Measuring Discrimination: The Impact of Host Race and Gender on Earnings from Airbnb*

Anya Marchenko

May 17, 2018

Abstract

Racial bias faced by Airbnb hosts is a topic largely underrepresented in economic literature. Earlier research has attempted to measure discrimination against New York City hosts, but was limited by a small sample size and narrow set of controls. I address this issue by using a previously unexploited dataset from a large webscrape of the Airbnb website to measure discrimination. Controlling for a large set of covariates, I estimate that black hosts earn \$5 - 7 less per night, and Asian hosts \$6 - 9 less per night (depending on the sex of the host) than white hosts who post a similar type of listing. I then explore various hypotheses for this effect. There is little evidence that these price disparities are due to minority hosts choosing to price their listings lower because of differences in their marginal cost, or offering their listing up for rent for shorter periods of time, than white hosts. I also find little evidence that this effect is due to minority hosts owning listings of worse quality. Overall, among the hypotheses I test, discrimination is the most convincing explanation for this persistent price disparity.

1 Introduction

African-Americans have experienced pervasively worse outcomes in the housing market, and historic and current racial discrimination is one major cause.? Even after the gains of the Civil Rights Era,

*I thank Steven D. Levitt for valuable advice. Kirthi Bellamkonda, Jacob Dorn, Michal Dzitko, Sonia Jaffe, Ezra Karger, Sylvia Klosin, Victor Lima, and Kotaro Yoshida provided excellent feedback and comments. Mahathi Ayyagari, Fong Chai, Joseph Day, Melody Jih, Spencer Kang, Cristian Raygoza, Tony Song, Lydia Sum, Putter Thepkanjana, Leah Umanskiy and provided research assistance. All remaining errors are my own. Data was provided by InsideAirbnb.com, founded and maintained by Murray Cox. Airbnb was not consulted in this research. Sentiment analysis would not be possible without Sentimentr and coding by Ben Chaimberg.

such as the landmark Fair Housing Act of 1968, discrimination in the housing market is widely documented by social scientists. Yinger (1986) measured discrimination using fair housing audits and found that black renters are told that there are 30% fewer available housing units than white renters.**??** African-American families face higher barriers when raising capital to purchase a home, since black applicants get fewer loans for the same application as white applicants.**?** [EXAMPLE 3]. Landlords renting out apartments discriminate both because of their own prejudice and in response to the prejudice of their white renters.[1] **????**

Economists have primarily studied discrimination against African-American tenants. There is little research on the other side of the market - when African-Americans are supplying, rather than demanding, housing.[2] Logistically, it is difficult for researchers to conduct an experiment to randomize ownership of property to landlords of different races. Without randomization, it is difficult to disentangle true discrimination from systematic differences in the housing owned by African-Americans and white landlords. **INSERT 'BAD' LANDLORD STUDIES HERE**. Some studies have found anecdotal evidence of discrimination against African-American landlords.

This paper leverages the recent rise of sharing economies and the data and standardization their platforms provide to credibly measure discrimination against minority landlords. Airbnb is a sharing economy platform that allows people to rent out their apartment, house, or single room to short-term lodgers. Since Airbnb simply provides an online platform for a market that already exists, it is reasonable to assume that agents will discriminate on Airbnb in a similar way that they discriminate in the real world. This means that studying discrimination in sharing economies could be an important way to learn about discrimination in traditional housing markets as well.

Measurements of discrimination on Airbnb are still potentially confounded by many other factors that affect a listing's price. Edelman and Luca (2014) estimated the effect of host race on the price of their listing in New York City with a small set of controls. Their results suggest that non-black hosts on Airbnb have prices roughly 12% higher than black hosts. However, they control for only a few property characteristics, the quality of the host's reviews, and a measure of the reliability of

[1] Some studies have found evidence that discrimination in rental markets is statistical in nature (that is, landlords use race as a proxy for income). For African-Americans who imply that they are of a higher social class when applying for an apartment, discrimination is virtually not present.**???**

[2] One would also expect black landlords to fare worse than white landlords in this area as well. Properties owned by African-Americans tend to be less expensive than those owned by white Americans. The average black household still has less mean wealth than a white household.**?** Even middle-class black and Hispanic households still live in neighborhoods with median incomes similar to those of very poor white neighborhoods.**?**

the host. They leave out many other unobservables that could be confounding the effect such as the location of the listing (an important factor, especially in New York), the type of listing (which can vary from a single room to an entire house), and proxies for the quality of the host themselves.

In this paper, I use a new, previously unexploited dataset from a webscrape of Airbnb to empirically measure discrimination. Using price and listing information for 70,000 Airbnb hosts throughout the country, I address the limitations of Edelman and Luca's research by controlling for location, additional property characteristics, comprehensive measures of listing size, and text analyses of host-written descriptions of the listing.

Crucially, the data allows for estimation of not just the price, but the vacancy rate of a listing, which is a measure of the quantity demanded. Knowing both quantity and price allows me to distinguish whether the price disparity between minority and white hosts on Airbnb is due to a demand shift (consistent with discrimination) or a supply shift (consistent with a difference in marginal cost for the hosts, or other hypothesis, discussed below).

I find that non-white hosts, both male and female, have lower prices than white hosts. The biggest effect is for Asian female hosts, whose prices are roughly \$9 less per day than white male hosts who own the same type of listing. The second biggest effect is for black males, with a coefficient of \$7, followed by black women and Asian men with coefficients of \$6 per day, and Hispanic females with a coefficient of \$5.[3] For Hispanic men the effect is small, around \$2, and is not statistically significant.

My coefficients are smaller than Edelman and Luca's, raising the concern that the price disparity measured could be erased altogether by adding more controls. I attempt to address this problem by adding measures of host quality and listing attractiveness that could differentiate one listing from another for a given property size and type. I find that the price disparity is stable to the addition of these controls, providing confidence that adding more variables of a similar nature would not eliminate the disparity.

After having addressed the empirical robustness of this main finding, I test several alternative hypotheses that could explain the price disparity. In contrast to previous research, my data on the vacancy rate allows me to distinguish between discrimination and several other mechanisms which

[3] This effect is statistically significant at the $p < .001$ level for black hosts and Asian women, the $p < .01$ level for Asian male hosts, and $p < .05$ level for Hispanic women.

could be causing the price disparity.

One alternative hypothesis for the price disparity is that minority hosts charge a lower price for their listing because it is cheaper for them, on the margin, to operate the same listing as compared to a white host. Since Black and Hispanic workers tend to earn less than their white counterparts, even for the same amount of education, they may have a lower opportunity cost of time.**????** Minority hosts would therefore have a lower marginal cost of managing their listing, and so would choose to set lower prices for their listings than white hosts with comparable listings. If the price disparity was due to supply differences, then basic supply and demand theory predicts that the quantity demanded of minority hosts' listings should be higher than white hosts'. To test this hypothesis, I consider the number of reviews that a host has as a proxy for the quantity demanded of that listing. I regress the number of reviews on host race, controlling for the preferred specification. I find that minority hosts have a lower, not higher, number of reviews than white hosts, suggesting that this hypothesis fails to explain the price disparity.[4]

One potential explanation for their lower number of reviews is that minority hosts make their listing available to guests less frequently than white hosts. A host controls how many days of the month they offer their listing for rent via an availability calendar on the listing page. When a guest books their listing, the booked days disappear from the availability calendar. Therefore, the measure of availability is actually a measure of true vacancy for the listing. If minority hosts have lower numbers of reviews, perhaps this is because they offer their listing for fewer days of the month than white hosts. To test this, I regress a listing's availability out of 30 days on host race, controlling for the preferred specification. Contrary to this hypothesis, regression results show that the listings of black hosts stay vacant 1 - 2 days per month longer than the listings of white hosts. However, there is evidence that Asian hosts do choose to make their listings available less frequently, which would contribute to their lower number of reviews.

A final, alternative, hypothesis that could account for the price disparity between minority and white hosts is that minority hosts own listings of worse quality, or are simply worse hosts. I consider the quality of a host's reviews as a proxy for the quality of the listing and host. I use the race and

[4]Throughout this analysis, I assume that guests review hosts of different race at equal rates. X and Y have found that rates of review tend to be the same/different across hosts. MELODY - FIND WHETHER WHETHER THERE'S LIT ON DIFFERENT REVIEW RATES ACROSS RACES. However, if reviewers systematically under-review minority hosts, this itself could be evidence of discrimination. For more discussion on this assumption, see Section 5, Part 1.

gender of the reviewer and the host to compare the sentiment (how favorable or unfavorable the review is) of the reviews that guests leave for white and for minority hosts.[5] Rather than observing that all minority hosts uniformly had lower quality reviews, which the hypothesis would predict, the significance of the result was either negligible, or depended on the demographics of the reviewer and host. While there is some evidence that male reviewers tend to rate male hosts higher, there is little within-race preference between reviewers and hosts. Taken as a whole, sentiment analysis suggests that minority hosts do not have lower quality reviews.

1.1 About Airbnb

Airbnb is an online marketplace founded in 2008 that allows hosts to rent their private dwellings to guests as temporary accommodation. As of 2017, it has more than 3 million listings, more than Marriott's 1.2 million rooms worldwide.**?** Just like traditional hotel chains, guests on Airbnb can browse listings by city and property type, and book a stay based on prices, location, past reviews, pictures of the listing, size, and amenities. Unlike traditional hotel chains, however, hosts create a profile for themselves and a page for each listing they are renting. Each listing page includes the name and picture of the host, the reviews left by previous guests, and those guests' profile pictures. Guests can therefore infer demographic information about the host through a host's picture and name, creating the opportunity for discrimination. Figures 1-5 present screenshots of a listing in a Chicago neighborhood, illustrating some of the information that would be available to a potential guest.

1.2 Previous Literature

Most relevant to this paper is Edelman and Luca (2014), the first paper to identify and measure the extent of anti-host discrimination on Airbnb.**?** They explore the effect of host race on the price of their listing using a snapshot of roughly 3,800 New York City hosts in 2012. Controlling for several confounders that influence price, their findings indicate that non-black hosts on Airbnb have prices roughly 12% higher than black hosts. I build on Edelman and Luca's research in several important ways. First, their sample was relatively narrow and confined to a single city. My sample includes

[5]Since it required hand-coding, demographic information of the reviewers is only available for a randomly-chosen subset of hosts in Chicago.

seven cities throughout America, which are all large urban centers, picked so that they cover most geographic regions in the US. It is important to have this variety because discrimination in a large, cosmopolitan city with a highly diverse population such as New York might look different from discrimination in Nashville, which is more racially homogenous.[6] Second, their set of controls was limited by the relatively sparse listing information available on the Airbnb website (Airbnb has added more comprehensive data on each listing page since 2013). Their covariates only included a listing's location, the number of people the listing accommodates, the rating, the number of bedrooms, and whether or not the whole apartment is available to the guest. After confirming their result when I run a regression using their controls, I then control for a more complete set of covariates (see Section 2.1 and Section 3.1).[7] Most importantly, I also test alternative hypotheses for these price disparities, which Edelman and Luca are unable to address.

