. Listings in higher income areas charge more (0.1% more per US$1000 median household income) and earn higher monthly revenues (0.1% per US$1000 median household income) This could reflect the fact that in most cases high income serves as a proxy for urban amenities and high-quality urban design. A 1% increase in the percentage of the non-Hispanic White population leads to an increase in average nightly price of 0.2% and increase in monthly revenue of 0.1%.

Ultimately the model explained 3.9% of all variation in prices due to borough variability, 11.7% due to census tract variability and 79.8% due to variability between listings. This establishes the important role the neighbourhood plays in influencing the average price per night of an Airbnb listing. In the revenue model, 0.6% of variation was explained due to borough variability, 1.4% due to variability between census tracts and 32.1% due to variation between individual listings.

DISCUSSION AND CONCLUSIONS

Our model demonstrates that locational factors and neighbourhood variation have a large impact on both the average price per night of Airbnb listings and their average monthly revenue. This is the first study to demonstrate the importance of neighbourhood variation, and not just straight-line distance to city centres or town halls, for Airbnb listing performance. Of these factors, transit accessibility to jobs has the most impact on both average nightly price and average monthly revenue. The premium paid by short-term guests for accessible homes, and the willingness of landlords and hosts to reap the financial benefits by catering to them, suggests that long-term residents of accessible neighbourhoods will face increasing displacement pressure and increasing risk of being relegated to less accessible neighbourhoods.

Moreover, the results of this model shed a considerable light on the practices of commercial Airbnb operators. The model suggests that the more professionally Airbnb listings are managed, the higher average monthly revenue they stand to make. Attributes that reflect host experience and savvy, such as superhost status and the number of reviews per listing, predict increased monthly revenue. Likewise, listings with instant booking earn significantly more revenue on average. Renting entire homes instead of private rooms unsurprisingly earns considerably more

revenue. Finally, hosts with multiple listings earn more per listing, an effect that increases when a host has 10 or more listings. Taken together, these facts suggest that hosts who treat their listings as de facto hotels rather than opportunities for part-time 'home sharing' are considerably more successful in the Airbnb marketplace. These entire home listings frequently booked are the most likely category to contribute to the removal of housing from the long-term rental market (Lee, 2016; Wachsmuth et al., 2017, 2018; Wachsmuth & Weisler, 2018). As such policy-makers should be concerned about the comparable success of professional hosts on the Airbnb platform, especially as the platform itself tries to position itself as helping average New Yorkers.

A related finding that emerges from the model is a cluster of independent variables that simultaneously predict lower nightly prices but higher monthly revenue. These are hosts with 10 or more separate listings, the presence of the instant booking feature and the number of reviews a listing has received. Together, these variables appear to indicate high-volume commercial operators who use low prices to achieve higher occupancy rates, and thus higher overall revenues. Since other research has suggested that the Airbnb market is increasingly dominated by a relatively small number of commercial operators (Wachsmuth et al., 2017), exploring in more detail the behaviour of this subset of hosts is an important task for future research.

As noted above, one potentially problematic implication of this study is that many of the same locational attributes are valued for Airbnb tourists and for long-term residents. In particular, Airbnb guests are willing to pay a very large premium for transit accessibility to jobs (not, presumably, because of their need to access employment opportunities per se, but rather because these opportunities are co-located with tourism amenities). This suggests that Airbnb hosts can earn a large revenue premium by converting long-term housing in high-accessibility residential neighbourhoods to de facto Airbnb hotels. The result would be that long-term tenants are forced into less accessible locations in the city. Previous studies have linked short-term rentals to gentrification and displacement (Mermet, 2017; Wachsmuth & Weisler, 2018); this study suggests long-term tenants may also face transportation disconnection. Accordingly, Airbnb may be causing long-term residents to have longer commutes, incur higher transportation costs and produce larger carbon footprints.

The present study used a multilevel random intercept modelling technique to investigate the impact of multiple locational factors on both the average price and average monthly revenue of Airbnb listings in NYC. Unlike other hedonic analyses of Airbnb that have used OLS techniques, the multilevel allowed for spatial autocorrelation to be taken into account, which allowed us to control for cultural neighbourhood aspects (at the census tract and borough levels) such as the 'cachet' or desirability of specific districts. These neighbourhood-level cultural aspects may be important for future research, as there is increasing evidence of tourists seeking 'authentic' experiences by staying outside of traditional tourist neighbourhoods (Füller & Michel, 2014; Gant, 2016). In a recent study, Wachsmuth and Weisler (2018) found the neighbourhoods most likely to be impacted by short-term rental-related gentrification are those that would be attractive to tourists, likely reflecting this cultural desirability. These cultural dimensions could be valuable for future analysis in order to determine if hosts and commercial operators are able to capitalize on these factors. The significant variation based on census tract variability (11.7%) highlights the likely uneven impact of Airbnb on New York's neighbourhoods.