Becker (1957) proposed the idea that discrimination against a group can be reflected in that group's market prices.**?** In the Airbnb market, Becker's market discrimination would be reflected in the price that the guest (buyer) pays to the host (seller) to stay with them. If the guest (buyer) is discriminating, then given two comparable listings, they would choose not to stay in the one owned by a minority host (seller). Responding to the lower demand, hosts in minority groups rationally post a lower price in a competitive market.

Becker (1957) was concerned with discrimination arising from face-to-face interactions between minority and majority groups. Since then, there has been a large amount of research indicating that Becker's theory holds for people participating in online labor, lending, rental, and other markets as well. In these cases, participants simply bring their prejudices online and use names and photos to discriminate. A canonical example is the Bertrand and Mullainathan (2004) study, which found that resumes with white sounding names received 50% more callbacks from potential employers than identical resumes with African-American sounding names.**?** Doleac and Stein (2010) examined market outcomes when selling an iPod on various online marketplaces. In some pictures, a dark-skinned hand was holding the iPod, signaling a black seller, while in others, a light-skinned hand was holding the iPod, signaling a white seller.**?** Hands which indicates a black seller received fewer and lower offers than white sellers. In sharing economies, a similar pattern occurs. Uber riders with

[6] According to the U.S. Census Bureau's QuickFacts, Nashvile is 60.4% white, 28.4% black, 10.0% Hispanic, and 2.5% Asian. New York City is 44% white, 25.5% black, 12.7% Asian, and 28.6% Hispanic.

[7] See Table 4 for the results of my regression using their covariates. See Section 3.1 for a discussion of my controls.

distinctively African-American names experience longer wait times and more frequent cancellations than riders who use distinctively white names.? A later study by Edelman and Luca (2016) found a similar result: guests with distinctively African-American names receive 16% fewer responses from Airbnb hosts than those with white names.? These examples suggest that users of online platforms transfer their biases from the real world into the online world.

Frenken, 2017

2 Data

2.1 Source

My data are taken from the website Inside Airbnb, which provides cleaned and aggregated data on Airbnb listings in 43 cities across the world.? The data provided on the website are sourced from a webscrape of publicly available information on the Airbnb website. Inside Airbnb is not run by or affiliated with Airbnb itself.[8] The intent of Inside Airbnb is to inform the public on how Airbnb competes with the residential housing market in their areas.

The scrape of the Airbnb website was conducted throughout 2015, and provides a point-in-time snapshot of all of the listings available in a particular city. This includes all of the information that would be available to an Airbnb user browsing through listings at the time of the scrape.

A total of 70,000 host pictures across seven US cities were coded - Chicago, Los Angeles, New York City, Austin, Washington, D.C., and New Orleans.[9] Large cities with racial diversity which were geographically dispersed were chosen. This approach limits the applicability of my findings to urban areas, discounting the roughly fifth of Airbnb's listings which are located in rural areas.[10] In addition to main host data, demographic data was also coded for 16,000 reviewers who stayed with a subset of the Chicago hosts.[11] For those hosts in Chicago, it is thus possible to study the

[8]Airbnb's host profiles and listings are publicly available information, and no private data was accessed in the scrape. The cleaned data is under a Creative Commons Public Domain Dedication.

[9]For every city but New York, every single Airbnb listing that existed in that city at the time of the scrape was coded. In New York, which had the most listings in the sample, half of the existing 40,000 listings were randomly chosen to be coded. Time, effort, and monetary constraints prohibited the coding of all 16 US cities whose data was available on InsideAirbnb.com.

[10]A 2017 report released by Airbnb stated that 18.4% of all active listings are located in rural areas, and there was 138% year-in-year growth in Airbnb guest arrivals at rural listings.

[11]This represents about 23% of the total number of reviewers in Chicago. Not all reviewers could be coded due to time and manpower constraints. A random subset of Chicago hosts was chosen such that the 16,000 reviews represent all of the reviews left for those hosts. Each review has a unique reviewer id, host id, listing id, the date of the review,

interaction of reviewer demographics, host demographics, and review quality.

2.2 Data Summary

Summary statistics of host demographic information and their listings are displayed in Table 2. There is significant variation in both sex and race of the hosts on Airbnb. Roughly a third of the sample are single females, and a third are single males, with the rest being couples or groups.[12] About two-thirds of the hosts are white (X%), and less than a tenth are black (X%), Hispanic (X%), or Asian hosts (X%).[13] Black hosts in the sample are underrepresented relative to the proportion of African-Americans in the national population (13%). Hispanic hosts are similarly very underrepresented relative to the proportion of Hispanics in the population (16%). One explanation for this could be that people self-identify as Hispanic for census data, while Airbnb hosts were identified by RAs who might have mistakenly coded Hispanic hosts as other categories. Asian-American hosts (X%) are overrepresented by a few percentage points relative to the 5.6% of Asian-Americans in the national average.**?**

The prices of listings owned by white hosts are dramatically higher than those of other hosts. The mean price per night of a listing is $178 for white hosts, down to $125 per night for black hosts, $160 per night for Hispanic hosts, and $131 per night for Asian hosts. Minority hosts also have lower median prices and lower standard deviations, indicating that not only do minority hosts own cheaper listings on average, but their listings are more concentrated around the lower mean.[14]

It is reasonable to expect that a large portion of the price differences described above are driven by differences in property characteristics. Table 2's *Listing characteristics* FIX WITH NEW TABLE shows that white hosts dominate the most expensive option in every single category of observable property characteristics. White hosts own the most houses and the fewest apartments or lofts. They have the most bedrooms, bathrooms, beds, and amenities in their properties, and lease out more of their property than minority hosts do. In most of these measures of property

the review text, and the coded race, sex, and age of the reviewer.

[12] 38% of the sample are single females, 31% are single males, 23% are couples, and the other 8% are groups > 2 or pictures without a face. Couples and other groups were not included in the final analysis.

[13] 64% of the hosts in the sample are white, 7% are black, 5% are Hispanic, and 9% are Asian. The rest of the profile pictures were either pictures of groups, pictures without a human face, or multiracial couples, all of which were put in the "Unknown/Multiracial" category in Table 2.

[14] The median price of a listing owned by a white hosts is $115 per night, $90 for black hosts, $99 for Hispanic hosts, and $90 for Asian hosts.

quality, the listings owned by Hispanic hosts come the closest in quality to white hosts, often only a few percentage points behind. Either black or Asian hosts have properties of the lowest quality as measured by these metrics, depending on the category.

While white hosts' listings are of higher quality in terms of property characteristics, this is not the case for host characteristics. Black hosts do well in categories where the host can personally influence their desirability: responding on time, writing the long descriptions, or making their listing available for more days out of the month. They have the highest response rate at 77%, with white and Hispanic hosts behind them at 75.6%. They make their listings available an average of 14 days a month, a full 4 days more than white hosts. However, black hosts have the lowest acceptance rate, accepting only 36% of guests who ask to stay with them. Hispanic hosts have the highest acceptance rate at nearly 50%.

Black and Hispanic hosts also do well in some of my constructed measures of "host quality". They describe their listings using as many or more good words like "airy", "beautiful", and "clean", an average of .04 words higher than white hosts. While white hosts write the longest descriptions in every host-written field, black hosts are, on average, only four words behind white hosts in this metric. Asian hosts write the shortest descriptions in every host written field. The difference between white and Asian hosts increases as the fields get less prominent on the profile. At most the difference in the length of descriptions that white and Asian hosts write is 13 words. See Figure 2 for an example of host-written descriptions on a real listing profile.

White hosts also have the highest number of reviews, and the highest review ratings. Airbnb also designates especially experienced, highly-rated hosts as "Superhosts". Users on Airbnb are willing to pay more to stay with a "Superhost", since Superhosts are likely perceived as more trustworthy or as owning listings of higher quality. Since Airbnb assigns Superhost status based on the number of stays a host has, the quality of their reviews, and their response rate, it is not surprising that white hosts are Superhosts most frequently: 13.4% of white hosts are Superhosts, while the next runner-up, Hispanic hosts, are at 10.8% Superhosts. FIX THIS WORDING?

The reviewers who stayed with the Chicago hosts have similar gender diversity as the overall host population but significantly less racial diversity. A third of the reviewers are female and a similar proportion is male. However, 67% of reviewers are white, with only 6% being African American and Hispanic (about 500 reviewers each), and 12% Asian. Importantly, the measure of review quality externally assigned by Sentimentr to the text of each review generally matches up with the numeric scores reviewers gave. While all hosts have on average very positive reviews, white hosts have the most positive review sentiment, followed by Hispanic, Asian, and black hosts, but the differences between them are not significant. STATISTICALLY OR MEANINGFULLY? - FIX

3 Results

3.1 Main result: Do minority hosts have lower prices than white hosts?

Before analysis, the data set used was restricted to hosts who have profile pictures and manage less than 20 listings, and listings priced at less than \$800 per night. 64,611 listings were left after restricting the data set. There were only 20 hosts who did not have profile pictures.

Table 3 presents OLS estimates of the effect of host race and gender on the listing price. The specification is of the form:

$$Price_{i,j} = \beta_1 Race_i \, X \, Sex_i + \beta_2 Age_i + \beta_3 x_{i,j}$$

The $Price_{i,j}$ is host i's price from their Airbnb listing j. For hosts with multiple listings, each listing is treated separately. The $Race_j \, X \, Sex_j$ is the interaction of the race and sex of the host. White males are the omitted category. Age_i is the age of the host (young, middle-aged, or senior). Young hosts are the omitted category for age. $x_{i,j}$ is vector of other covariates that grows from left to right in the columns of Table 3. The columns are additive in their covariates, so each column controls for everything in the previous columns, plus a new set of covariates. Standard errors are clustered by neighborhood throughout.

The first column, Model 1, in Table 3 presents the raw effect of host race and sex on the price of a listing. These are consistent with the mean listing prices by race presented in Table 2, except now also broken down by male and female hosts within each racial category.

Model 2 adds city and neighborhood fixed effects.[15] Location is an important proxy for income

[15] Neighborhoods are assigned in accordance with each city's designations. In Chicago, for example, the fixed effect granularity is at the level of locating a listing within Hyde Park vs. Woodlawn (The University of Chicago is located in Hyde Park, and Woodlawn is directly adject to it).

levels, crime rates, and distance from downtown, which are all predictors of property prices and therefore listing prices on Airbnb. As expected, controlling for location substantially decreases the estimated racial gaps in prices. The coefficients for minority hosts decrease from a range of $20-40 to a range of $10-20 (these are all negative, and relative to white male hosts). I observe the largest decrease in the coefficients on black hosts, which go down from $40 to roughly $15. Coefficients of Hispanic hosts decrease by around $10; Asian hosts by about $20.

It is well-documented that blacks in urban populations are nearly four times more likely than whites to live in neighborhoods where the poverty rate is 40% or higher.**?** In fact, minorities at every income level live in poorer neighborhoods than do whites with comparable incomes. For example, a black household earning $75,000 a year resides in a higher-poverty neighborhood than a white household with earnings of less than $40,000 a year.**?** It is therefore expected that a large part of the variation in Airbnb prices between those groups can be explained by their listing's location. The coefficients of white females, on the other hand, persist at around $4 with the addition of location controls. This is most likely because white females tend to live in the same areas as white males and therefore have little to no variation in price that can be explained by differences in neighborhood.

Model 3 adds controls for listing-specific characteristics. Listing characteristics include fixed effects for the property type and room type, the listing's duration on the market, the number of guests the listing accommodates, the number of bathrooms, bedrooms, and beds, the bed type, the number of amenities, the number of minimum nights, any extra fees, whether the listing is instantly bookable, and the cancellation policy.[16] Controlling for these listing characteristics decreases all effects to $5-10, depending on the race of the host. Asian female hosts have the largest decrease in coefficient after controlling for listing characteristics, which indicates that a substantial part of their effect is driven by owning properties with worse observable characteristics. The effects on middle-aged and senior hosts are almost eliminated by controlling for property characteristics, indicating that their higher listing prices are primarily driven by better observable characteristics. The effects for Hispanic males and white women largely disappear with the addition of property controls.

In general, from Model 1 to Model 3, coefficients steadily decrease in magnitude and the R^2

[16]The listing's duration on the market is proxied by fixed effects for the month and year of the listing's first review.

increases from .166 with neighborhood controls to .621 with listing controls. Most of the variation in price between minority hosts and white male hosts can be explained by either the property's location or observable property characteristics. The R^2 jumps substantially to .621 in Model 3, so adding property characteristics explains much more of the total price variation than the location. This might be because Airbnb listings tend to be more concentrated in certain areas of each city (North Side in Chicago, lower and middle Manhattan in New York City, etc). If listings in a city cluster together instead of being uniformly dispersed, then controlling for location won't explain as much of the variation as controlling for property characteristics. LISTING OR PROPERTY PICK ONE - FIX

Model 4 in the last column presents my full, preferred specification. It adds host-specific characteristics to Model 3, including the host response time and the host response rate, whether the host is a Superhost, whether the host identity was verified by Airbnb, and if the host requires a guest's profile picture or phone to book.

Importantly, Model 4 also controls for variation in host effort. I attempt to account for the idea that some hosts may have higher prices not because of better observable characteristics, but just because they are better hosts. There are several host-written fields on each listing page, the "Summary", "Description", "Space", "Neighborhood Overview", "Transit", and "Notes". By filling out these fields, hosts not only describe their listing, but have the opportunity to provide guests with helpful tips and information about the surrounding area. How well a host writes these descriptions is an indication of how much effort they are willing to put into hosting. To this end, I construct three variables to measure host effort. My first variable simply measures the length of each of these fields. Presumably, the longer the description, the more effort the host put into writing it. My second variable measures whether these fields had mostly long words or short words, so that a description that uses shorter words, such as "My house is nice", would be counted as lower quality than "My house is gorgeous".

My third measure of host effort is a rudimentary sentiment analysis of the "Description" field. Hu and Liu (2004) create a list of 2,006 positive words that commonly appear in customer reviews to aid in sentiment categorization.**?** I only include words that have substantial variation in the description, meaning that more than 5% of descriptions had these words. This narrowed the list of viable words significantly. I take 7 positive words from that list that would be most relevant for Airbnb listings: “spacious”, “beautiful”, “clean”, “comfort”, “great”, “love”, and “quiet”. I then added a covariate for the number of these “good words” in the host’s “Description” field. Together, these three “host effort” variables control for hosts who write longer descriptions, use longer words in those descriptions, and put more words that are associated with positive reviews in their descriptions.

After controlling for my final specification, I estimate that, across the board, minority hosts earn lower prices from their Airbnb listing than white hosts. The biggest effect is for Asian female hosts, whose prices are roughly $9 less per day than white male hosts who own the same type of listing. The second biggest effect is for black males, with a coefficient of $7. The coefficients on black women and Asian men are $6 per day each, Hispanic females is $5. This effect is statistically significant at the $p < .001$ level for black hosts and Asian women, the $p < .01$ level for Asian male hosts, and $p < .05$ level for Hispanic women. For Hispanic men the effect is around $2 and is not statistically significant. There is little effect for white females, and a small effect that’s not statistically significant for middle-aged and senior hosts. An F-test shows that host race is jointly significant for price at the $p < .001$ level after controlling for both property and host characteristics. CHECK THIS FTEST My results are stable to the addition of host characteristic controls while still clustering standard errors at the neighborhood level. The inclusion of these host characteristics does not improve the fit of the model substantially. Property characteristics and location still explain more of the variation in price than host characteristics.

My results are consistent with Edelman and Luca’s findings, but I find smaller effects (they found a 12% price disparity, I found about a 7% price disparity). This is most likely because I control for a larger set of covariates. To confirm this, I run a regression on listings in New York City, controlling for the same covariates that Edelman and Luca used in their main result. The results, presented in Table 4, show that I get the same coefficient as the one they found - an $18 (X%) price difference between black hosts and white hosts. This indicates that my main results in Table 3 were smaller because I controlled for more variation, not because of a structural change in the extent of discrimination in Airbnb.[17]

If one believed the price difference was driven by unobserved characteristics, one might have

[17] Airbnb has changed their user interface in the past four years, so I approximated several of their regressors with the closest variable available in my data. For example, instead of whether the host had social media accounts, I controlled for whether the host’s contact information was verified by Airbnb.

expected that the price gap between white and minority hosts would disappear with the addition of more controls. This is true up to a point, since when I add more covariates my coefficients shrink relative to Edelman and Luca's. However, after that, my coefficient of interest is stable to the addition of controls - adding host-specific controls does not substantially change any of the effects. As one might expect, the R^2 goes up to .621 with the addition of location and property controls, but adding my host controls increases the R^2 by only .006.

There are a few possible sources of unexplained variation in the price of the listing - variation in the real, physical qualities of the listing that wasn't captured by the property controls, and variation in the quality of the listing's profile that was not captured by the host controls. Since I was able to control for all of the property-quality variables that Airbnb offers on a listing page, it is unlikely that there are unobserved property characteristics driving the price differences. Since adding host controls explained very little variation in the price, increasing the R^2 by only .006, it is unlikely that adding more sophisticated measures of host quality or effort would significantly help explain price disparities. While this does not eliminate the possibility that there is a set of controls not related to property type or host type that would have increased the R^2 drastically, this is still good evidence to believe that the price difference I estimate is a real difference, rather than purely caused by endogeneity.

4 Robustness checks of main result

In this section, I explore if the effects on price persist when I break up the data by city and property type.

Effects persist between large cities

Edelman and Luca found much larger effects of host race just in New York City than I did in data that included all seven cities. A reasonable hypothesis for this is that New York City is driving all of the variation in price, and when other cities are included, where discrimination might not exist, the effects get smaller. To test this, I broke up the effects of host race on listing price by city and controlled for my preferred specification. The results are in Table 6. In general, no single large city in my data set is driving all of the variation in my data. The effects on price are mostly negative for minority hosts, with a few positive coefficients in cities with fewer observations.

As expected, New York City and Los Angeles, the cities with the most host diversity and largest sample size, most closely resemble the coefficients from my main result in Table 3. In smaller cities, more than half of the negative effects are significant to various levels, and none of the positive effects are significant.

However, there are a few outlier coefficients that are most likely driven by low sample size in smaller cities. The coefficients for black hosts are fairly consistent with the combined data in all cities but New Orleans. In New Orleans, a black host is estimated to earn $18 less for the same kind of listing as a white host, an effect that is statistically significant at the $p < .05$ level. The coefficients on Hispanic hosts are mixed - in LA, NYC, and Chicago, the coefficients on Hispanic hosts are approximately the same as the combined analysis, while in Austin, New Orleans, and DC, the coefficients are slightly positive. The outlier coefficient is in Nashville, where Hispanic hosts are estimated to earn $39 less per day than a comparable white host. However, there are only 21 Hispanic hosts in Nashville, so this result is not very generalizable. In LA and NYC, the coefficients of Asian hosts are consistent with the combined data in sign and magnitude. However, they have large coefficients of -$18 to $28 in Chicago and Austin, respectively, both of which are significant. The reasoning is similar to Hispanic hosts.

Generally speaking, the price difference in New York City and Los Angeles is relatively the same. While there are outlier coefficients in smaller cities, it is unlikely that discrimination against Asian and Hispanic hosts in those cities is actually 8 times higher than New York City or Los Angeles. Rather, those cities often have less than 50 hosts in a particular racial category, so any outliers have the potential to skew the coefficients to a much larger degree.

Effects persist between listing types

Table 7 presents the effects of host race on price, broken down by various listing characteristics such as price, time on market, and property type. I break up the listings by price in the following way: separately for Los Angeles and New York City, I predict the price of a listing based on its property and host characteristics, without host race. I then use this predicted price to break up listings into higher than, and lower than, the mean predicted price in each city. I control for my preferred specification.

I find that there is much greater price disparity between white and minority hosts among high-priced listings rather than low-priced listings. Column 1 of Table 7 considers price disparities only in listings priced below-average in Los Angeles, and Column 2 considers only above-average priced listings in Los Angeles. The price disparities are much larger for the expensive listings - \$10 for black hosts, \$15 for Hispanic hosts, and \$18 for Asian hosts. By contrast, the coefficients are much smaller, only \$2-3, for the cheaper listings. This pattern in price disparities indicates that discrimination is more pronounced against minority hosts who own more expensive properties than against minority hosts who own cheaper properties. In New York City (Columns 3 and 4), the coefficients for expensive listings are larger than coefficients for cheap listings by about \$6, so this effect is not limited to one city or driven exclusively by city-specific characteristics. One hypothesis for this effect is that all of the discrimination is statistical, in which case the host race isn't as much as a proxy for property value for guests for cheap listings as it is for expensive listings, which are owned primarily by white hosts. Another explanation is that if a guest is expecting to put up a larger financial investment, they are more selective about which listings they stay at, so any existing discrimination is exacerbated.