The number and complexity of both independent and dependent variables of this study greatly expanded on previous works; these included multiple locational factors, as well as an estimate of listings' average monthly revenue. Future studies could investigate the link between variables suggesting increasing professionalization and increases in monthly profit. One such variable might be Airbnb's new 'Business ready' feature, which requires hosts to provide a set of amenities including laptop-ready workspaces and self-check-in. Further research in the New York area would ideally include adjacent areas of New Jersey where significant Airbnb activity occurs but current locational data were limited. While additional variables and data might be included in

future research, our model highlights the clear connection between locational variables and both average price per night and average monthly revenue.

ACKNOWLEDGEMENTS

FUNDING

ORCID

Danielle Jane Kerrigan ⓘ http://orcid.org/0000-0002-5877-1698
David Wachsmuth ⓘ http://orcid.org/0000-0001-5689-9527
Ahmed El-Geneidy ⓘ http://orcid.org/0000-0002-0942-4016

REFERENCES

Airbnb. (2018). Airbnb unveils roadmap to bring magical travel to everyone [Press release]. Retrieved from https://press.atairbnb.com/airbnb-unveils-roadmap-to-bring-magical-travel-to-everyone/

Armstrong, R. J., & Rodriguez, D. A. (2006). An evaluation of the accessibility benefits of commuter rail in eastern Massachusetts using spatial hedonic price functions. *Transportation*, *33*(1), 21–43.

Ashworth, G., & Page, S. J. (2011). Urban tourism research: Recent progress and current paradoxes. *Tourism Management*, *32*(1), 1–15.

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, *58*(304), 933–942.

Barron, K., Kung, E., & Proserpio, D. (2017). The sharing economy and housing affordability: Evidence from Airbnb.

Benjamin, J. D., & Sirmans, G. S. (1996). Mass transportation, apartment rent and property values. *Journal of Real Estate Research*, *12*(1), 1–8.

Bull, A. O. (1994). Pricing a motel's location. *International Journal of Contemporary Hospitality Management*, *6*(6), 10–15.

Bull, P., & Church, A. (2001). Understanding urban tourism: London in the early 1990s. *International Journal of Tourism Research*, *3*(2), 141–150.

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219.

Chatman, D. G., Tulach, N. K., & Kim, K. (2012). Evaluating the economic impacts of light rail by measuring home appreciation: A first look at New Jersey's River line. *Urban Studies*, *49*(3), 467–487.

Chen, C.-F., & Rothschild, R. (2010). An application of hedonic pricing analysis to the case of hotel rooms in Taipei. *Tourism Economics*, *16*(3), 685–694.

El-Geneidy, A., Lierop, D. V., & Wasfi, R. (2016). Do people value bicycle sharing? A multilevel longitudinal analysis capturing the impact of bicycle sharing on residential sales in Montreal, Canada. *Transport Policy*, *51*, 174–181.

Franklin, J. P., & Waddell, P. (2003). *A hedonic regression of home prices in King County, Washington, using activity-specific accessibility measures*. Paper presented at the Proceedings of the transportation research board 82nd annual meeting, Washington, DC.

Füller, H., & Michel, B. (2014). Stop being a tourist! New dynamics of urban tourism in Berlin–Kreuzberg. *International Journal of Urban and Regional Research*, *38*(4), 1304–1318.

Gant, A. C. (2016). Holiday rentals: The new gentrification battlefront. *Sociological Research Online*, *21*(3), 1–9.

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., & Goodwill, A. (2018). Pricing in the sharing economy: A hedonic pricing model applied to Airbnb listings. *Journal of Travel & Tourism Marketing*, *35*(1), 46–56.

Glaesener, M.-L., & Caruso, G. (2015). Neighborhood green and services diversity effects on land prices: Evidence from a multilevel hedonic analysis in Luxembourg. *Landscape and Urban Planning*, *143*, 100–111.

Gottlieb, C. (2013). Residential short-term rentals: Should local governments regulate the 'industry'? *Planning & Environmental Law*, *65*(2), 4–9.