Columns 5 and 6 show that the price disparities are similar for both old and new listings (old listings are defined as those which have been on the market for more than two years). This suggests that the effects of discrimination are *not* erased out by spending a longer time on the market; in other words, minority hosts do not simply "catch up" to white hosts after a few years. Columns 7-9 break price disparities up by property type. The effects for black hosts across property types get more pronounced the more expensive the property type. While black hosts who own an apartment expect to earn \$5 less per day than white hosts, that number increases to an \$11 loss for black hosts who own houses. This is consistent with the results in Columns 1-4 and with the statistical discrimination hypothesis - if hosts are using race as a proxy for property value, then we should expect guests to discriminate more against minority hosts who own more expensive properties, such as houses.

5 Testing Alternative Hypotheses

Hypothesis 1: Are prices lower due to supply-side effects?

Market discrimination laid out by Becker rests on the idea that groups who are discriminated against see lower demand in the market, which drives down their prices. I have estimated that prices are lower, but there could be multiple explanations in addition to discrimination for this result. One hypothesis is that, for the same type of listing, minority hosts *choose* to price their listings lower than white hosts. This might happen if minority hosts earn lower wages because they experience discrimination in the labor market. A lower opportunity cost of time would mean they have a lower marginal cost of putting up and managing their listing than a white host who owns a similar property. If minority hosts choose to price lower for every quantity, this is, in effect, equivalent to the supply curve being lower for minority hosts relative to white hosts. We therefore know that we can test this hypothesis by looking at the quantity demanded. If prices are low because the supply curve is lower, then minority hosts would have a higher quantity demanded. Conversely, if the prices are low because the demand curve is lower - which would be in line with the presence of discrimination - then the quantity demanded should be lower than it is for white hosts.

In order to test this hypothesis, I use number of reviews as a proxy for quantity demanded. In Table 6, I regress the number of reviews on host race, controlling for the same set of models as Table 3. I find that minority hosts have either the same or lower review numbers than white hosts for a similar listing that spent the same amount of time on the market. Most coefficients are either roughly zero, or negative and in the range of 1-2 reviews less than white hosts. The results are significant for white females and black hosts. While the coefficients were significant for Asian hosts under the less robust specifications, under the full specification the coefficient is not significant, but still slightly negative. In general, my results suggest that non-white male hosts do not see a higher quantity demanded than white male hosts. This is evidence against the supply-side explanation.

It is important to keep in mind that this conclusion is only salient if the total number of reviews is a reasonable proxy for the demand of a listing. Yet, one can imagine that if reviewers systematically under-review minority hosts relative to white hosts, these groups would have lower numbers of reviews that do not necessarily represent a lower quantity demanded. There is no way to tell apart these mechanisms in my data, and no previous research has been done on race-based differences in review rates.

My working assumption is that even if not every guest leaves a review, the review proportion is similar across host race, and a lower number of reviews therefore indicates a real difference between quantity demanded of minority hosts and white hosts. I substantiate this assumption with another supply-side metric, a measure of listing vacancy that I explore in the next section, that provides more evidence that lower prices are not driven by supply-side effects. Taken together, these two measures will provide strong evidence to reject the hypothesis that lower prices are due to supply-side effects.

Hypothesis 2: Are number of reviews lower because minority hosts offer their listings a fewer number of days?

In the previous section, I argued that minority hosts had a lower quantity demanded than comparable white hosts. However, they may have lower number of guests because they offer up their listing for fewer days of the month, not because people don't want to stay with minority hosts. In order to test this, I regress the availability of the listing out of 30 days on host race, controlling for my preferred specification. The availability of a listing is controlled by the host, who can update their availability calendar on their listing page. Potential guests can then see on which days the listing is available and book accordingly. When a guest books an available day, that day is removed from the availability calendar. Therefore, the availability out of 30 days measure is a true measure of the *vacancy* rate of a listing.

The results, presented in Table 7, are striking. I find that the listings of black hosts spend about 20% more time on the market vacant than the listings of white males. The effect is statistically significant, and amounts to about 2-3 days per month in real units. Interestingly, white females make their listing less available than white males, with approximately a .9 of a day statistically significant difference. This might explain why white females don't have lower prices than white males, but did have lower numbers of reviews. Perhaps white females simply offer their listing for fewer days than their white male counterparts. There were no statistically significant effects in availability for Hispanic hosts, like for most of my measures. Asian female hosts actually had a lower vacancy rate than white male hosts, which could help explain why they have a lower number of reviews.

Overall, there is strong evidence that even though black hosts offer their listing for more days, they have less people staying with them than white hosts. This is significant evidence to reject the supply-side hypothesis for black hosts. This is not the case for Asian hosts - the coefficients for Asian hosts were negative, indicating that they actually are putting up their listing for rent less often than white hosts. This means that lower availability is another possible explanation, in addition to discrimination, for why Asian hosts have a lower number of reviews. However, I am not able to further distinguish between these two hypotheses in my data.

Hypothesis 3: Are their prices lower because minority hosts have worse reviews?

Reviews are often critical for the decisions guests make about the listings they book. It is reasonable to expect the quality of a listing's reviews influences the demand, and therefore the price, for that listing. Previous analyses, including Edelman and Luca (2014), involved controlling for the numeric review score of the listing as a proxy for listing "quality". However, the numeric review score on a listing often carries little information about the real quality of the listing because there is very little variation in the numeric score.[18] In my data, 50% of listings had an average review of > 96 out of 100, and 75% had an average review score above 91 out of 100.[19] An Airbnb guest, seeing little variation in the number of stars different hosts have, may instead rely on the text of the reviews to make their booking decision. Since review text allows guests more flexibility in the feedback they give, it may provide a more accurate and nuanced picture of the guest's experience.[20] For this reason I use review text instead of the numeric score in the analysis.

In order to do this, the race, sex, and age of 16,000 reviewers who left reviews for a subset of the Chicago hosts was coded.[21] For each sentence of each review, I used a sentiment-analysis algorithm called Sentimentr to evaluate how positive or negative the sentence is. Sentimentr uses a dictionary of positive and negative words to assign each sentence a sentiment score from -1 to 1, where 1 is a positive sentence, -1 is a negative sentence, and 0 is a neutral sentence carrying no emotion. Unlike other sentiment analysis programs, Sentimentr doesn't merely count the number of good or bad words in a sentence. It also takes into account valence shifters, or words that affect the sentiment-carrying word in the sentence. For example, the algorithm assigns "I like the listing", "I *really* like the listing", and "I like the listing, *but*..." different valence scores because of the presence of valence-shifting words like "really" and "but". Sentimentr calculates the correct sentiment 60-70% of the time as compared to a human grader. One limitation of conducting sentiment analysis in

[18]A low share of guests who review may be a more accurate proxy for low quality, because many users prefer to leave no review rather than a negative review. Review share information, however, is not available, so instead I use Sentimentr to measure quality from reviews.

[19]This is the case for most online marketplaces. Fradkin, Grewal, and Holtz (2017) study the determinants of review informativeness on Airbnb and find that most reviews, both numeric and text, are positive. In general, reviews tend to reflect real experience of the user.**?**

[20]This is because a guest who leaves a text review have the opportunity to use qualifiers like "but", or "except", strengthening words like "really" or "a lot", etc.

[21]16,000 is 23% of the total amount of Chicago reviewers in the data set

this way, however, is that not every review that a human would consider bad or good carries a sentiment word that the algorithm would pick up. For example, "The apartment had cockroaches" is certainly a horrible review, but would be given a score of 0 by Sentimentr because it contains no emotion-laden words.

In Table 8, I regress this sentiment score on the host race, controlling for my preferred specification from Model 4 in Table 3. This means that each coefficient should be read as the standardized review quality, relative to white males, that reviewer of type A gave host of type B.[22] I break up my regressions by the race and sex of the reviewer, varying across the columns of Table 8. The race and sex of the host varies by row. I therefore am able to see any trends in the quality of the review different types of reviewers gave to different types of hosts.

I find that results were mixed. Overall, white reviewers show little evidence of systematic bias against minority hosts, as measured by the sentiment scores provided by Sentimentr. The reviews that white guests leave for minority hosts do not significantly differ in quality from those they leave for white male hosts. There are stronger effects when considering the quality of the review minority reviewers gave to minority hosts. Black male guests rate Asian hosts almost 4-8 standard deviations above the mean, but rate black women 3 standard deviation lower than the mean. All minority female reviewers, including black females, rate black men worse than they would rate white men who own a similar type of listing. However, all minority male reviewers rate black men anywhere from .5-2 standard deviations higher than they do white men. This suggests that there is some gender-based favoritism between minority reviewers and black male hosts. However, it is important to keep in mind that some of these large, very significant coefficients are suspicious because of small sample sizes - in several thousand Chicago host and reviewer pairs, there are simply not enough black men who stayed with Asian women to be representative of the overall distribution.

In general, there is no one minority group that uniformly has lower quality reviews. Some groups do tend to give other groups far better reviews, but there is no larger pattern of within-gender or within-race bias between hosts and guests that holds for more than one host-guest pair. Overall, there is not enough evidence to substantiate that minority hosts have systematically lower review quality that can explain lower prices.

[22] Review quality was standardized with mean 0 and standard deviation of 1.

6 Conclusion & Further Work

INSERT SOME 'WHY THIS MATTERS' HERE

To estimate the revenue loss that would result from the price differences found in the previous section, I construct a measure of revenue equal to:

$$Total\ Revenue\ =\ Price\ Per\ Day\ *\ Reviews\ Per\ Month\ *\ 12$$

The estimates of the effect of host race on host revenue are in Table 5.[23] Consistent with my prediction, all of the estimates are negative and in the range of $100-300 dollars. The biggest yearly revenue loss in the entire sample is for black females at $300, or about a 12% loss. Black males and Asian women lose about $160-180 throughout the year. Notably, white females, who had no statistically significant effects on their price in Table 3, have a significant revenue coefficient of $144. This is because of reasons inherent to the definition of revenue. Even if white women didn't have a statistically significant difference in price from their male counterparts, they do have a lower number of reviews. In the next two sections, I present and discuss evidence that *both* the price and the number of reviews for minority hosts are lower than for white male hosts. That means that when I multiply these two values together, two complementary effects both lower the total revenue of hosts. Overall, however, the same groups which had significantly lower prices also have lower revenues - black males, black females, Asian females, and white females all have significant effects in the range of several hundreds of dollars.