Guttentag, D. (2015). Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, *18*(12), 1192–1217.

Guttentag, D. A., & Smith, S. L. (2017). Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. *International Journal of Hospitality Management*, *64*, 1–10.

Herrin, W. E., & Carvell, S. A. (1990). Pricing in the hospitality industry: An implicit markets approach. *Hospitality Review*, *8*(2), 27.

Hill, D. (2015). How much is your spare room worth? *IEEE Spectrum*, *52*(9), 32–58.

Horn, K., & Merante, M. (2017). Is home sharing driving up rents? Evidence from Airbnb in Boston. *Journal of Housing Economics*, *38*, 14–24.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157.

Lee, D. (2016). How Airbnb short-term rentals exacerbate Los Angeles's affordable housing crisis: Analysis and policy recommendations. *Harvard Law and Policy Review*, *10*, 229–254.

Lee, D., Hyun, W., Ryu, J., Lee, W. J., Rhee, W., & Suh, B. (2015). *An analysis of social features associated with room sales of Airbnb*. Paper presented at the ACM Conference Companion on Computer Supported Cooperative Work & Social Computing, Vancouver, British Columbia.

Lee, S. K., & Jang, S. (2011). Room rates of US airport hotels: Examining the dual effects of proximities. *Journal of Travel Research*, *50*(2), 186–197.

Lewis-Workman, S., & Brod, D. (1997). Measuring the neighborhood benefits of rail transit accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, *1576*, 147–153.

Li, M. M., & Brown, H. J. (1980). Micro-neighborhood externalities and hedonic housing prices. *Land Economics*, *56*(2), 125–141.

Li, J., Moreno, A., & Zhang, D. J. (2015). Agent behavior in the sharing economy: Evidence from Airbnb. *Ross School of Business Working Paper Series, 1298*, 2015.

Mermet, A.-C. (2017). Airbnb and tourism gentrification: Critical insights from the exploratory analysis of the 'Airbnb syndrome' in Reykjavik. In M. Gravari-Barbas & S. Guinnad (Eds.), *Tourism and gentrification in contemporary metropolises* (pp. 52–74). Reykjavik: Routledge.

Orford, S. (2000). Modelling spatial structures in local housing market dynamics: A multilevel perspective. *Urban Studies*, *37*(9), 1643–1671.

Oskam, J., & Boswijk, A. (2016). Airbnb: The future of networked hospitality businesses. *Journal of Tourism Futures*, *2*(1), 22–42.

Pivo, G., & Fisher, J. D. (2011). The walkability premium in commercial real estate investments. *Real Estate Economics*, *39*(2), 185–219.

Rauterkus, S. Y., & Miller, N. (2011). Residential land values and walkability. *Journal of Sustainable Real Estate*, *3*(1), 23–43.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55.

Teubner, T., Hawlitschek, F., & Dann, D. (2017). Price determinants on Airbnb: How reputation pays off in the sharing economy. *Journal of Self-Governance & Management Economics*, *5*(4), 53–80.

Wachsmuth, D., Chaney, D., Kerrigan, D., Shillolo, A., & Basalaev-Binder, R. (2018). The high cost of short-term rentals in New York City. *A report from the Urban Politics and Governance research group, School of Urban Planning, McGill University.* Retrieved from http://www.sharebetter.org/wp-content/uploads/2018/01/High-Cost-Short-Term-Rentals.pdf,2,2018.

Wachsmuth, D., Kerrigan, D., Chaney, D., & Shillolo, A. (2017). Short-term cities: Airbnb's impact on Canadian housing markets. *Policy report. Urban Politics and Governance research group, School of Urban Planning, McGill University.*

Wachsmuth, D., & Weisler, A. (2018). Airbnb and the rent gap: Gentrification through the sharing economy. *Environment and Planning A: Economy and Space.* doi:10.1177/0308518x18778038

Wang, D., & Nicolau, J. L. (2017). Price determinants of sharing economy based accommodation rental: A study of listings from 33 cities on Airbnb.com. *International Journal of Hospitality Management*, *62*, 120–131.

Zervas, G., Proserpio, D., & Byers, J. W. (2017). The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of Marketing Research*, *54*(5), 687–705.

Zhang, Z., Ye, Q., & Law, R. (2011). Determinants of hotel room price: An exploration of travelers' hierarchy of accommodation needs. *International Journal of Contemporary Hospitality Management*, *23*(7), 972–981.