In this paper, I estimate that minority hosts on Airbnb, especially black and Asian hosts, earn $5-$9 less per day for the same type of listing as white hosts. This amounts to a yearly loss in revenue of $100 - $300. Importantly, I rule out several alternative hypotheses that could be driving these results, and show that racial discrimination is the explanation most consistent with my findings.

Discrimination on the platform is pertinent as subject of research because Airbnb itself can do much to address these issues. In response to media outcry about allegations of discrimination, Airbnb updated its Discrimination Policy in September 2016, increasing instant bookings (the op-

[23] I acknowledge that there are potential problems associated with using number of reviews as a proxy for demand, which are briefly mentioned in the introduction. I fully discuss them in Section 5.

portunity for guests to book without waiting for host approval) and making host profile pictures smaller. Evaluating Airbnb's efforts to address discrimination is therefore a relevant extension of this research. Since InsideAirbnb.com is continually being updated, there is now data available from webscrapes of listings after Airbnb's new discrimination policy took effect in September 2016. Future work can explore whether the policy helped curb discrimination on the platform by measuring the extent of discrimination before and after the policy took effect. If user interface really does influence the extent of discrimination in Airbnb, then the prices of minority and white hosts should start converging, suggesting that discrimination was causing the price disparities in the first place.

References

Frenken, K., 2017. Political economies and environmental futures for the sharing economy. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 375 (2095).
URL http://rsta.royalsocietypublishing.org/content/375/2095/20160367

Table 1: Summary Statistics by Race: Host and Listing Characteristics

	(1) Full sample	(2) White	(3) Black	(4) Hispanic	(5) Asian
Outcome variables					
Price	167.37	178.62	125.95	160.39	131.06
($)	(277.7)	(289.4)	(208.1)	(275.0)	(242.1)
Host listings count	5.53	5.50	10.5	3.16	2.68
	(33.0)	(31.2)	(60.3)	(17.8)	(3.62)
Number of reviews	16.57	17.14	15.06	16.46	14.08
	(30.8)	(31.9)	(27.2)	(29.7)	(27.6)
Selected covariates					
Host demographics					
Race					
White	.637	1.00	0	0	0
Black	.073	0	1.00	0	0
Hispanic	.051	0	0	1.00	0
Asian	.085	0	0	0	1.00
Unknown/Multiracial	.152	0	0	0	0
Sex					
Male	.309	.356	.354	.417	.367
Female	.378	.427	.541	.426	.476
Unknown/Two people	.312	.216	.104	.156	.157
Age					
Young (< 30)	.427	.469	.514	.481	.587
Middle-aged	.421	.491	.470	.490	.379
Old (> 65)	.018	.026	.004	.009	.009
Unknown	.133	.013	.011	.018	.024
Listing characteristics					
Property Type					
Apartments/Lofts	.598	.590	.654	.625	.601
Townhouses/Condominiums	.042	.039	.041	.041	.055
Houses	.321	.336	.279	.289	.311
Other	.039	.035	.026	.045	.033
Room Type					
Entire home/apt	.577	.607	.449	.510	.418
Private room	.384	.363	.483	.434	.530
Shared room	.04	.029	.067	.056	.052
Accommodates	3.26	3.36	2.90	3.17	2.89
	(2.3)	(2.3)	(2.1)	(2.4)	(2.1)

	(1) Full sample	(2) White	(3) Black	(4) Hispanic	(5) Asian
Bedrooms	1.30	1.33	1.20	1.25	1.20
	(.88)	(.92)	(.72)	(.90)	(.76)
Bathrooms	1.27	1.29	1.20	1.26	1.21
	(.66)	(.68)	(.52)	(.69)	(.58)
Beds	1.73	1.76	1.63	1.74	1.59
	(1.32)	(1.31)	(1.22)	(1.60)	(1.21)
Availability	11.30	10.9	14.4	11.46	10.88
(out of 30 days)	(10.99)	(10.85)	(11.54)	(11.03)	(11.05)
Number of Amenities	.80	.81	.77	.80	.75
	(1.10)	(1.11)	(1.05)	(1.10)	(1.13)
Host characteristics					
Review value	93.61	94.08	91.89	92.81	92.24
(out of 100)	(8.00)	(7.49)	(9.42)	(8.72)	(9.27)
Host is a superhost	.124	.134	.084	.108	.097
	(.329)	(.341)	(.277)	(.310)	(.296)
Response rate	.756	.756	.771	.756	.744
	(.391)	(.393)	(.368)	(.386)	(.399)
Acceptance rate	.453	.463	.357	.494	.446
	(.463)	(.463)	(.451)	(.466)	(.467)
Total “good” words	.655	.656	.686	.677	.604
	(.857)	(.843)	(.882)	(.867)	(.826)
Length of “Summary”	208.67	210.20	203.25	206.74	205.81
	(64.99)	(64.08)	(70.59)	(65.62)	(65.87)
Short words in “Summary”	.182	.185	.187	.175	.175
	(1.19)	(1.15)	(1.24)	(1.26)	(1.32)
Reviewer characteristics (Chicago data only)					
Host race	1.00	.738	.099	.079	.083
Reviewer race	1.00	.759	.041	.047	.153
Review sentiment	.510	.512	.503	.509	.506
	(.261)	(.254)	(.258)	(.276)	(.287)
Listing sentiment	.507	.509	.502	.499	.506
	(.072)	(.067)	(.089)	(.096)	(.094)
Observations (full sample)	68,983	43,988	5,023	3,524	5,893

	(1) Full sample	(2) White	(3) Black	(4) Hispanic	(5) Asian

Notes: The values in the table are means and standard deviations of host-level data in my full sample. Summary statistics for selected covariates are listed in the table. Categorical variables such as race, sex, age, property type, and room type do not have standard deviations. White refers only to Non-Hispanic Whites. Property types are explicitly listed if more than 1.5% of listings are that type. Length of "Summary" and proportion of short words in the "Summary" refer to my constructed measures of host quality. These two measures were also calculated for the description, space, neighborhood overview, notes, and transit fields, but were not included in the table for the sake of clarity and because they follow a similar pattern as the "Summary" field. The bottom panel of the table is the Chicago Reviewer data. The host race and the reviewer race in that panel is the proportion of each race that are included in the Reviewer data. The review sentiment is the sentiment of each review, the listing sentiment is the average sentiment per listing.

Table 2: Main result: Estimates of effect of host's race and gender on price

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
White Male	0	0	0	0
	(.)	(.)	(.)	(.)
White Female	-3.727*	-3.714*	-0.827	-0.243
	(1.770)	(1.562)	(1.049)	(1.053)
Black Male	-39.43***	-14.24***	-6.891***	-6.917***
	(4.058)	(3.483)	(1.994)	(1.963)
Black Female	-41.69***	-11.20***	-6.271***	-5.793***
	(4.315)	(2.793)	(1.577)	(1.586)
Hispanic Male	-20.59***	-7.247**	-2.537	-2.165
	(3.727)	(2.759)	(2.051)	(2.073)
Hispanic Female	-23.05***	-11.38***	-5.284*	-5.022*
	(4.426)	(3.219)	(2.110)	(2.078)
Asian Male	-27.42***	-12.08***	-5.834**	-6.338**
	(4.954)	(3.553)	(2.206)	(2.185)
Asian Female	-39.18***	-21.20***	-8.611***	-8.604***
	(4.463)	(2.592)	(1.586)	(1.599)
Middle-aged	12.21***	10.62***	1.724	1.702
	(2.126)	(1.281)	(0.913)	(0.907)
Old (> 65)	8.145	3.664	-1.752	-2.239
	(5.936)	(5.339)	(3.271)	(3.237)
Constant	147.4***	54.75***	-33.93***	-0.742
	(5.015)	(1.506)	(5.507)	(6.764)
Observations	45072	45072	45072	45072
Adjusted R^2	0.019	0.166	0.621	0.627

Standard errors in parentheses

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

Notes: The dependent variable is the price of the listing. All race coefficients are relative to White Males. The omitted category for age is "Young (< 30)". The unit of observation is a listing. The sample is the full sample of listings across 7 US cities. Model 1 controls for host demographics. Model 2 controls for listing location to the neighborhood level. Model 3 adds listing characteristics such as property type, time on market, number of bedrooms, availability, etc. Model 4 adds host characteristics such as response and acceptance rates, measures of host effort, Superhost status, etc. See Section 3.1 for a discussion of my covariates.

Table 3: Robustness check with controls from Edelman & Luca (2014), data from NYC

	(1)
	Price per night
Black	-18.11***
	(1.813)
Accommodates	12.84***
	(0.488)
Bedrooms	33.60***
	(1.227)
Review Scores Location	-74.66***
	(7.363)
Review Scores Location Squared	5.407***
	(0.421)
Review Scores Checkin	-1.268
	(1.157)
Review Scores Communication	-1.226
	(1.218)
Review Scores Cleanliness	3.454***
	(0.706)
Review Scores Accuracy	-1.479
	(0.973)
Host verified	1.945
	(1.357)
Private room	-71.14***
	(1.400)
Shared room	-102.8***
	(3.109)
Observations	11999
Adjusted R^2	0.526

Standard errors in parentheses

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

Notes: This table presents the effect on price of controlling for Edelman & Luca's (2014) full specification using my NYC data. My results are nearly identical to theirs (their coefficient on Black hosts was -17.8) when controlling for similar covariates in the same city. The omitted category for race is White hosts. The omitted category for room type is Entire Apartment. I could not control for host social media accounts as a proxy for host reliability like Edelman & Luca did, because Airbnb no longer provides this information. Instead, I controlled for "host verified", a boolean for whether Airbnb has the host's phone number and email. I was not able to control for "picture quality" either, but picture quality did not significantly influence price in Edelman & Luca's regression.

Table 4: Estimates of effect of host's race and gender on yearly revenue

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
White Male	0	0	0	0
	(.)	(.)	(.)	(.)
White Female	-199.0***	-156.4***	-151.9***	-144.1***
	(48.54)	(46.80)	(39.72)	(40.63)
Black Male	-655.0***	-329.5***	-261.8***	-182.9**
	(98.27)	(96.16)	(59.77)	(59.93)
Black Female	-814.7***	-365.0***	-319.5***	-298.3***
	(96.68)	(78.69)	(51.94)	(46.80)
Hispanic Male	-209.0	-44.58	-25.42	-6.391
	(112.5)	(97.48)	(88.24)	(83.84)
Hispanic Female	-280.8*	-79.78	-119.0	-68.88
	(140.1)	(120.7)	(108.2)	(104.6)
Asian Male	-360.5**	-95.77	-15.85	-22.78
	(129.4)	(115.4)	(88.42)	(84.11)
Asian Female	-676.6***	-329.2***	-183.6**	-162.6**
	(98.12)	(74.93)	(62.31)	(60.08)
Middle-aged	18.31	98.64*	-33.37	-121.1***
	(55.99)	(44.60)	(36.71)	(36.42)
Old (> 65)	222.9	249.5	-73.13	-266.0*
	(158.6)	(135.3)	(113.0)	(112.5)
Constant	2301.2***	2384.3***	2009.5***	758.0**
	(109.2)	(42.90)	(187.8)	(235.3)
Observations	45072	45072	45072	45072
Adjusted R^2	0.006	0.082	0.350	0.401

Standard errors in parentheses

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

Notes: The dependent variable is a constructed measure of yearly host revenue, as measured by (price * number of reviews per month * 12) for each listing. The omitted category for race is White Males, so all coefficients are relative to that group. The omitted category for age is "Young (< 30)". The unit of observation is an Airbnb listing, so hosts who have multiple listings are treated separately each time. The sample is the full sample of listings across 7 US cities. The specification is the same as Table 4. See Section 3.1 for a discussion of my covariates.

Table 5: Robustness checks by city

	(1) LA	(2) NYC	(3) Austin	(4) Chicago	(5) New Orleans	(6) DC	(7) Nashville
White	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Black	-5.156*	-3.977*	-6.284	-2.942	-18.45*	-7.426	-4.754
	(2.144)	(1.692)	(10.75)	(3.181)	(8.203)	(4.872)	(8.193)
Hispanic	-5.621*	-1.246	0.877	-0.807	4.109	3.264	-38.58***
	(2.197)	(1.937)	(5.212)	(5.180)	(10.77)	(4.739)	(9.458)
Asian	-5.585**	-5.975**	-27.66***	-17.64***	3.805	-5.880	10.50
	(1.785)	(2.043)	(7.763)	(4.353)	(13.36)	(3.131)	(21.29)
Observations	16825	14765	3636	3255	2563	2285	1747
Adjusted R^2	0.684	0.616	0.611	0.613	0.568	0.586	0.670

Standard errors in parentheses

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

Notes: The dependent variable is the price of a listing. This table breaks down the combined effects shown in the last column of Table 3 by city. The omitted category for race is White hosts, so all coefficients are relative to that group. For ease of reading, I did not include the gender of the host. I control for my preferred specification (referred to as Model 4 in Table 3) that includes host demographics, listing location, listing characteristics, and host characteristics. See Section 3.1 for a full discussion of the covariates included. Low number of observations for Black, Hispanic, and Asian hosts contribute to imprecise estimates in cities with less than 5,000 Airbnb hosts (New Orleans and Nashville have less than 100 Hispanic and Asian hosts, DC and Austin have less than 200 Hispanic and Asian hosts).

Table 6: Robustness checks by listing characteristics

	(1) Low $ LA	(2) High $ LA	(3) Low $ NYC	(4) High $ NYC	(5) Older listings	(6) Newer listings	(7) Apartment	(8) Condo	(9) House
White	0	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
Black	-2.241	-12.05	0.499	-10.27**	-8.925***	-7.256***	-4.875***	-7.660	-11.74**
	(1.834)	(6.999)	(1.110)	(3.753)	(2.156)	(1.363)	(1.438)	(7.996)	(3.693)
Hispanic	-3.345**	-14.91*	-1.307	0.286	-3.783	-3.089	-2.881	-8.052	-6.157
	(1.140)	(6.929)	(1.792)	(4.032)	(3.207)	(1.725)	(1.528)	(9.087)	(3.866)
Asian	-3.019**	-17.77**	-3.749*	-8.495*	-6.602**	-6.214***	-6.884***	-18.25*	-6.895*
	(1.077)	(6.005)	(1.578)	(3.487)	(2.124)	(1.743)	(1.501)	(7.687)	(2.803)
Observations	12357	4468	8383	6382	9847	25883	28410	1854	13510
Adjusted R^2	0.376	0.554	0.320	0.489	0.667	0.667	0.557	0.605	0.689

Standard errors in parentheses

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

Notes: I break down my combined data by price, time on market, and property type. The categories, from left to right, are: listings in Los Angeles and New York whose price is below vs. above the mean predicted price in those cities; listings in the entire data set which have been on the market for more than 2 years vs. less than 2 years; and listings of different property types, including apartments (includes apartments and lofts), condos (includes condos and townhouse), and houses. I do not break up data by high/low prices for the other cities in my data because smaller sample sizes lead to skewed and less informative coefficients in those cities. I control for my preferred specification throughout. The outcome variable is price of the listing.

Table 7: Estimates of effect of host's race and gender on number of reviews

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
White Male	0	0	0	0
	(.)	(.)	(.)	(.)
White Female	-0.804	-0.560	-1.420***	-1.362***
	(0.540)	(0.496)	(0.344)	(0.338)
Black Male	-2.049	-1.494	-2.034**	-1.287*
	(1.137)	(0.970)	(0.617)	(0.599)
Black Female	-2.153	-1.439	-2.523***	-2.330***
	(1.125)	(1.004)	(0.559)	(0.536)
Hispanic Male	-1.404	-0.168	-0.183	-0.00232
	(1.258)	(1.178)	(0.817)	(0.797)
Hispanic Female	-0.443	0.805	-0.867	-0.462
	(1.119)	(1.042)	(0.749)	(0.704)
Asian Male	-2.856**	-1.054	-1.024	-1.030
	(0.913)	(0.832)	(0.606)	(0.557)
Asian Female	-3.112**	-0.818	-1.201*	-0.940
	(0.952)	(0.770)	(0.598)	(0.565)
Middle-aged	4.358***	5.298***	0.634	-0.101
	(0.585)	(0.562)	(0.371)	(0.361)
Old (>65)	13.30***	14.75***	2.328	0.741
	(2.130)	(1.981)	(1.358)	(1.265)
Constant	15.49***	27.80***	130.1***	114.8***
	(0.559)	(0.474)	(2.409)	(2.853)
Observations	45072	45072	45072	45072
Adjusted R^2	0.009	0.049	0.398	0.443

Standard errors in parentheses

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

Notes: The dependent variable are the number of reviews of the listing. The omitted category for race is White Males, so all coefficients are relative to that group. The omitted category for age is "Young (< 30)". The unit of observation is an Airbnb listing, so hosts who have multiple listings are treated separately each time. The sample is the full sample of listings across 7 US cities. The specification is the same as Table 4. See Section 3.1 for a discussion of my covariates.

Table 8: Effect of host's race on listing availability out of 30 days

	(1) Number of vacant days out of 30
White Male	0
	(.)
White Female	-0.861***
	(0.114)
Black Male	2.317***
	(0.308)
Black Female	1.785***
	(0.277)
Hispanic Male	-0.154
	(0.331)
Hispanic Female	-0.0906
	(0.341)
Asian Male	-0.195
	(0.299)
Asian Female	-1.191***
	(0.259)
Observations	45779
Adjusted R^2	0.215

Standard errors in parentheses

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

Notes: This table presents the effect of host race on listing availability out of 30 days, controlling for my preferred specification in Table 3, Model 4. When a listing is booked, this availability metric is updated on the Airbnb website to reflect that booking. Therefore, this measure actually represents the number of days out of the total available days that listings were vacant, relative to a White male host.

Table 9: Estimates of effect of host demographics on review sentiment, by reviewer demographics

	(1) White M Rev.	(2) White F Rev.	(3) Black M Rev.	(4) Black F Rev.	(5) Hisp. M Rev.	(6) Hisp. F Rev.	(7) Asian M Rev.	(8) Asian F Rev.
White Male	0	0	0	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
White Female	-0.0280	0.0871	1.034*	-0.277	0.189	0.0536	0.132	0.0961
	(0.0716)	(0.0484)	(0.454)	(0.283)	(0.326)	(0.202)	(0.126)	(0.111)
Black Male	0.0490	0.0513	2.389	-0.0865	1.267	-3.605**	0.323	-1.272***
	(0.228)	(0.329)	(1.409)	(1.079)	(0.819)	(1.136)	(0.365)	(0.267)
Black Female	-0.118	0.0165	-2.719**	-0.0154	0.232	-0.0402	0.968***	-0.370
	(0.160)	(0.103)	(0.916)	(0.666)	(0.548)	(0.743)	(0.249)	(0.205)
Hispanic Male	0.00202	0.109	-0.345	0.852	-0.459	-0.521	-0.0522	0.0696
	(0.0956)	(0.126)	(0.905)	(0.616)	(0.573)	(0.906)	(0.198)	(0.287)
Hispanic Female	0.0442	-0.0668	4.569*	-0.867	-1.141	1.364**	0.0841	-0.146
	(0.325)	(0.0823)	(1.819)	(2.900)	(0.823)	(0.485)	(0.192)	(0.464)
Asian Male	-0.270	-0.185	3.609***	0.373	-0.0566	-1.121	0.741**	0.229
	(0.235)	(0.167)	(0.546)	(0.819)	(1.064)	(1.815)	(0.251)	(0.273)
Asian Female	-0.163	-0.135	7.498***	0.946	-0.633	-0.573	0.0775	-0.397
	(0.159)	(0.127)	(1.754)	(0.603)	(0.958)	(0.523)	(0.227)	(0.310)
Observations	2690	2557	124	171	201	145	487	537
Adjusted R^2	0.007	0.001	0.271	0.194	0.102	0.136	0.021	-0.006

Standard errors in parentheses

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

Notes: This table measures the quality of a review that reviewers (each column represents a type of reviewer) leave for hosts (each row is a different type of host) in Chicago. The outcome variable is the sentiment of the review. Each coefficient is the standardized sentiment of a review (standardized with a mean 0 and a standard deviation of 1). Review sentiment was coded by an R script, SentimentR, and measures how positive or negative the review is. Reviews that are numerically positive are of positive sentiment and numerically negative are negative sentiment, relative to the mean sentiment score for each host type. The demographics of the reviewers are the columns (Male is “M”, Female is “F”), and the demographics of the host are the rows. The unit of observation is a single review. The data is a subsample of the Chicago hosts and their reviewers. I control for my preferred specification throughout (referred to as Model 4 in Table 3) that includes host demographics, listing location, listing characteristics, and host characteristics. See Section 3.1 for a full discussion of the covariates included.

7 Appendix

7.1 Data Appendix

Inside Airbnb also provides some time-series information on prices, but since the each listing's price was not scraped daily, there are often week-long or month-long gaps in the time-series price data. A cursory glance at the time-series prices reveals that hosts do not change prices often, and if they do, they often reflect predictable weekend or holiday seasonality. There is therefore reason to believe that the prices posted at the time of the scrape are representative of a listing's price throughout the year. Because of the incompleteness of the time-series data set, I focus on the cross-sectional data for the main analysis.

The data set does not include Airbnb's original neighborhood designations "due to inaccuracies". Instead, the scraper assigned neighborhoods to each listing by comparing the geographic coordinates of the listing with each city's neighborhood designations.[24] Figure 6 presents a map of Chicago's neighborhoods to give the reader a sense of the granularity of the neighborhood controls.

Airbnb does not provide the demographic information of their users, so research assistants manually coded the hosts' demographic information. Research assistants were provided a link to the host profile picture and host name, and coded each picture according to the host's sex, race, and age. Only hosts with single-person pictures who were identifiably white, black, Asian, or Hispanic were included in the main analysis.[25] Importantly, listings that no longer existed at the time of coding were also excluded.[26]

Each RA was compensated based on the quantity of the listings they coded. This could create the incentive to code for speed rather than accuracy, so a simple double-checking process was put in place to check codings. For hosts whose picture was ambiguous on any of the dimensions of race, sex, or age, RAs were instructed to flag the listing. I subsequently coded each flagged picture and checked RA work. Due to manpower constraints, one RA coded each picture.[27]

[24] Location information for listings is anonymized by Airbnb, and no exact address is provided for any listing. The location for a listing could be 0-150 meters from the actual address.

[25] All other types of profile pictures, including couples, groups of more than two people, children, pictures without a human face, or hosts of ambiguous race were dropped from the main analysis.

[26] If certain groups of hosts systematically exited the Airbnb market between the time of the scrape and the time of the coding, dropping those listings could bias the results. Unfortunately, there is no way to verify the demographics of the hosts who dropped out, since Airbnb takes down their profile picture.

[27] It is important to note that the coding need not reflect the actual demographics of the host. Rather, it is sufficient that they are coded with the race, sex, and age that the average user on Airbnb would assume after looking at the profile picture. However, one limitation of this method is that the average University of Chicago undergraduate might not be representative of the average guest on Airbnb. With more resources, a more rigorous coding process could have been conducted. In future research, it would be preferable for two people to code each picture, and a third person to mediate any disagreement.

Table 10: Coding categories

	Sex	Race	Age
1	Male	White	Young (< 30)
2	Female	Black	Middle-aged
3	Two males	Hispanic	Old (>65)
4	Two females	Asian	Unknown
5	Two people, different sex	Multiracial	
6	Unknown	Unknown	

Notes: This table presents the categories according to which Research Assistants coded the race, sex, and age of the hosts and reviewers. Each host was assigned one category from each column. White refers only to non-Hispanic Whites. The Unknown categories are for profile pictures that are non-human, had more than one person, had only children, or did not contain a face. Multiracial is for pictures with two people of different race. For my main results, only Male/Female and White/Black/Hispanic/Asian were included, as interactions.

7.2 Figures

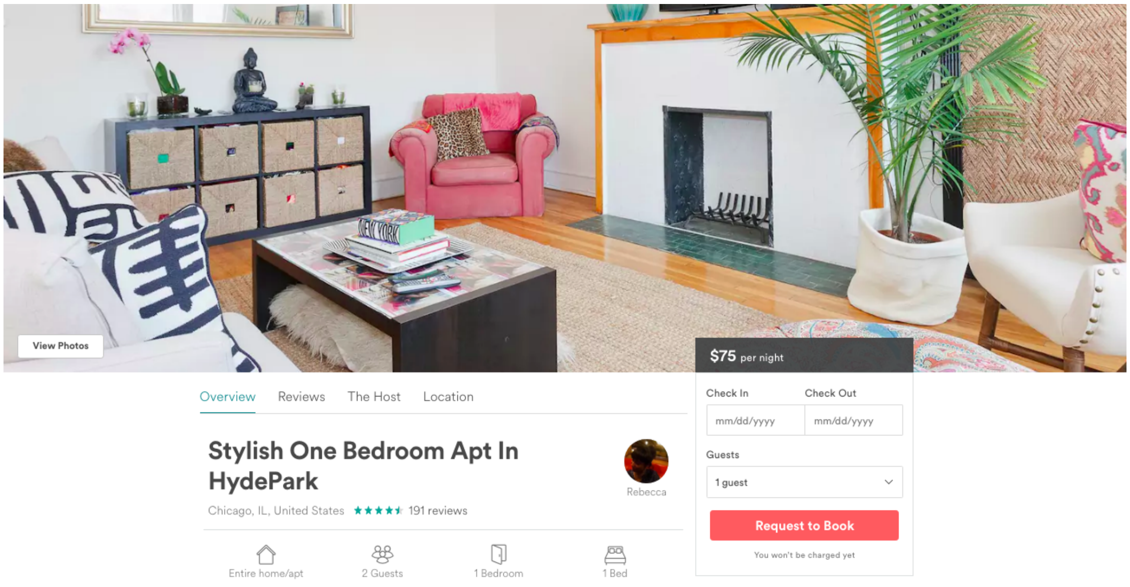

Figure 1: Sample listing profile from Hyde Park, Chicago

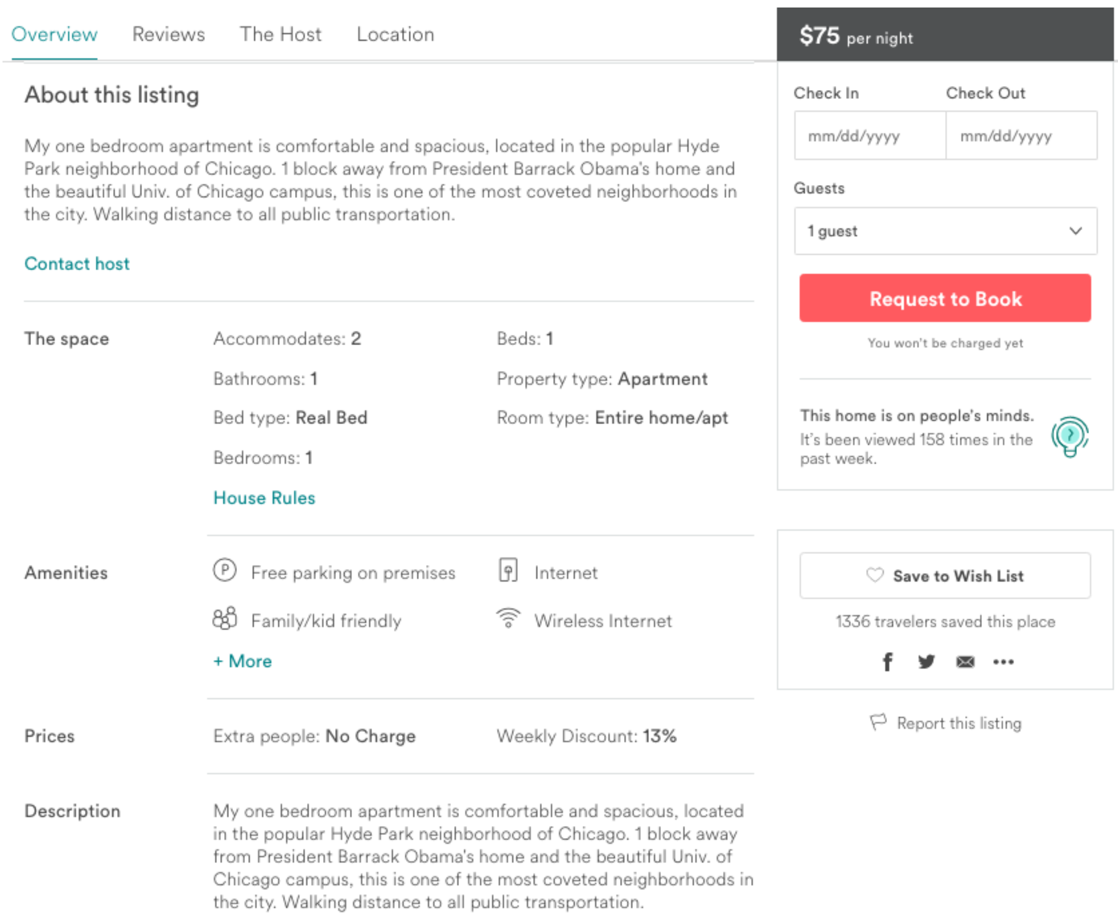

Figure 2: Sample property characteristics

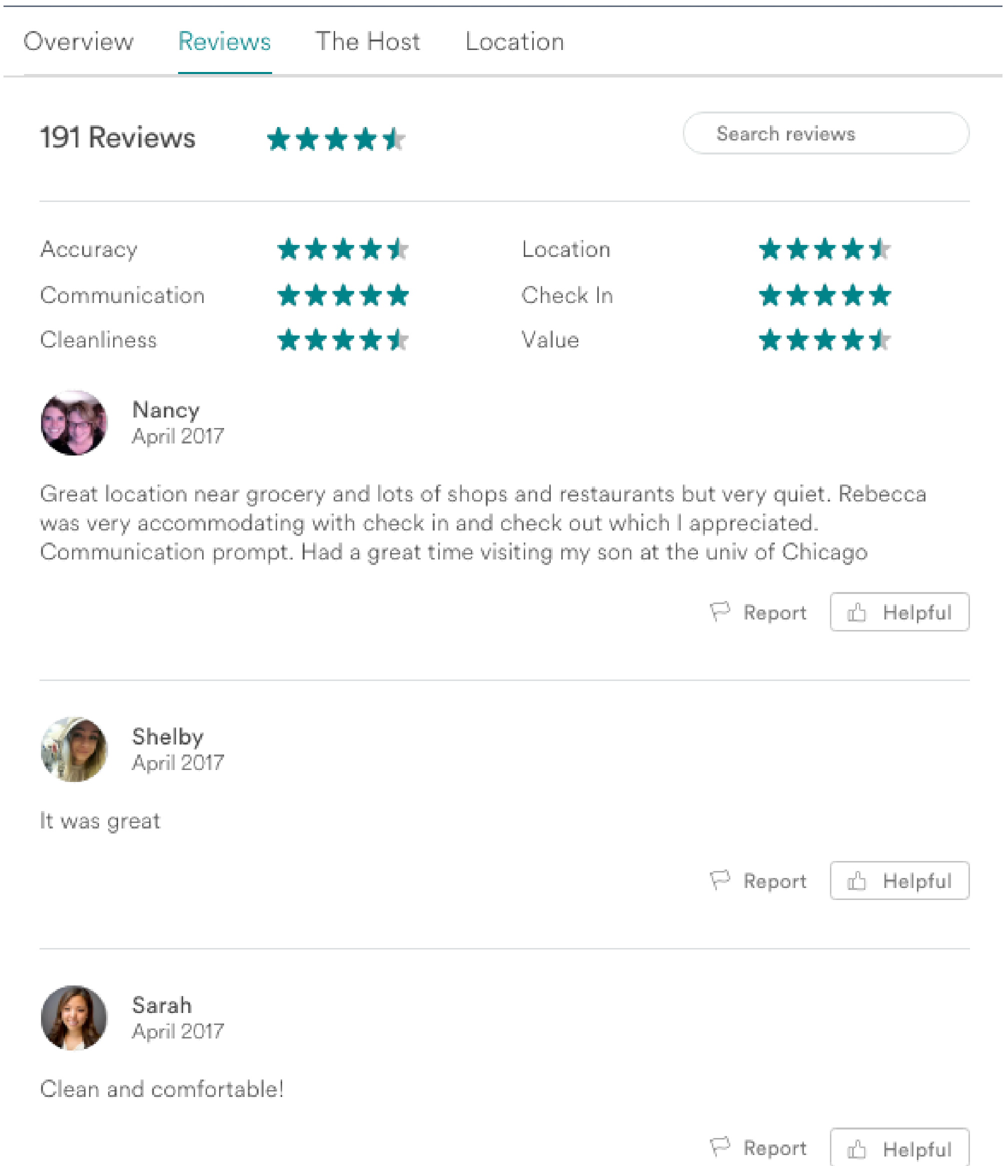

Figure 3: Sample review information

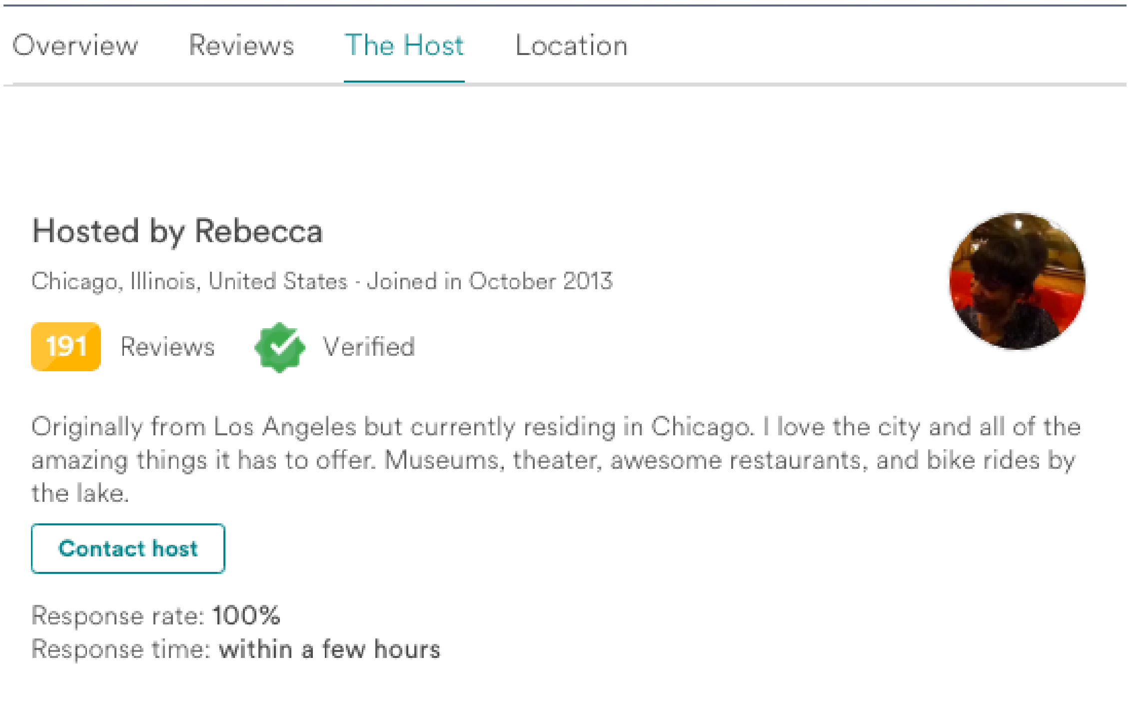

Figure 4: Sample host information available. Note that this Figure reflects the changes made to the website after Airbnb updated its discrimination policy. The guests in my data would have seen a larger host profile picture than shown here.

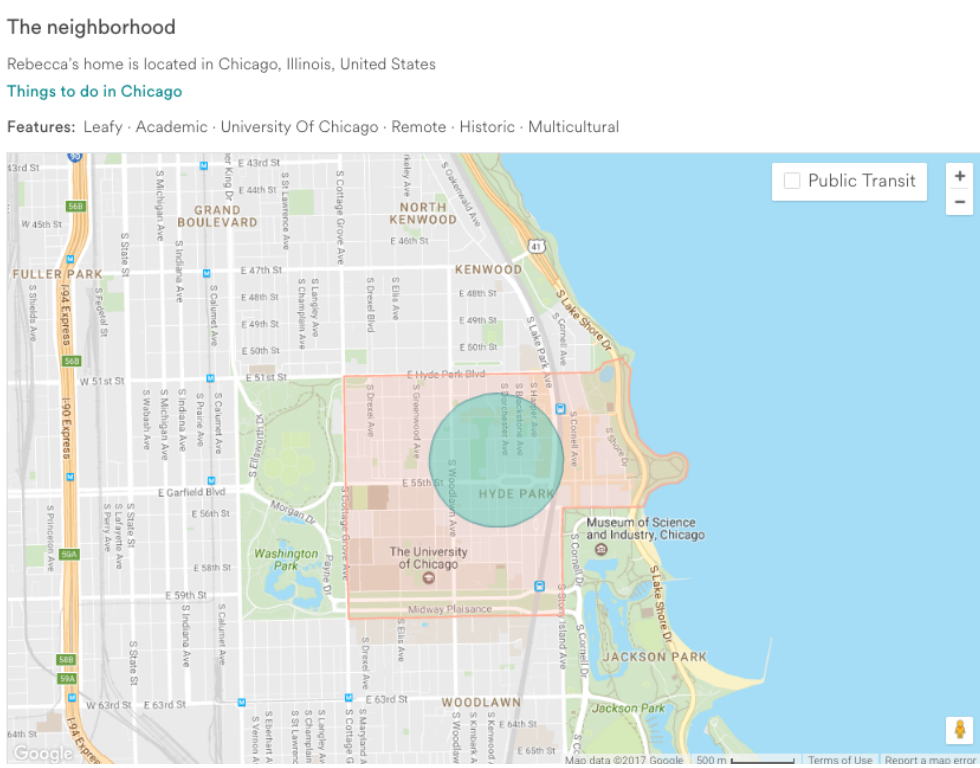

Figure 5: Sample location information

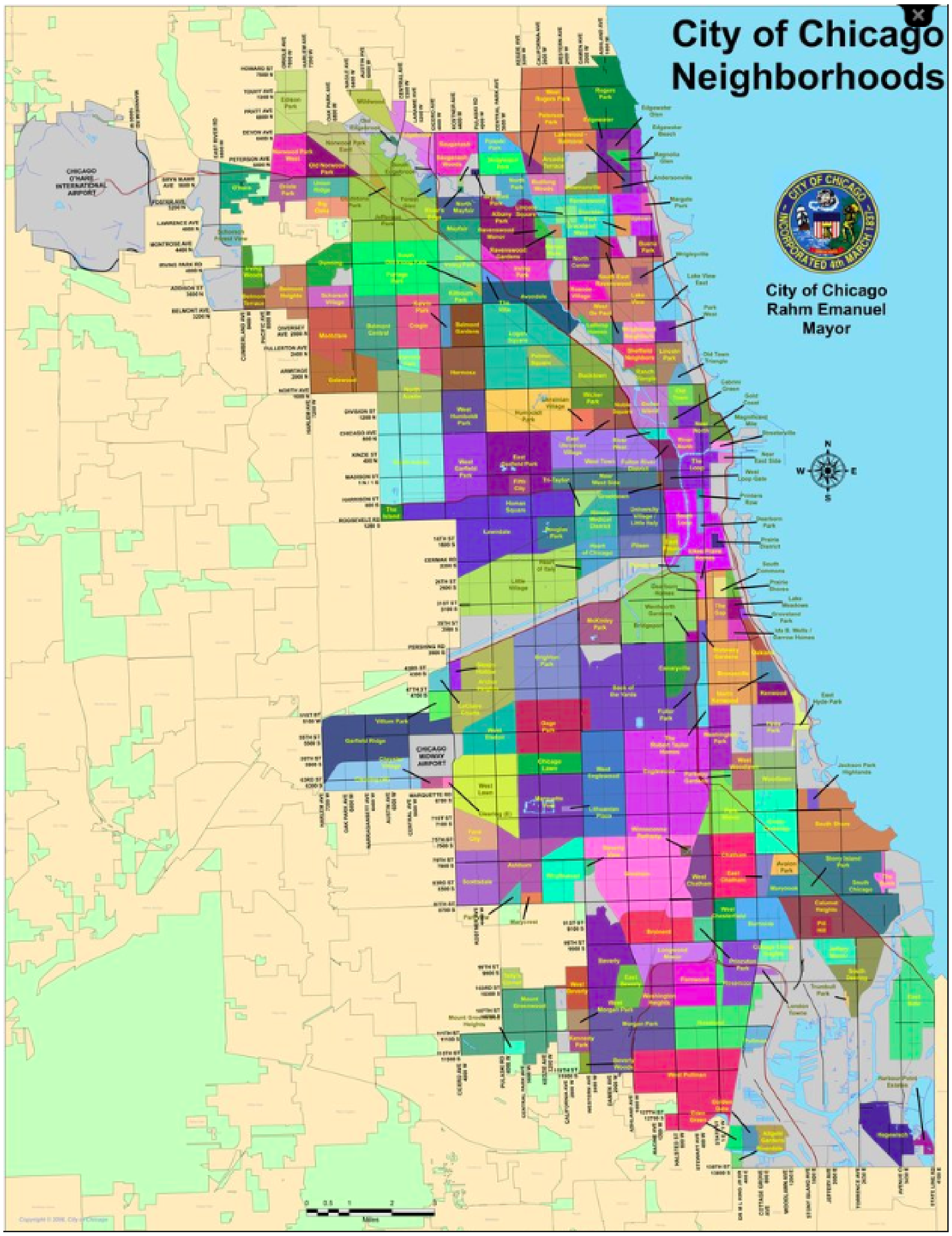

Figure 6: City of Chicago neighborhoods, showing level of granularity of neighborhood controls